



ERSTELLUNG VON OBJEKT-ORIENTIERTER
ANALYSE-SOFTWARE UND
BESTIMMUNG VON JET-KORREKTUREN
ZUR REKONSTRUKTION VON
TOP-QUARKS AM LHC

CHRISTOPHER JUNG

DIPLOMARBEIT

AN DER FAKULTÄT FÜR PHYSIK
DER UNIVERSITÄT KARLSRUHE

*Referent: Prof. Dr. G. Quast
Institut für Experimentelle Kernphysik*

*Korreferent: Prof. Dr. M. Feindt
Institut für Experimentelle Kernphysik*

23. JANUAR 2004

Contents

Introduction	1
1 The CMS detector at LHC	3
1.1 LHC	3
1.2 CMS	4
1.2.1 Magnet System	5
1.2.2 Tracking	5
1.2.3 Calorimetry	6
1.3 Muon System	8
1.4 Triggering	8
2 The Top Quark	11
2.1 The Standard Model and its Particles	11
2.2 Special Features of the Top Quark	12
2.3 Top Mass	13
2.4 Decay	13
2.5 Effects of the Short Lifetime	14
2.6 Top Pair Production	15
2.7 Top Pair Decay	17
2.8 Single Top Production	17
2.9 Top Quark, W Boson and Higgs Boson	18
3 Jet Algorithms	21
3.1 Iterative Cone Jet Algorithm	21
3.2 k_T Algorithm	22
3.2.1 Inclusive Mode	22
3.2.2 Exclusive Mode	23
3.3 Jet Algorithm Implementation in ORCA	24
3.4 Using Jet Algorithms	24
4 Production and Analysis Tools	25
4.1 Software	25
4.2 Production of $t\bar{t}$ Samples	30
4.3 PAX	31
4.3.1 PAX features	31
4.3.2 The PAX n-tuple Filler	32
4.3.3 The PAX Combined Objects Class	32

5	The W Mass Reconstruction	37
5.1	Event Topology	37
5.2	Reconstruction using Monte Carlo Knowledge	37
5.2.1	Jets at Generator and Reconstruction Level	39
5.2.2	Using the Monte Carlo Knowledge	41
5.2.3	Jet Energy Corrections	44
5.2.4	Best Mass Approach	52
5.3	Realistic Reconstruction	52
5.3.1	Iterative Cone Algorithm with Cone Size 0.5	53
5.3.2	Iterative Cone Algorithm with Cone Size 0.6	56
5.3.3	Iterative Cone Algorithm with Cone Size 0.7	60
5.3.4	Iterative Cone Algorithm with Cone Size 0.8	62
5.3.5	k_T Algorithm	64
5.3.6	Comparing the Results of the Jet Algorithms	69
6	Top Mass Reconstruction	71
6.1	Reconstruction using Monte Carlo Knowledge	71
6.1.1	Using the Jet Energy Corrections of the W -Jets	71
6.1.2	Using the b -jet Energy Corrections	71
6.2	Realistic Reconstruction	74
6.2.1	Best Mass approach	74
6.2.2	Using the Angles	75
7	Conclusion and Outlook	77
	Zusammenfassung	79
	Danksagung	91

List of Figures

1.1	The LHC tunnel and the four LHC experiments.	4
1.2	The pixel detectors in a schematic view.	5
1.3	The ECAL in a three-dimensional view.	7
1.4	A muon traversing the CMS detector.	7
1.5	The data flow in the CMS Trigger/Data Acquisition system . . .	9
2.1	Top Quark mass measurements	14
2.2	The “off-diagonal” basis of a $t\bar{t}$ pair	15
2.3	Feynman diagrams for top pair production	15
2.4	Parton distribution function for a proton	16
2.5	Feynman diagrams for electroweak production of single top quarks in leading order	18
2.6	Comparison of the prediction of the top mass and the W boson mass from electroweak precision measurements to their measured values.	19
3.1	A jet with a cone size of 0.4 in the η - ϕ -plane	23
4.1	A simple visualisation of the production chain used	29
4.2	Schematic overview of a PAX event interpretation.	32
4.3	Schematic overview of PAX relation management.	33
4.4	A track and the calorimeter fulfilling the distance criterion. . . .	34
5.1	Sketch of the semileptonic $t\bar{t}$ decay.	38
5.2	Number of generator jets using an iterative cone algorithm with a cone size of 0.5.	38
5.3	Number of reconstructed jets using an iterative cone algorithm with a cone size of 0.5.	39
5.4	A lego plot of the number of generator jets and reconstructed jets for the iterative cone algorithm with a cone size of 0.5. . . .	40
5.5	The distribution of the true W mass, taken from the Monte Carlo data.	41
5.6	Reconstructed W masses from generator jets, using Monte Carlo data to select the right jets.	42
5.7	W mass distribution from reconstructed jets, using Monte Carlo data to select the right jets.	43

5.8	Number of the reconstructed jets originating from the hadronically decaying W bosons as a function of $ \eta $ and the energy . . .	44
5.9	'Stamp plots' of the correction factor for different $ \eta $ -energy intervals	47
5.10	'Eta slices' for different values of $ \eta $, showing the dependence of the correction factor on the energy of the reconstructed jet. . . .	48
5.11	The correction factors depending on the absolute value of the pseudo-rapidity, $ \eta $, and the jet energy, for a cone of size 0.5. . . .	49
5.12	Fitting of the parameters of the function defined in (5.1) as polynomials of degree 3.	50
5.13	Reconstructed W masses from reconstructed jets after jet energy corrections, using MC data for the selection of the right jets. . .	51
5.14	Top mass approach for obtaining the W mass from reconstructed jets	52
5.15	W mass reconstructed from five jet events, assuming a perfect b - and muon-tagging.	54
5.16	W mass reconstructed from events with five to twelve jets, assuming a perfect b - and muon-tagging.	55
5.17	Jet multiplicities of reconstructed jets using an iterative cone algorithm with a size of 0.6.	56
5.18	Jet energy corrections for an iterative cone algorithm with a cone size of 0.6.	57
5.19	Realistically reconstructed W masses from five jet events, assuming perfect b - and muon-tagging, using an iterative cone algorithm with cone size 0.6.	58
5.20	W mass reconstructed from events with five to twelve jets, assuming a perfect b - and muon-tagging, using an iterative cone algorithm with a cone size 0.6.	59
5.21	Jet multiplicities of reconstructed jets using an iterative cone algorithm with a cone size of 0.7.	60
5.22	W mass reconstructed from events with five to twelve jets, using an iterative cone algorithm with cone size 0.7, assuming a perfect b - and muon-tagging.	61
5.23	Jet multiplicities of reconstructed jets using an iterative cone algorithm with a cone size of 0.8.	62
5.24	W mass reconstructed from events with five to twelve jets, using an iterative cone algorithm with cone size 0.8, assuming a perfect b - and muon-tagging.	63
5.25	Box plot of the number of reconstructed jets vs. the number of generator jets, using the inclusive k_T algorithm and tight matching to Monte Carlo data.	64
5.26	The correction factors depending on the energy and the absolute value of the pseudo-rapidity, $ \eta $, for the inclusive k_T algorithm, using the ΔR scheme as the angular definition and the E scheme as the recombination scheme.	65
5.27	W mass obtained from reconstructed jets, using Monte Carlo knowledge, for the inclusive k_T algorithm.	66

5.28	W mass realistically reconstructed from reconstructed jets, for the inclusive k_T algorithm.	67
5.29	W mass reconstructed from jets found by the exclusive mode of the k_T algorithm, assuming a perfect b - and muon-tagging. . . .	68
6.1	Top mass reconstructed using tight Monte Carlo matching for finding the b -jet belonging to the hadronically decaying W , assuming perfect b - and muon-tagging.	72
6.2	Top mass reconstructed using Monte Carlo information for finding the b -jet belonging to the hadronically decaying W , using b -jet energy corrections.	73
6.3	Top mass reconstructed without using Monte Carlo information, using b -jet energy corrections, choosing the best combination by comparing the invariant mass to the world average.	74
6.4	Top mass reconstructed using b -jet energy corrections, choosing the best combination by comparing the angles between the reconstructed W and the b -jets.	75
7.1	'Eta-Scheiben' für verschiedene Werte von $ \eta $	81
7.2	Realistisch rekonstruierte W -Boson-Masse aus Ereignissen mit fünf bis zwölf Jets, die vom iterativen Cone Algorithmus mit Cone Größe 0,5 gefunden wurden.	82
7.3	Top-Quark-Massenrekonstruktion unter Annahme von perfektem b - und μ -Tagging. Zur Auswahl des b -Quarks werden die Winkel zwischen rekonstruiertem W -Boson und den b -Jets benutzt. . . .	84

List of Tables

2.1	The interactions and their bosons (without gravitation)	12
2.2	The quarks and their quantum numbers.	12
2.3	The leptons and their quantum numbers.	12
4.1	An extract from a typical PYTHIA output	27
4.2	Versions of the software used for $t\bar{t}$ production	30
4.3	The CMKIN datacard used for producing the $t\bar{t}$ data sample . .	31
5.1	Bin sizes for the energy, used for all the eta slices.	45
5.2	Realistically reconstructed W masses, assuming perfect b - and muon-tagging, using an iterative cone algorithm with cone size 0.5.	54
5.3	Realistically reconstructed W masses, assuming perfect b - and muon-tagging, using an iterative cone algorithm with cone size 0.6.	57
5.4	Reconstructed W masses for jets found by an iterative cone algo- rithm with a cone size of 0.7. Perfect b -tagging and muon-tagging are assumed.	61
5.5	Reconstructed W masses for jets found by an iterative cone algo- rithm with a cone size of 0.8. Perfect b -tagging and muon-tagging are assumed.	63
5.6	Comparison of the results of the different jet algorithms	69
7.1	Vergleich der Resultate der verschiedenen Jet-Algorithmen. . . .	83

Introduction

Quarks and leptons are the fundamental building blocks of matter. Six quarks and six leptons are expected to exist within the framework of the Standard Model. They are arranged in three families of doublets. The top quark was the last to be discovered. It is for sure the most interesting of the quarks, having a mass much higher than all the others. The top quark is an important ingredient in the Standard Model of particle physics, e.g. radiative corrections strongly depend on its mass.

At LHC, top quarks will be produced with enormous statistics. This enables physicists to measure the top quark parameters with small errors.

Also, the top quark will be very important in the hunt for the Higgs boson. If the Higgs boson exists, there will be associated Higgs production with a top pair, the $t\bar{t}H$ channel. Furthermore, the top quark pair production constitutes a significant source of the background events, e.g. $t\bar{t}$ decay will be a background to $H \rightarrow ZZ^* \rightarrow 4\mu$.

In associated Higgs production, top quarks will be used for tagging purposes. For a top decay into three jets, the tagging depends on b -tagging and on the reconstruction of the invariant mass. Because of calorimetric smearing effects, energy leakage outside the predefined cone size of certain definitions of jet clustering schemes, and because of noise and pile-up effects, the jet energies measured in the calorimeters must be corrected in order to correspond roughly to the energy of the original quark.

In this thesis, a jet energy correction scheme for typical high energetic jets, arising from top decays, will be developed. This is done on a large data sample, which has been specially produced for this purpose, with full detector simulation. The jet energy correction scheme is tested by reconstructing the masses of hadronically decaying W bosons and top quarks.

Chapter 1

The CMS detector at LHC

Hadron colliders have been very successful in high energy physics discoveries. For example, the $Spp\bar{S}$ (CERN Super proton antiproton Synchrotron) was built in order to discover the intermediate gauge bosons W and Z . This goal was attained in 1983.

The Tevatron Accelerator at FNAL¹ was built to measure the W and Z masses directly with higher statistics and did so very well. Tevatron's greatest success was the discovery of the top quark and the determination of its mass in 1995.

The next step in this chain is LHC², currently being built to discover the mechanism that is responsible for the masses of the elementary particles. The most popular theory for this is the Higgs mechanism.

1.1 LHC

The LHC is a proton-proton accelerator and collider, currently being installed in the former LEP³ tunnel near Geneva. The ring has a circumference of approximately 27 km. The collider is supposed to be brought on line in 2007.

About 10^{11} protons will form a bunch. There will be about 3000 bunches in the ring. They will be accelerated and brought to collision. The protons will be kept on track by magnetic fields of about 8 T, and will collide head on at a frequency of 40 MHz at four interaction points. The total centre-of-mass energy of the proton-proton collisions will be $\sqrt{s} = 14$ GeV, which is nearly a magnitude higher than the centre-of-mass energy at Tevatron⁴.

There will be four experiments situated at LHC (see figure 1.1). Two experiments will be general purpose detectors, ATLAS⁵ and CMS⁶, one will be a detector purpose-built for heavy ion collisions during heavy ion runs, ALICE⁷,

¹Fermi National Accelerator Laboratory

²Large Hadron Collider

³Large Electron Positron Collider

⁴ $\sqrt{s} = 1.96$ TeV

⁵A Toroidal LHC Apparatus

⁶Compact Muon Solenoid

⁷A Large Ion Collider Experiment

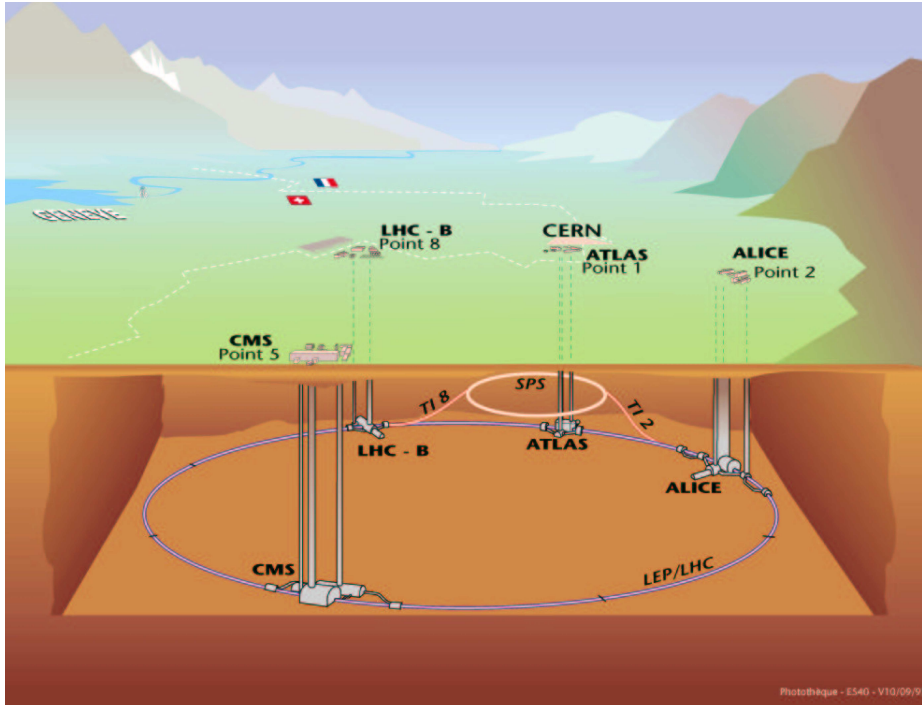


Figure 1.1: The LHC tunnel and the four LHC experiments.

and the fourth one will be a detector purpose-built for b-physics, LHCb⁸.

LHC will not only have a very high collision energy but also a huge design luminosity of $L = 2 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1}$ during the first few years and even $L = 10^{34} \text{cm}^{-2} \text{s}^{-1}$ during the later years of operation. This results in 3.5 and 17.5 interactions per crossing on average, respectively. One expects beam lifetimes of up to 10 hours.

LHC experiments will be able to discover new physics and to study them with sufficient statistics. They will also measure known physics with enormous statistics. In the field of new physics, the hunt for the Higgs boson and its mass is clearly the most interesting issue. There are also other theories, e.g. supersymmetry or technicolour theories, and the so called exotic theories, which will be proved or excluded during the LHC runs. Many precision measurements will be done at LHC, e.g. the determination of the W boson mass and the top quark mass. These measurements themselves will be important for the determination and the discovery of new physics, e.g. for constraining SUSY⁹ parameters.

1.2 CMS

The CMS detector consists of several components which will be discussed in the following sections. A much more detailed description can be found in [CMSOutreach].

⁸LHC beauty experiment

⁹Super Symmetry

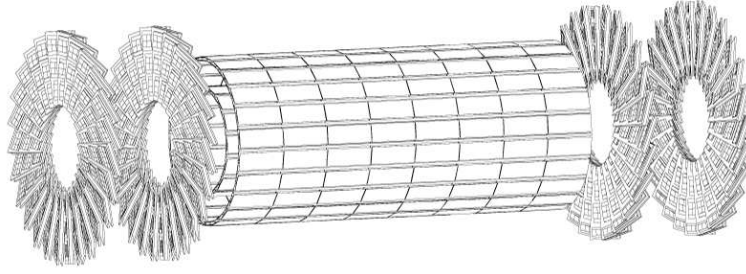


Figure 1.2: The pixel detectors in a schematic view.

1.2.1 Magnet System

The main feature of the CMS detector is its uniform magnetic field of 4 Tesla. The magnet system mainly consists of the superconducting coil, the magnet yoke and the vacuum tank. The total weight is about 12000 metric tons. Because of the relatively 'small' volume¹⁰ the detector is called 'compact'.

1.2.2 Tracking

CMS uses an all-silicon approach for tracking. The total area of silicon strip sensors is 210 square meters. The tracking will be done by silicon strip detectors and by pixel detectors.

Pixel detectors

The silicon pixel detectors (for a schematic view see figure 1.2) are located at distances of 4 cm and 7 cm at low luminosity and 7 cm and 11 cm at high luminosity from the beam line. The two endcap disks at each end of the detector cover the region for radii from 6 cm to 15 cm.

The cell size in the Pixel detectors is 0.025 mm^2 . Being n-on-n devices, the response in the barrel is affected by the 34° Lorentz angle of the drift electrons. The spatial resolutions are about $10 \text{ }\mu\text{m}$ and $14 \text{ }\mu\text{m}$ in the ϕ and z coordinates respectively.

These pixel detectors allow the measurement of secondary vertices which are mainly the result of the decay of long-living particles, e.g. B mesons.

Silicon Strip Detectors

The four inner barrel layers are assembled in shells. Layer 1 and 2 are even double-sided. The two inner endcaps consist of three small discs.

In the outer barrel the modules are arranged in six concentric layers with the two inner layers being double-sided, too.

¹⁰The ATLAS detector is about eight times larger.

Measuring the Momentum of Charged Particles

In a homogeneous magnetic field, the transverse momentum, p_T , of a charged particle can be calculated from measuring the radius of its trajectory

$$p_T = 0.3 \frac{\text{GeV}}{c} \cdot \frac{1}{\text{m} \cdot \text{T}} \cdot B \cdot R,$$

with B being the magnetic field and R being the radius.

In the central region $|\eta| \leq 1.6$, high momentum isolated tracks can be reconstructed with a transverse momentum resolution of better than

$$\frac{\delta p_T}{p_T} \approx \left(15 \frac{p_T}{\text{GeV}/c} + 0.5 \right) \%.$$

This degrades gradually to

$$\frac{\delta p_T}{p_T} \approx \left(60 \frac{p_T}{\text{GeV}/c} + 0.5 \right) \%,$$

with the absolute value of the pseudo-rapidity, $|\eta|$, approaching 2.5.

1.2.3 Calorimetry

The basic idea of calorimetry is to measure the energy and position coordinates from high energy interactions by total-absorption methods. An incoming particle interacts in a large calorimeter. This process generates secondary particles which will generate tertiary particles and so on. In the end, the incident energy appears as ionisation or excitation in the calorimeter.

There are two layers for calorimetry, the electromagnetic one (ECAL) and the hadronic one (HCAL). They stop the photons, electrons, positrons and hadrons and measure their energy.

ECAL

The electromagnetic calorimeter uses $PbWO_4$ scintillating crystals of high density in order to get a good energy resolution and to have a very compact calorimetry system. A sketch of the ECAL is given in figure 1.3.

HCAL

The hadronic calorimeter is a sampling calorimeter with 5 cm thick copper absorber plates, which are interleaved with 4 mm thick scintillators. Strong interacting particles will deposit their energy in the HCAL. In order to have a good resolution for missing energy a calorimetry coverage up to a pseudo-rapidity $|\eta| = 5$ is needed.

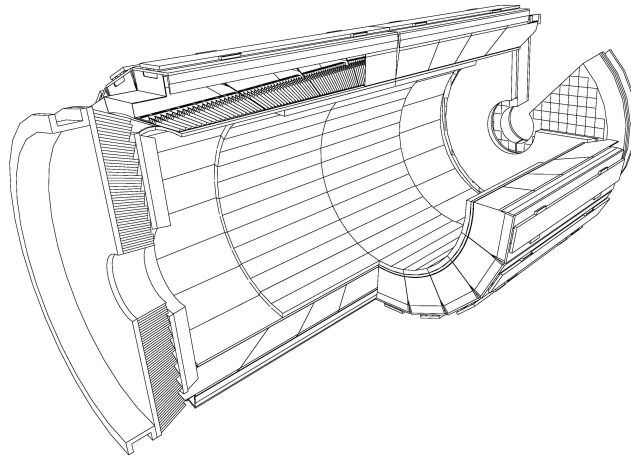


Figure 1.3: The ECAL in a three-dimensional view.

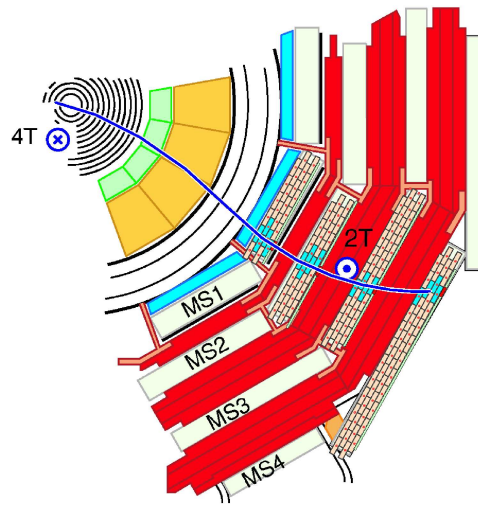


Figure 1.4: A muon traversing the CMS detector.

1.3 Muon System

The muon detectors consist of four muon stations interleaved with the iron return yoke plates, arranged in concentric cylinders around the beam line in the barrel region and in disks perpendicular to the beam line in the endcaps.

The absorber material, iron, cannot be traversed by any particles but muons, neutrinos and WIMPs¹¹. The identification process requires hits in at least two of the four muons stations.

The measurement of the muon momentum uses the bending of charged tracks in the magnetic field produced by the coil and conducted by the return yoke.

Muons will play a very important role in several physics channels, e.g. the golden channel $H^0 \rightarrow ZZ^{(*)} \rightarrow 4\mu$, since they provide very clean signals.

A sketch of a muon traversing the muon system is shown in figure 1.4.

1.4 Triggering

At LHC, there will not only be high statistics of 'interesting' physics, but also much more statistics in physical processes we are not interested in. The challenge of selecting the interesting physics processes is complicated by pile-up.

There are two types of pile-up:

1. **„In-Time” pile-up**

This pile-up consists of particles coming from the same crossing but originating from a different proton-proton interaction.

2. **„Out-of-Time” pile-up**

Since the distance between two bunches is $\sim 7.5\text{m}$ (being $\sim 25\text{ ns}$ in time), there will be signals left over from interactions in previous cross. These are called „out-of-Time” pile-up.

CMS starts selecting events by using the Level-1 trigger. This trigger is very fast (runs less than a μs). It uses special hardware processors that seek simple signs of an interesting event, e.g. muon chamber hits lying on a given path. The Level-1 trigger selects the best 100,000 events each second. If an event is accepted, the data is stored in 500 independent memories, called RDPMs¹². Each of these memories is connected to a different part of the detector.

The next step is event building, i.e. assembling the data corresponding to various pieces of the detector. There will be a large readout switch, connecting all 500 RDPMs to a farm of computers, where the test on the high trigger levels (Level-2 and Level-3) will be run.

Now Level-2 has more detailed detector information and also more time for deciding. It can take roughly a millisecond to decide.

At Level-3, the full event has been assembled. This allows looking for complex signatures, e.g. identifying photons.

The data flow is visualised in figure 1.5.

¹¹weakly-interacting massive particles

¹²Read-out dual-port memory

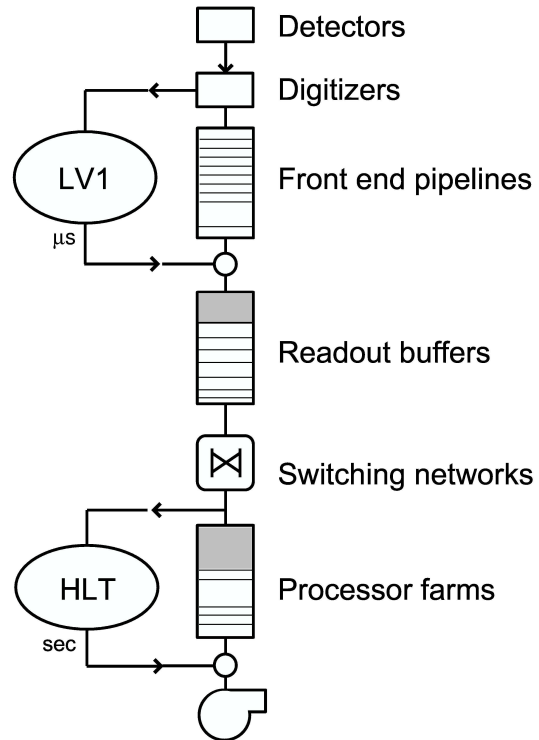


Figure 1.5: The data flow in the CMS Trigger/Data Acquisition system. CMS uses only two physical levels, opposed to the traditional approach using three physical level.

Chapter 2

The Top Quark

In 1964, Gell-Mann and Zweig independently proposed that all hadrons are composed of more elementary particles, which Gell-Mann called quarks. At that time, only three quark flavors were known: up, down and strange.

Ten years later, the groups of Ting and Richter independently discovered the J/ψ particle¹. It is a bound state of the fourth quark, called charm. Actually, it already had been postulated many years earlier by Bjorken and Glashow.

In the summer of 1977, the group of Leon M. Lederman, working at Fermilab, discovered the Upsilon. It was quickly recognised as being the bound state of a fifth quark, called bottom.

At once, particle physicists began searching for the sixth quark, already called top (or truth).

The following is a condensed summary of descriptions given in [Daw], [ChaKon] and [Wil2].

2.1 The Standard Model and its Particles

Although bearing a rather simple name, the Standard Model (SM) of particle physics is a powerful model describing elementary particles and their interactions. Its theoretical concepts are well described in [Gri], [Schm] and [CotGre]. In the following a very brief overview will be given.

Today, we know four types of interactions in Nature. The Standard Model only considers the weak interaction, the electromagnetic interaction and the strong interaction (see table 2.1). Gravitation is described by the General Theory of Relativity, its forces are insignificant at the level of particle physics. The strong interaction is described by QCD², the weak interaction and the electromagnetism have been unified by Glashow, Weinberg and Salam to the electroweak interaction.

There are two kinds of particles: fermions and bosons. Fermions carry a half-integer spin, bosons have an integer spin. In the Standard Model there are two groups of fermions: quarks and leptons. Each of those groups has a substructure of three families (see tables 2.2 and 2.3).

¹This is the only particle to have two names. Ting named it J , Richter called it ψ .

²Quantum Chromo Dynamics

interaction	boson	el. charge	spin	mass [GeV]
weak interaction	$W^\pm, Z$	$\pm 1, 0$	1	80.423, 91.188
electromagnetic interaction	photon	0	1	0
strong interaction	gluons	0	1	0

Table 2.1: The interactions and their bosons (without gravitation). The electric charge is given in units of $|e|$.

first family	second family	third family	isospin T^3	charge Q	hypercharge Y
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\frac{1}{3}$
u_R	c_R	t_R	0	$\frac{2}{3}$	$\frac{4}{3}$
d_R	s_R	b_R	0	$-\frac{1}{3}$	$-\frac{2}{3}$

Table 2.2: The quarks and their quantum numbers.

The mass eigenstates are not equivalent to the flavour eigenstates for electroweak processes and quarks. The mass eigenstates d , s and b and the flavour eigenstates d' , s' and b' are connected according to the unitary CKM³-Matrix, i.e.

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

The 90 % confidence limits on the magnitude of the elements of this complex matrix are [PDG]

$$\begin{pmatrix} 0.9741 & \text{to} & 0.9756 & 0.219 & \text{to} & 0.226 & 0.0025 & \text{to} & 0.0048 \\ 0.219 & \text{to} & 0.226 & 0.9732 & \text{to} & 0.9748 & 0.038 & \text{to} & 0.044 \\ 0.004 & \text{to} & 0.014 & 0.037 & \text{to} & 0.044 & 0.9990 & \text{to} & 0.9993 \end{pmatrix}$$

2.2 Special Features of the Top Quark

The top quark was discovered at CDF⁴, only as recently as in 1995. It is for sure the most interesting of the six known quarks. Because its parameters have

³Cabibbo, Kobayashi, Maskawa

⁴Collider Detector at Fermilab

first family	second family	third family	isospin T^3	charge Q	hypercharge Y
$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L$	$\begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	-1
e_R	μ_R	τ_R	0	-1	-2

Table 2.3: The leptons and their quantum numbers.

not been fully measured, we do not know if it is just a 'normal' Standard Model particle. The most intriguing property of the top is its mass. It is much heavier than all the other quarks. The large mass difference between top (174.3 GeV) and bottom quark (4.0 to 4.5 GeV) causes the top to decay quickly. Actually, its lifetime is so short that the strong interaction is unable to depolarise it.

The LHC will not only be a Higgs factory (if the Higgs boson exists) but also a top factory. This will enable physicists to get hold of all the top parameters and to get a thorough understanding of this sixth quark.

But why is it so important to understand the top and its interaction, especially at the LHC? Because one of LHC's main goals is to discover and precisely measure the parameter of the Higgs boson. There will be top quarks both in some discovery channels and in background. So without a firm knowledge of the processes involving top quarks, there will be severe difficulties to get a certain hold on the Higgs boson. Also, since the mass difference between top and b is very large, it plays an important role in radiative corrections.

2.3 Top Mass

Up to now, only CDF and D0⁵ have measured the top mass directly. Their Run I results from the 'single lepton +jet'-channel are:

	<i>mass</i>			stat. error		syst. error	
CDF	176.1	±	4.8	±	5.3	GeV	
D0	173.3	±	5.6	±	5.5	GeV	

The systematic errors are mainly due to uncertainties in the jet energy scale (CDF: ± 4.4 GeV, D0: ± 4.0 GeV).

The combination of all channels yields (see figure 2.1)

$$M_t = 174.3 \pm 5.1 \text{ GeV.}$$

During Run II both collaborations were able to decrease their statistical and systematical errors. For LHC, errors of less than 1.5 GeV are predicted.

2.4 Decay

The CKM matrix element $|V_{tb}|$, describing the possibility for a top quark to decay into a bottom quark, is only slightly less than 1. Hence, the top decays nearly exclusively via W boson emission into a bottom quark.

CDF Run I measurements yielded the branching ratio

$$R_{tb} = \frac{\Gamma(t \rightarrow Wb)}{\Gamma(t \rightarrow Wq)} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} = 0.94^{+.31}_{-.24}.$$

This result was obtained by comparing $t\bar{t}$ events with one b -tagged jet to $t\bar{t}$ events with two b -tagged jets.

$$R_{tb} = \frac{\text{number of } t\bar{t} \text{ events with 2 } b\text{-tagged jets}}{\text{number of } t\bar{t} \text{ events with at least one } b\text{-tagged jet}}$$

⁵D0 is no acronym. It is the position of the experiment at the ring. CDF has position B0.

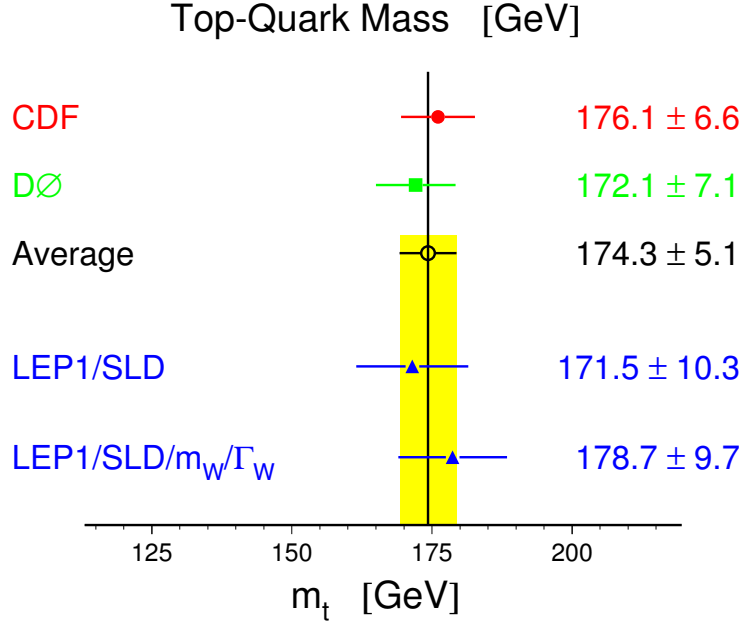


Figure 2.1: Top quark mass measurements at Fermilab and indirect predictions from precision measurements. The errors given are a combination of statistical and systematic errors. [LEPEWWG].

Assuming that there are only three quark families and further assuming unitarity of the CKM matrix, the CKM matrix element is

$$|V_{tb}| = 0.97^{+0.15}_{-0.12}.$$

2.5 Effects of the Short Lifetime

Because of the large mass difference between top and bottom quark, the top lifetime is very short. A simple comparison to the average interaction time of QCD

$$\tau_t \sim 5 \times 10^{-25} \text{s} < \tau_{\text{QCD}} \sim 3 \times 10^{-24} \text{s}$$

shows that there is no time for QCD to take effect on the top. Therefore the top is the only Standard Model quark not to form any bound states.

The strong interaction is also not able to depolarise the top quark. So its spin is observable via the angular distribution of its decay products.

The top quark spin in top pair production is correlated with the spin of the top antiquark. In the “off-diagonal” basis, the correlation is 100% for all energies (see figure 2.2). Instead of the direction of motion, this basis uses another direction, which makes an angle ψ with respect to the beam

$$\tan \psi = \frac{\beta^2 \sin \theta \cos \theta}{1 - \beta^2 \sin^2 \theta},$$

with θ being the scattering angle and β being the velocity of the quarks in the centre-of-momentum frame. [Wil2]

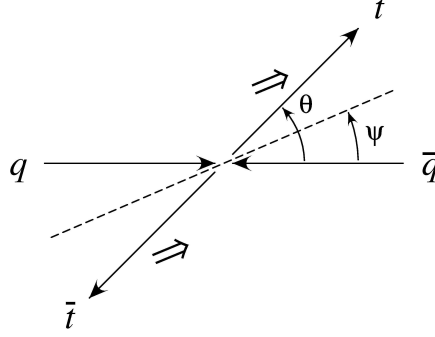


Figure 2.2: The “off-diagonal” basis of a $t\bar{t}$ pair. The spins of the top quark and antiquark point in the same direction. [Wil2]

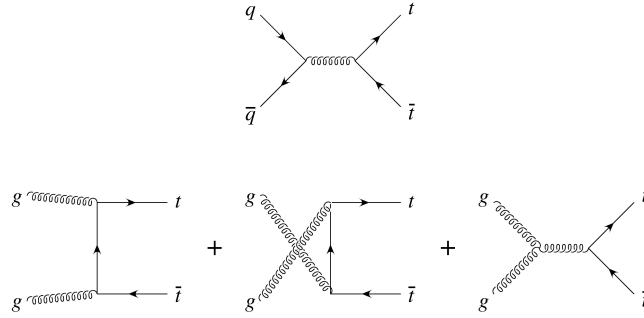


Figure 2.3: Feynman diagrams for top production via strong interaction. The upper diagram shows quark-antiquark annihilation, the lower ones show gluon fusion.

2.6 Top Pair Production

There are two ways to produce top pairs at hadron colliders. These are quark-antiquark and gluon-gluon fusions (see figure 2.3).

In the parton model, the protons are viewed as collections of quarks, antiquarks and gluons, each carrying some fraction, x , of the proton’s four-momentum, P .

The square of the total energy in the centre-of-momentum frame

$$S := (P_1 + P_2)^2 \cong 2P_1 P_2 \quad (2.1)$$

and the square of the total energy $\hat{s}$ of the partonic subprocess are connected via

$$\hat{s} := (x_1 P_1 + x_2 P_2)^2 \cong 2x_1 x_2 P_1 P_2 \stackrel{(2.1)}{=} x_1 x_2 S. \quad (2.2)$$

For top pair production, there has to be sufficient energy to produce a $t\bar{t}$ pair at rest

$$\hat{s} \geq 4m_t^2.$$

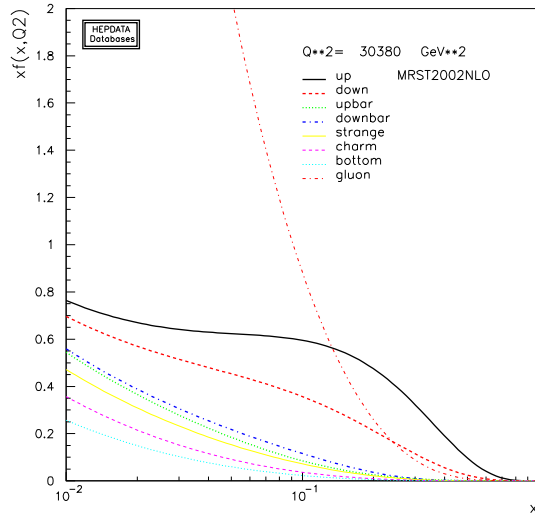


Figure 2.4: Parton distribution function for a proton, which has been generated at [Durham]. Actually, it is not $f(x)$ but $xf(x)$ what is plotted.

This implies

$$x_1 x_2 = \frac{\hat{s}}{S} \geq \frac{4m_t^2}{S}. \quad (2.3)$$

Finding a quark of momentum-fraction x in the proton becomes less probable for increasing X , so normally $x_1 x_2$ is near the threshold for $t\bar{t}$ production. Assuming $x_1 \approx x_2 =: x$ and using equation (2.3) yields

$$x \approx \frac{2m_t}{\sqrt{S}} \quad (2.4)$$

as the typical value of x for $t\bar{t}$ production.

The probability of finding a parton of a given species and carrying a momentum fraction between x and $x + dx$ is $f(x)dx$, with $f(x)$ being the parton distribution function (see figure 2.4) for that parton species. This parton distribution function also depends on the relevant scale μ of the process, which is of order m_t for top quark production.

Equation (2.4) yields a typical value $x \approx 0.18$ for the Tevatron and $x \approx 0.025$ for the LHC. Taking a look at figure 2.4, it turns out that quark-antiquark annihilation reigns at Tevatron (about 85% of top pairs produced in Run II), but gluon fusion prevails at LHC (approximately 90% of the top pairs produced).

The LHC will be a $t\bar{t}$ factory. The $t\bar{t}$ cross section will be [ChaKon]

$$\sigma_{t\bar{t}} = 825^{+58}_{-43} \text{pb}$$

at NLO⁶+NLL⁷, which is about 2.5 times the cross section for single top production (see section 2.8).

⁶Next-To-Leading Order

⁷Next-to-Leading Logarithmic Order

2.7 Top Pair Decay

Because of $|V_{tb}| \approx 1$ almost all top pairs decay via

$$t\bar{t} \rightarrow W^+W^-b\bar{b}.$$

An easy way to classify these events is to take a look at the decay of the W^+W^- pair.

1. Di-lepton events:

Both W bosons decay into $e^\pm$ or $\mu^\pm$. The probability is $2/9 \times 2/9 \approx 4.9\%$. Those events are clean and can be triggered efficiently. One year of LHC at low luminosity ($\int L dt = 20 fb^{-1}$) will deliver about 800.000 of these events.

2. Single lepton plus jet events:

One W boson decays into a $e^\pm$ or to a $\mu^\pm$, the other one into $q\bar{q}$. The probability is $2 \times 2/9 \times 6/9 \approx 29.6\%$. These events are fully reconstructable and have a small background. There will be nearly 5.000.000 of these events in one year of LHC at low luminosity.

3. Multi jet events:

Both W 's decay into jets, yielding a probability of $6/9 \times 6/9 \approx 44.4\%$. Although there are many of these events (7.400.000 for one year of LHC at low luminosity), they are the least interesting top pair decay events: there is a huge background of QCD multi-jet events, making triggering difficult.

Of course there will be di-lepton events with one or both leptons being a tau and single lepton events with the lepton being a muon, but since the tau is much harder to trigger, these have not been included in the previous numbers.

2.8 Single Top Production

There are three channels for producing single top quarks. They are denoted with their corresponding next-to-leading-order in QCD cross section ([SmiWil], [StSuWi], [Zhu]).

- s-channel: $q + \bar{q} \rightarrow t + \bar{b}$ (see figure 2.5(a))
 $\sigma = 10.6\text{pb} \pm 5\%$
- t-channel: $q + b \rightarrow q + t$ (see figure 2.5(b),(c))
 $\sigma = 250\text{pb} \pm 5\%$
- Associated production of W and t : $b + g \rightarrow W + t$ (see figure 2.5(d),(e))
 $\sigma = 75\text{pb} \pm 10\%$

Single top physics is an interesting and important topic for many reasons:

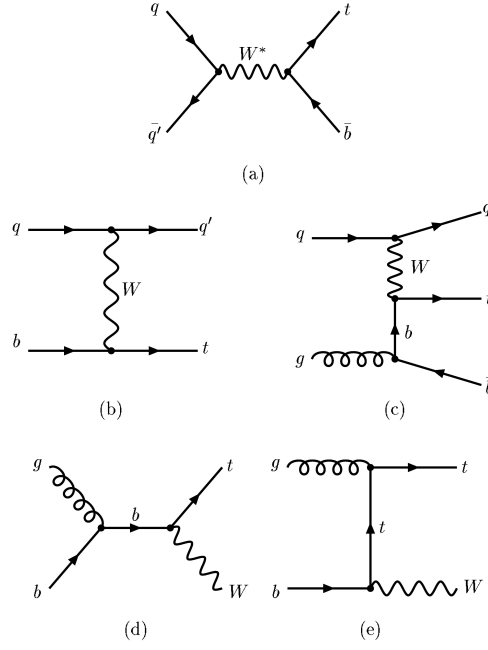


Figure 2.5: Feynman diagrams for electroweak production of single top quarks in leading order. (a) shows the s-channel, (b) and (c) show the t-channel and (d) and (e) show the associated production of W and t .

- In many 'new physics' models there are extra sources for single top production. If there is a measurement of a cross section larger than the Standard Model prediction, this will give strong evidence for physics beyond the Standard Model.
- At the LHC, measuring $|V_{tb}|$ directly will be possible for the very first time. Both $t\bar{t}$ and single top events have to be measured. A precision of 5% is expected. The procedure is explained in detail in [ChaKon].
- Having a polarisation of nearly 100% after the production, top quarks can be used to test the V-A-structure of the charged currents in the electroweak theory.

2.9 Top Quark, W Boson and Higgs Boson

One of the most interesting subjects in radiative corrections is the dependency of the W boson mass on the masses of Z boson, top quark and Higgs boson

$$M_W^2 = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{\sin^2 \Theta_W (1 - \Delta r)}$$

with $\sin^2 \Theta_W := 1 - \frac{M_W^2}{M_Z^2}$ ($= 0.2228(4)$), α being the fine structure constant and G_F being the Fermi coupling constant. Δr represents 1-loop-corrections.

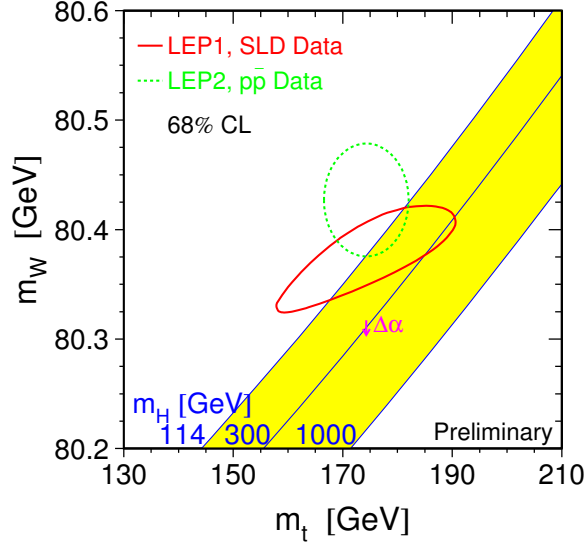


Figure 2.6: The prediction of top quark mass and the W boson mass from electroweak precision measurements at 68 % confidence level (green, dashed ellipse), using SLD and LEP-2 data. Also shown are the results of the direct measurement of both masses at Tevatron and LEP-2. The yellow band shows the constraint between the two masses within the Standard Model, depending on the Higgs boson mass, and also, but only to a small extent, on the hadronic vacuum polarisation (small arrow labelled $\Delta\alpha$). [Grü], [LEPEWWG]

The corrections for top quark and Higgs boson are [Wil2]

$$(\Delta r)_{\text{top}} \approx -\frac{3 G_F m_t^2}{8\sqrt{2}\pi^2} \frac{1}{\tan^2 \Theta_W}$$

and

$$(\Delta r)_{\text{Higgs}} \approx \frac{11 G_F M_Z^2 \cos^2 \Theta_W}{24\sqrt{2}\pi^2} \ln \frac{m_h^2}{M_Z^2},$$

where m_t is the top quark mass, m_H is the Higgs boson mass. Being logarithmically dependent, the Higgs mass correction is much smaller than the top mass correction.

By combining these formulas and LEP and SLD⁸ measurements, a well-known plot (see figure 2.6) can be obtained. Both contour curves prefer a low value for the mass of the Higgs Boson in the Standard Model.

If the Higgs boson exists, its mass will surely be measured at LHC. This result will be a very good test on the current knowledge of electroweak corrections.

⁸SLAC Large Detector, SLAC is the Stanford Linear Accelerator

Chapter 3

Jet Algorithms

Jets are collections of objects considered to be very close to each other and to originate from a common source according to some distance measure.

A jet is defined via its distance measure and its recombination algorithm. The distance measure determines the group of objects which are in the jet, the recombination algorithm constructs kinematical variables of the jets based on the momenta of the objects.

Traditionally, the cone algorithms are used at hadron colliders. Today, also the k_T algorithm is used.

For a good overview on jet algorithms, see [KtJet] and [Kun]. The following sections are a condensed summary of these two references.

3.1 Iterative Cone Jet Algorithm

Describing the kinematics in terms of rapidities and azimuthal angles, the two-dimensional $\eta \times \phi$ lego plot of the energy depositions is analysed (see figure 3.1).

1. The particle with maximal transverse energy from the particle list¹ is searched. Its direction (η^C, ϕ^C) is used as a seed.
2. A trial cone with radius of R is drawn in the η - ϕ plane around the direction (η^C, ϕ^C) of the seed. Now for all particles within the cone the following condition holds:

$$\sqrt{(\eta^C - \eta_i)^2 + (\phi^C - \phi_i)^2} \leq R.$$

3. The trial cone is considered to be stable if the 'physical' centre of the cone defined by the recombination algorithm

$$\eta_T^R = \frac{\sum_i \eta_T^i E_T^i}{\sum_i E_T^i}, \quad \phi_T^R = \frac{\sum_i \phi_T^i E_T^i}{\sum_i E_T^i}$$

¹One can also use a list of calorimeter energies or tracks.

coincides with its geometrical centre (η^C, ϕ^C) . In order to decide on this, the distance $\Delta R = \sqrt{(\eta^C - \eta^R)^2 + (\phi^C - \phi^R)^2}$ in η - ϕ plane is calculated.

4. • If ΔR is larger than a given value, usually 0.01, the trial cone is unstable. The physical centre of the cone is used as a new seed

$$(\eta^C, \phi^C) := \left(\frac{\sum_i \eta_T^i E_T^i}{\sum_i E_T^i}, \frac{\sum_i \phi_T^i E_T^i}{\sum_i E_T^i} \right).$$

Now go back to 2.

- If ΔR is less than the given value, the trial cone is stable. The kinematical variables of the jet can be constructed by using a recombination scheme, e.g.

$$E_{x,y,z}^J = \sum_i E_{x,y,z}^i$$

$$\Theta^J = \arctan \frac{E_T^J}{E_z^J}, \quad \phi^J = \arctan \frac{E_x^J}{E_y^J},$$

$$\eta^J = -\log \left(\tan \frac{\Theta^J}{2} \right), \quad E_T^J = E^J \sin \Theta^J.$$

The particles in the cone are taken from the particle list and the cone is stored in the jet list. If there are still particles on the particle list, go to 1. Else the algorithm ends.

3.2 k_T Algorithm

The main idea for the k_T algorithm was to avoid jet overlaps.

3.2.1 Inclusive Mode

Starting with the initial list of particles and an empty list of jets, the algorithm assigns the particles to the jets in such a way that in the end we have an empty list of particles and a filled list of jets.

The regrouping procedure of the particles into the jets is done iteratively in the following steps.

1. Every particle i and every particle pair (i, j) have a distance value, d ,

$$d_i = p_{i,T}^2 \quad d_{ij} = \min(p_{i,T}^2, p_{j,T}^2) \frac{\Delta R_{ij}^2}{D^2},$$

with D being a free parameter, usually set to 1.0.

2. Calculate $d_{min} = \min_{i,j}(d_i, d_{ij})$.
3. • If $d_{min} = d_{ij}$, the particle pair (i, j) will be replaced with a pseudoparticle², combining the particle pair (i, j) according to a user specified recombination scheme.

²which is from now on treated as a particle

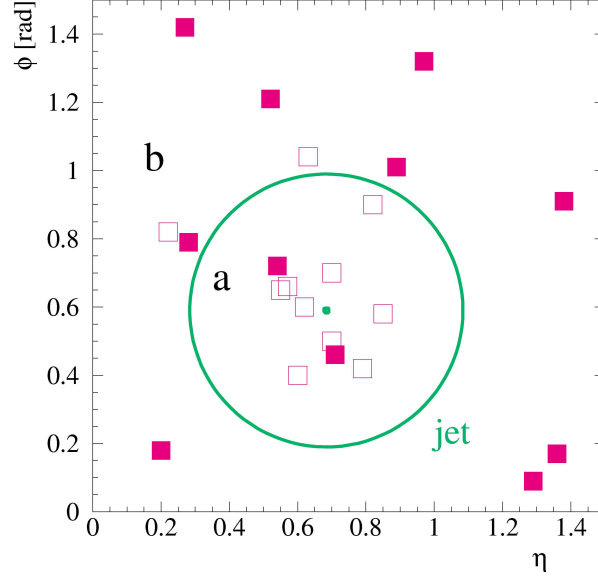


Figure 3.1: A jet (circle) with a cone size of 0.4 in the η - ϕ -plane. The empty boxes represent energy depositions, the full boxes energy from other sources. The η - ϕ -plane is divided into two regions: a (inside the jet) and b (outside the jet). [Dro2]

- If $d_{min} = d_i$, the particle i will be removed from the list of particles and will be added to the list of jets.
4. If the particle list is non-empty, the algorithm starts over from 1 again.

3.2.2 Exclusive Mode

The algorithm separates the 'hard final states' from the 'soft beam remnants' explicitly in this mode.

The definition of the hard final state is done using a stopping parameter, d_{cut} .

1. d_k and d_{kl} are defined in the same way as in section 3.2.1.
2. Calculate $d_{min} = \min_{i,j}(d_i, d_{ij})$.
3.
 - If $d_{min} < d_{cut}$, all remaining objects will be classified as jets. The algorithm is finished for this event.
 - If $d_{min} \geq d_{cut}$:
 - If $d_{min} = d_{ij}$, the particle pair (i, j) will be replaced with a pseudoparticle. This pseudoparticle is constructed using the recombination scheme.
 - If $d_{min} = d_i$, the particle i will be removed from the list of particles and will be included in a 'beam jet'.

4. If the particle list is non-empty, the algorithm starts over from 1 again.

There is another possibility for the use of the exclusive mode. One can choose to stop merging when a given number n of jets is reached instead of using a stopping scale d_{cut} .

3.3 Jet Algorithm Implementation in ORCA

In the CMS reconstruction software ORCA (see page 28), the jet algorithms mentioned above are implemented. Also a simple cone algorithm, using no iterative process, exists. The k_T jet implementation in ORCA uses the KtJet package by J. M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh [KtJet].

One can apply these algorithms to Monte Carlo (MC) data, to ECAL and HCAL towers and to tracks. Soon a new class will be implemented, allowing to apply the algorithms to a combination of tracks with calorimetric information. [Hei]

Although using ORCA version 7.2.4, the jet algorithm were taken from a later version, ORCA 7.4, for this thesis.

3.4 Using Jet Algorithms

For a thorough and careful analysis, one has not only to consider which jet algorithm to use, but also several other aspects. These include preclustering and the role of the underlying events. Also effects from merging and splitting can have some impact on the result of the analysis, e.g. a smaller cone size can split a jet into two jets.

Chapter 4

Production and Analysis Tools

Analysing high energy physics at a computer needs interplay of several programs. Most of them are subject to continuing improvements and extensions.

During the course of this thesis, important contributions to a higher-level analysis software called PAX (see section 4.3) were made.

4.1 Software

The software used for this diploma thesis will be listed and explained in the following subsections.

PYTHIA

The PYTHIA/JETSET package consists of the two Fortran programs PYTHIA and JETSET. Being used together in most cases, they are normally referred to as PYTHIA¹.

PYTHIA assumes jet universality. Jet universality means that fragmentation is fundamentally the same for an e^+e^- or a pp event. The only difference depends on the parton-level processes involved. Furthermore, PYTHIA can generate collisions between leptons, hadrons and photons.

The fragmentation is implemented using a phenomenological model of string fragmentation. The model used by default is the “Lund model”.

Emission of strongly interacting particles off quarks and gluons is very prolific. A single initial parton can easily give rise to a whole bunch of partons in the final state. Photon emission may cause considerable effects, too. Most of the bremsstrahlung corrections depend only on a few key numbers, e.g. on the momentum transfer scale of the process. The short distance interactions of quarks, leptons and gauge bosons are described by a perturbative approach. The branching ratios for decays of fundamental resonances ($Z^0, W^\pm$) are calculated dynamically. For decays which are not isotropic, the matrix elements, if available, are taken into account.

The chronological order of the calculations PYTHIA performs is as follows: in the beginning, one shower initiator from each beam starts off a sequence of

¹this will be done in this document from now on

branchings, which build up an initial-state shower. Then one incoming parton from each of the two showers enters the hard process. There the outgoing partons, usually two, are produced. The outgoing partons can branch, too, building up final-state showers. After taking a shower initiator out of a beam particle, the beam remnant is left behind. The non-observability of outgoing quarks and gluons is guaranteed by QCD confinement mechanism. Quarks and gluons fragment to colour neutral hadrons. The remaining instable hadrons decay further, unless this decay has been switched off.

An example of PYTHIA output is shown in Table 4.1. All history information on the particles is given in this output. The particles are numbered consecutively (column I), their name is given (column `particle/jet`), their status² is given (column KS), a particle identification is given (column KF), the line number of the parent particle or jet is given (column `orig`), the four-vector and the mass of the particle are given (columns `p_x`, `p_y`, `p_z` and `E`). Additional information can be found in an unnamed column between status and particle identification. Particles with an A, an I or a V belong to the same string parton system³. [Gui], [PYTHIA]

CMKIN

CMKIN is a CMS-specific wrapper package for several different event generators, e.g. Herwig [HERWIG], PYTHIA, Isajet [ISAJET] and CompHep [CompHep].

It offers a standard way to interface those generators with GEANT in CMSIM or OSCAR⁴, which is going to be the successor of CMSIM. In order to do so it uses a common block HEPEVT⁵ which is a standard format to store particle kinematics information for one event. In the end, the HEPEVT common block is converted to HBOOK ntuples for input/output. This allows events to be analysed using PAW⁶.

CMSIM and GEANT

The event generator describes the particle collisions. We further need to know how the particles interact with the detector.

The CMSIM package is an application of GEANT 3. GEANT 3 is a general system of detector description and simulation tools. It simulates the particle transport through matter, i.e through the detector. Today, GEANT is not only used in High Energy Physics, but also in medical and biological sciences or astronautics to describe the transport of particles and energy deposits matter.

In the end, results are stored in ZEBRA⁷ banks. Interfaces in GEANT allow

²Status: stable 1, decayed particle or fragmented jet 11, fragmented jet followed by more jets in the same colour-singlet jet 12, documentation line used to give a compressed story of the event at the beginning of the event-record 21

³A, I and V give a “poor man” representation of an arrow.

⁴Object oriented Simulation for CMS Analysis and Reconstruction

⁵High Energy Physics event

⁶or ROOT, if one converts the HBOOK ntuple to a ROOT ntuple via h2root.

⁷The ZEBRA is an extension to the FORTRAN programming language, designed to provide techniques for dynamic storage and data handling.

Event listing (summary)									
I	particle/jet	KS	KF	orig	p_x	p_y	p_z	E	m
1	!p+!	21	2212	0	0.000	0.000	7000.000	7000.000	0.938
2	!p+!	21	2212	0	0.000	0.000	-7000.000	7000.000	0.938
=====									
3	!ubar!	21	-2	1	0.467	-0.521	100.921	100.923	0.000
4	!g!	21	21	2	-1.226	-1.401	-964.599	964.601	0.000
5	!ubar!	21	-2	3	2.768	1.493	91.653	91.707	0.000
6	!u!	21	2	4	-1.545	-2.057	-164.363	164.383	0.000
7	!Z0!	21	23	0	-23.125	40.410	-105.636	148.632	93.620
8	!Z0!	21	23	0	24.349	-40.974	32.926	107.458	90.506
9	!tau-!	21	15	7	-50.329	-3.662	-36.851	62.511	1.777
10	!tau+!	21	-15	7	27.204	44.072	-68.784	86.121	1.777
11	!tau-!	21	15	8	28.544	21.993	24.859	43.813	1.777
12	!tau+!	21	-15	8	-4.195	-62.967	8.067	63.645	1.777
=====									
13	(Z0)	11	23	7	-23.125	40.410	-105.636	148.632	93.620
14	(Z0)	11	23	8	24.349	-40.974	32.926	107.458	90.506
15	(tau-)	11	15	11	28.544	21.993	24.859	43.813	1.777
16	(tau+)	11	-15	12	-4.195	-62.967	8.067	63.645	1.777
17	(tau+)	11	-15	10	27.204	44.072	-68.784	86.121	1.777
18	(tau-)	11	15	9	-50.328	-3.662	-36.851	62.510	1.777
19	gamma	1	22	9	0.000	0.000	0.000	0.000	0.000
20	p+	1	2212	1	0.625	0.474	1881.693	1881.694	0.938
21	(ubar)	A 12	-2	4	0.003	0.075	-51.776	51.777	0.330
22	(g)	I 12	21	4	0.448	-0.478	-11.437	11.456	0.000
23	(g)	I 12	21	4	-1.069	0.111	-43.486	43.500	0.000
24	(u)	V 11	2	2	0.743	0.627	-247.747	247.750	0.330
25	(d)	A 12	1	0	-1.008	-4.600	89.792	89.916	0.000
26	(g)	I 12	21	0	-3.302	-2.689	34.426	34.688	0.000
27	(g)	I 12	21	0	-1.748	-3.385	13.461	13.990	0.000
28	(dbar)	V 11	-1	0	1.008	4.600	4.643	6.613	0.000
29	(u)	A 12	2	0	-2.516	-2.891	-6.926	7.916	0.000
30	(g)	I 12	21	0	3.087	0.709	-1.011	3.325	0.000
31	(ubar)	V 11	-2	0	2.516	2.891	-2.430	4.538	0.000
32	(d)	A 12	1	0	1.295	3.476	7.792	8.630	0.000
33	(dbar)	V 11	-1	0	-1.295	-3.476	2833.715	2833.717	0.000
34	(u)	A 12	2	0	-3.442	0.073	62.811	62.906	0.000
35	(ubar)	V 11	-2	0	3.442	-0.073	6.737	7.566	0.000
36	(u)	A 12	2	0	1.159	3.165	17.688	18.006	0.000
37	(g)	I 12	21	0	-3.087	-0.709	0.956	3.309	0.000
38	(ubar)	V 11	-2	0	-1.159	-3.165	-1.628	3.744	0.000
39	(u)	A 12	2	1	-1.093	0.047	1639.492	1639.493	0.330
40	(g)	I 12	21	3	-2.316	-2.025	8.555	9.091	0.000
41	(g)	I 12	21	0	-7.292	1.252	221.973	222.096	0.000
42	(g)	I 12	21	0	7.292	-1.252	52.248	52.769	0.000

Table 4.1: An extract from a typical PYTHIA output

the users to access data structures without knowing details of the ZEBRA memory management system. [CMSIM], [GEANT]

SCRAM

In order to enable groups around the world to work together on complicated software projects like the CMS software, SCRAM (Software Configuration, Release And Management software) was developed. SCRAM has many functionalities, especially configuration and resource management.

SCRAM has been designed to make sure that all developers use the same consistent set of libraries, software environment, source codes and external products. [SCRAM]

COBRA

COBRA is the Coherent Object-oriented Base Reconstruction and Analysis software. It is a general framework with the purpose to provide object-oriented tools to ORCA and OSCAR.

ORCA

ORCA is an acronym for the Object Oriented Reconstruction for CMS Analysis. This software framework is used for reconstruction studies and global detector performance evaluations.

ORCA is based on CARF⁸. This framework implements two principles: event driven notification and action on demand. The latter saves CPU time by ensuring that objects are reconstructed only if necessary and only once, e.g. detector hits are only reconstructed if an algorithm asks for them. Event driven notification means that so-called observers are notified when a new event arrives. Now these observers take appropriate actions, e.g. the analysis is started.

ORCA is intended to be used for final detector optimisations, global detector performance evaluation and trigger studies. It has many subsystems, e.g. Calorimetry, Tracker, Vertex Reconstruction and Jet Finders. [ORCA]

ROOT

ROOT⁹ [ROOT] is an object-oriented interactive data analysis system, written in C++, succeeding PAW, which was written in FORTRAN. ROOT provides a basic framework for data analysis, e.g. histogramming, curve fitting and minimisation.

One of the special features of ROOT is the built-in CINT interpreter, which allows users to interactively analyse data without compiling/linking. [ROOT]

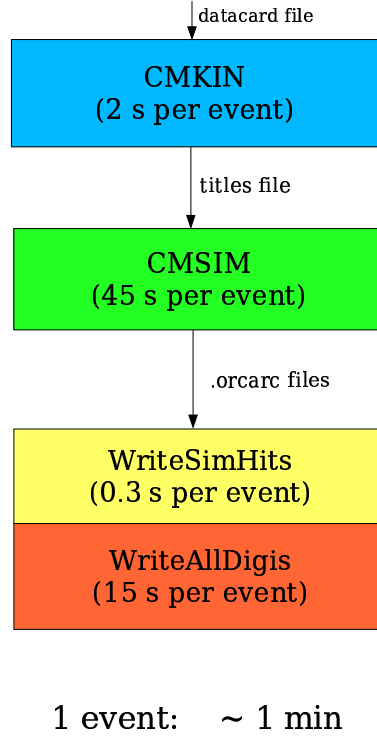


Figure 4.1: A simple visualisation of the production chain used. The times noted are approximations of the CPU times the $t\bar{t}$ -sample production needed at the GridKa. The production ran on Xeon and Pentium III processors. The versions of the software used are given in table 4.2.

4.2 Production of $t\bar{t}$ Samples

At first, the physical process itself is simulated with PYTHIA. PYTHIA allows the user to choose the physical process he/she is interested in. There are two ways of choosing a specific sub-process. The first one is to switch off unwanted decays, e.g. the user lets the W boson decay exclusively to $\mu + \nu_\mu$ and nothing else. Unfortunately, this is of no use in the semileptonic $t\bar{t}$ channel, because we want one W boson to decay to quarks and one to $\mu + \nu_\mu$. A CMKIN 'datacard', defining the physical process, is shown in table 4.3.

In this case a FORTRAN routine called KIS_USER is used. It has to be written by the user, who defines criteria the event and its particles must fulfil. In the end, PYTHIA just selects the events satisfying these criteria and discards the others.

Now the interactions of the particles in the event with the detector hardware are simulated by CMSIM. The ZEBRA banks written out by CMSIM have to be converted to an ORCA format by using an ORCA program called writeHits. Then the ORCA program writeDigis calculates the detector response to the energies deposited in the detector. It also allows to add pile-up events, which has not been done for this thesis.

Having produced a complete database, the user can start to analyse his data. Since ORCA programs are rather slow, it is a good approach to write out all the relevant data into ROOT trees, allowing to access them much faster. This was done for this thesis. A visualisation of the production chain is given in figure 4.1.

Producing a large dataset of completely simulated events is a time-consuming job. About 150 days of CPU time was needed to produce the 216,00 events for this thesis. Also other problems occur, e.g. programs crashing, computers not running or insufficient local disk space on the computing nodes. Of course, such a dataset needs much disk space: 1.3 GB for the Ntuple files, 240 GB for the ZEBRA banks and 230 GB for the ORCA databases.

⁸CMS Analysis and Reconstruction Framework

⁹ROOT is astonishingly not an acronym. The meaning of its name is that it gives a solid root on which other systems can grow.

software	version
PYTHIA	6.215
CMKIN	1.3.0
CMSIM	133
ORCA	7.2.4

Table 4.2: Versions of the software used for $t\bar{t}$ production

```

MRPY 1= 0001
KSEL 0
CFIL 'EVT0', 'ttbar_semileptonic.ntpl'
MSEL = 6
MSTJ 11 = 3
MSTP 51 = 4
MSTP 81 = 1
MSTP 82 = 4

```

Table 4.3: The CMKIN datacard used for producing the $t\bar{t}$ data sample. The most important commands are KSEL 0 (calling KIS_USER) and MSEL=6 (choosing $t\bar{t}$ production using massive matrix elements). The MSTJ command and the MSTP commands choose the fragmentation function, the multiple parton interactions and the multi-parton model. MRPY 1 sets the seed for the random number generator.

4.3 PAX

PAX (for an extensive description see [PAX]), the Physics Analysis Expert toolkit, is a data analysis utility working on a high level of abstraction.

4.3.1 PAX features

PAX accesses data via interfaces to the detector reconstruction or the Monte Carlo generator. This allows to protect the physics analysis code against changes in the detector reconstruction layer. Furthermore, it enables the user to run an analysis program on both Monte Carlo data and on reconstructed data without much extra effort.

One of the main premises of PAX is to give the user full control of every step of the analysis process. This work is facilitated by a simple and intuitive programming interface. Supporting modular physics analysis structures, PAX supports team work.

Using PAX does not force the user to use PAX objects only for his analysis, but it allows the user to access the original data of the experiment in order to use experiment-dependent information and related methods.

The main idea of PAX is to use event interpretations, i.e. there can be different interpretations of one event, e.g. if there are several possibilities to reconstruct a W boson from jets. The user can do the analysis on all the event interpretations in one step¹⁰. Therefore, he does not have to impose strict cuts the event has to fulfil but may at the end of his analysis take a look at all the possible event interpretations.

An event interpretation is a container for storing all the information describing one event (see figure 4.2). It contains four-vectors, vertices, collisions, their relations, and additional values needed in user's analysis. These physics quantities of course have relations, which are handled by a relation manager

¹⁰Of course, he is still restricted by CPU time and memory.

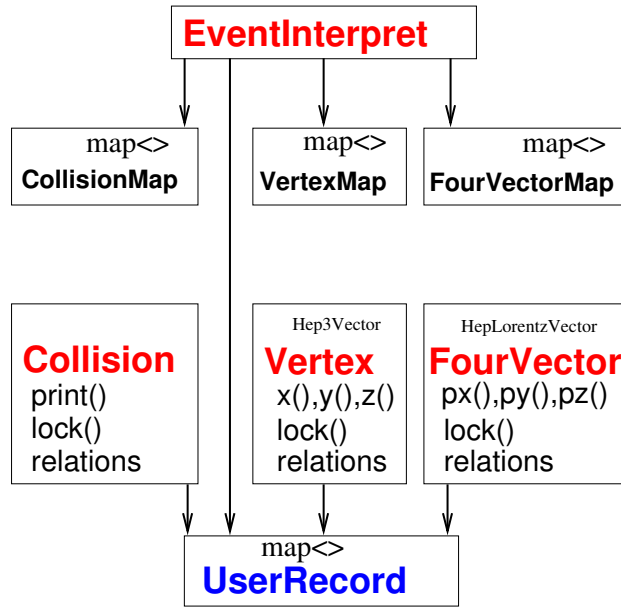


Figure 4.2: This schematic overview shows a PAX event interpretation, *Pax-EventInterpret*, together with the classes for collisions, four-vectors and user records.

(see figure 4.3). Hence the user can easily access the four-vectors which originate from a certain vertex or the connections between multiple collisions and their vertices.

An important feature of the relation manager is the possibility to 'lock' away certain parts of the event from the analysis. When excluding a certain vertex from an event interpret, all of its outgoing four-vectors and so on are excluded as well.

In order to use the advantages of object-orientation, PAX is coded in C++. This facilitates combining PAX with the reconstruction software of modern physics experiments, which are usually written in C++ as well.

4.3.2 The PAX n-tuple Filler

The class `PaxNTPLFill` allows to fill PAX event interprets from Monte Carlo data stored in NTUPLE files.

At first, this class converts the NTUPLE file into a ROOT file, which is then read out. The advantages of using ROOT files are their small file sizes¹¹ and the higher rate for accessing data from the file (more than a factor of two faster).

4.3.3 The PAX Combined Objects Class

The PAX combined objects class is an algorithm for energy measurement. It aims at improving energy measurement and reducing the need for energy correc-

¹¹a ROOT file has about half the size of a NTUPLE file

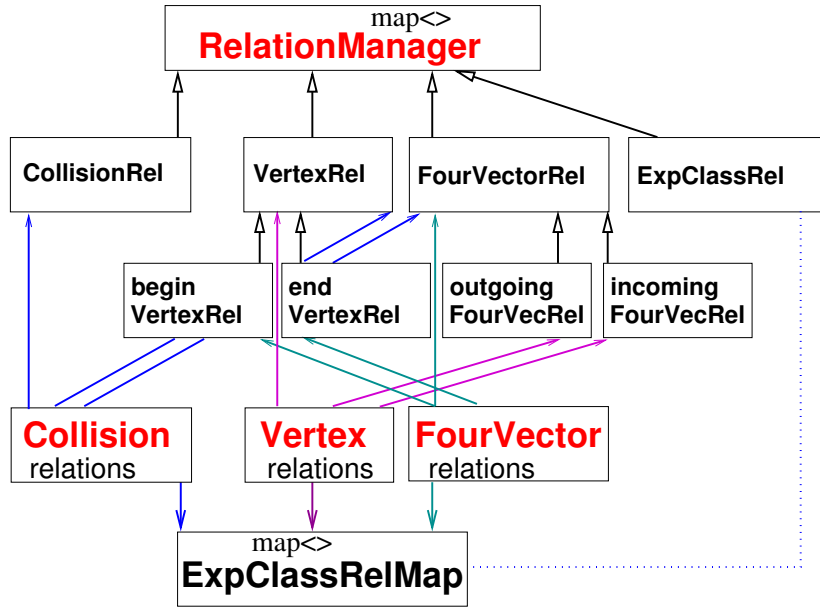


Figure 4.3: The PAX relation management allows access to experiment specific classes, to the decay trees and to the records of analysis history. The user also can exclude parts of the event from the analysis.

tions afterwards. This is done by using both calorimeter information and track information. A detailed description and benchmarks for the CDF environment will be given in [Kir].

Every experiment-specific class inherits from the abstract base class, *Pax-CombinedObjects*. The purely virtual experiment-specific functions have to be reimplemented.

The algorithm works in the following way:

1. The algorithm has to be given four event interprets as input: one containing calorimeter information, one containing track information, an empty one in which the extrapolated tracks will be stored, and one in which the final combined objects information will be stored.
2. All the four-vectors in the track event interpret which have enough p_T to reach the calorimeter surface are extrapolated to the surface of the calorimeter and they are stored in the event interpret for extrapolated tracks.
3. For each track, a check is done whether it fulfils so-called good track criteria. These are experiment-specific, e.g. they can depend on the p_T .
4. Then the algorithm iterates over all good extrapolated four-vectors in a specific sequence, e.g. starting with the highest p_t and ending with the lowest p_t . Now the actual combination starts:

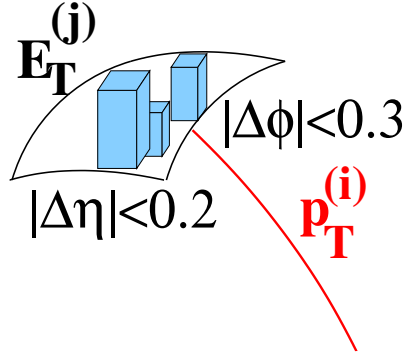


Figure 4.4: A track and the calorimeter fulfilling the distance criterion.

- (a) Sum up all calorimeter four-vector energies which fulfil a certain distance criterion to the track four-vector, e.g. $|\Delta\phi| < 0.3$ and $|\Delta\eta| < 0.2$ (see figure 4.4).
- (b)
 - If the sum of the transverse energies of the calorimeter towers within the given distance is larger than the transverse momentum of the track,

$$\sum_{\text{in distance}} E_{T,calo} > p_{T,track},$$

or the track is not a good track, loop over calorimeter four-vectors. This sequence is experiment-specific, e.g. starting with the four-vector which has the E_T closest to the track p_T .

- i. If the E_T of the calorimeter four-vector is smaller than the p_T of the track four-vector, copy the calorimeter four-vector into the combined objects event interpret, lock the calorimeter four-vector and store the remaining track p_T in the user record of the track. Go to the next calorimeter four-vector.
- ii. If the E_T of the calorimeter four-vector is larger than the p_T of the track four-vector, copy the calorimeter four-vector into the combined objects event interpret and store the remaining energy $E_t - p_T$ of the calorimeter in the user record of the calorimeter tower. Go to the next track.
- If $\sum_{\text{in distance}} E_{T,calo} \leq p_{T,track}$, loop over calorimeter four-vectors. Copy each track four-vector into the combined object event interpret. Lock each calorimeter four-vector of the loop. Store the remaining transverse momentum $p_T - \sum_{\text{in distance}} E_{T,calo}$ of the track in the user record of the track.

5. Copy each remaining calorimeter four-vector tower and each 'good' track four-vector, which is in forward direction or in the central region, but too weak to reach the calorimeter, into the combined objects event interpret.

The combined objects can now be used as an input for jet algorithms. The combined objects algorithm yields a better reconstruction of the direction and of the energy of the jets.

For this thesis, the PAX Combined Objects Class has not been used. This will be done in a later study, hopefully yielding better results.

Chapter 5

The W Mass Reconstruction

For this analysis (and for further analyses) a dataset of about 216,000 completely simulated $t\bar{t} \rightarrow \mu + jets$ events has been produced. Half of them were used for deriving the energy corrections, the rest for analysing the performance of the procedure.

5.1 Event Topology

In case of a semi-leptonic $t\bar{t}$ decay, both top quarks decay via W emission to b quarks. Further, one of the two W bosons decays into two quarks, while the other one decays into a charged lepton and a neutrino. In the end, there are 4 jets, a charged lepton and missing energy (see also figure 5.1).

In the data sample produced for this study, the lepton was chosen to be a muon.

5.2 Reconstruction using Monte Carlo Knowledge

In this section, the data sample is analysed using the iterative cone algorithm with a cone size of 0.5 and a seed E_T cut of 2.0 GeV. Trying to reconstruct the W mass from reconstructed data, we see the necessity for using jet energy corrections. An energy correction scheme, depending on twelve parameters in the barrel range of the detector and twelve parameters in the endcap range of the detector, respectively, is introduced and its parameters are adjusted to the data sample and the jet algorithm.

In the following, 'Monte Carlo data' denotes the particles of the hard process produced by PYTHIA, 'generator data'¹ denotes the particles produced by PYTHIA in the hadronization process and 'reconstructed data' denotes the calorimetric data reconstructed within the trigger code by ORCA.

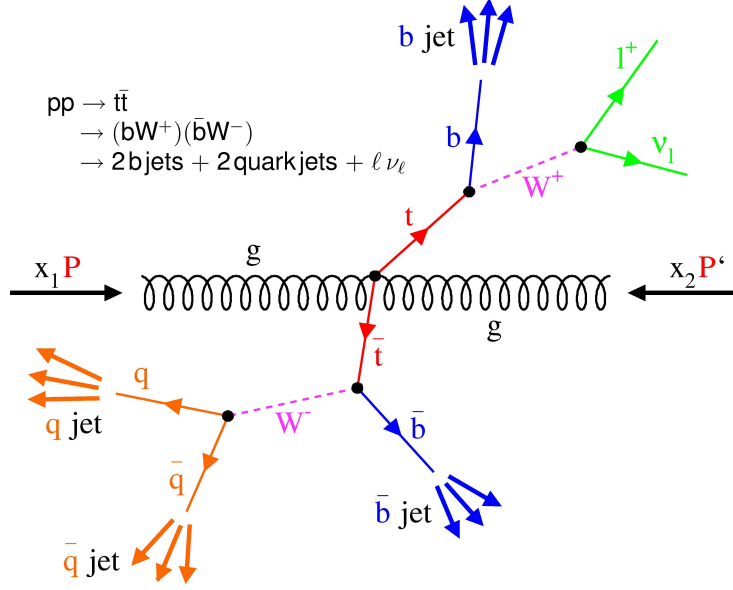


Figure 5.1: A sketch of a semileptonic $t\bar{t}$ decay. The top quark decays via emission of W^+ into a b quark, with the W^+ decaying to a charged anti-lepton and a neutrino. The anti top quark decays via W^- emission into a $\bar{b}$ quark, with the W^- decaying to a quark and an anti quark. The final signature is four jets, a charged lepton and missing energy. [Son]

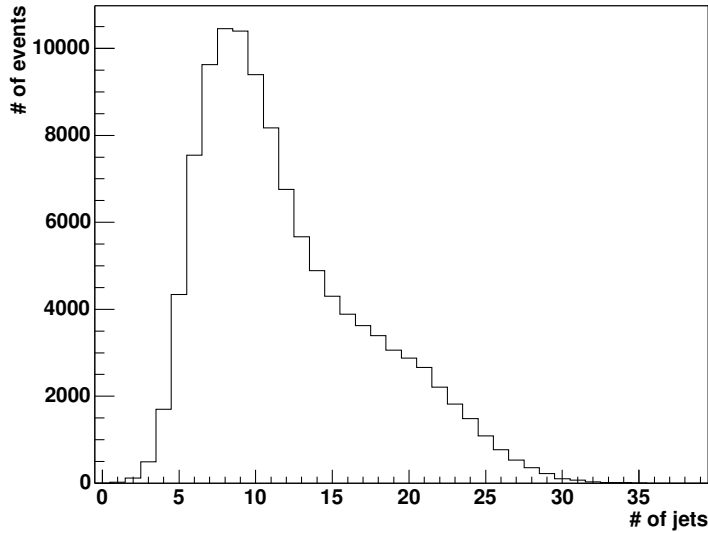


Figure 5.2: Number of generator jets in a $t\bar{t}$ event with $t \rightarrow Wb$, $W \rightarrow q\bar{q}'$, $\mu\nu_\mu$, using an iterative cone algorithm with a cone size of 0.5. The mean number of jets is 12.1, the RMS is 5.5 and the number of events is 112,100.

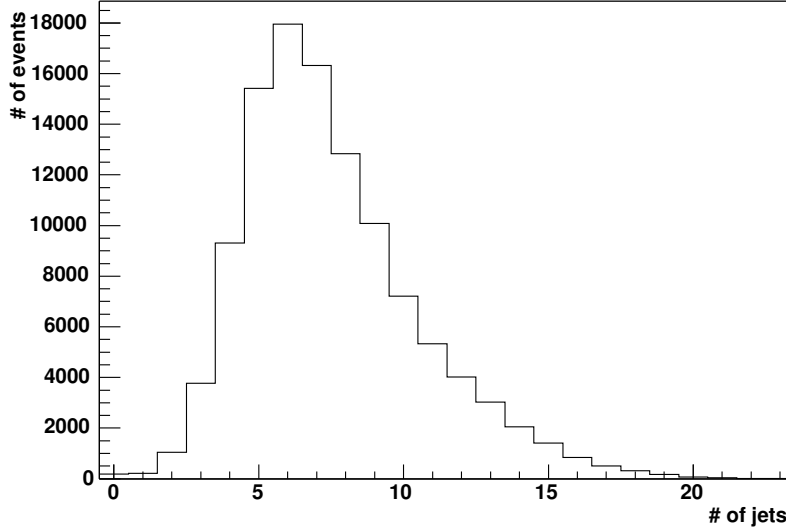


Figure 5.3: Number of reconstructed jets using an iterative cone algorithm with a cone size of 0.5. The mean number of jets is 7.534, the RMS is 3.047 and the number of events is 112,100.

5.2.1 Jets at Generator and Reconstruction Level

In the semi-leptonic decay of $t\bar{t}$ one expects at least five jets to be found: two b -jets, two jets coming from a W , and a 'muon-jet', i.e. the algorithm takes the muon to be a jet.

On average, there are twelve jets at generator level (see figure 5.2). There are also a few events with less than five jets. A reason is that there were no eta cuts applied when producing the data samples. So it can happen that one or several of the jets and/or the muon are outside the detector range. Also, this can be caused by overlapping jets, or by the muon having the same direction as a real jet.

There are many events with more than five jets. This is mainly due to the rather small cone size, leading to splitting of the jets, and to final state radiation.

Next, the algorithm is used to analyse the reconstructed data. Cone size and seed E_T cut are the same as at generator level, in addition a tower E_T threshold of 0.5 GeV is used, suppressing noise in the calorimeters. There are 7.5 jets on average (see figure 5.3).

Obviously, there are fewer jets at reconstruction level than at generator level. This is due to the calorimeter smearing.

The number of jets on generator level and the number of jets at reconstruction level are correlated, as shown in figure 5.4.

¹We consider only those particles, which are within the detector range of $|\eta| < 2.5$ and which are not neutrinos, since these cannot be detected.

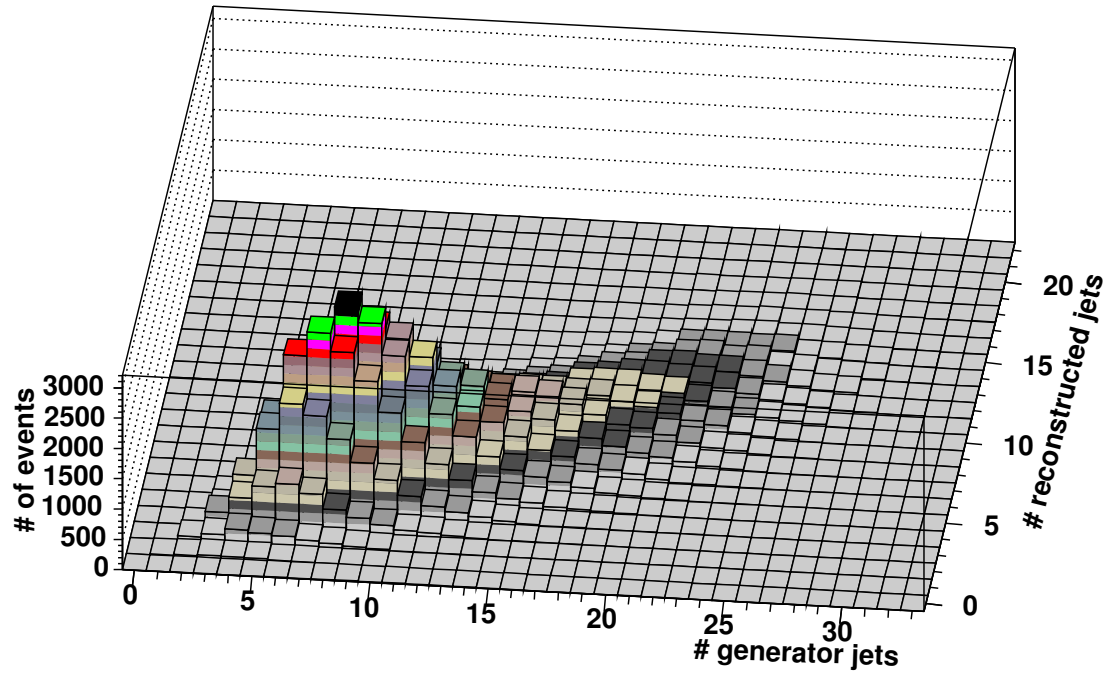


Figure 5.4: A lego plot of the number of generator jets and reconstructed jets for the iterative cone algorithm with a cone size of 0.5. The means and the RMS's for the axis are the same as in figure 5.2 and figure 5.3, respectively. The covariance is 13.35, the correlation coefficient is 0.79.

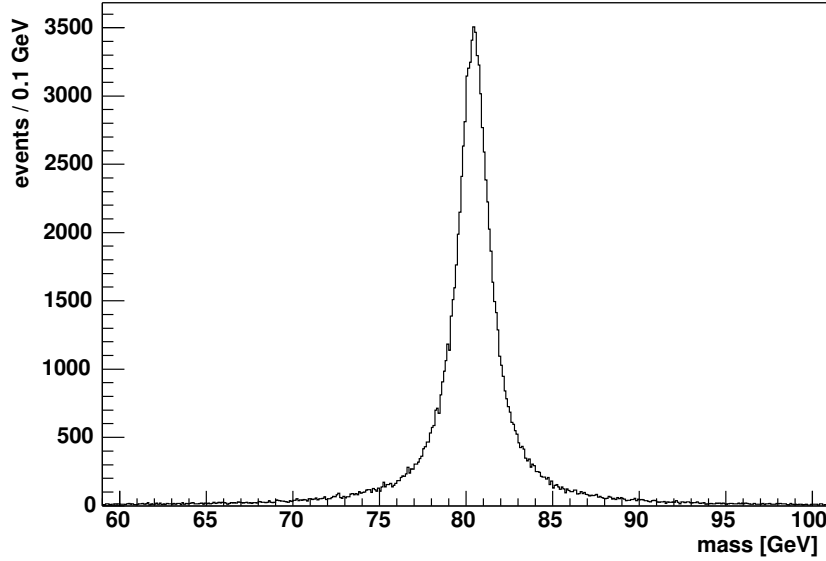


Figure 5.5: The distribution of the true W mass, taken from the Monte Carlo data. The histogram has 112,100 entries, the mean value is 80.25 GeV. The distribution is a Breit-Wigner distribution.

5.2.2 Using the Monte Carlo Knowledge

Having gained knowledge on jet multiplicities, we are interested in reconstructing W masses from jets.

In figure 5.5 the Monte Carlo masses of the hadronically decaying W bosons are shown.

A first approach for reconstructing W masses from generator jets is to match the Monte Carlo quarks originating from the hadronically decaying W to the generator jets. Then the W mass will be reconstructed if two generator jets can be matched to the two Monte Carlo quarks.

The matching criterion taken here was a tight distance criterion in the η - ϕ plane ($\Delta R \leq 0.16$). Comparing the W mass distribution from generator jets (see figure 5.6) to the W mass distribution taken from the Monte Carlo data (see figure 5.5), there are some obvious differences. By choosing this tight matching criterion, about half of the events is lost. This rather strict matching has an effect which can be described as 'trying to turn off the final state radiation afterwards'. That means if a jet originating from the W emits a gluon jet and therefore changes its η - ϕ coordinates, its distance to the Monte Carlo quark it originates from will be too large to be matched.

Tails to both sides are bigger and longer than the Monte Carlo W mass distribution (see figure 5.5). The left tail is due to jets which are not completely included in the cone, e.g. because of soft gluon emission, and due to neutrinos in the jets. The right tail is caused by jets 'catching' particles which do not belong to them and by mismatching because of hard gluon emission.

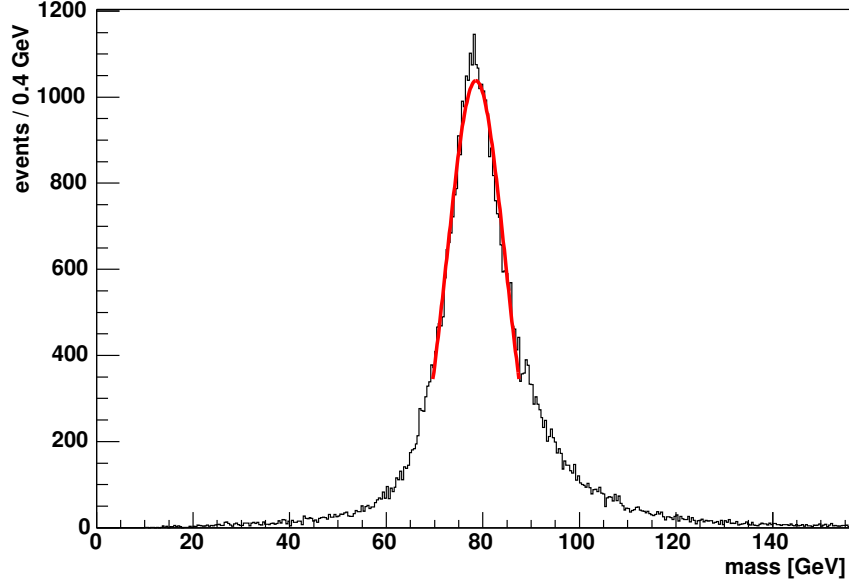


Figure 5.6: Reconstructed W masses from those generator jets, which can be matched tightly to Monte Carlo quarks of the hadronically decaying W . The matching criterion is $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \leq 0.16$. Only events with at least five jets were considered. The plot contains 51693 events, has a mean of 81.51 GeV and an RMS of 16.7 GeV. The curve (in red) is a Gauss fit in the range within ± 1.5 RMS around the mean. Its mean is 78.57 GeV and its standard deviation is 6.006 GeV .

However, we are much more interested in determining the W boson mass from reconstructed data than from generator data. So we match reconstructed jets to the Monte Carlo quarks originating from the hadronically decaying W . Again, the distance criterion $\Delta R \leq 0.16$ in the η - ϕ plane is used. This yields the W mass distribution in 5.7. In order to gain a quantitative description of the width, we do a '1.5- σ Gauss fit'. That is a Gaussian fit in the range within ± 1.5 RMS around the mean. The mean of the 1.5- σ Gauss fit is at 56.18 GeV and not in the region around 80 GeV. The ratio $\frac{\text{standard deviation}}{\text{mean}}$ for the reconstructed jets is $\frac{10.49 \text{ GeV}}{56.18 \text{ GeV}} = 0.187$, which is more than twice the ratio for the generator jets (see figure 5.6), $\frac{6.006 \text{ GeV}}{78.57 \text{ GeV}} = 0.076$. This is due to calorimeter smearing and the E_T cut.

In order to improve the detector resolution, jet energy corrections will be introduced in the following section.

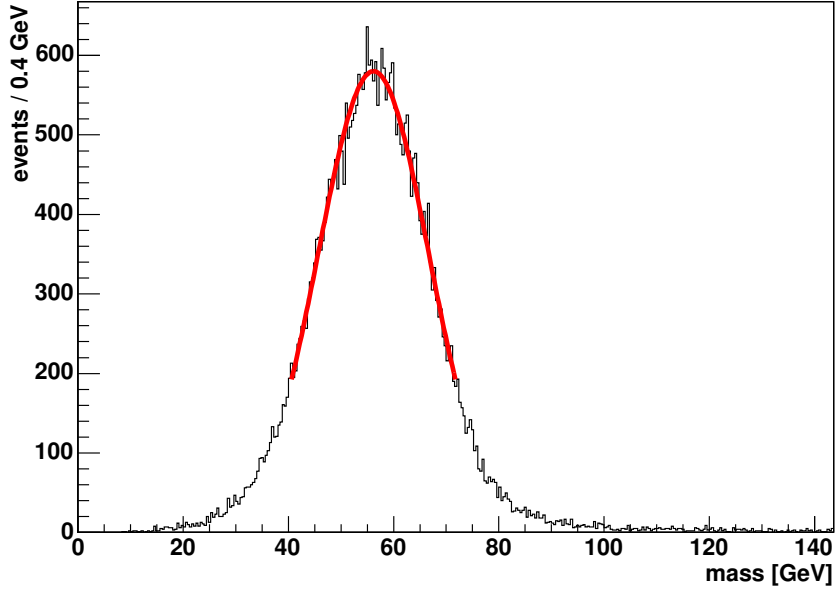


Figure 5.7: Reconstructed W masses from reconstructed jets, using the same matching criterion ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \leq 0.16$) as used in reconstructing the W mass from generator jets (see figure 5.6). About 35.5% of the events pass the matching criterion. There are 39759 events in the plot, having a mean of 56.78 GeV and an RMS of 13.03 GeV. The '1.5- σ Gauss fit' (this is a Gauss fit in the range within $\pm 1.5\sigma$ around the mean), shown as the smooth line (in red), has a mean of 56.18 GeV and a standard deviation of 10.49 GeV.

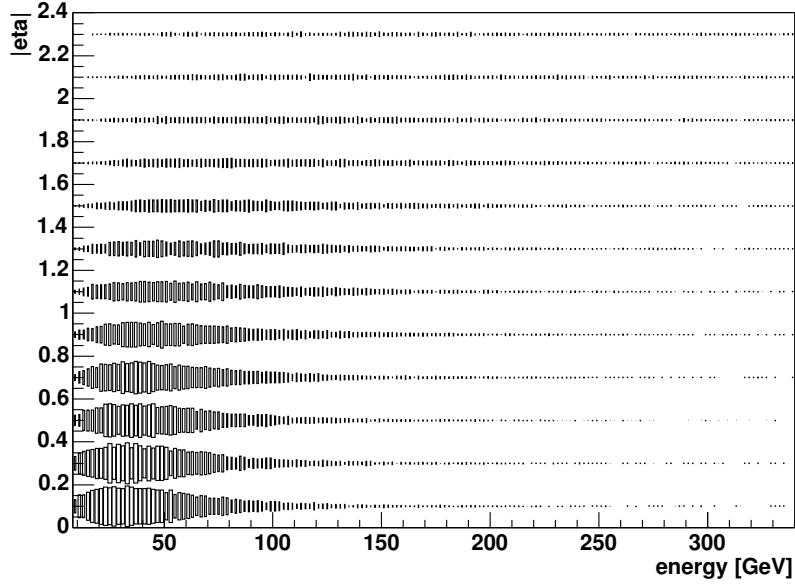


Figure 5.8: This box plot shows the number of the reconstructed jets originating from the hadronically decaying W bosons as a function of the absolute value of the pseudo-rapidity, $|\eta|$, and the energy. It has bins of size 2 GeV for the energy and bins of size 0.2 for $|\eta|$. There are 119818 entries, with a mean of 85.58 GeV and an RMS of 64.34 GeV for the energy and a mean of 0.9539 and an RMS of 0.6337 for $|\eta|$. The box sizes are proportional to the number of reconstructed jets. The maximum number of reconstructed jets in a $|\eta|$ -energy bin is 350.

5.2.3 Jet Energy Corrections

The most naive approach for jet energy corrections is to correct the energy and the three components of the momentum p^μ by the factor

$$f_{\text{naive}} = \frac{m_{W,\text{MC data}}}{m_{W,\text{reconstructed jets}}} = \frac{80.25 \text{ GeV}}{56.78 \text{ GeV}} = 1.413.$$

This for sure yields a very good mean value, but also the standard deviation grows by a factor f_{naive} . An 'intelligent' jet energy correction scheme has to provide a better sigma/mean ratio.

A constructive approach is to make the corrections depend on the pseudo-rapidity, η , and the energy, E , of the jet. Because of symmetry, we will use $|\eta|$ instead of η . At first we are interested in the energy- $|\eta|$ distribution of the reconstructed jets originating from the W . Taking a glance at figure 5.8, it is easy to see that for increasing $|\eta|$ and increasing energy the number of reconstructed jets decreases. In order to have sufficient statistics in most bins, the bin sizes for the energy are changed according to table 5.1, the bin sizes for $|\eta|$ are kept.

A more sophisticated approach for jet energy corrections is to calculate a correction factor for each η -energy interval. In order to do so, we match the reconstructed jets to the quarks from the hadronic W decay (once again using

8 - 20 GeV:	3	bins à	4 GeV
20 - 60 GeV:	20	bins à	2 GeV
60 - 100 GeV:	10	bins à	4 GeV
100 - 150 GeV:	5	bins à	10 GeV
150 - 250 GeV:	5	bins à	20 GeV
250 - 340 GeV:	3	bins à	30 GeV

Table 5.1: Bin sizes for the energy, used for all the eta slices.

the $\Delta R \leq 0.16$ criterion) and take a look at their distributions in each η -energy interval.

The distributions cannot be Gaussian because the correction factors have to be positive. Therefore we will estimate the correction in an interval by the mean of the corrections in this interval. Its error will be estimated by the RMS divided by the square root of the number of entries in the interval.

Taking a look at the different correction factors for one value $|\eta|$ but differing energies, we get so-called 'eta slices' (an example slice is shown in figure 5.10). One can fit a function

$$f = a + b \cdot \left(\frac{E}{\text{GeV}} \right)^c \quad (5.1)$$

using a χ^2 -fit in this eta slice, with E being the energy of the reconstructed jet and a , b and c being the dimensionless parameters to be fitted.

Now we have twelve (one for each $|\eta|$ slice) of these fitting functions. Of course it would be much easier and more convenient if we only had one or two fit functions for all parameters of $|\eta|$ and energy.

Considering that we have a barrel calorimeter and an endcap calorimeter, we fit the parameters for pseudo-rapidities $0 < |\eta| < 1.4$ and for $1.4 < |\eta| < 2.4$, respectively, in order to describe the $|\eta|$ -dependence of the fitted parameters. Each of the three parameters a , b and c is going to be fitted by a polynomial of degree three. The fitting is done separately for each parameter.

1. $a(|\eta|)$ is fitted by a polynomial of degree three.
2. Each eta slice is refitted for a fixed a . This value is calculated by using $a_3 \cdot |\eta|^3 + a_2 \cdot |\eta|^2 + a_1 \cdot |\eta| + a_0$ at the centre of the η interval. This yields a new set of 12 parameters c , depending on the absolute value of the pseudo-rapidity, $|\eta|$.
3. $c(|\eta|)$ is fitted by a polynomial of degree three, using the new set of parameters c .
4. Again, the eta slices are refitted, this time with a and c fixed in each slice. c is given in the same way as a was in 2.
5. Finally b is fitted, by a polynomial of degree three.

In total, the function for the absolute value of the pseudo-rapidity $|\eta|$ and the energy dependence of the energy corrections is now described by twelve

parameters:

$$f(|\eta|, E) = a_3 \cdot |\eta|^3 + a_2 \cdot |\eta|^2 + a_1 \cdot |\eta| + a_0 \quad (5.2)$$

$$+ (b_3 \cdot |\eta|^3 + b_2 \cdot |\eta|^2 + b_1 \cdot |\eta| + b_0) \cdot \left(\frac{E}{\text{GeV}} \right)^{c_3 \cdot |\eta|^3 + c_2 \cdot |\eta|^2 + c_1 \cdot |\eta| + c_0}.$$

The correction factors are visualised in figure 5.11.

When taking a look at the polynomial fitting of the parameters (see figure 5.12), one clearly sees that the fit functions of the parameters are not continuous in the overlap region of the barrel and endcap around $|\eta| \approx 1.4$. Jets with their axis in the proximity of this pseudo-rapidity cannot have a clear jet energy correction and therefore should not be considered in an analysis.

As a check of the parametrisation, a χ^2 -fit is applied. χ^2 is given by

$$\chi^2 = \sum_{\text{all bins}} \left(\frac{f(\text{bin}) - F_{\text{bin}}}{\Delta F_{\text{bin}}} \right)^2$$

with

$$F_{\text{bin}} = \frac{1}{N_{\text{bin}}} \sum_{i=1}^{N_{\text{bin}}} \frac{E_{\text{gen}}^i(\text{bin})}{E_{\text{rec}}^i(\text{bin})},$$

$$\Delta F_{\text{bin}} = \frac{1}{\sqrt{N_{\text{bin}}}} \text{RMS}(F_{\text{bin}})$$

and N_{bin} being the number of events in a bin. A bin is determined by its $|\eta|$ and energy range.

In the barrel range $\chi^2 = 453.95$ for 322 bins and twelve free parameters, in the endcap range $\chi^2 = 301.19$ for 228 bins and twelve free parameters. Obviously $\chi^2 \approx N_{\text{bins}} - \# \text{ free parameters} = \# \text{ degrees of freedom}$. There are a few reasons explaining this. First of all, the distributions in the eta slices are not Gaussian and mostly rather unsymmetrical. Also, the iterative fitting method does not take the correlation completely into account. And finally, there are effects in the border areas of barrel and endcap due to the fact that only a part of the jet is still in the calorimeter (this effect causes the behaviour of the fitting curve of parameter b in the $|\eta| > 2.2$ region in 5.9).

Applying the parametrisation of the jet energy corrections to tightly matched jets ($\Delta R \leq 0.16$), which do not point to the barrel/endcap region, before reconstructing the W mass (see figure 5.13) we see a tremendous improvement with respect to figure 5.7. The sigma/mean ratio in figure 5.13 is $\frac{10.55 \text{ GeV}}{81.69 \text{ GeV}} = 0.129$, which is approximately 31% lower than the sigma/mean ratio in figure 5.7, which is $\frac{10.49 \text{ GeV}}{56.18 \text{ GeV}} = 0.187$.

As we will see in further cross-checks, the energy corrections work well.

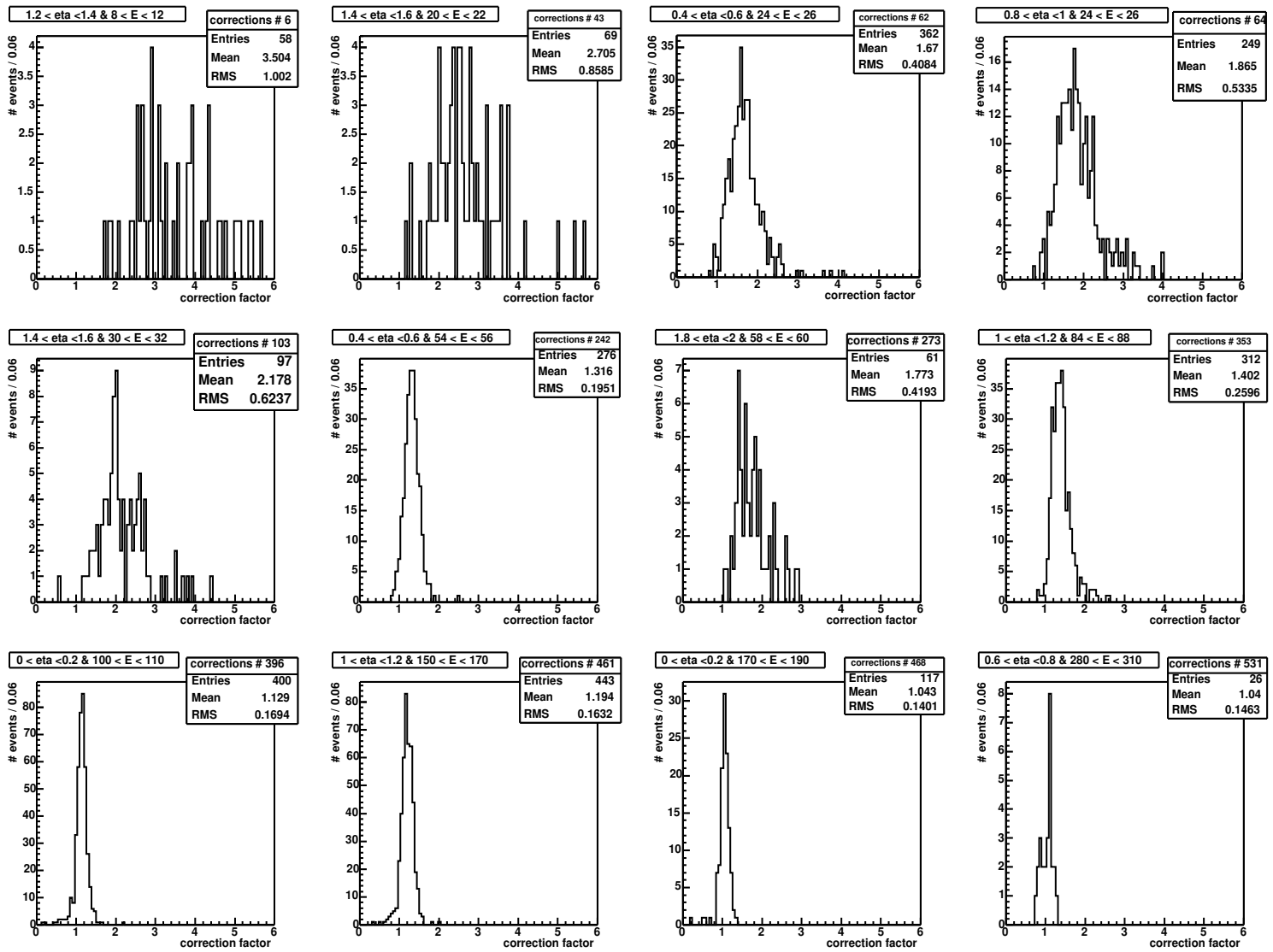


Figure 5.9: These ‘stamp plots’ show some distributions of the correction factor in different $|\eta|$ -energy intervals. The distributions are clearly non-Gaussian. They differ very much in RMS and in statistics. In the upper left corner, the energy- $|\eta|$ interval is given, in the upper right corner, the number of entries, the mean and the RMS are given.

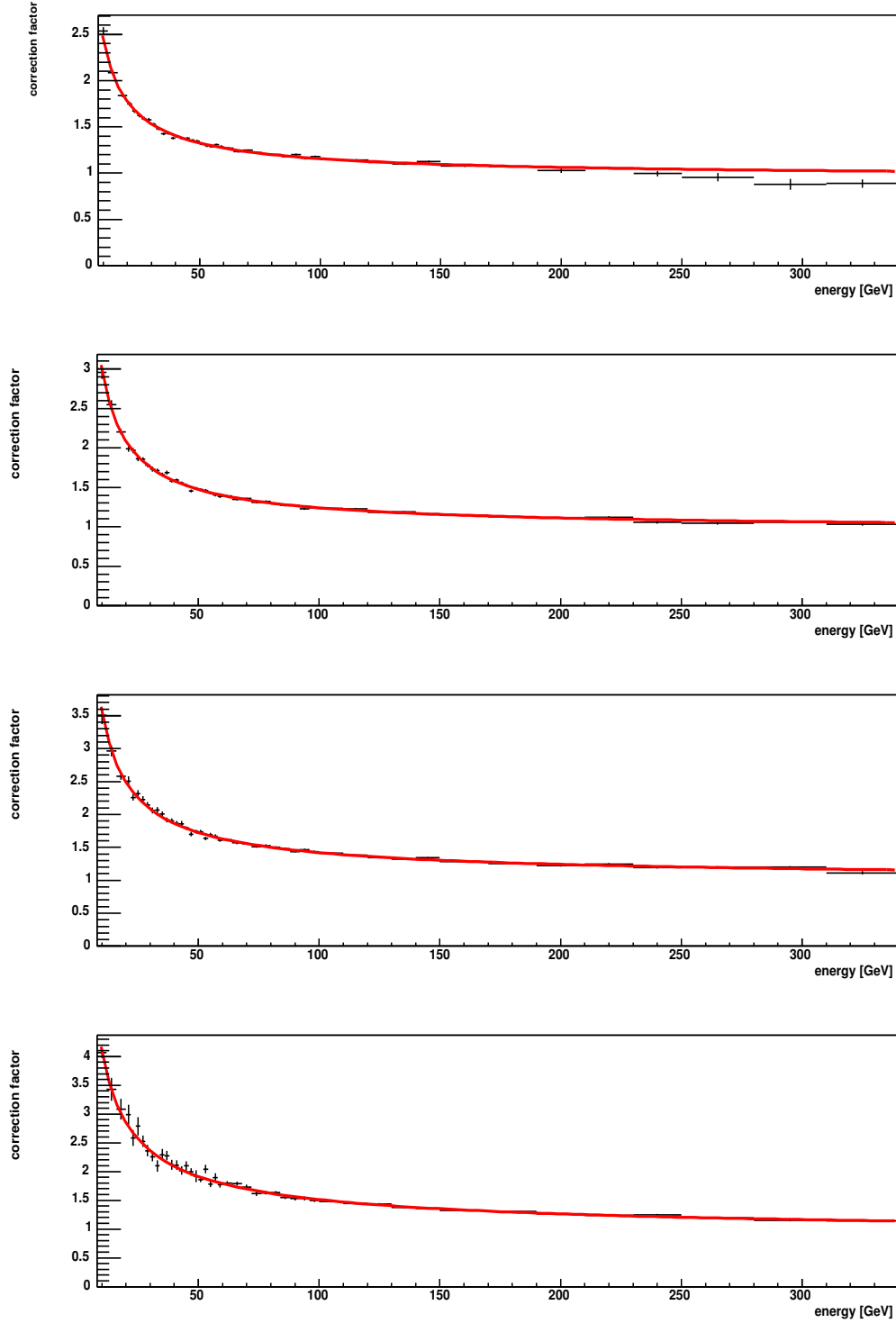


Figure 5.10: 'Eta slices' for different values of $|\eta|$, showing the dependence of the correction factor on the energy of the reconstructed jet. From up to down: Inner barrel regions $0.2 < |\eta| < 0.4$ and $0.8 < |\eta| < 1.0$, outer barrel region $1.2 < |\eta| < 1.4$ and inner endcap region $1.8 < |\eta| < 2.0$. The fitting function is defined in equation (5.1).

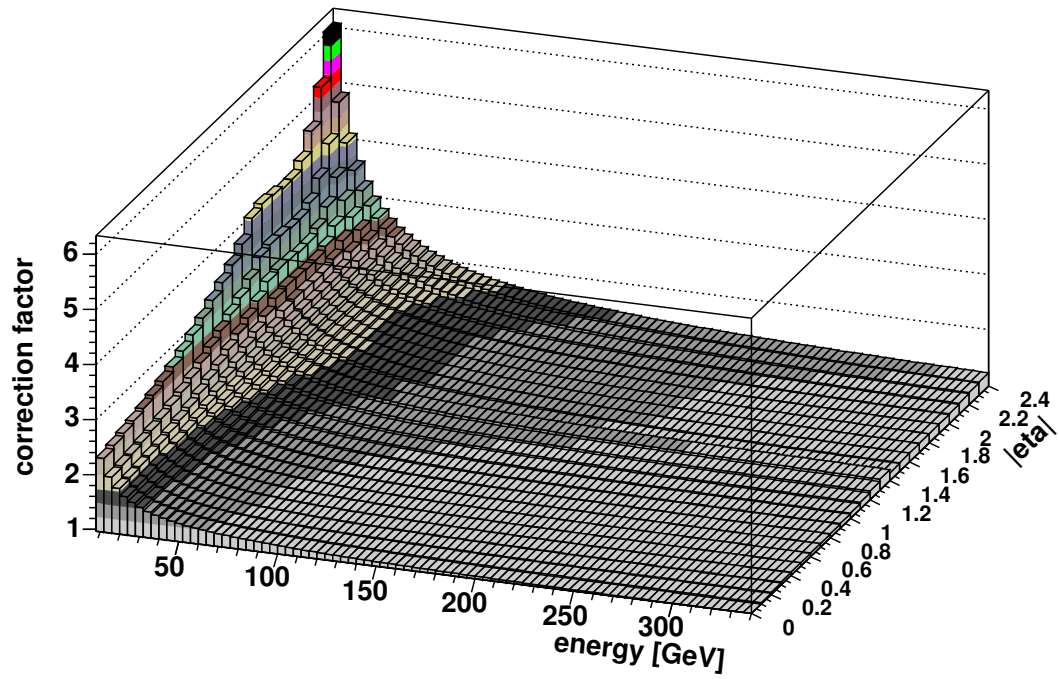


Figure 5.11: The correction factors in dependence on $|\eta|$ and the jet energy for a cone of size 0.5.

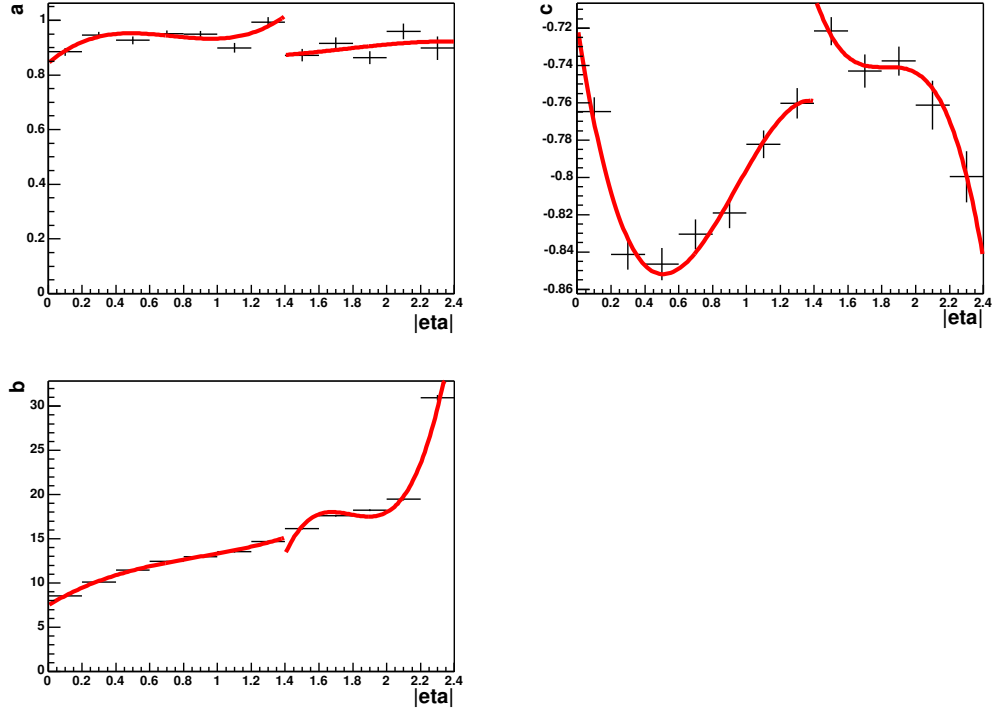


Figure 5.12: Fitting of the parameters of the function defined in (5.1) as polynomials of degree three. The plots are ordered according to the sequence in which they were fitted. The vertical error bars are the errors of the parameters fitted in the η slices. A detailed description of the way the fitting was done is given in subsection 5.2.3 on page 45.

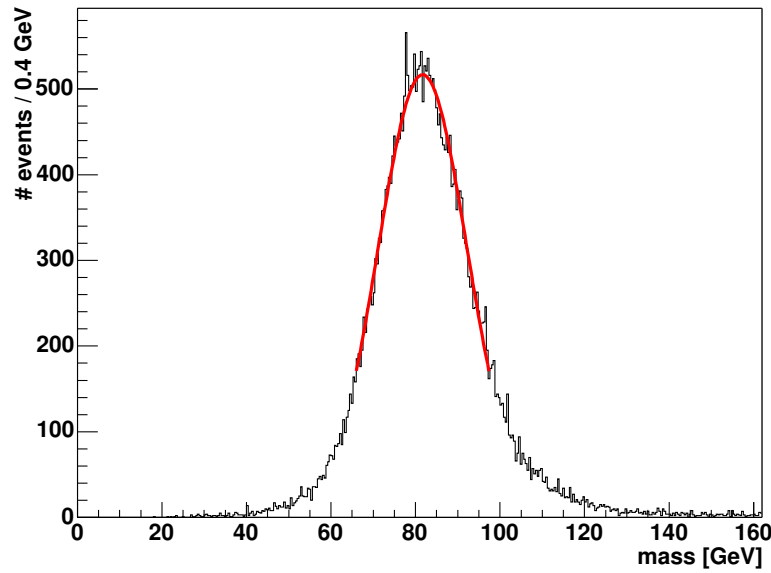


Figure 5.13: Reconstructed W masses from reconstructed jets after using jet energy corrections. Again, the reconstructed jets have been matched with the MC data using the tight matching criterion ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \leq 0.16$). There are 36901 events in this plot. The mean is 83.53 GeV and the RMS is 15.3 GeV. In the $1.5\text{-}\sigma$ Gauss fitting range (drawn in red) the mean is 81.69 GeV and the standard deviation is 10.55 GeV.

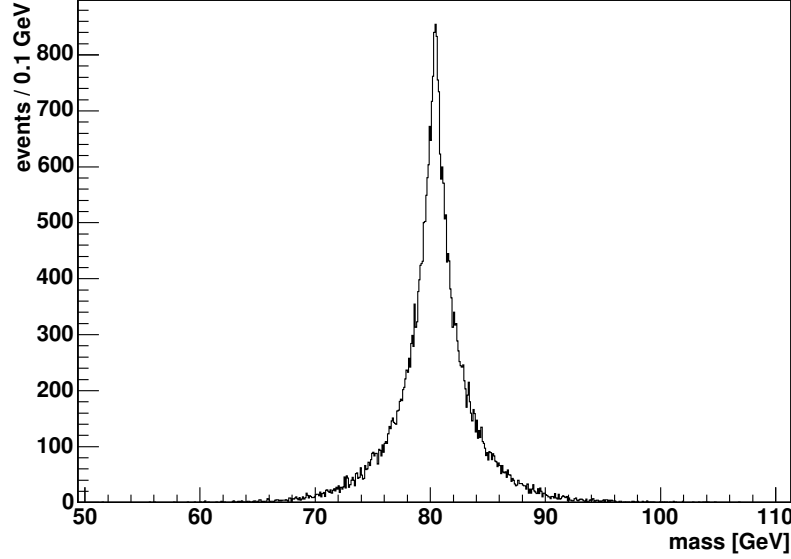


Figure 5.14: W mass reconstructed from reconstructed jets, using the jet energy corrections. The W mass is calculated for every possible combination of jets and the mass nearest to the Monte Carlo value is taken. The events are pulled towards the W mass. There are 34219 entries, yielding a mean of 80.35 GeV and a RMS of 3.565 GeV.

5.2.4 Best Mass Approach

A very simple approach for selecting the two jets originating from the hadronically decaying W is to reconstruct a mass from every combination of two jets. Then we select the mass which is closest to the Monte Carlo value. The results of this approach can be seen in figure 5.14.

Many events in the plot are pulled towards the W mass. Because of the many possible combinations, it is not very hard to have at least one pair of jets yielding a reconstructed mass near the Monte Carlo value. In addition, the shape is clearly non-Gaussian.

Unfortunately, this approach cannot be used in a realistic analysis, because most of the combinations do not use both W -jets for the W boson mass reconstruction. Our result is biased very much by MC value of the W boson mass.

5.3 Realistic Reconstruction

In the real experiment, the W bosons have to be reconstructed without any Monte Carlo information, as was done above in section 5.2.3.

5.3.1 Iterative Cone Algorithm with Cone Size 0.5

Given a semi-leptonic $t\bar{t}$ event with five jets, one will be able to find the jets originating from the W . First, it is easy to find the 'muon-jet' as muons can be reconstructed very well. Second, there is b -tagging. b -tagging yields a probability for a jet to be a b -jet as well as a probability for mistagging. One working point for b -tagging at CMS is an efficiency of 67.6%, with about 22.8 % of the c -jets and about 3.7 % of the light quark and gluon jets being mistagged as b -jets. [Wei]

In the following, we will assume perfect muon reconstruction and a perfect b -tagging, i.e. all b -jets and 'muon-jets' will be recognised as not coming from the hadronically decaying W . This will be done by matching the reconstructed jets to the Monte Carlo data, again using the distance in the η - ϕ plane, but this time using a less strict $\Delta R \leq 1.25$. A very good approach to get hold of the W -jets is to select the two jets of high energy after having done the perfect b - and muon-tagging.

At first, we take a look at the five jet events (5 jets = 2 b -jets + 2 W -jets + 1 'muon-jet'). Since there are only two jets left after perfect tagging, we reconstruct the W mass from these two jets. Taking a look at figure 5.15, we see a tail to the left and a tail to the right. The tail to the right is due to using the wrong jets for reconstructing the W mass, e.g. if the muon-jet and a jet originating from the hadronically decaying W overlap and a gluon jet emitted by a b -jet is taken to be a jet from the W . The tail to the left is due to gluon radiation. To obtain a measure for the width of the reconstructed W mass distribution, we fit a Gaussian on a flat distribution

$$c \cdot e^{-\frac{(m-\mu)^2}{2\sigma^2}} + s \cdot (m - \mu) + b \quad (5.3)$$

in range of $\pm 2\sigma$ around the peak, with m being the reconstructed mass, c being a constant, s being the slope of the line, and b being the intercept of the line at $m = \mu$.

The reconstruction can also be done for events with more than two remaining jets, selecting the two jets with highest energy as the W -jets. It is possible to fit signal and background reasonably for events with up to nine remaining jets. Though having a large number of events in total, there is no sufficient statistics for events with more than nine remaining jets. The results are summarised in table 5.2.

A statistically more robust way for fitting is to plot all W masses reconstructed from events with up to nine remaining jets in one histogram and to fit function 5.3 in this plot. That is done in figure 5.16. There are 12077 events in the $2\text{-}\sigma$ range of the fit, yielding an efficiency of 0.44.

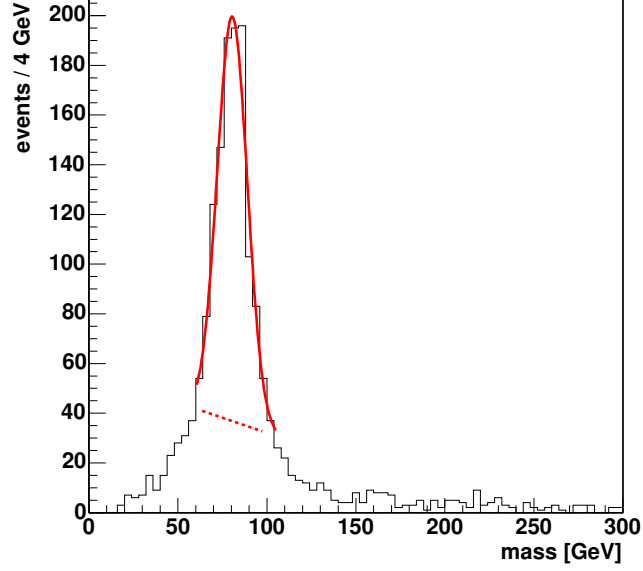


Figure 5.15: W mass reconstructed from five jet events, assuming a perfect b - and muon-tagging. Only events in which the two remaining jets had their axis outside the barrel/endcap overlap region and each had a reconstructed energy of more than 25 GeV were taken. The fitting (in red) was done in the $2\text{-}\sigma$ range, the background is fitted by the red, dashed line. There are 1742 entries in this plot, the mean is 89.73 GeV and the RMS is 39.76 GeV. The fit yields a Gaussian with a constant factor of 162.8, a mean of 80.54 GeV and a standard deviation of 8.43 GeV while the red line has a slope of $-0.25 \frac{\text{events}}{\text{GeV}}$ and an intercept of 36.8 events per 4 GeV at the mean of the Gaussian.

number of re- maining jets	reconstructed W mass [GeV]	sigma [GeV]	events in the $2\text{-}\sigma$ range	total events
2	80.54	8.43	1170	1742
3	80.02	9.82	2278	3877
4	80.03	9.32	2325	4960
5	80.09	10.4	2093	4760
6	81.47	9.20	1499	4144
7	82.62	8.51	1082	3214
8	80.54	13.0	1035	2602
9	84.18	10.2	711	2099

Table 5.2: Realistically reconstructed W masses, assuming perfect b - and muon-tagging, using an iterative cone algorithm with cone size 0.5. 'Number of remaining' jets denotes the number of jets after taking away the tagged jets. Of those remaining jets, the two highest energy ones had to have their jet axis out of the barrel/endcap overlap region, and had to have a reconstructed energy of more than 25 GeV. The W mass was reconstructed from these two jets.

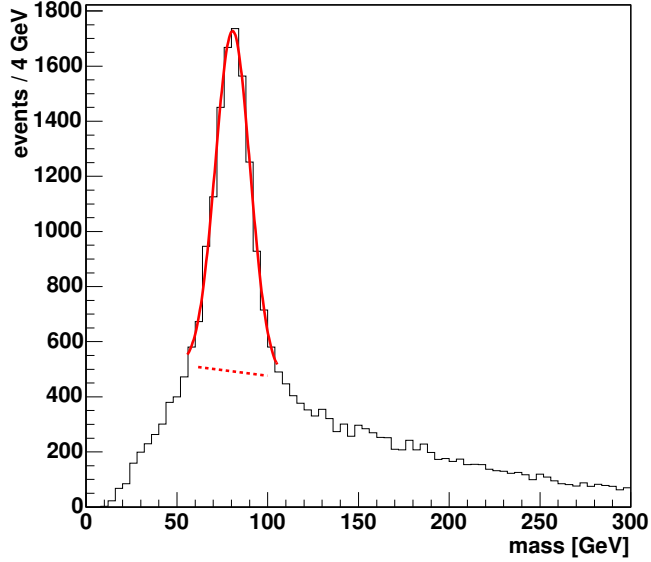


Figure 5.16: W mass reconstructed from events with five to twelve jets, assuming a perfect b - and muon-tagging. After tagging, the two remaining highest energy jets were taken for reconstructing the W mass, in case they both were outside the barrel/endcap overlap region and they both had a reconstructed energy of more than 25 GeV. There are 27398 events (1926 overflow) in this plot, yielding a mean of 111.2 GeV and an RMS of 59.26 GeV. The fitting (in red) in the $2\text{-}\sigma$ range yields a constant factor of $1236 \frac{\text{events}}{4 \text{ GeV}}$, a mean of 80.68 GeV, a sigma of 9.54 GeV, a slope of $-0.83 \frac{\text{events}}{\text{GeV}}$ and an intercept at mean of $492 \frac{\text{events}}{4 \text{ GeV}}$.

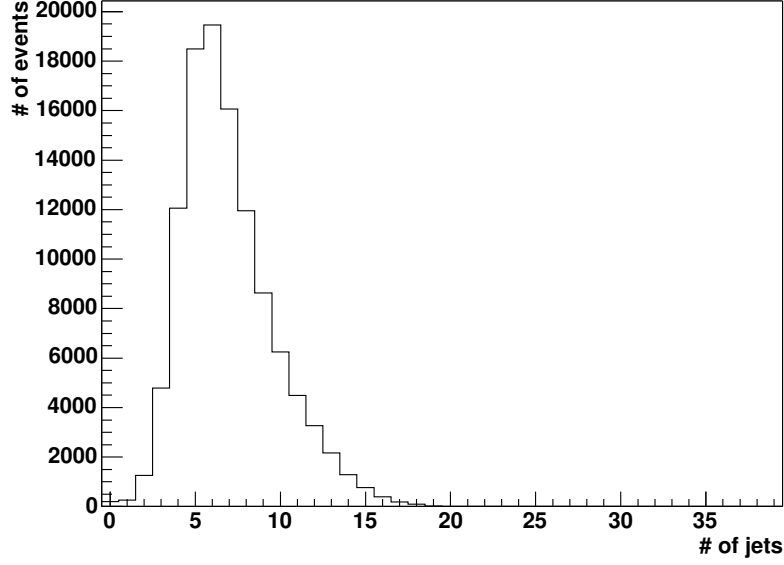


Figure 5.17: Jet multiplicities of reconstructed jets using an iterative cone algorithm with a size of 0.6. There are 112,100 events with a mean of 6.957 jets and an RMS of 2.742 jets.

5.3.2 Iterative Cone Algorithm with Cone Size 0.6

In this and the following subsections we use increased cone sizes for the iterative cone algorithm. Increasing the cone size will decrease the number of gluon jets, because soft gluon jets will now be more often in the cone of the original jet. But we will also have the problem of overlap. Two jets, which were originally separated, can now happen to be considered as one jet.

Having reconstructed the W mass from jets found by an iterative cone algorithm with cone size 0.5, we want to examine the results of other cone sizes. We are interested in the possibilities of getting a lower sigma/mean ratio and/or getting a higher efficiency. Taking a look at the number of reconstructed jets using a cone of size 0.6, we see that the number is about 0.6 smaller than that for a cone of size 0.5 (jet multiplicities are shown in figure 5.17). The reason for this is that soft gluon radiation jets are now more often in the cone of the original jet.

Of course, jet energy corrections have to be computed again for this new cone size. This is done in the same way as with the cone size of 0.5 (the result can be seen figure 5.18).

Reconstructing the W mass from the two highest energy jets in the respective event yields the numbers shown in table 5.3. The W mass plot for just two remaining jets after b -tagging and muon-tagging is given in figure 5.19. When taking all events with up to nine remaining jets together, we get figure 5.20.

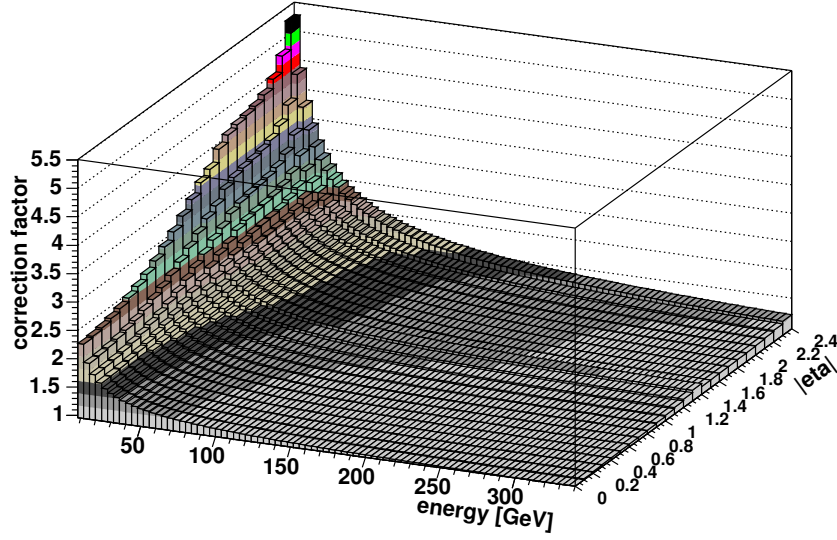


Figure 5.18: Jet energy corrections for an iterative cone algorithm with a cone size of 0.6.

number of re- maining jets	reconstructed W mass [GeV]	sigma [GeV]	events in the $2\text{-}\sigma$ range	total events
2	80.25	7.41	1439	2418
3	79.10	9.32	2447	4583
4	79.60	8.66	2173	5340
5	80.56	10.9	1932	4695
6	81.41	12.1	1496	3778
7	82.03	10.5	1013	2968
8	85.61	9.49	670	2331
9	84.46	9.55	533	1778

Table 5.3: Realistically reconstructed W masses, assuming perfect b - and muon-tagging, using an iterative cone algorithm with cone size 0.6. The W mass reconstructed from the two highest energy jets after taking away the tagged jets, in case their axis do not point to the barrel/endcap region, and their reconstructed energies are higher than 25 GeV.

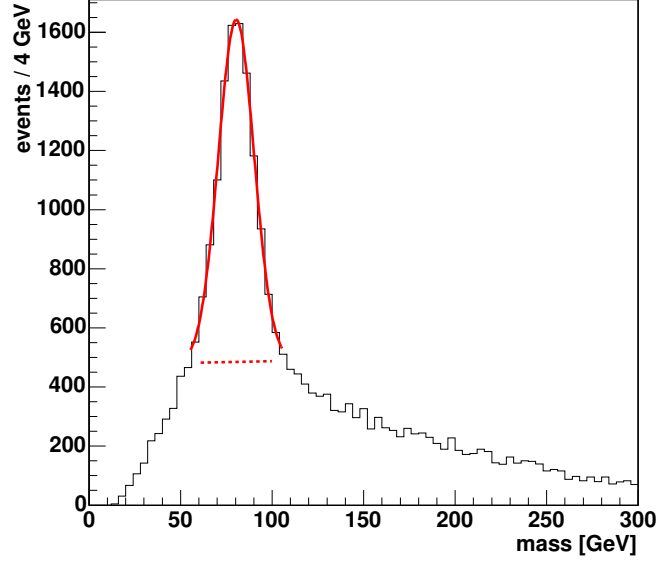


Figure 5.19: Realistically reconstructed W masses from five jet events, assuming perfect b - and muon-tagging. The jet algorithm used is an iterative cone algorithm with cone size 0.6. The W was reconstructed from the two remaining jets after taking out of the tagged jets, in case those two remaining jets did not have their axis in the barrel/endcap overlap region. There are 5800 entries in this plot. The mean is 78.78 GeV and RMS is 39.88 GeV. The fitting (in red) in the $2\text{-}\sigma$ range yields a constant factor of $1477 \frac{\text{events}}{4 \text{ GeV}}$, a mean of 79.06 GeV, a sigma of 10.0 GeV, a slope of $-1.57 \frac{\text{events}}{\text{GeV}}$ and an intercept of $128 \frac{\text{events}}{4 \text{ GeV}}$ at mean.

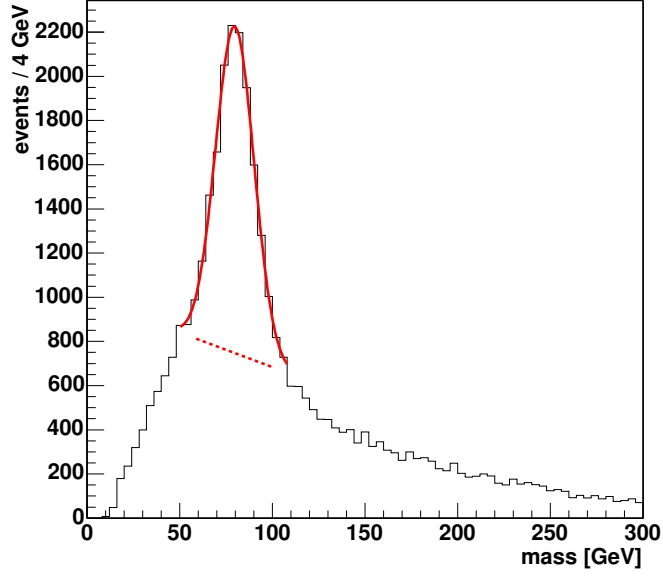


Figure 5.20: W mass reconstructed from events with five to twelve jets, assuming a perfect b - and muon-tagging, using an iterative cone with cone size 0.6. After taking away the tagged jets, the two highest energy jets of the remaining jets had to be outside the barrel/endcap overlap region and had to have reconstructed energies lower than 25 GeV to be used for reconstructing the W mass. The 27891 events (2171 overflow events) have a mean of 115.2 GeV and an RMS of 60.58 GeV. In the $2\text{-}\sigma$ range we have a constant of $1158 \frac{\text{events}}{\text{GeV}}$, a mean of 80.38 GeV and a sigma of 9.749 GeV. The slope of the dashed line is $-3.117 \frac{\text{events}}{\text{GeV}}$ and has an intercept at the mean of $747.5 \frac{\text{events}}{4 \text{ GeV}}$. There are 11499 events in the $2\text{-}\sigma$ range of the fit.

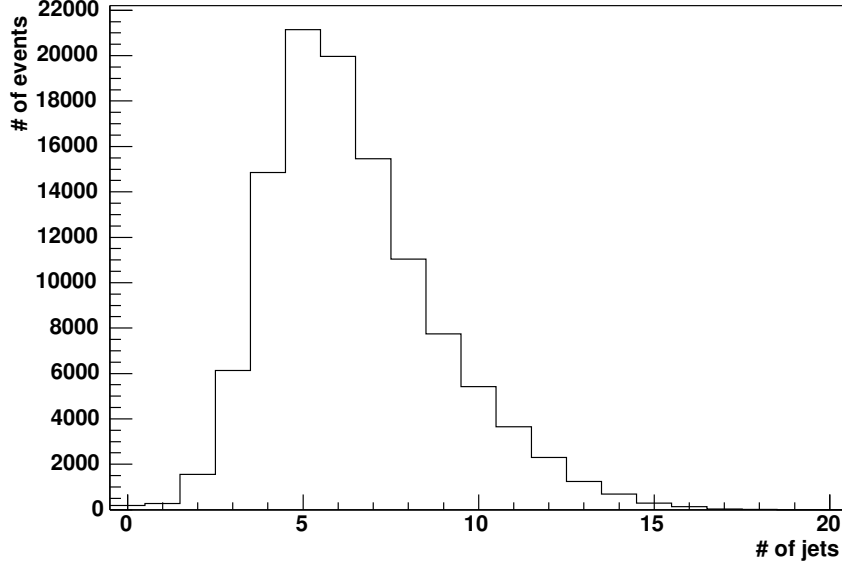


Figure 5.21: Jet multiplicities for an iterative cone algorithm with a cone size of 0.7. The 112,100 entries yield a mean of 6.49 jets and a RMS of 2.47 jets.

5.3.3 Iterative Cone Algorithm with Cone Size 0.7

In similar fashion, one can treat the use of the iterative cone with a cone size of 0.7.

Because of the increased cone size, the average number of reconstructed jets, 6.49, is about 0.5 smaller than with a cone size of 0.6 (see figure 5.21).

After calculating the jet energy corrections, one can reconstruct the W mass from those events with altogether five to twelve jets. The results for reconstructing the W mass depending on the number of jets is shown in table 5.4, the plot for five to twelve jets combined is given in figure 5.22.

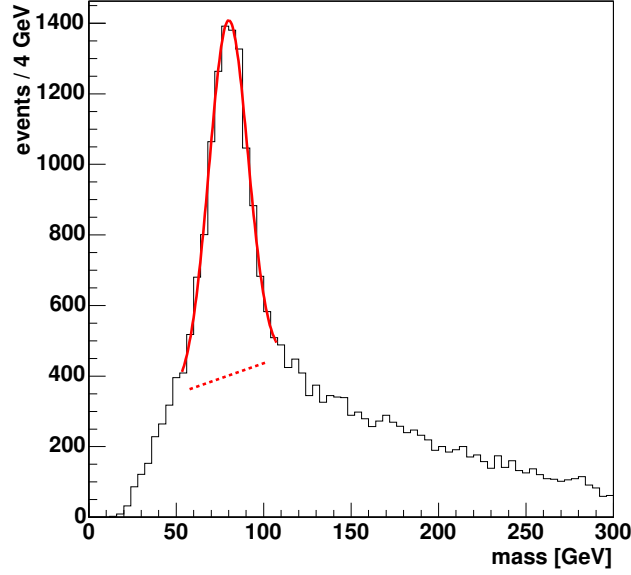


Figure 5.22: Using an iterative cone algorithm with cone size 0.7, the W mass reconstructed from events with five to twelve jets, assuming a perfect b - and muon-tagging. The axis of the two remaining jets with the highest energies had not to point into direction of the barrel/endcap overlap region, also these two jets each had to have a reconstructed energy of more than 25 GeV. There are 26884 events (2199 overflow events) in this histogram, having a mean of 119.4 GeV and a RMS of 61.99 GeV. In the $2\text{-}\sigma$ range, the fit yields a constant factor of $1007 \frac{\text{events}}{4 \text{ GeV}}$, a mean of 79.86 GeV, a standard deviation of 11.1 GeV, a slope of $1.75 \frac{\text{events}}{\text{GeV}}$ and an intercept of $401 \frac{\text{events}}{4 \text{ GeV}}$ at the mean of the fit. There are 10502 events in the range of the mean plus/minus 2 standard deviations of the fit.

number of re- maining jets	reconstructed W mass [GeV]	sigma [GeV]	events in the $2\text{-}\sigma$ range	total events
2	78.32	6.81	1506	2894
3	78.76	10.8	2606	5036
4	79.48	9.62	2090	5210
5	79.65	11.8	1680	4362
6	83.86	11.7	1191	3448
7	84.57	11.7	905	2740
8	86.39	10.6	547	1958
9	86.23	14.82	419	1236

Table 5.4: The W masses were realistically reconstructed from jets found by an iterative cone algorithm of cone size 0.7. The two highest energy jets had to have their axes not to point into the direction of the overlap region of the barrel and th endcap and also had to have energies of more than 25 GeV. The W mass is reconstructed from the two highest energy jets.

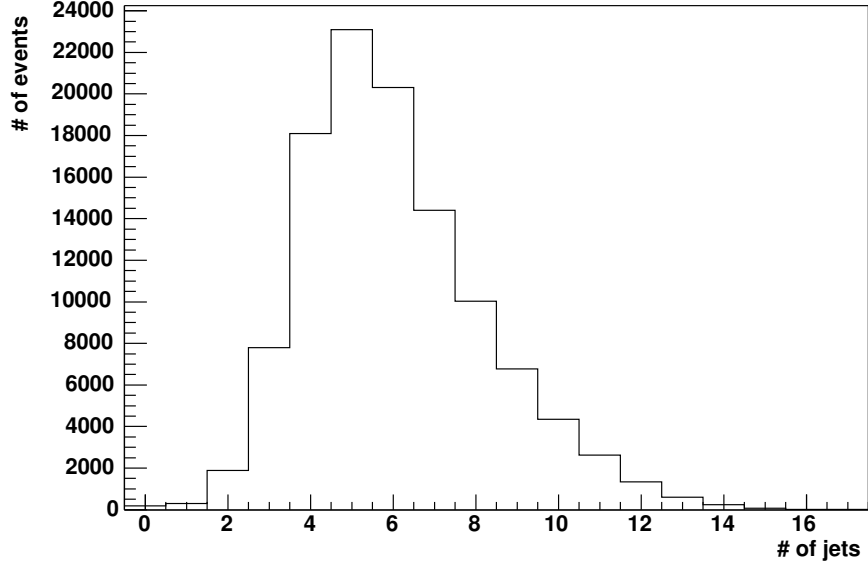


Figure 5.23: The jet multiplicities for an iterative cone algorithm with a cone size of 0.8. The 112,100 entries yield a mean of 6.066 jets and a RMS of 2.235 jets.

5.3.4 Iterative Cone Algorithm with Cone Size 0.8

The last cone size to be considered in this section is 0.8.

The average number of jets (see figure 5.23) found in an event, 6.066, is about 1.5 jets per events lower than the average number of jets found in an event when using an iterative cone algorithm with a cone size of 0.5.

Again, we reconstruct the W mass from events, which have five to twelve jets. Yet again, we assume perfect b -tagging and muon-tagging. After excluding the tagged jets, we only use those events which have their highest energy jets' axis in the barrel/endcap overlap region.

Table 5.24 shows the reconstructed masses W for events with five to nine jets, figure 5.5 shows the W mass reconstructed from events with five to twelve jets together.

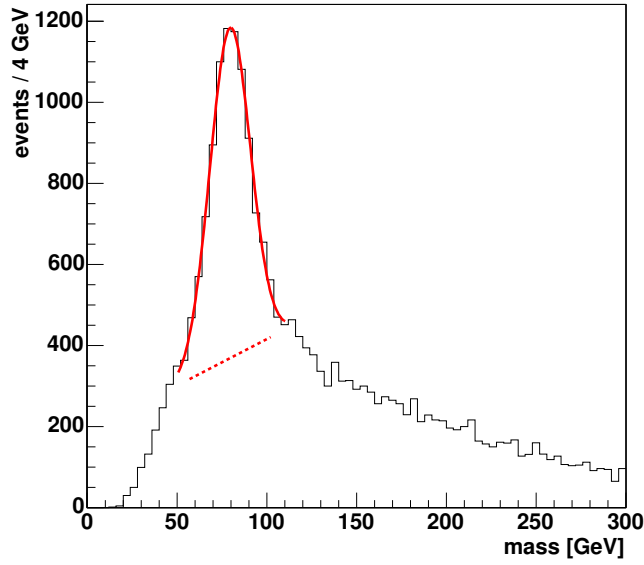


Figure 5.24: Realistically reconstructed W masses, using events with five to twelve jets found by the iterative cone algorithm with a cone size of 0.8. Only those events were used, which had the two highest energy jets' axis pointing outside the barrel/endcap overlap region and their energies larger than 25 GeV, after excluding the tagged jets. There are 24982 entries (2255 overflow entries), the mean is 123.4 GeV and the RMS 63.53 GeV. The fit in the $2\text{-}\sigma$ range yields a Gaussian with a constant of $814.7 \frac{\text{events}}{4 \text{ GeV}}$, mean of 79.59 GeV and a standard deviation of 11.27 GeV and a line with a slope of $-2.294 \frac{\text{events}}{\text{GeV}}$ and an intercept at the mean of $368.9 \frac{\text{events}}{4 \text{ GeV}}$. There are 9645 events in the $2\text{-}\sigma$ range.

number of re- maining jets	reconstructed W mass [GeV]	sigma [GeV]	events in the $2\text{-}\sigma$ range	total events
2	78.28	9.14	1812	3218
3	77.95	10.4	2365	5203
4	78.99	10.72	1905	5045
5	79.52	15.55	1609	4006
6	85.16	10.65	956	3213

Table 5.5: W boson masses realistically reconstructed from jets found by an iterative cone algorithm of cone size 0.8. The two highest energy jets had to have their jet axes pointing outside the overlap region of the barrel and endcap and had to have jet energies larger than 25 GeV. The W mass is reconstructed from those two highest energy jets. There was not sufficient statistics for fitting the Gaussian plus line curve to events with seven or more remaining jets.

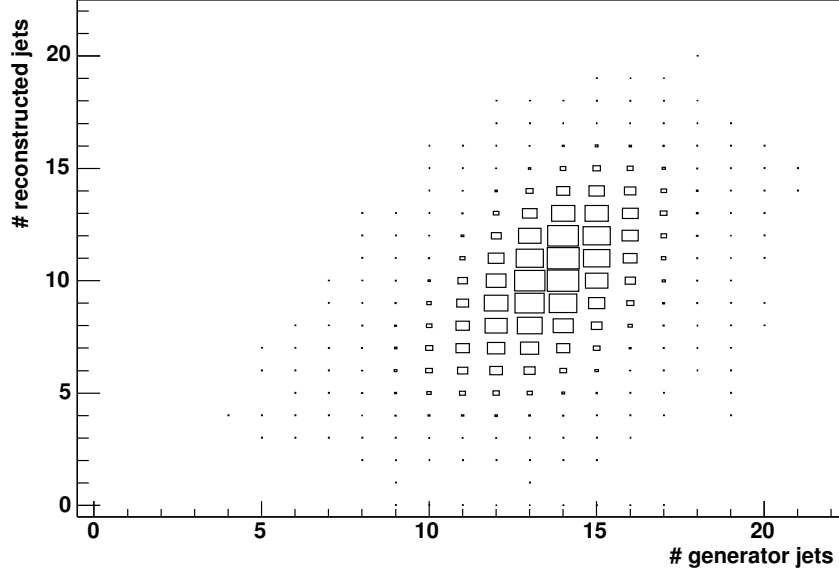


Figure 5.25: Box plot of the number of reconstructed jets vs. the number of generator jets, using the inclusive k_T algorithm with the ΔR scheme as the angular definition and the E scheme as the recombination scheme. The average number of generator jets is 13.55 with an RMS of 1.856, the average number of reconstructed jets is 10.02 with an RMS of 2.71. The Covariance is 2.51, the correlation coefficient is 0.50.

5.3.5 k_T Algorithm

The k_T Algorithm has been implemented in ORCA recently² and is not yet fully tested and understood. Its implementation is based on KtJet ([KtJet]).

We will take a short look at the results of the k_T algorithm for both the inclusive and exclusive modes.

Inclusive Mode

The default configuration of the k_T algorithm in ORCA is the inclusive k_T algorithm with the ΔR scheme as the angular definition and the E scheme as the recombination scheme.

Looking at figure 5.25, it can be seen that the average number of jets found is larger than the average number of jets found by the iterative cone algorithm with a cone size of 0.5. The number of five jet events is very small, especially at the generator level.

The jet energy corrections are calculated in the way described in subsection 5.2.3. The corrections are shown in figure 5.26.

Reconstructing the W mass from the reconstructed jets with Monte Carlo knowledge (matching criterion $\Delta R \leq 0.16$), but without jet energy corrections,

²Actually, there has been an implementation of the k_{Tt} jet algorithm, which did not work properly.

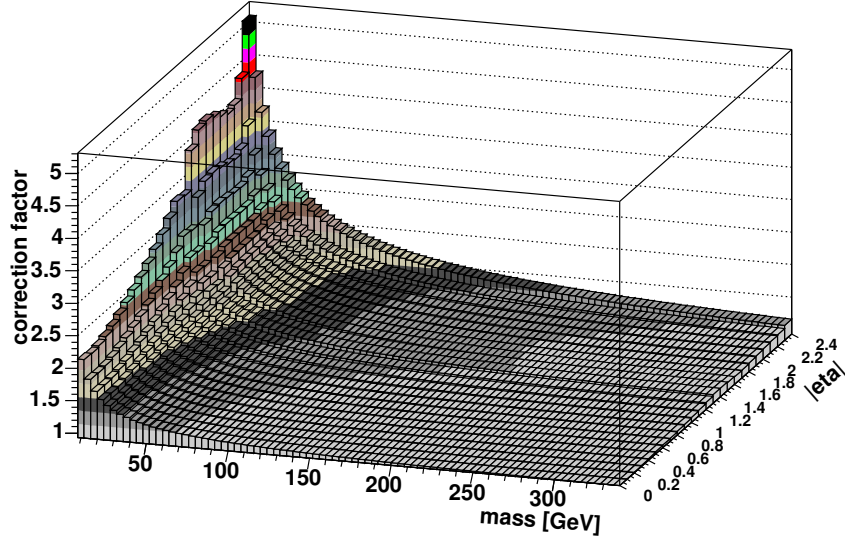


Figure 5.26: The correction factors depending on the energy and the absolute value of the pseudo-rapidity, $|\eta|$, for the inclusive k_T algorithm, using the ΔR scheme as the angular definition and the E scheme as the recombination scheme.

yields the results shown in figure 5.27. The ratio of standard deviation to mean in the $1.5\text{-}\sigma$ fitting range is 0.1836. The number of events in which two jets could be matched to the Monte Carlo quarks from the hadronically decaying W is 24981. That is quite smaller than the number for the iterative cone algorithm with cone size 0.5, which is 39579.

Reconstructing the W mass realistically, we again use the perfect b -tagging and muon-tagging via comparison to Monte Carlo data (distance criterion $\Delta R \leq 1.25$ in the η - ϕ plane). The tagged jets (two b -jets, one muon-jet) are removed from the event. If the two highest energy jets are both not in the barrel/endcap overlap region and they both have reconstructed energies of more than 25 GeV, the W mass is reconstructed from them, using the jet energy corrections. Using the events with five to fifteen overall jets, there are 11959 events in the 2 standard deviation range (see figure 5.28). This is an efficiency of 0.391.

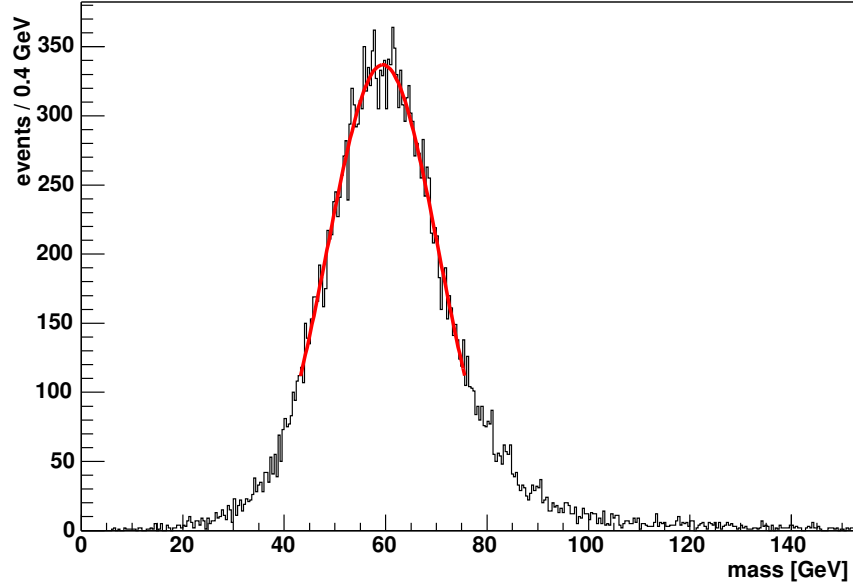


Figure 5.27: W mass obtained from reconstructed jets, using Monte Carlo data for matching the jets to the quarks from the hadronically decaying W boson, for the inclusive k_T algorithm using the ΔR scheme as the angular definition and the E scheme as the recombination scheme. There are 24981 events in the histogram, yielding a mean of 61.73 GeV and a RMS of 16.15 GeV. In the $1.5\text{-}\sigma$ fitting range, the mean is 59.43 GeV and the standard deviation is 10.91 GeV.

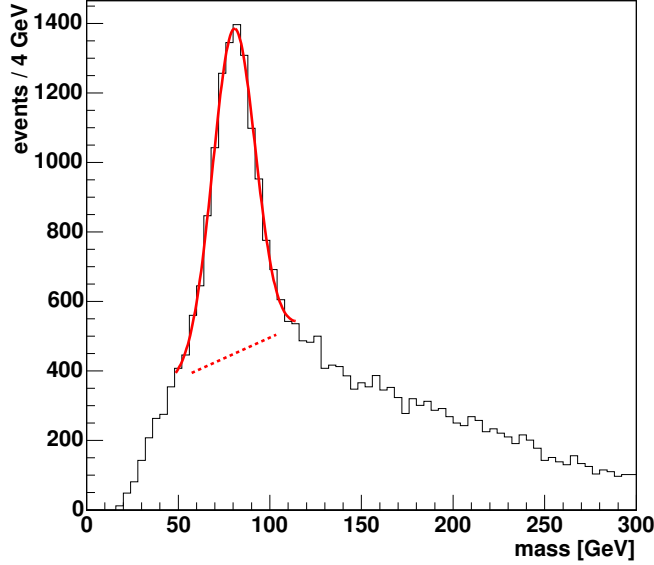


Figure 5.28: W mass realistically reconstructed from events with five to fifteen reconstructed jets, for the inclusive k_T algorithm, assuming perfect b -tagging and muon-tagging. After tagging, the tagged jets were excluded. Of the remaining jets, the two highest energy jets were used for reconstructing the W mass, using the jet energy corrections. If one or both of the two highest energy jets were inside of the barrel/endcap overlap region and if one or both jets have reconstructed energies higher than 25 GeV, those events were not used for reconstruction. There are 30555 entries (2779 overflow) in this histogram, with a mean of 124.4 GeV and a RMS of 63.69 GeV. The Gaussian fit in the $2\text{-}\sigma$ range yields a constant of $934.8 \frac{\text{events}}{\text{GeV}}$, a mean of 80.41 GeV and a standard deviation of 11.55 GeV. The dashed line has a slope of $2.391 \frac{\text{events}}{\text{GeV}}$ and an intercept at the mean of the Gaussian of $449.3 \frac{\text{events}}{4 \text{ GeV}}$.

Exclusive Mode

Using the k_T algorithm in the exclusive mode, one can choose to stop merging jets when a given number of jets is reached.

Expecting five jets in our semileptonically decaying $t\bar{t}$ sample, we choose to stop merging when five jets have been found.

The jet energy corrections are calculated in the same way as done in the previous sections. We use them to reconstruct the W (see figure 5.29). There is a second peak in the plot at 10 to 20 GeV. This is mainly caused by a W -jet being at a small distance from a b -jet or the muon-jet. In this case, the algorithm considers them to be one jet and a wrong jet, e.g. a low-energy gluon jet, is taken for reconstructing the W mass.

There are 7626 events in the 2 standard deviation range. This yields the high efficiency of 0.5140.

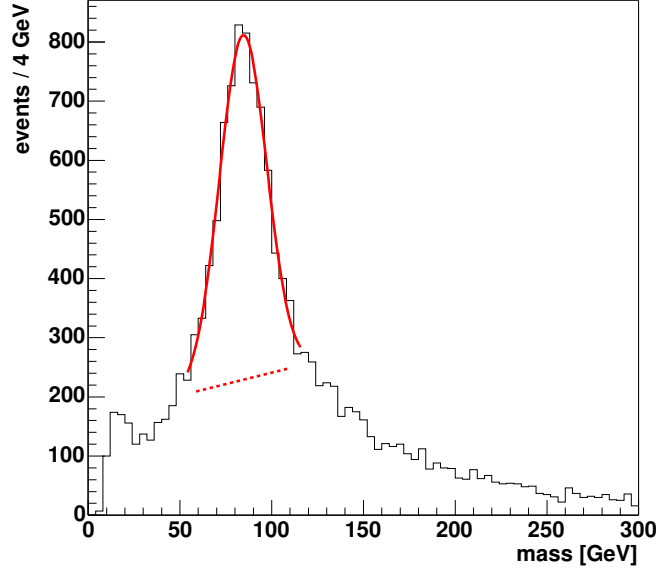


Figure 5.29: W mass reconstructed from jets found by the exclusive mode of the k_T algorithm, assuming a perfect b - and muon-tagging. The tagged jets were excluded. The two remaining highest energy jets were taken for reconstructing the W mass, if their axes pointed outside the barrel/endcap overlap region and they each had a reconstructed energy of at least 25 GeV. There are 14838 events (604 overflow) in this plot, yielding a mean of 104.01 GeV and a RMS of 55.23 GeV. The fitting (in red) in the $2\text{-}\sigma$ range yields a constant of $582.7 \frac{\text{events}}{4 \text{ GeV}}$, a mean of 84.44 GeV, a sigma of 12.84 GeV, a slope of $0.7786 \frac{\text{events}}{\text{GeV}}$ and an intercept of $228.9 \frac{\text{events}}{4 \text{ GeV}}$ at mean.

5.3.6 Comparing the Results of the Jet Algorithms

In table 5.6 the results of the different jet algorithms are listed. The efficiency is given by

$$\text{efficiency} = \frac{\text{number of events in the region mean} \pm 2 \text{ standard deviations}}{\text{number of total events}}. \quad (5.4)$$

An event had to fulfil these criteria for being taken into account:

- It had to have five to twelve jets for the cone algorithm or a five to fifteen jets event for the inclusive k_T algorithm, respectively. (For the exclusive mode all events were five jets events.)
- Matching of b -jets and muon-jet to Monte Carlo data had to be possible.
- After excluding the tagged jets, the two jets with highest energy of the remaining jets had to fulfil:
 - Their axis did not point to the barrel/endcap overlap region, i.e. the absolute value of the pseudo-rapidity, $|\eta|$, was less 1.3 or larger 1.5.
 - Their reconstructed energies (before jet energy correction) were larger than 25 GeV.

The k_T algorithm in the exclusive mode, forcing the final state to decompose into five jets, yields the highest efficiency, but also the highest standard deviation/mean ratio. The iterative cone algorithm with cone size 0.5 has the smallest standard deviation/mean ratio and the second highest efficiency.

Increasing the cone size for the iterative algorithm does not increase the efficiency but decreases it slightly. Also the standard deviation/mean ratio increases when increasing the cone size.

The k_T algorithm has to be further investigated, since it is relatively new in the ORCA package.

algorithm	$\frac{\text{standard deviation}}{\text{mean}}$	reconstructed W mass [GeV]	efficiency
iterative cone 0.5	0.118	80.7	0.421
iterative cone 0.6	0.121	80.4	0.413
iterative cone 0.7	0.139	79.9	0.391
iterative cone 0.8	0.142	79.6	0.386
inclusive k_T	0.143	80.4	0.391
exclusive k_T (5 jets)	0.152	84.4	0.514

Table 5.6: Comparison of the results of the different jet algorithms. The efficiency is defined in equation (5.4). Cone size 0.5 yields the smallest standard deviation/mean ratio, the exclusive k_T algorithm (5 jets) has the highest efficiency.

The iterative cone algorithm with cone size 0.5 is established at CMS and has proven to be of value in this study. It will be used for the reconstruction of the top mass in the following chapter.

A study with a cone size of 0.4 could not be performed for this thesis because of technical reasons.

Chapter 6

Top Mass Reconstruction

6.1 Reconstruction using Monte Carlo Knowledge

Having reconstructed the W mass from the hadronically decaying W , we now want to reconstruct the top mass. We will do so for the jets found by the iterative cone algorithm using the cone size 0.5.

We need to know which b -jet is the partner of the hadronically decaying W , i.e. the b quark originating from the top decay which emitted the (later) hadronically decaying W . At first, we use the Monte Carlo information to find it, which is done via the matching procedure already described.

6.1.1 Using the Jet Energy Corrections of the W -Jets

At first, we use jet energy corrections we obtained from the W -jets for the b -jets. Our analysis is partially realistic and partially based on MC. To reconstruct the top mass, we use the reconstructed W bosons in the 2 standard deviation range from the realistic reconstruction of the W mass using cone size 0.5 (see section 5.3). The corresponding quark is selected by using Monte Carlo data: in the MC data we look up which b belongs to the hadronically decaying W .

Using this information, we obtain figure 6.1. The reconstructed top mass is 166.1 GeV.

6.1.2 Using the b -Jet Energy Corrections

Another approach for correcting the energy of the b -jets is to use special b -jet energy corrections. These are calculated in the very same way as the corrections for jets originating from the hadronically decaying W boson.

Having obtained these, reconstruct the top mass in the same way as in the previous section. The results can be seen in figure 6.2.

The ratio $\frac{\text{standard deviation}}{\text{mean}}$ is 0.1175, a little less than the one without using b -jet corrections, 0.1191. The reconstructed top mass is by far closer to the Monte Carlo value when using b -jet corrections.

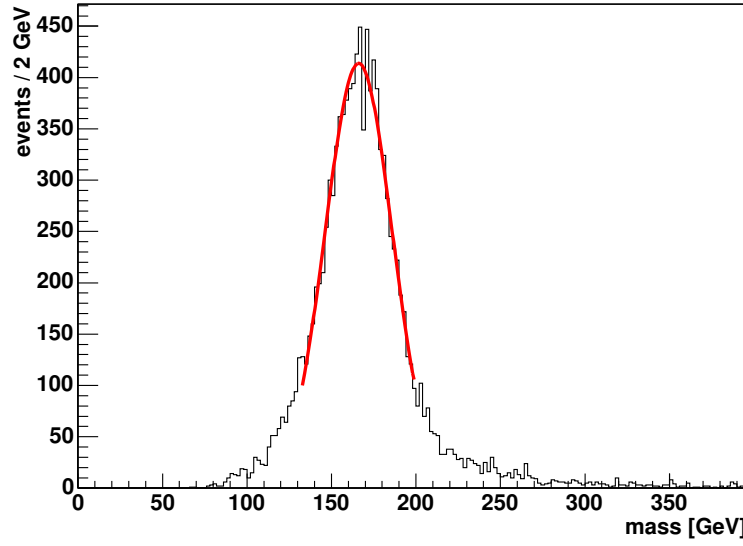


Figure 6.1: Top mass obtained using tight Monte Carlo matching for finding the b -jet belonging to the hadronically decaying W , assuming perfect b - and muon-tagging. The b -jet energy was corrected using the jet energy corrections obtained from the hadronically decaying W 's. Only events with a realistically reconstructed W mass in the 2 standard deviation range are taken. There are 11752 events. In the $1.5\text{-}\sigma$ range a Gaussian was fitted. It has a mean of 166.1 GeV and standard deviation 19.78 GeV.

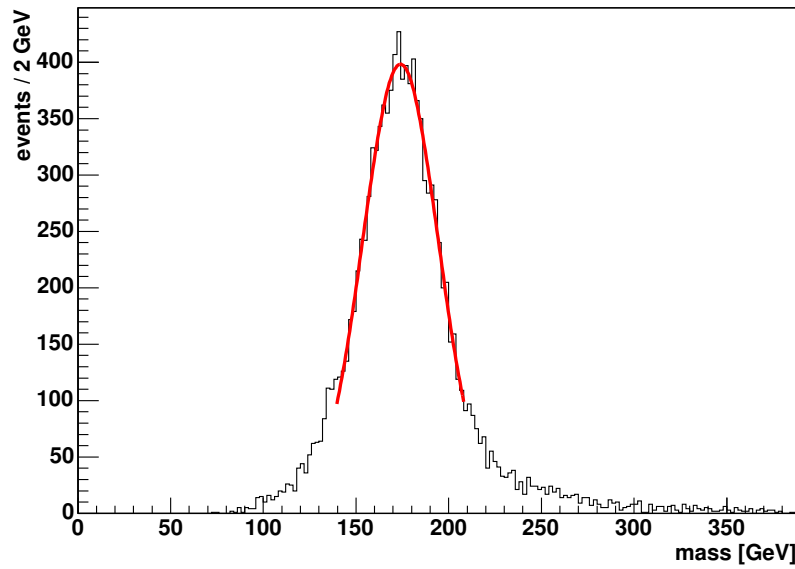


Figure 6.2: Top mass reconstructed using Monte Carlo information for finding the b -jet belonging to the hadronically decaying W , assuming perfect b - and muon-tagging. b -jet corrections are used. Only events with a realistically reconstructed W mass in the 2 standard deviation range are taken. There are 11752 events. In the $1.5\text{-}\sigma$ range a Gaussian was fitted. It has a mean of 174.0 GeV and standard deviation 20.44 GeV.

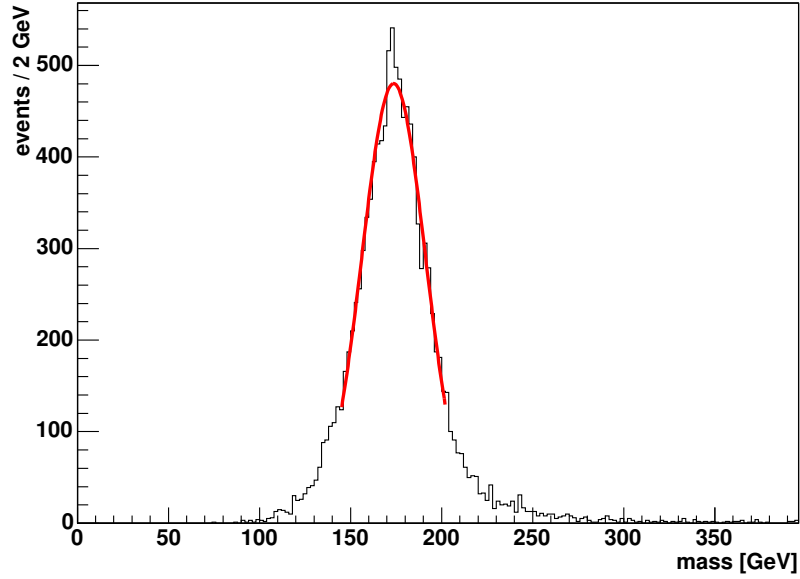


Figure 6.3: Top mass reconstructed without using Monte Carlo information, using b -jet energy corrections, choosing the best combination by comparing the invariant mass to the world average. Only events with a realistically reconstructed W mass in the 2 standard deviation range are taken. There are 11752 events. In the $1.5\text{-}\sigma$ range a Gaussian was fitted. It has a mean of 173.7 GeV and standard deviation 17.49 GeV.

6.2 Realistic Reconstruction

After having reconstructed the hadronically decaying W boson, we can choose between two b -jets for reconstructing the top.

Two approaches for doing so are given in the following subsections¹.

6.2.1 Best Mass approach

A rather naive approach is to calculate the invariant top masses from each combination and to choose that combination that is closer to the Monte Carlo value.

Using this approach we get figure 6.3. As can be clearly seen, this approach yields 'better results' than using Monte Carlo information. The plot is pulled around 174 GeV, because it is obviously biased the top mass world average. This approach cannot be used for a realistic reconstruction.

¹Working on a $t\bar{t}$ sample, there is the possibility of reconstructing the second top quark. This offers the criterion of taking the combination with the smallest mass difference between the two top quarks. Because the leptonically decaying W was not reconstructed in this analysis, this will not be done.

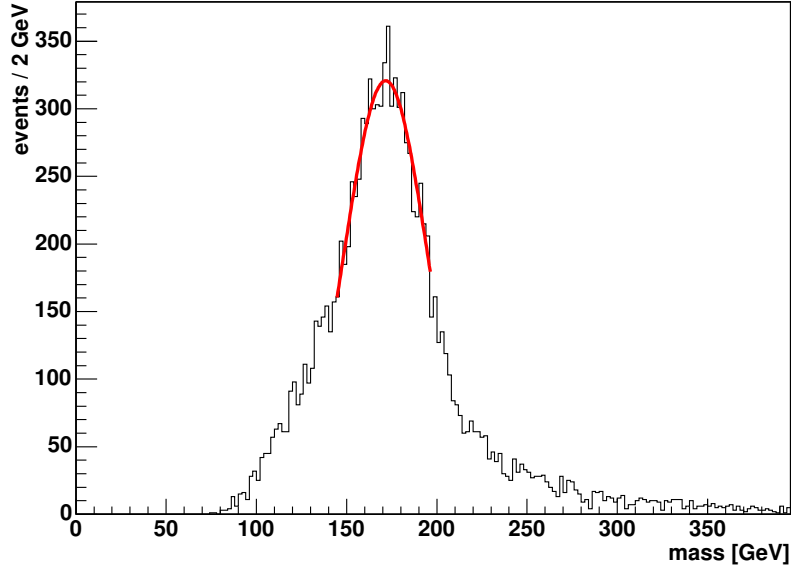


Figure 6.4: Top mass reconstructed, using b -jet energy corrections. The b -jet corresponding to the reconstructed W is chosen by using the b -jet having the smallest angle to the reconstructed W . Only events with a realistically reconstructed W mass in the 2 standard deviation range are taken. There are 11752 events. In the $1\text{-}\sigma$ range a Gaussian was fitted. It has a mean of 171.6 GeV and standard deviation 22.94 GeV.

6.2.2 Using the Angles

Recalling the event topology (see figure 5.1 on page 38), one expects that the angle between the hadronically decaying W and the corresponding b -jet is smaller than the angle between the hadronically decaying W and the other b -jet. So we use this as a criterion to decide which b -jet to use for reconstructing the top mass.

The resulting plot of this approach is figure 6.4. In the $1\text{-}\sigma$ range, we get a top mass of 171.6 GeV and a standard deviation of 22.94 GeV. This results in a standard deviation/mean ratio of 0.133.

Chapter 7

Conclusion and Outlook

In this thesis, a jet energy correction scheme, relying on only 12 parameters for barrel and endcap, respectively, was presented. The energy corrections depend on the absolute value of the pseudo-rapidity, $|\eta|$, and on the energy at reconstruction level. The scheme is general enough to be applied to all kinds of jet algorithms. In this thesis, it was applied to jets of hadronically decaying W bosons and to b -jets from $t\bar{t} \rightarrow WWb\bar{b} \rightarrow \mu + \nu_\mu + jets$. The jet algorithms considered were the iterative cone algorithm with cone sizes of 0.5, 0.6, 0.7 and 0.8 and the k_T algorithm in the inclusive and in the exclusive modes.

The large data sample of more than 200,000 $t\bar{t}$ events was produced on the computing facilities at the Forschungszentrum Karlsruhe. Half of the events were used for deriving the jet energy corrections, the other half was used for reconstruction of W and top mass. This special production with complete detector simulation was necessary for a realistic study, because no official dataset of similar size for deriving jet energy corrections existed so far.

The energy corrections were used to reconstruct the W mass from jets tightly matched to the quarks of the Monte Carlo. Here the iterative cone algorithm with a cone size of 0.5 was used. Without the use of the corrections, the standard deviation/mean ratio of the W mass was 19%, using the corrections the resolution shrank to 13%, which is about 31 % lower. These results can be compared with an earlier study done by Volker Drollinger (see [Dro1]). There the fast detector simulation CMSJET, which includes energy and spatial smearing, instead of a full detector simulation, was used. Using a tight matching criterion and a cone algorithm with cone size 0.4, a standard deviation/mean ratio for the corrected W mass of 11% was obtained. This result is slightly smaller than in this thesis, but it is essentially confirmed.

Tight matching of the Monte Carlo quarks and the jet directions suppresses events with gluons radiated off quarks. In order to have a more realistic approach, the W mass is reconstructed without any Monte Carlo constraints on the jets. However, perfect tagging of the μ and the b quarks was assumed. Using the jet energy corrections, the W mass was reconstructed of the two highest jets which had not been tagged. This was done for the jet finding algorithms given above. Overall, the best algorithm was the iterative cone algorithm with cone size of 0.5, which resulted in a reconstructed W mass within the error

given by the available Monte Carlo statistics, a standard deviation/mean ratio of 11.8% and an efficiency of 42.3 %, which is the percentage of events with a reconstructed W mass within two standard deviations of the reconstructed W mass.

Finally, the top mass was reconstructed realistically, using the iterative cone algorithm with a cone size of 0.5, again assuming perfect b - and muon-tagging. If the two highest energy jets which were not tagged could be tagged as W -jets, i.e. their reconstructed mass was within two standard deviations of the reconstructed W mass, this event was used for reconstructing the top. The b with the smaller angle in respect to the reconstructed W -jet was taken for reconstructing the top. A top within the error given by the available Monte Carlo statistics and a standard deviation/mean ratio of 13.3 % were obtained.

These checks show that the jet energy correction scheme can be used tagging W bosons and top quarks.

In the future it should be investigated, which improvements on reconstructing the top quark mass can be achieved, when using the jet energy correction scheme developed in this thesis in combination with PAX.

Zusammenfassung

Quarks und Leptonen sind die fundamentalen Bausteine der Materie. Im Standardmodell werden sechs Quarks und sechs Leptonen erwartet. Sie sind jeweils in drei Familien von Dubletts eingeteilt. Das Top-Quark, das als letztes entdeckte Quark, ist mit Sicherheit das interessanteste Quark. Es unterscheidet sich insbesondere durch seine große Masse von den anderen Quarks. Der große Massenunterschied zum Bottom-Quark macht es besonders wichtig bei Strahlungskorrekturen. Auf der Suche nach dem letzten fehlenden Teilchen des Standardmodells spielt das Top-Quark auch eine wesentliche Rolle sowohl bei den Signalereignissen, z. B. bei der assoziierten Higgs-Produktion im $t\bar{t}H$ -Kanal, als auch in Hintergrundereignissen, z.B. im sogenannten goldenen Kanal $H \rightarrow ZZ^* \rightarrow 4\mu$.

In dieser Arbeit werden Jet-Energie-Korrekturen für hochenergetische Jets, die aus Top-Quark-Zerfällen kommen, entwickelt. Dies wird auf einem großen, vollständig simulierten Datensatz durchgeführt. Dieser Datensatz wurde speziell für diese Diplomarbeit produziert.

Im ersten Kapitel werden der großen Hadronen-Beschleuniger LHC sowie der CMS-Detektor beschrieben. Dabei werden die verschiedenen Detektorkomponenten sowie das Triggering erläutert.

Im zweiten Kapitel wird ausführlich auf das Top-Quark eingegangen. Auf eine kurze historische Einordnung folgt eine knappe Beschreibung des Standardmodells der Teilchenphysik. Danach werden folgende Themen erläutert: Masse des Top-Quarks, Zerfall des Top-Quarks, die Effekte der kurzen Lebensdauer sowie die Top-Quark Erzeugung in Paaren als auch einzeln. Für diese Arbeit ist besonders der darauf folgende Abschnitt über Zerfälle von Top-Quark-Paaren wichtig, die sich in drei Kanäle (di-leptonisch, semi-leptonisch und Multi-Jet) aufteilen. Abschließend wird auf radiative Korrekturen sowie den Zusammenhang zwischen Top-Quark-Masse, W -Boson-Masse und Higgs-Boson-Masse eingegangen.

Kapitel drei beschreibt die beiden wichtigsten Jet-Algorithmen für Hadronenbeschleuniger, den iterativen Cone Algorithmus und den k_T Algorithmus. Beide Algorithmen werden in dieser Arbeit benutzt. Zudem wird auf die Implementierung der Algorithmen in der CMS-Software ORCA eingegangen.

In Kapitel vier werden die Software-Komponenten der CMS-Distribution, die zur aufwändigen Produktion der vollständig simulierten Ereignisse gebraucht werden, erläutert. Danach wird auf das Datenanalyse-Paket PAX, welches auf einer hohen Abstraktionsebene arbeitet, eingegangen, da im Verlaufe dieser Arbeit wichtige Beiträge dazu erbracht wurden. Ausführlich wird die Klasse der

'Kombinierten Objekte' besprochen, die Daten aus dem Tracker und den Kalorimetern kombiniert. Diese Klasse wird in Zukunft eine Möglichkeit sein, die Auflösung des Detektors weiter zu verbessern.

In Kapitel fünf werden nun die Jet-Energie-Korrekturen eingeführt. Die produzierten Ereignisse sind $t\bar{t}$ -Ereignisse mit semi-leptonischem Zerfall, d.h. das eine Top-Quark zerfällt unter Aussendung eines W -Bosons in ein Bottom-Quark, wobei das W -Boson in Jets zerfällt, das andere Top-Quark zerfällt durch Aussendung eines W -Bosons in ein Bottom-Quark, wobei das W -Boson in ein geladenes Lepton und ein entsprechendes Neutrino zerfällt. In dieser Studie ist das Lepton ein Myon.

Zuerst wird die W -Boson Masse sowohl aus den generierten Jets (das sind die Jets auf Generatorniveau) als auch aus den rekonstruierten Jets (das sind die Jets auf Rekonstruktionsniveau) berechnet. Um jeweils die richtigen Jets zu benutzen, wurde ein enges Matching ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \leq 0,16$) an die Monte Carlo Quarks angewendet. Die so erhaltene W -Boson-Masse beträgt 78,57 GeV mit einer Standardabweichung von 6 GeV auf Generatorniveau und 56,18 GeV mit einer Standardabweichung von 10,49 GeV auf Rekonstruktionsniveau. Vergleicht man das Verhältnis der Standardabweichung zur Masse, so ist der Wert auf Rekonstruktionsniveau mehr als doppelt so groß wie auf Generatorniveau. Um die Auflösung auf Rekonstruktionsniveau zu verbessern, werden die Jet-Energie-Korrekturen eingeführt.

Diese Korrekturen werden durch den Vergleich der Energie des Jets mit der Energie des eng gematchten Quarks für Intervalle in Energie und in dem Betrag der Pseudo-Rapidity, $|\eta|$, berechnet. In sogenannten 'Eta-Scheiben' werden die durchschnittlichen Korrekturen in den einzelnen Intervallen für einen bestimmten $|\eta|$ -Bereich nach der Energie aufgetragen; mehrere 'Eta-Scheiben' in verschiedenen Detektorbereichen sind in Abbildung 7.1 gegeben. Daran wird nun eine Kurve von folgendem Typ angepasst:

$$f = a + b \cdot \left(\frac{E}{\text{GeV}} \right)^c. \quad (7.1)$$

Dies wird für jede Eta-Scheibe wiederholt. In einem iterativen Verfahren werden dann die Parameter a , b und c jeweils für Barrel- und Endkappen-Bereich durch Polynome dritten Grades in $|\eta|$ gefittet. Somit ergibt sich für die beiden Bereiche jeweils eine von zwölf Parametern abhängige Funktion zur Beschreibung der Jet-Energie-Korrekturen.

Nun kann die W -Boson-Masse mit engem Matching der rekonstruierten Jets und den Jet-Energie-Korrekturen berechnet werden. Dabei ergibt sich ein Wert von 81,69 GeV mit einer Standardabweichung von 10,55 GeV. Dies ist eine um 31% bessere Auflösung wie ohne die Anwendung der Jet-Energie-Korrekturen.

Natürlich sind wir aber an einer realistischen Methode zur Rekonstruktion des hadronisch zerfallenden W -Bosons interessiert. Im folgenden werden wir die Monte Carlo Informationen ausschließlich zur Simulation eines perfekten b - und Myon-Taggings benutzen. Das Myon-Tagging ist deshalb notwendig, weil die Jet-Algorithmen es als einen Jet betrachten.

Das Verfahren für die realistische Rekonstruktion ist wie folgt: Wir betrachten die ungetaggten Jets des Ereignisses. Zeigen die beiden höchstenergetischen

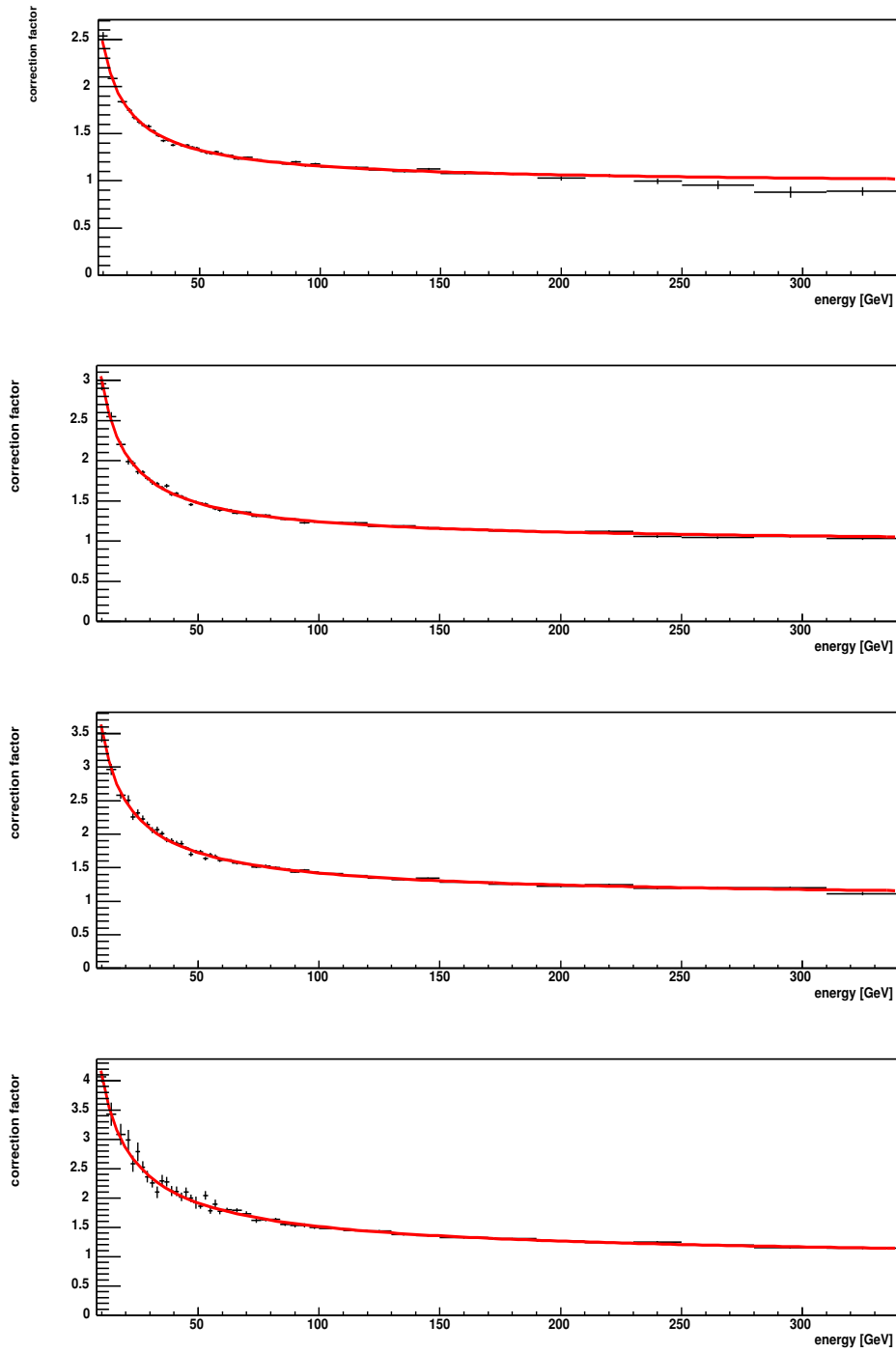


Abbildung 7.1: 'Eta-Scheiben' für verschiedene Werte von $|\eta|$. Die verschiedenen Regionen sind (von oben nach unten): Innerer Barrel-Bereich $0,2 < |\eta| < 0,4$ und $0,8 < |\eta| < 1,0$, äußerer Barrel-Bereich $1,2 < |\eta| < 1,4$ und innerer Endkappen-Bereich $1,8 < |\eta| < 2,0$.

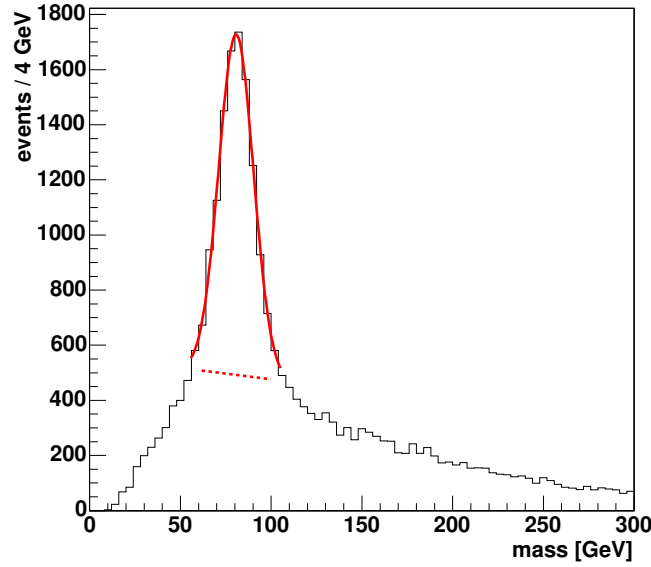


Abbildung 7.2: Realistisch rekonstruierte W -Boson-Masse aus Ereignissen mit fünf bis neun Jets, die vom Iterativen Cone Algorithmus mit Cone Größe 0,5 gefunden wurden. Es wurde ein perfektes b - und Myon-Tagging durch enges Matching an die MC-Daten simuliert. Falls von den nicht getaggtten Jets die zwei energiereichsten jeweils eine Energie von mehr als 25 GeV hatten und zudem jeweils ihre Richtung nicht auf den Barrel/Endkappen-Überlagerungsbereich zeigt, dann wird aus ihnen nach Anwendung der Jet-Energie-Korrekturen die W -Boson-Masse bestimmt. Das Histogramm hat 27398 Einträge. Die Gaußkurve hat ihren Mittelwert bei 80,68 GeV und eine Standardabweichung von 9,54 GeV. Die Gerade hat die Steigung $1236 \frac{\text{Ereignisse}}{\text{GeV}}$ und einen Wert von $492 \frac{\text{Ereignisse}}{4 \text{ GeV}}$ am Mittelwert der Gaußverteilung.

Jets jeweils nicht in den Überlagerungsbereich von Barrel und Endkappe und haben sie zudem eine rekonstruierte Energie von mehr als 25 GeV, so wenden wir die Jet-Energie-Korrekturen auf die beiden Jets an und rekonstruieren dann aus ihnen die W -Boson-Masse. Dieses Verfahren wird mit dem iterativen Cone Algorithmus für unterschiedliche Cone-Größen und dem k_T Algorithmus im inklusiven und exklusiven Modus durchgeführt. Als Beispiel ist eine Verteilung der rekonstruierten W -Masse für den Iterativen Cone Algorithmus mit Cone Größe 0,5 in Abbildung 7.2 dargestellt.

Im Bereich um den Peak wird die Funktion

$$c \cdot e^{-\frac{(m-\mu)^2}{2\sigma^2}} + s \cdot (m - \mu) + b \quad (7.2)$$

an die Daten angepasst. Diese Funktion stellt eine Gaußkurve auf einer Geraden dar.

In Tabelle 7.1 werden die Ergebnisse für verschiedene Cone Größen beim

Iterativen Cone Algorithmus, sowie für inklusiven und exklusiven¹ Modus des k_T Algorithmus aufgelistet.

Der iterative Cone Algorithmus mit Cone Größe 0,5, der sich bereits bei CMS bewährt hat, schneidet insgesamt am besten ab. Er wird deshalb auch im anschließenden Kapitel verwendet.

In Kapitel sechs wird die Rekonstruktion von Top-Quark-Massen besprochen. Dazu werden spezielle Korrekturen für b -Jets eingeführt und getestet. Bei einer Rekonstruktion der Top-Quark-Masse mit engem Matching an Monte Carlo Daten stellt sich heraus, dass bei ungefähr gleichem Verhältnis von Standardabweichung zur rekonstruierten Masse die Top Masse um 8 GeV besser rekonstruiert wird, wenn man für die b -Jets die entsprechend bestimmten Korrekturen² statt der Korrekturen, die man für die W -Jets berechnet hat, benutzt.

Schließlich wird die Top-Quark-Masse unter Annahme von perfektem b - und μ -Tagging rekonstruiert. Dabei werden die Ereignisse benutzt, bei denen die realistische W -Boson-Massenrekonstruktion eine Masse im Bereich von 2 Standardabweichungen um die in Kapitel fünf rekonstruierte W -Boson-Masse erbracht hat. Der b -Jet, der einen kleineren Winkel mit dem rekonstruierten W -Boson einschließt, wird zur Top-Quark-Massenrekonstruktion genutzt. Man erhält somit Abbildung 7.3. Das Verhältnis von Standardabweichung zur rekonstruierten Masse ist 13,3%.

In Kapitel sieben werden die Ergebnisse der Arbeit zusammengefasst. Zudem wird festgestellt, dass das von Volker Drollinger in einer Studie mit schneller Detektorsimulation (siehe [Dro1]) erhaltene Verhältnis von Standardabweichung zur rekonstruierten W -Boson-Masse unter Benutzung von engem Monte Carlo-Matching hier grundsätzlich bestätigt wird.

Es wäre interessant zu untersuchen, welche weiteren Verbesserungen in der

¹Es gibt im exklusiven Modus die Möglichkeit, eine bestimmte Anzahl von Jets zu verlangen. In dieser Analyse wurden fünf Jets gefordert.

²Das Verfahren ist das gleiche wie bei der Bestimmung der Korrekturfaktoren der W -Jets und ist schnell durchgeführt.

Algorithmus	<u>Standardabweichung</u> rekonstruierte Masse	rekonstruierte W -Boson Masse [GeV]	Effizienz
Iterativer Cone 0,5	0,118	80,7	0,421
Iterativer Cone 0,6	0,121	80,4	0,413
Iterativer Cone 0,7	0,139	79,9	0,391
Iterativer Cone 0,8	0,142	79,6	0,386
Inklusiver k_T	0,143	80,4	0,391
Exklusiver k_T (5 Jets)	0,152	84,4	0,514

Tabelle 7.1: Vergleich der Resultate der verschiedenen Jet-Algorithmen. Beim iterativen Cone Algorithmus werden Ereignisse mit fünf bis zwölf Jets, beim inklusiven k_T Algorithmus Ereignisse mit fünf bis fünfzehn Jets betrachtet. Beim inklusiven k_T Algorithmus werden alle Ereignisse betrachtet.

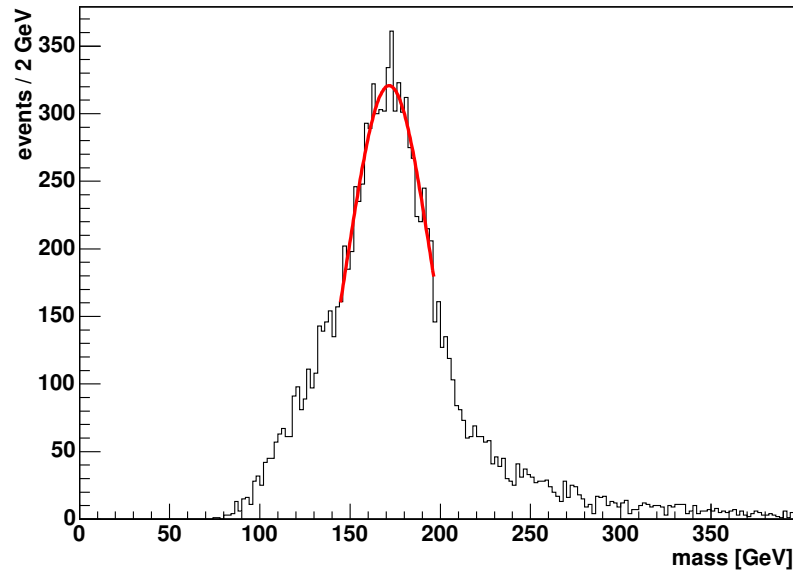


Abbildung 7.3: Top-Quark-Massenrekonstruktion unter Annahme von perfektem b - und μ -Tagging. Zur Auswahl des b -Quarks werden die Winkel zwischen rekonstruiertem W -Boson und den b -Jets benutzt. Der b -Jet, der mit dem rekonstruierten W -Boson den kleineren Winkel einschließt, wird gewählt. Das Histogramm hat 11752 Einträge, die Gaußkurve (in rot) hat ihren Mittelwert bei 171,6 GeV und eine Standardabweichung von 22,94 GeV.

Top-Massen-Rekonstruktion durch die Kombination der hier entwickelten Jet-Energie-Korrekturen mit PAX erreicht werden können.

Bibliography

- [BloLoh] V. Blobel und E. Lohrmann, “Statistische und numerische Methoden der Datenanalyse”, Teubner Studienbücher, 1988.
- [Bro] G. Brooijmans [CDF and D0 Collaborations], “Top quark mass measurements at the Tevatron”, arXiv:hep-ex/0005030.
- [ChaKon] D. Chakraborty, J. Konigsberg and D. Rainwater, “Review of top quark physics”, arXiv:hep-ph/0303092.
- [CMSIM] CMS detector simulation software group, “CMS Simulation Package CMSIM, Users’ Guide and Reference Manual”, <http://cmsdoc.cern.ch/cmsim/cmsim.html>
- [CMSOutreach] CMS Outreach, “Compact Muon Solenoid Outreach Activities”, <http://cmsinfo.cern.ch>
- [CompHep] A. Pukhov *et al.*, “CompHEP: A package for evaluation of Feynman diagrams and integration over multi-particle phase space. User’s manual for version 33”, arXiv:hep-ph/9908288.
- [CotGre] W. Cottingham and D. Greenwood, “An introduction to the Standard Model of Particle Physics”, Cambridge University Press, 1998.
- [Daw] S. Dawson, “The top quark, QCD, and new physics”, arXiv:hep-ph/0303191.
- [Dro1] V. Drollinger, “Jet Energy Corrections”, CMS Note 1999/066, 1999.
- [Dro2] V. Drollinger, “Reconstruction and Analysis Methods for Searches of Higgs Bosons in the Decay Mode $H^0 \rightarrow b\bar{b}$ at Hadron Colliders”, PhD Thesis, 2001.
- [Durham] Durham Database Group at Durham University(UK), “The Durham HEP Databases”, <http://durpdg.dur.ac.uk/hepdata/>
- [GEANT] “GEANT - Detector Description and Simulation Tool”, <http://wwwasd.web.cern.ch/wwwasd/geant/>
- [GridKa] Grid Computing Center Karlsruhe, “GridKa homepage”, <http://grid.fzk.de/>

- [Gri] D. J. Griffiths, “Introduction To Elementary Particles”, John Wiley and Sons, 1987.
- [Grü] M. W. Grünewald, “Electroweak precision data: Global Higgs analysis”, arXiv:hep-ex/0304023.
- [Gui] J. Guillaud, “PYTHIA miniguide”,
available at <http://lappc-in39.in2p3.fr/Pythia/>
- [HalMar] F. Halzen and A. D. Martin, “Quarks And Leptons: An Introductory Course In Modern Particle Physics”, John Wiley & Sons, 1984.
- [Hei] A. Heister, “Status: ORCA jet subsystems”, talk given at the JetMET Meeting in December 2003
- [HERWIG] G. Corcella *et al.*, “HERWIG 6.5 release note”, arXiv:hep-ph/0210213.
- [ISAJET] H. Baer, F. E. Paige, S. D. Protopescu and X. Tata, “ISAJET 7.69: A Monte Carlo event generator for p p, anti-p p, and e+ e- reactions”, arXiv:hep-ph/0312045.
- [Kir] M. Kirsch, Diploma thesis, in preparation
- [KtJet] J. M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh, “Kt-Jet: A C++ implementation of the K(T) clustering algorithm”, Comput. Phys. Commun. **153** (2003) 85 [arXiv:hep-ph/0210022].
- [Kun] Z. Kunszt, “Jet physics at hadron colliders”, Nucl. Phys. A **715** (2003) 180 [arXiv:hep-ph/0212013].
- [LEPEWWG] The LEP Electroweak Working Group, “LEP EWWG homepage”, <http://lepewwg.web.cern.ch/LEPEWWG/>
- [ORCA] “CMS OO Reconstruction”, <http://cmsdoc.cern.ch/orca/>
- [PAX] M. Erdmann *et al.*, “Physics analysis expert PAX: First applications,” eConf **C0303241** (2003) THLT008 [arXiv:physics/0306085].
- [PDG] Particle Data Group, “Particle Physics Booklet”, 2002.
- [Per] D. H. Perkins, “Introduction To High-Energy Physics”, Cambridge University Press, 4th edition, 2000.
- [PYTHIA] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, “High-energy-physics event generation with PYTHIA 6.1”, Comput. Phys. Commun. **135** (2001) 238, [arXiv:hep-ph/0010017].
- [ROOT] R. Brun, F. Rademakers, S. Panacek, I. Antcheva and D. Buskulic, “ROOT Users Guide 3.05”, <http://root.cern.ch/>
- [Ros] J. L. Rosner, “The standard model in 2001”, arXiv:hep-ph/0108195.

- [Schm] P. Schmüser, “Feynman-Graphen und Eichtheorien für Experimentalphysiker”, Springer, 1995.
- [SCRAM] S. Ashby, “SCRAM User Manual”, 2003. See also <http://spi.cern.ch/cgi-bin/scrampage.cgi>
- [Sim] E. H. Simmons, “The top quark: Experimental roots and branches of theory. ((U)) ((V))”, arXiv:hep-ph/0211335.
- [SmiWil] M. C. Smith and S. Willenbrock, “QCD and Yukawa Corrections to Single-Top-Quark Production via $q \bar{q} \rightarrow t \bar{b}$ ”, Phys. Rev. D **54** (1996) 6696 [arXiv:hep-ph/9604223].
- [Son] L. Sonnenschein, “Top mass determination in the $t\bar{t}$ semileptonic decay channel”, CMS Note 2001/001, 2001.
- [StSuWi] T. Stelzer, Z. Sullivan and S. Willenbrock, “Single top quark production at hadron colliders”, Phys. Rev. D **58** (1998) 094021 [arXiv:hep-ph/9807340].
- [Wei] C. Weiser, “Status of b-tagging”, talk given at the PRS b/τ meeting, October 2001
- [Wil1] S. Willenbrock, “The road ahead”, arXiv:hep-ph/0307324.
- [Wil2] S. Willenbrock, “The standard model and the top quark. ((U))”, arXiv:hep-ph/0211067.
- [Wil3] S. Willenbrock, “Thinking about top within the standard model,” arXiv:hep-ph/9905498.
- [Zhu] S. Zhu, “Next-To-Leading Order QCD Corrections To $B \bar{G} \rightarrow T W$ -At The Cern Large Hadron Collider”, Phys. Lett. B **524** (2002) 283 [Erratum-ibid. B **537** (2002) 351].

Danksagung

Zuerst gilt mein Dank natürlich Herrn Prof. Dr. Günter Quast, bei dem ich mich immer sehr gut betreut gefühlt habe. Er ließ mir Platz für eigene Ideen, wusste aber immer auch interessante Aufgaben zu stellen. In keinem anderen Jahr des Studiums habe ich soviel gelernt wie im letzten.

Ebenso danke ich Herrn Prof. Dr. Michael Feindt für die Übernahme des Korreferats, sowie für die hilfreichen Kommentare zu dieser Arbeit.

Mein Dank gilt selbstverständlich auch Herrn Dr. Klaus Rabbertz, der sich nicht nur um eine funktionsfähige und aktuelle CMS-Software kümmerte, sondern auch bei physikalischen Fragestellungen ein guter Ansprechpartner war.

Meine Kollegen haben dafür gesorgt, dass die tägliche Arbeit nicht zum täglichen Arbeitsfrust wurde, sondern dass ich gerne zur Arbeit gegangen bin. Natürlich sollen ihre Namen nicht ungenannt bleiben: Ullrich Felzmann, Sabine Feyen, Deborah Miksat, Christian Piasecki, Sven Schalla und Alexander Schmidt.

Zu unseren Meetings gehörten immer Telefonate mit unseren Kollegen vom CERN. Dies brachte immer einen wertvollen, physikalischen Input. Danke an Steffen Kappler, Christian Weiser und Joanna Weng.

Vielen Dank an alle Korrekturleser! Die meisten sind schon erwähnt worden, die 'externen' möchte ich noch namentlich erwähnen: Ingo Giehl, Matthias Kirsch, meine Freundin Simone Klotzbach, Martin Schellhas, Dirk Schuricht und Antje Stotz.

Die Systemadministratoren am EKP haben einen harten Job. Dafür, dass sie die Computer so gut am Laufen halten (es gibt ja Einrichtungen, da läuft es ganz anders), gebührt ihnen mein Dank: Yves Kemp, Ulrich Kerzel, Jens Rehn, Christian Sander, Thorsten Scheidle, Patrick Schemitz und Marcel Stanitzki.

Besonders möchte ich meinen Eltern dafür danken, dass sie mir dieses Studium ermöglicht haben und mich immer moralisch und finanziell unterstützt haben.

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Hilfsmittel verwendet zu haben.

Christopher Jung

Karlsruhe, den 23. Januar 2004