Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

IEKP-KA/2008-1

Studien zur Messung der W-Boson-Helizität im Top-Quark-Zerfall mit dem CMS Experiment

Georg Sartisohn

Diplomarbeit

bei Priv.-Doz. Dr. W. Wagner
Institut für Experimentelle Kernphysik

Korreferent Prof. Dr. Th. Müller
Institut für Experimentelle Kernphysik

an der Fakultät für Physik
der Universität Karlsruhe (TH)

27. November 2007

Zusammenfassung

Das schwerste bisher bekannte Elementarteilchen, das Top-Quark, wurde 1995 von den Experimenten CDF und DØ am Proton-Antiproton Beschleuniger Tevatron entdeckt. Das Tevatron befindet sich am Fermilab in der Nähe von Chicago und ist, bis zur Inbetriebnahme des Large Hadron Colliders (LHC) am CERN in der Nähe von Genf, der Teilchenbeschleuniger mit der zur Zeit höchsten Schwerpunksenergie von 1.96 TeV. Die Masse des Top-Quarks wurde bereits mit sehr hoher Präzesion gemessen, jedoch sind die Messungen weiterer Eigenschaften des Top-Quarks immer noch statistisch limitiert. Dies wird sich nach dem Start des LHC ändern, da der Wirkungsquerschnitt für die Erzeugung von Top-Antitop-Quark-Paaren bei der Schwerpunksenergie von 14 TeV des LHC 833 pb beträgt und damit um einen Faktor 100 größer ist als am Tevatron.

Die große Masse des Top-Quarks ist der Grund dafür, dass es zerfällt bevor es Bindungszustände mit anderen Quarks eingehen kann. Diese Tatsache erlaubt es, die Eigenschaften eines nahezu freien Quarks studieren zu können, da diese Eigenschaften direkt auf die Zerfallsprodukte übertragen werden. Es ist sogar möglich, die Spinausrichtung des Top-Quarks aus den Eigenschaften seiner Zerfallsprodukte zu rekonstruieren.

Der Zerfall des Top-Quarks verläuft in nahezu 100% aller Fälle über ein b-Quark und ein reelles W-Boson. Dieser Zerfall wird im Standardmodell der Elementarteilchenphysik mit Hilfe der schwachen Wechselwirkung beschrieben. Da diese nur an linkshändige Teilchen und rechtshändige Antiteilchen koppelt, sind in diesem Zerfall nicht alle Spinkombinationen von b-Quark und W-Boson erlaubt. Die Richtung des Spins eines Teilchens im Bezug auf seine Bewegungsrichtung wird durch die Händigkeit dieses Teilchens beschrieben. Bei rechtshändigen Teilchen ist der Spin parallel zur Bewegungsrichtung und bei linkshändigen Teilchen antiparallel zur Bewegungsrichtung eingestellt. Für massebehaftete Spin 1 Teilchen, wie zum Beispiel das W-Boson, gibt es noch eine weitere Möglichkeit für die Einstellung des Spins, nämlich senkrecht zur Bewegungsrichtung. Diese Teilchen werden als longitudinal polarisiert bezeichnet.

Das Standardmodell macht Vorhersagen für die Anteile der unterschiedlichen Polarisationen der W-Bosonen im Top-Quark-Zerfall. Für eine Top-Quark-Masse von 173,7 GeV/c^2 beträgt der Anteil longitudinal polarisierter W-Bosonen genau F_0 70% und der Anteil linkshändiger W-Bosonen F_- 30%, während der Anteil F_+ an rechtshändi-gen W-Bosonen aufgrund des großen Masseunterschieds zwischen dem Top-Quark und dem b-Quark stark unterdrückt ist. Vernachlässigt man die deutlich geringere Masse des b-Quarks von 4,2 GeV/c^2, so sind keine rechtshändigen W-Bosonen erlaubt. Gravierende Abweichungen von diesen Vorhersagen wären starke

Hinweise auf neue Physik jenseits des Standardmodells. Die Messung dieser sogenannten Helizitätsanteile stellt somit einen wichtigen Test zur Überprüfung des Standardmodells und der darüber hinausgehenden Theorien dar.

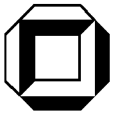
Das Ziel dieser Diplomarbeit ist es eine Messung dieser Helizitätsanteile am CMS Experiment vorzubereiten. Dazu werden von Monte-Carlo Ereignis-Generatoren erzeugte und mit der CMS-Software vollständig simulierte Top-Antitop-Quark Ereignisse verwendet. Diese simulierten Monte-Carlo Daten werden benutzt um eine Messmethode zu entwickeln. Im ersten Schritt werden jedoch Top-Antitop-Quark Ereignisse zweier Monte Carlo Ereignis Generatoren (Pythia und Herwig++) miteinander verglichen, um zu prüfen, ob große Unterschiede für die Kinematik von Top-Antitop-Quark Ereignissen bei den Schwerpunksenergien von 1,96 TeV und 14 TeV vorhergesagt werden. Bis auf den größeren Boost des Top-Antitop-Quark Paares am LHC werden von den Monte-Carlo Ereignis-Generatoren keine gravierenden Unterschiede für die Kinematik dieser Ereignisse vorhergesagt.

Da für die Messung der Helizitätsanteile der Kosinus des Zerfallswinkels θ^* benutzt wird, welcher als der Winkel zwischen dem Impuls des geladenen Leptons im Ruhesystem des W-Bosons und dem Impuls des W-Bosons im Top-Quark-Ruhesystem definiert ist, müssen aus den Top-Antitop-Quark Ereignissen diejenigen ausgewählt werden, bei denen mindestens eins der entstandenen W-Bosonen leptonisch, d.h. in ein geladenes Lepton (Elektron oder Myon) und ein Neutrino, zerfällt. Um die Top-Antitop-Quark Ereignisse gut rekonstruieren zu können, werden nur "lepton+jets" Ereignisse selektiert, bei denen eines der beiden W-Bosonen leptonisch und das andere hadronisch, d.h. in ein Quark q und ein Antiquark $\bar{q}'$, zerfällt. Bei diesen Ereignissen entsteht nur ein Neutrino, welches als fehlende Transversalenergie nach-gewiesen werden kann. Die Signatur dieser Ereignisse besteht folglich aus genau einem geladenen Lepton, fehlender Transversalenergie, zwei leichten Jets und zwei b-Jets. Diese Jets entstehen durch Hadronisation der Quarks in farbneutrale Objekte. Der Anteil von "lepton+jets" Ereignissen bezogen auf alle Top-Antitop-Ereignisse beträgt 29% und wird durch das Ausnutzen ihrer Signatur auf 85% nach der Selektion erhöht.

Um den Zerfallswinkel θ^* zu bestimmen müssen sowohl der Vierervektor des Top-Quarks als auch die Vierervektoren des W-Bosons und des geladenen Leptons rekonstruiert werden. Alle selektierten Ereignisse werden mit Hilfe ihrer Zerfallsprodukte vollständig rekonstruiert. Dabei ergeben sich mehrere Möglichkeiten die gemessenen Jets den Quarks aus dem Top-Quark Zerfall zuzuordnen, was für jedes Ereignis zu einer Vielzahl möglicher Hypothesen für die Vierervektoren der Top-Quarks führt. Um für jedes Ereignis eine Hypothese auszuwählen, die der wahren Ereignis-Topologie möglichst gut entspricht, wird jeder Hypothese eine Zahl Ψ zugewiesen. Bei der Berechnung dieser Zahl werden Eigenschaften der Topologie des Top-Quark-Zerfalls ausgenutzt, folglich stellt Ψ ein Maß dar, welches angibt, wie nahe die Hypothese an die Top-Paar-Zerfallstopologie herankommt.

Jedes rekonstruierte Ereignis trägt somit zu der $\cos\theta^*$-Verteilung bei. Diese Verteilung unterscheidet sich aber von der theoretisch erwarteten Kurve. Die Unterschiede sind durch die Akzeptanz der Selektion und durch Rekonstruktionseffekte bedingt. Diese Effekte müssen berücksichtigt werden um aus der erhaltenen $\cos\theta^*$-Verteilung die Werte für F_0 und F_+ bestimmen zu können. Für die Messung der Werte wird die Verteilung in Abschnitte unterteilt und für jeden dieser Abschnitte wird die Anzahl der erwarteten Ereignisse unter Berücksichtigung der gennannten Effekte berechnet. Anschließend wird dieser Verteilung mit einer Likelihood-Methode eine Kurve angepasst, um die beiden Parameter F_0 und F_+ zu bestimmen. Für das Wertepaar, bei dem die Likelihood-Funktion ihr Maximum erreicht, beschreibt die angepasste Kurve die gemessene $\cos\theta^*$-Verteilung am besten. Somit ist dieses Wertepaar das Ergebnis der Messung und kann mit der Theorie verglichen werden.

Diese ersten Studien zur Messung der W-Boson Helizität im Top-Quark-Zerfall am CMS Experiment zeigen, dass die vorgestellte Messmethode eine hohe Sensitivität aufweist. Für eine integrierte Luminosität von $\mathcal{L} = 1$ fb^{-1}, wird für die Messung von F_0 ein relativer statistischer Fehler von 5% erwartet. Die Erwartung für die statistische Abweichung für den angenommenen Wert von $F_+ = 0.0$ ist ± 0.016.



Universität Karlsruhe (TH)
Research University · founded 1825

Studies for the Measurement of the W Boson Helicity in Top Quark Decays with the CMS Experiment

Georg Sartisohn

Diploma Thesis

at the Faculty for Physics
of the University of Karlsruhe (TH)

Referee: Priv.-Doz. Dr. W. Wagner
Institut für Experimentelle Kernphysik

Co-Referee: Prof. Dr. Th. Müller
Institut für Experimentelle Kernphysik

November 27, 2006

Introduction

The most massive elementary particle known till now, the top quark, was discovered in 1995 by the CDF and DØ collaborations [1, 2] at the Tevatron proton-antiproton collider at Fermilab. Its mass is measured with high precision [3, 4] but the measurements of several other properties are still statistically limited. This will change after the start of the Large Hadron Collider at CERN. The top-quark production cross-section at the center-of-mass energy $\sqrt{s} = 14$ TeV of the Large Hadron Collider is $\sigma_{t\bar{t}} = 833$ pb [5], that is 100 times larger than at the Tevatron.

Due to their large masses top quarks decay before they can hadronize, giving the unique opportunity to study a nearly bare quark. Top quarks decay with nearly 100% into a real W boson and a b quark. The V–A structure of the electroweak interaction implemented in the Standard Model predicts that the W bosons from this decay are dominantly longitudinally polarized or left-handed, while right-handed W bosons are heavily suppressed due to the large mass difference between the top quark and the b quark.

Under the assumption of a massless b quark and a top quark mass of 173.7 GeV/c^2 the Standard Model predicts the fraction F_0 of longitudinal polarized W bosons to be 0.70. The fraction F_- of left-handed W bosons is predicted to be 0.30, while the prediction for the fraction F_+ of right-handed W bosons is zero. The changes of these values due to next-to-leading order corrections are very small. A significant deviation from the non zero value of F_+ would indicate new physics. The measurement of the fraction of right-handed W bosons is a test of the V–A structure of the electroweak interaction. A possible V+A coupling would lead to non vanishing F_+ values on the expense of the F_- fraction as predicted in some left-right symmetric models.

This thesis uses Monte Carlo generated top-antitop-quark events to prepare a measurement of F_0 and F_+ for the CMS experiment at the Large Hadron Collider. The helicity fractions are measured in top-antitop-quark events where one top quark decays semileptonically, while the other one decays hadronically. The signature of this channel, called lepton+jets channel, is one charged lepton, missing transverse energy, and at least four jets. The measurement method uses the cosine of the decay angle θ^*, which is defined as the angle between the momentum of the charged lepton in the W-boson rest-frame and the W boson momentum in the top-quark rest-frame. The used Monte Carlo samples are fully simulated using the CMS detector simulation in order to receive a realistic detector response.

The thesis is organized as follows. The first chapter gives a brief overview over the Standard Model of elementary particle physics. It emphasizes the production and

decay of top quarks and discusses the possibility to measure the helicity fractions of W bosons from the top-quark decay using the cosine of the decay angle θ^*. Chapter two describes the Large Hadron Collider and the CMS detector at CERN. The third chapter presents a comparison of $t\bar{t}$ events at center-of-mass energies of the Tevatron $\sqrt{s} = 1.96$ TeV and the Large Hadron Collider $\sqrt{s} = 14$ TeV. The selection of lepton+jets events is described in chapter four. Chapter five presents the full reconstruction of the kinematic quantities in the selected events. Finally, the measurement method is described in chapter six and then checked in chapter seven.

Contents

List of Figures

List of Tables

Chapter 1

Theory

This chapter describes the most important properties of the Standard Model of elementary particle physics, as well as the properties of the top quark, the heaviest particle yet known. In addition, we discuss the possibility to measure the W boson helicity in top quark decays.

1.1 The Standard Model of Elementary Particle Physics

The Standard Model of Elementary Particle Physics describes the nature at subatomic scales with fundamental particles and their interactions [6,7]. In the Standard Model (SM) matter is formed out of spin $\frac{1}{2}$ particles, called fermions, while interactions between these fermions are mediated by spin 1 particles, the gauge bosons. There are 12 fermions divided into two groups quarks and leptons, each containing six particles.

The six different quark flavors are up u, down d, charm c, strange s, top t, and bottom b. In addition to the spin quarks carry further quantum numbers, electric charge $+\frac{2}{3}(u, c, t)$ or $-\frac{1}{3}(d, s, b)$ of the elementary charge and one color charge of the three types (red, blue, or green). The color charge is the quantum number of the strong interaction, which has a special property. Colored particles are not observed, i.e. quarks always form colorless particles. They are confined into mesons, consisting of two quarks, or baryons, containing three quarks. All the quarks in baryons carry different colors, which leads in total to a colorless object, while mesons are build of a quark and an antiquark, carrying a certain color and the corresponding anticolor.

The second group of fermions is formed by leptons. In contrast to quarks leptons do not underly the strong interaction and hence do not carry a color charge. Only the electron e, the muon μ, and the tau τ have an elementary electric charge, while the corresponding neutrinos ν_e, ν_μ, and ν_τ are neutral.

name	symbol	el. charge $[e]$	mass $[\mathrm{MeV}/c^2]$
up quark	u	$\frac{2}{3}$	$(1.5 - 3.0)$
down quark	d	$-\frac{1}{3}$	$(3.0 - 7.0)$
electron	e	-1	0.511
e-neutrino	ν_e	0	$< 2 \cdot 10^{-6}$
charm quark	c	$\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
strange quark	s	$-\frac{1}{3}$	(95 ± 25)
muon	μ	-1	106
μ-neutrino	ν_μ	0	< 0.190
top quark	t	$\frac{2}{3}$	$(172.5 \pm 2.3) \cdot 10^3$
bottom quark	b	$-\frac{1}{3}$	$(4.20 \pm 0.07) \cdot 10^3$
tau	τ	-1	1777
τ-neutrino	ν_τ	0	< 18.2

Table 1.1: Properties of the fermions (spin-$\frac{1}{2}$ particles) [8]. The electric charge is given in units of the electron charge, the mass in units of MeV/c^2.

Table 1.1 shows all 12 elementary particles, ordered in three families or generations. To each mentioned particle corresponds one antiparticle, with same properties like mass, spin, and lifetime but opposite charge. The SM does not predict the masses of these particles. All ordinary matter is made of first-generation particles (up quark, down quark, electron and electron neutrino). Particles from the other generations decay quickly into the first-generation ones and can only be observed in high-energy interactions as they occur in particle experiments, in the cosmos or in the atmosphere. In particle experiments electrons or protons are accelerated and then brought to collision. The kinetic energy of the colliding particles can be transformed into mass and heavier particles may be produced. In collisions of compound particles, proton-proton collisions for example, the interaction is between the constituents. This means that the kinetic energy of the constituents has to be taken into account. The proton is composed of three valence quarks which are bound together by gluons. These gluons can split up into quark-antiquark pairs, the see quarks. All constituents also called partons share the momentum of the proton. The parton distribution function (PDF) $f_{i,p}(x_i, \mu^2)$ describes the probability density to find a parton i inside the proton p carrying a momentum fraction $x_i = \frac{p_i}{p_p}$. It depends on the factorization scale μ of the considered interaction. Figure 1.1 shows the parton distribution function CTEQ5M1 [9] for protons at a factorization scale $\mu^2 = (175\ \mathrm{GeV})^2$.

The SM describes also three of the four known forces, the electromagnetic force, the weak force, and the strong force. Gravitation is described by the Theory of General Relativity. Since the theories, describing the forces covered by the SM, are gauge theories, the particles transmitting these forces are called gauge bosons. The charge- and massless photon γ is the gauge boson of the electromagnetic force. The gluons g, which transmit the strong force, are also massless but they have two

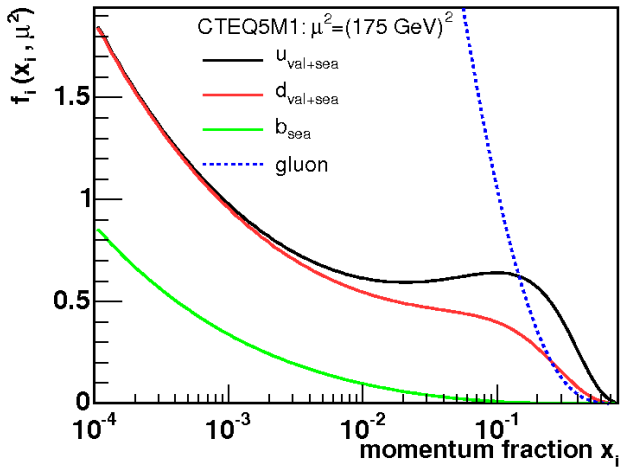


Figure 1.1: The CTEQ5M1 parton distribution function at a factorization scale $\mu^2 = (175\ \text{GeV})^2$.

color charges, while the gauge bosons of the weak force, W and Z, are massive. W bosons carry an electric charge, while Z bosons are neutral. Table 1.2 presents the properties of these bosons.

name	force	symbol	el. charge $[e]$	mass $[\text{MeV}/c^2]$
gluon	strong	g	0	0
photon	electromagnetic	γ	0	0
W boson	weak	$W^{\pm}$	± 1	80.403 ± 0.029
Z boson	weak	Z^0	0	91.188 ± 0.002

Table 1.2: Properties of the gauge bosons (spin-1 particles) [8]. The electric charge is given in units of the electron charge, the mass in units of GeV/c^2.

An important aspect of the weak force is that the eigenstates of the weak interaction are not the same as the mass eigenstates of the quarks. This provides the only way to change the particle flavour. The transformation of mass eigenstates into flavour eigenstates is accomplished by multiplying with the Cabibbo-Kobayashi-Maskawa (CKM) matrix [10, 11].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

The coupling strength of two quarks i and k to a W boson is proportional to the corresponding squared element $|V_{ik}|^2$ of the CKM matrix. The values of individual matrix elements can in principle all be determined from weak decays of the relevant quark, or, in some cases, from deep inelastic neutrino scattering. The 90%

confidence limits on the magnitude of the elements of the complex matrix are [8]:

$$\begin{pmatrix} 0.9739 \text{ to } 0.9751 & 0.221 \text{ to } 0.227 & 0.0029 \text{ to } 0.0045 \\ 0.221 \text{ to } 0.227 & 0.9730 \text{ to } 0.9744 & 0.039 \text{ to } 0.044 \\ 0.0048 \text{ to } 0.014 & 0.037 \text{ to } 0.043 & 0.9990 \text{ to } 0.9992 \end{pmatrix} \tag{1.2}$$

Particles can only interact via a force, if they carry the corresponding charge. Interactions between two charged particles take place by emission and reabsorption of gauge bosons. Feynman diagrams visualize such processes (figure 1.2). Following the Feynman rules, one can translate diagrams for a specific physical process into the corresponding formula. Only fundamental particles appear in Feynman diagrams, fermions are represented by straight lines and bosons by wavy lines.

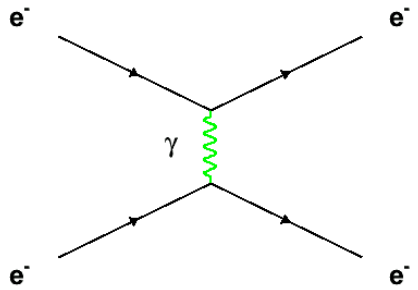


Figure 1.2: Feynman diagram of electron-electron scattering (Møller scattering). On the left-hand side are the two incoming initial state electrons. They interact via a virtual photon (γ) and the scattered final state electrons phase out to the right.

One characteristic feature of the weak interaction are the massive gauge bosons W and Z. Since all theories in the SM are gauge theories the corresponding gauge bosons should be massless. This is mandated by local gauge invariance, which would be violated by explicit mass terms for gauge bosons in the Lagrangian. To explain the mass of the W and Z bosons an additional scalar field [12], the Higgs field, is needed. Particles then get their masses by interacting with that field. The quantum of the Higgs field is the Higgs boson, which is also massive due to its selfinteraction. This is the only particle predicted by the SM but not yet discovered.

1.2 The Top Quark

The top quark, the heaviest fundamental particle known, was discovered in 1995 by the CDF and DØ experiments [1,2] at the Tevatron Collider. Until the commissioning of the Large Hadron Collider (LHC) at CERN the Tevatron is the only collider that provides enough energy ($\sqrt{s} = 1.96$ TeV) to produce top quarks, mostly via quark-antiquark annihilation (figure 1.3 a). While at the LHC, which is a proton-proton collider with a center of mass energy of $\sqrt{s} = 14$ TeV, the dominant pro-

duction process will be the gluon fusion (figure 1.3 b,c) with an expected fraction of about 90%, the quark-antiquark annihilation will contribute about 10%. Most top quarks will be produced as $t\bar{t}$ pairs. The $t\bar{t}$ production cross section at a center-of-mass energy of $\sqrt{s} = 14$ TeV and an assumed top quark mass of $m_t = 175$ GeV/c^2 is estimated to be $\sigma_{t\bar{t}} = 833$ pb [5] at next to leading order.

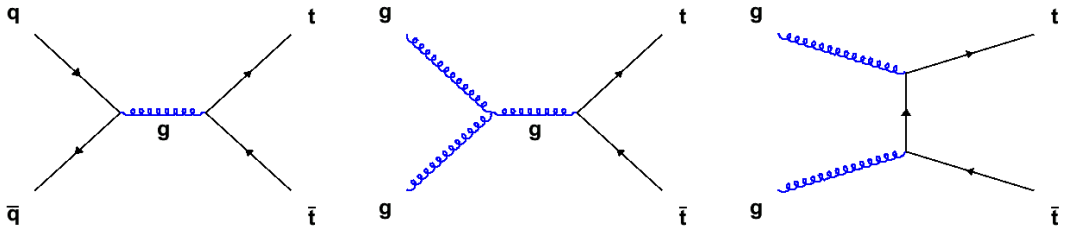


Figure 1.3: Leading-order Feynman diagrams of top quark pair production: (a) quark-antiquark annihilation and (b), (c) gluon fusion.

The top quark lifetime is predicted to be about 10^{-25}s [8] by the SM. Due to its huge mass the top quark decays into a W boson and a b quark, before it can hadronize to colorless objects. This gives us the unique opportunity to study a nearly bare quark by observing its decay particles to which the top quark transmits its properties. The decay into lighter quarks is suppressed by the small matrix elements V_{td} and V_{ts}. While the b quark hadronizes and subsequently decays to many secondary particles, the W boson has two modes to decay via the weak interaction. The first one is the decay into a light quark-antiquark pair and the second one is the decay to a charged lepton and the corresponding antineutrino. Both decay modes are depicted in figure 5.1.

In figure 1.4 all possible decay channels for the combination of both W bosons decays in a $t\bar{t}$ event are summarized:

- Dilepton events: both W bosons decay into a lepton-neutrino pair (highlighted in blue).
- Hadronic events: both W bosons decay into a quark-antiquark pair (highlighted in brown).
- Lepton+jets events: one W boson decays into a lepton-neutrino pair and the other into a quark-antiquark pair (highlighted in green).

Due to the three different quark color charges, the decay into quark-antiquark pairs is enhanced by a factor of three, leading to 81 possible final states. The branching ratios can be obtained by dividing the specified numbers in figure 1.4 by 81.

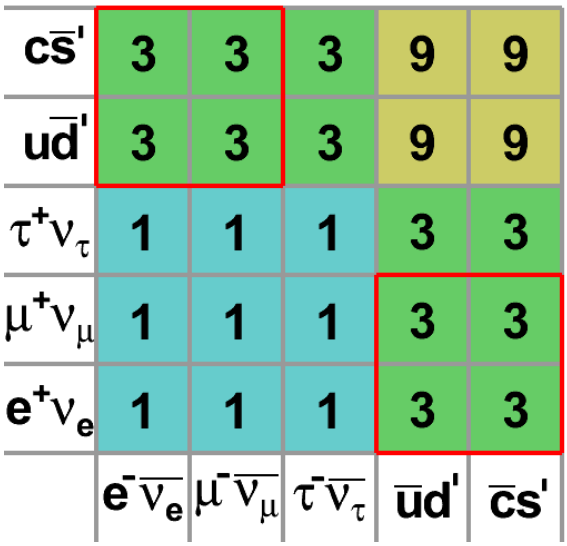


Figure 1.4: Possible decay channels for $t\bar{t}$ events. The branching ratios can be obtained by dividing the specified number by 81, the number off all possible final states.

1.3 W- Helicity in Top Quark Decays

Helicity is defined by the projection of the spin $\vec{S}$ of a particle on the axis given by the direction of motion $\hat{p}$. The helicity operator $\vec{S} \cdot \hat{p}$ thus projects out two physical states, with the spin along or opposite the direction of motion. If the spin is projected parallel on the direction of motion, the particle is of positive helicity, if the projection is antiparallel to the direction of motion, the particle is of negative helicity.

Using the Feynman rules for the weak top decay we can write:

$$\frac{g}{\sqrt{2}}|V_{tb}|\gamma^\mu \frac{1}{2}(1-\gamma^5). \tag{1.3}$$

where $|V_{tb}|$ denotes the CKM matrix element, γ^μ the Dirac matrices, γ^5 the chirality operator, and g the coupling constant of the electroweak interaction. Equation 1.3 implements the vector $(\bar{\psi}\gamma^\mu\psi)$ minus axial vector $(\bar{\psi}\gamma^\mu\gamma^5\psi)$ structure, or V−A structure, of the weak interaction in the SM. The operators

$$P_- = \frac{1}{2}(1-\gamma^5) \tag{1.4}$$

$$P_+ = \frac{1}{2}(1+\gamma^5) \tag{1.5}$$

perform a left-handed and accordingly a right-handed projection. In the extreme relativistic limit, the chirality operator is equal to the helicity operator the two helicity modes are also called right handed or left handed, respectively. For massive bosons with spin 1 there is an a additional helicity mode, which is called longitudinally polarized, when the spin is perpendicular to the momentum of the particle.

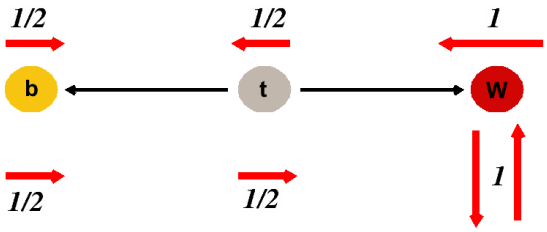


Figure 1.5: Possible W helicity modes in top decays assuming a massless b quark. The spin configuration for left-handed W bosons is illustrated by the arrows above the drawn particles, while the configuration for longitudinal polarized W bosons is shown below.

In the top quark decay we can assume a massless b quark, which is justified due to the large masses of the top quark and the W boson. Due to the V−A structure of the electroweak theory the b has to be left-handed. As a fermion the top quark has just two possibilities to adjust its spin, which limits the W boson, to be left-handed W_L or longitudinally polarized W_0. These two helicity modes are depicted in figure 1.5 by the arrows above and below the particles drawn in the top quark rest frame. Right-handed W_R bosons in top decays are not allowed by the V−A structure. Therefore, the fraction F_+ of right-handed W bosons in top decays is exactly zero in the SM.

$$L = \frac{g}{\sqrt{2}}[W_\mu^- \bar{b}\gamma^\mu(f^L P_- + f^R P_+)t] + \text{C.C.} \tag{1.6}$$

Equation 1.6 denotes a Lagrangian [13] of the top quark decay including possible right-handed couplings. W_μ^- is the effective vector field of the W boson, while t and b denote the fermion fields of the top and bottom quark, respectively. f^L and f^R denote the form factors of the left-handed and right-handed couplings of the top quark and C.C. means the same term but complex conjugated. In the SM the form factor f^L is given by the CKM matrix element V_{tb} and is therefore equal to one, while f^R is equal to zero, leading to the V−A structure of the SM. The SM prediction for the fraction F_0 of longitudinally polarized W bosons in top quark decays is a function of the top quark and W boson masses [13].

$$F_0 = \frac{\Gamma(t \to W_0 b)}{\Gamma(t \to W_L b) + \Gamma(t \to W_0 b) + \Gamma(t \to W_R b)} = \frac{m_t^2}{2m_W^2 + m_t^2} \tag{1.7}$$

Where $\Gamma(t \to W_R b)$ was set to zero. Using $m_W = 80.403$ GeV/c^2 [8] and $m_t = 172.5$ GeV/c^2 [8], one obtains $F_0 = 0.697$. Under the assumption of a massless b quark the fraction of left-handed W bosons is

$$F_- = \frac{2m_W^2}{2m_W^2 + m_t^2} = 0.303 \tag{1.8}$$

An additional right-handed coupling f^R, i.e. a V+A current in top quark decay, would lead to a nonzero fraction of right-handed W bosons at the expense of F_-, while F_0 would remain unchanged. The form factors of the left-handed and right-handed coupling and the helicity fractions of the W boson are connected via:

$$\frac{F_+}{F_-} = \left(\frac{f^R}{f^L}\right)^2 . \tag{1.9}$$

Next to leading order corrections, considering electroweak and finite W width effects as well as QCD-loops and a nonzero b quark mass were performed [14–23] leading only to a small influence on the fraction of longitudinally polarized W bosons F_0 of about one percent and an even ten times smaller influence on F_+.

A significant deviation from the predicted value for F_0 or a nonzero value for the right-handed fraction F_+ could indicate new physics. Left-right symmetric models [24], for example, lead to a significant right-handed fraction of W bosons in top quark decays. Such a right-handed component (V+A coupling) would lead to a smaller left-handed fraction, while the longitudinal fraction F_0 would change insignificantly. Since the decay rate to longitudinal polarized W bosons depends on the Yukawa coupling of the top quarks [25, 26], the measurement of F_0 is sensitive to the mechanism of electroweak symmetry breaking. Alternative models of electroweak symmetry breaking, such as topcolor-assisted technicolor models, can lead to an altered F_0 fraction [27, 28].

We now turn the leptonically decaying W boson. Since the SM considers neutrinos to be massless, there exist only left-handed neutrinos, thus the neutrino from the W decay has to be left-handed. As a consequence of momentum conservation and spin conservation the charged lepton in the W decay has then always to be right-handed. Considering the decay angle θ^*, which is defined as the angle between the momentum of the charged lepton in the W boson rest frame and the W boson momentum in the top quark rest frame, it is possible to measure the W helicity in top quark decays. Figure 1.6 illustrates the extent of θ^* for the three different

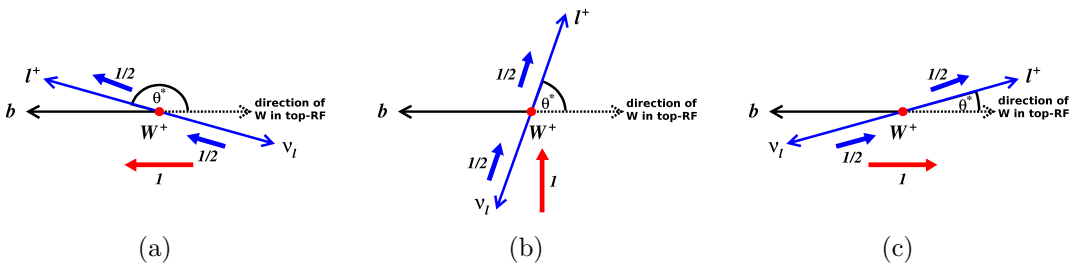


Figure 1.6: Illustration of the decay of a polarized W boson into a charged lepton and a neutrino in the W boson rest frame. In (a) the W boson is left-handed, in (b) the W boson is longitudinally polarized, and in (c) the W boson is right-handed. The dotted black arrow represents the direction of the W boson momentum in the rest frame of the top quark.

W polarizations. Charged leptons from left-handed W bosons are likely to fly in the direction of the b quark, leading to large values for θ^* (a), while charged leptons from right-handed W bosons fly mostly in opposite direction, leading to small values of θ^* (c). For longitudinally polarized W bosons the θ^* distribution has its maximum at 90^0 (b). The cosine of that decay angle θ^* is a very suitable distribution to distinguish between the three different helicity modes. The $\cos\theta^*$ distribution is proportional to $(1-\cos\theta^*)^2$ for left-handed W bosons. For longitudinally polarized W it is proportional to $(1-\cos^2\theta^*)$, and to $(1+\cos\theta^*)^2$ for right-handed W bosons. The general $\cos\theta^*$ distribution is then given by:

$$\frac{dN}{d\cos\theta^*} = F_- \cdot \frac{3}{8}(1-\cos\theta^*)^2 + F_0 \cdot \frac{3}{4}(1-\cos^2\theta^*) + F_+ \cdot \frac{3}{8}(1+\cos\theta^*)^2. \quad (1.10)$$

Figure 1.7 shows the $\cos\theta^*$ distributions for the three helicity modes as well as for the SM expectation which is derived by inserting the SM values for F_0, F_-, and F_+ in equation 1.10.

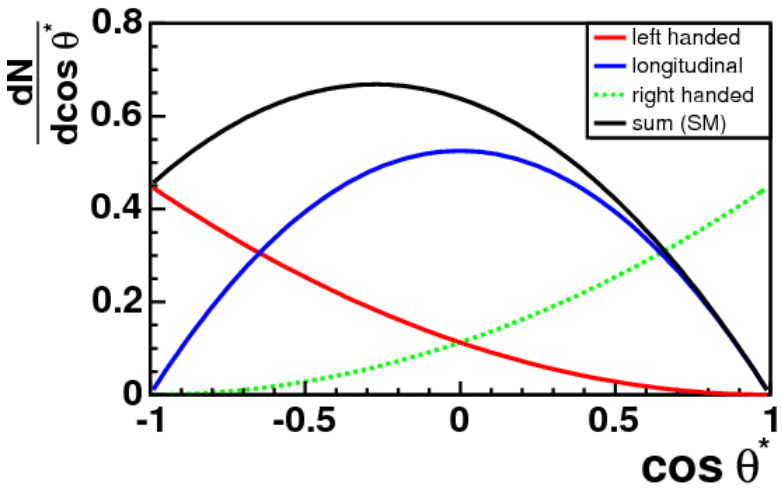


Figure 1.7: Theoretical $\cos\theta^*$ distribution for left-handed (red), right-handed (green dashed) and longitudinally polarized (blue) W bosons. The distribution predicted by the SM is indicated by the black solid line.

Chapter 2

The CMS Experiment at the Large Hadron Collider

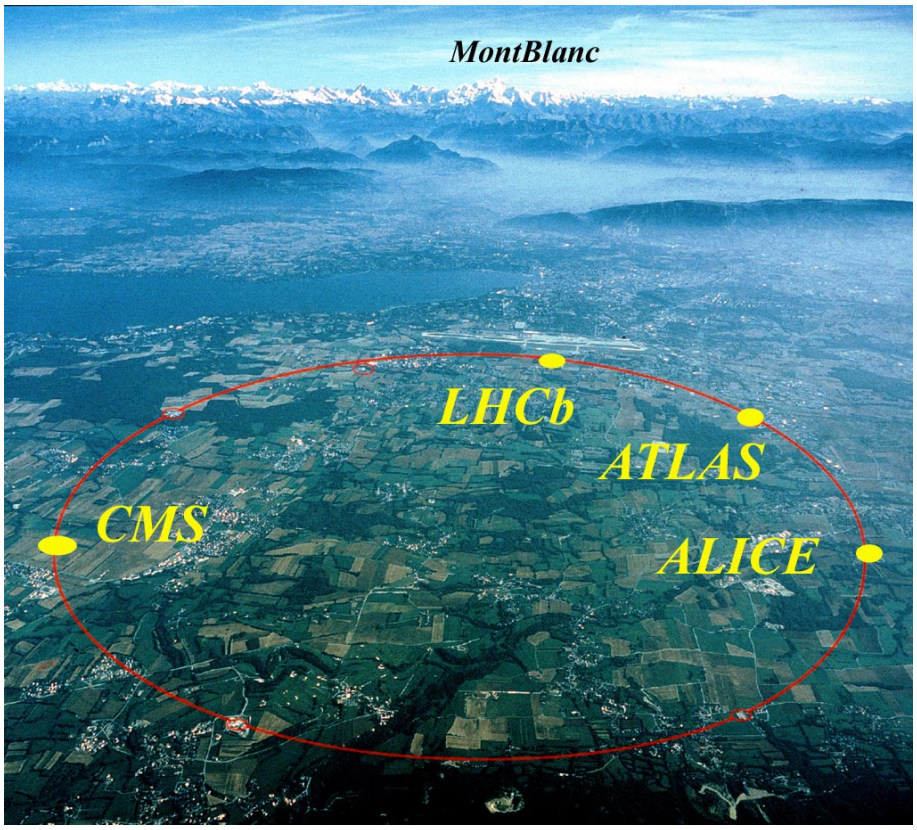


Figure 2.1: View on the CERN arial. The 27 km long LHC tunnel is highlighted in red and the position of the experiment in yellow.

This chapter describes the Large Hadron Collider (LHC) at CERN near Geneva. Figure 2.1 shows a view on the CERN arial. The 27 km long LHC tunnel is highlighted in red and the position of the experiment in yellow. The main goal of the LHC is the discovery of the Higgs Boson and the investigation of its properties. This

particle is the only Standard Model particle not yet observed, it plays the key role in electroweak symmetry breaking and is responsible for the masses of the elementary particles. The CMS experiment is also described in the second part.

2.1 The Large Hadron Collider (LHC)

The LHC is a proton-proton collider installed in the former tunnel of the Large Electron-Positron collider (LEP). The LEP was operated at a center-of-mass energy of up to209 GeV, and was limited by energy loss ΔE caused by synchrotron radiation.

$$\Delta E \approx \frac{1}{R}\left(\frac{E}{m}\right)^4 \tag{2.1}$$

R denotes the radius of the track, on which the particles are circulating, E the particle energy and m the particle mass. There are two possibilities to reduce the energy loss. The first one is to increase the radius, but therefore a completely new infrastructure is needed. The second possibility is to use heavier particles. Protons are 2000 times heavier than electrons, so energy loss due to synchrotron radiation is $1.6 \cdot 10^{13}$ times smaller.

The beam energy at the LHC is 7 TeV, it is limited by the strength of the magnetic dipole field 8.33 T, which is produced by superconducting niobium-titanium magnets and is necessary to keep the protons on their circular track. Such high center-of-mass energy of 14 TeV is required because the particles brought to collisions are protons, which are not elementary particles but compound objects of quarks and gluons. The interaction in collisions is between these elementary particles. Hence, these constituents carry only a fraction of proton momentum, beam energies considerably above the scale up to which one would like to observe new phenomena are needed.

A duoplasmatron is used as proton source. It operates as follows: hydrogen is introduced into a vacuum chamber where it ionized though interactions with electrons emitted by a cathode. The resulting protons are accelerated by an electric field and injected into the Radio Frequency Quadrupole with an energy of 92 keV. The Radio Frequency Quadrupole is is only 1.75 m long, its four electrodes provide a quadrupole radio frequency field to accelerate and focus the protons. This alternating electrical field of 200 MHz forms the continuous proton beam to bunches, since the protons are only accelerated by the adequate half-wave of the alternating field. After this short acceleration distance the bunches reach an energy of 750 keV and enter the 30 m long Linear Accelerator. Here they are accelerated to 50 MeV. The Linear Accelerator is followed by the Proton Synchrotron Booster. It is the first of altogether four ring accelerators. To increase its capacity the Proton Synchrotron Booster is equipped with four beam pipes which lie on top of each other. The Proton Synchrotron Booster accelerates the bunches up to 1.4 GeV. Now the bunches are transfered to the Proton Synchrotron with the nominal bunch spacing of 25 ns.

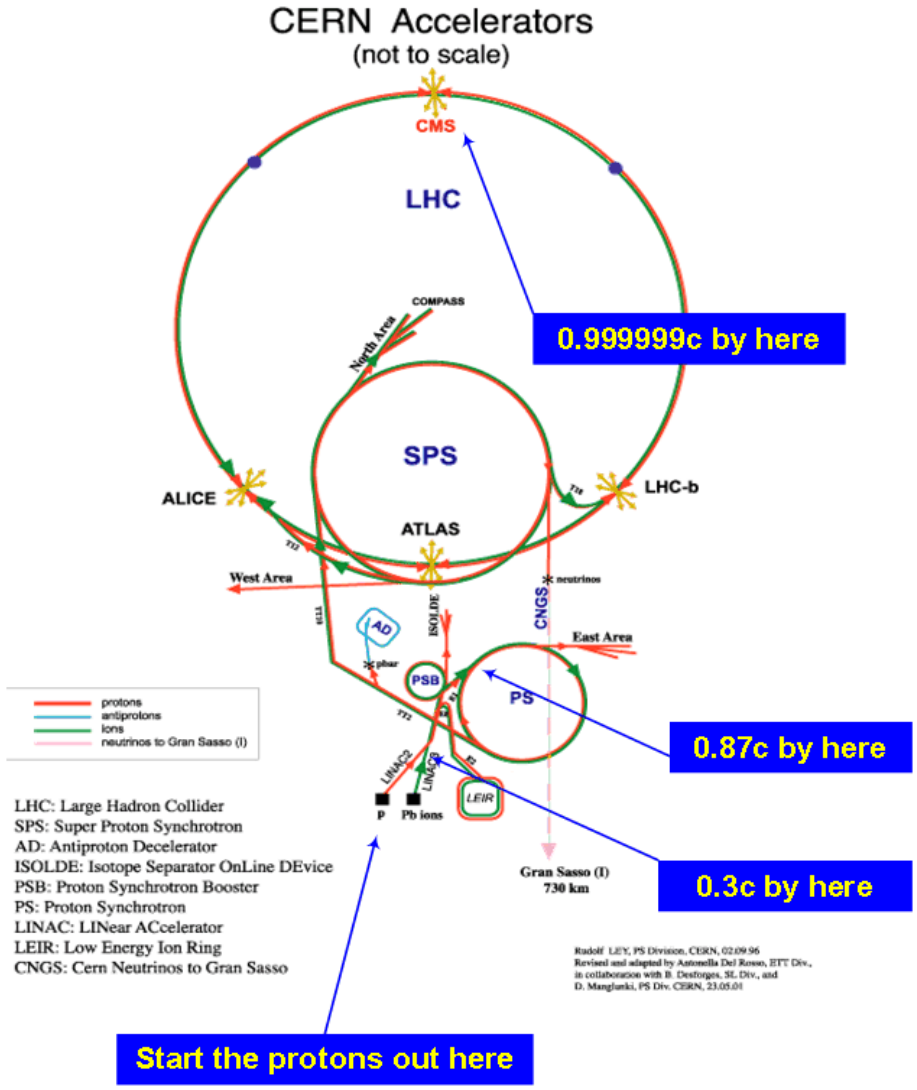


Figure 2.2: Schematic overview of the CERN accelerator complex, with proton velocities.

After reaching an energy of 25 Gev the bunches are injected to the Super Proton Synchrotron. Four charges of the Proton Synchrotron are needed to fill the Super Proton Synchrotron which has a circumference of 6912 m. The bunches are accelerated to 450 GeV and then transfered to the LHC. This time twelve charges of the Super Proton Synchrotron are necessary to fill the LHC. There are two points where the bunches are injected into LHC, to obtain two beams one circulating clockwise and the other counterclockwise. In the two beam pipes of the LHC the bunches reach the nominal energy of 7 TeV. In figure 2.2 the different accelerator stages and proton velocities are shown. Further information can be found in [29].

Since LHC is looking for processes with very small cross sections σ, interesting events will be rare. In order to observe enough events it is important to increase the event rate

$$n = \mathcal{L} \cdot \sigma \tag{2.2}$$

and therefore the luminosity $\mathcal{L}$, i.e. to increase the chance for a proton to collide with another one. The design luminosity of LHC is $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}, so each of the two beams have 2808 bunches with the nominal intensity of $1.15 \cdot 10^{11}$ protons per bunch and a nominal bunch spacing of 25 ns. This corresponds to a collision frequency of 40 MHz, which is a challenge to the five LHC detectors.

The detectors are located at four interaction points, where the bunches are brought to collision. There are three specialized experiments ALICE, LHCb, TOTEM and two complementary multipurpose detectors ATLAS and CMS. TOTEM shares the interaction point with CMS and detects particles flying in very forward direction in order to measure the total pp cross section.

2.2 The Compact Muon Solenoid (CMS)

The CMS experiment is a multipurpose detector implementing the following requirements:

- Efficient muon identification and good muon momentum resolution over a wide geometric coverage.
- The best possible electromagnetic calorimeter providing an accurate diphoton and dielectron invariant mass resolution and efficient π^0 rejection.
- An inner tracking detector with good momentum resolution and reconstruction efficiency of charged particles. Efficient triggering and offline vertex reconstruction which requires pixel detectors close to the interaction point.
- A hadron calorimeter with almost hermetic (4π) coverage to allow an accurate determination of missing transverse energy, and a fine lateral segmentation for a good dijet mass resolution.

The emphasis on muons in the CMS experiment relies on their large penetrating power which allows an efficient and precise identification and reconstruction of these particles even at high luminosities.

The detector layout is driven by the choice of the magnetic field configuration, to allow good muon momentum resolution. So the heart of the detector is a superconducting niobium-titanium solenoid [31], which produces a magnetic field of 4 T. It is 13 m long, with an inner diameter of 5.9 m and stores a total energy of 2.7 GJ. The return field is large enough to saturate 1.5 m of iron, allowing four muon stations to be integrated. The bore of the magnet is large enough to host the tracker, the electromagnetic and the hadronic calorimeters, each of them designed to measure different types of particles. They are positioned in concentric layers around the collision point. The different components are depicted in Figure 2.3 and will

be described in detail in the following sections. Together with the endcaps and the muon chambers the CMS detector is nearly 22 m long with a width of 14.4 m which results in a total weight of 12500 tons. Reference [30] gives a more detailed overview of the CMS experiment and its components.

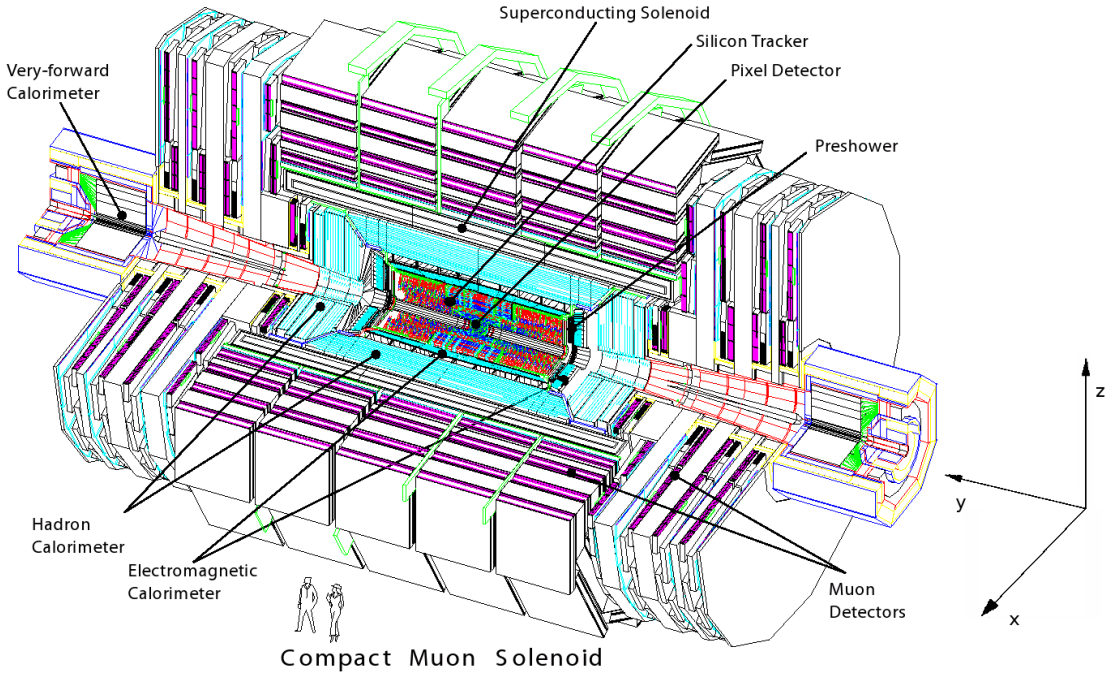


Figure 2.3: A schematic view of the CMS detector [30].

2.2.1 Coordinate conventions

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the detector, the y-axis pointing vertically upward, the x-axis pointing radially inward toward the center of the LHC, and the z-axis along the beam direction. The azimuthal angle ϕ is measured from the x-axis in the $x-y$ plane. The polar angle θ is measured from the z-axis. Instead of θ it is often more handy to use the rapidity y defined as

$$y = \frac{1}{2} ln \left(\frac{E + p_z}{E - P_z} \right) \tag{2.3}$$

which is a Lorentz invariant quantity. But in the majority of cases the pseudo-rapidity defined as

$$\eta = -ln \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{2.4}$$

is used. For massless particles the rapidity y is equal to the pseudorapidity η.

Thus, the angular distance between two point objects, as observed from the origin of the CMS detector, is expressed as the size

$$R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.5}$$

The momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components. The transverse momentum p_T is the projection of the momentum p onto the $x-y$ plane

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{2.6}$$

The transverse energy E_t is the fraction of energy perpendicular to the z axis.

$$E_T = \sqrt{E^2 \cdot \frac{p_T^2}{p_x^2 + p_y^2 + p_z^2}} \tag{2.7}$$

2.2.2 Tracking system

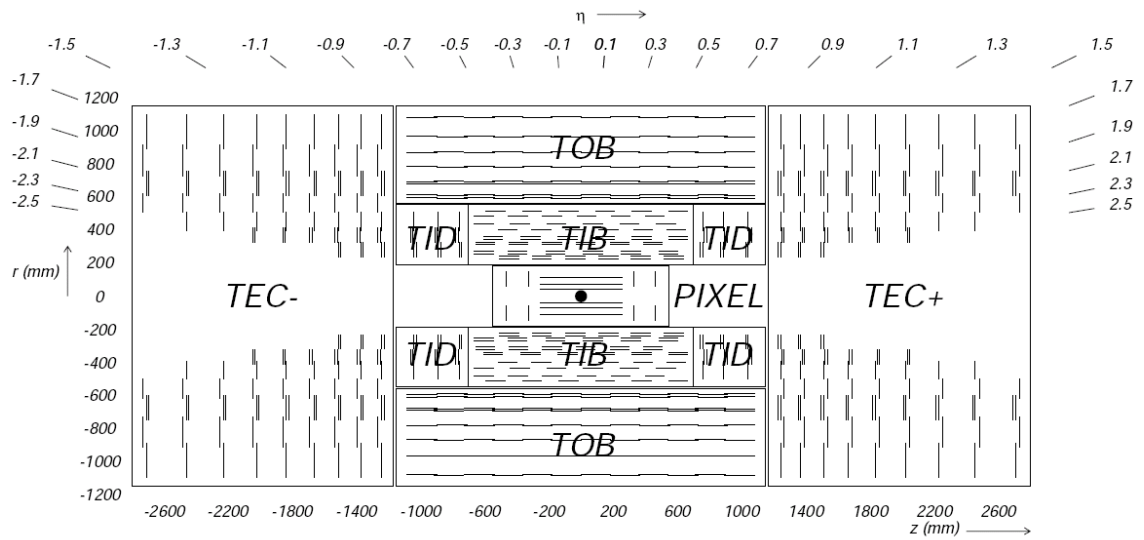


Figure 2.4: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits [36].

The tracker (figure 2.4) is the innermost part of the detector, situated around the interaction point. It is used to reconstruct particle tracks and to measure their momentum and vertex of origin with high resolution. This allows to define isolated objects coming from gauge boson decays, to tag particles coming from secondary vertices (b-tagging) and to distinguish particles from pile-up events. Detailed information can be found in [35].

The tracking system consists of a pixel tracker which is surrounded by a strip tracker. The pixel tracker comprises 66 million silicon pixels which cover a surface

of about 1 m^2. Together with the 9.6 million silicon strips from the strip tracker an area of 200 m^2 is covered. With an outer radius of 110 cm and a length of 540 cm it is the biggest silicon detector of the world. The endcaps provide an pseudorapidity coverage up to $|\eta| = 2.4$. One of the two tracker endcaps was build in Germany. The Institut für Experimentelle Kernphysik (IEKP) participated in the assembly and the quality control of the endcaps modules.

Pixel tracker

The silicon pixel detector consists of three barrel layers and two layers in the endcap disks on each side. The barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, and have a length of 53 cm. The endcap discs, extending from 6 to 15 cm, are placed on $|z| = 34.5$ cm and 46.5 cm. In order to achieve an optimal vertex position resolution, a design with an almost square pixel shape of 100×150 μm^2 has been adopted. The spatial resolution is measured to be about 10 μm for the $r-\phi$ measurement and about 20 μm for the z measurement.

Strip tracker

In the barrel part, the silicon microstrip detectors can be divided into two parts:

- The Tracker Inner Barrel (TIB), with four shorter layers of microstrips, covers up to $|z| < 65$ cm with a single-point resolution of between 23-34 μm in the $r-\phi$ direction and 230 μm in z.

- The Tracker Outer barrel (TOB) comprises the remaining six layers of microstrips and covers up to $|z| < 110$ cm. The resolution varies from 35 μm to 52 μm in the $r-\phi$ direction and 530 μm in z.

The endcaps are also divided into two regions, the Tracker Inner Discs (TID) and the Tracker End Cap (TEC). The TID consists of three small discs to fill up the gap between the TIB and the TEC, which comprises nine disks that extend into the region 120 cm $< |z| <$ 280 cm. Table 2.1 summarizes energy and transverse momentum resolutions of the tracker.

p_T [Gev/c]	η	σ_E [GeV]	σ_E/E [%]	σ_ϕ [mrad]	σ_{p_T}/P_T [%]
10	0.17-0.26	0.68	6.7	4.24	5
10	1.26-1.35	0.78	4.0	3.39	5
40	0.17-0.26	1.04	2.5	3.22	20
40	1.26-1.35	1.59	2.0	2.96	20

Table 2.1: Energy and transverse momentum resolutions of the tracker system [35].

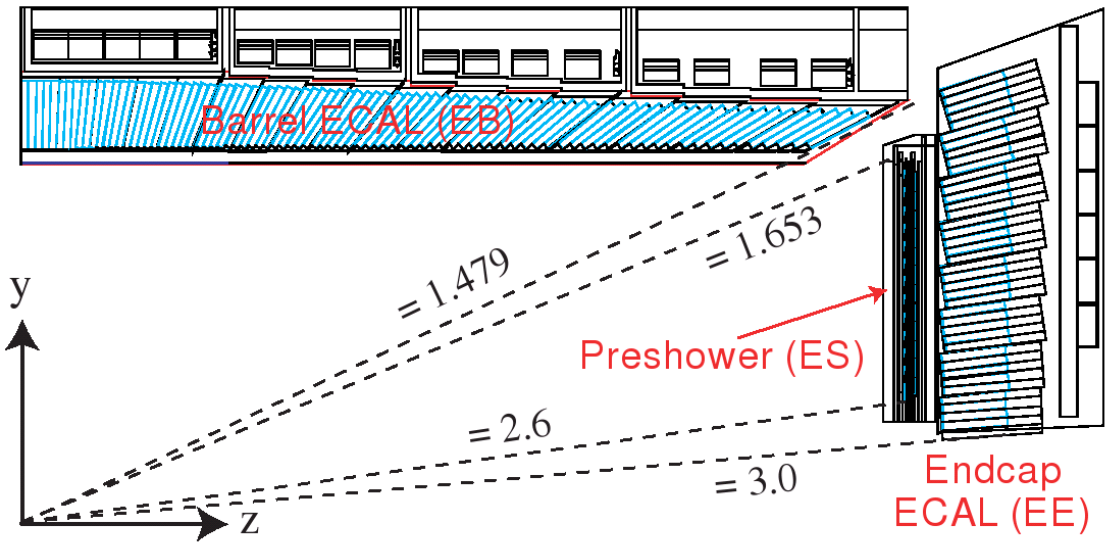


Figure 2.5: Transverse section through the ECAL, showing the geometrical configuration [30].

2.2.3 Electromagnetic calorimeter (ECAL)

The electromagnetic calorimeter is placed around the silicon tracker to measure the energy of electrons, photons, and , in conjunction with the hadron calorimeter, jets with high precision. A particle passing through the calorimeters interacts with the calorimeter material, producing particle showers. These showers lead to ionization in the calorimeter, which can be measured and used to determine the energy and direction of the particle. An overview of the ECAL [34] is given in figure 2.5.

CMS has chosen lead tungstate ($PbWO_4$) scintillating crystals for its ECAL. These crystals have a short radiation length ($\chi_0 = 0.89$ cm) and a small Moliere radius (2.2 cm), are fast (80% of the light is emitted within 25 ns) and radiation hard (up to 10 Mrad). However, the relatively low light yield requires use of photodetectors with intrinsic gain.

In the barrel region the crystals are grouped into 36 identical supermodules covering a pseudorapidity interval of $0 < |\eta| < 1.479$. The crystals have a front face cross-section of about 22×22 mm^2 and a length of 230 mm. The crystals in each of the two endcaps have a cross-section of 28.6×28.6 mm^2, a length of 220 mm and are grouped into units of 5×5. In the endcaps the ECAL covers an η region of $1.479 < |\eta| < 3.0$. The energy resolution of the ECAL is shown in table 2.2.

Contribution	Barrel ($\eta = 0$)	Endcap ($\eta = 2$)
stochastic term	$2.7\%/\sqrt{E}$	$5.7\%/\sqrt{E}$
constant term	0.55%	0.55%
noise (low luminosity)	155 MeV	770 MeV ($E_T = 205$ MeV)
noise (high luminosity)	210 MeV	915 MeV ($E_T = 245$ MeV)

Table 2.2: Energy resolution of the ECAL [34].

2.2.4 Hadronic calorimeter (HCAL)

The ECAL is completely surrounded by the hadronic calorimeter [32], which is used to measure the energy and direction of hadronic particle jets and $\not{E}_T$. The HCAL (figure 2.6) consists of plastic scintillator tiles. Brass has been chosen as absorber material as it has a reasonably short interaction length and is non-magnetic. The hadron barrel (HB) is placed inside the coil and consists of 2304 towers with a segmentation $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, covering the pseudorapidity region $|\eta| < 1.4$. To reduce the tails in the energy resolution function and to improve the missing transverse energy resolution, an additional layer of scintillators, referred to as hadron outer detector (HO), is lined outside of the coil ($\eta < 1.26$). The hadron endcap (HE)

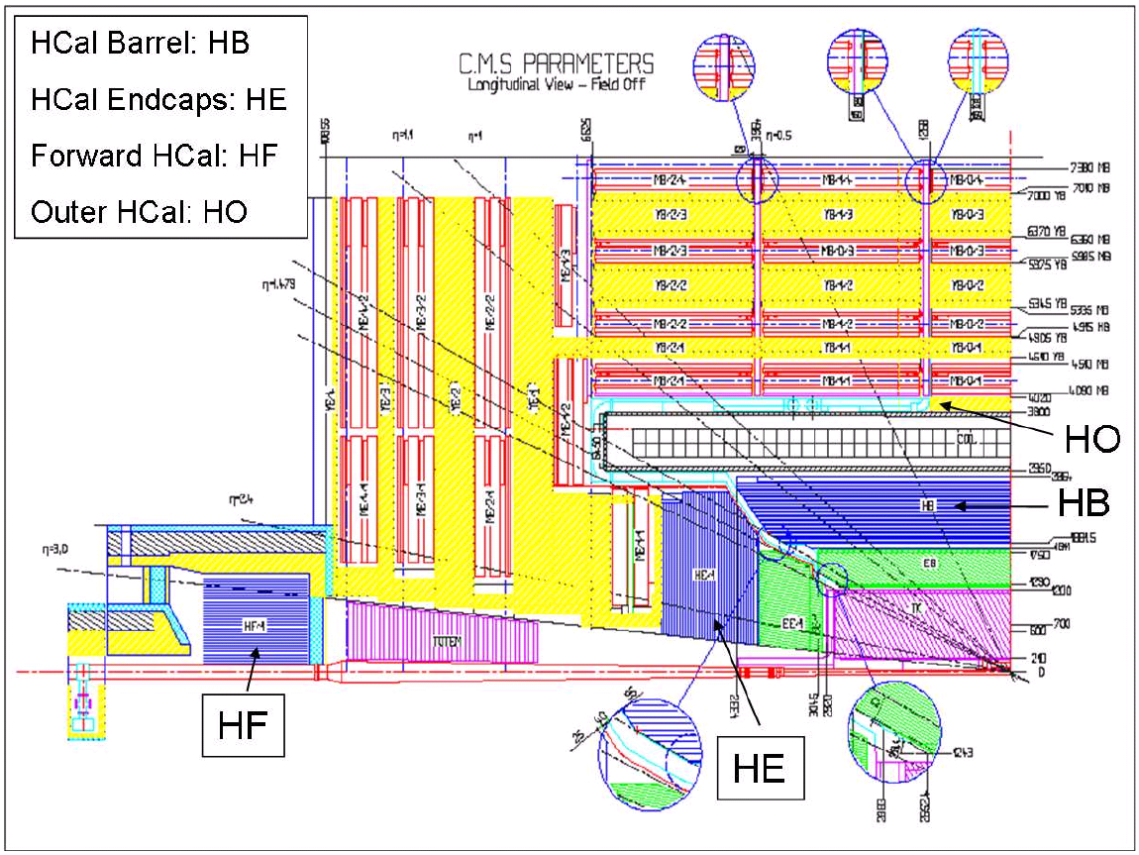


Figure 2.6: Transverse section through the HCAL, showing its geometrical configuration.

detector covers a pseudorapidity interval of $1.3 < |\eta| < 3.0$, which is extended up to $\eta = 5$ by two forward calorimeters (HF) in steel/quartz fiber technology. The energy resolution of the CMS Hadron calorimeters is expected to have a constant term of about 5%.

2.2.5 Muon system

As muons provide the cleanest experimentally measurable signatures, the ability to identify, to trigger, and to reconstruct them with high precision is a central concept of the CMS experiment. Centrally produced muons are measured three times: in the inner tracker, after the coil, and in the return flux.

Three types of gaseous detectors are used to identify and measure muons. Drift tube (DT) chambers are used in the barrel region ($|\eta| < 1.2$), while in the two endcaps cathode strip chambers (CSC) are deployed to cover up to $|\eta| < 2.4$. In addition to this, resistive plate chambers (RPC) are used in both the barrel and the endcap regions [33] (figure 2.7). RPCs provide a fast response with good time resolution but a lower spacial resolution. The DTs or CSCs and the RPCs operate within the first level trigger system, providing two independent and complementary

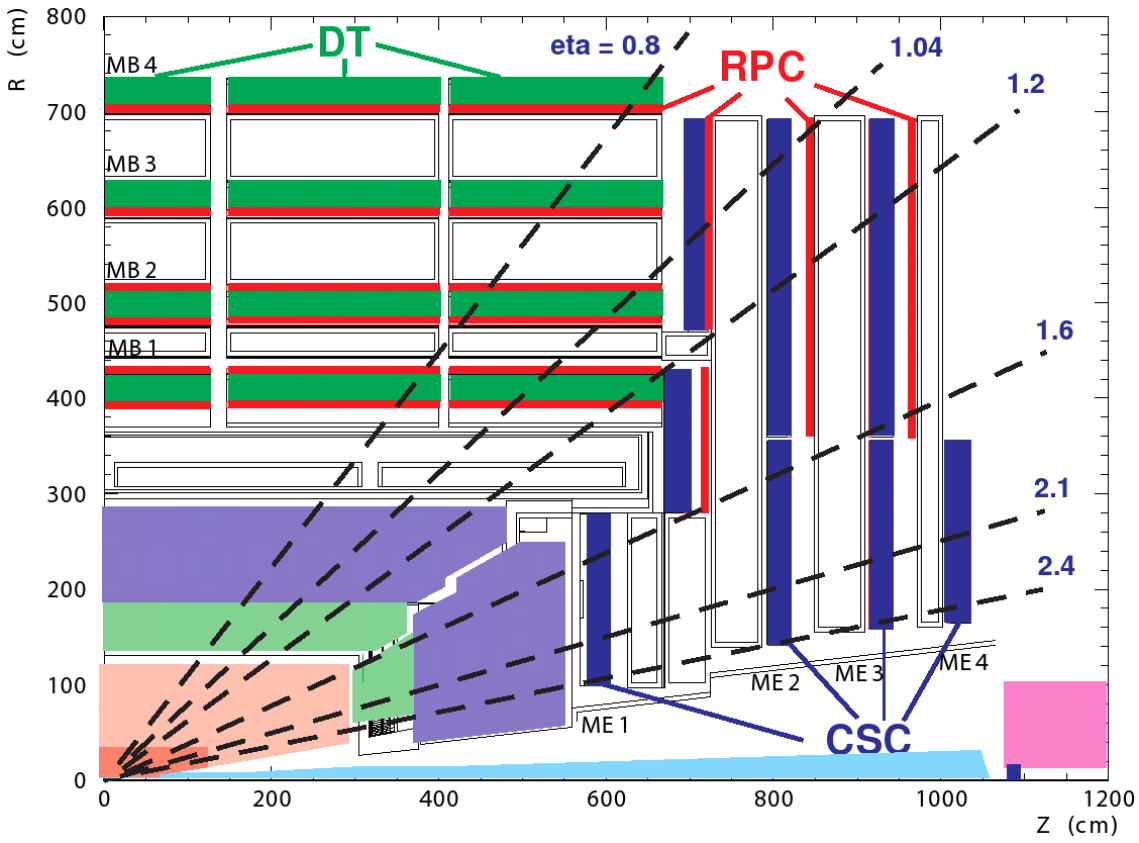


Figure 2.7: Layout of one quarter of the CMS muon system for initial low luminosity running. The RPC system is limited to $|\eta| < 1.6$ [30].

sources of information.

The Barrel Detector (DT) consists of 250 chambers organized in four layers. The maximum drift length is 2.0 cm and the single-point resolution is about 200 μm. Each CSC is trapezoidal in shape and consists of six gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. The spatial resolution provided by each chamber from the strips is typically about 200 μm. The RPCs are covering the region up to $|\eta| < 1.6$ and will be extended to $|\eta| < 2.1$ later. Table 2.3 summarizes the transverse momentum resolution of the muon system.

p_T [GeV/c]	Barrel region $\Delta p_T/p_T$ [%]	Endcap region $\Delta p_T/p_T$ [%]
1000	18.6	37.9
500	15.2	-
300	12.3	-
100	9.3	19.4
10	8.9	15.2

Table 2.3: Transverse momentum resolution of the muon system [33].

2.2.6 Trigger and data acquisition system

The LHC bunch crossing rate of 40 MHz leads to 10^9 interactions per second at design luminosity. But only 10^2 events per second can be written to archival media. So the trigger system has to filter the most interesting events. The CMS trigger and data acquisition system consists of four parts: the detector electronics, the Level-1 trigger processors (calorimeter and muon system), the readout network, and an online event filter system (processor farm) that executes the software for the High-Level Triggers (HLT).

The Level-1 decision is based on the presence of objects such as photons, electrons, muons, and jets above defined E_T or p_T thresholds, and global sums of E_T and $\not\!\!E_T$. Much of the logic in the trigger system is contained in custom Application Specific Integrated Circuits and Field Programmable Gate Arrays. The Level-1 trigger calculations are computed in less than 1 μs. The transit time for signals from the front-end electronics to the Level-1 trigger and back is 3.2 μs. During this time, the detector data is held in buffers, and is then, according to the Level-1 trigger decision, discarded or placed in dual-port memories for access by the data acquisition system. This reduces the event rate to 16 kHz, at startup of the LHC. The design value of 100 kHz is set by the average time to transfer full detector information through the readout system.

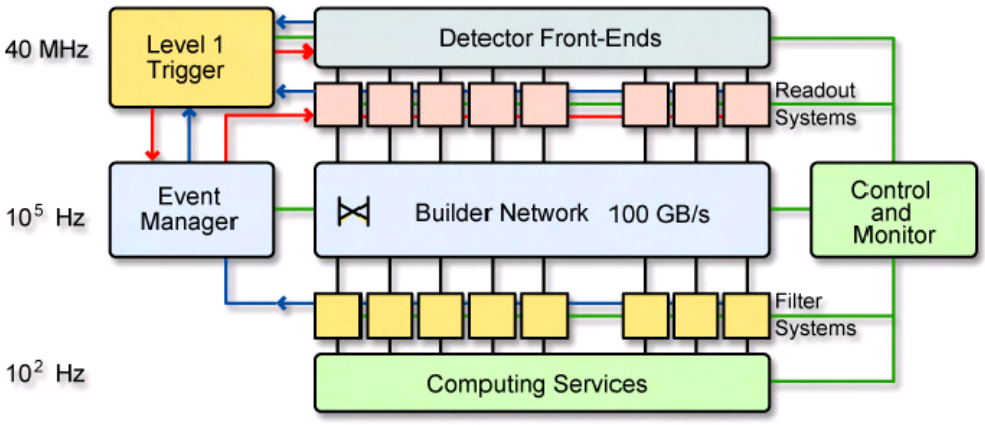


Figure 2.8: General architecture of the CMS data acquisition system [30].

The events, with a size of about 1.5 MB, are then transfered to the processor farm and the HLT software is executed to reduce the Level-1 output rate of 100 kHz to 100 Hz for mass storage. The use of a processor farm allows maximal benefit to be taken from the evolution of computing technology, as well as the sophistication of the selection algorithms. [37] and [38] provide more detailed information. An overview of the general architecture of the CMS data acquisition system is shown in figure 2.8.

Chapter 3

Comparison of Monte Carlo event generators used for signal modeling

The first section of this chapter gives an overview of Monte Carlo event generators, and describes briefly the Monte Carlo method. The second section presents the comparison of two Monte Carlo event generators. The comparison is performed using $t\bar{t}$ events at center-of-mass energies of $\sqrt{s} = 1.96$ TeV and $\sqrt{s} = 14$ TeV reached by Tevatron and LHC, respectively.

3.1 Monte Carlo Event Generators

Monte Carlo (MC) event generators simulate events in high energy physics and therefore help to design the detectors and to define the experimental strategies. In order to generate a MC event, $pp \to t\bar{t}$ for example, it is necessary to know the parton distribution function, the cross section of this process, and the angular distributions of the resulting top quarks. These distributions define the phase space, that contains all degrees of freedom of the considered process. To cover the phase space uniformly distributed random numbers are thrown and event candidates are obtained. Using the cross section it is possible to calculate the probability for the event candidates to occur. Therefore, MC event generators produce events with the frequency predicted by the used theory and the individual events represent what might be observed in an experiment.

Figure 3.1 illustrates the general structure of the showering and hadronization process, which are simulated by the MC event generators. The example shows a proton-proton collision. We now describe the labeled parts:

Parton Distributions

The parton distribution functions $f_{i,p}(x_i, \mu^2)$ describe the probability density to find a parton i inside a proton p carrying a momentum fraction $x_i = \frac{p_i}{p_p}$ (see section 1.1).

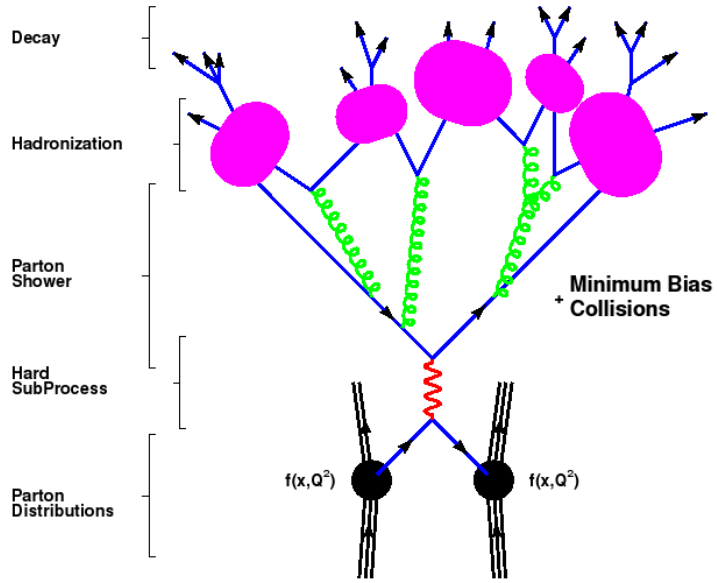


Figure 3.1: Basic structure of the showering and hadronization process simulated by MC event generators [39]. The labeled parts are described in the text. [39]

Hard subprocess

In this example, a valence quark is resolved in the proton shown on the left, while an anti-quark is resolved from the proton on the right. The quark and antiquark annihilate into a resonance, Z^0 for example, denoted by a wavy line. This can be computed exactly to lowest order in perturbation theory. The hard momentum transfer scale Q together with the color flow of the subprocess set the boundary conditions for the initial- and final-state parton-showers, if there are any.

Heavy object decays

Massive produced objects such as top quarks or electroweak gauge bosons can decay on timescales that may be shorter than or comparable to that of the QCD parton showers. Depending on their nature and the decay mode, they may also initiate parton showers before and/or after decaying. In the example the resonance decays into a quark antiquark pair.

Initial- and final-state parton-showers

A parton constituent of an incident beam hadron with low spacelike virtuality can radiate timelike partons. In the process it decreases its energy to a fraction x of that of the beam and increases its spacelike virtuality, which is bounded in absolute value by the scale Q of the hard subprocess. Such initial-state radiation (ISR) is shown in figure 3.2. This figure also shows the final-state radiation (FSR). The

outgoing partons, the quark antiquark pair in our example, can radiate a gluon, with large timelike virtuality, which then generate a shower of partons with lower virtuality. The amount of emission depends on the upper limit on the virtuality of the initiating parton, which is again controlled by the momentum transfer scale Q of the hard subprocess. The coherence of soft gluon emission from different parton showers is controlled by the color flow of the subprocess.

Hadronization process

In order to construct a realistic simulation one needs to transform the partons into hadrons. Outgoing products of parton showering give rise to hadronic jets. This hadronization process takes place at a low momentum transfer scale, for which the strong coupling is large and perturbation theory is not applicable. In the absence of a firm theoretical understanding of nonperturbative processes, it must be described by a phenomenological model, which can be different for the various generators.

Hadron decays

After hadronization, many short-lived hadrons are present. They are decayed into lighter hadrons in this simulation step.

Underlying event

In the underlying event the colored beam remnants, which are left behind when the parton participating in the hard subprocess is pulled out, are included in the same hadronization system, since they are color connected to the hard subprocess. Multiple parton-parton interactions, wherein more than one pair of partons from the beam protons interact, are also accounted for.

Pile-up

In a final step, pile-up from other proton-proton collisions in the same bunch crossing are added to the event. In figure 3.1 pile-up is referred to as minimum bias collisions.

Different types of MC event generators have been developed:

- Matrix Element Event Generators
  (like AlpGEN [47], MadEvent/MadGraph [44, 45], TopReX [48])
  These event generators describe a specific final state to lowest order in pertubation theory by computing all relevant tree-level matrix elements. Next to leading order (NLO) matrix elements are not included to avoid complications involving the regularization of these matrix elements. Hadronization processes are also not included, thus the final states consist of bare quarks and gluons.

- General Purpose Showering and Hadronization Event Generators
  (like Pythia [40], Herwig [42], Sherpa [49])
  In contrast to Matrix Element Event Generators these programs include higher order effects by evolving the event using the parton shower, which allow partons to split into pairs of other partons. The resulting partons are then hadronized and grouped together to color-singlet hadrons. Subsequently resonances are decayed. The underlying event, interactions from other partons, and the pile up, collisions between other hadrons, are also generated.

A more detailed list of different MC generators and their properties is given in [39] as well as a short introduction to MC generator techniques. Here we describe only the main difference between Herwig and Pythia, which is the hadronization model.

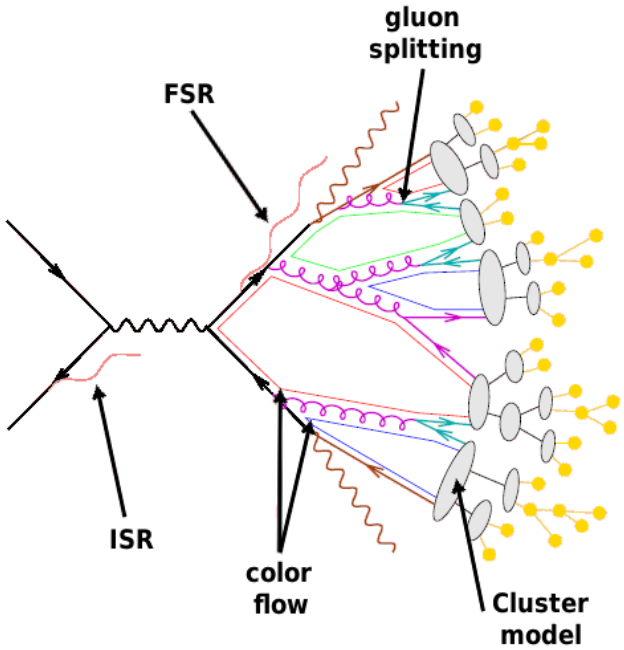


Figure 3.2: Showering model of the MC event generator Herwig: initial-state radiation (ISR) and final-state radiation (FSR) are illustrated as well as gluon splitting and the color flow. The decay of clusters is also shown.

Herwig uses the cluster hadronization model [42] which is based on the color preconfinement property of the parton shower. After this perturbative showering process all outgoing gluons are split non-perturbatively into light quark-antiquark pairs. Quarks are then combined with their nearest neighbours to form color singlet states. If the resulting cluster is too light to decay it represents the lightest hadron of its flavor else it is decayed into lighter hadrons. Figure 3.2 illustrates the gluon splitting, the cluster hadronization model, and the color flow.

Pythia implements the string fragmentation or Lund model [40]. The string model 3.3 is based on the assumption of the linear confinement of Quantum Chromo Dynamics. If two partons forming a color singlet, $q\bar{q}$ for example, move apart from their common production vertex a color flux tube is being stretched between them. The potential energy stored in this tube or string increases linearly with the distance between the partons and the string may break producing a new quark antiquark pair $q'\bar{q}'$. This forms two new color singlets $q\bar{q}'$ and $\bar{q}q'$. The process continues until only on-mass-shell hadrons remain.

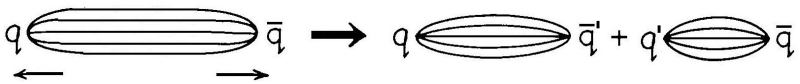


Figure 3.3: String fragmentation model: a color flux tube is stretched between two quarks q and $\bar{q}$ which are moving apart. The energy in the string increases and produces a new quark antiquark pair $q'\bar{q}'$.

For studies made in this thesis Pythia and Herwig++ were used to determine whether one has to expect differences between $t\bar{t}$ events at a center-of-mass energy of 1.96 TeV and 14 TeV, reached by Tevatron and LHC respectively (next section) and to study differences in the two event generators. To study the W helicity in top quark decays we used MC samples generated by the CMS collaboration. These samples were generated with TopRex and AlpGen. Further MadEvent/MadGraph was used to generate $t\bar{t}$ events with different fractions of left handed, right handed, and longitudinally polarized W bosons. As already mentioned, the final states generated by this MC generator consists of bare quarks and gluons, which were passed to Pythia for showering. The later samples were used to prepare the analysis code (with the MadEvent/MadGraph sample containing W boson helicity fractions predicted by the Standard Model) and to check the obtained method (by applying it to samples containing other fractions).

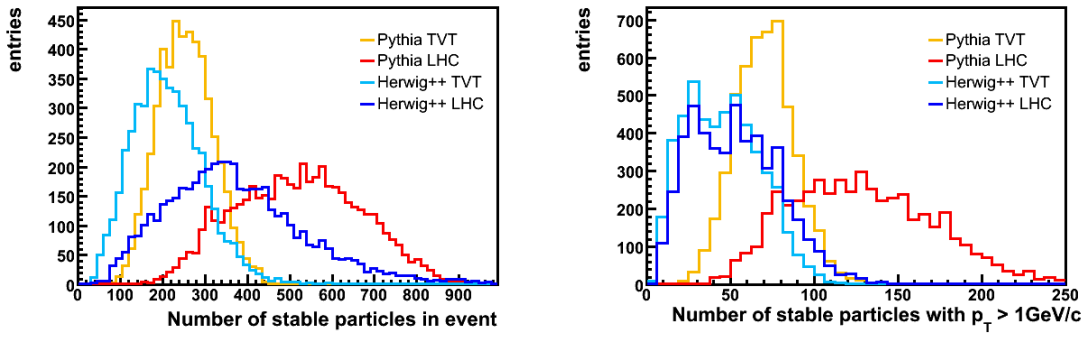
3.2 Pythia in comparison to Herwig++

In this section the prediction of the MC event generators Pythia and Herwig++ for the $t\bar{t}$ production at the LHC takes center stage. The LHC will have a seven times higher center-of-mass energy than the Tevatron. The influence of this difference on $t\bar{t}$ events is probed. In addition, a comparison between the two different event generators is done.

Herwig++ is a completely new event generator, written in C++. It is built on the experience gained with the well-known event generator Herwig. We used Herwig++ 2.0.0 and compared it with Pythia 6.319[1]. For this purpose 5000 $t\bar{t}$ events

[1]Now a newer Herwig++ version 2.0.3 is available. Pythia is also ported to C++ in version 8.1.

at 1.96 TeV ($p\bar{p}$) and 14 TeV (pp) were generated using each program.



(a) (b)

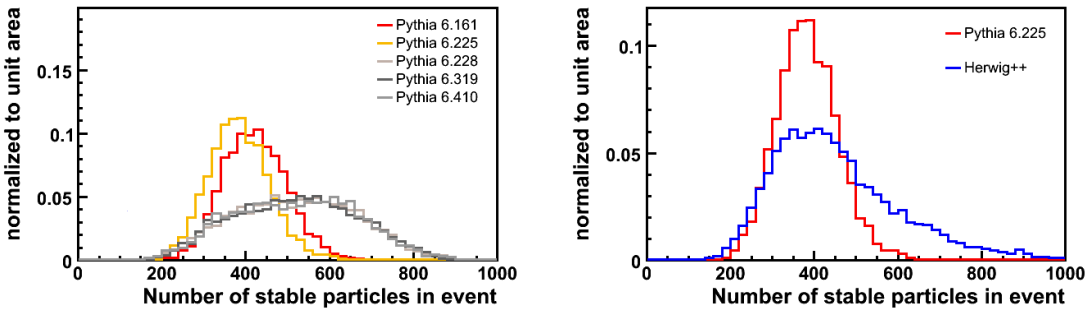
Figure 3.4: Number of stable particles in $t\bar{t}$ events without cuts (a) and with a $p_T > 1$ GeV/c cut (b). TVT denotes a center-of-mass energy of 1.96 TeV reached by the Tevatron collider, and LHC labels a center-of-mass energy of 14 TeV.

Before studying the $t\bar{t}$ pair and its decay products we take a look at some general distributions. Both MC generators predict more stable particles per event at the LHC than at the Tevatron (figure 3.4). This was expected due to the higher center-of-mass energy at the LHC. Figure 3.4 a shows the first discrepancy, at both center-of-mass energies Pythia produces more stable particles than Herwig++. The difference is even bigger when a $p_T > 1$ GeV cut is applied to select particles with higher transverse momentum (figure 3.4 b). Further studies showed that the parameterization used in Pythia produces more mesons, like pions and kaons, baryons (protons and neutrons), and photons than Herwig++.

This difference is caused by ISR and FSR settings in Pythia, the so called Tune A, which was made by Rick Field in order to describe data taken at the Tevatron more precisely. Figure 3.5 a shows the distribution of stable particles for different Pythia versions, version 6.161 and 6.225 are without Tune A. In figure 3.5 b the comparison between Pythia 6.225 and Herwig++ is shown. Here, both event generators produce on average about 400 stable particles per event. The distribution obtained with Herwig++ has a tail to more stable particles, while the Pythia distribution is more balanced. Not until data is taken at the LHC one will know which generator makes the better prediction. Probably a new tune will be needed.

Now we turn to the $t\bar{t}$ pair and its properties. In figure 3.6 one can see the invariant mass which is defined as

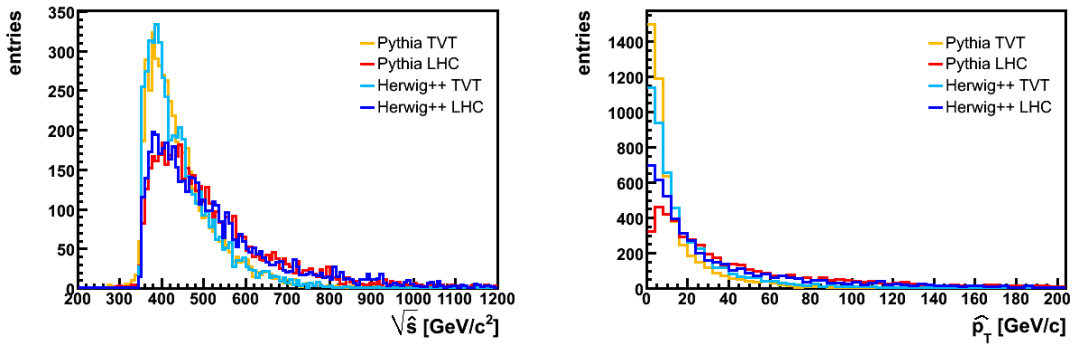
$$m_{t\bar{t}} = \sqrt{\hat{s}} = \sqrt{\frac{E^2}{c^4} - \frac{p^2}{c^2}} \tag{3.1}$$



(a) (b)

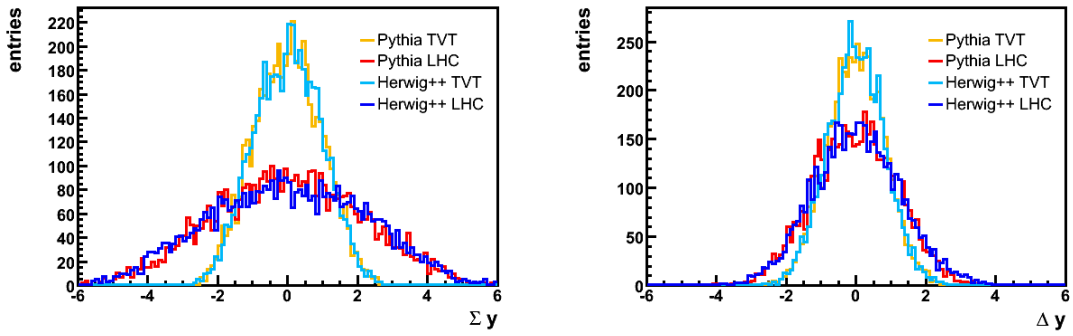
Figure 3.5: Number of stable particles in $t\bar{t}$ events for pp collisions at 14 TeV (LHC) using different Pythia versions. Pythia 6.161 and 6.225 are without Tune A, while the tune is applied in the newer version (a). Pythia 6.225 and compared to Herwig++ (b).

and the transverse momentum $\hat{p}_T$ of the $t\bar{t}$ pair. The distribution of the invariant mass of the $t\bar{t}$ system starts nearly at the same point for both center-of-mass energies. At the LHC as well as at the Tevatron the $t\bar{t}$ pair is produced at its threshold of about 350 GeV/c^2, but the distributions for the LHC have larger tails. While Pythia and Herwig++ nearly agree in the prediction of the invariant mass, the prediction for the transverse momentum is different. The $\hat{p}_T$ distribution obtained from Herwig++ is softer than the distribution obtained from Pythia in the case of the LHC but harder in the case of the Tevatron.



(a) (b)

Figure 3.6: (a) shows the invariant mass of the $t\bar{t}$ system and (b) its transverse momentum. TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.



(a) (b)

Figure 3.7: Sum (a) and difference (b) of the rapidity y of the $t\bar{t}$ system.
TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

Figure 3.7 illustrates the agreement of Pythia and Herwig++ for the $t\bar{t}$ rapidity y defined in equation 2.3. The Σy distribution is the sum of the rapidity of the top and antitop quark, it indicates the larger boost of $t\bar{t}$ pairs at the LHC, the difference of the rapidities Δy is also broader. In figure 3.8 the rapidity of the individual top quarks is shown.

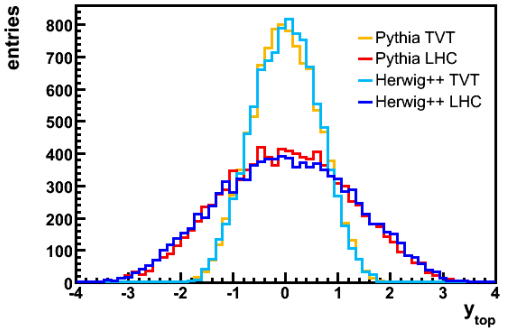
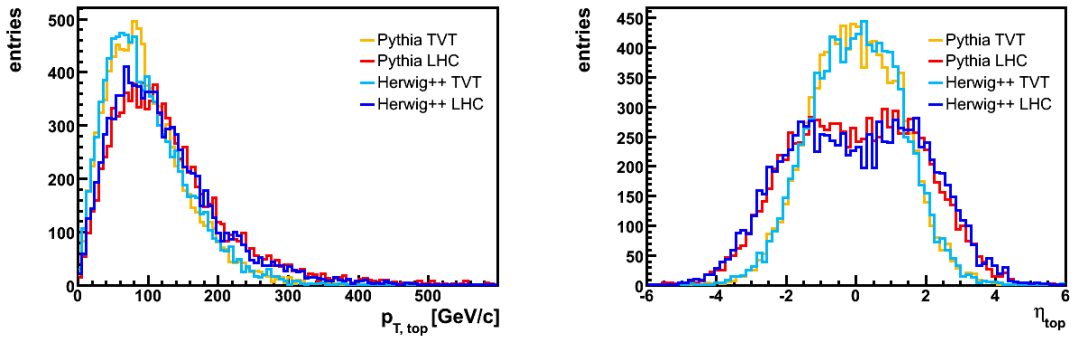


Figure 3.8: Rapidity of individual top quarks TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

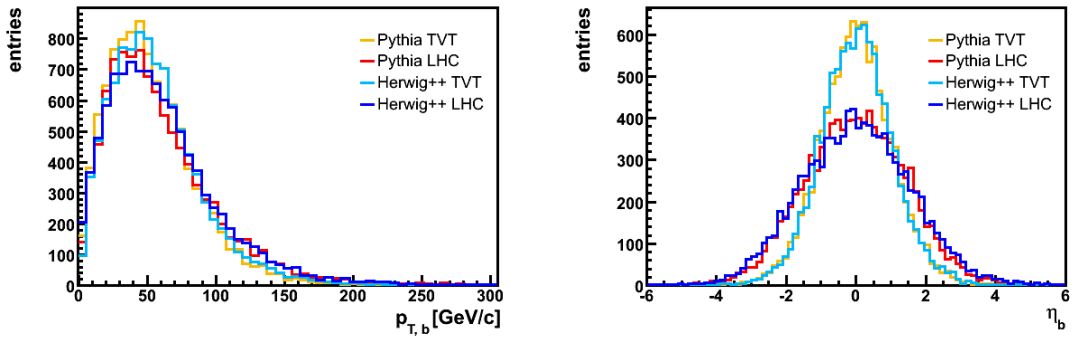
Further distributions of top quark properties can be found in figure 3.9. The transverse momentum p_T at LHC is harder than at the Tevatron and the pseudorapidity (η) distribution is widened. There are no discrepancies between the predictions of Pythia and Herwig++ for LHC. Both generators also agree at the center-of-mass energy of Tevatron.



(a) (b)

Figure 3.9: Transverse momentum (a) and η distribution (b) of individual top quarks
TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

This agreement continues in the p_T and η distributions for the decay products: the b quark (figure 3.10) and the charged lepton (figure 3.11). The difference between LHC and Tevatron nearly disappears in the transverse momentum, while it remains in the η distributions.

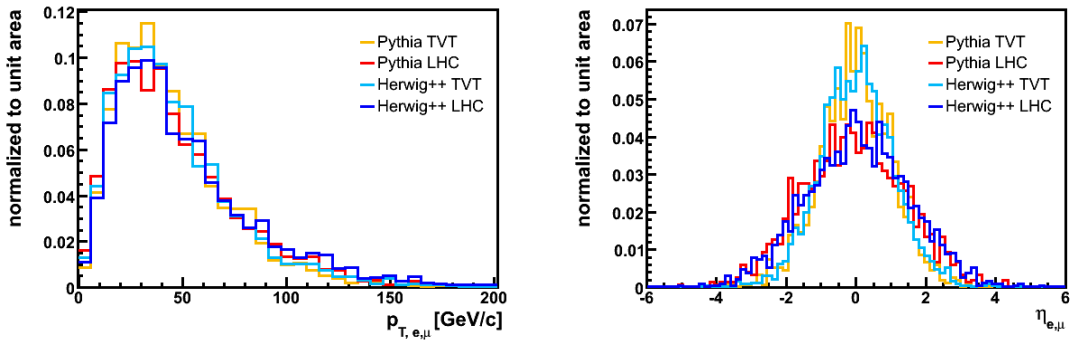


(a) (b)

Figure 3.10: Transverse momentum (a) and η distribution (b) of b quarks
TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

The final hadrons were clustered to form jets. Therefore the K_T jet clustering algorithm [52, 53] was applied. The initial parameters of the K_T algorithm were set to:

- The y_cut parameter was set to 0.013.
  This parameter determines the termination of the clustering.



(a) (b)

Figure 3.11: Transverse momentum (a) and η distribution (b) of leptons TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

- The cone_size was set to 0.54.
  This parameter defines the cone of the jets and is equivalent to a cone of 0.4 in R (see equation 2.5)

As input we used all stable particles with $p_T > 1\ GeV$ and $\eta < 5$, except neutrinos, charged leptons and their children. The result for lepton+jets events is shown in figure 3.12. Events generated with Pythia have the expected n jet distribution, more jets per event at the LHC which results from initial and final state radiation due to the higher center-of-mass energy.

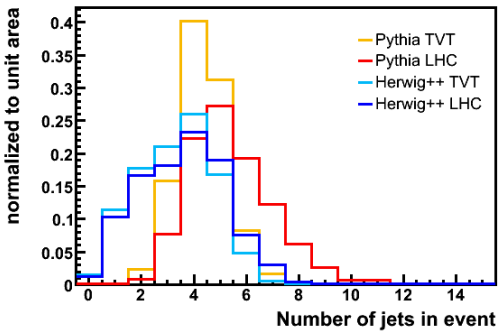


Figure 3.12: Number of jets in lepton+jets events.TVT: $\sqrt{s} = 1.96$ TeV; LHC: $\sqrt{s} = 14$ TeV.

3.2.1 Conclusion

Comparison between Tevatron and LHC

- Only small difference in p_T distributions.
- The rapidity and pseudorapidity distributions are broader in the case of the LHC, which indicates the larger boost of $t\bar{t}$ pairs at the LHC.

Comparison between Pythia and Herwig++

- Excellent agreement in distributions related to the hard process. This was expected, since the used matrix elements are the same.
- Different settings for initial state radiation as can be seen from the $\hat{p}_T$ distribution depicted in figure 3.6b.
- Large difference in hadronization if Tune A is used. Smaller difference if Tune A is not used.
- Large difference in the n jet distribution. Pythia generates the expected n jet distributions for Tevatron and LHC, while the distributions obtained from Herwig++ are very similar.

Chapter 4

Selection of Candidate Events

This chapter describes briefly the used Monte Carlo samples. The event selection which is needed to extract the candidate events is also explained, as well as the most important backgrounds. The last section gives a short description of the CMS software and the LHC Computing Grid which were used to develop and to execute the event selection.

4.1 Monte Carlo Samples

To develop the analysis method we generated two MC samples one with Pythia and the other with MadEvent/MadGraph. Each sample contains one million $t\bar{t}$ events with W boson helicity fractions as predicted by the Standard Model. Events from the later sample were handed to Pretty Good Simulation (PGS) [54] to simulate the CMS detector. The PGS package is a simple simulation package for generic collider detectors and was originally designed for the detectors at the Tevatron collider.

Then the measurement method is validated by applying it to simulated events with different W boson helicity fractions. In total 13 different samples, each with one million $t\bar{t}$ events, were generated using MadEvent/MadGraph. These samples were also simulated with PGS.

To modify the fraction F_0 of longitudinally polarized W bosons the mass of the generated top quarks was changed according to equation 1.7. The fraction F_+ of right-handed W bosons is adjusted by changing the form factors of the left-handed (f^L) and right-handed (f^R) couplings. Equation 1.9 denotes the connection between these form factors and the helicity fractions of the W boson. The fraction F_+ of right-handed W bosons increases at the expense of the fraction F_- of left-handed W bosons, while F_0 remains unchanged. Table 4.1 summarizes the used MC samples and their tuning parameters.

To be comparable with other CMS analyzes we finally use the $t\bar{t}$ and background samples generated by the CMS collaboration (Table 4.2). The $t\bar{t}$ events were generated with TopRex and the W plus n jets and $Wb\bar{b}$ plus n jets samples

Sample	F_-	F_0	F_+	m_t [Gev/c^2]	f^L	f^R
$t\bar{t}$_pythia	0.3	0.7	0	175	1	0
$t\bar{t}$_madevent	0.3	0.7	0	175	1	0
hel_125_00	0.45	0.55	0	125	1	0
hel_140_00	0.40	0.60	0	140	1	0
hel_155_00	0.34	0.65	0	155	1	0
hel_170_00	0.31	0.69	0	170	1	0
hel_195_00	0.25	0.75	0	195	1	0
hel_225_00	0.20	0.80	0	225	1	0
hel_175_00	0.30	0.7	0	175	1	0
hel_175_05	0.25	0.7	0.05	175	1	$\sqrt{0.2}$
hel_175_10	0.20	0.7	0.10	175	1	$\sqrt{0.5}$
hel_175_15	0.15	0.7	0.15	175	1	1
hel_175_20	0.10	0.7	0.20	175	1	$\sqrt{2}$
hel_175_25	0.05	0.7	0.25	175	1	$\sqrt{5}$
hel_175_30	0	0.7	0.30	175	0	1

Table 4.1: Monte Carlo samples and their helicity fractions. In the last three columns the tuning parameters used to create the samples are denoted. F_-, F_0 and F_+ denote the fraction of left-handed, longitudinal and right-handed polarized W bosons, respectively. f^L and f^R are the corresponding form factors.

with AlpGen[1]. These samples[2] were simulated, digitized, and reconstructed. All steps from the event generation to the reconstruction are done in the CMS software framework. The first step, the MC event generation, was described in the previous chapter. Now we dwell on the following three steps.

The simulation of the CMS detector is based on the GEANT4 [64] package. It describes the interactions between the particles and the detector material. CMSSW provides the geometry of the detector and GEANT uses it to model the propagation of the particles through the material. Deflection of charged particles in the magnetic field, electromagnetic and hadronic showering, as well as ionization and scattering processes are simulated.

In the digitization step energy depositions by the mentioned processes are then translated into electronic signals to simulate the response of the various detector components. The resulting signals from the tracker, the electromagnetic and hadronic calorimeter, and the muon chambers are comparable to signals from real events.

[1] AlpGen v2.05 with CMS settings.

[2] All MC samples presented in Table 4.2 were generated in the Spring07 MC production. The ttbar and W+n jets samples were generated, simulated, digitized, and reconstructed in CMSSW 1.3.1, while CMSSW 1.3.5 was used for the Wbb+n jets samples. The bug concerning the electrons in the CMSSW 1.3.1 Spring07 production was fixed by applying the appropriate algorithm before the selection of candidate events.

Sample	Process	Cross Section [pb]	Generator	Events
ttbar	$pp \to t\bar{t}$	830	TopRex	2.7M
W4j	$pp \to W + 4jets$	174	AlpGen	72K
W5j	$pp \to W + 5jets$	45	AlpGen	55K
W6j	$pp \to W + 6jets$	31	AlpGen	20K
Wbb2j	$pp \to W + b\bar{b} + 2jets$	3.6	AlpGen	148K
Wbb3j	$pp \to W + b\bar{b} + 3jets$	3.8	AlpGen	23K

Table 4.2: Monte Carlo samples

In the last step, the reconstruction, the digitized signals are used to create more complex objects like particle tracks, vertices, and jets. Tracks are obtained by fitting the measured hits in the tracker system. Vertices are intersections of such tracks. Jets are computed using certain algorithms which take the energy deposits in the calorimeters (calo towers) as input.

4.2 Selection of Candidate Events

We want to select lepton+jets events, since they have some advantages over dilepton events, namely higher branching ratio and only one neutrino. The advantages over hadronic events are lower background and fewer possible jet combinations in the reconstruction of the $t\bar{t}$ pair. The full event reconstruction is explained in chapter 5. The signature of a lepton+jets event is:

- one isolated charged lepton,
- missing transverse energy,
- and at least four jets, two of them from b quarks.

The neutrino from the leptonic W boson decay appears as missing transverse energy, since it does not interact with the detector material. We concentrate on the events where the charged lepton is an electron or a muon, because these particles especially the muon can be well measured. Events where the charged lepton is a tau are not considered. Tau leptons decay inside the detector into a neutrino and a W boson. This leads to different event topologies with two additional jets, if this W boson decays hadronically, or to two neutrinos and a second charged lepton, if the W boson decays leptonically. All possible final states with the described signature are highlighted using red rectangles in figure 1.4, their branching ratio is 24/81.

In order to select the desired lepton+jets events and to suppress the possible backgrounds, which are described in the next section, some selection cuts were applied. Thereby the signature of these events was exploited.

Lepton Requirements

Since the charged lepton we want to select comes from a W boson decay it has to satisfy the following criteria. It has to be an isolated electron or muon. A lepton is considered to be isolated when the isolation transverse momentum p_T^{iso} does not exceed 0.01 GeV/c. The isolation transverse momentum is calculated in the following way. At the primary vertex a cone of opening angle ΔR=0.3 is constructed around the lepton. Tracks are considered associated to this primary vertex if their point of closest approach along the z-direction is below 2 mm and closer than to any other vertex. The scalar sum of the transverse momentum of all tracks associated, minus the transverse momentum of the lepton track is defined as the isolation transverse momentum [55]. The second requirement is a value called combined likelihood ratio, which has to exceed 0.5. The combined likelihood ratio is computed to identify leptons from W boson decays, therefore five different variables are considered.

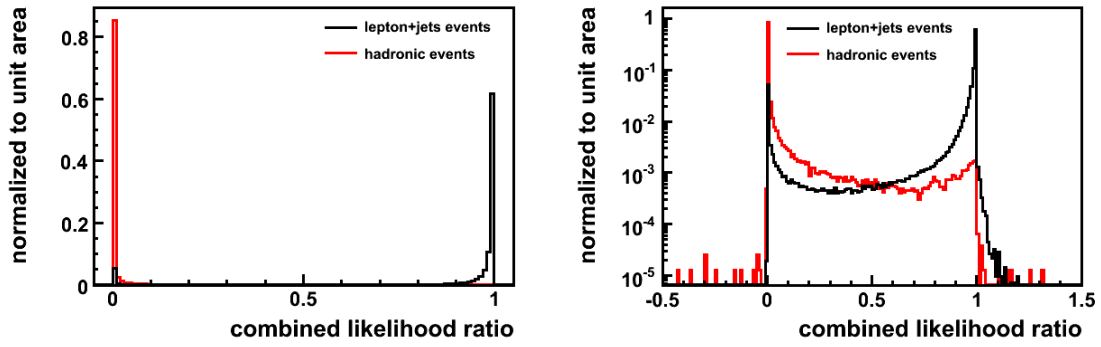
- Isolation energy E_{iso}: the sum of all energy deposits in the calorimeter which are in a cone of opening angle $\Delta R = 0.3$ around the lepton direction.
- Isolation angle α_{iso}: the smallest angle among all angles between the reconstructed lepton candidate and all reconstructed jets.
- The logarithm of the transverse momentum $\log(p_T)$ of the reconstructed lepton candidate.
- Primary vertex significance S_z: The significance of the distance along the z-direction between the reconstructed and fitted primary vertex and the lepton's point of closest approach to this vertex.
- The logarithm of the isolation transverse momentum $\log(p_T^{iso})$ which is already described above.

For each variable the likelihood ratio is calculated. Then these ratios are combined to obtain a single value. A detailed description can be found in reference [55]. Figure 4.1 shows the combined likelihood ratio for lepton+jets events in black and hadronic events in red. This requirement rejects nearly all hadronic events.

The transverse momentum of electrons is required to be $p_T > 26$ GeV/c. The pseudo rapidity range for electrons is $|\eta| < 2.4$. For muons these requirements are slightly different. We use $p_T > 19$ GeV/c and $|\eta| < 2.1$ [55]. A dilepton veto ensures that events with more than one leptons fulfilling these requirements are discarded.

Jets

The final state in the lepton+jets channel contains of four quarks which hadronize to many secondary particles forming jets. These jets are measured as energy deposits in the electromagnetic and hadronic calorimeters. We use the iterative cone



(a) (b)

Figure 4.1: The combined likelihood ratio separates leptons from W boson decay (lepton+jets events) from fake leptons which are not from W boson decay (hadronic events). (a) shows the combined likelihood ratios for lepton+jets event (black) and for hadronic events on a normal scale. About 5% of lepton+jets events are assigned as hadronic events. (b) shows both distributions using a logharitmic scale to zoom in the region with small values.

algorithm [57, 58] to reconstruct these jets. In the iterative cone algorithm, an E_T-ordered list of input calorimeter towers is created. The list does not contain towers below a certain E_T threshold. The E_T threshold values depend on the η region and are specified in [57]. A cone of size R in the $\eta - \phi$ plane is cast around the input object having the largest transverse energy. The towers inside the cone are used to calculate a proto-jet direction and energy. This is done by summing up the transverse energy of the towers. The computed direction is used to seed a new proto-jet. The procedure is repeated until the energy of the proto-jet changes by less than 1% between iterations and the direction of the proto-jet changes by $\Delta R < 0.01$. When a stable proto-jet is found, all towers in the proto-jet are removed from the list of input objects and the stable proto-jet is added to the list of jets. The whole procedure is repeated until the list contains no more towers. Then a merging/splitting algorithm is applied to obtain separated jets.

In the event selection we use at least four jets with a cone size of R = 0.5. To avoid too many jet combinations in the reconstruction of $t\bar{t}$ pairs, events with more than six jets are not taken into account. Due to several calorimeter and physics effects the energy of the jets has to be corrected. The jet calibration techniques implemented in the CMS reconstruction software were studied with high-statistics Monte Carlo samples of QCD dijet events. Generated events were passed through a full detector simulation of the CMS detector. The jets were reconstructed using the iterative cone algorithm (R = 0.5). Particle-level and reconstructed jets were found by applying the same jet algorithm to stable particles (excluding neutrinos and muons) and calorimeter cells, respectively. A matching criterion, based on the distance $\Delta R < 0.2$, was used to associate particle-level and reconstructed jets. The

calibration coefficients were then determined using the ratio $r_{\text{jet}} = E_T^{Rec}/E_T^{MC}$ for 16 different η regions. This Monte Carlo calibration of jet response is described in reference [57]. We apply these jet corrections and select jets with a transverse momentum exceeding 30 GeV/c and $|\eta| < 2.4$. For the used $t\bar{t}$ sample the jet corrections are not yet very suitable. Figure 4.2 illustrates the difference between the transverse momenta of reconstructed and generated jets, normalized to the transverse momentum of the generated jet. The same matching criterion $\Delta R < 0.2$ was used to associate generated and reconstructed jets. The transverse momenta of the reconstructed jets have on average larger values than the transverse momenta of generated jets. This has no influence on the $\cos\theta^*$ angle, but we have to bear this in mind for the reconstruction of $t\bar{t}$ pairs.

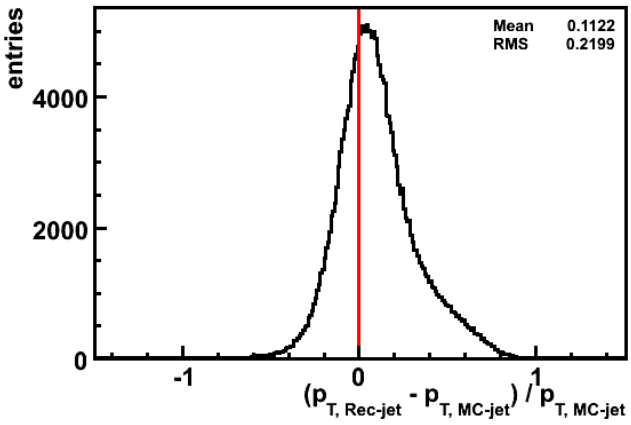


Figure 4.2: Comparison between the reconstructed jets with applied corrections and the generated MC jets.

There is one last item concerning jets. Since leptons are also measured in the hadronic calorimeter, they have to be removed from the list of jets. This is called jet cleaning. Jets inside a cone of R = 0.5 around an isolated lepton are omitted.

Missing Transverse Energy

The event topology contains a neutrino. Since neutrinos do not interact with the detector material they are only measured as missing transverse energy $\not\!\!E_T$. The missing transverse energy vector is calculated by summing individual calorimeter towers having energy E_n, pseudorapidity η, and azimuthal angle ϕ [56]:

$$\vec{\not\!\!E}_T = -\sum_n \left(\frac{E_n \cos\phi_n}{\cosh\eta_n}\hat{x} + \frac{E_n \sin\phi_n}{\cosh\eta_n}\hat{y} \right). \tag{4.1}$$

The magnitude of this vector is defined as $\not\!\!E_T = |\vec{\not\!\!E}_T|$. Since this calculation is based on the calorimeter towers, $\not\!\!E_T$ has to be adjusted for the effect of the jet corrections.

Reconstructed muons are taken into account by replacing the expected calorimeter deposit (about 4 GeV) with the reconstructed track p_T. In our analysis the corrected missing transverse energy in an event is required to be at least 30 GeV.

b-tagging

The identification of jets containing b quarks relies on the properties of b hadrons decays. b hadrons have a lifetime of about 1.6 ps and they produce, on average, 5 charged particles per decay. The lifetime of the b hadron and its high momentum lead to an observable traveled distance. Therefore, tracks caused by decay products of a b hadron intersect in a secondary vertex, which is displaced from the primary interaction vertex. This information can be exploited using the impact parameter. The impact parameter is the distance of the point of closest approach of a track (perigee) to the primary vertex. As shown in figure 4.3, tracks originating from b decays have large impact parameters, since they come from a displaced vertex, while the impact parameters of tracks coming from the primary vertex are compatible with the tracking resolution.

We use the track probability algorithm [59]. First the algorithm selects tracks by applying some criteria like thresholds for the transverse momentum of the tracks and the number of hits in the tracker. Then for each selected track the probability to stem from the primary vertex is computed, using its impact parameter. Finally these probabilities are combined to provide the jet probability.

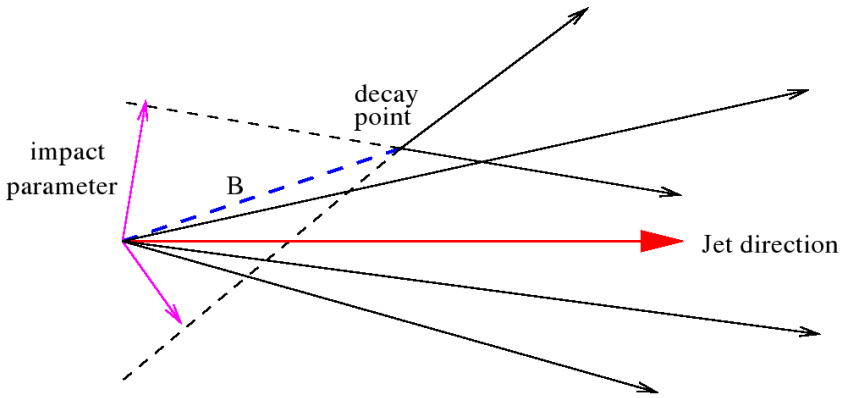


Figure 4.3: Representation of an hadronic jet originating from a b quark [59]. The decay point of the b hadron represents the secondary vertex with two tracks, while the other three tracks intersect in the primary vertex. The impact parameter is also shown.

Jets with a jet probability exceeding 0.4 are selected as b jets. Since the b-tagging efficiency is about 50% we require at least one b-tagged jet, although two b-tagged jets are expected in $t\bar{t}$ events.

Summary

Table 4.3 shows the cut flow for the TopRex $t\bar{t}$ sample. The lepton requirements cause the largest reduction of the initial number of events, because they are the first applied cuts. Only 20% of events remain although the fraction of generated events with at least one lepton is 54% as can be seen in figure 4.4. The following cuts on missing transverse energy and number of b tags reject less events. In 60% of the remaining events the observed number of jets is lower than four. These events are also rejected since we need at least four jets to reconstruct the $t\bar{t}$ pair.

Cut	0 jets	1 jets	2 jets	3 jets	4 jets	≥5 jets	all
Total	15855	106256	331719	612836	730995	965665	2763326
lepton requirements	0	44810	129531	186251	133792	66111	560495
$\not\!\!E_T > 30$ GeV	0	41257	116187	163025	115971	57900	494340
b tag ≥ 1	0	20221	80495	129743	98631	50505	379595

	4 jets	5 jets	6 jets	
$4 \leq$ jets ≤ 6	98631	37497	10244	146372

Table 4.3: Cut flow table of $t\bar{t}$ event selection for events of the $t\bar{t}$ signal MC sample generated with TopRex. The different cuts are described in the text.

Figure 4.4 shows the fractions of the different $t\bar{t}$ decay channels in the TopRex sample before (on the left) and after (on the right) the selection. Nearly 85% of the selected events are lepton+jets events. The selection efficiency for lepton+jets events is 30%. The largest fraction before the selection, the hadronic events, is reduced to 1.38% especially by the cut on the combined likelihood ratio. The tau+jets and the dilepton fractions are also reduced. We combine all three non lepton+jets channels and refer to them as $t\bar{t}$-other. This events are considered as background.

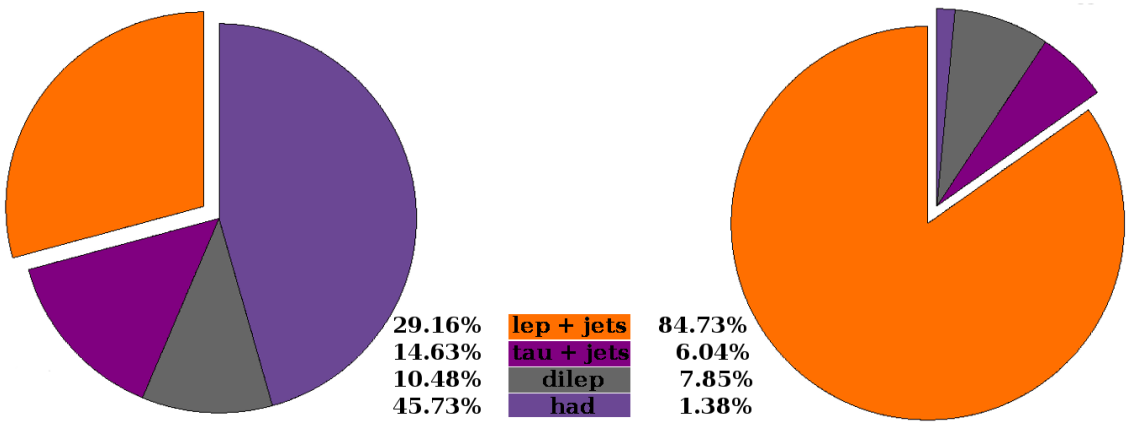


Figure 4.4: Fractions of the different $t\bar{t}$ decay channels in the TopRex sample before (on the left) and after (on the right) the selection.

4.2.1 Expected Numbers of Signal and Background Events

We now compute the expected numbers for the signal and background events for an integrated luminosity of 1 fb^{-1}. We multiply the cross section of the considered process ($pp \to t\bar{t}$ or $pp \to W + jets$) with the ratio of selected to generated events. For the signal (lepton+jets events) and the $t\bar{t}$ background ($t\bar{t}$-other) the branching ratios of 24/81 and 57/81 respectively are also included. Table 4.4 shows the obtained results. More than 50% of the background events are $t\bar{t}$-other events, the remaining fraction consists of $W + n$ jets events. In total 23% of the selected events are background events. Therefore further analysis will try to reduce especially the fraction of selected $t\bar{t}$-other events.

Signal		
sample	cross section [pb]	events
lepton+jets	$833 \cdot 24/81$	37961
Background		
sample	cross section [pb]	events
$t\bar{t}$-other	$833 \cdot 57/81$	6538
$W + 4$ jets	174	239
$W + 5$ jets	45	156
$W + 6$ jets	31	82
$W + b\bar{b} + 2$ jets	3.6	48
$W + b\bar{b} + 3$ jets	3.8	104
Total background		11475

Table 4.4: Expected events for an integrated luminosity of 1 fb^{-1}

4.3 CMS Software and LHC Computing Grid

The event selection introduced in the last section is implemented as a module for the CMS software framework (CMSSW) and it uses another module called Top Quark Analysis Framework[3]. In this section we briefly describe the CMS software framework, the Top Quark Analysis Framework, and the LHC Computing Grid on which the selection is executed.

CMS Software framework

CMSSW is a software framework used to analyze data in CMS. CMSSW consists of over a thousand subpackages which have been created to provide an extensive toolkit for users to carry out analyses of data and perform other software-related tasks with only a small contribution of code by themselves. The overall collection of software is built around a Framework, an Event Data Model (EDM), and Services needed by

[3]The software versions we used are CMSSW 1.3.6 and TQAF_136_070822.

the simulation, calibration and alignment, and reconstruction modules that process event data. The primary goal of the framework is to facilitate the development and deployment of reconstruction and analysis software.

The CMSSW framework implements a software bus model wherein there is one executable, called cmsRun, and many plug-in modules which run algorithms. The same executable is used for both detector and Monte Carlo data, it is configured at run time by the user's job-specific configuration file. This file tells cmsRun which data to use, which modules to run, which parameter settings to use for each module, and in what order to run the modules. Required modules are dynamically loaded at the beginning of the job.

The EDM is centered around the concept of an Event as a C++ object container for all RAW and reconstructed data pertaining to a physics event. During processing, data are passed from one module to the next via the Event, and are accessed only through the Event. All objects in the Event may be individually or collectively stored in ROOT files.

Top Quark Analysis Framework

Since many analyses are to be performed on the same top quark events (single top or top quark pairs) and are often based on the same reconstruction tools and final states, a common framework is being developed: the Top Quark Analysis Framework (TQAF). The analysis code is structured in three layers:

1. reconstruction of general-purpose TopObjects (final state and analysis independent);
2. construction of event solutions with these TopObjects (final state dependent but analysis independent);
3. the actual analyses to perform physics measurements (final state and analysis dependent).

This structure allows the use of the same framework modules in a whole range of physics analysis. Only the first layer of the TQAF is used for the analysis presented in this thesis. This means that all selected objects are comparable to the ones of other analysis, which also use the TQAF. But the full reconstruction of $t\bar{t}$ pairs is completely independent.

LHC Computing Grid

The LHC Computing Grid, further referred to as Grid, is the next step in exploiting networked computer power. Currently the Internet and World Wide Web allow us to share information and transfer data quickly and easily around the world. In the future the Grid will let us share computer processing power, software packages, and data storage space.

The Grid has many applications, but its first major application is to allow users to share global computing power to manage and process the huge quantities of data that will be produced by the LHC. By linking desktop computers in a global network, managed by so-called middleware, the Grid brings supercomputing power to desktops. Figure 4.5 shows a snapshot of grid with some active connections.

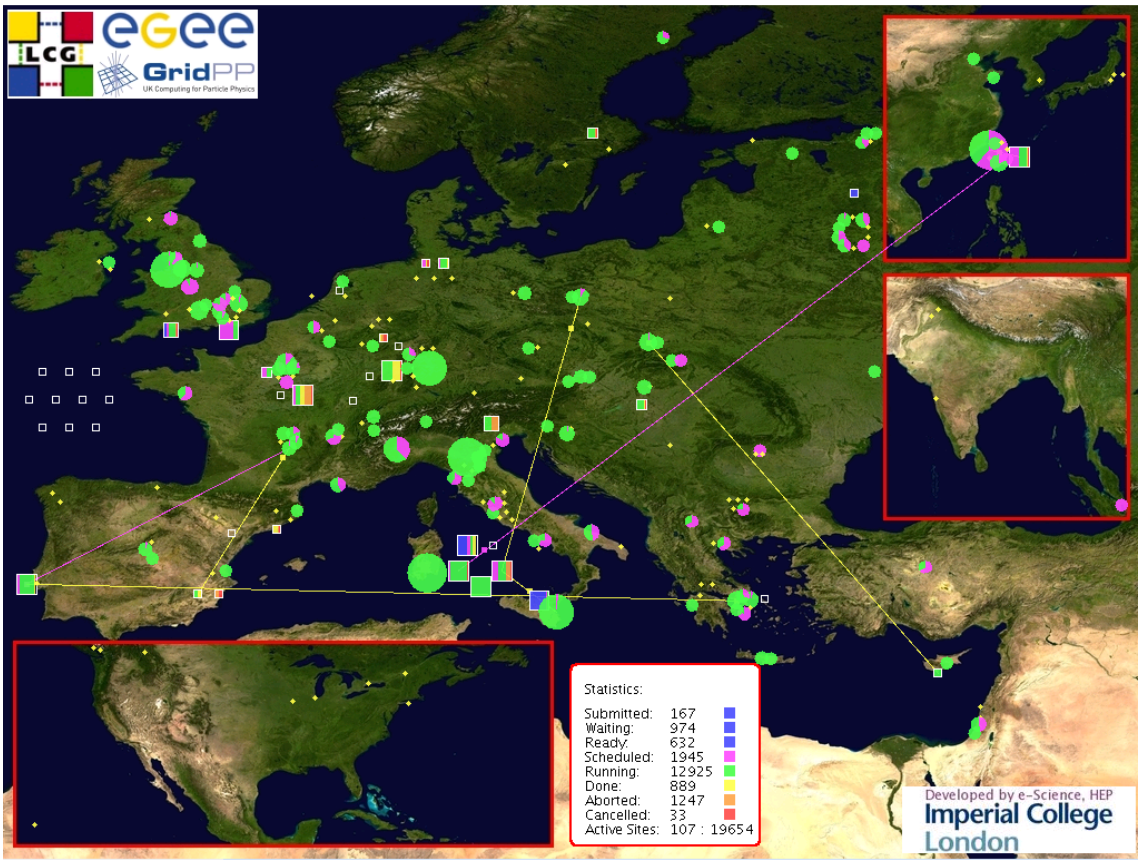


Figure 4.5: Grid snapshot with some active connections.

The event selection of the signal sample was processed on the following sites: at CERN, in France, and at the Fermi National Accelerator Laboratory in the USA. The background samples were processed at a cite in Belgium, since these MC samples are stored only there.

The Grid itself is implemented as a hierarchic four tiered model:

- One single Tier-0 center located at CERN records the raw event data as emerging from the experiment high-level triggers. These data is written to tape after having performed a first-pass reconstruction. A second copy is transfered to one of the Tier-1 centers associated with the experiment. According to the policy of the experiment, reconstructed data are also transfered to the Tier-1 center.

- The Tier-1 centers are interconnected with CERN using high-speed networks. Their role differs slightly for the different LHC experiments, but in general they perform the reconstruction of the raw data and manage the permanent storage of raw, reconstructed, and simulated data.

- The role of the Tier-2 centers is mostly to provide sufficient computational resources for theoretical simulations and physics analysis. Reconstructed data are received from allocated Tier-1 centers and simulated data are sent to these Tier-1 centers for permanent storage.

- Other laboratories or universities might take part in analyzing LHC data as Tier-3 facilities. These sites might make significant contributions to the experiment's needs as they provide computing resources for interactive user analysis or theoretical simulations.

To ease the use of the grid, members from the CMS group of the IEKP developed a tool called grid-control. This tool is used to submit the event selection to the grid. The further execution chain is controlled by the grid applications. Figure 4.6 illustrates the several steps. Grid-control plays the role of the application which submits the job to the grid. In a first step the Planer checks the identity of the user (Policy/Security), and the availability of the dataset by using the Catalog Services. If this check is passed the Executer looks for Compute Resource at the sites where the data are stored. After the execution of the program the resulting output is transfered back to the user.

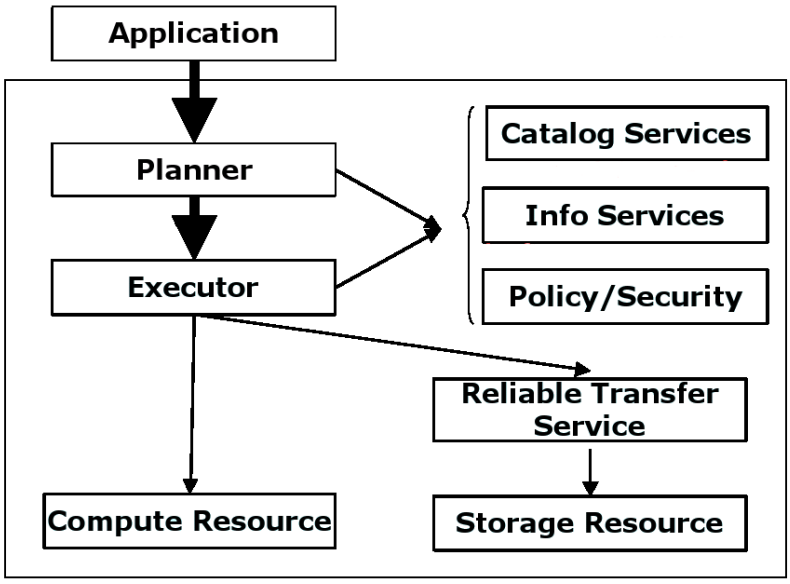


Figure 4.6: Schematic overview of the execution process on the Grid.

Chapter 5

Full Reconstruction of $t\bar{t}$ Events

Having done the selection we can now reconstruct $t\bar{t}$ pairs. From here on the analysis is independent from CMSSW and Grid resources. This chapter describes the full reconstruction wherein several ambiguities occur. As a consequence, mostly 24 different interpretations or hypotheses for a four jet $t\bar{t}$ event are obtained. The challenge is to select the correct event interpretation. We specify the selection method and review it using the information of the Monte Carlo events.

5.1 Reconstruction Algorithm

To reconstruct $t\bar{t}$ events means to assign the objects measured by the detector, in this case the objects after the detector simulation, to the final-state particles. The Feynman diagram in figure 5.1 visualizes the final-state quarks and leptons in the top quark decay.

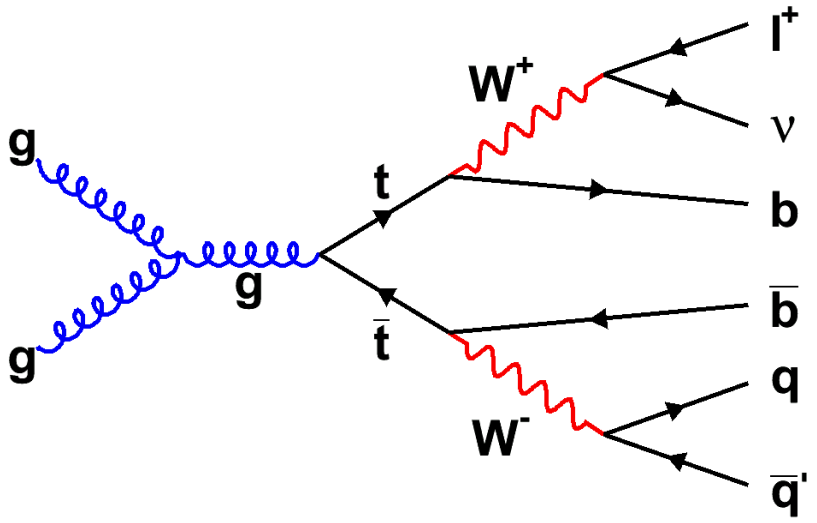


Figure 5.1: Feynman diagram of $t\bar{t}$ production via gluon fusion and decay in the lepton+jets channel.

The charged lepton, in our selected events an electron or muon, can be identified by its signature in the detector components, energy deposits in the electromagnetic calorimeters and the muon chambers. Hence, the assignment of this object is unambiguous. The neutrino, which does not interact with the detector material, can only be measured as missing transverse energy. This means we have to reconstruct the z component of its momentum. As described later on this leads to two different solutions. The remaining particles are two b quarks and two light quarks. Due to confinement quarks are not observed as free quarks but hadronize and therefore appear in the detector in the form of jets. Therefore, it is not possible to distinguish between quarks and antiquarks or the quark flavors. For this reason a jet can be assigned to each of the quarks leading to several event interpretations. Certainly, there is one exception where it is possible to assess the quark flavor namely b quarks. We will use the b tagging information when selecting one of the obtained event interpretations. The reconstruction is done in several steps.

Leptonically decaying W Boson: $W \to \ell\nu_\ell$

The reconstruction of the $t\bar{t}$ pair starts with the charged lepton which can be reconstructed almost perfectly. As a second object the neutrino is reconstructed. The x and y components of the neutrino momentum $\vec{P}_{\mathrm{T},\nu}$ are given by $\not\!\!\vec{E}_T$. Since the neutrino comes from the W boson decay, we can use the well known mass of the W boson $m_W = 80.4$ GeV/c^2 as a constraint. This leads to a quadratic equation for the z component $P_{z,\nu}$ of the neutrino momentum.

$$P_{z,\nu}^2 - 2 \cdot \frac{\mu \cdot P_{z,\ell}}{E_\ell^2 - P_{z,\ell}^2} \cdot P_{z,\nu} + \frac{E_\ell^2 \cdot P_{\mathrm{T},\nu}^2 - \mu^2}{E_\ell^2 - P_{z,\ell}^2} = 0 \tag{5.1}$$

$$\text{with} \qquad \mu = \frac{m_W^2}{2} + \cos(\Delta\Phi) \cdot P_{\mathrm{T},\ell} \cdot P_{T,\nu} \tag{5.2}$$

Here $P_{z,\ell}$ and E_ℓ denote the z component of the momentum and the energy of the charged lepton (electron or muon) respectively. The quantity $\Delta\Phi$ is the azimuthal angle difference between the momentum of the charged lepton and $\not\!\!\vec{E}_T$. In general, a quadratic equation leads to two solutions. We store both solutions and decide after the full reconstruction which to choose. In about 41% of all cases there is a complex solution. In these cases the measured transverse mass of the W boson, which is defined as

$$M_{\mathrm{T},W\to\ell\nu_\ell}^2 = (P_{\mathrm{T},\ell} + P_{\mathrm{T},\nu})^2 - (P_{x,\ell} + P_{x,\nu})^2 - (P_{y,\ell} + P_{y,\nu})^2 \quad , \tag{5.3}$$

is larger than $M_W = 80.4$ GeV/c^2 as demanded using the W mass constraint. Choosing just the real part of the $P_{z,\nu}$ solution, leads to bigger reconstructed W boson mass, since the real part does not obey equation 5.1. Figure 5.2 shows the differences between the generated and reconstructed transverse W boson masses. The reason is obviously a mismeasurement of $\not\!\!\vec{E}_T$. To avoid such a complex solution we fix the transverse mass $m_{\mathrm{T},W}$ to 80.4 GeV/c^2, and obtain a quadratic dependence

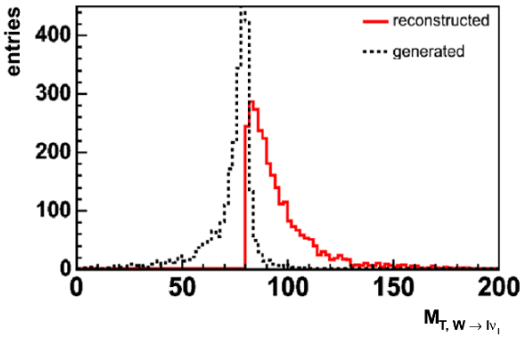
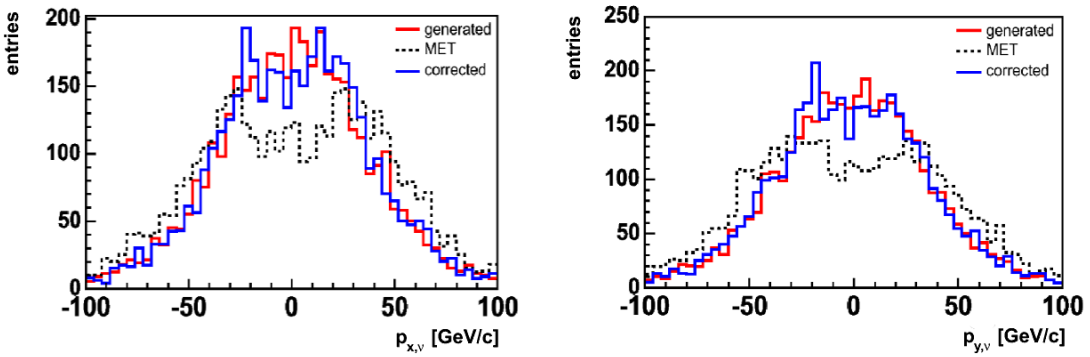


Figure 5.2: Difference between the generated and reconstructed transverse W boson masses. The dashed line depicts the transverse masses of all generated W bosons, while the red line depicts the reconstructed transverse masses in cases with a complex solution using only the real part of equation 5.1.

of $P_{y,\nu}$ on $P_{x,\nu}$. We now want to correct these values. Under the assumption that the $\vec{\not{E}}_T$ measurement is not absolutely wrong we define the difference

$$\delta(P_{x,\nu}) = \sqrt{(P_{x,\nu} - \not{E}_{T,x})^2 + (P_{y,\nu}(P_{x,\nu}) - \not{E}_{T,y})^2} \quad . \tag{5.4}$$

This difference is minimized with respect to $P_{x,\nu}$, in a range where $P_{y,\nu}$ does not become complex. Since there are in general two solutions for $P_{y,\nu}$, we may find two values for δ. In these cases we take the solution which is closest to the measured $\vec{\not{E}}_T$, that means the smaller δ value. Figure 5.3 illustrates the resulting $P_{x,\nu}$ and $P_{y,\nu}$ values as well as the generated and measured $\vec{\not{E}}_T$ components. The corrected $P_{x,\nu}$

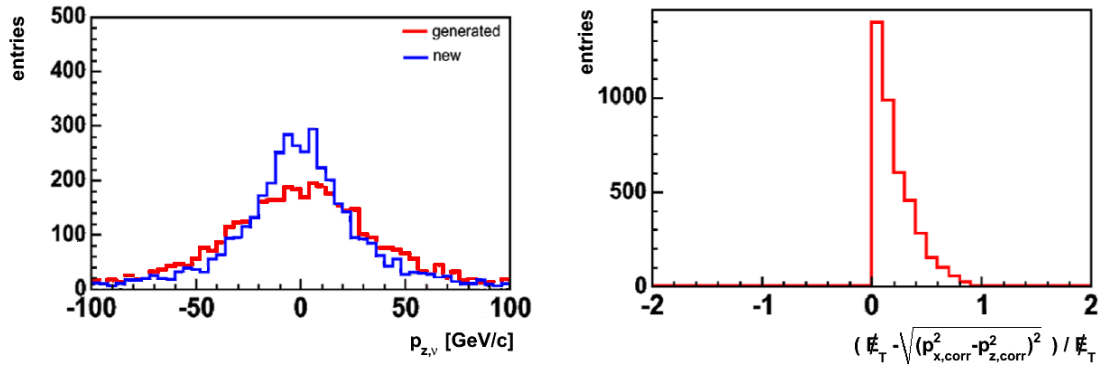


(a) (b)

Figure 5.3: The corrected $P_{x,\nu}$ and $P_{y,\nu}$ distributions (blue) follow the generated (red) distributions, while the measured $\not{E}_{\mathrm{T},x}$ and $\not{E}_{\mathrm{T},y}$ distributions (dashed) are broader.

and $P_{y,\nu}$ distributions (blue) follow the generated (red) distributions well, while the measured $\not\!\!E_{T,x}$ and $\not\!\!E_{T,y}$ distributions (dashed) are broader.

The distribution of the reconstructed $P_{z,\nu}$ component is shown in figure 5.4a. The deviation between the reconstructed and generated $P_{z,\nu}$ component is comparable with the deviation of the $P_{x,\nu}$ and $P_{y,\nu}$ distributions. The difference between the corrected and measured $\not\!\!E_T$ is depicted in 5.4b. The corrected $\not\!\!E_T$ values are always lower or equal to the measured ones.



(a) (b)

Figure 5.4: (a) the reconstructed $P_{z,\nu}$ component is broader than the generated one. (b) the corrected $\not\!\!E_{\mathrm{T}}$ values are always lower or equal to the measured ones.

The four-momentum of the leptonically decaying W boson is reconstructed by adding the four-momenta of the charged lepton and each neutrino hypothesis, respectively.

Semileptonically decaying *Top* Quark: $t \to b\ell\nu_\ell$

In order to reconstruct the semileptonically decaying top quark all hypotheses obtained by adding the four-momentum of each selected jet and the four-momentum of the leptonically decaying W boson are considered. The number of possible hypotheses up to this step is equal to the number of selected jets N_{jets} times the number of solutions for the z component of the neutrino momentum.

Hadronically decaying W Boson: $W \to jj$

The reconstruction of the four-momentum of the hadronically decaying W boson is done by combining the four-momenta of two of the remaining jets, which are not assigned to the semileptonically decaying top quark. Hence there are $(N_{\mathrm{jets}} - 1) \cdot (N_{\mathrm{jets}} - 2)$ combinations to reconstruct the hadronically decaying W boson. The order in which the two four-momenta of the jets are added does not change the four-momentum of the W boson. Therefore, the combinatorics is reduced by a factor of two.

Hadronically decaying *Top* Quark: $t \rightarrow bjj$

So far three of the final-state quarks are covered. In this last step a jet is assigned to the b quark from the hadronically decaying top quark. The four-momentum of one of the $(N_{\text{jets}} - 3)$ remaining jets is added to the four-momentum of the hadronically decaying W boson to obtain the four-momentum of the hadronically decaying top quark.

Considering only the jets, the assignment leads to $N_{\text{jets}} \cdot (N_{\text{jets}} - 1) \cdot (N_{\text{jets}} - 2) \,/\, 2 \cdot (N_{\text{jets}} - 3)$ possible hypotheses for the complete reconstruction of the kinematics in $t\bar{t}$ events, and to $N_{\text{jets}} \cdot (N_{\text{jets}} - 1) \cdot (N_{\text{jets}} - 2) \cdot (N_{\text{jets}} - 3)$ possible hypotheses for $t\bar{t}$ events with two solutions for the z component of the neutrino momentum. Considering an event with four jets and two neutrino solutions we obtain 24 different event interpretations. To avoid large amounts of hypotheses, we select events with maximum six jets. This is justified, since the fraction of discarded events with more than six jets is less than 2%, as can be derived from table 4.3. The numbers of possible hypotheses are depicted in figure 5.5.

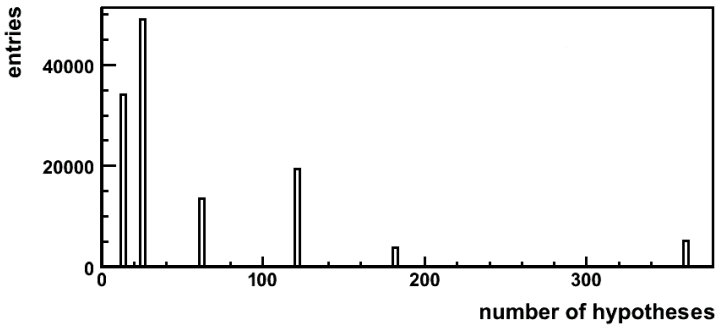


Figure 5.5: Number of reconstructed event hypotheses. There are 12 possible interpretations for an event with four jets and one neutrino solution and 360 event interpretations in events witch six jets and two neutrino solutions.

5.2 Best Possible Event Interpretation

In simulated Monte Carlo events it is possible to find that hypothesis with the smallest deviation to the generated $t\bar{t}$ pair. This is done by computing the distance ΔR in the η-ϕ–plane (see equation 2.5) between the reconstructed and generated W bosons and top quarks. We define the best possible hypothesis as the hypothesis for which the sum of the distances of the leptonically decaying W boson $\Delta R_{W \rightarrow \ell\nu}$, the semileptonically decaying top quark $\Delta R_{t \rightarrow b\ell\nu}$, the hadronically decaying W boson $\Delta R_{W \rightarrow jj}$, and the hadronically decaying top quark $\Delta R_{t \rightarrow bjj}$ to the corresponding generated particles has a minimum value:

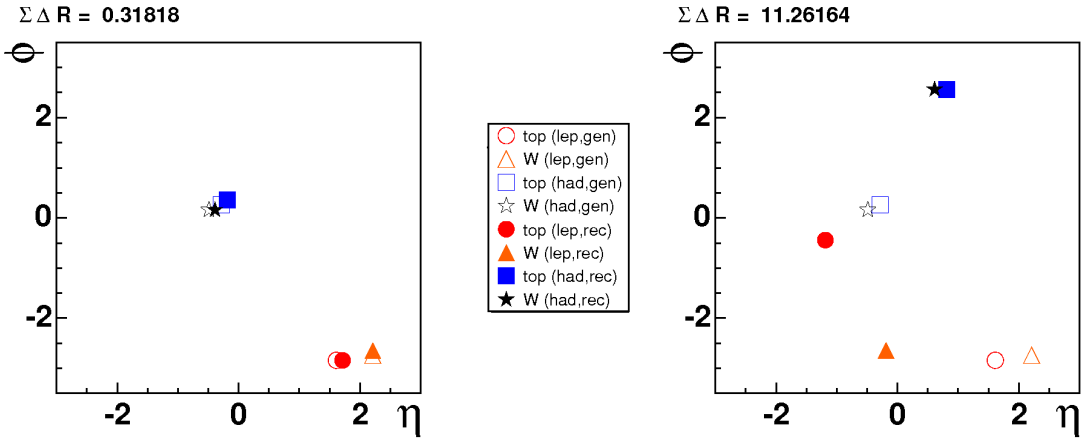


Figure 5.6: Position of the generated (referred to as "gen") and reconstructed (referred to as "rec") four-vectors of the two W bosons (the leptonically decaying W boson is indicated with "lep", the hadronically decaying with "had") and top quarks (same notation as for the W bosons) in the η-ϕ–plane for the best possible hypothesis on the left and for one of the other 23 event interpretations of the same event on the right. The corresponding values for $\sum \Delta R$ are also shown (for explanation see equation 5.5).

$$\sum \Delta R = \Delta R_{W\to\ell\nu} + \Delta R_{t\to b\ell\nu} + \Delta R_{W\to jj} + \Delta R_{t\to bjj} \quad . \tag{5.5}$$

This equation uses the fact that for every event there exists one hypothesis which is as close as possible to the generated MC event. It is always possible to arrange the four-vectors of the final state particles in such a way that the reconstructed event topology is closest to the true values. This is the case not only for not well measured jet directions but also for events where one of the selected jets is from gluon radiation. In such events one of the four jets from the final state quarks is not measured or not selected by the applied cuts. Figure 5.6 illustrates two hypotheses with different ΔR values for a four-jets event. The positions of the two reconstructed W bosons and top quarks are marked in the η-ϕ–plane as well as the positions of the generated ones.

5.3 Method to select one Event Interpretation in measured Events

For the further analysis it is necessary to select the one event interpretation for each single event. This section describes the selection method. To select an interpretation we assume that the reconstructed events are real data and use therefore no information that is available from Monte Carlo true values. The obtained method is then checked using the Monte Carlo information.

For each hypothesis we determine a quantity Ψ which gives a quantitative esti-

mate how well the hypothesis matches the $t\bar{t}$ pair assumption. We construct Ψ in such way that it reaches a minimum value for a hypothesis that matches the $t\bar{t}$ pair assumption well, while it becomes larger for a hypothesis that does not agree to the $t\bar{t}$ pair assumption. Therefore, we consider the following items:

- the probability P_ν to choose the correct neutrino solution,
- the b-likeness P_b of the jets assigned to the b quarks,
- a constraint on the mass of the hadronically decaying W boson,
- a constraint on the difference between both reconstructed top quark masses,
- and a constraint on the sum of the reconstructed transverse energy of the two top quarks which should in leading order calculation be equal to the transverse energy of the event.

The last three items are combined in a χ^2 function. Equation 5.6 denotes the quantity Ψ and its components which are now described in detail.

$$\Psi = P_\nu \cdot P_b \cdot \chi^2 \tag{5.6}$$

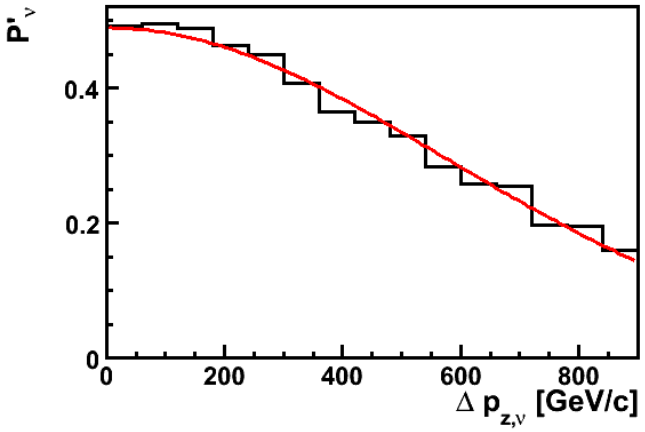


Figure 5.7: Probability P'_ν for the larger $P_{z,\nu}$ solution to be correct.

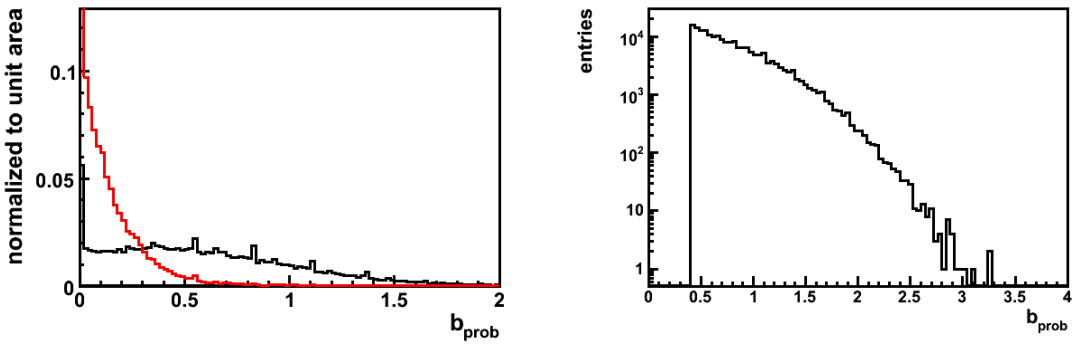
Contribution of the neutrino solution: P_ν

In about 59% there are two possible solutions for the z component of the neutrino momentum. To select one of them we determine a function P'_ν which takes the difference $\Delta P_{z,\nu}$ of the two solutions and returns a probability for the solution with the larger $P_{z,\nu}$ value to be the correct one. We exploit Monte Carlo information for this evaluation. The fraction with correct larger $P_{z,\nu}$ solutions is fitted to obtain the

P'_ν function (figure 5.7). For events with a small difference $\Delta P_{z,\nu}$ the probability for the larger $P_{z,\nu}$ solution to be the correct one is about 50%. This is expected since both solutions are very similar for these events. But for events with larger differences $\Delta P_{z,\nu}$ this probability decreases. Since Ψ should reach small values for the correct hypothesis we define $P_\nu = 1 - P'_\nu$ for hypotheses with the larger $P_{z,\nu}$ solution and $P_\nu = P'_\nu$ for hypotheses with the smaller $P_{z,\nu}$ solution.

Use of the *b*-tag information: P_b

So far the b-tag information was not used. To prefer hypotheses, in which a b-tagged jet is assigned to a b quark, we use the b-tag value b_{prob} from the track probability algorithm see section 4.2. Figure 5.8a shows the b_{prob} distributions for jets containing b quarks (black line) and for jets without b quarks (red line), while figure 5.8b depicts the b_{prob} distribution of all selected b jets on a logarithmic scale. Since we defined b jets as jets with $b_{\text{prob}} > 0.4$ this cut is reflected by this distribution. Jets with larger b_{prob} value are more likely b jets than jets with smaller b_{prob} value. We transform the b_{ptob} distribution to the range from zero to one, and define $P_b = 1 - \frac{2}{7} \cdot b_{\text{prob}}$. For hypotheses, in which none of the jets assigned to the two b quarks is b-tagged, P_b is set to 1. Hypotheses with one b jet assigned to one of the b quarks have a P_b value of $(1 - \frac{2}{7} \cdot b_{\text{prob}})$. In the case of two b jets assigned correctly to both b quarks P_b becomes $(1 - \frac{2}{7} \cdot b_{\text{prob,j1}}) \cdot (1 - \frac{2}{7} \cdot b_{\text{prob,j2}})$.



(a) (b)

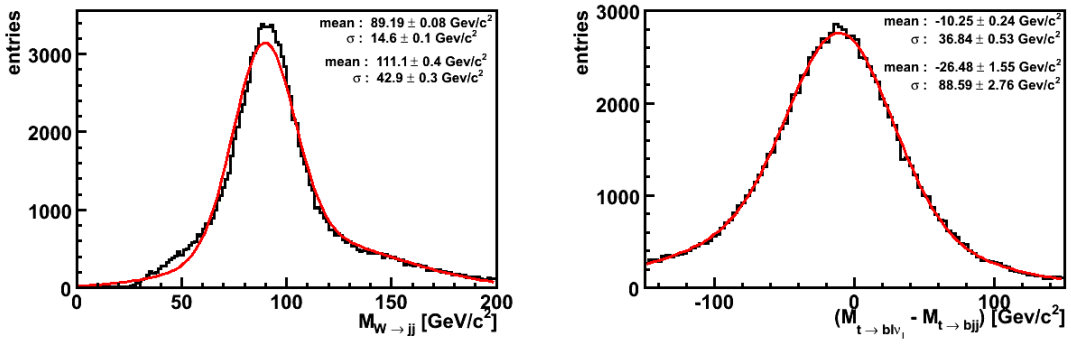
Figure 5.8: b_{prob} distributions (a) for jets containing b quarks black line and for jets without b quarks red line and b_{prob} distribution of all selected b jets (b).

Mass and Energy constraints: χ^2

The χ^2 function is defined via:

$$\chi^2 = \frac{(M_{W\to jj} - \overline{M}_{W\to jj})^2}{\sigma^2_{M_{W\to jj}}} + \frac{(\Delta M_t - \overline{\Delta M_t})^2}{\sigma^2_{\Delta M_t}} + \frac{(P_{\text{energy}} - \overline{P}_{\text{energy}})^2}{\sigma^2_{P_{\text{energy}}}} \quad . \tag{5.7}$$

The first term represents the constraint on the mass of the hadronically decaying W boson $M_{W\to jj}$. It should be equal to the mean value $\overline{M}_{W\to jj}$ of the $M_{W\to jj}$ distribution within the resolution $\sigma_{M_{W\to jj}}$. In the second term ΔM_t denotes the difference between the reconstructed mass of the semileptonically decaying top quark $M_{t\to b\ell\nu_\ell}$ and the mass of the hadronically decaying top quark $M_{t\to bjj}$. This difference and the mean value $\overline{\Delta M_t}$ are expected to be zero since both tops are identical particles. $\sigma_{\Delta M_t}$ is the width of the ΔM_t distribution. Each mean value $\overline{M}_{W\to jj}$, $\overline{\Delta M_t}$ and width $\sigma_{M_{W\to jj}}$, $\sigma_{\Delta M_t}$ is obtained from a double Gaussian fit to the corresponding mass distribution at the best possible event interpretation. The values are then obtained from the Gaussian with the smaller width. These distributions are presented in figure 5.9.



(a) (b)

Figure 5.9: A double Gaussian fit is performed to both distributions. Due to the jet corrections the fitted values for the reconstructed mass $M_{t\to bjj}$ of the hadronically decaying W boson is larger than expected (a) and the difference $(M_{t\to b\ell\nu_\ell} - M_{t\to bjj})$ between the masses of the leptonically and hadronically decaying top quarks (b) is not zero.

$$\overline{M}_{W\to jj} = 89.2\ \text{GeV}/c^2 \qquad \overline{\Delta M_t} = -10.3\ \text{GeV}/c^2$$
$$\sigma_{M_{W\to jj}} = 14.6\ \text{GeV}/c^2 \qquad \sigma_{\Delta M_t} = 36.8\ \text{GeV}/c^2$$

Both distributions and therefore the fitted mean values differ from the expectation of 80.4 GeV/c^2 and 0, respectively. These are consequences of the jet corrections. As mentioned in section 4.2 the corrected jets have more energy as they actually should. Therefore, the reconstructed mass of the hadronically decaying W boson is larger. That applies also to the masses of the top quarks. The hadronically decaying top quark is reconstructed from three jets, therefore the influence of the jet corrections on it is larger than on the semileptonically decaying top quark where only one jet enters the reconstruction. This leads to a negative mean for the difference of the top quark masses. The asymmetric shape of the ΔM_t distribution is caused by

the different detector performance in the energy resolution between jets and leptons.

The last term in equation 5.7 considers the constraint on the transverse energy of the event. Here P_{energy} is the sum of the transverse energies of the two top quarks divided by the total transverse energy of the event including missing transverse energy. $\overline{P}_{\text{energy}}$ is the mean value of the P_{energy} distribution.

$$P_{\text{energy}} = \frac{\sqrt{p^2_{\text{T,t}\to\text{b}\ell\nu_\ell} + M^2_{t\to b\ell\nu_\ell}} + \sqrt{p^2_{\text{T,t}\to\text{bjj}} + M^2_{t\to bjj}}}{\sum_{\text{jets}} p_{\text{T,jet}} + \not\!\!E_T + E_{\text{T},\ell}} \tag{5.8}$$

The reconstructed transverse momenta of the semileptonically and hadronically decaying top quark are represented by $p_{\text{T,t}\to\text{b}\ell\nu_\ell}$ and $p_{\text{T,t}\to\text{bjj}}$, while $M_{t\to b\ell\nu_\ell}$ and $M_{t\to bjj}$ represent the top quark masses. $p_{\text{T,jet}}$ denotes the transverse momentum of a jet, $\not\!\!E_T$ the measured missing transverse energy and $E_{\text{T},\ell}$ is the transverse energy of the charged lepton. For this distribution the fit of a double Gaussian returns a mean value of $\overline{P}_{\text{energy}} = 1.08$ for the Gaussian with the smaller width (figure 5.10). The fitted values for the width is $\sigma_{P_{\text{energy}}} = 0.11$.

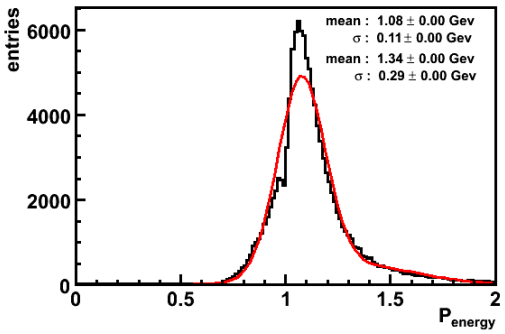


Figure 5.10: A double Gaussian is fitted to the P_{energy} distribution to determine the input parameters for the χ^2 function.

5.4 Performance of our Selection

The selection method presented in the last section was applied to the TopRex $t\bar{t}$ Monte Carlo sample. We now check the quality of the selection method by comparing the selected hypothesis with the best possible one. Table 5.1 presents the fraction of selected hypotheses which are also best possible as well as the fractions of selected hypotheses within a certain distance $\sum \Delta R$ from the Monte Carlo truth. In about 22% the selected hypotheses is exactly the best possible one and in about 76% of all cases its distance to the MC truth is below 6.

	selected $t\bar{t}$ hypothesis [%]
best possible	22.3
$\sum \Delta R < 2$	30.6
$\sum \Delta R < 4$	57.2
$\sum \Delta R < 6$	75.6

Table 5.1: Quality of the hypothesis selection. The fraction of selected hypotheses which are also best possible as well as the fractions of selected hypotheses within a certain distance $\sum \Delta R$ from the Monte Carlo truth.

We now present some plots for different quantities to compare the best possible, the selected and all remaining not selected hypotheses. The remaining hypotheses, referred to as all others, are weighted by the number of event hypotheses subtracted by one (the selected hypothesis). We exclude the selected hypothesis in order to compare it with all others, i.e. the best possible hypothesis is among the all other hypotheses in the cases where the selected hypothesis is not the best possible one.

Figure 5.11 shows the $\sum \Delta R$ values for all three categories. The distance of the best possible hypothesis to the MC truth is mostly about 1 but it can also reach values up to 5. The selected hypothesis have a large tale up to 11. The all others distribution is much broader, since it includes hypotheses with completely wrong combinations of jets and leptons as well as hypotheses with partial correct assignments in which for example only one top quark is correctly reconstructed.

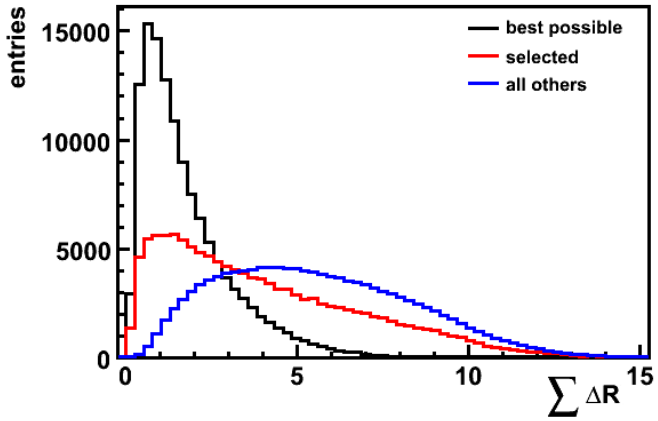


Figure 5.11: Performance of the hypothesis selection. The $\sum \Delta R$ distributions for the best possible hypothesis, the selected hypothesis, and the remaining not selected "all others" hypotheses are shown.

Before comparing the distributions for some kinematic quantities the performance of the b jet assignment is presented in table 5.2. For the assignment a

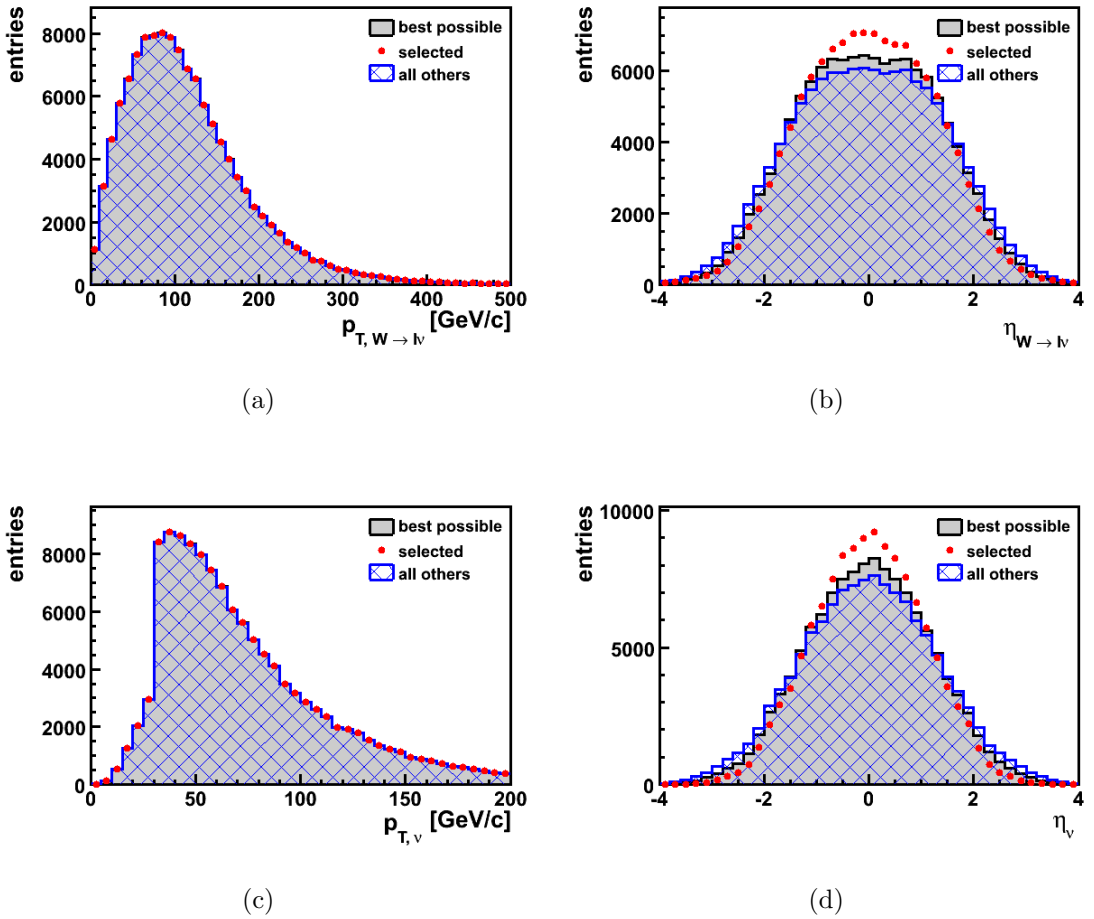
matching in R was performed to find out whether the b tagged jets are assigned to the b quarks or not. The matching criterion is fulfilled if the distance ΔR between the b tagged jet and the generated b quark is lower than 0.5 (the size of the used jet cone). As expected the assignment of b jets in the selected hypothesis is not as good as in the best possible hypothesis. In about 7% of the cases the b jets are missmatched, i.e. the b jet from the semileptonically decaying top quark was assigned to the hadronically decaying top quark and the b jet from the hadronically decaying top quark to the semileptonically decaying top quark. This fraction is not large. The main difference between the selected hypothesis and the all others hypotheses is the correct assignment of two b jets, it is also the main difference between the selected and best possible hypothesis.

	2 matched b jets [%]	1 matched b jet [%]	0 matched b jets [%]	mismatched b jets [%]	b jets not matchable [%]
best possible	63.4	25.2	5.2	1.6	5.6
selected	27.3	30.6	29.2	7.3	5.6
all others	4.2	30.5	52.0	7.7	5.6

Table 5.2: Performance of the b jet assignment in the different hypotheses best possible, selected, and all others.

Figure 5.12 shows the transverse momentum and pseudorapidity distributions of the leptonically decaying W boson and the neutrino. For the W boson and the neutrino there are no differences in the distributions of the transverse momentum between the best possible, selected, and all others. Since in the case of the neutrino the two differences $P_{z,\nu}$ do not influence the transverse momentum. The W boson is obtained by adding four-momenta of the charged lepton and the neutrino. Hence the neutrino and the charged lepton are the same for all hypotheses, the transverse momentum of the W is also the same for all hypotheses. The cut on the missing transverse energy ($\not\!\!E_T > 30$ GeV), represented in figure 5.12c as transverse momentum of the neutrino, is washed out due to the corrections on the $P_{x,\nu}$ and $P_{y,\nu}$ components mentioned in section 5.1. In case of the pseudorapidity of the neutrino a bias of the distribution obtained from the selected hypotheses towards smaller absolute values of η can be seen. This effect is due to the preference of the neutrino solution with the smaller magnitude of $P_{z,\nu}$ and therefore also occurs in the η-distribution of the leptonically decaying W boson.

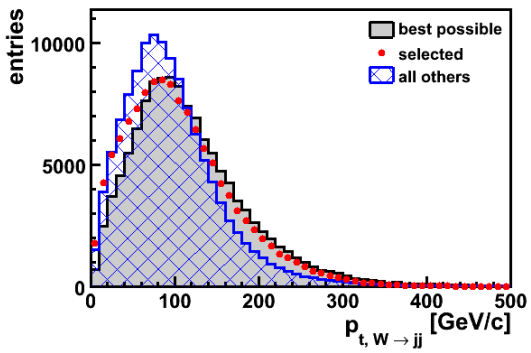
Figure 5.13 presents the transverse momentum, pseudorapidity, and the reconstructed mass of the hadronically decaying W boson. In general, the distributions obtained from the selected hypotheses follow well the distributions of the best possible hypotheses. In case of the W boson mass the shape of the distribution for all other hypotheses is substantially broader than that of the best possible hypothesis, while the shape of the selected hypothesis is narrower. This effect originates from the mass constraint.



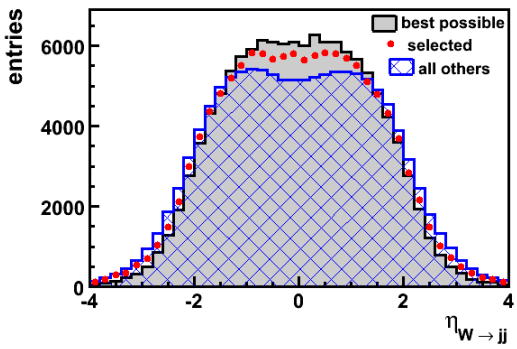
(a) (b)

(c) (d)

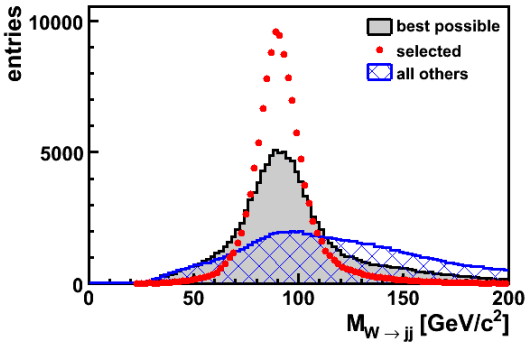
Figure 5.12: Transverse momentum (a) and pseudorapidity (b) distributions of the leptonically decaying W boson as well as the transverse momentum (c) and pseudorapidity (d) distributions of the neutrino. The best possible (grey), selected (red), and all others hypotheses are explained in the text.



(a)

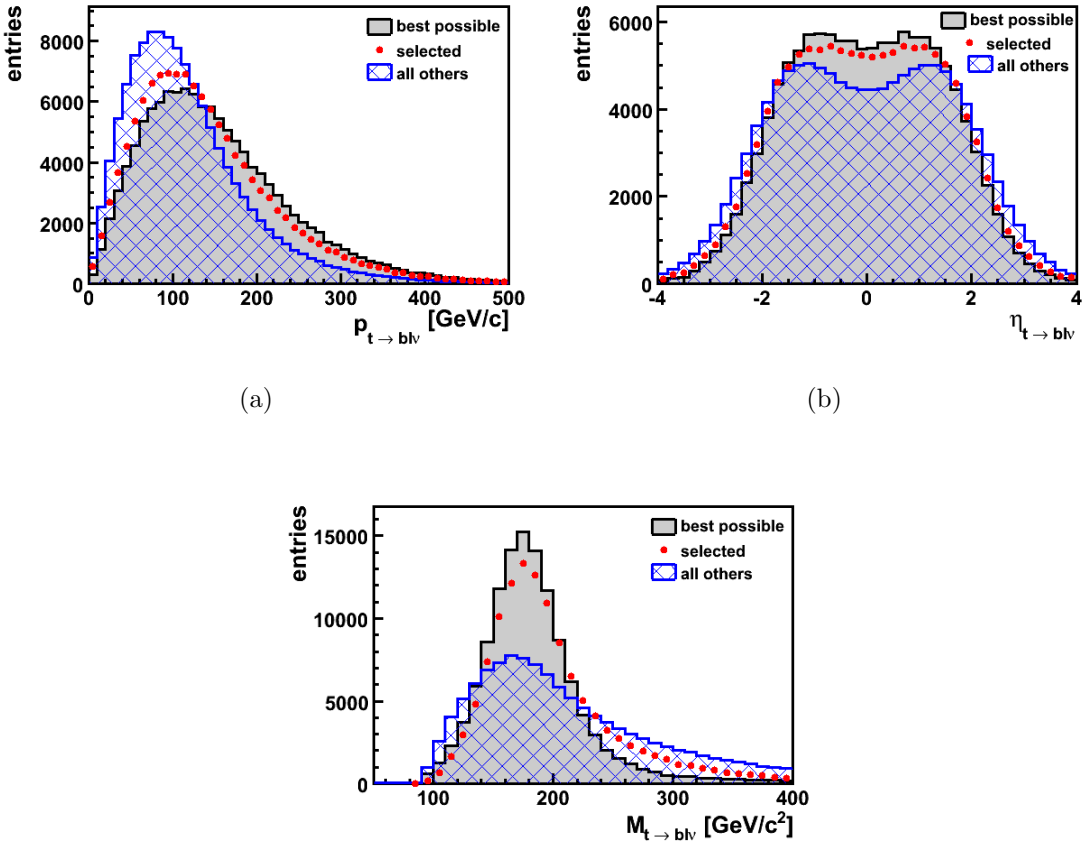


(b)



(c)

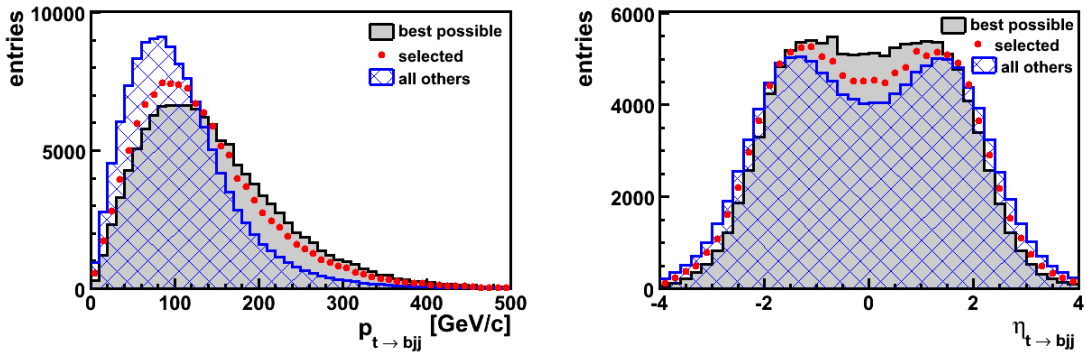
Figure 5.13: Transverse momentum (a), pseudorapidity (b), and mass distributions of the hadronically decaying W boson. The best possible (grey), selected (red), and all others hypotheses are explained in the text.



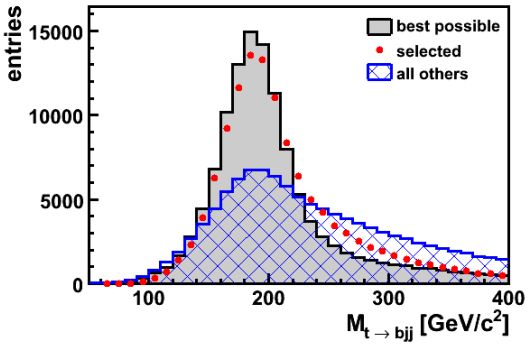
(a) (b)

(c)

Figure 5.14: Transverse momentum (a), pseudorapidity (b), and mass distributions of the leptonically decaying top quark. The best possible (grey), selected (red), and all others hypotheses are explained in the text.



(a) (b)



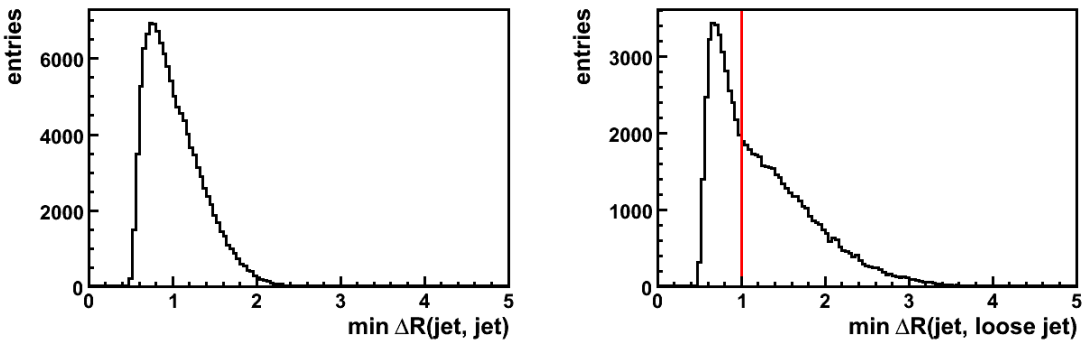
(c)

Figure 5.15: Transverse momentum (a), pseudorapidity (b), and mass distributions of the hadronically decaying top quark. The best possible (grey), selected (red), and all others hypotheses are explained in the text.

The same three distributions are shown for the leptonically decaying top quark and the hadronically decaying top quark in figure 5.14 and figure 5.15, respectively. Again the shape of the best possible hypothesis is well approximated by the shape of the selected hypothesis. The largest difference between the best possible hypothesis, the selected hypothesis and the all other hypotheses is in the distributions of the masses. The shape of the all others distribution is again much broader the shape of the best possible and selected hypotheses.

5.5 Full Reconstruction including final state radiation

A further attempt to improve the reconstruction method is done by using additionally jets, called loose jets, under the assumption that they originate from final state radiation. Jets with transverse momentum within a range of 20 – 30 GeV/c are considered. Figure 5.16a provides the smallest distance in R between two jets, while in figure 5.16b the smallest distance between a jet and the loose jet, with the lowest transverse momentum, is shown. Due to the used jet cone of $R = 0.5$ the smallest distances between two jets are larger than 0.5. From figure 5.16b can be seen that in many events this distance is smaller than one. The reconstruction was modified to use loose jets with a distance of $R < 1$ to one of the jets. We illustrate the modified reconstruction by using an event with four jets, two neutrino solutions, and one appropriate loose jet. As usual the 24 possible hypotheses are reconstructed in the first step. Then in the second step the loose jet is added to the matching jet and 24 new hypotheses are reconstructed. The best possible hypotheses and the selected one are then obtained from all 48 hypotheses by applying the $\sum \Delta R$ and Ψ criteria.



(a) (b)

Figure 5.16: Smallest distance in R between two jets (a) and smallest distance in R between a jet and the loose jet with the lowest transverse momentum (b).

This modified reconstruction was tested with up to two matchable loose jets. Table 5.3 summarizes the fraction of events where it was possible to reconstruct such new hypotheses as well as the cases where one of the new hypotheses was selected by the Ψ criterion.

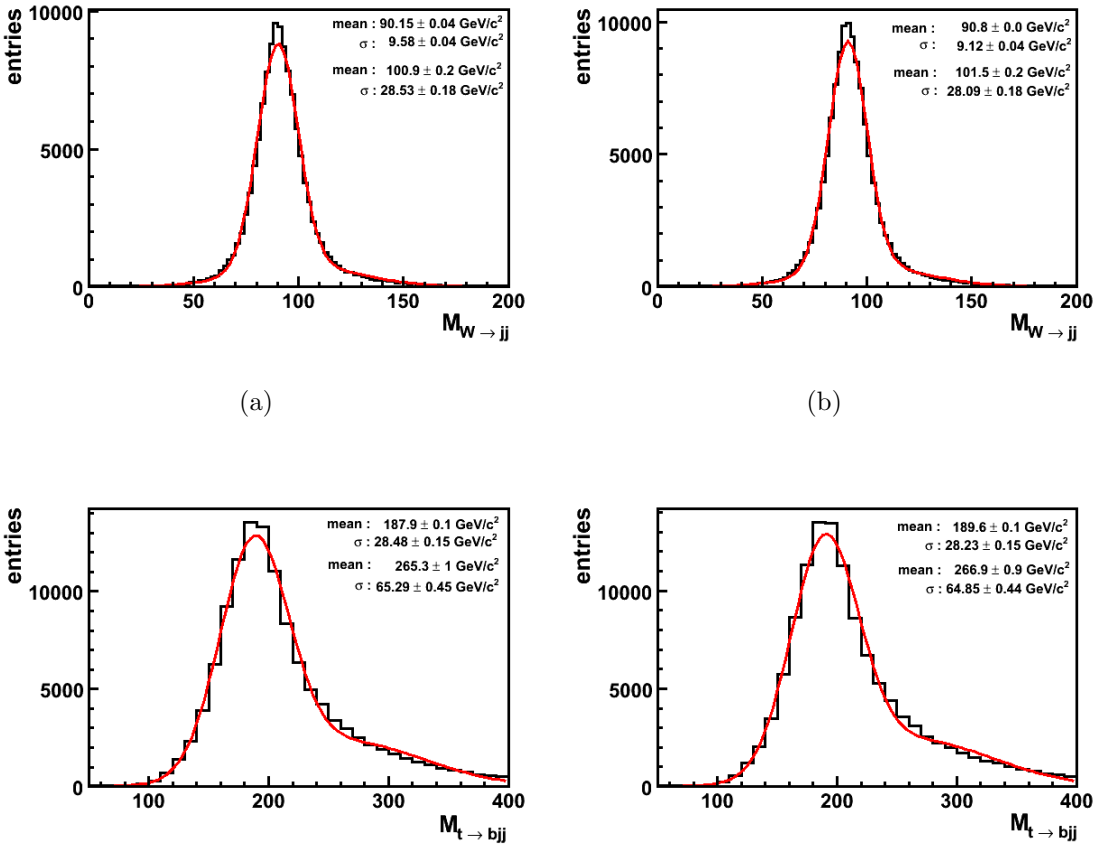
total events 124465	normal hypothesis [%]	1 loose jet hypothesis [%]	2 loose jets hypothesis [%]
reconstructed	100.0	30.3	5
selected	87.7	11.2	1.1

Table 5.3: Fraction of events where the selected hypothesis contains loose jets.

In about 11% the selected hypotheses contained one matched loose jet and the fraction of events where the selected hypothesis contained two loose jets only 1%. The effect on the reconstructed masses of the hadronically decaying W boson and top quark is illustrated in figure 5.17. On the left-hand side both masses for the normal reconstruction are shown. The distributions on the right-hand side include the new hypotheses with additional loose jets. All distributions are fitted with a double Gaussian. The mean of the Gaussian with the narrower width is taken to be the reconstructed mass. The values for the mean and the width of the W boson and top quark masses are:

without loose jets		with loose jets	
$\overline{M_{W\to jj}}$	$= 90.15 \pm 0.04$ GeV/c^2	$\overline{M_{W\to jj}}$	$= 90.8 \pm 0.0$ GeV/c^2
$\sigma_{M_{W\to jj}}$	$= 9.58 \pm 0.04$ GeV/c^2	$\sigma_{M_{W\to jj}}$	$= 9.12 \pm 0.04$ GeV/c^2
$\overline{M_{t\to bjj}}$	$= 187.9 \pm 0.1$ GeV/c^2	$\overline{M_{t\to bjj}}$	$= 189.6 \pm 0.1$ GeV/c^2
$\sigma_{M_{t\to bjj}}$	$= 28.48 \pm 0.15$ GeV/c^2	$\sigma_{M_{t\to bjj}}$	$= 28.23 \pm 0.15$ GeV/c^2

The W boson and top quark masses, reconstructed with the modified method, are only slightly heavier and their distributions are also slightly narrower. The additional loose jets with transverse momenta from 20 GeV/c^2 to 30 GeV/c^2 do not deteriorate the reconstruction but result in a slight improvement. Nevertheless this modified method was not used. One reason for this are the jet corrections which cause already too large masses for the reconstructed W bosons and top quarks, using additionally jets would increase them further. The second reason is that the loose jet is mostly used to reconstruct the hadronically decaying top quark, since the probability that a loose jet is closer to one of the three jets from the hadronically decaying top quark is 3:1. Therefore, the influence of loose jets on the helicity angle θ^*, which is measured between the products of the semileptonically decaying top quark, is not significant.



(a) (b)

(c) (d)

Figure 5.17: Comparison between the usual reconstruction (on the left) and the reconstruction which includes final state radiation (on the right). The reconstructed mass distributions of hadronically decaying W boson (first row) and top quark (last row) are shown together with a double Gaussian fit.

Chapter 6

Extraction of the Helicity Fractions

This chapter describes the measurement of the helicity fractions of the W boson, for which a fit to the reconstructed $\cos\theta^*_{\text{rec}}$ distribution is utilized. In the first section the impacts of the event selection and reconstruction on the $\cos\theta^*$ distribution are discussed as well as their consequences for the fit. The following sections describe the method used to fit the $\cos\theta^*_{\text{rec}}$ distribution and the obtained fit results. Finally, the reconstructed distribution is unfolded in order to compare it with the theoretical prediction.

6.1 Impacts of the reconstruction on the $\cos\theta^*$ distribution

As described in section 1.3 the sensitive variable to measure the helicity fractions is the cosine of the decay angle θ^*. This is the angle between the momentum of the charged lepton in the W boson rest frame and the momentum of the W boson in the top quark rest frame. Figure 6.1 provides a schematic view of that angle.

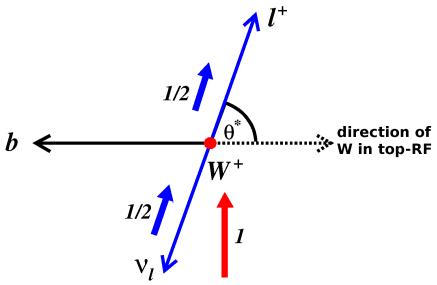


Figure 6.1: Schematic view of the decay angle θ^* in the rest frame of the W boson. The dotted black arrow represents the direction of the W boson in the rest frame (RF) of the top quark.

Since we apply several selection cuts to select candidate events, it is necessary

to check the influence of this selection on the $\cos\theta^*$ distribution. The detector, simulated by the CMS Software, and the reconstruction of the selected events have also an influence which must be considered. Figure 6.2 shows the changes of the $\cos\theta^*$ distribution. The black curve is the generated $\cos\theta^*$ distribution with the Standard Model helicity fractions $F_- = 0.3$, $F_0 = 0.7$, and $F_+ = 0$. This distribution was obtained by using the four vectors of the particles from the Monte Carlo true values. The blue distribution is also obtained from the MC true values but after the selection, i.e. the difference in the shape is caused by the selection cuts. The red distribution is computed using the reconstructed four vectors. It illustrates the influence of the simulated detector effects and the reconstruction on the $\cos\theta^*$ distribution. The reasons for these changes are now discussed in detail:

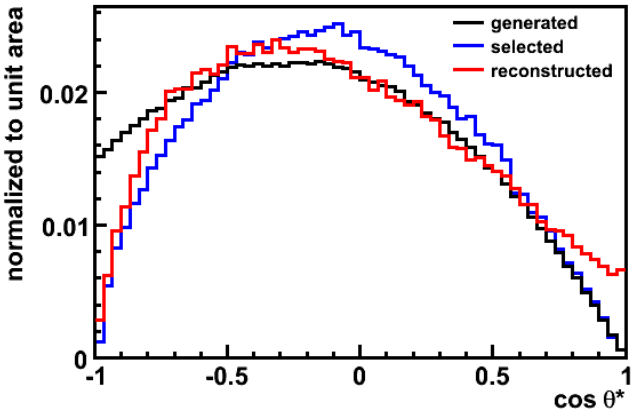


Figure 6.2: Impact of event selection and reconstruction on the $\cos\theta^*$ distribution obtained from the TopRex MC sample. The black line represents the distribution before applying any selection cuts, the distribution for the selected events is shown in blue. Both distributions are obtained from the four-vectors of the generated particles. The red line represents the distribution obtained from fully reconstructed events (for each event the hypothesis with the smallest value of Ψ is chosen).

Impact of the Event Selection

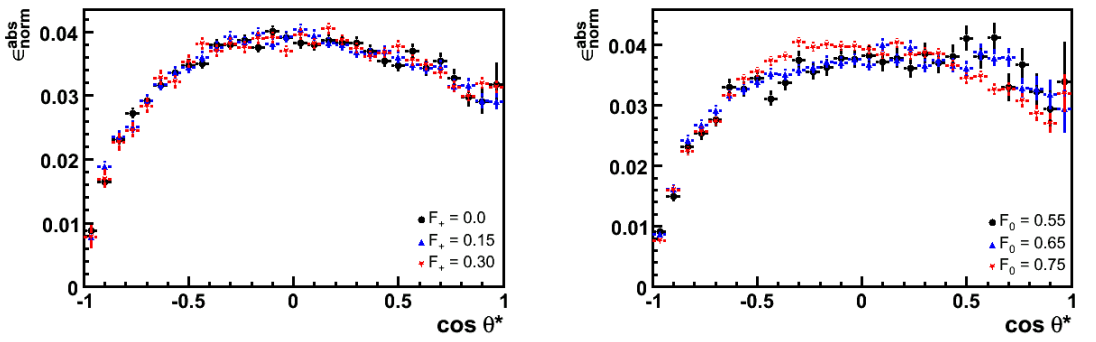
The difference in the shape of the generated and selected $\cos\theta^*$ distribution is mainly caused by the isolation cut on the charged lepton. Events where the charged lepton flies in the same direction as the b quark (see figure 6.1), which corresponds to $\cos\theta^*$ values close to -1, are therefore much stronger rejected by the isolation requirement than events with different values for $\cos\theta^*$. The second reason for the strong rejection of events with values for $\cos\theta^*$ close to -1 is due to the softer p_{T} spectrum of leptons being emitted in backward direction with respect to the W boson direction of motion. Thus, these events are more likely to fail the lepton p_{T} cut in the event selection than events with larger values for $\cos\theta^*$. While events with higher values of $\cos\theta^*$ are more likely to pass the lepton p_{T} cut since the lepton is emitted in

the direction of motion of the W boson. For events with values for $\cos\theta^*$ close to $+1$ $\not\!\!E_T$ decreases with increasing values for $\cos\theta^*$. In these events the neutrino is emitted in backward direction with respect to the W boson direction and therefore its transverse energy is smaller than in the opposite case. Thus, for large positive values of $\cos\theta^*$ the events are more likely to pass the lepton p_{T} requirement, but also more likely to fail the $\not\!\!E_T$ cut.

To account for these effects we define the absolute efficiency for each bin of the $\cos\theta^*$ distribution as:

$$\epsilon_i^{\mathrm{abs}} = N_i^{\mathrm{sel}}/N_i \quad , \tag{6.1}$$

where N_i^{sel} denotes the number of selected events in bin i and N_i the number of generated events in the same bin. The selected $\cos\theta^*$ distribution is therefore obtained by multiplying the generated distribution with the absolute efficiency in each bin $N_i^{\mathrm{sel}} = N_i \cdot \epsilon_i^{\mathrm{abs}}$. This definition is possible since the selection only reduces the number of the events. We assume, that the shape of the distribution of the absolute efficiency as a function of $\cos\theta^*$ is independent on the helicity of the W bosons. This assumption is checked using the MadEvent MC samples with different W boson helicity fractions. Figure 6.3 illustrates the normalized efficiencies ϵ_{norm}^{abs} which allow direct comparison between the shapes of efficiency distributions for different settings of helicity fractions. On the left (figure 6.3a) the efficiencies for three different fractions F_+ of right-handed W bosons are shown. The fraction F_0 of longitudinal polarized W bosons is set to the SM value of 0.7. On the right (figure 6.3b) one can see the normalized efficiencies obtained from MC samples with different F_0 fractions while F_+ is set to zero. The variation of F_+ and F_0 respectively is compensated by the variation of the F_- fraction.



(a) (b)

Figure 6.3: Normalized efficiencies obtained from MC samples with different F_+ fractions, while the fraction F_0 is set to a constant value of 0.7 (a), and normalized efficiencies from samples with different F_0 fractions but constant F_+ fractions, namely zero (b). The variation of F_+ and F_0 respectively is compensated by the variation of the F_- fraction.

The efficiencies for the three different F_+ fractions are in good agreement, this is also the case for the other MC samples mentioned in section 4.1. We do not plot the efficiencies for all seven samples since then it is not possible to recognize the difference between the curves. The efficiencies of the different F_0 fractions shows a few outlier, while most of the values are within the uncertainties.

Impact of the Event Reconstruction

The event reconstruction also has an impact on the $\cos\theta^*$ distribution. Several effects are the reason for the deviation of the reconstructed $\cos\theta^*_{\text{rec}}$ distribution compared to the distribution after the event selection. These effects can be classified in three categories.

- Physical Effects
  The final-state quarks from the top-quark decay hadronize into colorless particles. In the hadronization process color flux tubes are stretched between the quarks as they move apart. The potential energy stored in these tubes or strings increases linearly with the distance between the quarks and the strings may break producing a new quark-antiquark pair. The break of the strings causes a variation in the direction of the created colorless particles which are measured as jets. Thus, the direction of the measured jet does not give exactly the direction of the final-state quarks in all events. Therefore, in some events it is impossible to reconstruct correctly the four vectors of the two top quarks from the detected objects. The final-state quarks may also radiate gluons and therefore change their direction and decrease their energy.

- Measurement Effects
  In the case of such soft gluon radiation the gluons do not fulfill all the jet requirements and are not selected as jets. As a consequence, the energy of the jet radiating a gluon does not reflect the energy of the original quark. In other cases not all selected jets originate from the top quark decay, i.e. additional jets from hard gluon radiation are selected. Therefore, it is impossible to reconstruct the $t\bar{t}$ pair correctly since not all components are available. A further contribution arises from the finite resolution of the energy and direction measurement of jets and particles itself. Due to the mentioned effects the reconstruction is not always correctly as can be seen in figure 5.11. Even the distribution of best possible event interpretations take large $\sum \Delta R$ values.

- Reconstruction Effects
  Several ambiguities occur during the reconstruction of $t\bar{t}$ event candidates which lead to numerous event interpretations. The fact that only in 22% of all events the best possible event interpretation is chosen contributes to the deviation of the reconstructed $\cos\theta^*$ distribution, hence in several cases the wrong combination of jets and leptons is chosen to reconstruct the top quarks and W bosons.

All mentioned effects cause a "migration" of the $\cos\theta^*$ values from the selected distribution to the reconstruct distribution, since all selected events are reconstructed. This impact is taken into account by the migration matrix S. The matrix element $S(i,k)$ gives the probability for an event with a true value for $\cos\theta^*$ in bin i to occur after the reconstruction in bin k of the $\cos\theta^*_{\mathrm{rec}}$ distribution. To take into account that all selected events in a certain bin i have to occur somewhere in the reconstructed distribution, the matrix is defined in such a way, that $\sum_k S(i,k) = 1$ holds for all bins i. The expected number of events in bin k after the reconstruction is thus given by the sum over the numbers of selected events in every bin multiplied with the corresponding probability to be reconstructed in bin k:

$$N_k^{\mathrm{rec}} = \sum_i N_i^{\mathrm{sel}} \cdot S(i,k) \quad . \tag{6.2}$$

Figure 6.4 gives an illustration of the migration matrix obtained from the simulated TopRex Monte Carlo sample. The generated distribution $\cos\theta^*$ is plotted against the reconstructed $\cos\theta^*_{\mathrm{rec}}$ distribution.

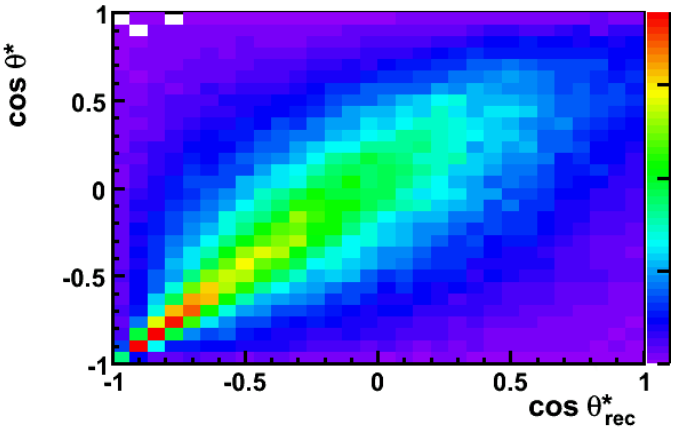
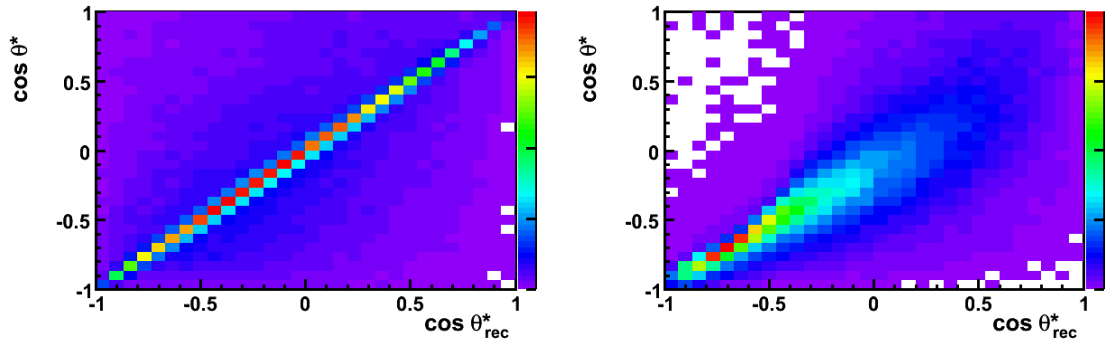


Figure 6.4: Migration matrix: The generated $\cos\theta^*$ distribution is plotted versus the reconstructed $\cos\theta^*_{\mathrm{rec}}$ distribution.

There are many entries beside the diagonal. To determine the reason of these entries we use the generated particles. The helicity angle θ^* is computed using the reconstructed four-vectors of the charged lepton, the leptonically decaying W boson, and semileptonically decaying top quark. As described in section 5.1 the W boson is reconstructed by adding the four-vectors of the neutrino and the charged lepton, while the four-vector of the top quark is obtained by adding the four-vectors of the W boson and the one of a b quark. We compute the $\cos\theta^*_{\mathrm{rec}}$ distribution once, using the four-vector of the generated neutrino and the reconstructed four-vectors of the charged lepton and the b quark, and a second time using the four-vector of the re-



(a) (b)

Figure 6.5: Influence on the migration matrix: (a) the $\cos\theta^*_{\text{rec}}$ distribution is calculated using the four-vector of the generated neutrino and the reconstructed four-vectors of the charged lepton and the b quark, while in (b) the four-vector of the reconstructed neutrino and the generated four-vectors of the charged lepton and the b quark are used.

constructed neutrino and the generated four-vectors of the charged lepton and the b quark. The migration matrix for the first case is shown in figure 6.5a. It is nearly diagonal, while the matrix computed with the reconstructed neutrino four-vector, shown in figure 6.5b, contains many entries with low $\cos\theta^*$ and high $\cos\theta^*_{\text{rec}}$ values. Hence the reconstructed neutrino, i.e. the measurement of $\not\!\!E_T$, causes the largest influence on the migration matrix.

We assume that like the efficiency also the migration matrix does not depend on the helicity fractions F_0 and F_+. This assumption is checked, using the MC sample with different F_0 and F_+ fractions. Figures 6.6 and 6.7 depict three out of 30 slices through the migration matrix. The slices on the left-hand side labeled with $S(5,k)$ illustrate the probability of the entries in the fivth bin of $\cos\theta^*$ to occur in one of the 30 bins of the $\cos\theta^*_{\text{rec}}$ distribution. The fivth bin corresponds to $\cos\theta^*$ values in the range of $-\frac{11}{15}$ to $-\frac{10}{15}$. The slices in the middle $S(15,k)$ and on the right-hand side $S(25,k)$ correspond to $-\frac{1}{15} < \cos\theta^* \leq 0$ and to $\frac{9}{15} < \cos\theta^* \leq \frac{10}{15}$, respectively. Figure 6.6 shows the three mentioned slices for three different F_+ fractions and figure 6.7 for three different F_0 fractions. The matrix elements show no deviation within the statistical uncertainties between the different helicity fractions.

The selection and reconstruction effects have to be taken into account in order to extract the helicity fractions from the reconstructed $\cos\theta^*_{\text{rec}}$ distribution. Hence, the efficiency and the migration matrix enter the fit templates, which are used in a binned likelihood fit. The calculation of these templates is presented in the next section.

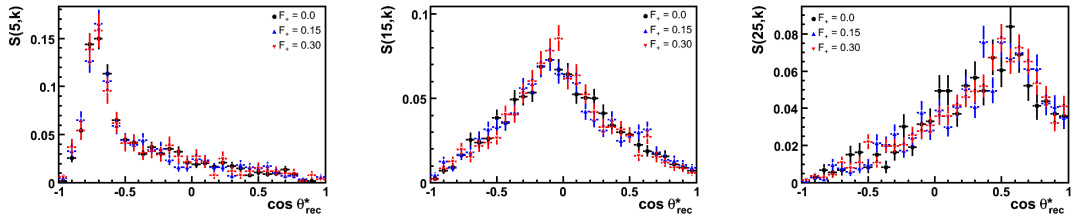


Figure 6.6: Some migration matrix elements $S(i,k)$ for different helecity fractions $F_+ = 0$ (black), $F_+ = 0.15$ (blue), and $F_+ = 0.30$ (red) of right-handed W bosons.

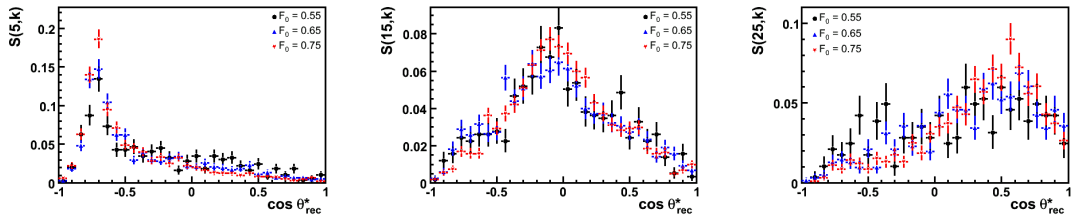


Figure 6.7: Some migration matrix elements $S(i,k)$ for different helecity fractions $F_0 = 0.55$ (black), $F_0 = 0.65$ (blue), and $F_0 = 0.75$ (red) of longitudinally polarized W bosons.

6.2 Calculation of the normalized Signal Templates

The fit templates $\mu_i^{\text{sig,obs}}(F_0, F_+)$ depend on the values of F_0 and F_+. They result in the number of expected events in each bin i for given (F_0, F_+) values. The binned likelihood fit varies both values to approximate the input distribution, and returns the pair of (F_0, F_+) values which leads to the best approximation for all bins. In this section we calculate the normalized signal templates $\hat{\mu}_i^{\text{sig,obs}}(F_0, F_+)$. The number of expected signal events in a certain bin $\mu_i^{\text{sig,obs}}(F_0, F_+)$ is then obtained by multiplying the normalized signal template with the total number of expected signal events: $\mu_i^{\text{sig,obs}}(F_0, F_+) = \hat{\mu}_i^{\text{sig,obs}}(F_0, F_+) \cdot N_{\text{sig}}$.

The starting point of the calculation of the normalized signal template $\hat{\mu}_i^{\text{sig,obs}}$ is the contribution of $t\bar{t}$ signal events $\hat{\mu}_i^{\text{sig}}$ in the i^{th} bin of $\cos\theta^*$ before applying any cuts. This distribution is obtained by integrating equation 1.10 for each bin separately and using the relation $F_0 + F_- + F_+ = 1$ to eliminate the fraction of left-handed W bosons:

$$\hat{\mu}_i^{\text{sig}}(F_0, F_+) \quad = \quad (1 - F_0 - F_+) \cdot f_i^- + F_0 \cdot f_i^0 + F_+ \cdot f_i^+ \tag{6.3}$$

where f_i^-, f_i^+ and f_i^0 are the fractions of events contained in bin i in case of left-handed, right-handed, and longitudinally polarized W bosons:

$$f_i^0 = \int_{a_i}^{b_i} \frac{3}{4}(1-\cos^2\theta^*)d\cos\theta^*, \tag{6.4}$$

$$f_i^- = \int_{a_i}^{b_i} \frac{3}{8}(1-\cos\theta^*)^2 d\cos\theta^*, \tag{6.5}$$

$$f_i^+ = \int_{a_i}^{b_i} \frac{3}{8}(1+\cos\theta^*)^2 d\cos\theta^*. \tag{6.6}$$

The fractions f^0, f^-, and f^+ are normalized to unit area, hence the relations $\sum_i^{N_{\text{bins}}} f_i^0 = 1$, $\sum_i^{N_{\text{bins}}} f_i^- = 1$, and $\sum_i^{N_{\text{bins}}} f_i^+ = 1$ hold. a_i and b_i denote the lower and upper edge of bin i and N_{bins} is the number of bins. Rewriting equation 6.3 points out the linear dependence of $\hat{\mu}_i^{\text{sig}}$ on F_0 and on F_+:

$$\begin{aligned}\hat{\mu}_i^{\text{sig}}(F_0, F_+) &= F_0 \cdot \left[f_i^0 - f_i^-\right] + F_+ \left[f_i^+ - f_i^-\right] + f_i^- \\ &= F_0 \cdot A_i + F_+ \cdot B_i + C_i \quad .\end{aligned} \tag{6.7}$$

Here the coefficients A_i, B_i and C_i of the i^{th} bin are defined as:

$$A_i = f_i^0 - f_i^- \quad , \tag{6.8}$$

$$B_i = f_i^+ - f_i^- \quad , \tag{6.9}$$

$$C_i = f_i^- \quad . \tag{6.10}$$

Now the efficiency ϵ_i^{abs} and the migration matrix $S(i,k)$ are considered to calculate the observed normalized signal templates $\hat{\mu}_i^{\text{sig,obs}}$ out of the theoretical normalized signal templates $\hat{\mu}_i^{\text{sig}}$. Since the aim of this section is to calculate normalized signal templates, only the shape of the ϵ_i^{abs} distribution matters. In other words, only the relations between the efficiencies in different bins and not the absolute values of the efficiencies in certain bins are of interest. Therefore, we can multiply ϵ_i^{abs} with an arbitrary factor that cancels later out due to the normalization. For simplicity we utilize relative efficiencies instead of absolute efficiencies. The relative efficiency ϵ_i^{rel} is defined as:

$$\epsilon_i^{\text{rel}} = \frac{N_i^{\text{sel}}/N^{\text{sel}}}{N_i/N} \tag{6.11}$$

and is obtained from ϵ_i^{abs} by dividing by the acceptance factor $\mathcal{A} = N^{\text{sel}}/N$:

$$\epsilon_i^{\text{rel}} = \frac{\epsilon_i^{\text{abs}}}{\mathcal{A}} \quad . \tag{6.12}$$

Substituting the absolute efficiency with the relative efficiency ϵ^{rel} the expected number of events in bin i after applying all selection cuts is defined as:

$$N_i^{\mathrm{sel}} = N_i \cdot \epsilon_i^{\mathrm{rel}} \cdot \mathcal{A} \quad . \tag{6.13}$$

The expected number of events in bin k after the reconstruction is given by the sum over the numbers of selected events in every bin multiplied with the corresponding migration matrix element $S(i,k)$:

$$N_k^{\mathrm{rec}} = \sum_i N_i^{\mathrm{sel}} \cdot S(i,k) \quad . \tag{6.14}$$

The normalized signal template $\hat{\mu}_k^{\mathrm{sig,obs}}$ which is defined as the fraction of the expected number of reconstructed events in bin k of the $\cos\theta_{\mathrm{rec}}$ distribution with respect to the total number of reconstructed events can now be calculated. Since every selected event is also reconstructed the total number of reconstructed events is equal to the total number of selected events.

$$\hat{\mu}_k^{\mathrm{sig,obs}}(F_0, F_+) \;=\; \frac{N_k^{\mathrm{rec}}}{N_{\mathrm{total}}^{\mathrm{rec}}} = \frac{\sum_i N_i^{\mathrm{sel}} \cdot S(i,k)}{\sum_k \sum_i N_i^{\mathrm{sel}} \cdot S(i,k)} \tag{6.15}$$

Using the expression for the number of selected events given in equation 6.13 this can be rewritten as:

$$\begin{aligned} \hat{\mu}_k^{\mathrm{sig,obs}}(F_0, F_+) &= \frac{\sum_i N_i \cdot \epsilon_i^{\mathrm{rel}} \cdot \mathcal{A} \cdot S(i,k)}{\sum_k \sum_i N_i \cdot \epsilon_i^{\mathrm{rel}} \cdot \mathcal{A} \cdot S(i,k)} \\ &= \frac{\sum_i N_i \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k)}{\sum_k \sum_i N_i \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k)} \quad . \end{aligned} \tag{6.16}$$

In the last step the decision to use relative efficiencies is vindicated, since the acceptance factor $\mathcal{A}$ occurs in the numerator as well as in the denominator and therefore cancels out. Using $N_i = N \cdot \hat{\mu}_i^{\mathrm{sig}}(F_0, F_+)$ and $\hat{\mu}_i^{\mathrm{sig}}(F_0, F_+) = F_0 \cdot A_i + F_+ \cdot B_i + C_i$ from equation 6.7 the above equation can be transformed to:

$$\hat{\mu}_k^{\mathrm{sig,obs}}(F_0, F_+) \;=\; \frac{\sum_i (F_0 \cdot A_i + F_+ \cdot B_i + C_i) \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k)}{\sum_k \sum_i (F_0 \cdot A_i + F_+ \cdot B_i + C_i) \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k)} \tag{6.17}$$

in which the total number of generated events N cancels out. Defining coefficients A_k^{obs}, B_k^{obs} and C_k^{obs} of the k^{th} bin of $\cos\theta_{\mathrm{rec}}^*$:

$$A_k^{\mathrm{obs}} \;=\; \sum_i A_i \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k) \quad , \tag{6.18}$$

$$B_k^{\mathrm{obs}} \;=\; \sum_i B_i \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k) \quad , \tag{6.19}$$

$$C_k^{\mathrm{obs}} \;=\; \sum_i C_i \cdot \epsilon_i^{\mathrm{rel}} \cdot S(i,k) \quad , \tag{6.20}$$

equation 6.17 yields:

$$\hat{\mu}_k^{\mathrm{sig,obs}}(F_0, F_+) \;=\; \frac{F_0 \cdot A_k + F_+ \cdot B_k + C_k}{\sum_k F_0 \cdot A_k + F_+ \cdot B_k + C_k} \quad . \tag{6.21}$$

For the calculation of the signal templates used later in the fit, the efficiency and migration matrix are obtained from MC samples. Since neither the efficiency nor the migration matrix depend on the W helicity fractions, for this purpose the TopRex $t\bar{t}$ sample is used. The relative efficiency and three out of 30 migration matrix elements obtained from this sample are presented in figures 6.8 and 6.9.

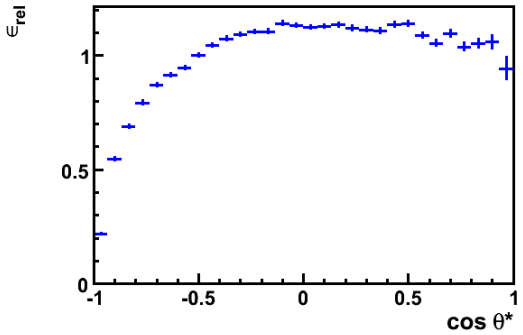


Figure 6.8: Relative efficiency as a function of $\cos\theta^*$ computed using the TopRex MC sample with SM values for F_0 and F_+.

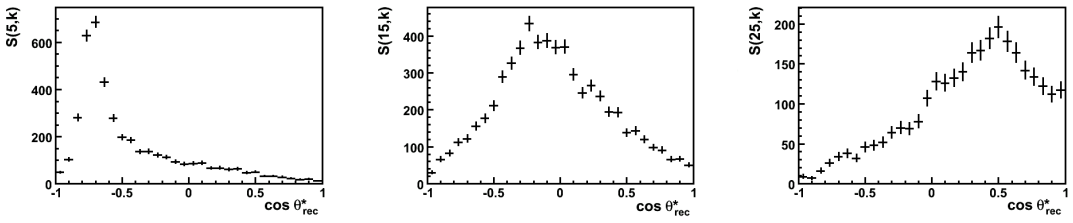


Figure 6.9: Some migration matrix elements $S(i,k)$ computed using the TopRex MC sample with SM values for F_0 and F_+.

The efficiency is determined in 30 bins and the migration matrix for 30×30 bins leading to 30 bins for the signal template. Since we use only six bins in our measurement, the number of bins is reduced after the calculation of the signal template with 30 bins by a factor of five. The decision to use six bins is motivated in the next section.

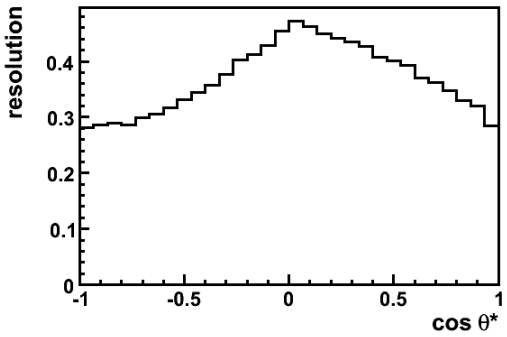


Figure 6.10: Resolution of the full reconstruction in bins of the $\cos\theta^*$ distribution. The single values represent the Gaussian widths of the $(\cos\theta^* - \cos\theta^*_{\mathrm{rec}})$ - distributions in each bin of the $\cos\theta^*$ distribution obtained from the TopRex MC sample.

6.3 Likelihood Fit and CMS Pseudo Experiment

To determine a proper value for the number of bins for the observed $\cos\theta^*$ distribution, we determine the resolution of the full reconstruction using the TopRex MC sample. For all 30 bins of the $\cos\theta^*$ distribution a Gaussian is fitted to the distribution of the difference between the $\cos\theta^*$ values obtained from generated and reconstructed four-vectors $(\cos\theta^*_{\mathrm{gen}} - \cos\theta^*_{\mathrm{rec}})$. The widths of these 30 Gaussians are shown in figure 6.10.

As can be seen in figure 6.11 it is possible to improve the resolution with an

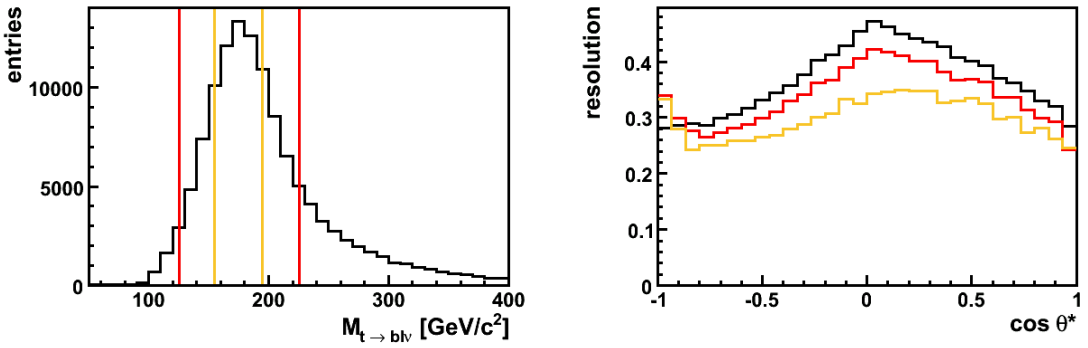


Figure 6.11: The resolution can be improved using cuts on the mass $M_{t\to W\ell\nu}$ of the leptonically decaying top quark. The cuts are indicated on the left-hand side, while the corresponding resolution is shown on the right-hand side. The black curve was obtained without cuts, while for the red and yellow curves the top mass was restricted to a range of 125-225 GeV/c^2 and 155-195 GeV/c^2, respectively.

additional cut on the reconstructed mass $M_{t\to b\ell\nu}$ of the leptonically decaying top quark. The cuts are indicated on the left-hand side, while the corresponding resolution is shown on the right-hand side. The black curve was obtained without cuts, while for the red and yellow curves the top mass was restricted to a range of 125-225 GeV/c^2 and 155-195 GeV/c^2, respectively. The deviation of the top-quark mass from the mean value in these events is mostly caused by the selection of wrong event hypothesis. The influence of this cut on the measurement and its sensitivity need to be studied in more details, since it reduces also the number of selected signal and background events. In the following we decided not to apply a cut on $M_{t\to b\ell\nu}$.

Since the resolution of the full $t\bar{t}$ reconstruction takes values between 0.2 and 0.4 we decide to use six bins for the observed $\cos\theta^*$ distribution.

We use a binned likelihood fit to extract the helicity fractions F_0 and F_+ from the reconstructed $\cos\theta^*_{\text{rec}}$ distribution. As mentioned in the previous section the fit uses the calculated signal template $\mu_i^{\text{sig,obs}}(F_0, F_+)$ and varies the (F_0, F_+) values simultaneously until a minimal deviation of the template from the fitted distribution is found. This deviation is described by the likelihood function.

$$L(F_0, F_+) = \prod_{k=1}^{N_{\text{bins}}} \frac{\mu_k^{\text{obs}}(F_0, F_+)^{n_k} \cdot e^{-\mu_k^{\text{obs}}(F_0,F_+)}}{n_k!} \quad . \tag{6.22}$$

Here, μ_k^{obs} denotes the number of events expected to be observed in bin k of the reconstructed $\cos\theta^*_{\text{rec}}$ distribution and n_k the measured number of events in the same bin. In order to get values for (F_0, F_+) we minimize the negative logarithm of the likelihood function $L(F_0, F_+)$ by varying the free parameters (F_0, F_+).

The expected number of events μ_k^{obs} in bin k is the sum of the expected number of $t\bar{t}$ signal events $\mu_k^{\text{sig,obs}}$ and the expected number of background events $\mu_k^{\text{BG,obs}}$:

$$\mu_k^{\text{obs}} = \mu_k^{\text{sig,obs}} + \mu_k^{\text{BG,obs}}. \tag{6.23}$$

The expected number of signal events in bin k is:

$$\mu_k^{\text{sig,obs}}(F_0, F_+) = N_{\text{data}} \cdot f_{\text{sig}} \cdot \hat{\mu}_k^{\text{sig,obs}}(F_0, F_+), \tag{6.24}$$

where $\hat{\mu}_k^{\text{sig,obs}}$ is the signal template for the k^{th} bin of $\cos\theta^*_{\text{rec}}$ as derived in the previous section (see equations 6.17 and 6.21). N_{data} is the number of observed events in the selected sample and $f_{\text{sig}} = 77\%$ the fraction of signal events, which is derived in section 4.2.1. Thus, the product $N_{\text{data}} \cdot f_{\text{sig}}$ gives the expected number of signal events in the selected sample.

Figure 6.12 illustrates the $\cos\theta^*$ distributions obtained from the signal and background MC samples as a stacked plot. The reconstructed $\cos\theta^*$ distributions from

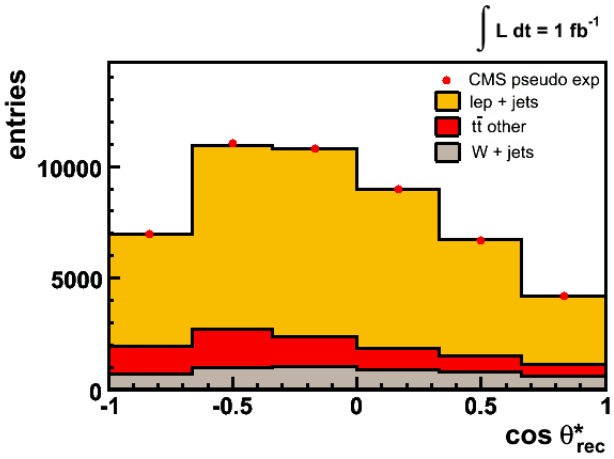


Figure 6.12: Illustration of $\cos\theta^*_{rec}$ distributions obtained from the signal and background MC samples. The red points indicate the generated CMS pseudo experiment for an integrated luminosity of $\mathcal{L}$ =1 fb^{-1}.

the $W+n$jets (with $n = 4, 5, 6$) and $W+b\bar{b}+n$jets (with $n = 2, 3$) background samples are summed up to the $W+jets$ distribution. In order to test the measurement method we simulate the outcome of the CMS experiment for an integrated luminosity of $\mathcal{L}$ =1 fb^{-1} using one pseudo experiment. The CMS pseudo experiment is generated by throwing a number of events for each sample according to a Poisson distribution with the mean of the expected background events which are presented in table 4.4. The different $\cos\theta^*$ distributions for each sample are obtained by filling histograms according the $\cos\theta^*_{\mathrm{rec}}$ distributions obtained from these samples. The the histogram representing the CMS pseudo experiment is the sum of all these histograms. This technique is described in more details in section 7.1. Figure 6.12 depicts the $\cos\theta^*_{\mathrm{rec}}$ distributions of the signal and background MC samples, scaled to the number of expected events, and the generated CMS pseudo experiment indicated by red points. Due to the large number of expected events, the deviation between the generated CMS pseudo experiment and the expected distributions is very small.

The result of a fit performed using the $\cos\theta^*$ distribution of the generated CMS pseudo experiment is presented in figure 6.13. The fitted values for the helicity fractions are:

$$\begin{aligned} F_0 &= 0.707 \pm 0.032 \\ F_+ &= -0.006 \pm 0.016 \ . \end{aligned}$$

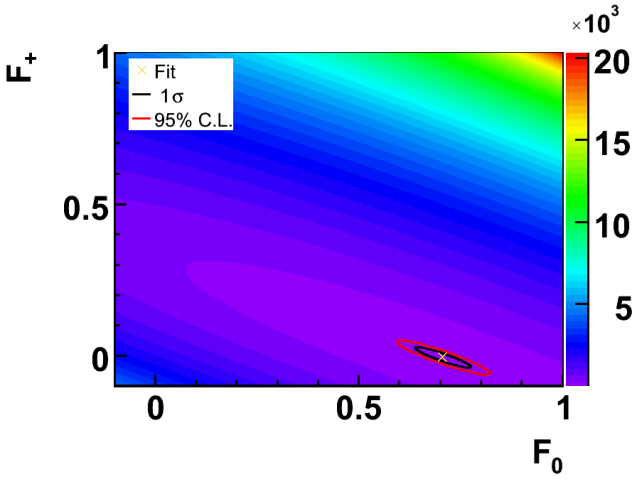


Figure 6.13: Negative log likelihood function as a function of F_0 and F_+. The fitted values of F_0 0.707 $\pm$ 0.032 and F_+ 0.006 $\pm$ 0.016 are marked by the cross and the 1σ and 95% C.L. regions are indicated by the black and red ellipses, respectively.

6.4 Unfolded $\cos\theta^*$ Distribution

In order to allow a direct comparison of the $\cos\theta^*_{\text{rec}}$ distribution obtained from the CMS pseudo experiment with the calculated distributions for the different W boson helicity modes presented in figure 1.7, the background is subtracted from the pseudo data. The shape of the $\cos\theta^*_{\text{rec}}$ distribution is then corrected for acceptance effects as well as for resolution effects applying the transfer function $\tau(F_0, F_+)$.

Since the transfer function explicitly depends on F_0 and F_+ for the correction of the pseudo data distribution the specific transfer function for the (F_0, F_+) values obtained by the likelihood fit have to be used. The value of the transfer function in the i^{th} bin of $\cos\theta^*$ is calculated from the normalized number of events $\hat{\mu}_i^{\text{sig}}$ before applying any selection cuts in the i^{th} bin of $\cos\theta^*$ and from the normalized number of events $\hat{\mu}_k^{\text{sig,obs}}$ after applying the selection cuts and performing the reconstruction in the k^{th} bin of $\cos\theta^*_{\text{rec}}$:

$$\tau_i(F_0, F_+) = \frac{\hat{\mu}_i^{\text{sig}}(F_0, F_+)}{\hat{\mu}_{k=i}^{\text{sig,obs}}(F_0, F_+)} \quad . \tag{6.25}$$

Multiplying the background subtracted number of events in bin k $(k = i)$ of $\cos\theta^*_{\text{rec}}$ with $\tau_i(F_0, F_+ = 0)$ and normalizing subsequently the corrected $\cos\theta^*$ pseudo data distribution to the theoretically calculated $t\bar{t}$ pair production cross section of $\sigma_{t\bar{t}} = 833$ pb [5], leads to the desired distribution which is directly comparable with the theory distributions.

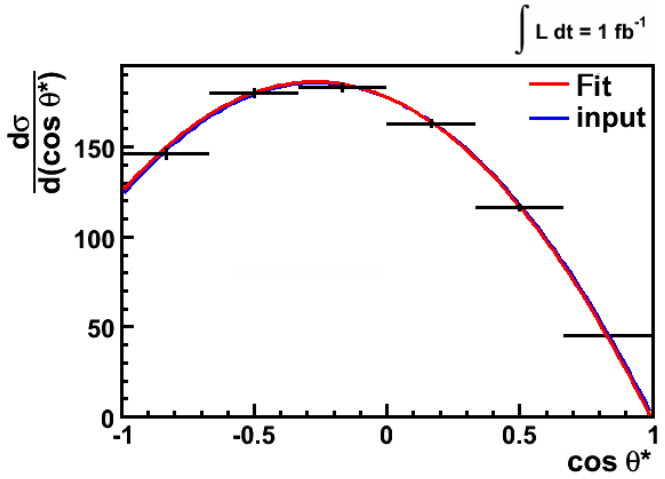


Figure 6.14: Unfolded $\cos\theta^*$ distribution from the CMS pseudo experiment normalized to the theoretically calculated $t\bar{t}$ pair production cross section of $\sigma_{t\bar{t}} = 833$ pb [5]. The distribution corresponding to the fitted values for F_0 and F_+ is shown in red as well as the distribution predicted by the SM, shown in blue.

In figure 6.14 the corrected $\cos\theta^*$ distribution of the pseudo data normalized to the $t\bar{t}$ pair production cross section is presented. The uncertainty in the pseudo data includes the uncertainty of the transfer function. The red curve represents the distribution corresponding to the fitted values for F_0 and F_+. In order to compare the results with the used Monte Carlo true values, the blue curve represents the distribution for the SM value of F_0. The agreement is well as expected.

Chapter 7

Verification of our Method

We verify our measurement method of the W boson helicity fractions using pseudo experiments. These are described in the first section and are then used to perform several checks for the measurement method introduced in the last chapter.

7.1 Pseudo Experiments

Pseudo experiments repeat a measurement of simulated data drawn from a MC sample with certain statistics. For each pseudo experiment we use the number of expected events for an integrated luminosity of 1 fb^{-1} as calculated in section 4.2.1. In total we perform 1000 pseudo experiments to obtain a whole series of simulated measurements and to calculate an average for the helicity fractions F_0 and F_+ as well as the expected sensitivity.

For each pseudo experiment we simulate the reconstructed $\cos\theta^*$ distributions for the signal and background events as shown in figure 6.12. Random numbers of background events are drawn according to a Poisson distribution with a mean of 6538 for $t\bar{t}$ other events and 4937 for $W + n$jets events. The number of lepton+jets events is obtained by subtracting the number of the background events from the number of expected events. Thus, this number also follows a Poisson distribution with a mean of 37961. These numbers of signal and background events are summarized in table 4.4. In the next step the $\cos\theta^*$ distribution for signal and background events is drawn according to the distributions obtained from the corresponding MC samples. We do not select randomly the thrown number of signal and background events from the signal and the background MC samples and reconstruct the $\cos\theta^*$ distribution from the selected events, since only the $\cos\theta^*$ distribution and no other kinematic quantities of the entire event are of interest for our studies.

Finally, the binned likelihood fit described in the last chapter is applied to the sum of generated $\cos\theta^*$ distributions. As a result, each pseudo experiment delivers a pair of the helicity fractions F_0 and F_+ and their uncertainties. The distributions of these values obtained from the whole ensemble of simulated experiments are used for further checks.

7.2 Consistency Check

For the consistency check the $\cos\theta^*$ distribution for each pseudo experiment is generated according the $\cos\theta^*$ distribution obtained from the TopRex MC sample, since this sample was used to determine the efficiency and migration matrix needed to calculate the fit templates. The fitted F_0 and F_+ values of all pseudo experiments are presented in figure 7.1.

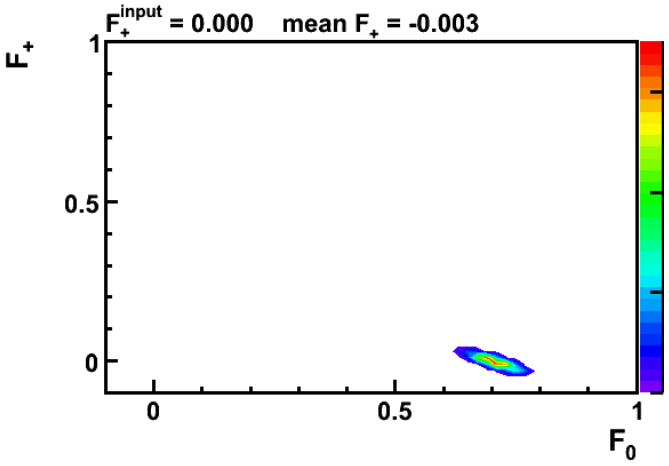


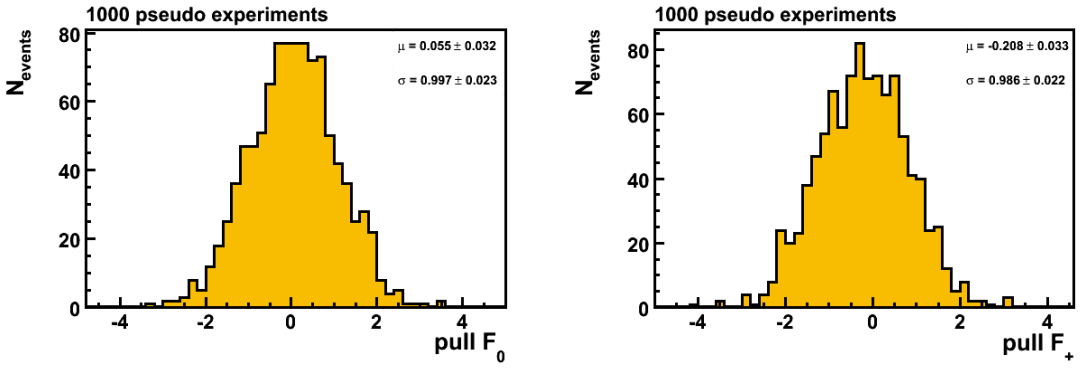
Figure 7.1: F_0 and F_+ values obtained from pseudo experiments. Mean $F_0 = 0.7$, mean $F_+ = -0.003$.

The obtained F_0 and F_+ distributions are very narrow. The quality of the measurement method is checked in detail using the pull distributions. The pull p is defined as the difference between the fitted values $F_{0/+}^{\text{fit}}$ and the input values $F_{0/+}^{\text{input}}$ divided by the fit uncertainty $\sigma_{F_{0/+}^{\text{fit}}}$.

$$p = \frac{F_{0/+}^{\text{fit}} - F_{0/+}^{\text{input}}}{\sigma_{F_{0/+}^{\text{fit}}}} \tag{7.1}$$

The mean μ of the pull distribution is expected to be zero and the width σ distribution is expected to be Gaussian. Figure 7.2 presents the obtained pull distributions for F_0 on the left-hand side and F_+ on the right-hand side. The mean is not zero as expected. The reason for this is that, due to the large number of expected events, the uncertainty of the fit result becomes very small as can be seen in figure 7.3. Since the uncertainty is in the denominator of equation 7.1, the resulting values for the pull get large even for small deviation of the fitted values from the input values.

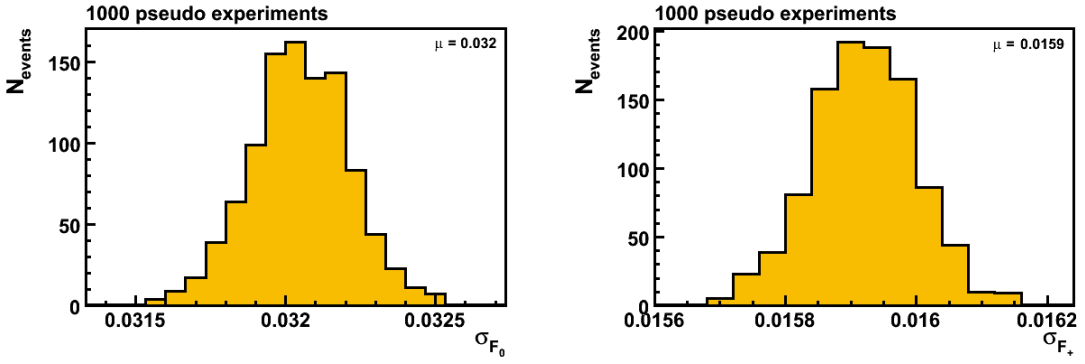
Figure 7.3 illustrates the distributions of the statistical uncertainties of the fitted F_0 and F_+ values obtained from the pseudo experiments. The expected sensitivity of the measurement method is given by the mean of these distributions. It allows



(a) (b)

Figure 7.2: Pull distributions of the F_0 (a) and F_+ (b) fractions obtained using pseudo experiments. The pull is defined as $(F_{0/+}^{\text{fit}} - F_{0/+}^{\text{input}})/\sigma_{F_{0/+}^{\text{fit}}}$.

to compare the statistical uncertainty of one measurement with the average uncertainty of many simulated experiments.



(a) (b)

Figure 7.3: Distributions of the expected sensitivity for the measurement of (a) F_0 and (b) F_+ obtained from pseudo experiments.

7.3 Linearity Check

In this section we check the measurement method using MC samples with different helicity fractions F_0 and F_+. Since these samples were generated with MadEvent we use the efficiency and migration matrix which is also obtained from a MadEvent

sample with Standard Model values $F_0 = 0.7$ and $F_+ = 0$, in order to avoid systematic effects arising from differences between the Monte Carlo event generators. The fit templates calculated using this efficiency and migration matrix are used to extract the F_0 and F_+ values from $\cos\theta^*$ distributions containing different helicity fractions.

The results are presented in figure 7.4 and 7.5. The plots show the extracted F_0 and F_+ fractions as well as the means of the pull distributions and their widths. In addition to the fit method used till now where both parameters F_0 and F_+ are varied we present the results of a fit where only one of parameter is varied, while the other is set to the constant value predicted by the Standard Model. This is possible since only one of the helicity fractions varies in the according MC samples, see 4.1. The results of the one parameter fit are depicted on the left-hand side, while the results obtained from the two parameter fit are shown on the right-hand side.

Figures 7.4a/b show the theoretical dependence of the fraction F_0 of longitudinal polarized W bosons on the top mass (black curve). The blue points indicate the fraction F_0 in the used MC samples obtained from a fit to the MC true values using equation 1.10. The red points are the fit results of our method. The results of our method agree well with the MC true values which were used as input, except in the last case where the generated top-quark mass is 225 GeV/c^2. In this case the MC true values do not match the theoretical curve, and the fit results of our method do not match the input values. This difference is the reason for the missing of the corresponding entry in the next plots 7.4c/d with the means of the pull distributions. As can be seen from the last two plots 7.4e/f the results of the two parameter fit have in average larger deviations to the input values than the results of the one parameter fit. Our measurement method consistently extracts the F_0 fraction from the different $\cos\theta^*$ distributions.

Figures 7.5a/b depict the fitted values for the helicity fraction F_+ as a function of the input values. The results of both fit methods are larger than the input values. As can be seen in figure 1.7 the F_+ fraction as a function of $\cos\theta^*$ has its maximum at $\cos\theta^* = 1$. To extract this fraction properly a good resolution for $\cos\theta^*$ values near one is needed, but we can only use six bins for the whole range of $\cos\theta^*$ values. This is one reason for the deviation of the fit results to the input values. The results of the two parameter fit have larger deviations from the input values and also larger uncertainties than the results from the one parameter fit, since in the two parameter fit the value of the F_+ fraction can be varied at the expense of the value of F_0, while this is not possible in the one parameter fit, where it is set to 0.7. This is also the reason why the pull values of the two parameter fit have larger deviations from zero and larger uncertainties than the pull values of the one parameter fit, see figure 7.5c/d. The widths of the pull distributions, figure 7.5e/f, obtained from the two parameter fit varies more than the corresponding widths obtained from the one parameter fit, like in the case of the F_0 fraction.

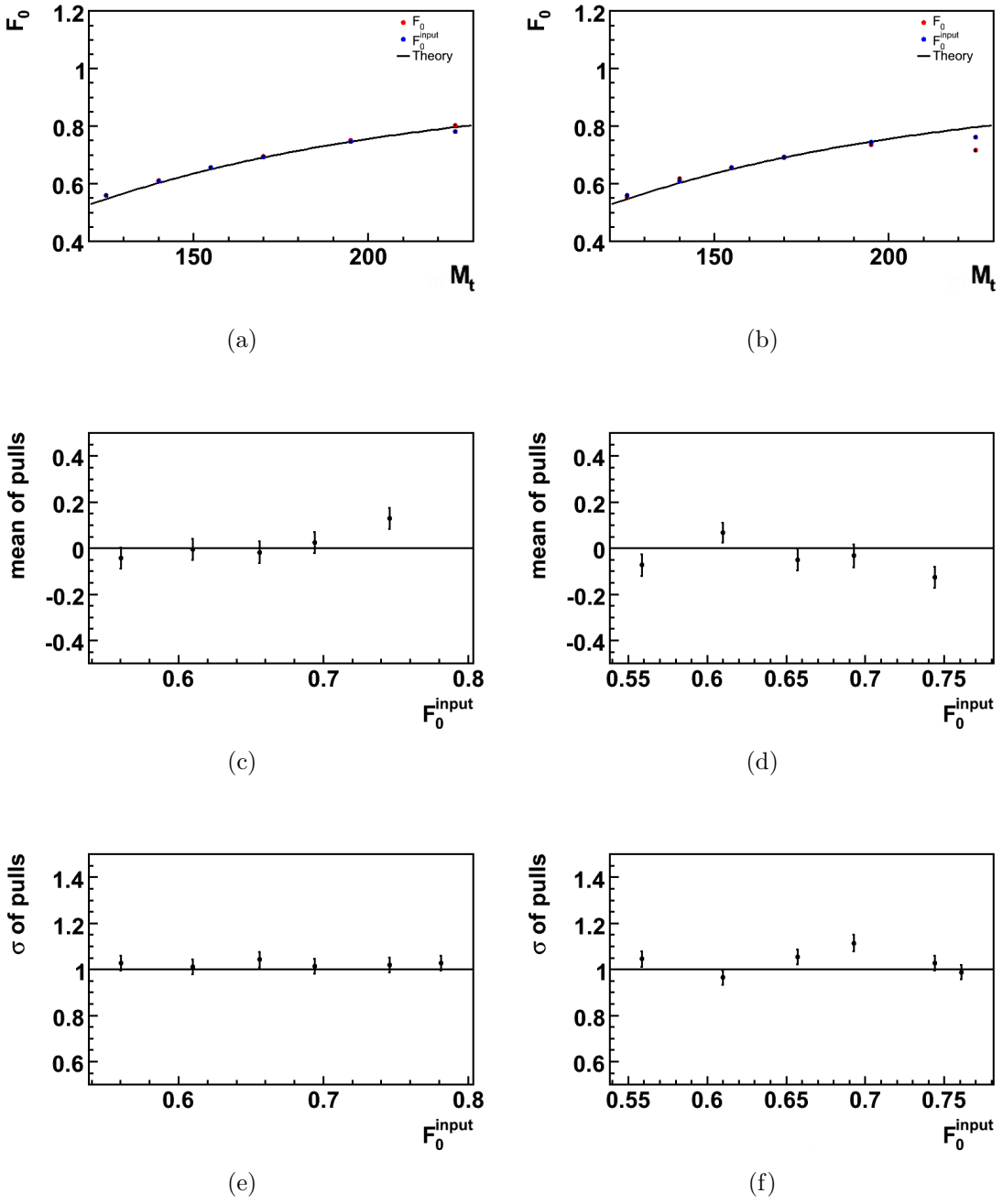


Figure 7.4: Linearity check of the measurement method using different F_0 helicity fractions. The fit results of a one parameter fit (F_0 is varied, while F_+ is set to zero) are shown on the left-hand side and the results of a two parameter fit (F_0 and F_+ are varied) on the right-hand side. (a) and (b) depict the theoretical dependence of the F_0 fraction on the top mass (black line) as well as the F_0 fractions in the MC samples (blue points) and the fit results (red points). (c) and (d) show the means of the pull distributions and (e) and (f) their widths.

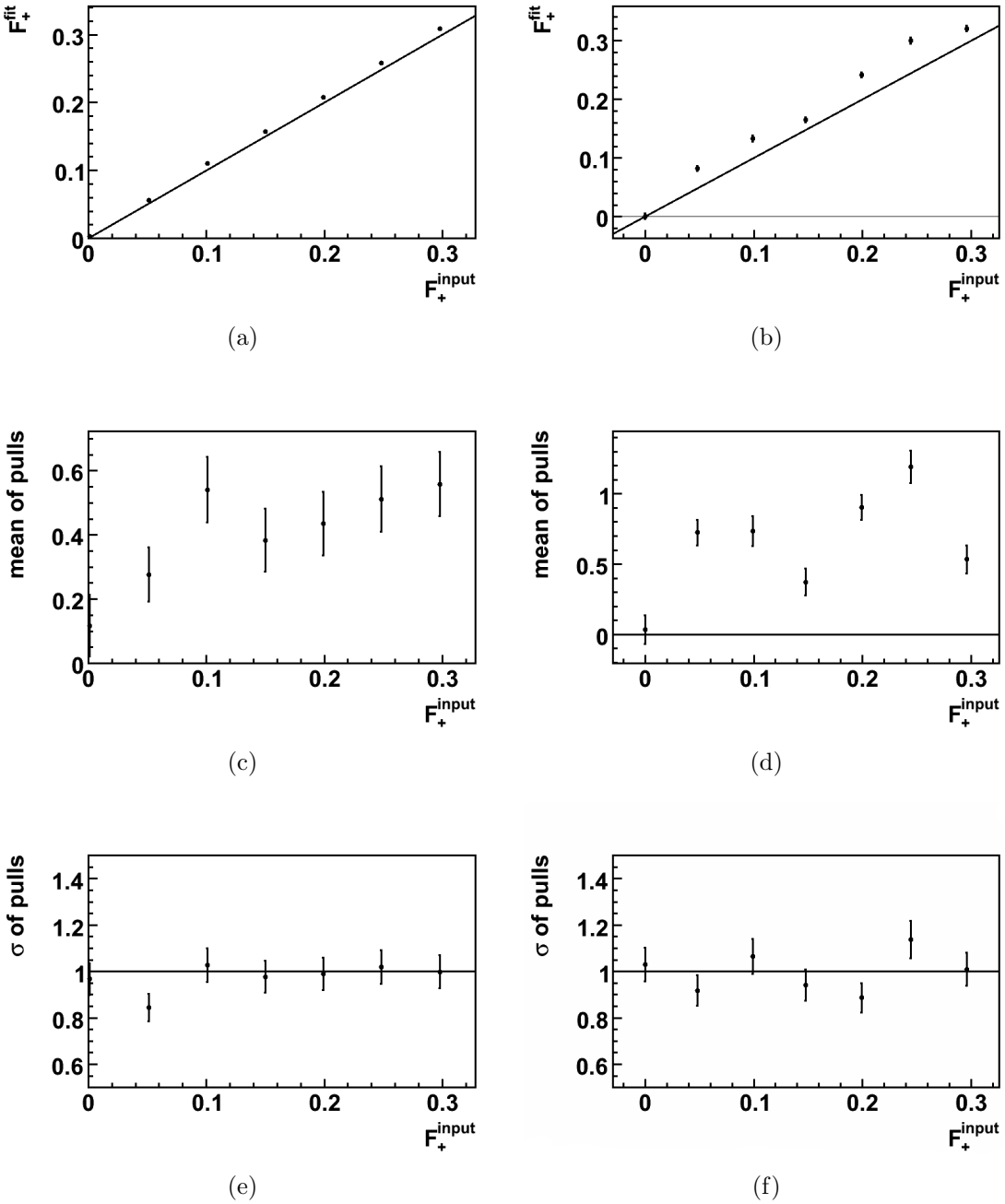


Figure 7.5: Linearity check of the measurement method using different F_+ helicity fractions. The fit results of a one parameter fit (F_0 is varied, while F_+ is set to 0.7) are shown on the left-hand side and the results of a two parameter fit (F_0 and F_+ are varied) on the right-hand side. (a) and (b) depict the F_+ and the fit results. (c) and (d) show the means of the pull distributions and (e) and (f) their widths.

Chapter 8

Summary

In this thesis Monte Carlo samples were used to prepare for the measurement of the W-boson helicity-fractions in top-quark decays with the CMS experiment at the Large Hadron Collider. In the first step two Monte Carlo event generators, Pythia and Herwig, were compared using $t\bar{t}$ events at center-of-mass energies of the Tevatron Collider $\sqrt{s} = 1.96$ TeV and the Large Hadron Collider $\sqrt{s} = 14$ TeV. Besides the different cross sections both Monte Carlo event generators predicted a larger boost of the $t\bar{t}$ pairs at the center-of-mass energy of the $\sqrt{s} = 14$ TeV. In a second step Monte Carlo events generated with MadEvent were used to determine proper cuts to select the lepton+jets $t\bar{t}$ events. Therefore, the signature of these events was exploited. These cuts were then applied to a fully simulated Monte Carlo sample. The event selection was performed on the Grid using CMS Software.

The cosine of the decay angle θ^* was used to measure the helicity fractions. To achieve this the kinematic quantities of the selected $t\bar{t}$ events were reconstructed. The full reconstruction led to different event interpretations. A criterion was applied to each event interpretation in order to select one that matches well the $t\bar{t}$ assumption. The values of $\cos\theta^*$ were calculated for each selected event interpretation. A binned likelihood fit was used to extract the helicity fractions from the obtained $\cos\theta^*$ distribution. The used fit templates were calculated from the theoretical $\cos\theta^*$ distribution considering the efficiency of the event selection and the migration effects of the event reconstruction. Finally, the obtained measurement method was checked using several Monte Carlo samples with different fractions of F_0 and F_+.

This first studies showed that the described method to measure the W boson helicity fractions in top quark decays with the CMS experiment works well. The expected sensitivity for an integrated luminosity of $\mathcal{L} = 1$ fb^{-1}, which corresponds to the expected integrated luminosity in the first year of the Large Hadron Collider, was determined by using pseudo experiments. The relative statistical uncertainty is expected to be 5%. The uncertainty from the assumed value $F_+ = 0.0$ is expected to be ± 0.016. Figure 8.1 depicts the expected sensitivity for higher integrated luminosities. Pseudo experiments with a scaled number of expected signal and background events were performed.

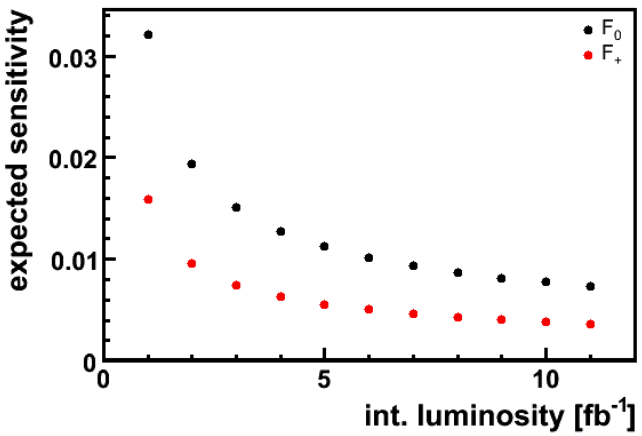


Figure 8.1: Expected sensitivities for higher integrated luminosities. The black points represent the statistic uncertainties for the measurement of F_0, while the red points represent the statistic uncertainties for the measurement of F_+.

Bibliography

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[2] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[3] T.E.W. Group (2006), hep-ex/0603039.

[4] A. Abulencia *et al.*, Phys. Rev. Lett. **96**, 022004 (2006).

[5] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, "NLL resummation of the heavy-quark hadroproduction cross-section," Nucl. Phys. B529 (1998) 424-450, arXiv:hep-ph/9801375.

[6] S.L. Glashow, Nucl. Phys. **22**, 579 (1961).

[7] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[8] W.-M. Yao *et al.*, J. Phys. G **33**, 1 (2006).

[9] H. L. Lai et al., Eur. Phys. G **33**, 1 (2000)

[10] N. Cabbibo, Phys. Rev. Lett. **10**, 531 (1963).

[11] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[12] P.W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).

[13] G.L. Kane, G.A. Ladinsky, and C.P. Yuan, Phys. Rev. D **45**, 124 (1992).

[14] M. Jezabek and J.H. Kühn, Nucl. Phys. **314**, 1 (1989).

[15] M. Jezabek and J.H. Kühn, Phys.Lett. **207**, 91 (1988).

[16] A. Czarnecki and K. Melnikow, Nucl. Phys. **544**, 520 (1999).

[17] K.G. Chetyrkin, R. Harlander, T. Seidensticker, and M. Steinhauser, Phys. Rev. **60**, 114015 (1999).

[18] A. Denner and T. Sack, Nucl. Phys. **358**, 46 (1991).

[19] R. Migneron, G. Eilam, R.R. Mendel, and A. Soni, Phys. Rev. Lett. **66**, 3105 (1991).

[20] M. Fischer, S. Groote, J.G. Korner, M.C. Mauser, and B. Lampe, Phys. Lett. B **451**, 406 (1999).

[21] M. Fischer, S. Groote, J.G. Korner, M.C. Mauser, Phys. Rev. D **63**, 031501 (2001).

[22] M. Fischer, S. Groote, J.G. Korner, M.C. Mauser, Phys. Rev. D **65**, 054036 (2002).

[23] H.S. Do, S. Groote, J.G. Korner, M.C. Mauser , Phys. Rev. D **67**, 091501 (2003).

[24] R.D. Peccei, S. Peris, and X. Zhang, Nucl. Phys. B **349**, 305 (1991).

[25] J.M. Cornwall, D.N. Levin, and G. Tiktopoulos, Phys. Rev. D **10**, 1145 (1974).

[26] B.W. Lee, C. Quigg, and H.B. Thacker, Phys. Rev. D **16**, 1519 (1977).

[27] X. Wang, Q. Zhang, and Q.Qiao, Phys. Rev. D **71**, 014035 (2005).

[28] C.-R. Chen, F. Larios, and C.-P. Yuan, Phys. Lett. B **631**, 126 (2005).

[29] LHC Design Report
`http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html`

[30] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 1," CERN/LHCC **2006-001** (2006). CMS TDR 8.1

[31] CMS Collaboration, "The Magnet Project Technical Design Report," CERN/LHCC **97-010** (1997). CMS TDR 1.

[32] CMS Collaboration, "The Hadron Calorimeter Technical Design Report," CERN/LHCC **97-031** (1997). CMS TDR 2.

[33] CMS Collaboration, "The Muon Project Technical Design Report," CERN/LHCC **97-32** (1997). CMS TDR 3.

[34] CMS Collaboration, "The Electromagnetic Calorimeter Technical Design Report", CERN/LHCC **97-033** (1997). CMS TDR 4.

[35] CMS Collaboration, "The Tracker Project Technical Design Report," CERN/LHCC **98-006** (1998). CMS TDR 5

[36] CMS Collaboration, "The CMS experiment at the CERN LHC," To be submitted to the Journal of Instrumentation (JINST)

[37] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems," CERN/LHCC **2000-38** (2000). CMS TDR 6.1.

[38] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger," CERN/LHCC **2002-26** (2002). CMS TDR 6.2.

[39] M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics," arXiv:hep-ph/0403045v2

[40] Torbjrn Sjstrand *et al.*, JHEP05(2006)026

[41] Pythia Webpage
`http://www.thep.lu.se/~torbjorn/Pythia.html`

[42] A. Kupco "Cluster Hadronization in HERWIG 5.9" arXiv:hep-ph/9906412v1

[43] Herwig++ Webpages
`http://projects.hepforge.org/herwig/talks`
`http://www.hep.phy.cam.ac.uk/theory/Herwig++/mainpage.html`

[44] F. Maltoni, T. Stelzer "MadEvent: Automatic Event Generation with MadGraph", arXiv:hep-ph/0208156v1.

[45] Johan Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation", arXiv:0706.2334v1.

[46] MadEvent/MadGraph Webpage
`http://madgraph.hep.uiuc.edu/`

[47] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions" arXiv:hep-ph/0206293v2.

[48] S.R. Slabospitsky, L. Sonnenschein "TopReX generator (version 3.25). Short manual", arXiv:hep-ph/0201292v1.

[49] J. Reuter, K. Hagiwara, W. Kilian, F. Krauss, T. Ohl, T. Plehn, D. Rainwater, S. Schumann "Next Generation Multi-particle event generators for the MSSM" arXiv:hep-ph/0512012v2.

[50] Homepage of the Sherpa Event Generator
`http://projects.hepforge.org/sherpa/dokuwiki/doku.php`

[51] S. Frixione, P. Nason, and B.R. Webber, "Matching NLO QCD and parton showers in heavy flavour production", arXiv:hep-ph/0305252v2.

[52] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406**, 187 (1993).

[53] J. M. Butterworth, J. P. Couchman, B. E. Cox, B. M. Waugh "KtJet: A C++ implementation of the Kt clustering algorithm" arXiv:hep-ph/0210022v1.

[54] PGS Webpage
`http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs.html`

[55] J. D'Hondt, S. Lowette, J. Heyninck, "Electron and muon reconstruction in single leptonic $t\bar{t}$ events" CMS NOTE **2006/024**

[56] P. Avery *et al.*, "Measurement of Missing Transverse Energy With the CMS Detector at the LHC", CMS NOTE **2006/035**

[57] A. Heister *et al.*, "Measurement of Jets with the CMS Detector at the LHC" CMS NOTE **2006/036**

[58] S. V. Chekanov, "Jet algorithms: a minireview", arXiv:hep-ph/0211298v1.

[59] A. Rizzi, F. Palla, G. Segneri "Track impact parameter based b-tagging with CMS", CMS NOTE **2006/019**

[60] J. D'Hondt, J. Heyninck, S. Lowette, "Measurement of the cross section of single leptonic $t\bar{t}$ events" CMS NOTE **2006/064**

[61] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 2," CERN/LHCC **2006-021** (2006). CMS TDR 8.2

[62] CMS Software Webpage
`http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms\&action=url\&urlkey=CMS_OFFLINE`

[63] Top Quark Analysis Framework Webpage
`https://twiki.cern.ch/twiki/bin/view/CMS/TWikiTopQuarkFramework`

[64] GEANT4 Collaboration, S. Agostinelli et al., "GEANT4: A simulation toolkit," Nucl. Instrum. Meth. **A506** (2003) 250-303.

Danksagung

An aller erster Stelle möchte ich mich bei Herrn Priv.-Doz. Dr. Wolfgang Wagner für die hervorragende Betreuung bedanken.

Ebenso bedanke ich mich bei Herrn Prof. Dr. Th. Müller für die Übernahme des Korreferats.

Dr. Dominic Hirschbühl und Thorsten Chwalek möchte ich besonders danken für die Unterstützung und die vielen Tipps während der gesamten Arbeit, sowie das aufmerksame Korrekturlesen und die zahlreichen Verbesserungsvorschläge.

Julia Weinelt und Philipp Sturm danke ich für das Korrekturlesen und die Hilfe mit CMSSW und dem Grid. Der gesamten Top-Gruppe möchte ich für die mehr als angenehme Arbeitsatmosphäre danken. Allen Mitgliedern des EKP im 8. und 9. Stock danke ich für die Zusammenarbeit vorallem dem Admin-Team sowie dem Sekretariat.

Meiner Schwester und meinen Eltern danke ich sehr herzlich für ihre Unterstützung und dafür dass sie mir das Studium ermöglicht haben.

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

Georg Sartisohn

Karlsruhe, 27. November 2007