GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

II. Physikalisches Institut

Studies with the Matrix Element Method using the example of a top quark mass measurement with the ATLAS experiment.

von

Andrea Knue

geb. in Meppen

The Matrix Element Method has been studied on the example of a top quark mass measurement in the electron + jets channel with the ATLAS experiment. Emphasis has been placed on the numerical implementation of additional parameters compared to existing implementations. Tests with smeared LO parton level events and NLO Monte Carlo data have been done. It was shown that the method in the current implementation provides an unbiased estimator of the top quark mass. Additional studies for the *in-situ* calibration of m_{top} using the JES have been performed. The light and the b quark JES can be estimated from top events without any bias.

GEORG-AUGUST-UNIVERSITÄT

GÖTTINGEN

II. Physikalisches Institut

Die Matrix Element Methode mit der Anwendung der Messung der Topquark Masse im semileptonischen Kanal am ATLAS Detektor

von

Andrea Knue

geb. in Meppen

Dieser Forschungsbericht wurde als Diplomarbeit von der Fakultät für Physik der Georg-August-Universität zu Göttingen angenommen.

Contents

1. Introduction

Elementary particle physics is a research field with a long tradition which is still developing and expanding. Since the discovery of the quarks and the gauge bosons, increasing efforts have been made in order to unreveal the structure of matter and its interactions. The Standard Model of particle physics comprises the elementary particles and fundamental forces between them. Although is has been successfully tested over the past years, still a lot of open questions remain. With the construction of the LHC, physicists hope to access to - so far - unexplored energy regions.

The most heavy particle of the SM is the top quark. It has been discovered in 1995 at the DØ and CDF detectors at the Tevatron collider. Due to its large mass of 173.1 GeV/c^2 (current world average) the Tevatron is the only collider where top quarks have been produced. The mass of the top quark is one of the most precisely measured parameters of the SM. The top quark has a very short lifetime and can therefore not form any bound states with other particles.

Recently, the production of single tops via the weak interaction has been observed. Single tops make it possible to further study important parameters, for example the CKM matrix element $|V_{tb}|$. One hopes to rediscover the top quark in the early data of the LHC. The detailed knowledge of its mass can be used to further constrain the Higgs mass.
A method for a top mass measurement is presented in this thesis. This measurement is done with the *Matrix Element Method* which has been developed by the DØ collaboration. The concept is based on a transition matrix element which is a prediction for the transition probability between initial and final states. The method uses the full kinematic information of the initial and final state particles and is therefore very precise.

The Standard Model of particle physics is introduced in Chapter 2 and the properties of the top quark are discussed. The experimental setup of the LHC and the ATLAS detector are presented in Chapter 3. The fourth Chapter introduces the Matrix Element Method. The method itself is explained first. The integration method is discussed and phase space transformations are performed. While knowledge about the detector resolution is crucial for precise measurements, transfer functions are used to map the parton energies to the measured energies. The concept of Markov Chain Monte Carlo is discussed in Chapter 6. Markov Chains make it possible to sample a function in the interesting regions of the phase space. The method has been applied to smeared parton level events (Chapter 7). In Chapter 8, the top quark mass is estimated from NLO MC data with a LO matrix element. The JES is the largest source of systematic uncertainties and has been studied in Chapter nine. An *in-situ* calibration of the top quark mass has been performed with a global and two separate JES factors. Finally, the results of the presented studies are summarized and an outlook on further studies is given.

2. The Standard Model

2.1. Elementary particles and fundamental forces

Particle physicists try to explore the fundamental particles and their interactions. Our current knowledge can be summarized in the so called *Standard Model of Particle Physics* (SM) [1, 2]. It contains the fundamental particles[1] and describes their interactions. The expression *elementary particle* referes to a pointlike particle, i. e. without substructure. The SM comprises twelve elementary particles with spin $\frac{1}{2}$ (fermions) as well as bosons with integer spin (Fig. 2.1). The fermions form three groups called *generations*. Each generation comprises two weak isospin doublets. One contains two quarks with electrical charge of $+\frac{2}{3}e$ and $-\frac{1}{3}e$, the other one a charged (-e) and an electrically neutral lepton (neutrino). Only left-handed doublets of leptons are observed in nature (parity violation). Right handed particles only exist in singlets. As the constituents of protons and neutrons, the quarks of the first generation ($m_{\mathrm{up,down}} \approx 1 \ \mathrm{MeV/c^2}$) and the associated leptons make up the visible matter in the Universe. The particles of the second and third generations have much higher masses and are therefore unstable. The two heaviest elementary particles are the top quark (173.1 GeV/c^2 [3]), the τ-lepton (1.8 GeV/c^2) and the b quark (5 GeV/c^2).
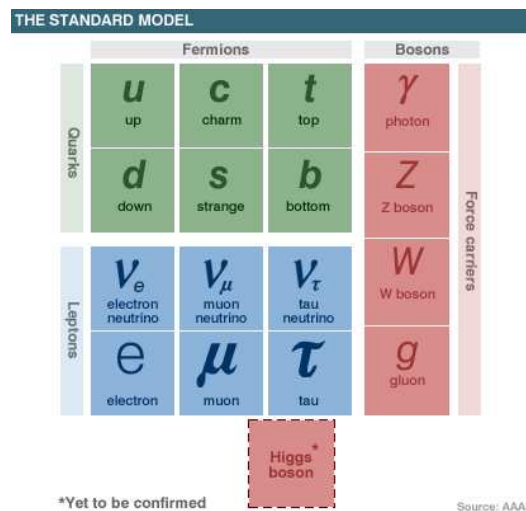
Figure 2.1: The three generations of elementary particles, the gauge bosons and the Higgs boson.

Each particle has a corresponding antiparticle. The antiparticles build a similar (right-handed) structure. The antiparticles are not quoted here separately. Quarks have an additional quantum number called colour charge (red, green, blue). The forces are mediated via gauge bosons. Each particle in the SM can be associated with a field. Theories which describe these particles and forces are therefore called *Quantum Field Theory (QFT)*.

---

[1]The number of particles is not predicted in the SM.

The field theories can be derived from the Lagrangian of the free fields while requiring local gauge invariance under the transformation of certain gauge groups. An interaction is then associated with a group.

Three fundamental forces are included in the SM: the strong force between quarks, described by the $SU(3)_C$ gauge group and mediated via eight gluons; the electromagnetic force between electrically charged particles ($U(1)$) and the weak force, which interacts with all fermions. The latter two can be unified to the electroweak (EW) force which is described by the groups $SU(2)_L \times U(1)_Y$. This was first proposed by Glashow, Weinberg and Salam (Nobel prize 1979). The electroweak force is mediated by the photon and three heavy gauge bosons, namely the $W^{\pm}$ and the Z^0. The fourth fundamental force is gravity and its hypothetical (while not observed) mediator is the graviton. Although gravity plays a crucial role in daily life, it is not yet fully understood on the quantum level and is consequently not included in the SM. The properties of the different particles and interactions are listed in Tab. 2.1.

Interaction	Gauge boson	Strength (rel. to the strong int.)	Range [m]
strong	8 gluons	1	10^{-15}
weak	W^+, W^-, Z^0	10^{-13}	10^{-18}
electromagnetic	γ	10^{-2}	∞
gravity	(graviton)	10^{-42}	∞

Table 2.1: The fundamental forces with the corresponding mediators, their rel. strength (cf. [2], p. 59) and range.

Since local gauge invariance is claimed, all particles like the photon and the eight gluons have to be massless. This does not hold true for the gauge group $SU(2)$ (weak interaction). From the experiment UA1 at the SPS (CERN, 1983) it is known that the W and Z bosons of the weak interaction are massive particles ($m_W = 80.4$ GeV/c^2, $m_Z = 91.2$ GeV/c^2 [4]). This discrepancy between theory and experiment could be solved by the mechanism of spontaneous symmetry breaking which leads to the so-called *Higgs Mechanism*.

In this theory [5] a scalar field is postulated which pervades the space and has a non-zero vacuum expectation value. A sketch of the potential of this field can be seen in Fig. 2.2.
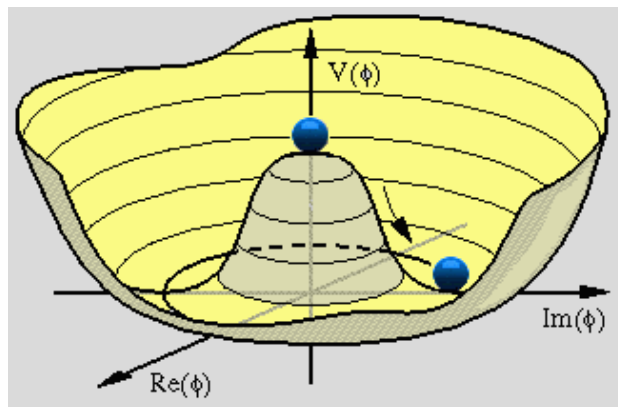

Figure 2.2: Schematic view of the potential of the Higgs field with a non-zero vacuum expectation value on a ring. [6]

The symmetry is spontaneously broken if one point for the potential has been chosen as ground state. This leads to mass terms in the Lagrangian density while gauge invariance is not

violated. The different masses are a result of the different couplings of the background field to the particles. Furthermore, a new boson with large mass is generated. This particle is called Higgs boson and arises from the self-coupling of the Higgs field. The Higgs is the only particle of the SM which has not been observed up to now.

2.1.1. Unanswered questions

The existence of the Higgs boson is not the only open question of the SM. Some further questions are discussed in the following:

- Since the introduction of the neutrino by Pauli in 1930, it was assumed to be massless. In the last years experimental results have shown, that the masses of the neutrinos are not equal to zero but very small. Up to now, only the differences of the neutrino masses have been measured [7].

- A further question is the matter-antimatter asymmetry after the Big Bang. Without this asymmetry, particles and antiparticles would have annihilated and only energy would have been left. CP violation (which has been observed in neutral kaon decays) could be an explanation for the observed asymmetry.

- Even the charge of particles raises still a lot of questions: Why has the proton, a particle with substructure and with a mass of about 1 GeV/c^2 just the opposite charge as an electron, which is an elementary particle with a mass of 0.511 MeV/c^2 ?

- Fritz Zwicky discovered already in 1933 that galaxies in the Coma cluster had a much larger rotational velocity than expected. This velocity could be explained by an additional fraction of mass, which has not been observed up to now. Therefore, it is called *Dark Matter* (DM). Although the observation was in 1933, the nature of the DM has not been identified yet. Some particles which could be responsible for this mass are embraced by the term "WIMPs" which is the abbreviation for *Weakly Interacting Massive Particles*. The fraction of missing matter is large. Massive particles could easily explain the difference. Also neutrinos are listed as a good candidate for hot DM. Being a light particle, the missing mass cannot be explained only by the neutrinos.

- Some of the questions raised above can be explained with an extension of the SM. One possible extension is the so called *Minimal Supersymmetric Standard Model* (MSSM, [8]). In this supersymmetric theory, each fermion of the SM has a bosonic partner and vice versa. These particles have the same quantum numbers except for the spin. Up to now, SUSY particles have not been found and thus the symmetry has to be broken.

2.2. Hadron collider physics

The LHC is a proton-proton collider. The interacting protons consist of valence quarks, sea quarks and gluons which carry a fraction x of the proton momentum. The value of x is not known *a priori* on an event by event basis. The momentum fraction cannot be precisely derived from pertubative QCD calculations but can be modeled by probability density functions from measurements of Deep Inelastic Scattering processes (DIS) [9].

Figure 2.3 (a) shows a schematic view of the interaction of two protons. The function $f(x_i, Q_i^2)$ is the probability density for quarks or gluons with a certain momentum fraction x_i and factorization scale Q_i^2 to be observed in the proton when probed at a scale Q_i^2 and are therefore called *Parton Density Functions* (PDFs). The PDFs are measured and their x and Q dependence can be parameterized. The two most commonly used parameterizations are MRST [10] and CTEQ [11]. Figure 2.3 (b) shows the PDFs for different partons at a factorization scale of 2 GeV from the CTEQ collaboration. For processes at low x, the largest contributions originate from gluons while at higher x, mainly up and down quarks contribute.

(a) Hard scattering process of a pp-interaction [12]

(b) Parton distribution functions (CTEQ 6M) for $Q = 2$ GeV ([11])
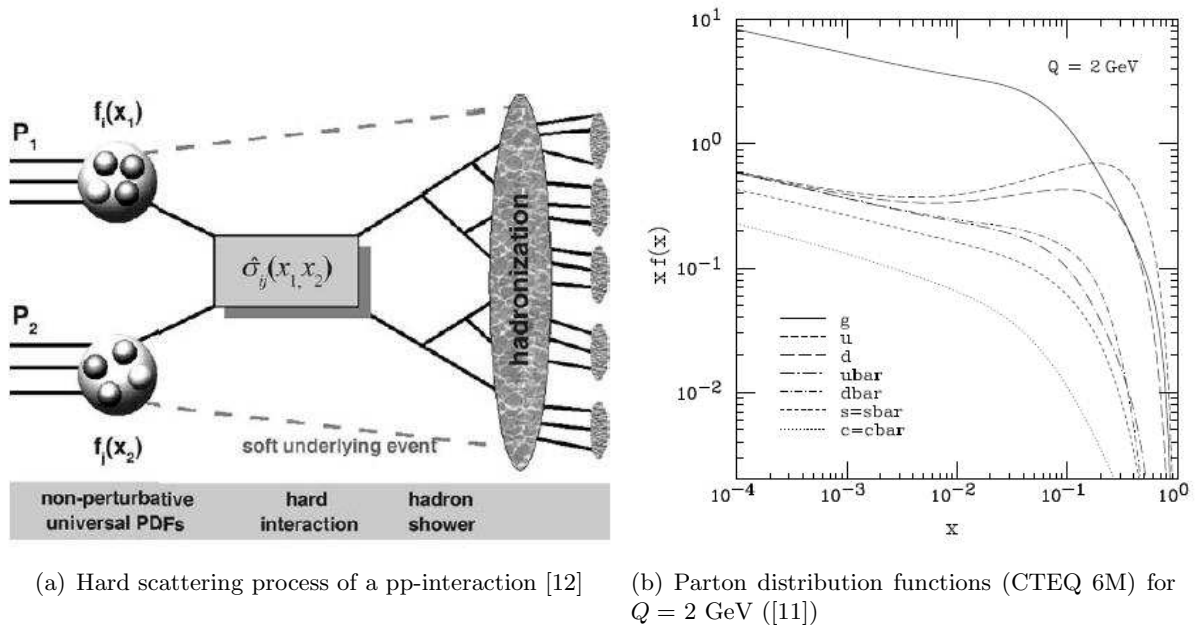
Figure 2.3: Hard scattering process of two protons. Partons which carry a fraction x of the proton momentum (right), interact and hadronize to jets (left)

The cross section $\hat{\sigma}_{i,j}$ of the hard scattering process has to be convoluted with these probability densities and summed over all possible inital state particles i and j in order to get the total hadronic cross section of the process (cp. [12] p.844):

$$\sigma_{\text{tot}} = \sum_{i,j} \int d\,x_i d\,x_j \, f_i(x_i, Q^2) \, f_j(x_j, Q^2) \, \hat{\sigma}_{i,j \to a,b}(x_i, x_j). \tag{2.1}$$

2.3. Top Quark Physics

The top quark plays an important role in the theory of elementary particles. It is the heaviest elementary particle and has a very short lifetime of $\tau \approx 0.42 \times 10^{-25}$ s. For this reason, the top quark decays before it can form any bound state (the hadronization time is $\approx 10^{-23}$ s). Due to its large mass and the short lifetime it is difficult to produce and observe it. It took several years after the discovery of the bottom quark in 1977 until its weak isospin partner had been found in 1995 at the CDF and DØ experiments at the Tevatron collider near Chicago.

In the following, the different production mechanisms of top pairs and single top quarks at hadron colliders are discussed. Subsequently, the different decay channels of the top quark are introduced. Finally, experimental aspects in top quark physics and features of the top quark mass are discussed.

2.3.1. Top Pair Production

Top quark pairs are produced via the strong interaction in two different processes at hadron colliders, $q\bar{q}$-annihilation and gluon fusion. In contrast to the Tevatron, the production via gluon fusion is the dominant channel (Tab. 2.2). This can easily be explained with the substructure of the proton (see previous section). According to the PDFs (Fig. 2.3), the quarks carry away a larger fraction of the proton momentum than the gluons. At the Tevatron, which has about $\frac{1}{7}$ of the center of mass energy compared to the LHC, the gluons rarely have enough energy to produce top quark pairs. At the LHC, the energy of the gluons is high enough and their contribution dominates at low x. Since the partons have less energy than the protons, the effective center-of-mass energy is smaller:

$$\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}, \tag{2.2}$$

with x_i being the momentum fraction of the ith parton. If the two partons have the same fraction x of the proton momentum, the minimum value to produce a top pair at threshold can be calculated as follows:

$$\sqrt{\hat{s}} \geq 2\, m_{\text{top}}\,, \tag{2.3}$$

$$\Rightarrow x_{\text{min}} \approx \frac{2\, m_{\text{top}}}{\sqrt{s}}\,. \tag{2.4}$$

	Tevatron		LHC
$\sqrt{s}$ [TeV]	1.8 (Run I)	1.96 (Run II)	14
x_{min}	0.192	0.177	0.025
$q\bar{q} \to t\bar{t}$ [%]	90	85	10
$gg \to t\bar{t}$ [%]	10	15	90

Table 2.2: The fraction of proton momentum needed for the production of a top quark pair with a top mass of 173.1 GeV/c² (Fig. 2.11)

.

The minimum x and the different center of mass energies of the Tevatron and the LHC are shown in Tab. 2.2. The top mass is assumed to be 173.1 GeV/c². At the LHC, a much smaller momentum fraction ($x_{\text{min}} \approx 0.025$) is necessary to produce top quarks than at the Tevatron Run I ($x_{\text{min}} \approx 0.192$) and Run II ($x_{\text{min}} \approx 0.177$).

This leads to an increase of the $t\bar{t}$ cross section to $\sigma_{t\bar{t}}^{\text{NLO+NLL}} = 908^{+82}_{-85}$ (scales) $^{+30}_{-29}$ (PDF) pb assuming a top quark mass of of 171 GeV/c^2 and using the PDFs provided by CTEQ 6.5 [13]. Even at a lower center-of-mass energy ($\sqrt{s} = 10$ TeV) in the early days of the LHC, the cross section will be very high ($\sigma_{t\bar{t}} \approx 400$ pb [14]). The Feynman diagrams of the leading order (LO) production processes are shown in Figs. 2.4 and 2.5.
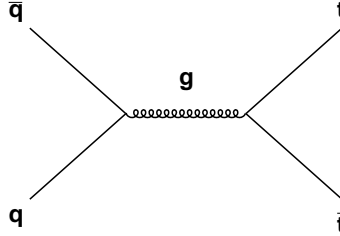
Figure 2.4: LO diagram of top pair production via $q\bar{q}$-annihilation.
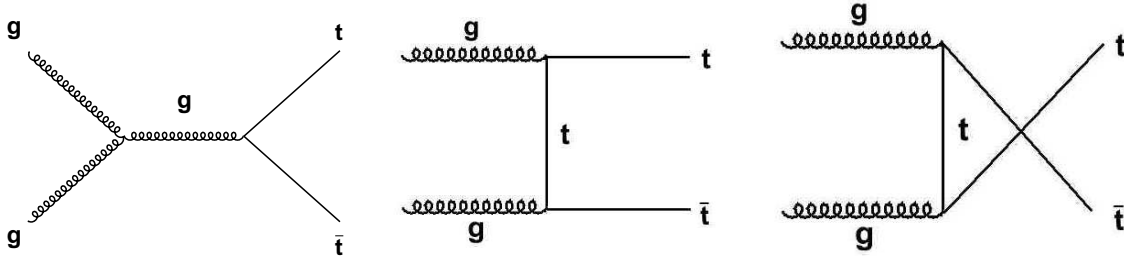
Figure 2.5: LO diagrams of top pair production via gluon fusion.

2.3.2. Single Top Quark Production

Top quarks can also be produced as single particles via the weak interaction. The single top production has been discovered recently by the DØ and CDF collaborations ([15], [16]). Three different channels are distinguished:

a The t-channel production is the most common one (see Fig. 2.6 (a)). Its expected cross section at the LHC is about 240 pb[1].

b The second most common channel is the associated production of a top quark and a W boson (see Fig. 2.6 (b)). While its production cross section is very small at the Tevatron (≈ 0.2 pb), it is expected to be about 66 pb at the LHC.

c The third channel is the s-channel production with a predicted cross section of 10 pb (Fig. 2.6 (c)).

---

[1]All cross sections are NLO predictions (cp. [17], p. 8)

(a) t-channel [18]   (b) associated Wt production [14]   (c) s-channel [18]
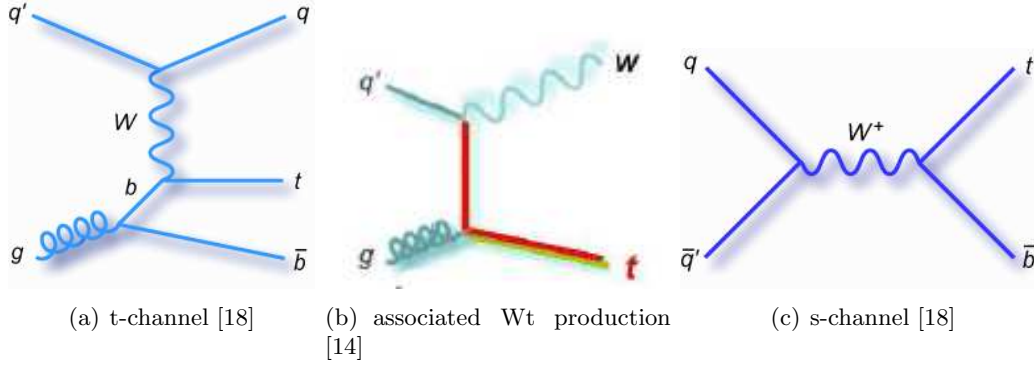
Figure 2.6: LO diagrams for single top production.

The production of single top quarks offers measurements of different top quark properties *e.g.* the CKM matrix element V_{tb}. The ratio of CKM matrix elements is expected to be:

$$R = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2} \approx 1 \ . \tag{2.5}$$

This holds true if only three generations of quarks and leptons exist. In single top events, a direct measurement of $|V_{tb}|$ is possible and could give hints to new physics. A significant deviation of $|V_{tb}|$ from 1 would for example indicate the existence of a fourth quark generation. However, the combined measurement of $t\bar{t}$ events at CDF and DØ yield the following result [19]:

$$R = 0.97^{+0.09}_{-0.08}(\text{stat.} + \text{syst.}). \tag{2.6}$$

The new limits at 95 % confidence level (C.L.) are $|V_{tb}| > 0.78$ (DØ, with $m_{\text{top}} = 170$ GeV/c^2 [15]) and $|V_{tb}| > 0.71$ (CDF, with $m_{\text{top}} = 175$ GeV/c^2 [16]). A precise measurement of the $t\bar{t}$ production cross section is important for the single top production since top pair production is one background for these processes.

2.3.3. Decay Channels

While the top quark decays almost exclusively into a b quark and a W boson, the event signature depends on the decay channels of the W boson.
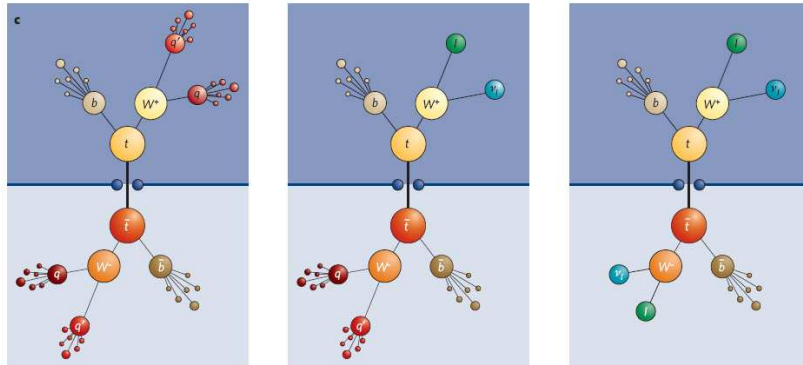
Figure 2.7: The decay channels of the top quark [20].

The W can decay into two jets (BR $\approx$ 67 %) or into a lepton and its corresponding neutrino (branching ratio (BR) $\approx$ 33 %) [4]. Therefore, there are three different signatures (Fig. 2.7):

a) the full hadronic channel ($t\bar{t} \rightarrow b\bar{b}q\bar{q}q\bar{q}$),

b) the semileptonic channel ($t\bar{t} \rightarrow b\bar{b}q\bar{q}l\nu$), and

c) the dileptonic channel ($t\bar{t} \rightarrow b\bar{b}l\bar{l}\nu\bar{\nu}$)

Due to lepton universality, the branching ratio for the three leptons is expected to be approximately the same, *i.e.* BR $\approx$ 11 % (cp. [4]).
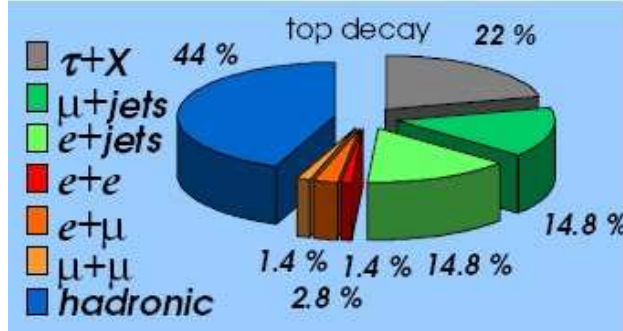
Figure 2.8: The different decay channels of the top quark.

The final state particles of the top pairs leave characteristic signatures in the detector. For the semileptonic channel, four jets and a high-p_T lepton (e or μ) are expected. Due to the undetectable neutrino, missing transverse energy ($E_{T_{miss}}$) can be observed. The tau-lepton is the heaviest lepton and can decay in other leptons or one or more hadrons. Therefore, it is not included in the semileptonic channel.

2.3.4. Background channels

Events with the same signature as the signal are called background. In the case of the semilep-tonic channel, background consists mainly of QCD multijet events and $W + 4\ jets$ events with a leptonically decaying W. QCD background consists of multijet events with misidentified elec-trons or muons with fake isolation. Additional jets come from underlying events or initial and final state radiation (ISR & FSR). The fraction of multijet events can only be measured from real data and is therefore not included in the following analysis. The semileptonic channel has a relatively small background and only one neutrino in the final state. Therefore, it is often called the *golden channel*. The Feynman diagrams for the $W + 1$ or $W + 2$ jets are shown in Fig. 2.9 and 2.10.

Figure 2.9: LO Feynman diagrams fo the production of W + 1 jet events.
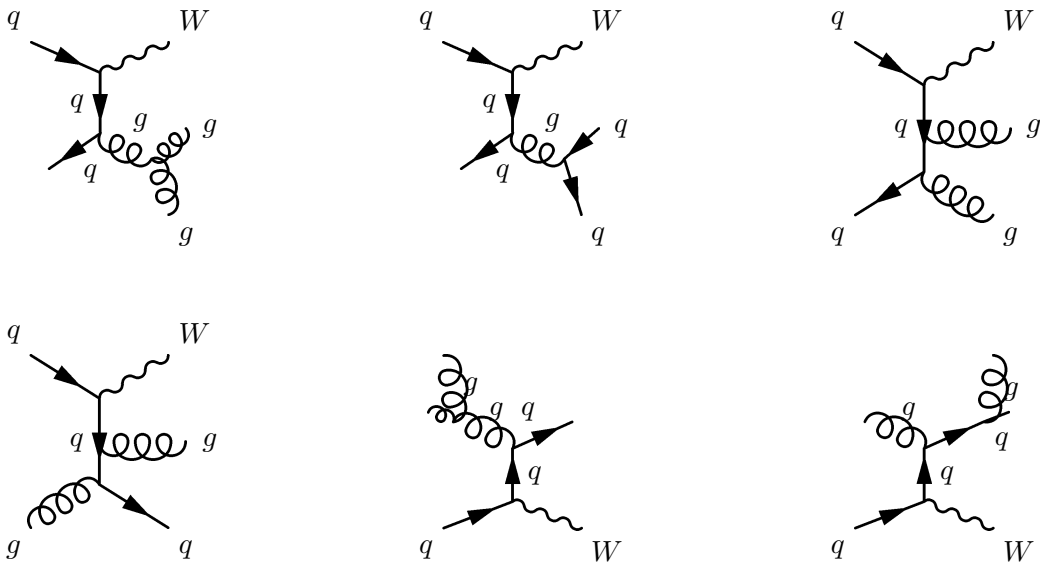
Figure 2.10: Background events with $W + 2$ jets (LO). The second jets come from the radiation of one or two gluons.

2.3.5. The Top Quark mass

The top quark is not only the heaviest particle in the SM, its mass is also the most precisely measured quark mass so far.
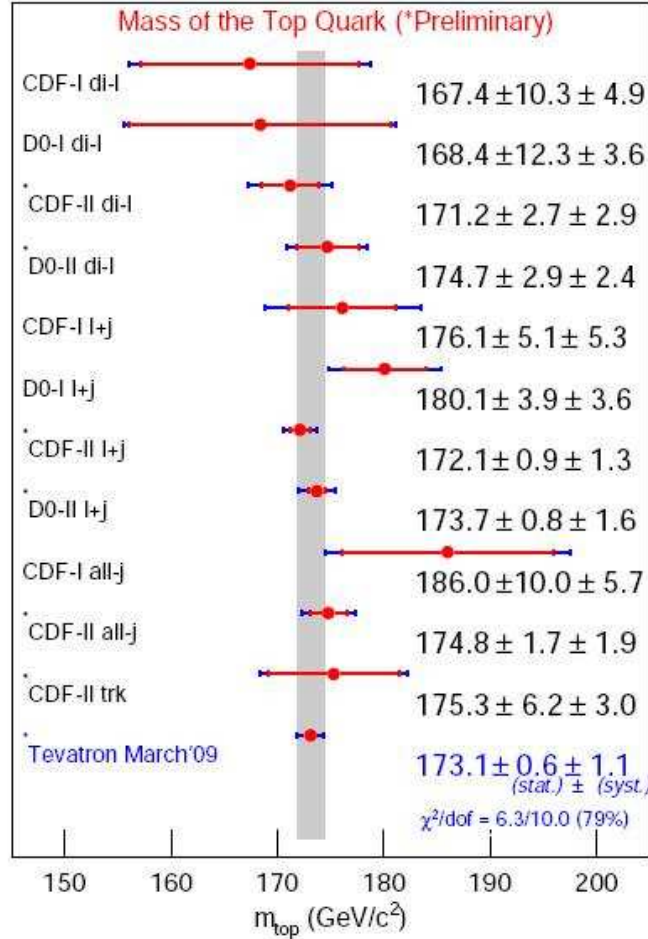
Figure 2.11: The current combination of the top mass measurements at CDF and DØ [3].

The current world average value from the Tevatron yields a top mass of [3]

$$m_{\text{top}} = 173.1 \pm 0.6 \text{ (stat)} \pm 1.1 \text{ (syst)} \text{ GeV}/c^2 \tag{2.7}$$

as one can see in Fig. 2.11. The overall uncertainty of the top mass is therefore 1.3 GeV/c^2 which is only 0.75 percent. These results show that the most precise measurements have been made in the lepton + jets channel.

This precision is very important to constrain the mass of the Higgs boson. In the following equation, the dependence of the W mass from loop radiative corrections, Δr, are shown.

$$M_W^2 = \frac{\hbar^3}{c} \frac{\pi\alpha}{\sqrt{2}G_{\text{F}}} \frac{1}{(1 - (\frac{M_W}{M_Z})^2)(1 - \Delta r)}. \tag{2.8}$$

These loop corrections have contributions from the top quark and the Higgs boson. The contribution from the top quark is quadratic in m_{top} and the one from the Higgs is logarithmic in $\frac{m_H^2}{m_Z^2}$.

$$(\Delta r)_{\text{top}} \propto m_{\text{top}}^2 \,, \tag{2.9}$$

$$(\Delta r)_{Higgs} \propto \ln \left(\frac{m_H^2}{m_Z^2} \right) \,. \tag{2.10}$$

Measurements of M_{W} and m_{top} to constrain the Higgs mass have been done. The correlation of these three masses is shown in Fig. 2.12.
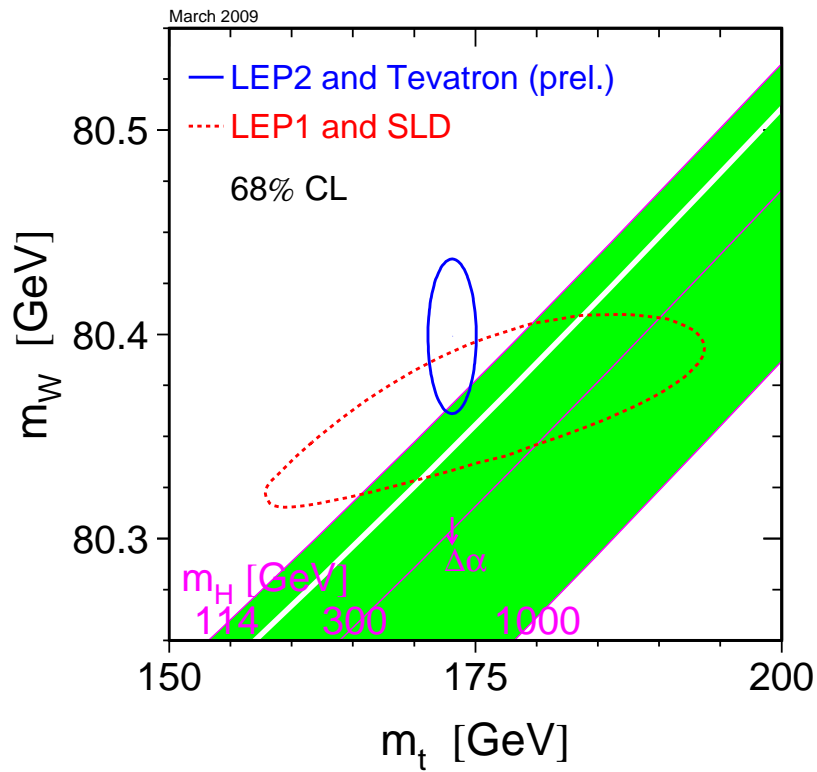
Figure 2.12: The correlation between the top and the W masses as well as a green band of Higgs mass assumptions are shown [21]. The solid and dashed lines show the results of the measurements performed in the runs of LEP1 (dashed) and LEP2 (solid) for a 68 percent confidence level. The Higgs mass limits are given due to direct searches (lower limit) and theoretical predictions (upper limit). The white line shows the mass region which has recently been excluded by Tevatron searches.

Jet Energy Scale (JES)   A very important task is the measurement of the uncertainty due to the jet energy scale (JES) which will be the largest source of uncertainty at the LHC. The JES is the generic term for jet energy mismeasurements. Several corrections have to be applied on the jet energies in order to estimate the parton energies correctly. The most important effects leading to a mismeasurement are discussed here.

The energy of a jet is measured by summing up the energy deposits in the calorimeter cells in a cone with a finite width. If the cone does not contain the complete shower or if low-p_T gluons are emitted, the jet energies are underestimated. Furthermore, energy may be lost in the services or other dead material. On the other hand, the energy can be overestimated due to overlapping jets from multiple parton interactions and underlying events. Further corrections have to be applied to correct for hadronization and fragmentation effects.

While the b quark has a high mass ($4.2^{+0.17}_{-0.07}$ GeV [4]), it decays within the detector and leaves a broader shower in the detector than the light quarks with negligible masses, one can distinguish between a JES for light and b jets. Top quark events can be used for the determination of the JES. As mentioned before, the top quark decays mainly in W bosons and b quarks. The hadronically decaying W boson allows to measure the light JES while the helicity angle distribution gives the dependence to the b JES.
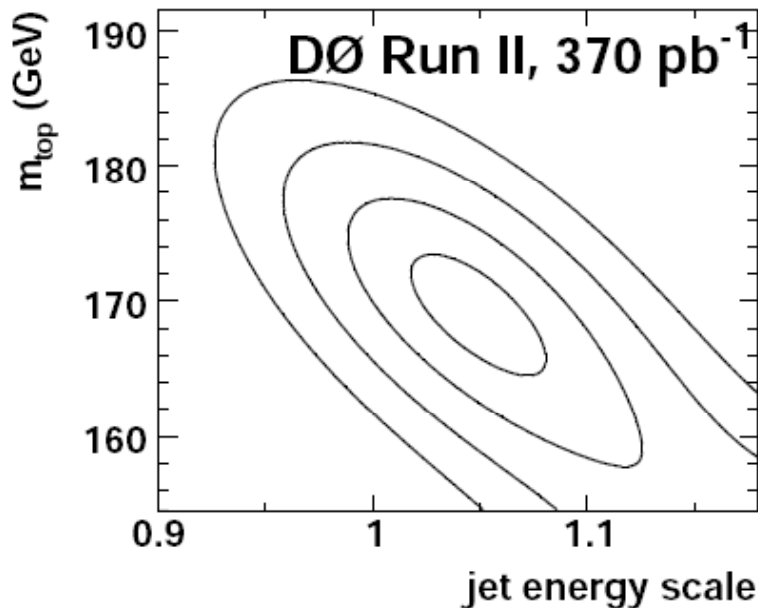
Figure 2.13: Measurement of the jet energy scale and the top quark mass with 370 pb^{-1} of data at DØ[22]. The measurement has been done in the semileptonic channel. Due to the hadronically decaying W boson and two b jets in the final state, the top quark mass is correlated to the JES and allows an *in-situ* calibration of both parameters.

3. The Experimental Setup

The experimental setup of the Large Hadron Collider (LHC) and the ATLAS experiment are described in the following Chapter. The different detector parts are discussed with an emphasis on the electromagnetic and hadronic calorimeters. The two different magnetic systems are compared as well. Finally, the trigger system and the data acquisition systems are explained.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is a proton-proton (Pb-Pb) collider with a center-of-mass energy of $\sqrt{s} = 14$ TeV ($\sqrt{s} = 5.5$ TeV) located at CERN, Geneva. The collider has a total length of 27 km and crosses the border between France and Switzerland 100 m underground. Four experiments are situated at the LHC, two of which are multipurpose experiments, CMS (*Compact Muon Solenoid*) and ATLAS (*A Toroidal Lhc ApparatuS*). Moreover, one experiment has been built to study properties of the quark-gluon plasma (Alice), and one to study *b*-physics (LHC-b).
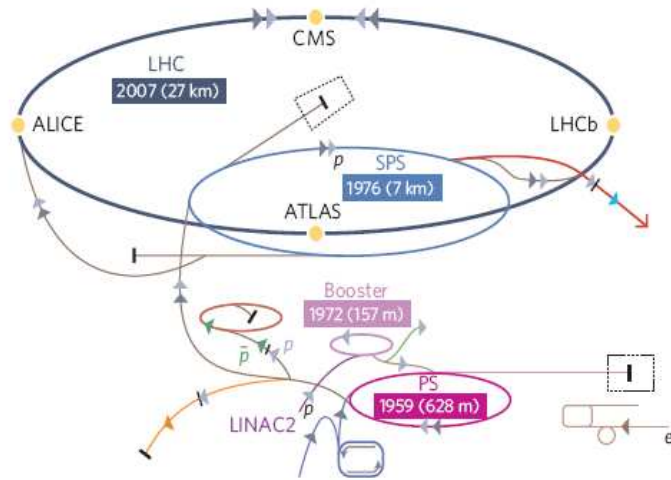
Figure 3.1: The LHC storage ring with its preaccelerators LINAC2, PS and SPS [23].

Compared to the $p\bar{p}$-collider Tevatron (Run II), the LHC has a seven times higher center of mass energy. The decision to use protons for each beam instead of protons and antiprotons has several advantages. First of all, the production of antiprotons is very expensive. As already mentioned in Chapter 2, the fraction of the proton momentum carried by the gluons increases with the center of mass energy. The gluon contribution to the production cross sections increases with the increased center of mass energy.

The protons are produced from hydrogen atoms by ionization. They are accelerated by the LINAC2 to an energy of 50 MeV. After a further acceleration in the booster (1.4 GeV) they are inserted in the Proton Synchrotron (PS, 25 GeV) and then sent to the SPS (Super Proton Synchrotron, 450 GeV). After that, they are injected into the LHC where they are accelerated for about 20 minutes to get their final energy of 7 TeV. They are brought to collision in one of the four interaction points [24]. A sketch of the LHC's preaccelerators is shown in Fig. 3.1.

The event rate is proportional to the cross section. The proportionality constant is named *luminosity* and is defined as number of particles per area and time:

$$\mathscr{L} = \frac{1}{\sigma}\frac{dN}{dt}. \tag{3.1}$$

The instantaneous design luminosity of the LHC is high ($\mathscr{L} = 10^{34}\frac{1}{\text{cm}^2\text{s}}$) compared the the luminosity at the Tevatron ($\mathscr{L} = 10^{32}\frac{1}{\text{cm}^2\text{s}}$). The protons are accelerated in 2,808 bunches of 10^{11} protons each. With a bunch spacing of 25 ns and 23 interactions per bunch crossing [25], this leads to 40 million collisions and a total amount of $\approx 10^9$ interactions per second. After restarting the LHC in October 2009, the LHC is planned to run for one year at a center of mass energy of 10 TeV.

3.2. The ATLAS Detector

The ATLAS detector is the largest experiment at the LHC. Due to the high interaction rate, the detector channels have to be read out very fast. Furthermore, the used material has to be radiation hard and the detector parts have to be highly segmented to allow tracking. ATLAS has two different types of magnets: a solenoid in the inner part of the detector and a toroid in the muon system. This multi-purpose detector has an onion-like structure with a total length of 44 m and a height of 25 m.
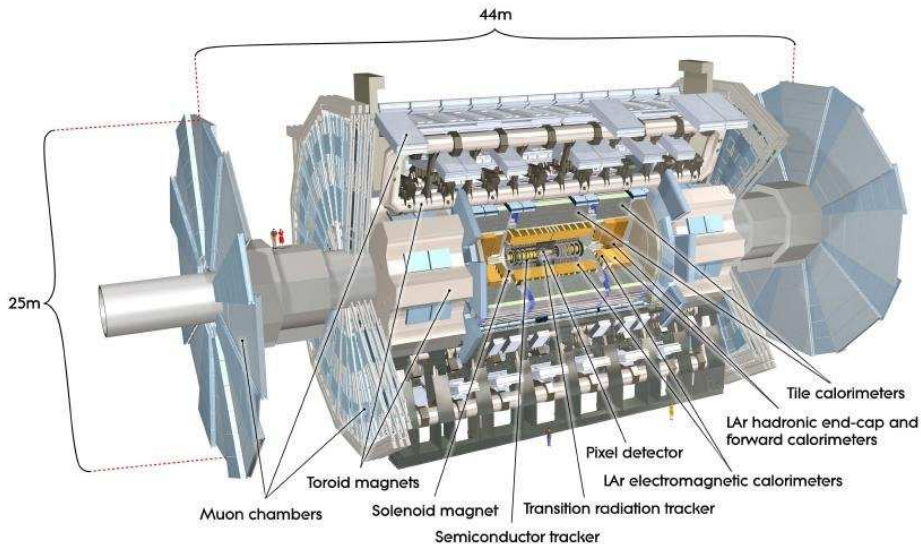

Figure 3.2: The ATLAS experiment and its subsystems, arranged in a typical onion-shell structure [26].

The innermost part consists of a tracking detector with three different subsystems: a pixel detector, a semiconductor tracker and a transition radiation tracker. This Inner Detector (ID) is contained in a solenoidal magnet. The curvature of charged particles and thus their momentum, is measured here. The energy of electrons, photons and jets are measured in calorimeters where they are stopped. Muons deposit only a small fraction of their total energy in these detectors. Their tracks are measured by the muon spectrometer, which is placed at the outer part of ATLAS.

The ATLAS coordinate frame and variables The ATLAS coordinate basis is right-handed with its origin in the interaction point. The x-coordinate points to the middle of the LHC whereas the y-coordinate points upwards. The z-component is the coordinate in the direction of the beam-pipe. The pseudorapidity, η, is defined as:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{3.2}$$

where θ is the polar angle in the rz-plane, ϕ is the azimuthal angle in the xy-plane. The transverse momentum, p_T, is defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}. \tag{3.3}$$

The energies of the particles are measured by summing up the energy within a cone of a finite size ΔR in the $\eta\phi$-plane:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.4}$$

Assuming the transverse momentum of the initial state particles to be zero, the transverse energy of the unobserved neutrino can be calculated from the four-momenta of the other final particles ($E_{T_{miss}}$). Energy can also be missing due to detector effects (energy loss in dead material, out of cone showering).

3.2.1. The Inner Detector

The Inner Detector (ID) is displayed in Fig. 3.3. It consists of a three layered pixel detector, a semiconductor tracker (SCT) and a transition radiation tracker (TRT). It is surrounded by a solenoidal magnet with a field strenght of 2 T. The design of the ID allows a measurement of the primary vertex and secondary vertices (b-tagging). It measures the particles in a range of $|\eta| < 2.5$.

b-tagging b quarks are heavy quarks with a mass of 5 GeV/c^2. Due to their large mass, b quarks decay within the detector. Their decay vertex is displaced from the primary vertex and can be used to indentify them. Although b-tagging is not used in this analysis, it could be used to associate *tagged* jets with b-quarks.

The Pixel Detector The pixel detector is a highly segmented semiconductor and built to achieve a good resolution for vertex measurements. The detector consists of 1,744 sensors with 47,232 pixels each [26]. Most of the pixels have a size of 50×400 μm^2 (10 % of the pixels have a size of 50×600 μm^2) and are connected with the sensors via bump bonds (made of In or PbSn). It has a total number of 80,4 million readout channels. The semiconductors consist of n-type

wafers with $n+$ doted electrodes [26]. The pixel detector is 1.3 m long and consists of three layers with radii of 50.5, 88.5 and 122.5 mm as well as three disks on each side with $z = 495$, 580 and 650 mm. Being this close to the interaction point, the pixel detector has to be radiation hard. It is designed to have a lifetime of about 10 years, whereas the innermost layer is planned to be replaced after 3 years of running. The spatial resolution is about 10 μm (in R-ϕ) and 115 μm (in z). To reduce thermal noise, the pixel detector is cooled down to a temperature of -5 - -10 °C [26].

The Semiconductor Tracker (SCT)   The SCT also consists of a barrel part with 2,112 modules and two endcaps. The modules are inclined by 40 mrad to measure both, R and ϕ. There are four stereo layers in the barrel region which are parallel to the beam pipe and two endcaps with nine disks each. The strips (with a pitch of 80 μm) have two sensors with a length of 6 cm. 15,912 sensors are being used, each of them being 285 μm thick. The spatial resolution of the different parts are shown in Tab. 3.1.
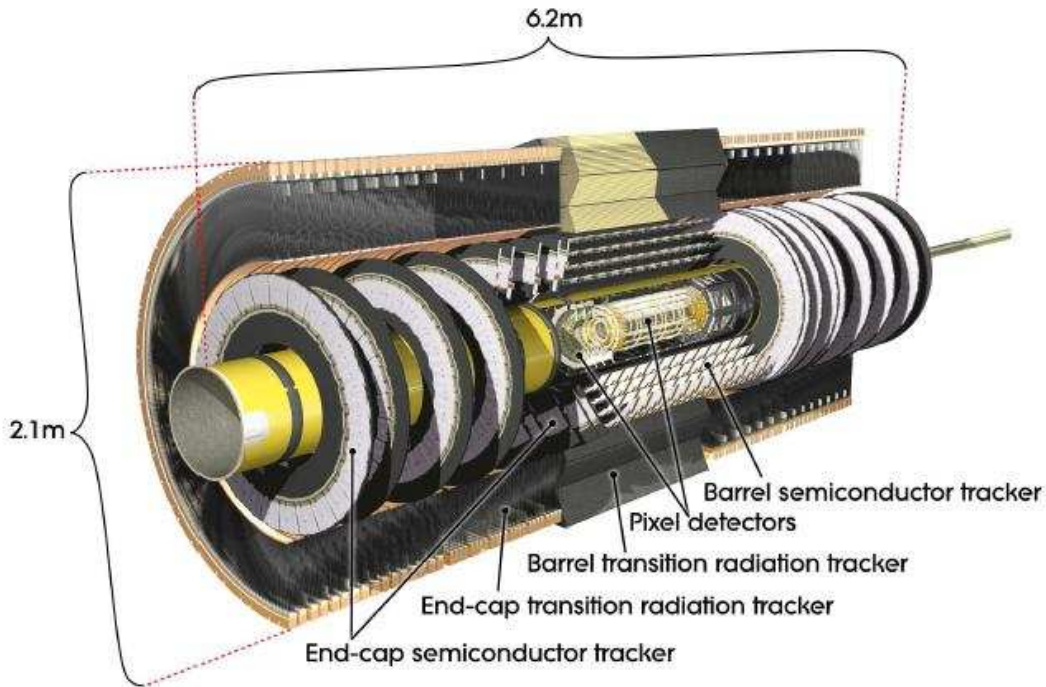
Figure 3.3: The structure of the Inner Detector with concentric layers in the barrel region and wheels in the forward direction [26].

The Transition Radiation Tracker (TRT)   The TRT comprises drift tubes (diameter: 4 mm) with a gold/carbon wire in the inside. The tubes are filled with a gas mixture of 70 % Xenon, 27 % CO_2 and 3 % O_2 at an overpressure of 5–10 mbar. The drift tubes are arranged in 36 layers parallel to the beam pipe. The barrel contains 96 modules with 52,544 straws each. The endcap tubes are oriented radially in twelve inner and eight outer wheels. The scintillating material is polypropylen. If a particle passes a region with a changing dielectric constant, photons are radiated at the boundary layers (see [27]). The particles and the generated photons ionize the gas. The drift time of electrons or ions can be measured and therefore the position of

the charge can be estimated. Electrons leave a stronger signal in the detector than pions, thus the detector is able to distinguish between electrons and pions. The spatial resolution is about 130 μm per straw. Since a maximum number of 36 space points can be measured in the barrel region per track, the TRT improves the measurement of the momentum. Compared to the Pixel Detector and SCT, the improvement is only small. In contrast to the pixel detector, the TRT works at room temperature [26].

Accuracy	Pixel Detector	SCT	TRT
barrel region: (R - ϕ) [μm]	10	17 (p. module)	130 (p. straw)
barrel region: (z) [μm]	115	580	
end cap wheels: (R - ϕ) [μm]	10	17	
end cap wheels: (z) [μm]	115	580	
space points per track	3	4	36

Table 3.1: Resolution of the subdetectors and the maximum number of available space points.

3.2.2. The Calorimeters

The Concept of Sampling Calorimeters

A sampling calorimeter consists of passive absorber material with high density (*e.g.* lead) and active material. The passive material interacts with the incoming particles and the active material is used to detect them. If electrons and photons enter the EM calorimeter, secondary particles are generated through pair production and bremsstrahlung (see [27]). These particles again produce further particles so that a *shower* develops in the detector. The shower development is interrupted when the particles have a critical energy, $E_{\text{crit.}}$, which is too small to create new particles. In a hadronic calorimeter, hadrons are produced until the particles reach the critical energy. The shower energy is measured in samples (hence the name "Sampling" Calorimeter). The energy is proportional to the number of particles in the shower, N, and follows therefore Poisson statistics [27]:

$$\frac{\sigma_E}{E}_{\text{sampl}} \propto \frac{1}{\sqrt{N}} \,. \tag{3.5}$$

In a simplified picture, the intrinsic energy resolution of such a calorimeter is:

$$\frac{\sigma_E}{E}_{\text{sampl}} = \frac{a}{\sqrt{E}}, \tag{3.6}$$

with a being a sampling term which depends on the material and the structure of the calorimeter.

The Electromagnetic Calorimeter

The ATLAS Electromagnetic Calorimeter (ECal) [26] is a sampling calorimeter with lead as passive and liquid argon (LAr) as active material. It has one barrel part and two endcaps. The η-coverage of the respective parts are shown in Tab. 3.4. The calorimeter has an accordeon-like shape which allows a measurement of the particle energy independent of the direction of the incoming particle. The signal is read out with copper wires in the LAr via capacitive couplings [26]. LAr does not only have an almost linear energy response. In addition, it is a radiation hard material [26]. The barrel is divided into two equal parts which have a distance of 4 mm around

the interaction point. Each part has 1,024 absorber plates which have the typical accordeon-like form. The half-barrel consists of 16 modules which are arranged around the beam pipe. The ECal is cooled via one barrel and two endcap cryostats to guarantee a constant temperature of 85 K.
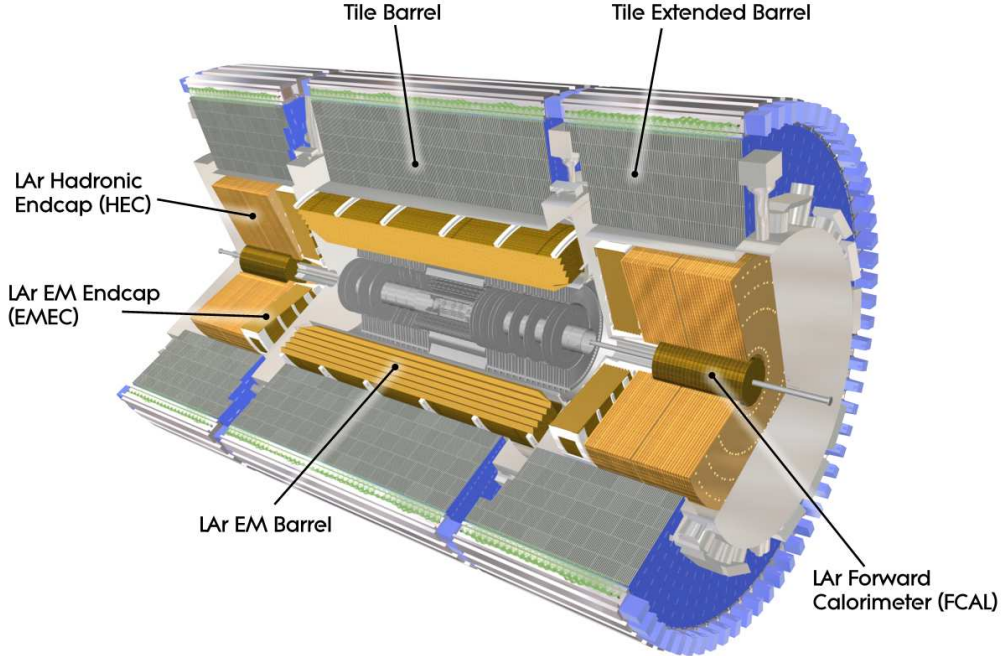
Figure 3.4: The calorimeter system of the ATLAS detector [26].

One characteristic quantity of an ECal is the radiation length, X_0. One radiation length is the distance an electromagnetic particle has to travel through a detector until its energy is decreased to $\frac{1}{e}$ of its original energy, E_0. The total thicknesses of the calorimeter parts are shown in Tab. 3.2 in units of X_0.

	η coverage	thickness
Barrel	$|\eta| < 1.475$	$> 24\ X_0$
Endcap (outer wheel)	$1.375 < |\eta| < 2.5$	$> 26\ X_0$
Endcap (inner wheel)	$2.5\ \ \ < |\eta| < 3.2$	$> 26\ X_0$

Table 3.2: Properties of the Electromagnetic Barrel and Endcap Calorimeter.

Charged particles transversing the detector produce secondary particles through bremsstrahlung, followed by e^+e^- pair production by the emitted photon. The number of produced particle generations in a shower, n, depends on the primary energy of the particle and the critical energy $E_{\text{crit.}}$ [27]:

$$n = \frac{\ln\left(\frac{E_0}{E_{\text{crit.}}}\right)}{\ln(2)}. \tag{3.7}$$

The same is true for the total number of particles in the shower maximum:

$$N_{\text{part.}} = 2^n = \frac{E_0}{E_{\text{crit.}}}. \tag{3.8}$$

The Moliere radius, R_{M}, (for LAr, $R_{\text{M}} \approx 10$ cm [27]) represents the transverse extension of an electromagnetic shower, 90 % of the shower energy is contained within a cylinder of the radius R_{M}:

$$R_{\text{M}} = 21 \ \text{MeV} \ \frac{X_0}{E_{\text{crit.}}}. \tag{3.9}$$

From these values, the expected energy resolution can be calculated. The energy resolution of the ATLAS ECal is [26]:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E(\text{GeV})}} \oplus 0.7\%. \tag{3.10}$$

The Hadronic Calorimeter

The hadronic calorimeter contains a barrel tile calorimeter and two extended barrels with a radius of 2.28 - 4.25 m. The corresponding η-regions are shown in Tab. 3.3. There are 64 modules in the azimuthal direction and three layers in the longitudinal direction. Between the barrel and the extended barrel is a gap for cables and services. To account for the energy lost in this gap, one uses the InterTileCal (ITC) which consists of scintillator plates.

The tile calorimeter is a sampling calorimeter with iron and 3 mm thick scintillator tiles which are read out via photomultiplier tubes (PMTs). The characteristic property of the HCal is the nuclear interaction length, λ. It is a measure for the probability for a particle to be absorbed in the detector. Of N particles traversing the detector, $\frac{N}{e}$ particles have not been absorbed after a distance of λ (for Pb: $\lambda = 18.5$ cm).

The HCal has to have a certain thickness to make sure that no hadrons enter the muon detectors. Due to material costs, the detector should be as small as possible. Therefore, the HCal has a maximum radial thickness of $\approx 11\lambda$.

The endcap calorimeters consist of two wheels on each side, segmented into 32 modules. The modules are made of Cu and LAr. The forward calorimeter, FCal, is also a sampling calorimeter with CuLAr (1 layer) and WLAr (2 layers). It covers the region close to the beam pipe ($3.1 < |\eta| < 4.9$) and has a thickness of 9.5 λ.

	η coverage	active material	passive material	thickness
Barrel	$|\eta| < 1.0$	Scintillator	Fe	9.2 λ
Extended barrel	$0.8 < |\eta| < 1.7$	Scintillator	Fe	
Endcap	until $|\eta| = 3.2$	LAr	Cu	

Table 3.3: Properties of the hadronic barrel calorimeter and the hadronic endcap calorimeter (HEC).

The HCal has the following resolutions [26]:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E(\text{GeV})}} \oplus 3.0\% \quad \text{(barrel and endcap)}, \tag{3.11}$$

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 10.0\% \quad \text{(forward)}. \tag{3.12}$$

While the hadronic showers become broader, the energy resolution decreases. Furthermore, the showers of several hadrons may overlap. Out-of-cone showering is one of the largest sources for systematic errors. To account for these effects, the jet energy scale has to be measured. The hadronic tile calorimeter has a longitudinal sementation of three samplings ([25]). Two of them have a granularity of 0.1×0.1 ($\Delta\eta \times \Delta\Phi$) and one of them has a granularity of 0.1×0.2. The hadronic endcap calorimeter has a longitudinal segmentation of four samplings with 0.1×0.1 ($1.5 < |\eta| < 2.5$) and 0.2×0.2 ($2.5 < |\eta| < 3.2$).

The particles entering the calorimeters have already lost a fraction of their energy in the ID. The thickness of the subdetectors can be seen in Fig. 3.5.
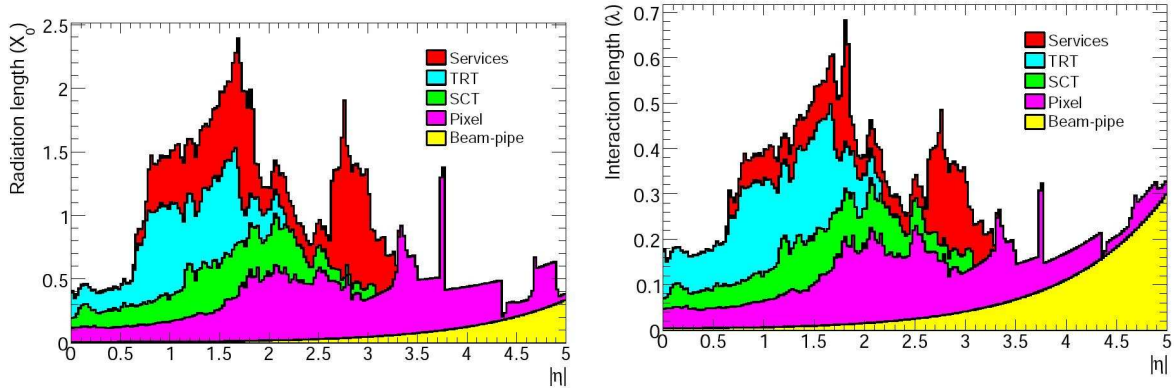

Figure 3.5: Thickness of the ID and services in radiation and interaction lengths depending on η [26].

3.2.3. The Muon Detectors

Muons leave energy in the detector only via the process of ionisation as described in the Bethe-Bloch equation. Muons with $\beta \times \gamma \approx 3.5$ are called minimum ionizing particles (MIPs). As MIPs, the muons only leave small fractions of their original energy in the detector. Therefore, the muon spectrometer is located at the outer part of the ATLAS detector. The so called stopping power is defined as the average energy loss per distance covered by the particle (Bethe-Bloch equation):

$$-\frac{dE}{dx} = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left(\frac{1}{2} \ln \left(\frac{2 m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right). \tag{3.13}$$

Z is here the atomic number of the absorber and A is the corresponding atomic mass, T describes the kinetic energy of the incident particle with a charge of z and I the mean excitation energy [28].
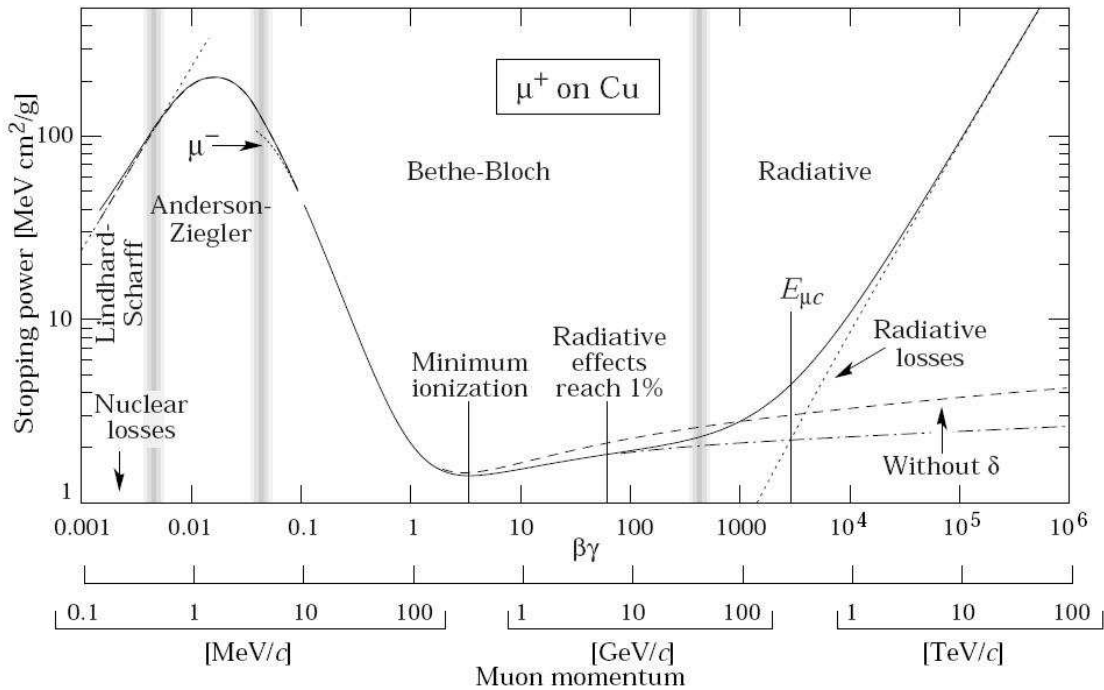
Figure 3.6: In this figure ([28]) the average stopping power of positively charged muons in Cu is shown. The stopping power is a function of $\beta\gamma$. In the different regions (see vertical lines) different approximations are used. MIPs have $\beta\gamma$-values of about three.

Muons are deflected in the solenoidal and toroidal magnetic fields. The sagitta of the track is measured in the solenoidal field to reconstruct the transverse momentum of the particles. The detector is designed to have a momentum resolution of 10 % if the track has a p_T of 1 TeV. Therefore, the uncertainty of the sagitta has to be smaller than 50 μm for a sagitta of 500 μm (see [26]). If the sagitta, s, is measured (see Fig. 3.7), the radius of the track, r, and thus the transverse momentum can be determined.

$$p_T = \underbrace{\frac{s^2 + \frac{d^2}{4}}{2s}}_{r} \, q \, B. \tag{3.14}$$
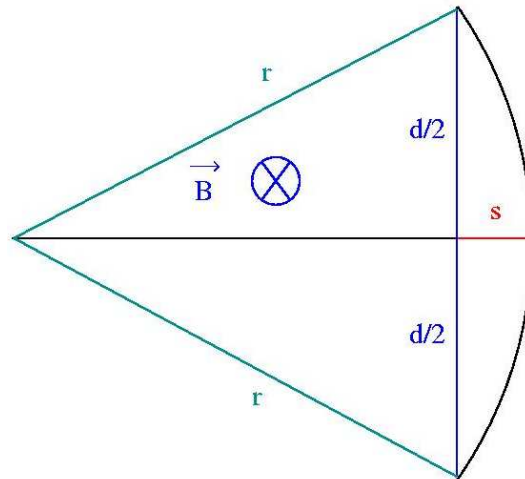
Figure 3.7: Measurement of the sagitta s for muons passing through a homogeneous magnetic field $\vec{B}$. s can be measured. The chord d and therefore the radius r can be calculated to determine the transverse momentum of the particle.

The muon system comprises different detector modules and a toroid magnet. The different detector modules are described in the following. Two coordinates of the muon tracks are measured: in the xy-plane (using Monitored Drift Tubes and Cathod Strip Chambers) and in the yz-plane (with the Resistive Plate Chambers and the Thin Gap Chambers).
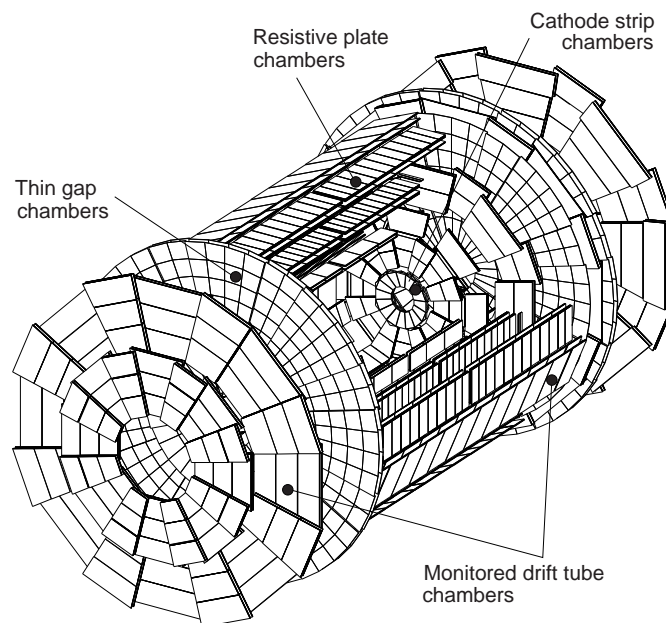
Figure 3.8: Sketch of the muon spectrometer with its different detector types in the barrel and forward region.

Monitored drift tubes (MDT) The monitored drift tubes cover the η - region up to 2.7. Each of the 1088 chambers has 3–8 layers of aluminium drift tubes which are filled with Argon (93 %) and CO_2 (7 %) at a pressure of three bar [26]. While the tube itself has a radius of 14.985 mm, the wire (Tungsten-Rhenium) has a radius of 25 μm. The spatial resolution is about 80 μm per tube and about 35 μm per chamber. The electrons in the MDT have a maximum drift time of 700 ns. The tubes are several meters long and have a sag of a few 100 μm. The MDTs are arranged around the beam pipe and form three layers with different radii.

Cathod strip chambers (CSC) The 32 CSCs are multi-wire proportional chambers located in the region of $2.0 < |\eta| < 2.7$ (7 m from the interaction point) and have a time resolution of 7 ns. The spatial resolution in the bending plane is much higher (40 μm) than in the transverse plane (5 mm). The wires are made out of Tungsten and Rhenium and have a diameter of 30 μm. The chambers are filled with Argon (80 %) and CO_2 (20 %) and form two disks with eight large and eight small chambers each. They overlap (see Fig. 3.8) to achieve full coverage in the azimutal angle ϕ.

Trigger Chambers The trigger chambers can be divided into resistive plate chambers (RPC) and thin gap chambers (TGC). The 544 RPCs are placed in the barrel part with $|\eta| < 1.05$ whereas the 3,588 TGCs cover $1.05 < |\eta| < 2.7$. They are used to trigger events containing muons and to measure the coordinate perpendicular to the radial direction. In the region of $1.3 \leq |\eta| \leq 1.65$, the field of the toroids in the barrel and in the endcaps overlap and form an inhomogeneous field. Due to this, the granularity and the design of this part depends of the pseudorapidity.

Resistive plate chambers (RPCs) The RPCs consist of two plates. The space between the plates is filled with a gas mixture of $C_2H_2F_4$ (94.7 %) Iso-C_4H_{10} (5 %) and SF_6 (0.3 %) and are read out via metallic strips.

Thin gap chambers (TGCs) The TGCs are also multi-wire proportional chambers filled with CO_2 (55 %) and n-C_5H_{12} (45 %). The wires in the TGCs have, adapted to η, a diameter of 10.8 - 55.8 mm. Both trigger chambers have a very good time resolution (RPC: 1.5 ns, TGC: 4 ns). Particles traversing the trigger chambers can ionize the gas-mixture. The electrons or ions drift to the wire. The position of the particle can be derived from the drift time. The time resolution has to be very good. If the particle needs too much time to reach the wire, new signals can occur after the first signal and it is not possible to assign the signal to the correct event.

3.2.4. The Magnets

The magnet system of the ATLAS detector comprises a solenoid magnet and a toroid magnet. As already shown in Fig. 3.7, they are used to deflect charged particles in order to measure their longitudinal momentum component and to identify the sign of their charge.
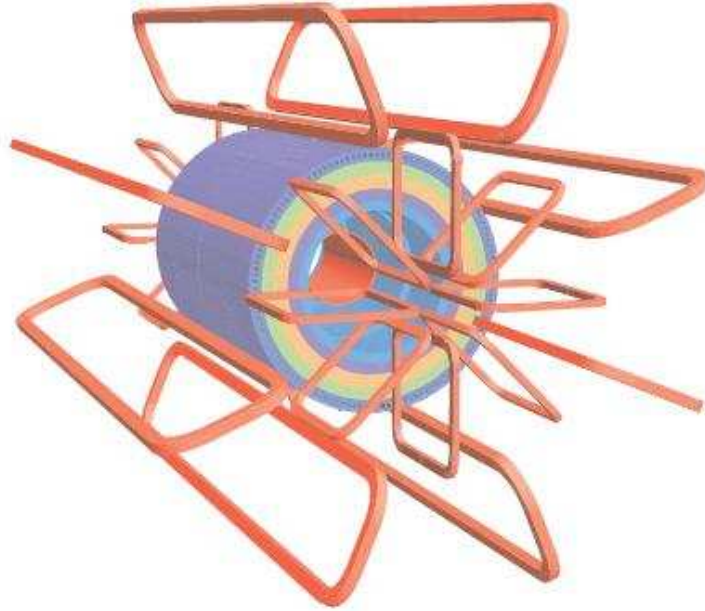
Figure 3.9: The magnet system of ATLAS, comprising a solenoid and a toroid magnet.

The solenoid magnet   The solenoid is built around the ID, 5.8 m long and has an outer radius of 1.28 m. The magnetic field has a strength of 2 T with a peak value of 2.6 T. The coils have a total weight of 5.4 tons and are connected with superconducting cables. With an operational current of 7,730 A, a total energy of 40 MJ can be stored in the magnetic field [26].

The toroid magnet   The toroid is divided into one barrel toroid and two endcaps where each of them has eight coils. The whole barrel system is 25.3 m long and has an outer radius of about 10 m. The used material for the coils is Nb/Ti/Cu. The superconducting coils are placed in a cryostat (4 K). The magnetic field of the toroid has a strength of 4 T (peak value: 4.7 T). With a current of 21,000 A, this magnet can store an energy of 1 GJ. It has an air core to reduce possible multiple scattering of the particles.

3.2.5. Trigger and Data Aquisition

As already discussed in the introduction the protons interact at a rate of 1 GHz at the LHC. The huge amount of data related to this interaction rate cannot be stored. Thus, efficient criteria are necessary to select only the interesting events. The event selection is done via a trigger consisting of three different stages: the Level 1 trigger (L1), and the High-Level trigger (HLT) which is divided into the Level 2 trigger (L2) and the Event Filter (EF).
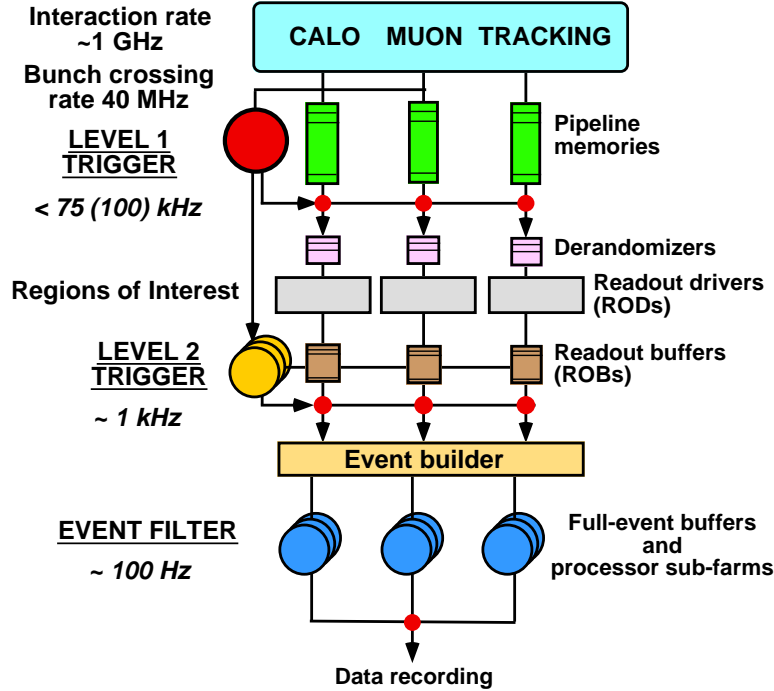
Figure 3.10: The three different stages of the trigger system

Level 1 trigger   The L1 gets information from the muon trigger chambers and subsystems of the calorimeters. It is sensitive to muons with high transverse momentum, large $E_{T_{miss}}$, large overall transverse energy ($\sum E_T$) as well as e^-/γ and jet signals. The decision has to be made very fast ($< 2.5\ \mu$s). Hence, only part of the information is used. After accepting the event, special regions of interest (ROI) are declared in the detector and the information is sent to the L2. The overall acceptance rate is max 75 kHz.

The High Level trigger   The HLT consists mainly of a large processing farm. The L2 reads all data available in the ROI and needs ≈ 40 ms to make its decision. In contrast to the L1, the L2 takes all information from the precision chambers (MDT and CSC) as well as the data from the tracking system into account. The data is reduced further to 1–3.5 kHz [26].
The complete event is being built in the Event Filter. Offline analysis is used to classify the events which takes about 4 s per event. If it is accepted, it is distributed to storage elements all over the world to prevent loss of the data and to make it accessible for the community over the GRID. Only 10 % of the data from the L2 stage is finally stored.

4. The Matrix Element Method

The *Matrix Element Method* is a useful tool to estimate parameters based on probability density distributions. The method has been developed by the DØ collaboration to provide a good estimator for the top quark mass [29]. A probability density is calculated with respect to parameters $\vec{\alpha}$ for each event. This probability density consists of a signal and a background fraction (f_{sig}, f_{bkg}).

For the generation of top quark pairs, the signal probability is proportional to the differential cross section of the $t\bar{t}$-production. The background probability depends on the production cross section of a semileptonic decaying $W + 4$ jets and is discussed in Chapter 4.5. QCD background is not considered in this analysis while the fraction of this background is difficult to estimate from Monte Carlo production. A probability density for the parameters under study, $\vec{\alpha}$, is calculated for each event for the signal and background hypotheses and weighted by the relative fractions:

$$P_{\text{ev}}\left(\vec{x}; \vec{\alpha}\right) = f_{\text{sig}} P_{\text{sig}}\left(\vec{x}; \vec{\alpha}\right) + f_{\text{bkg}} P_{bkg}(\vec{x}) \, , \tag{4.1}$$

with $f_{bkg} = 1 - f_{sig}$. The calculation is done while using the measured momenta and energies represented by $\vec{x}$. Each event probability depends on a set of parameters $\vec{\alpha}$. The Likelihood is defined as the product of event probabilities:

$$L\left(\vec{x_1}, ..., \vec{x_n}; \vec{\alpha}\right) = \prod_{i=1}^{n} P_{ev}\left(\vec{x_i}; \vec{\alpha}\right), \tag{4.2}$$

$$\ln \, L\left(\vec{x_1}, ..., \vec{x_n}; \vec{\alpha}\right) = \sum_{i=1}^{n} \ln \, P_{ev.}\left(\vec{x_i}; \vec{\alpha}\right) \tag{4.3}$$

To find the best set of parameters for a number n of events, the maximum value of the Log-Likelihood has to be found.

4.1. Calculation of the signal probability based on the differential cross section

First, the signal probability density is calculated which is proportional to the differential production cross section ([30], [31]) . The four-momenta of the initial partons are represented by Q_1 and Q_2 whereas the final state particles are represented by p_i. Fermi's *Golden Rule* is used to calculate the cross section of the hard scattering process:

$$d^n \hat{\sigma}_{\text{hs}}(\vec{x}; \vec{\alpha}) = \frac{(2\pi)^4 |\mathscr{M}(\vec{x}; \vec{\alpha})|^2}{4\sqrt{(Q_1 \cdot Q_2)^2 - m_1^2 m_2^2}} d\,\text{LIPS} \tag{4.4}$$

The matrix element, $\mathscr{M}$, is a measure for the transition probability between the initial and the final states. The matrix element is a function of the four-vectors $\vec{x}$ of the final state particles and a set of parameters, $\vec{\alpha}$. d $LIPS$ is the Lorentz Invariant Phase Space factor and the denominator is the flux factor.

$$d\,LIPS(Q_1 + Q_2; p_1, ..., p_n) = \delta^4(Q_1 + Q_2 - \sum_{i=1}^{n} p_i) \prod_{i=1}^{n} \frac{d^3\,p_i}{(2\pi)^3\,2\,E_i}. \tag{4.5}$$

Equation (4.4) only holds true for the hard scattering process of partons (gluons and quarks). In order to obtain the cross section for protons, $\hat{\sigma}_{\text{hs}}$ has to be convolved with the PDFs mentioned in chapter 2:

$$d^n\sigma(\vec{x}; \vec{\alpha}) = \sum_{flav} \int_{x_1} \int_{x_2} d^n \hat{\sigma}_{\text{hs}}(\vec{x}; \vec{\alpha}) d\,x_1 d\,x_2\, x_1\, f_{flav}(x_1)\, x_2\, f_{flav}(x_2), \tag{4.6}$$

All possible flavours have to be taken into account since it is not possible to reveal the identity of the partons. The total cross section of the process, $\sigma(\vec{\alpha})$, is obtained by integrating $d^n\sigma(\vec{x}; \vec{\alpha})$ over the given phase space:

$$\sigma(\vec{\alpha}) = \int d^n\sigma(\vec{x}; \vec{\alpha}). \tag{4.7}$$

4.1.1. Acceptance factors

In general, additional normalization factors have to be applied to account for the acceptance of top quark events. The acceptance increases with increasing top quark mass. This can be explained by the energies of the final state particles. The decay products of top quarks with lower masses are less central due to lower p_{T}. Therefore, less events are detected and the acceptance decreases. Since only samples of NLO events at two different top pole masses are treated in this analysis, it is not possible to derive the acceptance factors from the ATLAS MC as a function of the top quark mass.

4.1.2. Transfer functions

Furthermore, the energies and momenta cannot be measured perfectly due to the resolution of the detector. This resolution is accounted for by the transfer functions, $W(\vec{x}, \vec{y})$, which describe the mapping between true ($\vec{x}$) and measured ($\vec{y}$) quantities. This mapping does not only depend of the type of particle (leptons, light quarks, b quarks) but also on the η-region of the detector (see Chapter 5).

Including these effects, the signal probability density yields:

$$p(\vec{x}) = \frac{1}{\int d\vec{y}\, d^n\sigma(\vec{x}, \vec{\alpha})\, Acc(\vec{x})} \int d^n\sigma(\vec{x}, \vec{\alpha})\, W(\vec{x}, \vec{y})\, d\vec{y} \tag{4.8}$$

The calculation is done for a single parameter ($\vec{\alpha} = m_{\text{top}}$) first and later extended by two different jet energy scales. Since the top pair production via $q\bar{q}$-annihilation covers only 10% of the total production rate, only the top pair production via gluon fusion is considered in the following.

4.2. Integration variables and techniques

Energy and momentum of the initial and final state particles have to be measured. While it is only possible to measure the momenta of the jets and of the charged lepton, the following assumptions and approximations are made in order to perform the integration.

- The p_T of the $t\bar{t}$-system is assumed to be zero. With this assumption, the transverse momentum of the neutrino can be calculated. The absence of initial and final state radiation is valid for calculations in leading order (LO) but not for higher order corrections.

- While the initial particles are assumed to have zero transverse momentum, the longitudinal momentum of the gluons can be calculated via energy and momentum conservation.

- The angles of the charged lepton and the jets are assumed to be perfectly measured.

The dimension of the integration is further reduced to four jet-energies, one lepton energy and the mass of the W boson (or equivalent: p_z^ν). Performing an integration over m_W, the two solutions for the p_z of the neutrino can be determined (Chapter 4.4). In the case of the mass integration, the differential cross sections for both solutions are calculated and added.

Since it is not possible to unambiguously assign the jets to the partons, 24 combinations are possible and have to be taken into account. This number can be reduced by a factor of 2 if the matrix element is symmetrized. The remaining combinations have to be summed and averaged. The numerical integration is done with the Bayesian Analysis Toolkit BAT [32]. This toolkit contains several integration methods. The one used here is the sampling with Markov Chain Monte Carlo (MCMC) which is explained in more detail in Chapter 6.

4.3. Matrix element calculations with MADGRAPH

The events used in the first part of the analysis are produced with MADGRAPH [33, 34]. MADGRAPH is a leading order (LO) event generator. Various processes are defined, such as SM processes as well as a selection of SUSY processes. Event generation can be done online or with an offline package. With the SM online calculator, the input files with the parameters (particle masses, widths and couplings) are generated. Internal cuts of the produced particles can be specified as well as the center of mass energy and the PDF sets. MADGRAPH uses the LO PDF CTEQ 6L1 which is also included in the program via the C++ - wrapper of the Les Houches Accord [35, 36].

MADGRAPH is not only used for event generation but also for matrix element calculations in the analysis itself. Therefore, an interface to MADGRAPH has been written. This interface needs the four momenta of the initial and final state particles as well as the top pole mass hypothesis and its corresponding width as input parameters. It uses the offline package of MADGRAPH in which only the mentioned parameters are changed according to the event kinematics. The helicity amplitudes are calculated while taking the colour factors into account. The squared matrix element is returned. The matrix element contains the propagator for the top quark which depends on the width of the top quark. Therefore, the top width has to be calculated as a function of the top quark mass.

Since the width of the top quark increases with increasing mass, it has to be calculated in each integration step. This is done via a parametrization. The online Standard Model calculator of MADGRAPH has been used to calculate the width as plotted in Fig. 4.1. The parametrization has been performed via a 2nd order polynomial (Fig. 4.2 (d)).
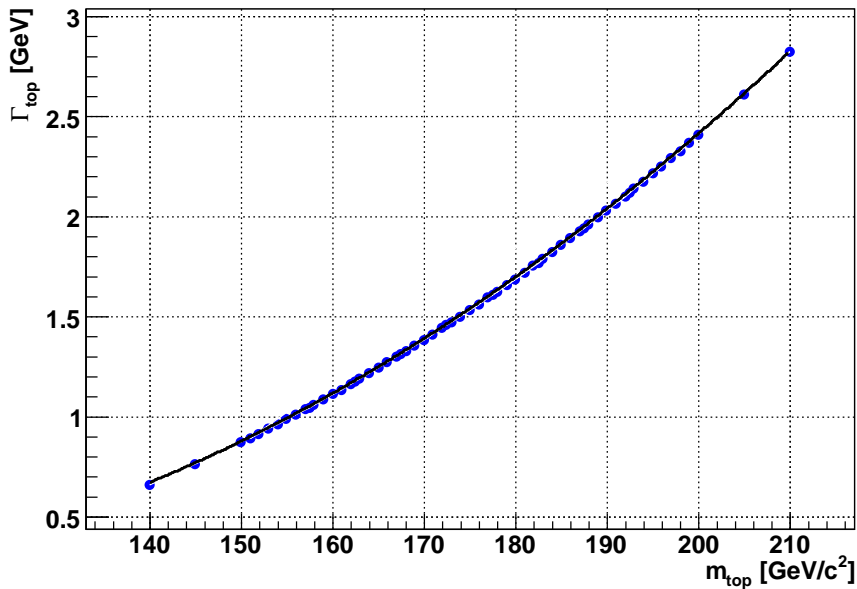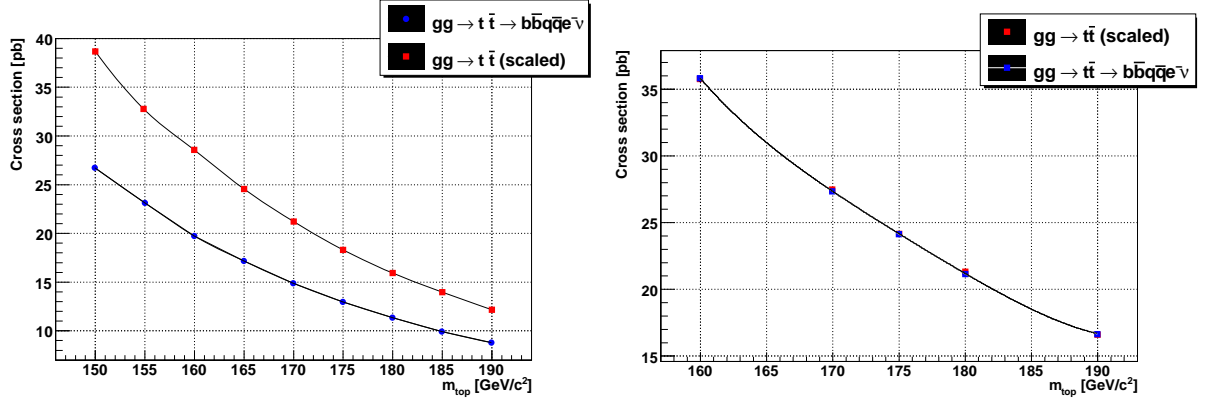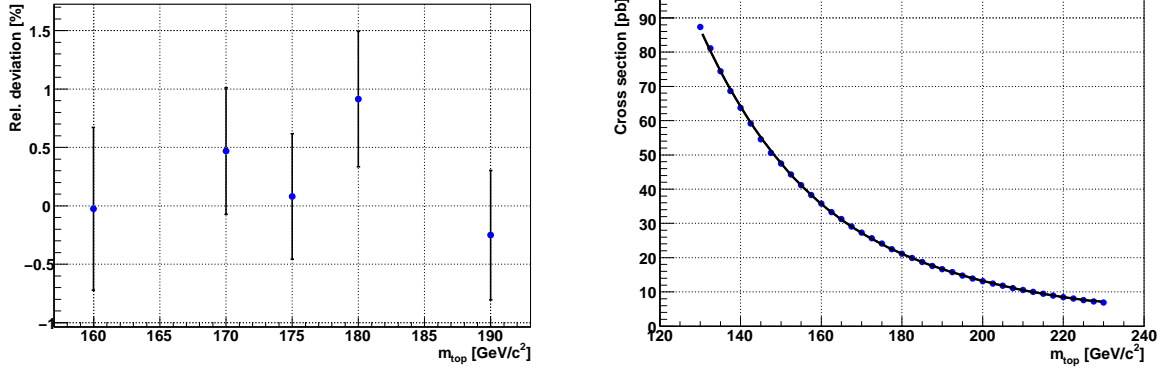

Figure 4.1: Parametrization of the top width Γ_{top} as a function of the top quark mass. The widths have been calculated with the MADGRAPH offline package.

The MADGRAPH offline package (version 4.4.5) has been used to calculate the total cross section of the process which is needed for the normalization of the signal probability.

(a) Cross sections from MADGRAPH with internal cuts

(b) Cross sections from MADGRAPH after removing all cuts

(c) Relative deviation of the cross section decay channel (blue) from the scaled total cross section (red).

(d) Parametrization of the total cross section

Figure 4.2: The total cross section of the $t\bar{t}$ production is parameterized as a function of the top quark mass with a fourth order polynomial. The errors are too small to be seen. A comparison of the scaled total cross section with the cross section of semileptonic decays reveals a bug in the calculation (a). The results of the corrected calculation is shown in (b) while the consistency of the two cross sections are displayed in (c). Fig. (d) shows the calculation of the cross section in steps of 2.5 GeV/c^2.

The total cross section for the process $gg \rightarrow t\bar{t}$ and the channel $gg \rightarrow b\bar{b}\,q\bar{q}\,e^-\,\nu_e$ are shown in Fig. (4.2 (a)), while the total cross section (upper curve) is scaled down to the branching ratio of the channel displayed in blue (lower curve). It is obvious that the results are inconsistent and the error is too small to be seen (Fig. 4.2 (a)). This can be explained with the internal cuts applied by MADGRAPH on η, p_T and ΔR. After removing the internal cuts, the cross sections agree (Fig. 4.2 (b)). The relative deviation of the two cross sections is shown in Fig. 4.2 (c). The deviations are consistent within the errors. Figure 4.2 (d) shows the values for the total cross section as a function of the top quark mass in steps of 2.5 GeV/c^2. A parametrization has been found using a fourth order polynomial.

4.4. Phase space transformations

Two different sets of integration variables are used in this analysis. First, the general phase space transformation with an integration over the five energies (see Chapter 5.1) and the longitudinal momentum of the neutrino is explained. Later, a transformation from $p_{z,\nu}$ to the W mass has been done. The comparison of these different integration variables is shown in Appendix C.

4.4.1. Integration over p_z^ν

The differential cross section is now integrated and convolved with the parton density functions and the transfer functions W. Here, x_1 and x_2 are the momentum fractions of the two gluons (p_1 and p_2), $\vec{x}$ are the true and $\vec{y}$ are the measured quantities.

$$\mathrm{d}^\mathrm{n}\sigma(\tilde{\mathrm{x}}; \mathrm{m_{top}}) = \int \mathrm{d}\,\mathrm{x}_1\,\mathrm{d}\,\mathrm{x}_2 \frac{(2\pi)^4 |\mathscr{M}(\tilde{\mathrm{x}}; \mathrm{m_{top}})|^2}{4\sqrt{(Q_1 \cdot Q_2)^2 - \mathrm{m}_1^2\,\mathrm{m}_2^2}}\, \mathrm{x}_1\,\mathrm{f_{fl}}(\mathrm{x}_1)\,\mathrm{x}_2\,\mathrm{f_{fl}}(\mathrm{x}_2)\,\mathrm{dLIPS}\,\mathrm{W}(\tilde{\mathrm{x}}, \tilde{\mathrm{y}}) \quad (4.9)$$

The masses of the initial state particles are negligible. Thus, the equation can be simplified to:

$$\mathrm{d}^\mathrm{n}\sigma(\tilde{\mathrm{x}}; \mathrm{m_{top}}) = \int \mathrm{dx}_1 \mathrm{dx}_2 \frac{(2\pi)^4 |\mathscr{M}(\tilde{\mathrm{x}}; \mathrm{m_{top}})|^2}{4(Q_1 \cdot Q_2)}\, \mathrm{x}_1\,\mathrm{f_{fl}}(\mathrm{x}_1)\,\mathrm{x}_2\,\mathrm{f_{fl}}(\mathrm{x}_2)\,\mathrm{dLIPS}\,\mathrm{W}(\tilde{\mathrm{x}}, \tilde{\mathrm{y}}) \quad (4.10)$$

The Lorentz-invariant phase space factor $dLIPS$ contains the 4-momentum conservation. As one can see in equation (4.12), energy and longitudinal momentum of the initial particles p_1 and p_2 can be calculated from this.

$$\mathrm{dLIPS} = \delta^4(Q_1 + Q_2 - \sum_{i=1}^{6} p_i) \prod_{i=1}^{6} \frac{\mathrm{d}^3\mathrm{p_i}}{(2\pi)^3 2E_i} \quad (4.11)$$

$$= \delta(E_{p_1} + E_{p_2} - \sum_{i=1}^{6} E_i)\delta(p_z^{p_1} + p_z^{p_2} - \sum_{i=1}^{6} p_z^i) \prod_{i=1}^{5} \frac{|\vec{p_i}|\,d\,E_i}{(2\pi)^3 2} \cdot \frac{d\,p_z^\nu}{(2\pi)^3 2\,E_\nu} \quad (4.12)$$

Here, p_1 and p_2 describe the initial state particles. Hence, the integration over x_1 and x_2 is an integration over δ functions. The transformation from $\mathrm{d}^3\mathrm{p_i}$ to $\mathrm{dE_i}$ can be found in Appendix A.

$$\mathrm{d}^\mathrm{n}\sigma(\tilde{\mathrm{x}}; \mathrm{m_{top}}) = \sum_{\text{flavour}} \int \frac{(2\pi)^4 |\mathscr{M}(\vec{y}; m_{top})|^2}{4\sqrt{(Q_1 \cdot Q_2)^2}}\, x_1\,f_{fl}(x_1)\,x_2\,f_{fl}(x_2) \frac{\mathrm{dp}_z^\nu}{(2\pi)^3 2\,E_\nu} \quad (4.13)$$

$$\times \prod_{i=1}^{5} W(E_i^{\text{meas}}, E_i) \frac{|\vec{p_i}|}{(2\pi)^3 2}\mathrm{d}\,\mathrm{E_i} \quad (4.14)$$

The numbering scheme for the transfer functions W_i is as follows: the numbers 1-4 describe the transfer functions and variables for the jets while number 5 describes the ones for the lepton.

4.4.2. Integration over m_W

It is also possible to integrate over the mass of the W boson. The procedure mentioned above can be repeated with only two changes: the first change is the factor dLIPS. For this integration, an additional transformation has to be made. The calculation can be found in Appendix A. Furtheron, the equation provides two solutions for the longitudinal neutrino momentum. Consequently, we have to calculate two solutions of $d^n\sigma(\vec{x}; m_{top})$. While these two solutions are correlated, the sum has to be taken for further calculations.

$$\text{dLIPS} = \delta^4(Q_1 + Q_2 - \sum_{i=1}^{6} p_i)\prod_{i=1}^{6}\frac{d^3 p_i}{(2\pi)^3 2E_i} \tag{4.15}$$

$$= \delta(E_{p_1} + E_{p_2} - \sum_{i=1}^{6} E_i)\delta(p_z^{p_1} + p_z^{p_2} - \sum_{i=1}^{6} p_z^i)\prod_{i=1}^{5}\frac{p_i dE_i}{(2\pi)^3 2}\frac{m_W}{(2\pi)^3(p_z^{\nu_i} E^l - p_z^l E^{\nu_i})}dm_W \tag{4.16}$$

4.5. The Background probability

Only W + jets background is taken into account in this analysis. The calculation could be performed similar to the signal process. However, the calculation of the matrix element for the W + 4 jets events includes a huge number of Feynman diagrams (in LO) and, as a result, is CPU time consuming. Therefore, a different approach was chosen.

The same matrix element as for the signal probability is used for the W + jets events. The bias on the top quark mass estimation for a varying number of background events is studied. As mentioned earlier, it is not possible to model the QCD multijet background in a proper way. Although the QCD background will dominate at the LHC, it cannot be included in our background calculations.

4.6. Description of the NLO Monte Carlo data

Signal events

The signal sample used in Chapter 6.3 is generated by the MC@NLO matrix element calculator ([37, 38]) with parton showering from Herwig [39]. Detector simulation is done with GEANT 4 ([40, 41]). Events can be selected according to the needs and stored in a flat ROOT n-tuple with TopView [42]. The data is stored in a tree and contains events from the semileptonic and dileptonic channel (about 461 pb). Although matching information is available in the tree it is not used in the analysis (see Chapter 6).

Background events

The background sample consists only of $W(e^-\nu)$ + 4 jets events and is generated with Alpgen [43]. The full simulation is done with HERWIG [39]. The total cross section for this sample is 45.303 pb. Unfortunately, only 111.84 pb^{-1} is available. As is shown in Chapter 6.3.1, only few background events survive the event selection. Therefore, detailed studies of the event purity are not possible.

5. The Transfer Functions

As already mentioned in Chapter 4, the detector resolution has to be taken into account when calculating the probability densities. The detector resolution is represented by the transfer functions $W(E_{\text{true}}, E_{\text{meas}})$ which give the probability that E_{true} is the true energy of the particle when E_{meas} was measured.

5.1. Gaussian transfer functions

Gaussian functions are used for a first tests of the implementation of the Matrix Element Method. According to the ATLAS TDR [26]. The electromagnetic calorimeter has a resolution of $\frac{10\%}{\sqrt{E}}$ and the hadronic barrel calorimeter has a resolution of $\frac{50\%}{\sqrt{E}}$. Hence, the transfer function for the electrons is a Gaussian with a σ of $0.1\sqrt{E(\text{GeV})}$ and the transfer functions for the jets have a σ of $0.5\sqrt{E(\text{GeV})}$.
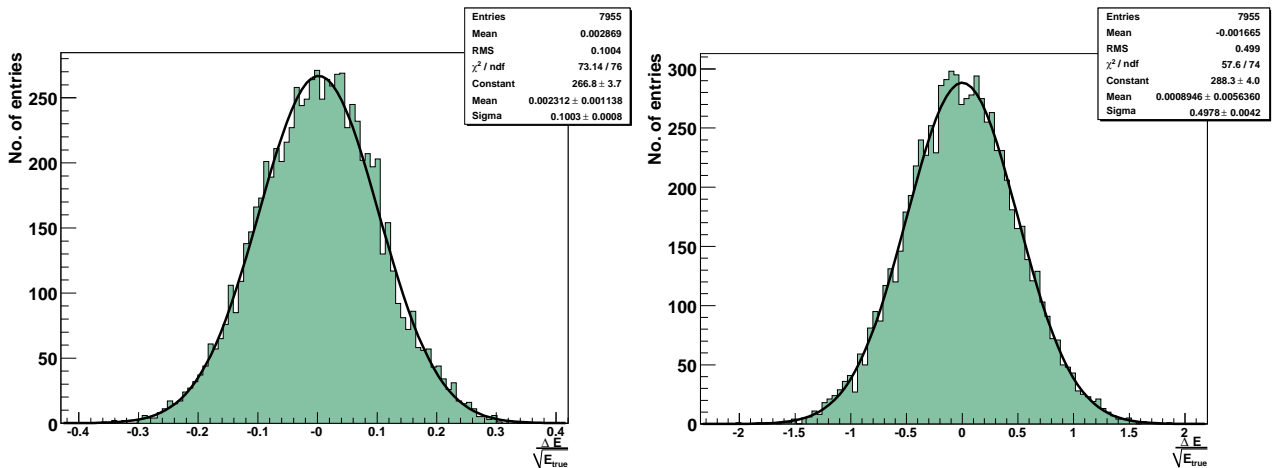

Figure 5.1: Gaussian transfer functions for electrons (left) and partons (right).

5.2. Detector resolution from ATLAS Monte Carlo

The assumption of symmetric transfer functions is not valid for a real detector. Particles passing regions of small pseudorapidity have passed less material than particles in high η regions. Thus, the resolution of the calorimeter increases with increasing η. Furthermore, particles with a large energy leave a broad shower in the detector. In contrast to the Gaussian functions in the previous Chapter, the real transfer functions are not symmetric. To get a more realistic estimation of the calorimeter resolution, the ATLAS-MC (NLO) is used to extract the transfer functions. In Fig. 5.2, one can see the difference between the true and the reconstructed energy for light jets and b-jets as well as for two different η regions.

One has to differentiate between light quarks and b quarks while the fragmentation functions of these two types are different. The parton and jets are matched with a ΔR criterion of 0.4. The distributions are shown for parton energies from 90 to 120 GeV.
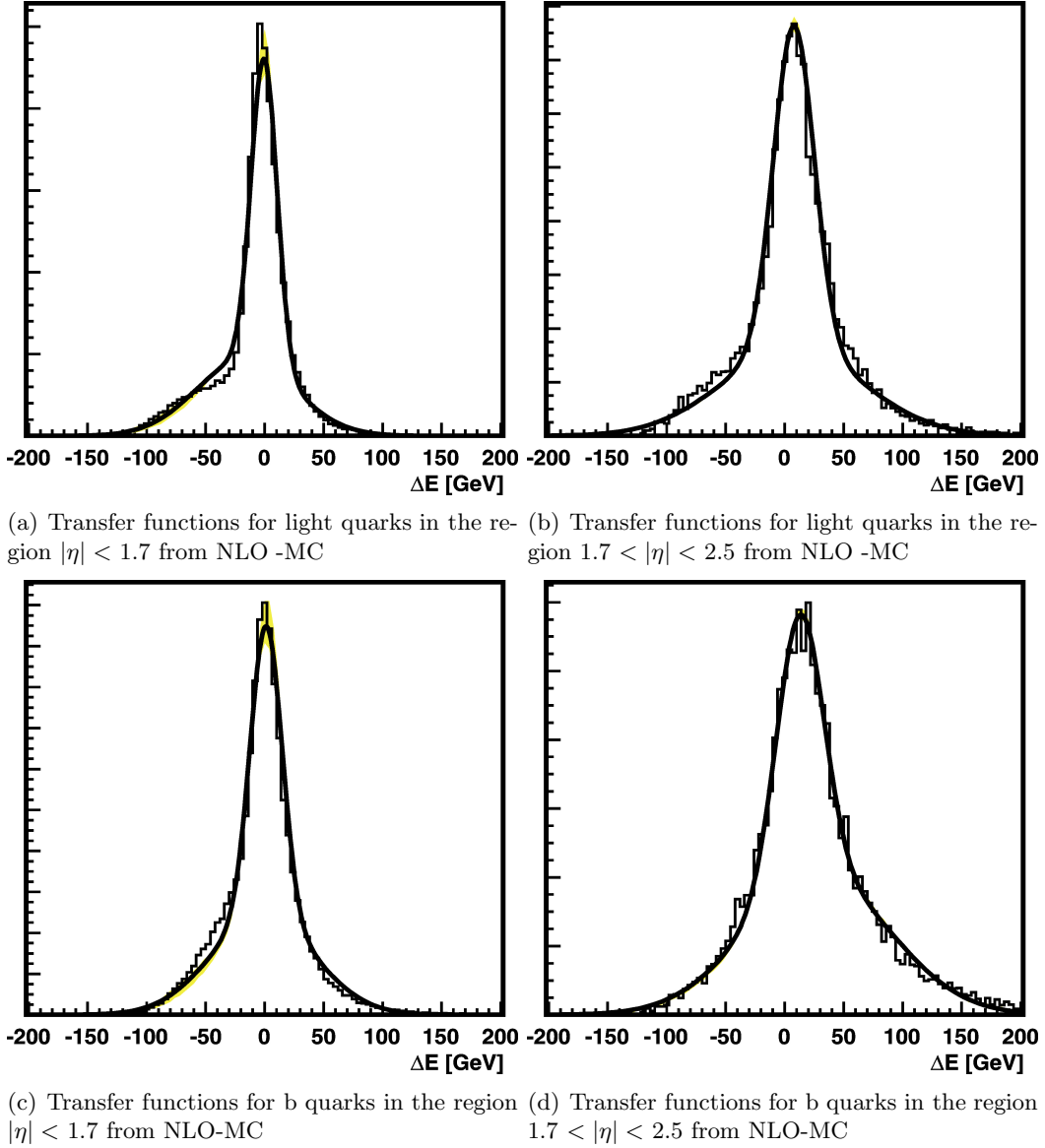

(a) Transfer functions for light quarks in the region $|\eta| < 1.7$ from NLO -MC

(b) Transfer functions for light quarks in the region $1.7 < |\eta| < 2.5$ from NLO -MC

(c) Transfer functions for b quarks in the region $|\eta| < 1.7$ from NLO-MC

(d) Transfer functions for b quarks in the region $1.7 < |\eta| < 2.5$ from NLO-MC

Figure 5.2: The transfer functions are shown for light and b quarks and for two different η regions.

To describe these energy resolutions properly, a good parametrization has to be found. The parametrization which is chosen in this analysis has been used in Tevatron analyses [30, 31]. The sum of two Gaussians is used whereas the second Gaussian accounts for the asymmetric tails:

$$W(E_{\text{part}}, E_{\text{jet}}) = \frac{1}{\sqrt{2\pi}(p_2 + p_3\,p_5)} \left[\exp(-\frac{(\Delta E - p_1)^2}{2\,p_2^2}) + p_3\,\exp(-\frac{(\Delta E - p_4)^2}{2\,p_5^2}) \right] \qquad (5.1)$$

with $\Delta E = E_{\text{jet}} - E_{\text{part}}$ and $p_i = a_i + b_i\,E_{\text{part}}$.

A total number of 40 parameters was estimated from MC events:

	light quarks		b quarks	
p_i	a_i	b_i	a_i	b_i
p_1	4.03668	-0.0359472	6.9543	-0.0450761
p_2	5.73034	0.0417829	7.89639	0.0444958
p_3	0.248581	0.0	0.222082	0.000874454
p_4	15.3105	-0.253933	29.3849	-0.248909
p_5	12.6232	0.234	16.9562	0.21453

Table 5.1: The parameters for the transfer functions in the central region $|\eta| < 1.7$.

	light quarks		b quarks	
p_i	a_i	b_i	a_i	b_i
p_1	12.2766	-0.035244	17.2693	-0.0272855
p_2	13.9903	0.0259577	15.8211	0.026291
p_3	0.261761	0.0	0.515255	0.0
p_4	24.6685	-0.142894	55.4544	-0.176426
p_5	46.9616	0.105927	44.1937	0.10971

Table 5.2: The parameters for the transfer functions in the central region $1.7 < |\eta| < 2.5$.
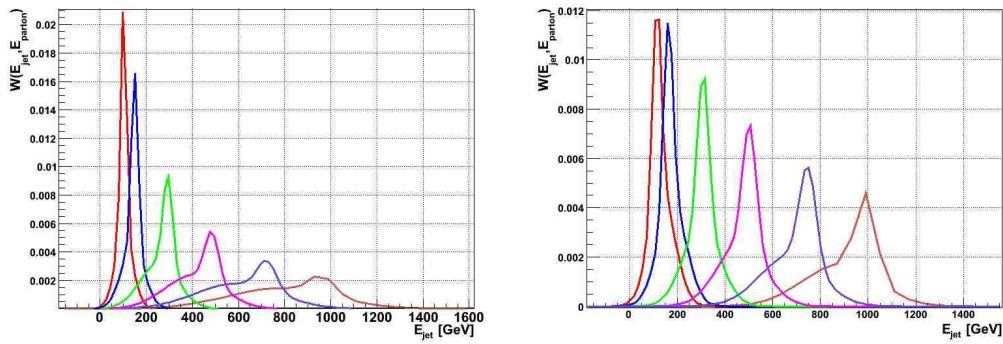
The transfer functions are shown for different parton energies in Fig.5.3 and 5.4. The general behaviour of the transfer functions does not change while the distributions get broader with increasing parton energy.
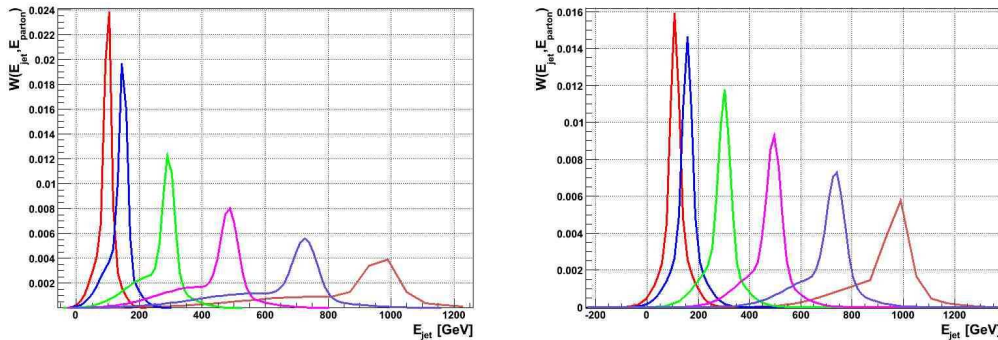
(a) Transfer functions for light quarks in the region $|\eta| < 1.7$

(b) Transfer functions for light quarks in the region $1.7 < |\eta| < 2.5$

Figure 5.3: The transfer functions are shown for parton energies of 100 GeV (red), 150 GeV (blue), 300 GeV (green), 500 GeV (pink), 750 GeV (light blue), 1000 GeV (brown).

All energy distributions are asymmetric and have a width depending on the original parton energy. The width of the functions increases with increasing parton energy.

(a) Transfer functions for b quarks in the region $|\eta| < 1.7$

(b) Transfer functions for b quarks in the region $1.7 < |\eta| < 2.5$

Figure 5.4: The transfer functions are shown for parton energies of 100 GeV (red), 150 GeV (blue), 300 GeV (green), 500 GeV (pink), 750 GeV (light blue), 1000 GeV (brown).

5.2.1. Normalization factors

The transfer functions parameterized in Eq. (5.1) used in this analysis are normalized to unity for a range of $-\infty$ to ∞. While the energies are only defined in a finite and positive region, an additional normalization factor has to be introduced:

$$N = \int_0^{1400\,\text{GeV}} W(E_{\text{part}}, E_{\text{jet}}) d\,E_{\text{jet}} \qquad (5.2)$$

and for b-quarks:

$$N = \int_{4.7\,\text{GeV}}^{1400\,\text{GeV}} W(E_{\text{part}}, E_{\text{jet}}) d\,E_{\text{jet}} \qquad (5.3)$$

The value of 1400 GeV has been chosen according to the highest energy in the data set. The normalization factor is shown in Fig. 5.5. This normalization factor is multiplied to the transfer functions in eq. (5.1).
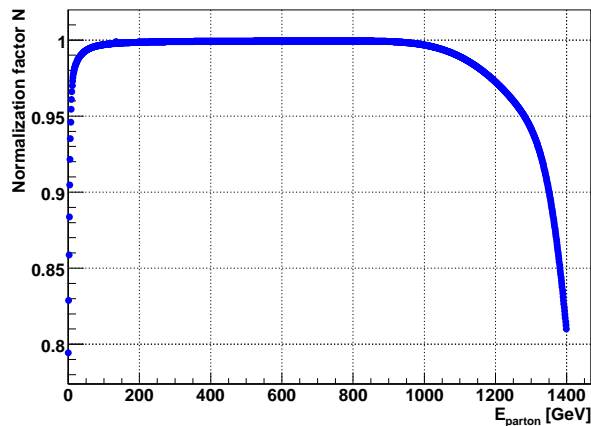

Figure 5.5: Additional normalization constant to account for the physical region.

6. Markov Chain Monte Carlo

The method of Monte Carlo sampling is applied on complex or analytically unsolvable problems. In this Chapter, the integration with Markov Chain Monte Carlo (MCMC) is explained. The MCMC is carried out with the BAT toolkit ($BAT \equiv$ *Bayesian Analysis Toolkit* [32]) which is based on Bayesian statistics. The integration method itself is introduced and the numerical precision and the convergence criteria for the Markov Chains are discussed.

6.1. Sampling with MCMC

The name *Markov chain* describes a series of random numbers $\vec{x}$ which have the *Markov property* that the value $\vec{x}_n$ depends only on the previous value $\vec{x}_{n-1}$ [32]. Markov chains are used to sample a distribution focusing on the interesting regions (*i.e.* regions with larger values of the function). First, a pre-run is carried out with a predefined number of chains (default: five chains). The chains start at different points in the phase space. A proposal point around the current point is chosen randomly by each chain. The decision to accept or reject a new value is made with the Metropolis Hastings algorithm. The efficiency for the acceptance is calculated. The pre-run continues until the chains converge and the efficiencies are between 15 and 50 %. The efficiency is updated once every 1,000 iterations. If it is too low or too high, the proposal function is adjusted by decreasing or increasing the vicinity, respectively. If neither the efficiency nor the convergence criteria are fullfilled, the chains run for the maximum number of iterations (default: 100,000 iterations).

After the pre-run, the *sampling and analysis run* ([32]) is carried out. The adjusted proposal function is used to sample the distribution. During the run, histograms are filled for all possible marginalized one and two dimensional distributions. The effective integration is done via marginalization of the distributions:

$$p(z) = \int dx \, dy \ p(x, y, z). \tag{6.1}$$

In this analysis, the distributions are marginalized with respect to all parameters but the top mass. The marginalized distribution is interpreted as a probability density function to the top pole mass.

6.1.1. Metropolis Hastings algorithm

In the BAT package, the Metropolis Hastings algorithm is used to make a decision about the acceptance or rejection of a new point. If the probability for the new point is larger than the previous probability ($\mathrm{p}(\vec{x}_{n-1}) < \mathrm{p}(\vec{x}_n)$), the new value is accepted. If $p(\vec{x}_{n-1}) > p(\vec{x}_n)$ the ratio of the two probabilities is calculated:

$$\alpha = \frac{p(\vec{x}_n)}{p(\vec{x}_{n-1})} \tag{6.2}$$

α is now compared with a new random number chosen from a flat distribution: $\alpha_{new} \in [0;1]$. If $\alpha < \alpha_{new}$, the new point $\vec{x}_n$ is rejected, if $\alpha > \alpha_{new}$ it is accepted. With this configuration, it is possible to increase the sampling in the regions of higher probability.
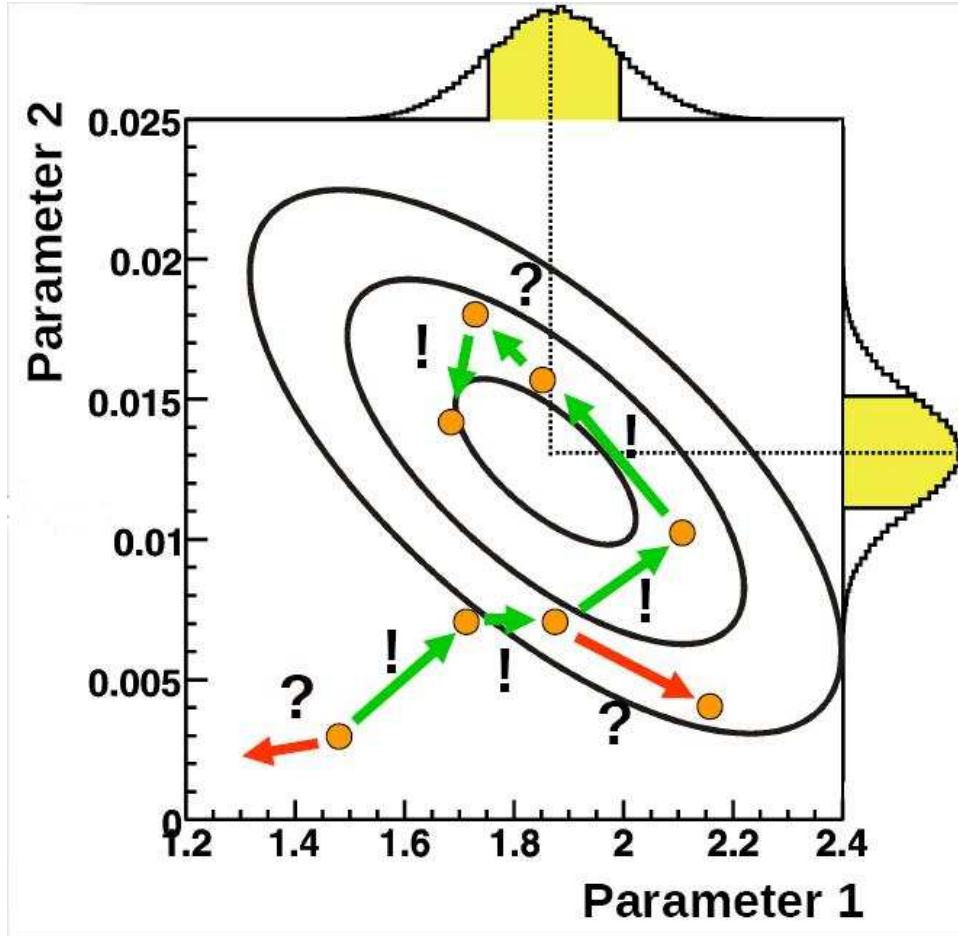

Figure 6.1: Schematic view of the walk of a Markov Chain in a 2 dim. probability distribution with the corresponding marginalized distributions [44].

A schematic view of a single chain in a 2-dimensional phase space is shown in fig 6.1. The accepted steps are marked with green arrows while the rejected steps are marked with red ones. As one can see in the sketch, the step size is adjusted. In contrast to other basic MC methods, mainly the interesting regions of the phase space are sampled.

6.1.2. Convergence criteria

A convergence criterion is defined in the pre-run according to [45]. More than one chain has to be used in order to check the correct behaviour of the integration. The chains are labeled as *converged* if the variance of the mean values for all chains is of the same order as the mean of the variances of the chains. This concept is shown in Fig. 6.2. Eight chains start from random points in the phase space. After the so called *burn-in* phase the chains sample in the same region.
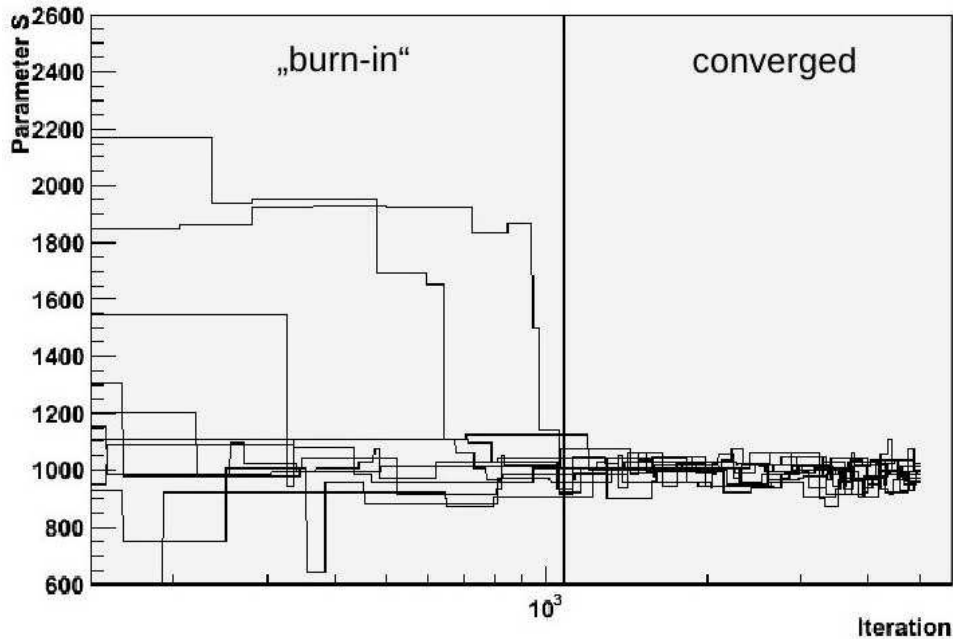
Figure 6.2: Schematic view of the convergence of eight separate chains. The chains start at randomly chosen points in the phase space. After the burn-in phase, all chains sample in the same region [46].

6.1.3. Numerical precision

The distributions created with Markov Chains are not always smooth but statistical fluctuations can occur. To improve the sampling and to smooth the curve, the number of integration steps as well as the number of chains can be increased. Consequently, the running time of the program increases as well. If the number of chains grows, it is more difficult to fulfill the convergence criterion mentioned above. This can be explained by the Metropolis Hastings Algorithm. If the new chosen value yields a lower probability than the value before, another random number is taken. Therefore it may happen, that the chain goes in a region of lower probability. The more chains are used, the higher is the probability that the chains are not located in the same corner of the parameter space. If the number of chains is considered in the calculation, the running time does not increase with the number of chains.

Further problems occur if more than one clear maximum is located in the phase space. If these maxima are in separate corners of the phase space the overall variance is higher than the variance of the single chains. Consequently, the chains do not converge. Nevertheless, the phase space can be sampled adequately. These effects have been studied and the results are shown in the following section. To illustrate that the method is able to find several maxima, a periodic function is sampled in a two dimensional phase space (see fig 6.3):

$$f(x,y) = \frac{\sin(x)}{x} + \frac{\cos(y)}{y}. \tag{6.3}$$

The two and three dimensional distributions are shown in the upper plots of Fig. 6.3. One can see the corresponding marginalized distributions in the lower plots. Five million iterations and 20 chains are used. All maxima are sampled correctly while the curves are smooth.
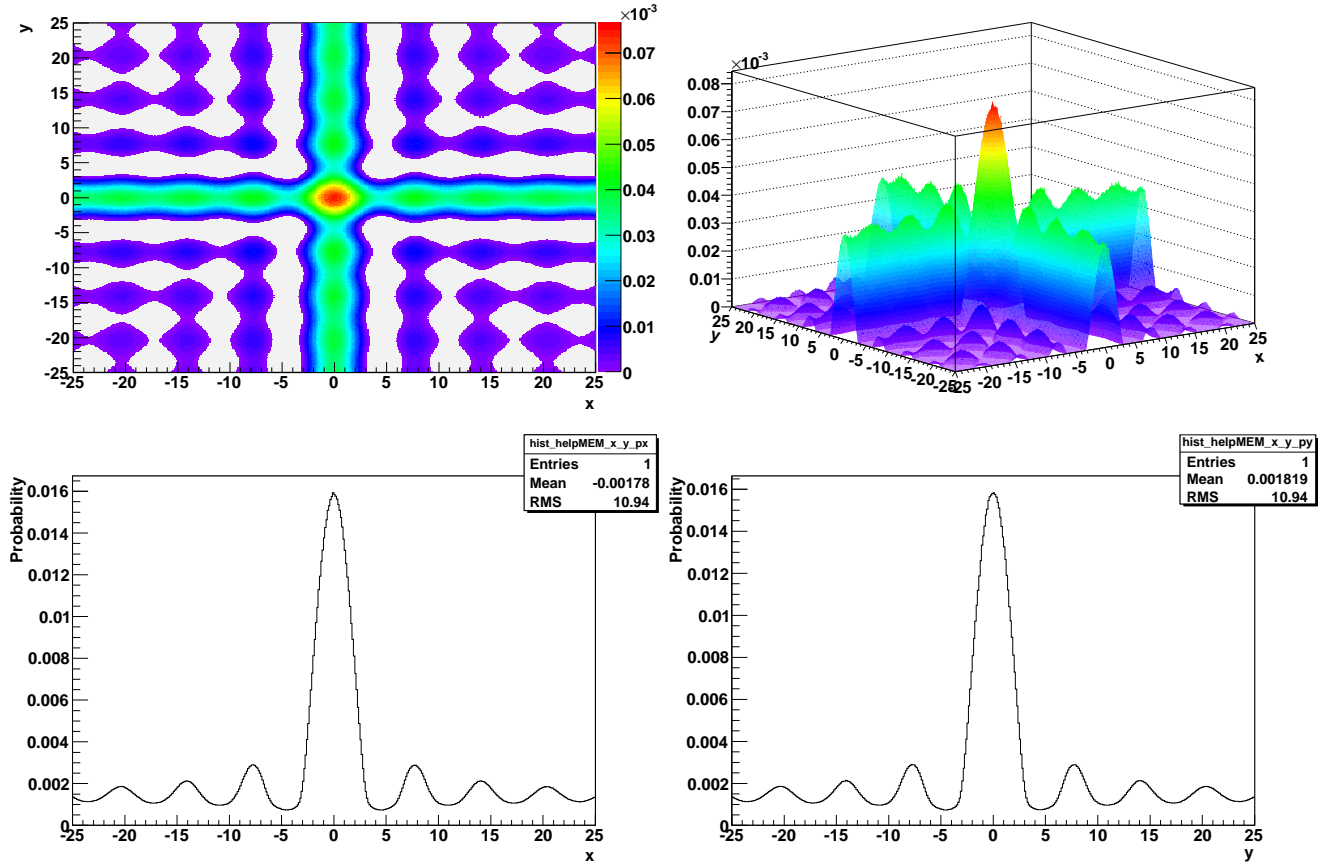
Figure 6.3: Sampling of a 2-dimensional periodic function with Metropolis Hastings Markov Chain Monte Carlo (MHMCMC).

6.2. Comparison of integration variables

The integration over the longitudinal momentum of the neutrino (p_z^ν) is compared to the integration over m_W in the following. Due to the large amount of CPU time, it is important to find the best integration variable to get the fastest results. Although the matrix element is not symmetrized, the effect of a permutation of the two light jets is assumed to be negligible. This results in twelve jet combinations which have to be considered in the integration over p_z^ν. For the integration over p_z^ν (black curve), the differential cross section is calculated via eq. (4.13) while for the integration over m_W (red curve), eq. (4.16) is used. The statistics per event have to be increased in order to smooth the curve. Nevertheless, statistical fluctuations occur.

6.2.1. Comparison on event-by-event basis

To get comparable results, the same number of iterations and parameter limits are used while the number of Markov Chains is set to five. The events used in this analysis are generated at $m_{\text{top}} = 190$ GeV/c^2. The parton energies are smeared according to the parametrization in Chapter 5.2.

One dimensional distributions

First, a one dimensional integration has been done to compare the distributions for p_z^ν. For the integration over m_W, both neutrino solutions are calculated and filled in the histogram weighted by the corresponding differential cross sections (Eq. (4.16)). Evidently, the main maxima can be found in both cases. Nevertheless, the distribution in Fig. (b) is smoother and the local maximum at -65 GeV is sampled more properly. Both distributions are normalized to unity. Due to different integration limits, the shown sections have different probabilities. For
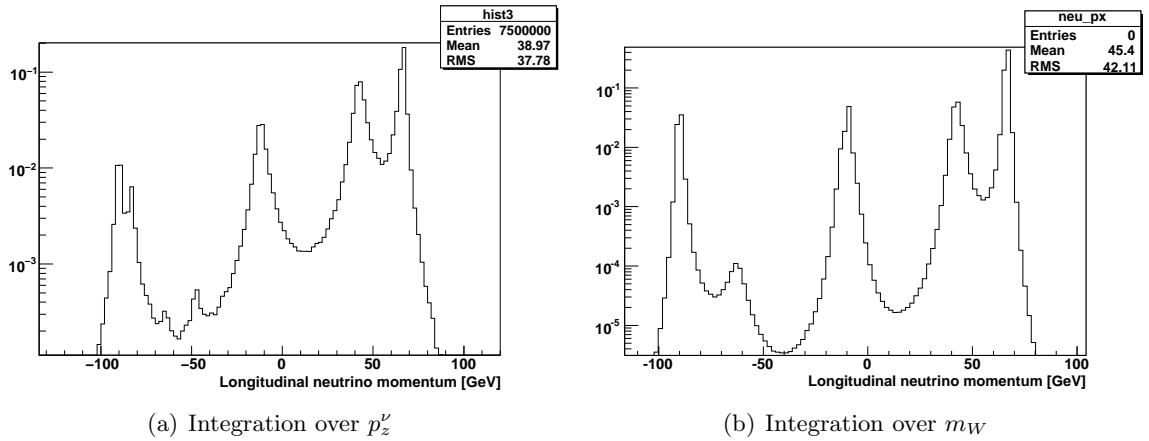

(a) Integration over p_z^ν (b) Integration over m_W

Figure 6.4: Comparison of the one-dimensional integration.

the integration over m_W, always both neutrino solutions are found. This cannot be guaranteed for the integration over p_z^ν. If one solution has very low probability, it may happen that only the one solution with high probability is found. This problem could be solved with a larger number of Markov Chains. As explained in the section above, an increasing number of chains means an increasing running time per event and in addition, the convergence criterion may not be fullfilled. Therefore, the integration over m_W seems to give the best estimation for this one dimensional probability density.

Comparison of multidimensional integration

In Fig. 6.5 and 6.6, single events are shown. The integration is carried out over six dimensions with respect to the top quark mass. The top mass distributions are compared for different integration variables. For the integration over m_W, a red curve is drawn. The black curve shows the distribution for the integration over p_z^ν. In some cases, both events (see Fig. 6.5 (a)-(c)) or no event (see Fig. 6.5 (d)) converges. Nevertheless, the distributions have a similar shape although statistical fluctuations still occur.
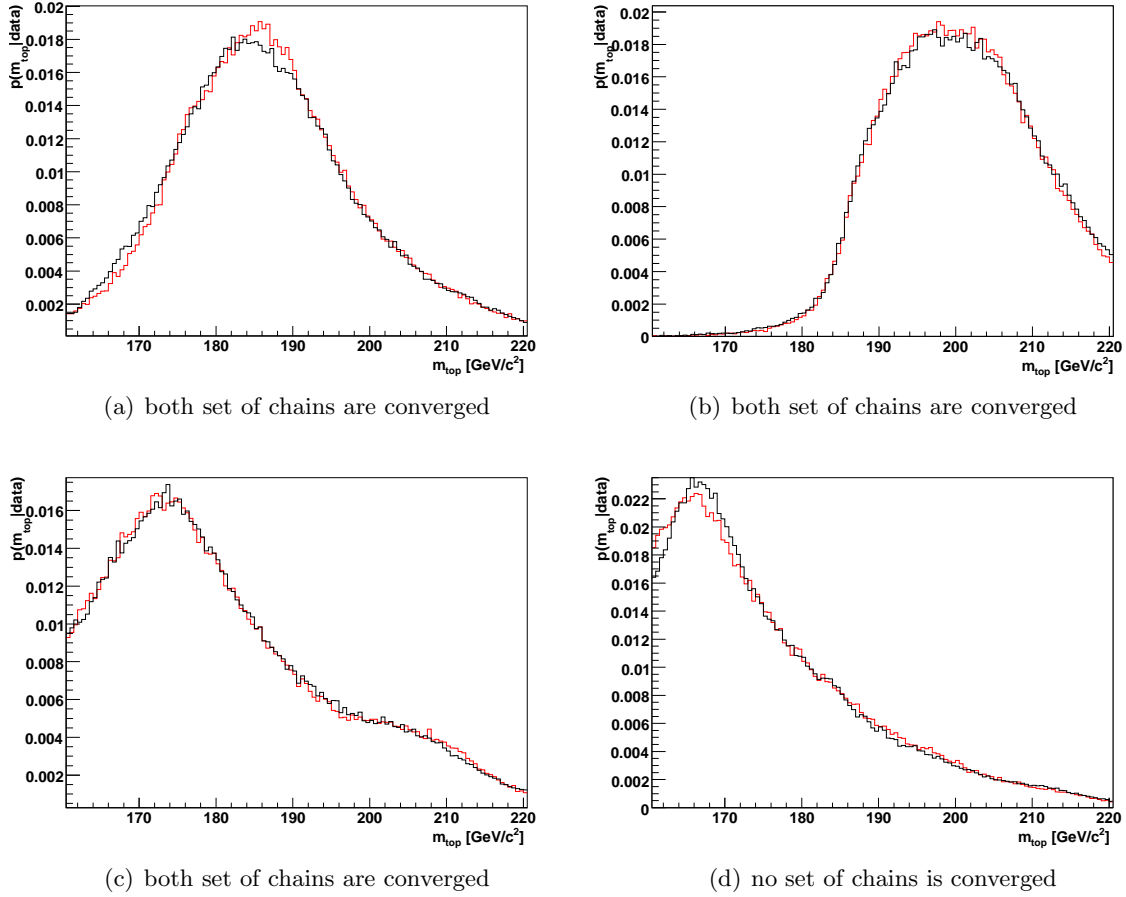

(a) both set of chains are converged

(b) both set of chains are converged

(c) both set of chains are converged

(d) no set of chains is converged

Figure 6.5: Comparison of events with different status of convergence. Red curve: integration over m_W, black curve: integration over p_z^ν.

Sometimes, only one event converges (Fig. 6.6). Even in this case, the distributions have a similar behaviour. The deviations in Fig. 6.6 (c) can be explained with a short running time and binning effects.
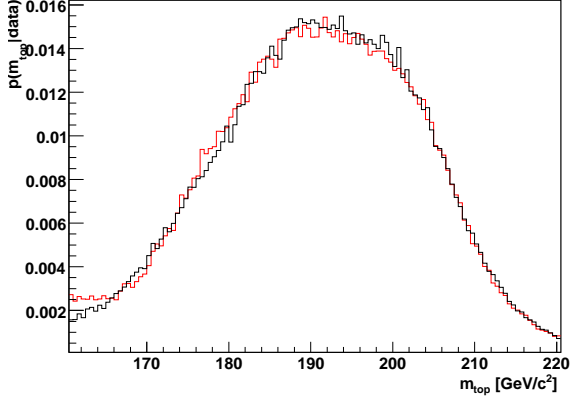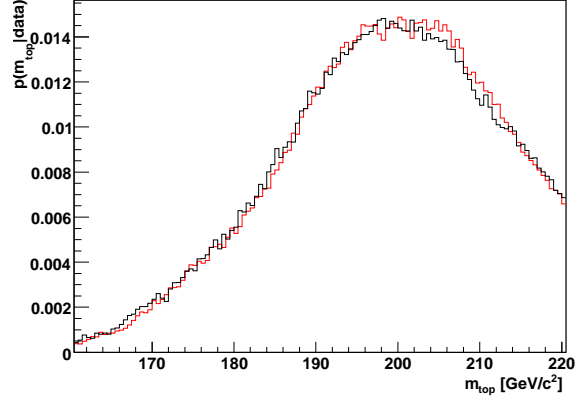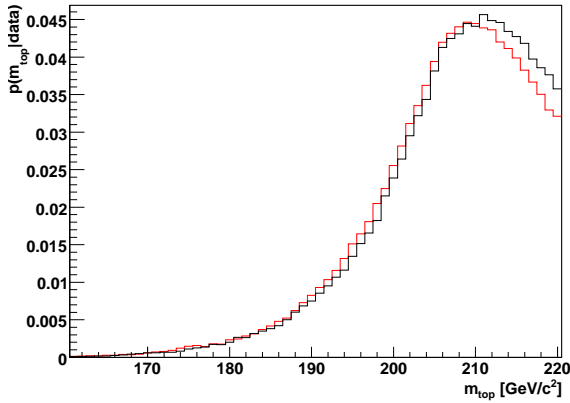

(a) Integration over m_W is converged.

(b) Integration over m_W is converged.

(c) Integration over p_z^ν is converged.

Figure 6.6: Comparison of converged to non converged events. Red curve: integration over m_W, black curve: integration over p_z^ν.

From the distributions above, one could conclude that the chains do not have to converge to provide a reasonable result. Nevertheless, a sufficient number of iterations has to be defined for the pre-run in order to get the correct efficiencies for the main run.

6.2.2. Comparison of ensembles

It is shown that the overall top mass measurement does not depend on the choice of the integration variables in the following. The same set of events is analyzed for each integration variable. The probability density functions for the top mass are shown in Fig. 6.7.
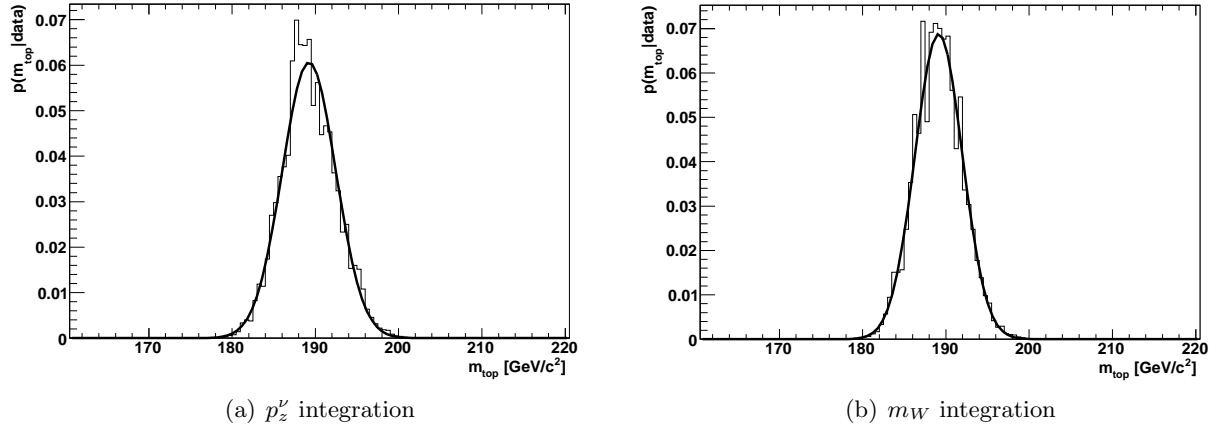

(a) p_z^ν integration

(b) m_W integration

Figure 6.7: Comparison of the combination of 36 events each for both integration variables independent of the status of convergence.

Both distributions have their peak around 190 GeV/c^2 and the curves have a similar shape. The statistical fluctuations are caused by the fluctuation of a single event. The combination of several events is explained in Chapter 7. With respect to the distributions above, one can conclude that the distributions are sampled correctly even if the chains do not converge. The number of iterations has to be large to get reasonable results. Despite the fact that the integration over m_W requires 24 matrix elements to be calculated (instead of 12), this variable has been chosen for the analyses in this thesis. Converging more easily, the pre-run will be shorter for each event and the number of converged events will be much larger. While a larger number of chains will make it more difficult for the chains to converge, five chains are used for the following studies.

6.3. Comparison of integration methods

The integration via the Markov Chain Monte Carlo is compared with the well established VEGAS package [47]. VEGAS is an integration method which uses an importance sampling algorithm [48]. It is implemented using the CUBA library [49]. A grid is defined in the parameter space while the integrand is evaluated at the intersection points. The grid is closer-meshed in the regions of higher probabilities in order to get a better precision in the interesting regions. The step size as well as the number of evaluation points and precision can be predefined. VEGAS performs the integration for a fixed top quark mass hypothesis. In the MCMC, the integration is done over the six variables with the top quark mass being the parameter α.

6.3.1. Comparison on event by event basis

The distributions for single events are shown in Fig. 6.8. The events are generated at a top pole mass of 160 GeV/c^2 and smeared according to the parametrization in Chapter 5.2. Only the correct jet–parton assignment has been used. In Fig. 6.8 (a) and (b), the same event is shown.

In the left plot, the integration with VEGAS is done with a precision of 1%. Ten Markov Chains have been used with ten million iterations in the main run. The right plot shows the same event for a precision of 0.5% in the VEGAS integration. The distributions converge and are consistent within the error bars. In Fig. 6.8 (c) and (d), further events are shown. The integration is done according to Fig. (b). Thus, VEGAS and MCMC provide the same results for high precision.
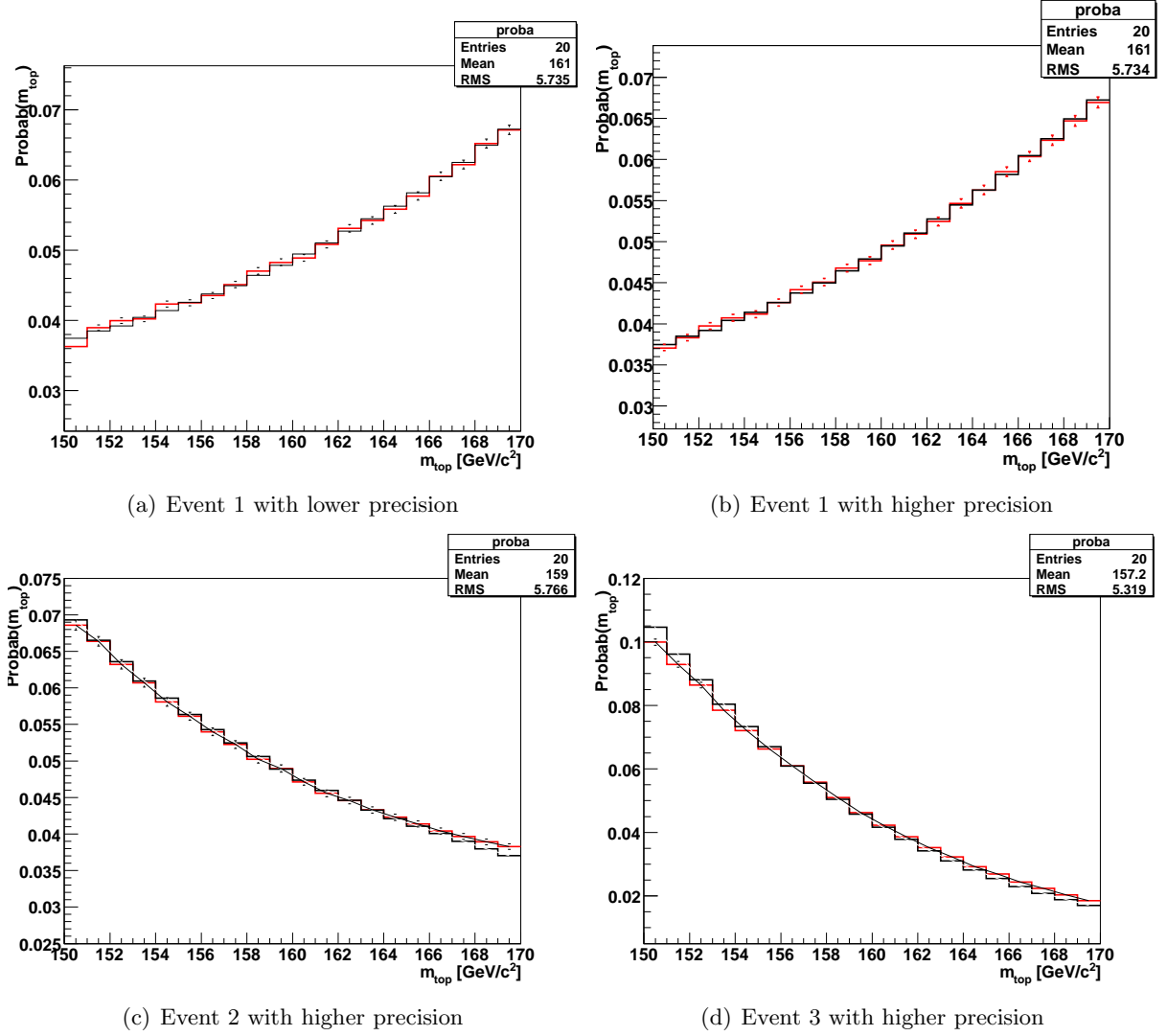

(a) Event 1 with lower precision

(b) Event 1 with higher precision

(c) Event 2 with higher precision

(d) Event 3 with higher precision

Figure 6.8: Comparison of events sampled with Vegas (red) and MCMC (black).

7. Parton Level Tests

To make sure that the method works in general, parton level tests are made with different models for the calorimeter resolution. The tests are done with smeared electron and parton energies. In the first approach, the transfer functions are simple Gaussian functions as explained in Chapter 5. Later, asymmetric transfer functions are used in order to get a better simulation of the real detector resolution. The events are generated with MADGRAPH. The top mass is estimated with the LO matrix element from MADGRAPH. Only semileptonic events are used where one top is decaying into an electron and its corresponding neutrino.

7.1. Tests with symmetric transfer functions

Five samples with different top pole masses (160, 170, 175, 180, 190 GeV/c^2) are generated for a center of mass energy of 14 TeV. The four-vectors are smeared with Gaussian functions. ISR and FSR effects are not taken into account. The probability density is calculated for the six variables (five energies and the W mass) using the top quark mass as a parameter α. The calculation of the top width and the total cross section is done according to the parametrization explained in Chapter 4.1.2.

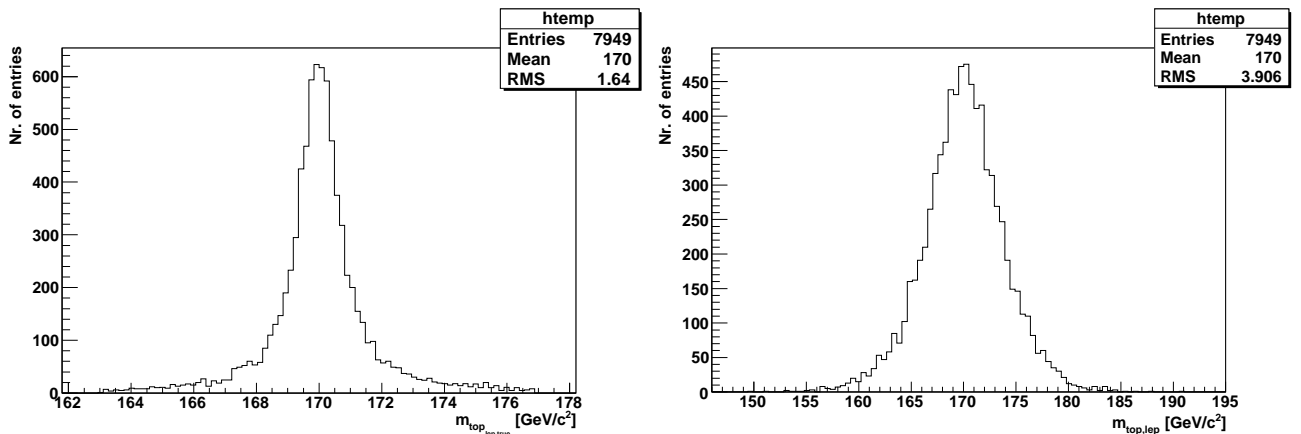**Top reconstruction from smeared parton level events**

Figure 7.1: Leptonic top mass distributions with reconstructed masses from true (left) and Gaussian smeared (right) four-vectors. The four-vector of the neutrino remains unchanged.

The hadronically decaying top quark has three jets in the final state while the leptonically decaying top quark has only one jet. While the neutrino is not smeared, the width of the leptonical distribution is too small. Therefore, the distribution for the hadronically decaying top is broader than the one from the leptonically decaying top quark (see Figs. 7.1, 7.2).
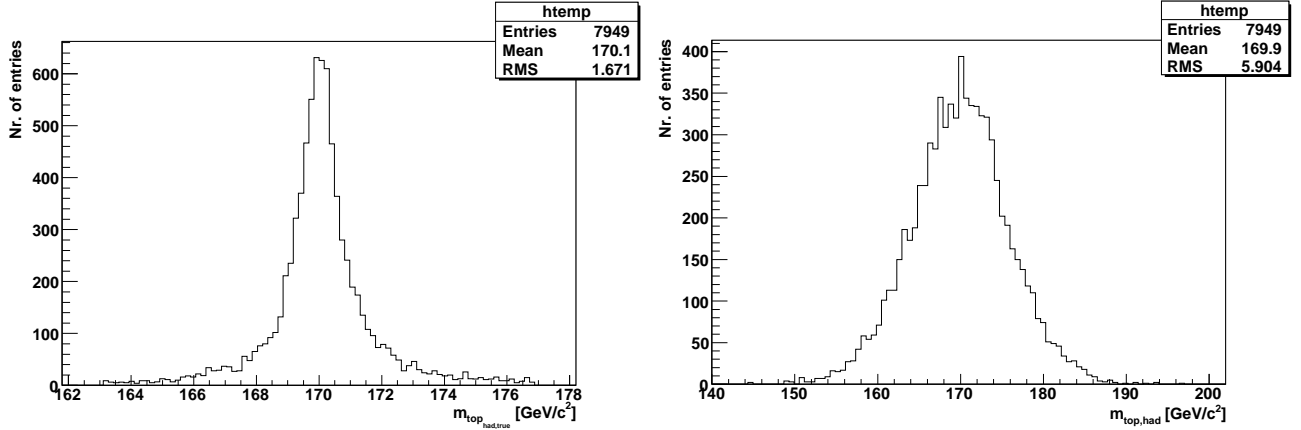
Figure 7.2: Hadronic top mass distributions with reconstructed masses from true (left) and smeared (right) four-vectors.

The sample is generated at a pole mass of 170 GeV/c^2. The distributions for the true values follow a Breit-Wigner function. The distributions for the smeared functions have the shape of a Breit-Wigner function convoluted with a Gaussian function.

Application on single events

The matrix element method is now applied to the generated data. One histogram is created for each event. Example distributions for single events are presented in Fig. 7.3 and 7.4. The events are generated at 170 GeV/c^2 and 190 GeV/c^2, respectively. The probability density distribution for one event is shown in the left plot. Only the correct parton-jet assignment is used. *Jet* in this context describes the smeared parton. The distribution with all twelve combinations is shown in the right plot. The distributions have a similar shape and seem to yield a similar mode. They are drawn with a bin size of 0.5 GeV/c^2. The estimated masses for the events and the corresponding errors are listed in Tab. 7.1.

gen. top mass [GeV/c^2]	# combinations	estim. top mass [GeV/c^2]
170	1	$171.00^{+2.69}_{-2.88}$
170	12	$170.75^{+2.88}_{-2.53}$
190	1	$190.25^{+3.17}_{-3.91}$
190	12	$189.75^{+3.65}_{-3.43}$

***Table 7.1:*** Estimated top quark mass with the corresponding errors for a different number of jet-parton combinations and combined events.

The uncertainties increase with increasing top mass. The top mass is correctly estimated for the two events. The results are consistent within the errors. One can therefore conclude, that the correct combination provides the largest contribution to the sum of the differential cross sections.
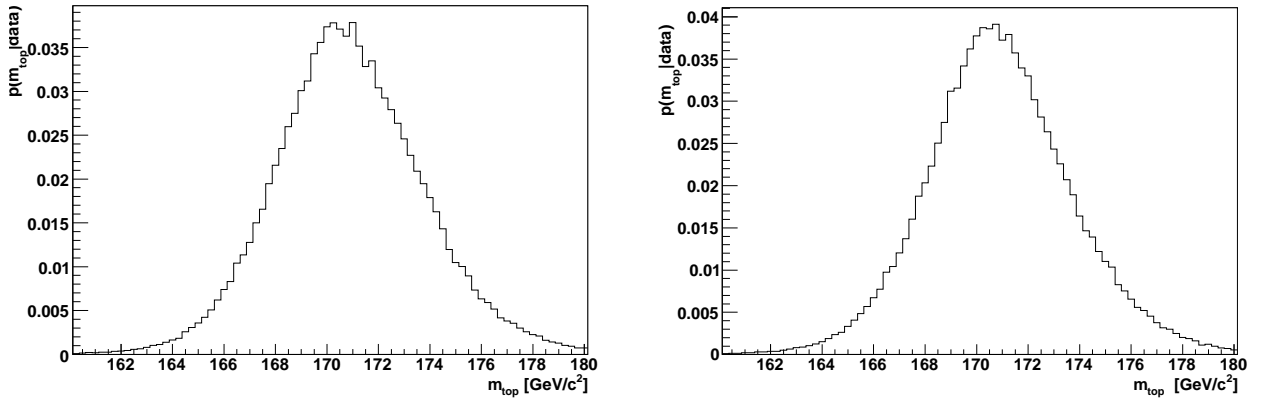
(a) Distribution for a single event with the correct combination

(b) Distribution for one single event with all combinations

Figure 7.3: Comparison of the distributions for one single event. The left plot shows the distribution for the correct combination only. In the right plot, all combinations are used. The distributions are normalized to 1.0. The event is generated with a top mass of 170 GeV/c^2.
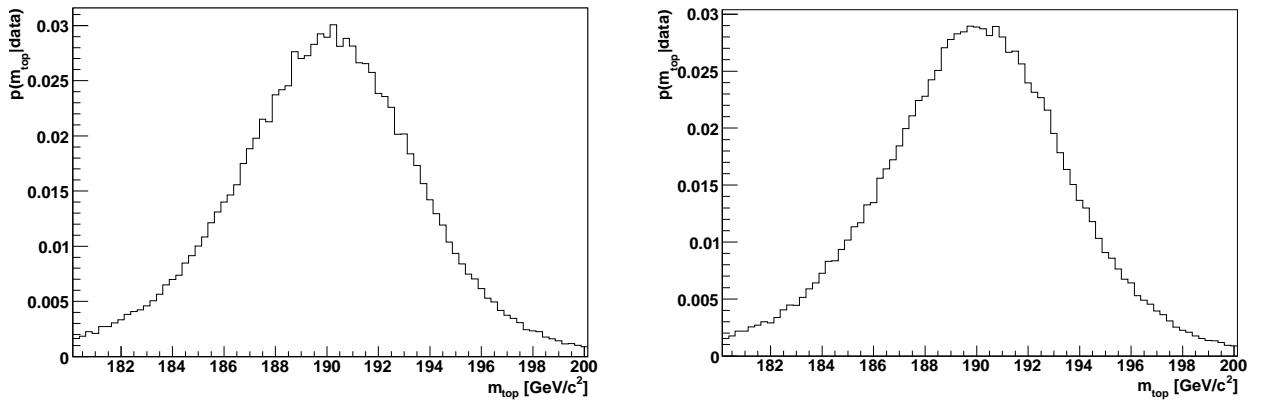
(a) Distribution for one single event with the correct combination

(b) Distribution for one single event with all combinations

Figure 7.4: Comparison of the distributions for one single event. The left plot shows the distribution for the right combination only. In the right plot, all combinations are used. The event is generated with a top mass of 190 GeV/c^2.

Top mass estimation with a set of events

The distributions are interpreted as probability densities for the top pole mass. They are normalized to unity for each event. For a combination of events, the distributions are multiplied bin-by-bin and then normalized to unity. The results of the combination of 200 events are shown in Fig. 7.5. Since the distributions are not symmetric, the mode is taken as the pole mass estimator and the uncertainties are defined as the differences between the mode and the quantiles at 84.16% and 15.84% respectively.

The top pole mass is estimated to be $170^{+0.13}_{-0.29}$ GeV/c^2 and $189.75^{+0.1}_{-0.39}$ GeV/c^2 in the two samples. The results are consistent within the uncertainties.

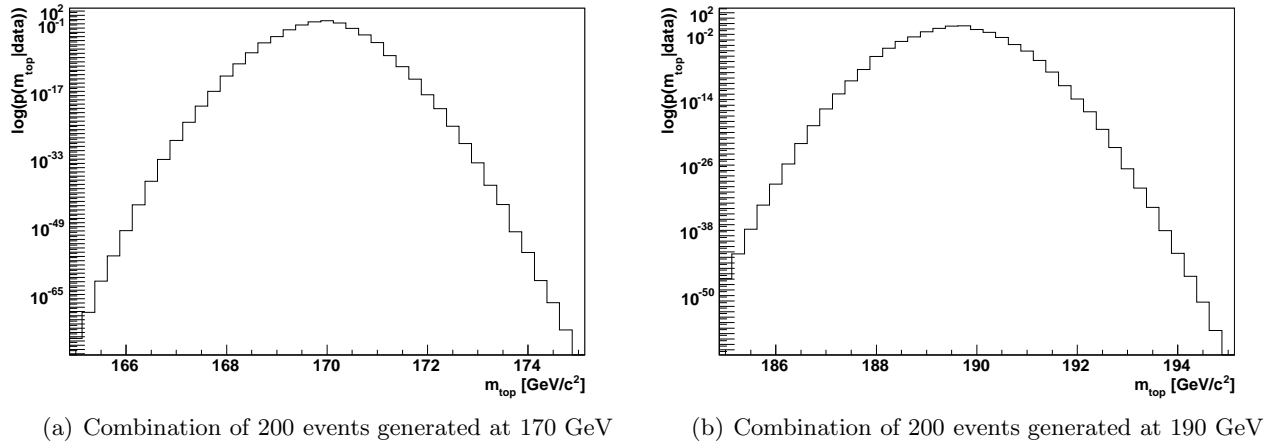(a) Combination of 200 events generated at 170 GeV   (b) Combination of 200 events generated at 190 GeV

Figure 7.5: Distributions of the combination of 200 events generated at different pole masses. Five Markov Chains have been used and all twelve combinations have been taken into account.

Calibration curve

The analysis has been applied on the five different Monte Carlo data sets mentioned above. Each data set contains 200 events. All twelve possible jet-parton combinations are taken into account for each event. Figure 7.6 shows the estimated pole mass from smeared MC events as a function of the five input pole masses. The graph has been fitted with a linear function yielding an offset of zero ($p_0 = -0.09 \pm 0.12$) and a slope of one ($p_1 = 0.99 \pm 0.01$). It is shown that the method yields an unbiased estimator of the top pole mass.
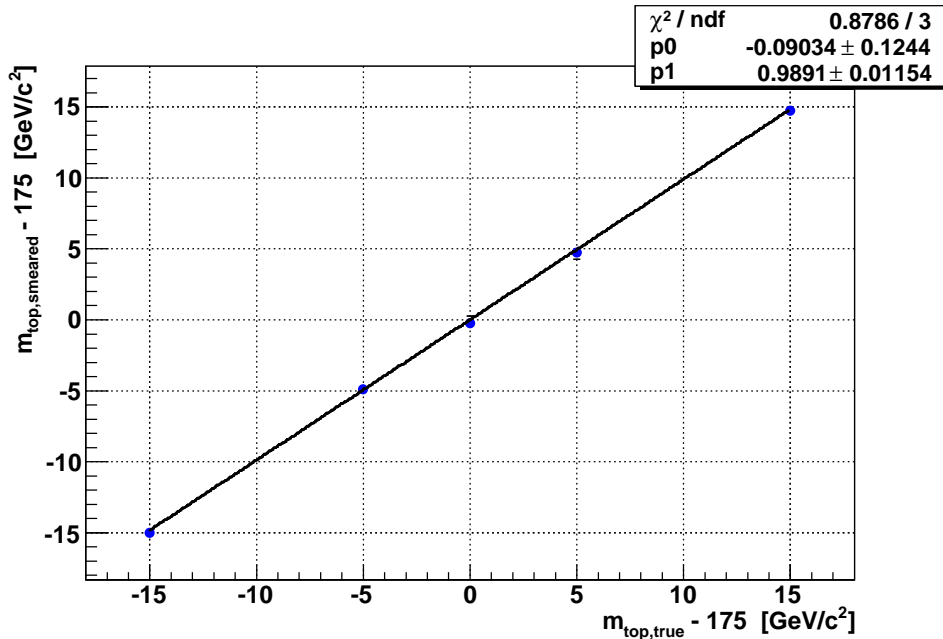

Figure 7.6: Calibration curve with 5 data sets and 200 events each. All twelve combinations have been taken into account.

Development of the statistical uncertainty

The estimated top mass with the corresponding uncertainties is shown as a function of the number of combined events in Fig. 7.7. The data set is again generated with a pole mass of 170 GeV/c^2. One can see that all values are consistent with a top pole mass of 170 GeV/c^2. The asymmetric errors decrease fast with an increasing number of events. Oscillations can be seen in Fig. 7.7. These are quantized due to the chosen binning.
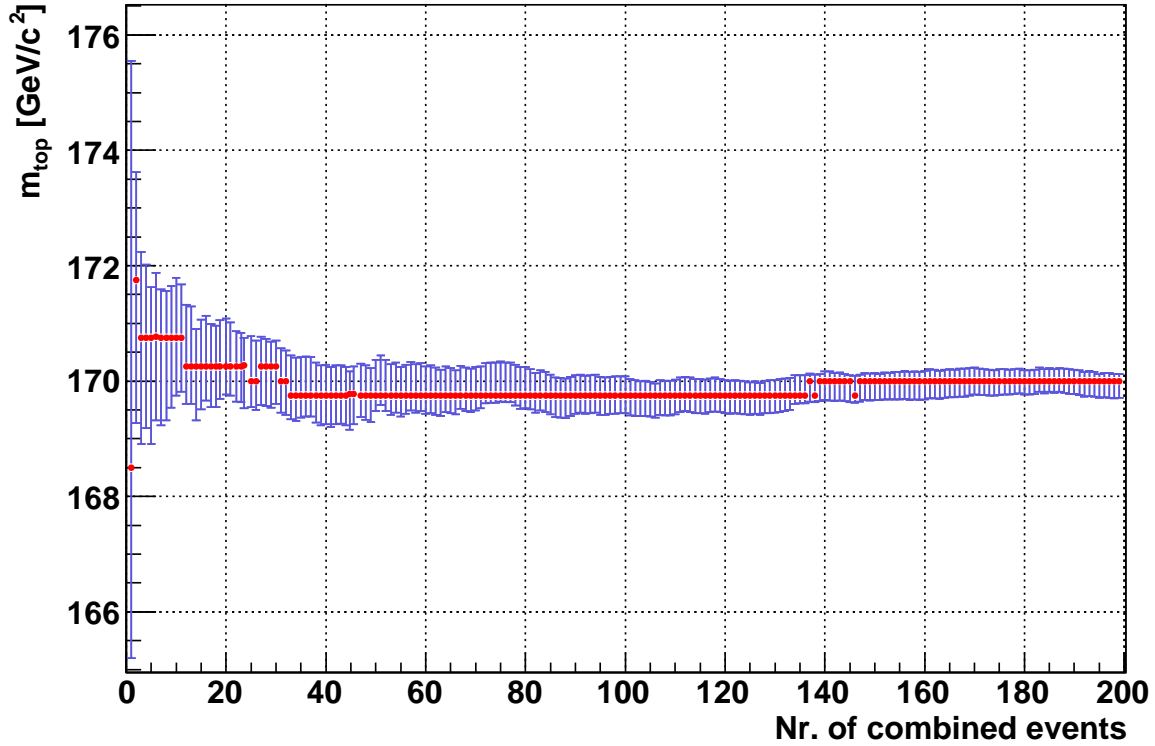

Figure 7.7: Estimated pole mass as a function of the number of combined events. The events are generated at a top mass of 170 GeV/c^2.

7.2. Tests with asymmetric transfer functions

Since the analysis works well for symmetric transfer functions, the next step is to include an ATLAS-like energy resolution. A parametrization has been derived and is described in Chapter 5. Again, five data samples with different top pole masses are generated with MADGRAPH and the four parton energies are smeared according to the parametrization. The electron energies are still smeared with a simple Gaussian with a width of 0.1 $\sqrt{E[\text{GeV}]}$.

The distributions in Fig. 7.8 show the mass of the top quark and W boson reconstructed from the smeared 4-momenta of the decay particles. The data set is generated at $m_{\text{top}} = 170$ GeV/c^2. On the left side, the mass of the leptonically decaying top quark is shown. On the right side one can see the distributions for the hadronically decaying particle. The distributions are broader than the ones generated with the Gaussian transfer functions.

Again, the leptonic distribution is more narrow than expected while the true neutrino energy and momentum are used. The distributions have the shape of a Breit-Wigner function convoluted with the double Gaussians. The mode of the reconstructed top quark mass distribution is larger than 170 GeV/c^2 (see Fig. 7.8).

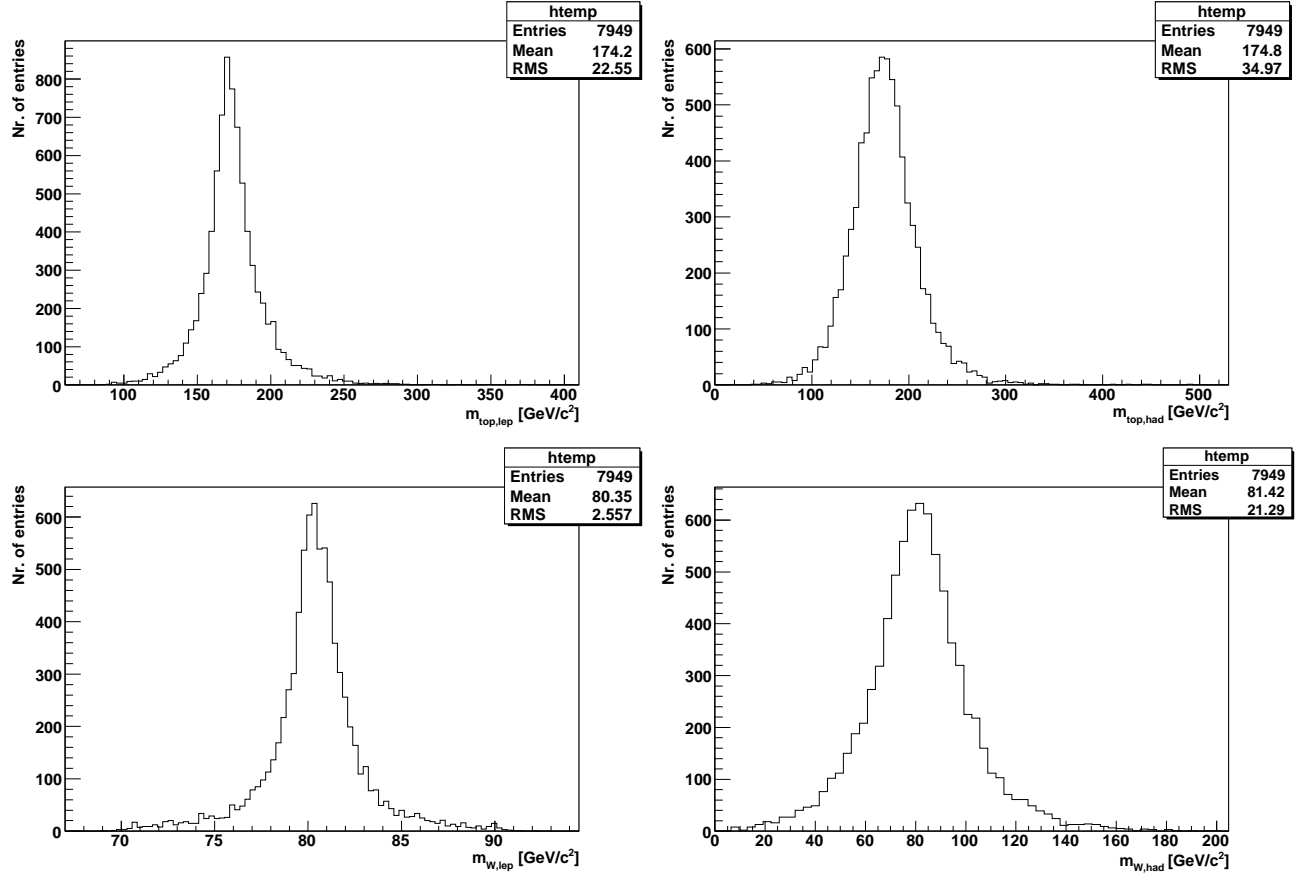***Figure 7.8:*** Mass of the leptonically (top left) and hadronically (top right) decaying top quark and of the leptonically (bottom left) and hadronically (bottom right) decaying W boson for asymmetric transfer functions.

7.2.1. Combinatorical background

While all possible jet-parton assignments have to be included in the analysis, combinatorical background can change the shape of the mass spectrum. To show the influence of wrong parton assignment, the distributions for the correct, wrong and half-correct combinations are displayed in Fig. 7.9. If the combination is called *half-correct*, one light and one b quark are assigned correctly while the other two changed places. For the wrong assignment, the light jets are taken as the b jets and vice versa. For each distribution, 200 events are used which are generated at 170 GeV/c^2.

(a) correct combination	(b) correct combination
(c) half-correct combination	(d) half-correct combination
(e) wrong combination	(f) wrong combination
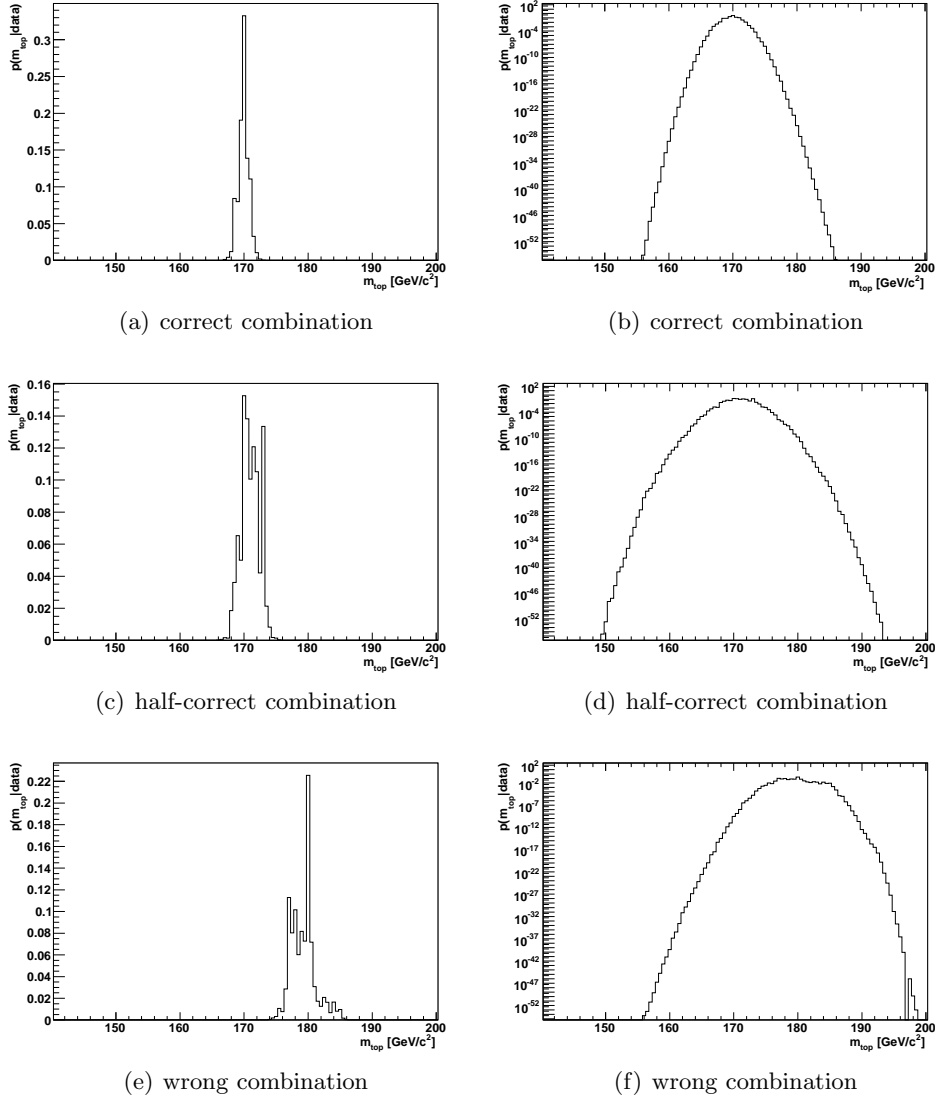
Figure 7.9: Combination of 200 events for the correct (top), half correct (middle) and wrong (bottom) jet assignment. In the left plots, a linear scale has been used while in the right plots the scale is logarithmic.

As one can see in the upper plots, the correct combination yields a mode of 170 GeV/c^2. In comparison, the middle distributions are broader and asymmetric while the wrong combination (bottom) results in a shift of the maximum and a broad, asymmetric shape.

7.2.2. Combination of a set of events

Figure 7.10: Distribution for a single event (left) and the combination of 375 events (right), generated at 170 GeV/c^2.

The distribution for a single event is shown in Fig. 7.10 (left). It is much broader compared to the result in Fig. 7.3 a and the mode is $176.5^{+11.98}_{-12.84}$ GeV/c^2. Due to limited computing power, only 375 events could be analyzed for each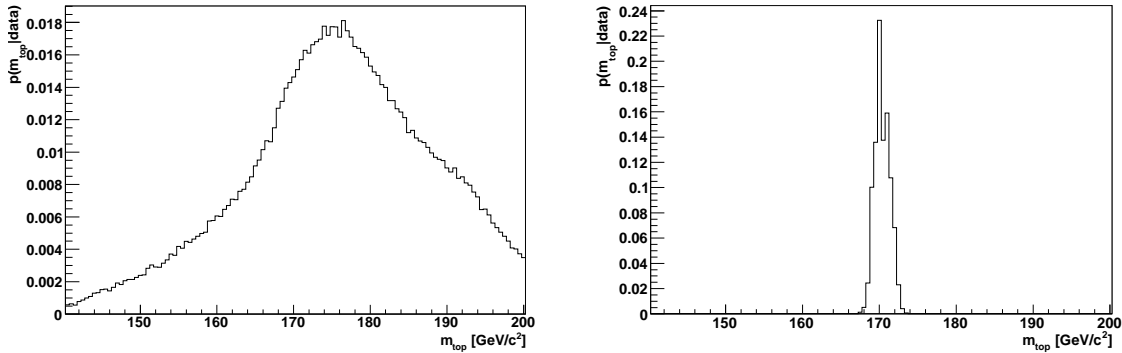 pole mass. The combination of these events is shown in Fig. 7.10 (right). The mode is $170^{+1.22}_{-0.53}$ GeV/c^2 but a second peak is visible around 171 GeV/c^2 due to statistical fluctuations. Therefore, the top mass can be estimated correctly.

7.2.3. Calibration curve

The analysis is applied to the five datasets with different pole masses and a calibration curve is shown in Fig. 7.11. The results of the linear fit are consistent with a slope of 1.0 and an offset of 0.0 within 2σ (see Fig. 7.11). Again, the analysis provides an unbiased top pole mass estimation. With a larger data set, the results are expected to improve.
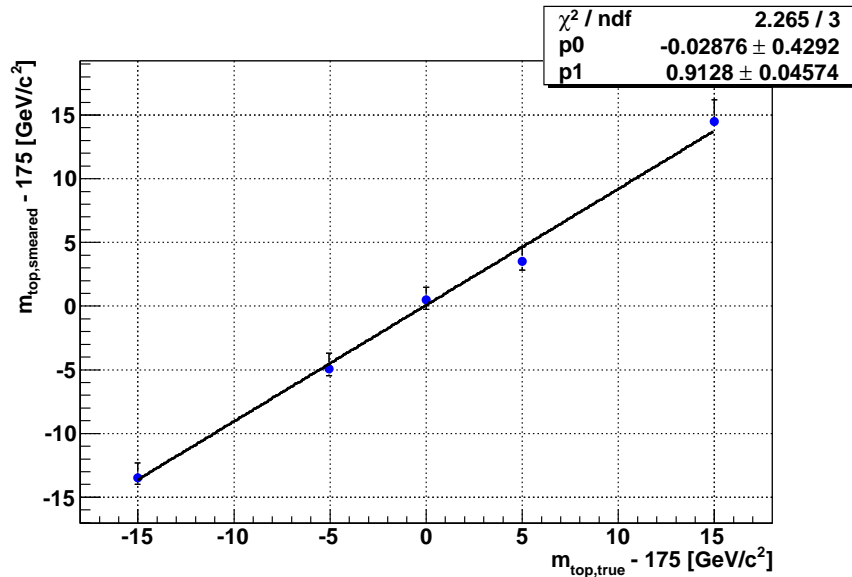
Figure 7.11: Calibration curve with five data sets and 375 events each. All twelve combinations have been taken into account.

7.2.4. Development of the statistical uncertainty

The estimated top pole mass and the uncertainties are shown in Fig. 7.12. Compared to the results from the Gaussian transfer functions, much more statistics is needed to get small uncertainties. Moreover, the errors are much larger and asymmetric. Even for a large number of combined events, oscillations can be seen.
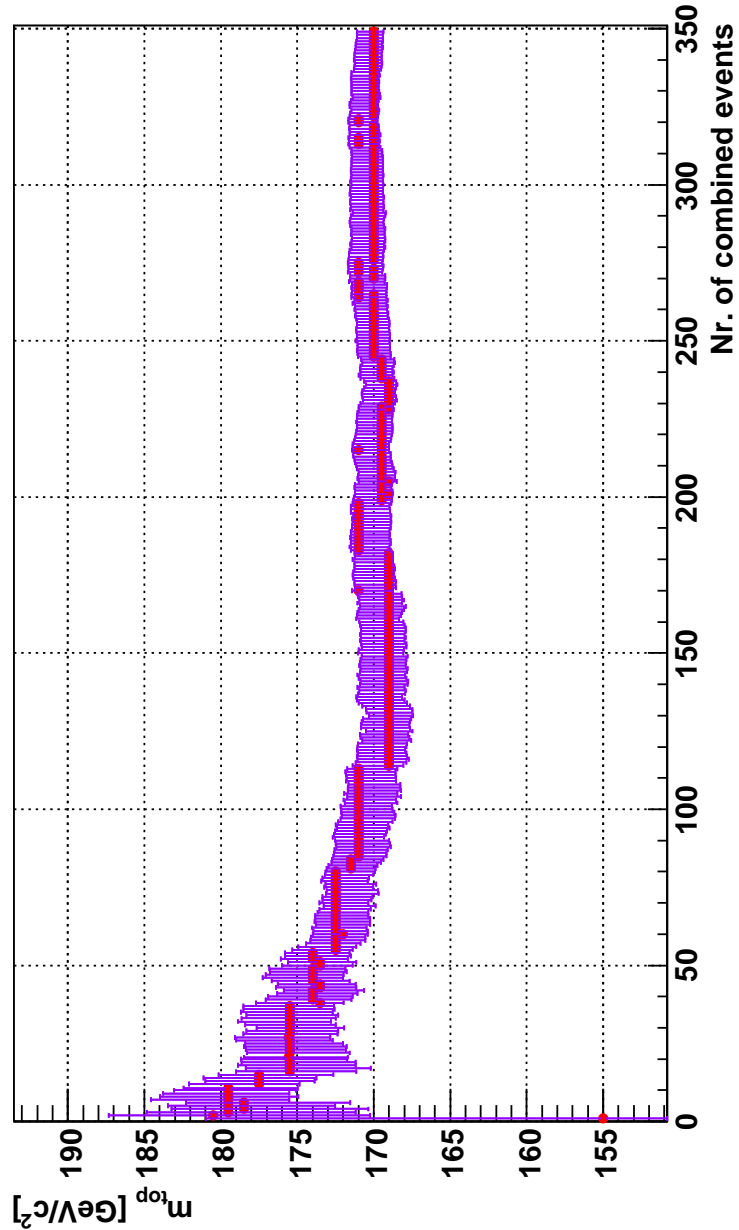
Figure 7.12: Uncertainty of the top quark mass estimation as a function of the number of combined events.

7.2.5. Ensemble testing

Ensemble tests are done to study statistical uncertainties with a limited number of events. In this case, only a signal sample with 400 events and a pole mass of $m_{\text{top}} = 170$ GeV/c^2 is used. A certain number of events is chosen randomly from this sample without being removed. Therefore, some events may be used several times in one ensemble and correlations may occur. The estimated pole mass and the uncertainties are histogrammed (Fig. 7.13).
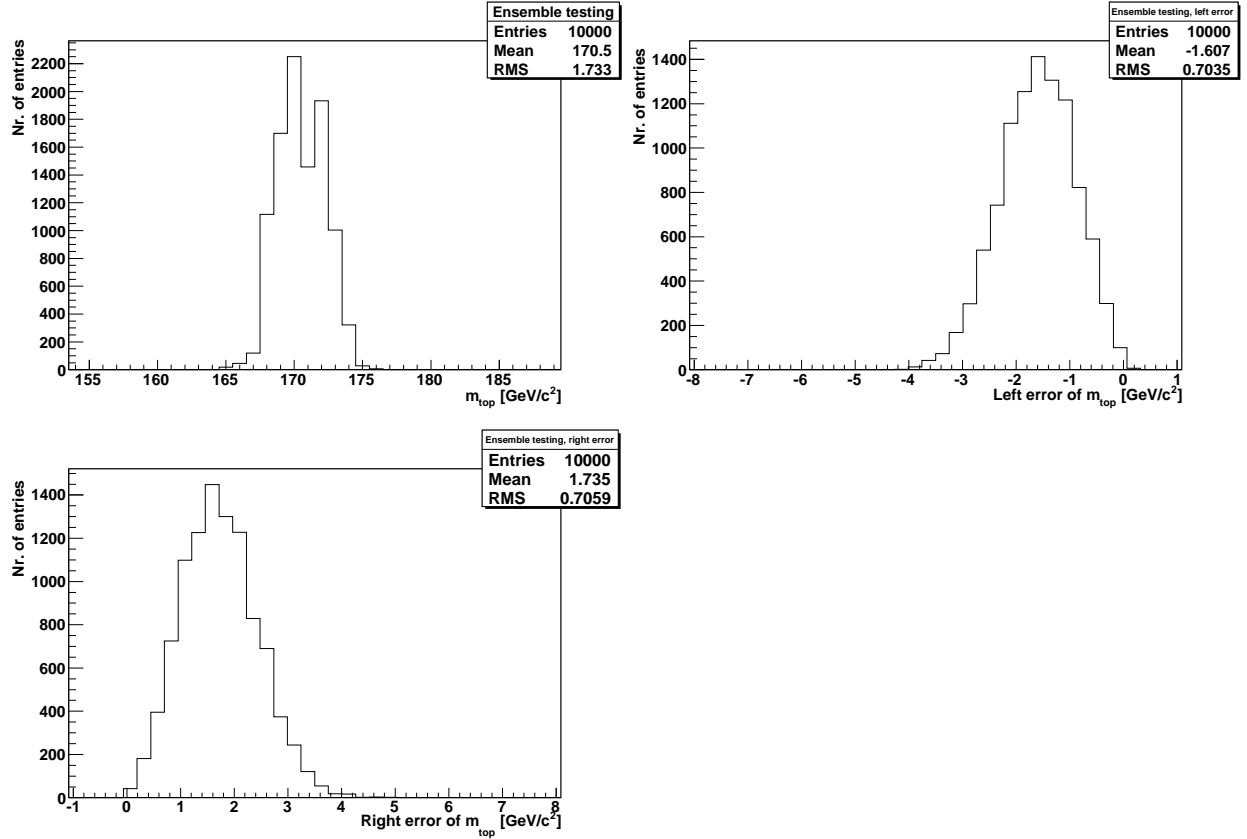
***Figure 7.13:*** Ensemble tests with 80 events per ensemble have been performed. In the top left plot, the distribution for the mode is shown. In the top right and bottom left plot, the uncertainties are drawn. The mode is found to be at 170 GeV/c^2.

The ensembles are built out of 80 events each. In the top quark mass distribution, the mode yields a value of 170 GeV/c^2. The second peak on the right side is assumed to be a statistical fluctuation caused by the small number of events in the sample. The uncertainty distributions are the 15.84 % and 84.16 % quantiles. They are smooth and only slightly asymmetric. The RMS value of m_{top} is consistent with the mean of the uncertainty distributions. Since the distributions are not Gaussian, a pull distribution could not be used to check the calibration of the method.

8. Application to ATLAS Monte Carlo

The method has been applied on NLO Monte Carlo events from the official ATLAS Monte Carlo production. The selection criteria for semileptonic events are discussed and a study on the selection efficiency is performed. Studies are performed for signal events first. Later, the influence of different background fractions on the precision of the top mass measurement is studied.

8.1. Event selection

The MC data sets described in Chapter 4.3.1 are analyzed and different cuts are applied in order to find good discriminators to reduce the background. To illustrate the different cuts, distributions of $E_{T_{miss}}$ and the jet multiplicities are shown in Fig. 8.1. The number of events used in the distributions are normalized to the respective cross sections. The cuts used here are slightly weaker than the final cuts in the CSC analyses [50]. If the cuts from the CSC book had been used, almost no background would have been left for further studies. Due to a misreconstruction of top events in the sample, a cut on the truth values of the top quark mass has to be done first with $|m_{\mathrm{reco}} - m_{\mathrm{true}}| < 1$ GeV/c^2. The event selection requires:

a) $E_{T_{miss}} > 20$ GeV

b) exactly one isolated electron with:

   − $p_T > 20$ GeV/c
   − $|\eta| < 2.5$
   − no overlap with muons

c) at least 4 **tight** jets:

   − $p_T > 20$ GeV/c, $|\eta| < 2.5$
   − no overlap with muons

d) at least 3 **good** jets:

   − $p_T >40$ GeV/c, $|\eta| < 2.5$
   − no overlap with muons

The $E_{T_{miss}}$ distributions from $t\bar{t}$ and $W + 4\ jets$ have a similar behaviour. Therefore, it is not expected to be a good variable for the reduction of the W+jets background. The isolation criterion for the lepton allows only an additional energy E < 6 GeV to be in a cone of 0.2 around the electron. Overlap removal of muons is done both for electrons and jets. In addition, a signal in the ECal is required. If more or less than one electron is reconstructed, the event is removed from the sample. The wrong reconstruction of electrons can have several reasons. If it is identified as a jet, no electron is reconstructed.
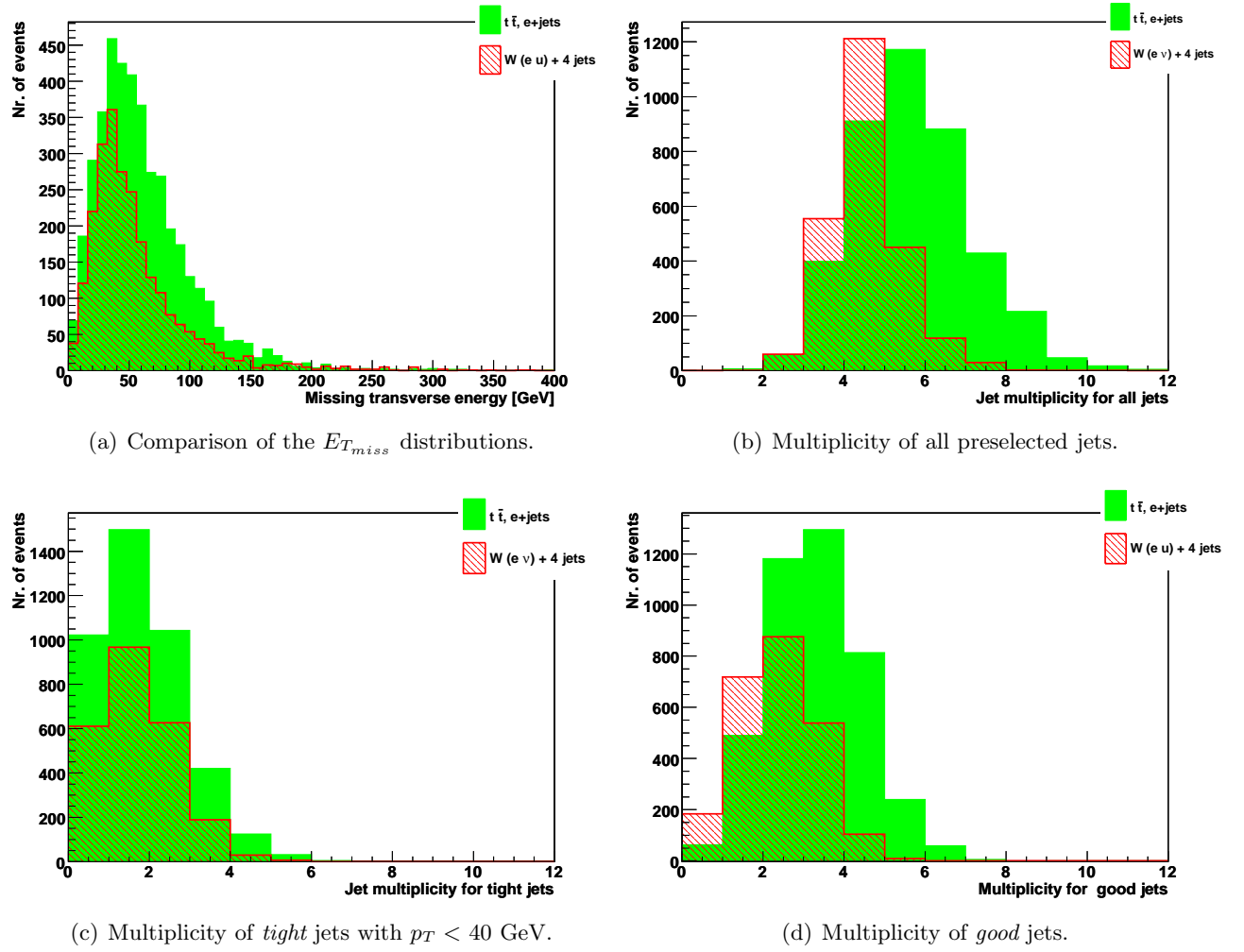
(a) Comparison of the $E_{T_{miss}}$ distributions.

(b) Multiplicity of all preselected jets.

(c) Multiplicity of *tight* jets with $p_T < 40$ GeV.

(d) Multiplicity of *good* jets.

Figure 8.1: Distributions for different discriminator variables. In the selection, events with at least 3 good jets are chosen. Most of the $W + 4\ jets$ events have less than three good jets and more than one tight jet with $p_T < 40$ GeV/c. Therefore, the jet multiplicity provides a good tool to reject $W + 4\ jets$ background. The distributions are normalized to the respective cross section mentioned in Chapter 4.6

Furthermore, pions could have been identified as electrons or additional electrons can be radiated from jets. In that case, too many electrons are found in the event.

The jet multiplicities are shown in Fig. 8.1 (b)-(d). Figure (b) shows the multiplicities for all jets in the event. Evidently, the multiplicity is higher for the signal events. The jets are classified with respect to their transverse momentum and η. They are labeled as **good** jets, if their transverse momentum is larger that 40 GeV/c with $|\eta| < 2.5$. The description **tight** jet is used here for jets with a $p_T > 20$ GeV/c and $|\eta| < 2.5$. In Fig. 8.1 (c), only the tight jets with $p_T < 40$ GeV/c are shown. In contrast to that distribution, the multiplicity of good jets is an excellent discriminator to improve the signal purity. Requiring at least 3 good jets leads to a significant decrease of the background fraction. The cuts mentioned above are applied and further distributions are made for a subset of these cuts. These distributions can be found in Appendix B.

8.1.1. The selection efficiencies

In addition to the object selection above, further criteria have been defined to select four jets. Only *good* or *tight* jets survive the preselection. All possible trijet combinations are compared and the combination with the highest p_T is chosen (m_{top} cut). One of the three possible dijet combinations of the top has to fullfill the equation $|m_W - m_{\text{jj}}| < 10$ GeV/c^2 (m_W - cut). If the event meets these requirements, a fourth jet has to be selected. If more than one additional jet has passed the preselection, the one with the highest p_T is chosen.

The efficiencies of the different cuts are presented in Tab. 8.1. The percentages correspond to the total number of events which contain only one electron (positron). The most effective cut to reduce the W+jets background is the requirement of at least 3 good jets in the event. Since the signal sample comprises dileptonic and semileptonic events, the first cut reduces the number of events by a factor of ≈ 4 according to the branching ratio of the e+jets channel.

	Signal events		Background events	
	Nr. of events	% (cumul.)	Nr. of events	% (cumul.)
total number	548,550		6,000	
only one e^-/e^+	146,262	100	2,431	100
Cut a)	145,360	91	2,181	89.72
Cut b)	92,250	63.46	1,697	69.81
at least 3 good jets	51,670	35.55	337	13.86
at least 1 nonloose jet (< 40 GeV)	37,237	25.62	337	13.86
m_W cut	24,190	16.64	116	4.77
only e^- events	12,053	8.24	55	2.26

Table 8.1: Comparison of cut efficiencies for signal and background events.

8.2. Application to NLO Monte Carlo events

The signal to background ratio is 8.34:1 after the event selection. The event selection mentioned above is also applied to a top sample with a pole mass of 170 GeV/c^2. It is generated according to the sample described in Chapter 4.6 with only small deviations. No information about jets or electrons overlapping with muons is available. Hence, no overlap removal can be done. Furthermore, the event weight for the selected events is zero. Due to this, the event weight is neglected for this data set. The selection efficiency is found to be 7.91 %. The signal samples are used to estimate the top pole mass from the NLO MC data with the LO matrix element from MADGRAPH.

The combination has to be done according to the event weight information from the MC production. This event weight accounts for regions in the phase space, where too many events have been generated. For positive weights, the entries are multiplied bin-by-bin. For negative weights, the entries are divided. Since only a small number of events is used in the analysis, the event weight may cause a bias.
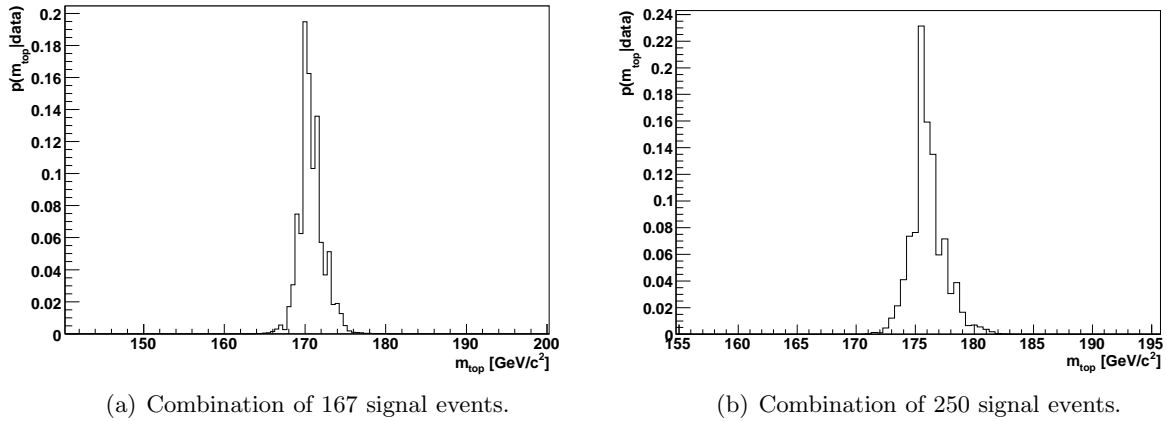
(a) Combination of 167 signal events.

(b) Combination of 250 signal events.

Figure 8.2: Estimation of the top quark mass using NLO MC data and a LO matrix element. A sample with a top pole mass of 170 GeV/c^2 (left) and one with 175 GeV/c^2 (right) have been used.

The top quark mass can be estimated correctly from the NLO MC data. The resulting mass of the sample generated at a pole mass of 170 GeV/c^2 yields $170^{+2.16}_{-0.57}$ GeV/c^2. For the sample generated at a pole mass of 175 GeV/c^2, the resulting mass is $175.5^{+1.97}_{-0.71}$ GeV/c^2 (Tab. 8.2).

8.2.1. Background contribution

To study the effect of background events, a combination of signal and background events is done for a signal to background ratio of 8.34:1 (see Fig. 8.3). The estimated top quark mass is biased (Tab. 8.2).

(a) Combination of 167 signal and 20 background events. (b) Combination of 250 signal and 30 background events.
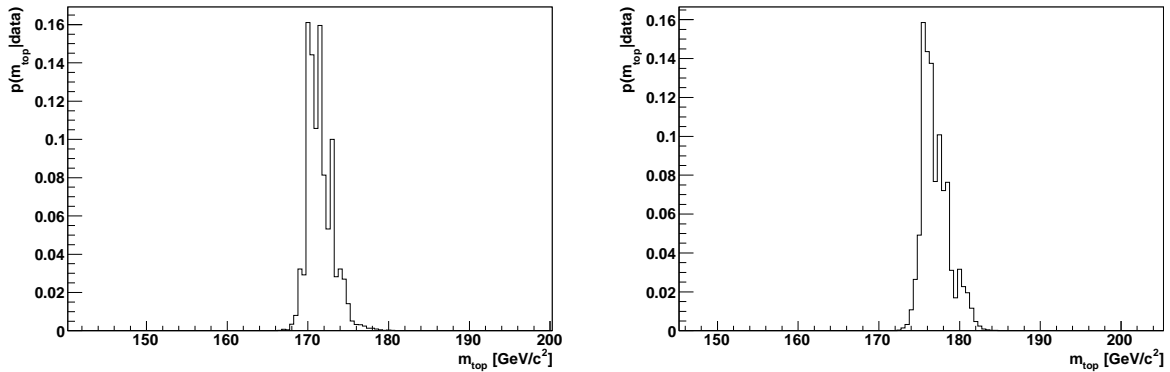
Figure 8.3: Estimation of the top pole mass including a background fraction for a pole mass of 170 GeV/c^2 (left) and 175 GeV/c^2 (right).

Results for the two data sets with respect to the expected background fraction are shown in Tab. 8.2. The mode and the respective differences of mode and quantiles are used to estimate the top quark mass and the corresponding uncertainties. The bins of the histograms have a width of 0.5 GeV/c^2. Due to this large binning, the modes of the distributions yield the same result with

gener. pole mass [GeV/c^2]	# Signal events	# Backgr. events	Estim. top mass [GeV/c^2]
170	167	0	$170^{+2.16}_{-0.57}$
170	167	20	$170^{+3.06}_{+0.01}$
175	250	0	$175.5^{+1.97}_{-0.71}$
175	250	30	$175.5^{+3.15}_{-0.04}$

Table 8.2: Results for pure signal samples and combinations of signal and background events. The fraction is built according to the expected number of signal and background events with respect to the production cross sections while all cuts are applied

and without using background events. Since the uncertainties of the top mass are asymmetric and the peak at 170 GeV/c^2 is very high, the lower uncertainty is positive for the distribution in Fig. 8.3 (b). Nevertheless, the top quark mass is estimated correctly for the signal sample with the LO MADGRAPH matrix element for both samples.

8.2.2. Purity studies

The effect of the sample purity on the top mass measurement is presented in Fig. 8.4. First, a pure signal sample has been used. A constant number of 100 events has been used while the background fraction has been increased about 5 % in each step. According to the previous event selection, the sample purity is about 0.89 ± 0.14. The systematic uncertainty on the top quark mass increases with increasing uncertainty of the sample purity. Due to the small number of remaining background events, the errors are very large. Up to a sample purity of 75 %, the result is still consistent with 1 σ. For larger background contributions, the top mass value is biased. This can be explained with the treatment of the background events. In this analysis, the matrix element for the background events is calculated like the matrix element for signal events. If the correct matrix element for W+jets production is used, no bias is expected. Nevertheless, more background events are needed to derive a conclusion for the used approach.
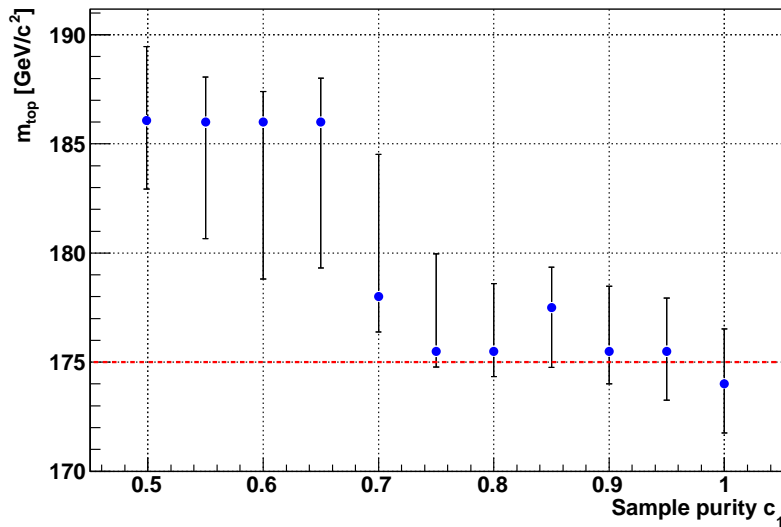

Figure 8.4: Studies of the top mass measurement depending on the sample purity.

8.2.3. Ensemble testing

Ensemble tests are done for the pure signal sample as well as for a signal sample with a certain background fraction. For the pure signal sample, 10,000 ensembles are built with 93 events each. The events are chosen randomly from a pool of 500 events. One event can be used several times in one ensemble. For the second ensemble test, 83 signal and 10 background samples are chosen randomly. These numbers refer to the remaining events after the selection cut and illustrate a sample purity of about 89 %.

In both top mass distributions, the mode is found to be at 176 GeV/c^2. The uncertainties are asymmetric and additional maxima occur. For the samples with background contribution, the additional peaks are much more pronounced. The distributions for the respective uncertainties do not change much. The distributions on the right side are only slightly broader than the ones on the left.

(a) Mode of ensembles with 93 signal events

(b) Left error of ensembles with 93 signal events

(c) Right error of ensembles with 93 signal events

(d) Mode of ensembles with 83 signal and 10 bkg. events

(e) Left error of ensembles with 83 signal and 10 bkg. events

(f) Right error of ensembles with 83 signal and 10 bkg. events
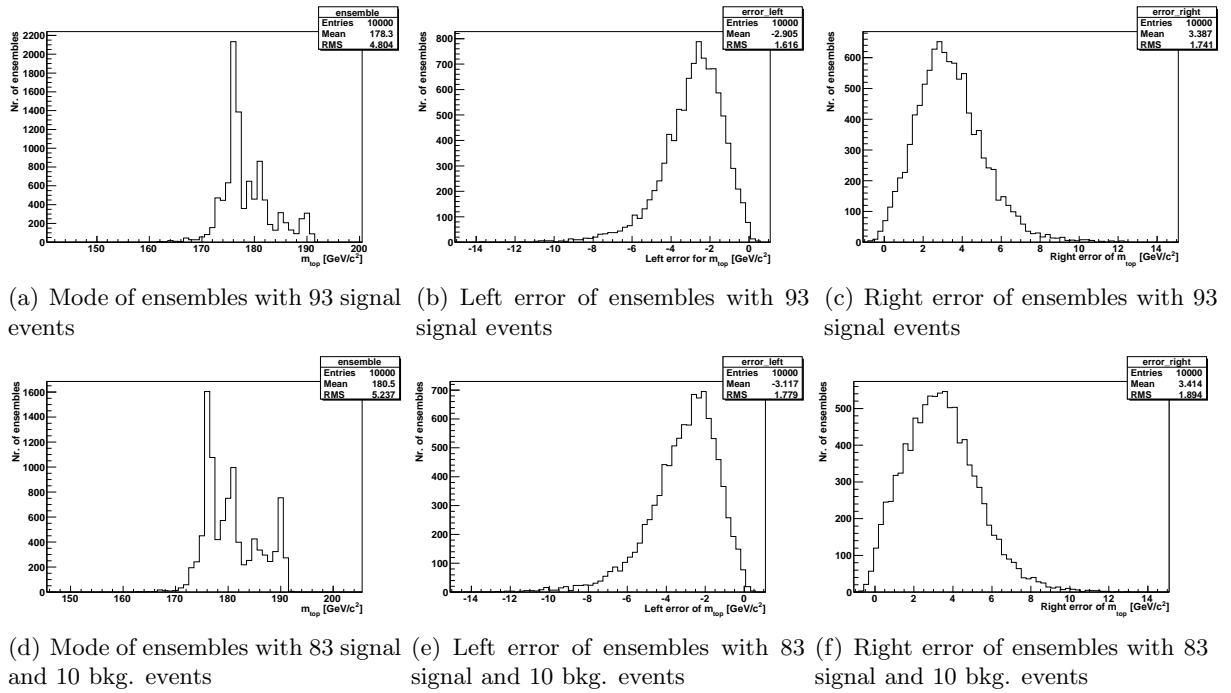
Figure 8.5: Ensemble testing for a pure signal sample (top) and a sample with 89 % purity (bottom). 93 events have been used for each ensemble. Due to the small number of events in the pool (500 signal and 55 background events), the correlation for the ensembles is high.

9. The Jet Energy Scale

Since the cross section for top pair production is very high at the LHC, the top quark mass measurement will only be limited by systematic uncertainties. The largest source of systematic uncertainty is the uncertainty on the *Jet Energy Scale (JES)* which has been discussed in Chapter 2. JES is only a generic term and consists of several contributions which can be dependent on the experiment or be physics related. Here, only a multiplicative factor will be applied without differentiation between the particular contributions. Studies on *in-situ* measurements of the JES have been done with smeared parton level events. The detector resolution is again parametrized by double Gaussians. A global JES is applied to all jets for the first tests. Later, the effect of two different JES for light quarks and b quarks is investigated.

9.1. Studies with a global JES

To study the effect of a global JES, the smeared four momenta of the partons from tree level events have been scaled with an additional multiplicative factor between 0.9 and 1.1 which is constant for one sample. In the analysis, the JES is introduced as an additional parameter of the likelihood. The transfer function has to be modified to [22]:

$$\frac{W(\frac{E_{\text{jet}}}{\text{JES}}, E_{\text{part}})}{\text{JES}} \qquad (9.1)$$

A two dimensional histogram of m_{top} vs. JES is created for each event. The results for the

(a) Correlation of m_{top} and the JES for one event   (b) Combination of 300 events
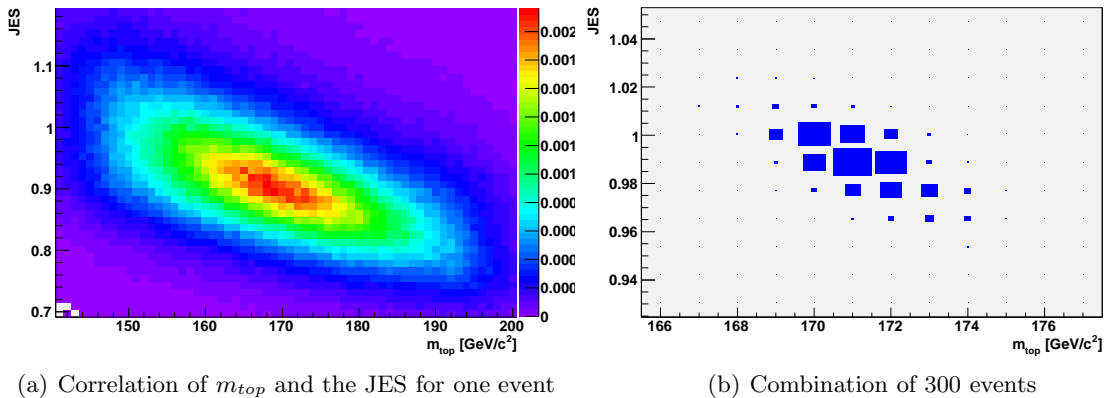
Figure 9.1: Two dimensional distributions for a single event (left) and the combination of 300 events. The sample is generated at a top mass of 170 GeV/c^2 and a JES of 1.0

combination of several events are obtained by normalizing the two dimensional distributions to unity and multiplying them bin-by-bin. In Fig. 9.1 a, one can see the probability distribution of the m_{top}-JES plane for a single event. The parameters are highly anti-correlated.

(a) Marginalized top mass distribution

(b) Marginalized JES distribution

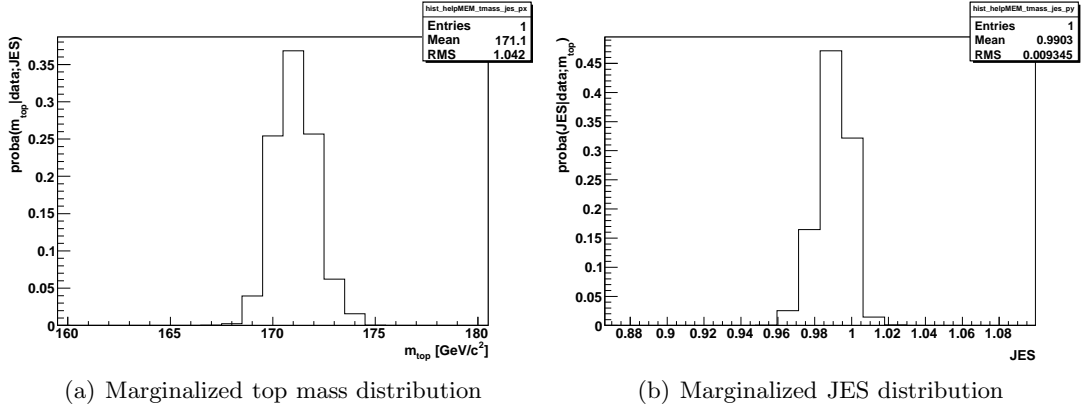***Figure 9.2:*** Marginalized distributions for a combination of 300 events (see Fig. 9.1). The sample is generated at a top mass of 170 GeV/c^2 and a JES of 1.0

The energies are scaled according to the factor in Eq. (9.1). If the JES is assumed to be larger than unity, one expects the real energies to be lower than the reconstructed ones.

Therefore, the top mass is estimated to be small. For JES smaller than unity, the estimated top mass is high. Figure 9.1 (b) shows the combination of 300 events of a data set generated at $m_{\text{top}} = 170$ GeV/c^2 and a JES of 1.0. In Fig. 9.2, the distribution is marginalized to the respective axes and the parameters are estimated from mode and the respective uncertainties. The resulting parameters yield:

$$m_{\text{top}} = 171^{+1.19}_{-1.05} \quad \text{GeV/c}^2 \tag{9.2}$$

$$\text{JES} = 0.988^{+0.012}_{-0.008} \tag{9.3}$$

9.1.1. Calibration curve

In Fig. 9.3, the calibration curve for the JES is shown. Three data sets have been generated at a pole mass of 170 GeV/c^2 and JES of 0.9, 1.0 and 1.1.
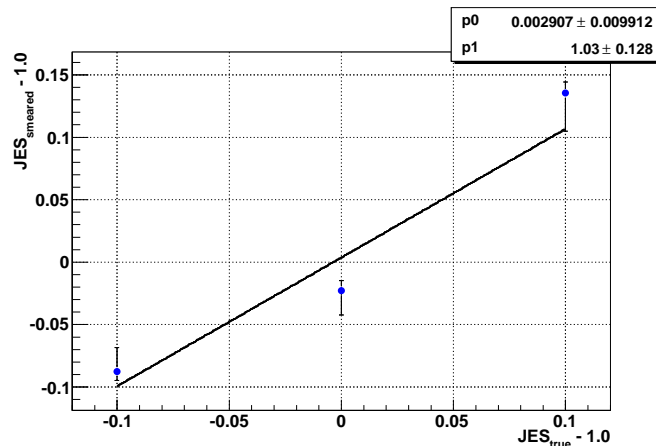

Figure 9.3: Calibration curve for the JES. A total of 125 events are used for each data set. Both, offset and slope are consistent within the uncertainties.

Due to limited computing power, only 125 events are used per data set. Although the uncertainties are very large, the parameters of the linear fit are consistent within the errors. The parameters of the calibration curve can be improved for larger statistics (see Fig. 9.2). These results show that the global JES seems to be estimated from the smeared parton level event without a bias. Nevertheless, more JES points should be used to get a more reliable result.

9.2. Studies of separate JES

Up to now, the same factor for the JES has been used for the light jets and the b jets. Different methods have been developed to measure the b-JES and the light JES separately (*e.g.* [51]). In this analysis, two independent multiplicative factors have been applied to smeared parton level events. Three samples are generated with JES from 0.9 – 1.1 while the same JES factor is used for the light JES and the b-JES.
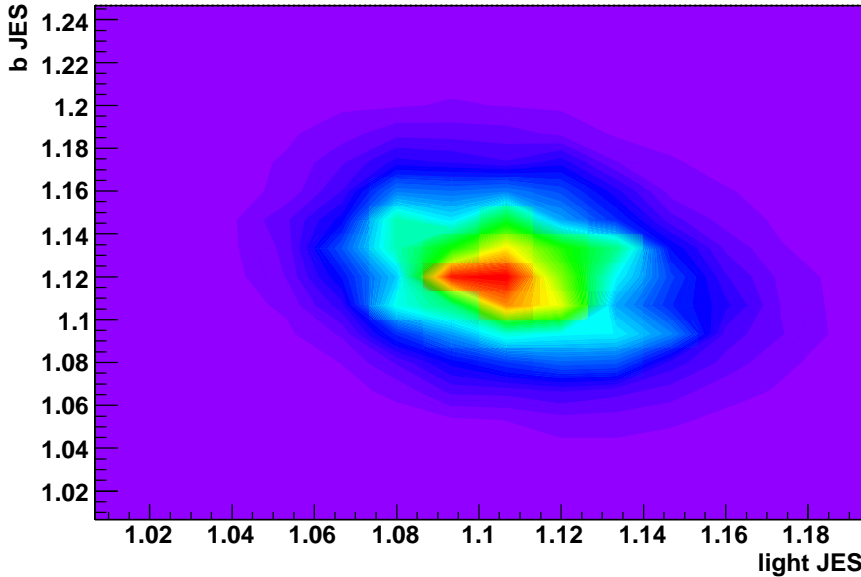
Figure 9.4: Correlation of light and b-JES for a combination of 150 events. The events are generated with JES = 1.1 for both parameters.

In Fig. 9.4, the resulting distribution from the combination of 150 events are shown. The sample is generated with a JES of 1.1. In Fig. 9.5, the marginalized distributions from Fig. 9.4 are presented: The parameters are estimated correctly to:

$$JES_{light} = 1.107^{+0.025}_{-0.022} \tag{9.4}$$

$$JES_b = 1.12^{+0.028}_{-0.024} \tag{9.5}$$

(a) (b)

Figure 9.5: Marginalized distributions from Fig.9.4. The events are generated with a JES of 1.1 for both parameters. From the marginalized distributions (left and right), the respective JES can be estimated.

9.2.1. Calibration curves

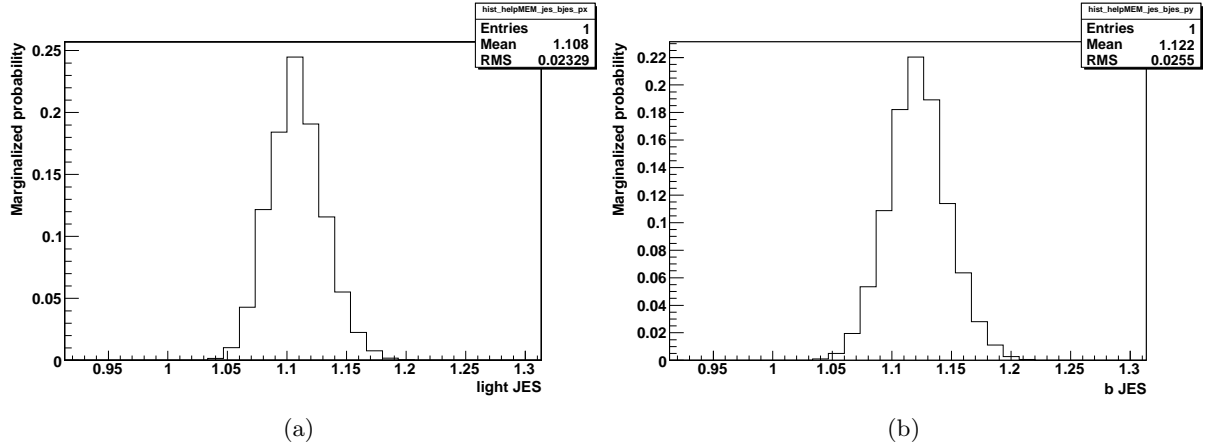Calibration curves are made for the two JES separately. The two dimensional JES distributions are normalized to unity and 150 events are combined for each data set. The samples are generated with a JES of 0.9, 1.0 and 1.1 both for the light and the b JES. The two dimensional distributions are marginalized to the respective axes. The mode and the quantiles have been used to estimate the JES.

(a) Calibration curve for the light JES (b) Calibration curve for the b JES

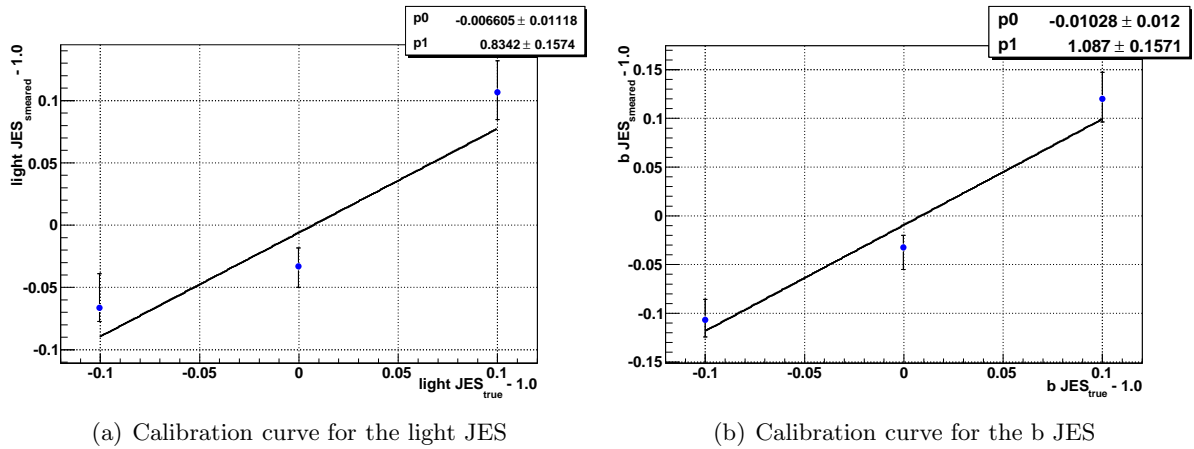***Figure 9.6:*** Calibration curves for the respective JES are made. For each point, 150 events are combined. The results from the linear fit are consistent with the expectation within two standard deviations (left) and one standard deviation (right)

The results of the respective fits are consistent within two (left) and one (right) standard deviations. As a result, the two JES can be estimated correctly from independent parameters.

10. Conclusion and Outlook

In the following, the results of the analysis are discussed and summarized. An outlook is given to further studies and improvements of the method.

10.1. Numerical integration

The Matrix Element Method was successfully implemented using Markov Chain Monte Carlo for the numerical integration. The method has been tested for estimations of the top quark mass. The integration via Markov Chains has been compared with VEGAS. The results showed a good agreement. The integration over two different integration variables has been performed. The compared distributions show that both integration methods are consistent.

10.2. Matrix element method and JES

The matrix element method has been tested on top quark events. LO Monte Carlo data has been analyzed with a LO matrix element for symmetric and asymmetric transfer functions. The estimated parameters are unbiased. NLO MC events have been successfully analyzed with a LO matrix element. The top quark mass has been correctly estimated from signal events with a LO matrix element. $W + 4\ jets$ background has been used to study the influence of different background fractions on the top quark mass estimation. For background contamination, a bias can be observed due to the treatment of the background events.

Studies on a global JES and two separate JES have been performed with smeared parton level events. The top quark mass and the JES are estimated without bias and an in-situ calibration can be done.

10.3. Outlook

The method has been successfully implemented. Its most important drawback is the large amount of CPU time which is needed to analyze the data. One reason for this is caused by the calculation via the MADGRAPH interface. The transition from C++ to Fortran and vice versa as well as the calculation of the helicity amplitudes requires most of the time. The code has to be tuned to be more efficient. A possible solution would be to write a program for this which is directly included in the method. Efforts on this have already been made but could not been finished up to now. Moreover, a variable transformation and further assumptions (e.g. narrow width approximation) would simplify the integration.

B-tagging information could be used to reduce the combinatorial background. The same could probably be done with the m_W cut applied in Chapter 7.3. The top with the highest transverse momentum is identified. If now the mass of one dijet combination fullfills the condition $|m_W - m_{jj}| < 10$ GeV, the remaining jet can be handled as the b-jet from the hadronically decaying top quark. The fourth jet is then treated as the b-jet from the leptonically decaying top quark.

If further data sets at different pole masses are provided, studies on the detector acceptance can be performed. Furthermore, a calibration curve can be made for the NLO data. To study the effect of the W+jets background in more detail, more background events are needed. Besides, a matrix element for the background events can be implemented.

A. Phase Space Transformation

In the method described in Chapter 5, the integration is carried out over the 3-momenta of the final state particles. While the angles are assumed to be perfectly measured, the number of parameters can be reduced to the energies of the five measured particles. Therefore, a variable transformation has to be made. Furthermore, the integration over the W mass results in two solutions for the longitudinal momentum of the neutrino. These are calculated in A.2.

A.1. Variable transformation

The absolute value of the three momentum is caluclated as:

$$|\vec{p}| = \sqrt{E^2 - m^2} \tag{A.1}$$

The derivative has to be made:

$$\frac{d|\vec{p}|}{d\,E} = \frac{E}{|\vec{p}|} \tag{A.2}$$

$$\Rightarrow \frac{1}{E} = \frac{dE}{d|\vec{p}||\vec{p}|} \tag{A.3}$$

With the volume element in the momentum space

$$dV = p^2 d\,p \, d\,\Omega \, 4\pi \tag{A.4}$$

one gets the following result:

$$\frac{d^3 p_i}{(2\,\pi)^3 2\,E_i} = \frac{|\vec{p}|}{(2\,\pi)^2}\,dE \tag{A.5}$$

A.2. Calculation of the two neutrino solutions

One of the W bosons decays in a charged lepton and its corresponding neutrinos. While only the charged lepton can be measured, the two neutrino solutions have to be measured from the W mass:

$$m_W^2 = (p_\nu + p_l)^2 \tag{A.6}$$

$$= 2\,E_\nu\,E_l - 2(p_x^\nu\,p_x^l + p_y^\nu\,p_y^l + p_z^\nu\,p_z^l) \tag{A.7}$$

The masses of the leptons are neglected here.

$$\underbrace{\frac{m_W^2}{2} + (p_x^\nu\,p_x^l + p_y^\nu\,p_y^l)}_{\alpha} = E_\nu\,E_l - p_z^\nu\,p_z^l \tag{A.8}$$

$$p_T^2 = p_x^2 + p_y^2 \tag{A.9}$$

The two solutions of the quadratic equation are:

$$p_z^\nu = \alpha \frac{p_z^l}{p_{T,l}^2} \pm \frac{\sqrt{E_l^2(\alpha^2 - p_{T,\nu}^2 p_{T,l}^2)}}{p_{T,l}^2} \tag{A.10}$$

For each solution, the corresponding energy has to be calculated.

B. Cut Efficiency

In the following distributions, the effect of different cuts are shown. The cuts are applied according to the definition in Chapter 7.3. If four preselected jets are required, the number of background events can be reduced (cp. Fig. B.1). In the top mass distribution, a clear peak can be seen in the signal distribution while the trijet combination of the background events is relatively flat.

(a) Comparison of the $E_{T_{miss}}$ distributions

(b) Multiplicity of all preselected jets

(c) Multiplicity of *nonloose* jets

(d) Multiplicity of *good* jets

(e) Distribution of the reconstructed m_{top}

(f) Distribution of the top p_T

Figure B.1: Distributions for missing transverse energy, jet multiplicities, reconstructed trijet mass and corresponding transverse momentum for preselected jets.

If the W mass cut is applied in addition, the background contribution decreases significantly. While in the signal sample a W boson decays hadronically, the jets in the background sample do not come from a W boson.

(a) Comparison of the $E_{T_{miss}}$ distributions

(b) Multiplicity of all preselected jets

(c) Multiplicity of *nonloose* jets

(d) Multiplicity of *good* jets

(e) Distribution of the reconstructed m_{top}
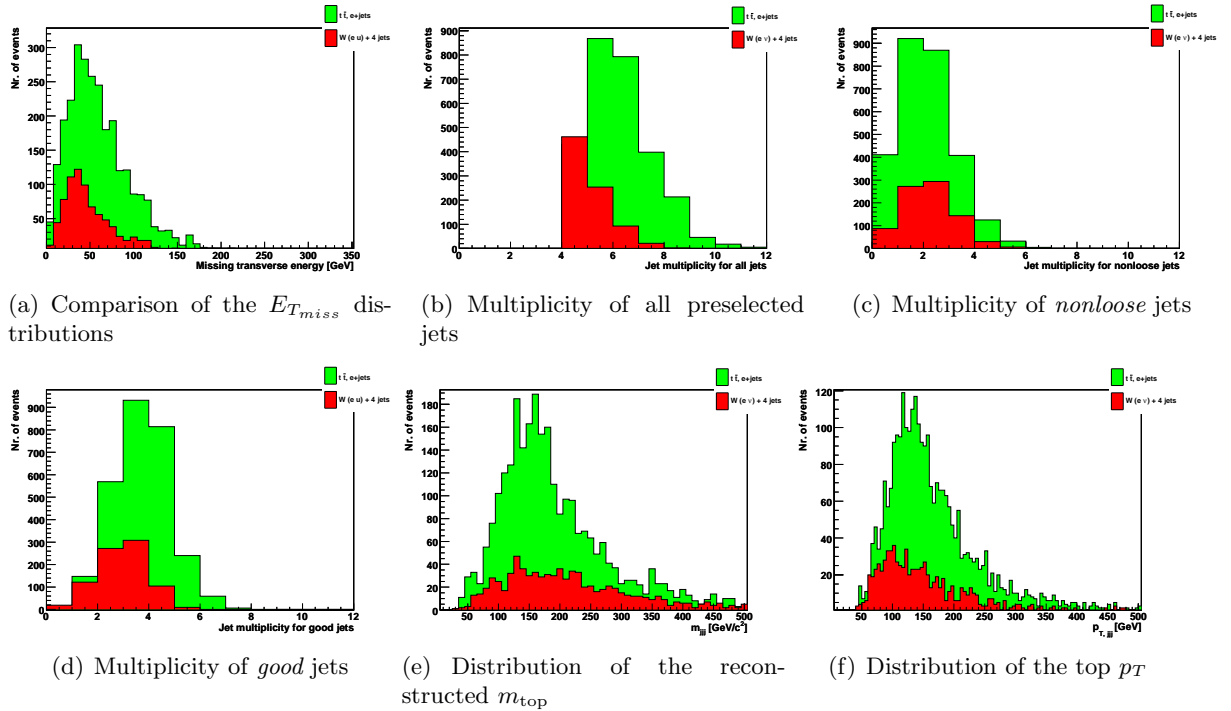
(f) Distribution of the top p_T

Figure B.2: Distributions for missing transverse energy, jet multiplicities, reconstructed trijet mass and corresponding transverse momentum for preselected jets. In addition, the W mass cut is applied.

(a) Comparison of the $E_{T_{miss}}$ distributions

(b) Multiplicity of all preselected jets

(c) Multiplicity of *nonloose* jets

(d) Multiplicity of *good* jets

(e) Distribution of the reconstructed m_{top}
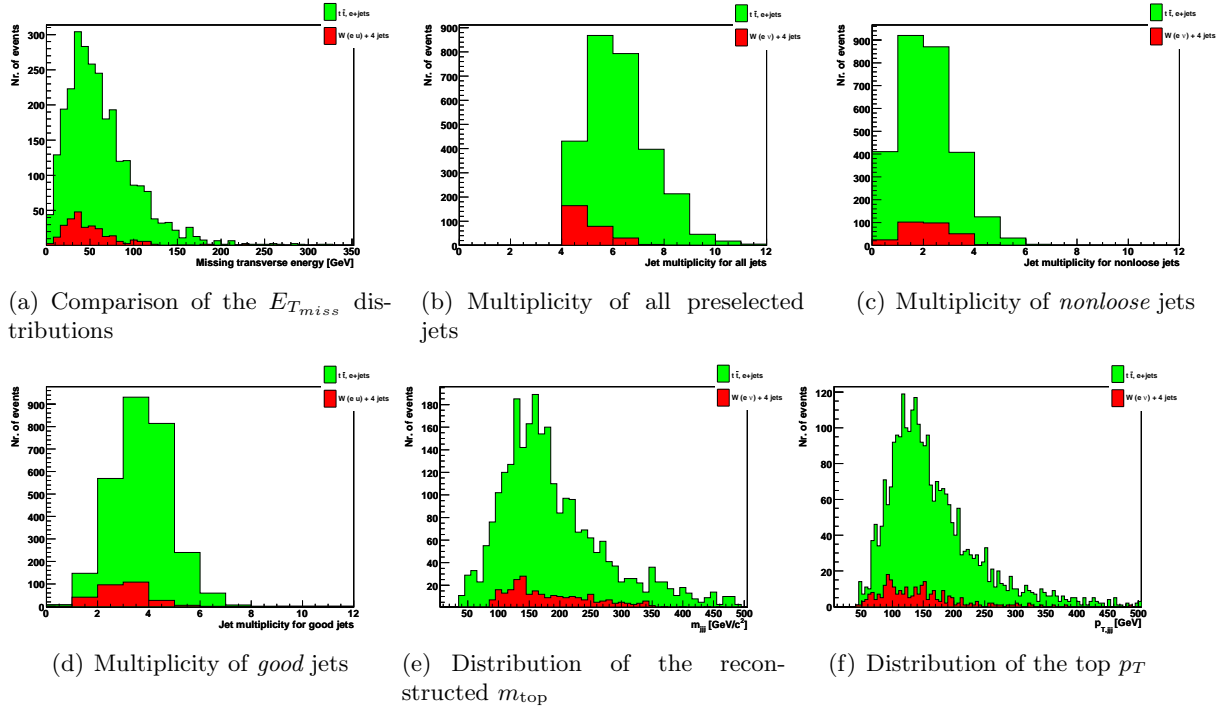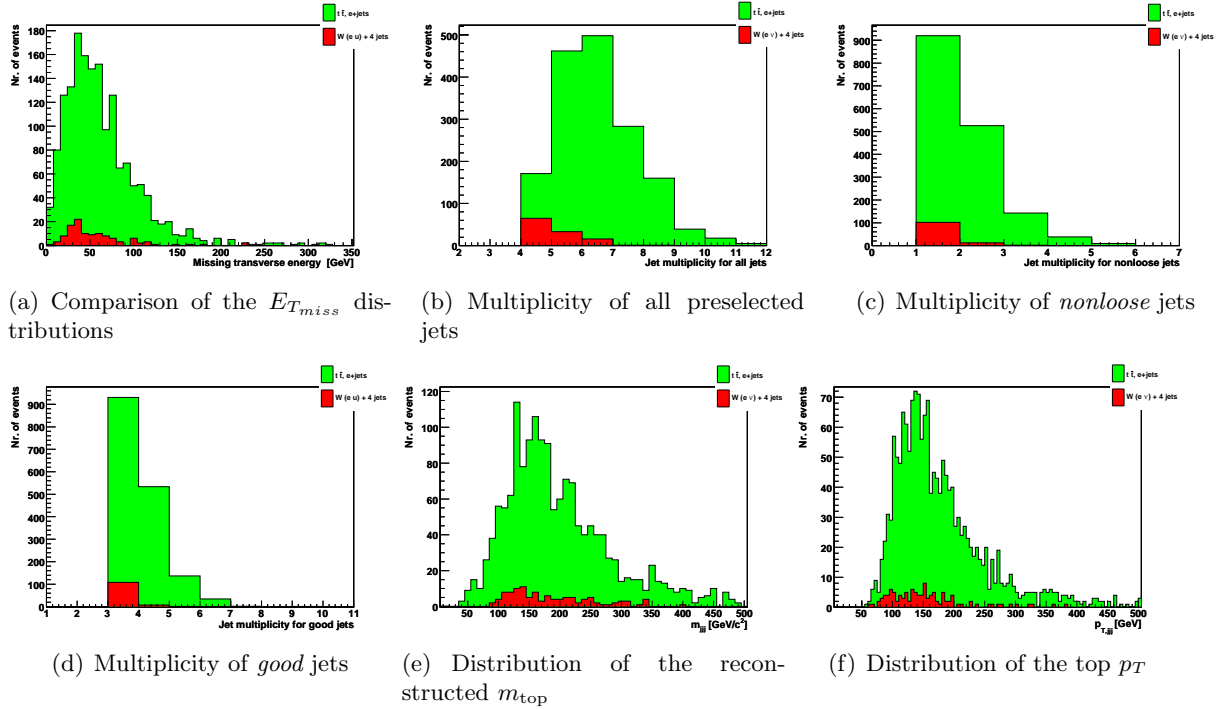
(f) Distribution of the top p_T

Figure B.3: Distributions for missing transverse energy, jet multiplicities, reconstructed trijet mass and corresponding transverse momentum for preselected jets. Three good jets are required and the W mass cut is applied. The number of background events is reduced significantly.

Bibliography

[1] W.N. Cottingham and D.A. Greenwood. An Introduction to the Standard Model of Particle Physics, 2007. Cambridge University Press.

[2] D. Griffith. Introduction to Elementary Particles, 2008.

[3] The Tevatron Electroweak Working Group, for the CDF, and DØ Collaborations. Combination of CDF and D0 Results on the Mass of the Top Quark. 2009. hep-ex/0903.2503.

[4] C. Amsler et al. (Particle data group). Review of Particle Physics 2008, 2008.

[5] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13(16):508–509, Oct 1964.

[6] http://www.hep.ph.ic.ac.uk/cms/physics/.

[7] K. Ichimura. Recent Results from KamLAND. 2008. hep-ex/0810.3448.

[8] S. P. Martin. A Supersymmetry Primer. 1997. hep-ph/9709356.

[9] J.M. Campbell et al. Hard interactions of quarks and gluons: a primer for lhc physics. *Rep. Prog. Phys.*, 70.

[10] R.S. Thorne, A.D. Martin, and W.J. Stirling. MRST parton distributions: Status 2006. 2006. hep-ph/0606244.

[11] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002. hep-ph/0201195.

[12] A. Quadt. Top quark physics at hadron colliders. *Eur. Phys. J.C*, 48:835–1000, 2006.

[13] Cacciari et al. Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC. *JHEP*, 09:127, 2008. hep-ph/0804.2800.

[14] P. De Jong. Top Physics at the LHC, 2008. ATL-PHYS-PROC-2008-017.

[15] V. M. Abazov et al. Observation of Single Top Quark Production. 2009. hep-ex/0903.0850.

[16] T. Aaltonen et al. First Observation of Electroweak Single Top Quark Production. 2009. hep-ex/0903.0885.

[17] M. Cristinziani and G. Petrucciani. Single top: prospects at lhc, 2008. hep-ex/0808.0565.

[18] http://www.fnal.gov/pub/presspass/images/single-top-images.html.

[19] M.-A. Pleier. Review of Top Quark Properties Measurements at the Tevatron, 2008. hep-ex/0810.5226.

[20] T. Wyatt. High-energy colliders and the rise of the standard model. *Nature*, 448(7151):274–280, 2007.

[21] `http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2009`.

[22] V.M. Abazov et al. Measurement of the top quark mass in the lepton + jets final state with the matrix element method. *Phys.Rev.*, D74:092005, 2006. hep-ex/0609053.

[23] P. Collier O. Brüning. Building a behemoth. *Nature*, 448(7151):285–289, jul 2007.

[24] `http://public.web.cern.ch/public/en/Research/AccelComplex-en.html`.

[25] ATLAS collaboration. ATLAS detector and physics performance. 1999. ATLAS TDR 14, CERN/LHCC 99-14;

[26] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[27] K. Kleinknecht. Detektoren für Teilchenstrahlung, Teubner. 4. Auflage.

[28] C. Amsler et al. Passage of particles through matter. `http://pdg.lbl.gov`, 2008.

[29] V.M. Abazov et al. (DØ-Collaboration). A precision measurement of the mass of the top quark. *Nature*, 429:638–642, 2004.

[30] K. Kröninger,. A measurement of the Top Quark Mass with the D0 Detector at $\sqrt{s} = 1.96$ TeV using the Matrix Element Method, 2004. Diploma thesis. University of Bonn.

[31] P. Schieferdecker. Measurement of the Top Quark Mass at D0 Run II with the Matrix Element Method in the Lepton+Jets Final State, 2005, PhD thesis, LMU Munich.

[32] A. Caldwell, D. Kollar, and K. Kröninger. Bat - the bayesian analysis toolkit, 2008. data - an/0808.2552.

[33] T. Stelzer and W. F. Long. Automatic generation of tree level helicity amplitudes. *Comput. Phys. Commun.*, 81:357–371, 1994. hep-ph/9401258.

[34] F. Maltoni and T. Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 02:027, 2003. hep-ph/0208156.

[35] M. R. Whalley, D. Bourilkov, and R. C. Group. The Les Houches Accord PDFs (LHAPDF) and Lhaglue. 2005. hep-ph/0508110.

[36] `http://public.web.cern.ch/public/en/Research/AccelComplex-en.html`.

[37] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002. hep-ph/0204244.

[38] S. Frixione, P. Nason, and B. R. Webber. Matching NLO QCD and parton showers in heavy flavour production. *JHEP*, 08:007, 2003. hep-ph/0305252.

[39] G. Corcella et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001. hep-ph/0011363.

[40] J. Allison, et al. Geant4 developments and applications. *Nuclear Science, IEEE Transactions on*, 53(1):207–278, 02 2006.

[41] S. Agostinelli et al. G4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research*, A 506, Issue 3:250–303, 2003.

[42] A. Shibata. Topview - atlas top physics analysis package. Technical Report ATL-SOFT-PUB-2007-002. ATL-COM-SOFT-2007-006. CERN-ATL-COM-SOFT-2007-006, CERN, Geneva, May 2007.

[43] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003. hep-ph/0206293.

[44] Kevin Kröninger. Bayesian Analysis Tool, Nov. 2008. Terascale Statistical Tools Group Meeting.

[45] A. Gelman and D. Rubin. Inference from iterative simulations using multiple sequences, 1992. Statistical Science.

[46] Kevin Kröninger. The Basics of the Bayesian Analysis Toolkit, 2009. Talk at the BAT day, 14. 01. 2009.

[47] G. P. Lepage. Vegas - an adaptive multi-dimensional integration program, 1978.

[48] V.Blobel/E.Lohrmann. Statistische und numerische Methoden der Datenanalyse, 1998. Teubner.

[49] T. Hahn. The CUBA library. *Nucl. Instrum. Meth.*, A559:273–277, 2006. hep-ph/0509016.

[50] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. 2009. 0901.0512.

[51] F. Fiedler. Independent measurement of the top quark mass and the light- and bottom-jet energy scales at hadron colliders. *Eur. Phys. J.*, C53:41–48, 2008. hep-ex/706.1640.

Acknowledgements

Last but not least I want to thank all people who supported and encouraged me in the last month.

I would like to thank Prof. Arnulf Quadt who managed to fascinate me for the exciting field of high energy physics and gave me the possibility of joining the unique experience of working in such an international environment. The discussions with him always helped me to see things in a wider context.

Special thanks goes to Dr. Frank Fiedler for being the coreferee for this thesis.

Moreover, I would like to thank Kevin Kröninger who supervised me in the last eighteen month. He was always ready to help and his patient and encouraging character was a great motivation. Without his support, this thesis would not have been possible.

Furthermore, I would like to thank the whole HEP working group for the extraordinary team spirit. Particularly I would like to thank Su-Jung Park and Johannes Erdmann for discussions and answering questions and Jörg Meyer for the support concerning the batch system especially in the last weeks.

I would like to thank all the people who supported me in the last five years, my friends at home and in university. Special thanks goes to Markus Osterhoff who worked with me in several lab courses and lectures and helped to review this thesis.

Above all, I would like to thank my parents and my sister Marion for supporting me financially and mentally. They were always interested in my work and helped me with constructive questions. Special thanks goes to my sister who helped me reviewing parts of this thesis.

Finally, I would like to thank my boyfriend Markus for being there for me all the time. Without his smiling and caring character, things would have been always much more difficult.