GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Bachelor's Thesis

II.Physik-UniGö-Bach2009/02

Studien zur Messung des Top-Antitop-Paarproduktionswirkungsquerschnittes im Lepton+Jets-Kanal bei ATLAS am LHC

Studies for the Measurement of the $t\bar{t}$ Production Cross Section in the Lepton+Jets Channel with ATLAS at the LHC

prepared at the II. Institute of Physics
by Henrike Fleischhack from Göttingen

Thesis period:	15th April 2009 until 20th July 2009
Supervisors:	Dr. Jörg Meyer, Dipl. Phys. Adam Roe
First Referee:	Prof. Dr. Arnulf Quadt
Second Referee:	Prof. Dr. Ariane Frey
External Referee:	Prof. Dr. Martine Bosman

Contents

1. Introduction **1**

2. Theoretical Overview **3**

2.1. The Standard Model of Particle Physics 3

2.2. The Proton 4

2.3. The Top Quark 4

2.3.1. Production 5

2.3.2. Decay 6

2.3.3. The Lepton+Jets Channel 6

2.3.4. Background 7

3. Experimental Overview **9**

3.1. Collider Physics 9

3.2. Important Physical Quantities 11

3.2.1. Luminosity, Counting Rate and Cross Section 11

3.2.2. Detector Coordinates 11

3.3. Experimental Setup 12

3.3.1. The LHC 12

3.3.2. The ATLAS Detector 12

4. Monte Carlo Samples **15**

5. Methods **19**

5.1. The Cut and Count Method for Cross Section Determination 19

5.1.1. Cut Efficiencies and their Uncertainties 20

5.2. The Matrix Method for QCD Background Determination 21

5.2.1. Uncertainties 22

6. Results **23**

6.1. Properties of $t\bar{t}$ Events and Dominant Backgrounds 23
6.2. Preselection Cuts 26
6.3. W Boson and Top Quark Mass Constraints 28
6.4. Jet-Parton Matching 30
6.5. Top Quark Candidate Algorithm 31
6.6. W Boson Candidate Algorithm 33
6.7. Re-evaluation of CSC preselection cuts 34
6.7.1. Uncertainties 39
6.8. Evaluation of QCD Background Using the Matrix Method 39
6.8.1. Muon Isolation Criteria 39
6.8.2. Composition of the Loose and Tight Sample 41
6.8.3. Determination of Isolation Efficiencies 42
6.8.4. Results 43

7. Summary and Outlook **45**

A. Tables **47**

B. Plots **51**

Bibliography **55**

1. Introduction

In the field of high energy physics, properties and interactions of elementary particles are studied. Particle accelerators are one set of the tools that can be used for this. In winter 2009, the new Large Hadron Collider (LHC) at CERN in Geneva, CH will start operating. It will reach energies up to seven times of what the Tevatron, the most powerful accelerator in use today, is capable of.
Top quark physics will play an important role during the whole runtime of the LHC. Rediscovering the top quark should only take a few weeks or months. During the early stages, the well-understood topology of top quark decays can be used to gain better understanding of the detector's behaviour and to calibrate some parts of the detector, e.g. determining the jet energy scale or the b-tagging efficiency.
Once detector and background processes are well understood, precision measurements of the top quark and some of its properties (e.g. mass and production cross section) can take place. Even with a few hundred inverse picobarns of data, it should be possible to determine these quantities. The measured values can then be compared to theoretical predictions or (in the case of the mass of the top quark) to previous measurements at the Tevatron. These measurements will be an important test of the Standard Model of particle physics and its predictions. Significant discrepancies will point at new physics beyond the Standard Model.
In this thesis, I present some simple studies for measuring the top quark pair production cross section. I will begin with a short summary of the Standard Model and its predictions regarding the top quark. The experimental setup at the LHC will be described. The ATLAS detector, one of the two multi-purpose detectors at the LHC, will be discussed in more detail.
As no real collision data is available from the LHC yet, the studies described here were done using Monte Carlo generated, simulated data. Doing these studies on simulated data at first is important as it will help understand the measured data later. I will describe two methods that can be used during cross section measurements, the cut-and-count method for cross section determination and the matrix

method for background estimation. I will show how these methods can be applied to simulated data samples. I will also test the performance of a simple algorithm for reconstructing the hadronically decaying top quark and the preselection cuts described in the CSC book [1].

2. Theoretical Overview

2.1. The Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a set of quantum field theories that describe subatomic particles and their interactions. Particles can be divided into bosons, with integer spin, and fermions, with half-odd-integer spin. Elementary fermions can again be divided into quarks, which carry colour charge, and leptons, which do not. We currently know three generations of fermions, each consisting of two quarks (one with charge $+\frac{2}{3}$ and one with charge $-\frac{1}{3}$), one charged lepton and one uncharged lepton (neutrino) (cf. fig. 2.1). For each particle there exists a corresponding anti-particle with the same mass but opposite internal quantum numbers, e.g. charge.
Forces between particles are mediated by bosons. The SM contains three fundamental forces: the electromagnetic force, mediated by the massless photon and acting on electrically charged particles, the strong force, mediated by eight massless gluons

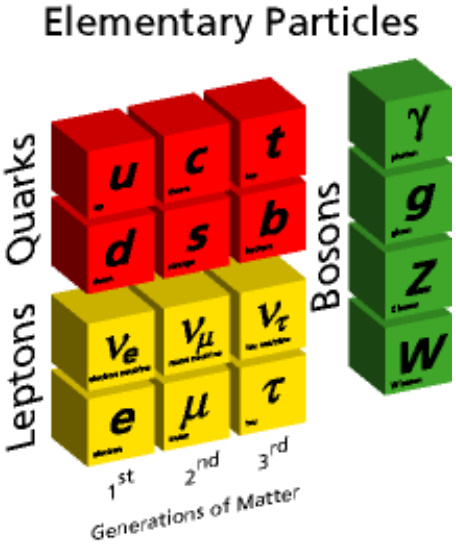

Figure 2.1.: Constituents of the Standard Model.

and acting on particles which carry colour charge (quarks and gluons), and the weak force, mediated by the massive W and Z bosons and acting on all fermions as well as W and Z bosons. All of the particles mentioned so far have been observed. There is however one other fundamental particle predicted by the SM, the Higgs boson, which has not been observed yet. It is needed to give mass to fermions as well as W and Z bosons. Searches for it are already being undertaken.
A more detailed introduction to the Standard Model and its predictions can be found in [2].

2.2. The Proton

As this thesis will deal with hadron collider physics, it is important to understand the structure of the protons that are used in the collisions.
Protons are composite particles, made up of a number of partons. The number and properties of the observed partons depend on the probe that is used to measure them, or rather on the momentum transferred between probe and proton. At low momentum transfers, the proton appears as an elementary particle: There is only one parton carrying the full proton momentum and energy. At higher momentum transfers one can observe three valence quarks (uud) as well as gluons and quark-antiquark pairs, the sea quarks. Each parton carries a fraction x of the proton's momentum and energy. The parton distribution functions which have been measured over a wide range of momentum transfers give the probability of finding a parton with momentum fraction x. Generally, gluons dominate at low x, valence quarks at high x.

2.3. The Top Quark

With a mass of $173.1 \pm 1.3\,\text{GeV}$ [3], the top quark is not only the heaviest quark, but the heaviest known fundamental particle. It has a charge of $+\frac{2}{3}$ and is a part of the weak isospin doublet also containing the bottom quark. The top quark was discovered in 1995 at the Tevatron [4, 5], which is the only collider to date where top quarks have been observed, but its existence had been predicted in 1973 [6]. It has not definitely been shown that the particle discovered at the Tevatron is indeed the top quark predicted by the Standard Model. Its charge has not been measured directly, so it could be an exotic particle with charge $-\frac{4}{3}$. Current measurements favor an SM top quark with charge $+\frac{2}{3}$ [7, 8].

The mean lifetime of the top quark has not been directly measured. The SM predicts a very short lifetime of about $\tau_{top} \approx 4 \cdot 10^{-25}\,\mathrm{s}$ [2]. This makes the top quark the particle with the shortest lifetime known so far. Its lifetime is too short to form bound states.

2.3.1. Production

At hadron colliders, top quark pairs are predominantly produced via the strong force. At the Tevatron, quark-antiquark annihilation is the dominant channel for $t\bar{t}$ production (cf. figure 2.3). At the LHC, gluon-gluon fusion (cf. figure 2.2) will dominate [9]. The expected cross section for $t\bar{t}$ pair production at the LHC at 10 TeV is about $401.6^{+3.7\%}_{4.3\%}(scales)^{4.6\%}_{4.5\%}(PDFs)$ pb [10]. At the Tevatron, $t\bar{t}$ pairs are produced with a much lower cross section, due to the lower available energy. A recent measurement is $\sigma_{t\bar{t}} = 8.18^{+0.98}_{-0.87}$ pb (combined DØ results) [11].
Single top quarks can be produced in hadron colliders via the electroweak force. This production mechanism was observed at the Tevatron and discovered in March 2009 [12, 13].

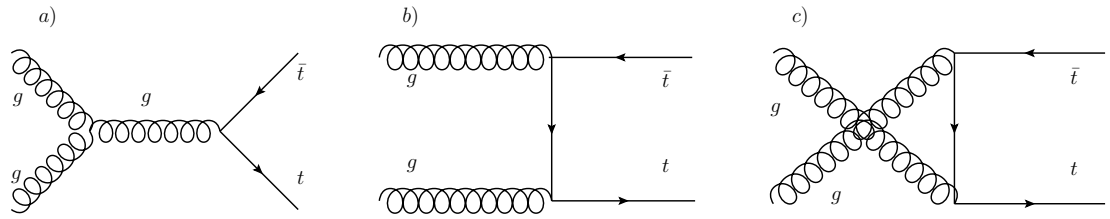

Figure 2.2.: $t\bar{t}$ production via gluon-gluon fusion. a) s-channel b) t-channel c) u-channel.

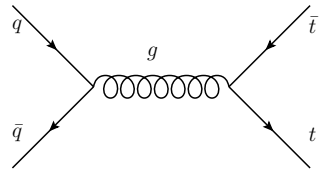

Figure 2.3.: $t\bar{t}$ production via $q\bar{q}$ annihilation.

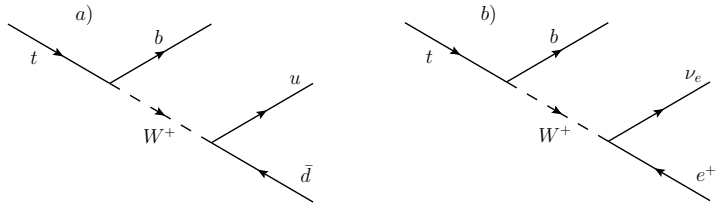

Figure 2.4.: a) hadronic, b) leptonic decay of the top quark.

2.3.2. Decay

Because of its large mass, the top quark decays weakly before hadronisation can take place. The overwhelmingly dominant channel is $t \to b\,W$. Decays into strange or down quarks are suppressed by the CKM matrix. The W boson can decay into a pair of quarks (usually ud or cs) or into a lepton and its corresponding neutrino (cf. figure 2.4).
For a $t\bar{t}$ pair, one therefore finds three classes of decay channels: The all jets channel, where both of the W bosons decay hadronically, the dilepton channel (both W bosons decay leptonically) and the semi-leptonic (lepton+jets) channel (one W boson decays hadronically, the other leptonically). Using the approximation $|V_{tb}| = 1$, one can calculate the branching ratios for each channel from the branching ratios of the W boson. Decay channels containing τ leptons play a special role: The τ lepton is so short lived that it decays before it enters the detector. As the τ can decay hadronically or leptonically, these decays then mimic different decay channels.

2.3.3. The Lepton+Jets Channel

Most of the following studies use events in the electron+jets channel, for which one expects a branching ratio of about $2 \cdot \frac{1}{9} \cdot \frac{2}{3} \approx 0.1482$. τ leptons decaying into an electron and two neutrinos were considered as electrons for this analysis, τ leptons decaying hadronically were not considered as jets. As $BR(\tau \to e\,\bar{\nu}_e\,\nu_\tau) = 0.1785$ [9], the total branching rate for our "signal" channel is $BR = 0.1746$.
The semi-leptonic channels are often used for analyses because (unlike the all hadronic channel) they provide a clean signature and have less ambiguities than the dilepton channels. The following signature is expected for e+jet events:

- One high-p_T electron from the W boson to trigger on
- 4 high-p_T jets (two b-jets and two light quark jets from the W boson)
- (real) missing energy due to the neutrino.

Due to imperfect detector acceptance and reconstruction efficiency, one might not see all of the above objects. There could also be more than four jets due to, for example, gluon radiation or recoil jets.
The signature of the μ+jets channel is similar to the above, with a high-p_T isolated muon instead of an electron.

2.3.4. Background

Even with the relatively clean signature of $t\bar{t}$ events in the lepton+jets channel, there are other types of events that mimic the $t\bar{t}$ signatures. These background events cannot be distinguished from the signal events on an event-by-event basis, but cuts can be introduced to filter out most of them. However, knowledge of the background processes is important as there will always be some events that cannot be filtered out (irreducible background). The most important sources of background are:

- ***W*+jets**: W bosons are produced along with QCD jets and decay into an electron (or muon) and a neutrino.
- ***Z*+jets**: Z bosons are produced along with QCD jets and decay into an electron-positron or a muon-antimuon pair. One of the leptons is not detected, resulting in missing energy.
- **single top** quarks decaying leptonically, with extra QCD jets
- $\boldsymbol{t\bar{t}}$ events in other decay channels, for example dileptonic decays where one of the leptons is not reconstructed properly.
- QCD **multijet** production with fake electrons or fake isolated muons (e.g. from b-jets).

For the studies in the electron+jets channel, W+jets events as well as $t\bar{t}$ events in the dilepton channel containing at least one electron (on Monte Carlo level) were considered as background. Those processes are the main sources of background. The combined cross section for the W+jets samples used ($W \rightarrow e\nu$+3, 4, or 5 partons) is about 337 pb [14], while the expected total cross section for top quark pair production is 401.6 pb [10]. The process $Z \rightarrow ee$+3, 4, or 5 partons has an expected total cross section of about 30 pb [14]. It was not considered as background for these studies as the contributions from $W \rightarrow e\nu$+jets events dominate.
QCD background in the μ+jets channel was studied using a matrix method. Data-driven methods to deal with this background are especially important as the current Monte Carlo models of QCD background with fake isolated muons are unreliable.

3. Experimental Overview

3.1. Collider Physics

Particle accelerators are an important tool for experimental particle physicists. They are used to accelerate beams of charged particles (usually electrons or protons as well as their anti-particles). These beams can be used for two different types of experiments: fixed target experiments and colliding beam experiments where two high-energy beams collide. The latter type of experiments needs more fine tuning to ensure that the beams collide at the right place. They are widely used as they allow for higher center-of-mass energies to be reached.
At the collision point, particles can scatter off each other, possibly producing other particles. These collisions usually happen inside a detector so that the reaction products can be detected and their properties (e.g. energy and momentum) can be measured. Most detectors that are in use today follow a layered design. The innermost layer (around the beam pipe) is used to track charged particles using, for example, wire chambers or silicon sensors. As the tracking system is usually surrounded by a magnetic field, the tracks can be used to determine the charge and momentum of the measured particles as well as the location of the interaction point. The tracking system is set up so that particles interact minimally with the detector and do not lose much energy.
The next layers are formed by the calorimeters, which are designed so that most particles deposit all of their energy there. The energy of the particle can be determined from the measured energy deposition. Most calorimeters are split into two parts. The first one is the electromagnetic calorimeter, in which electrons, photons, and neutral pions (which quickly decay into photons) deposit most of their energy. Electrons emit bremsstrahlung (photons) when interacting with a nucleus. Photons split into electron-positron pairs which radiate photons and so on. This process results in a shower of electrons and photons, growing exponentially, and stopping when the photons reach energies of $2 \cdot m_e$ or less and do not have enough energy

for pair production. The remaining electrons and photons still deposit energy via bremsstrahlung, ionization and Compton scattering. The resulting radiation can be detected and amplified, e.g. using photo-multipliers.
Hadrons are heavier than electrons and take longer to deposit their energy via bremsstrahlung (neutral hadrons do not emit any bremsstrahlung at all, of course). They deposit most of their energy in the second part of the calorimeter, the hadronic calorimeter, where they form similar showers to the electrons. The shapes of the showers differ, though, because hadrons mainly interact strongly with the atomic nuclei. During those interactions, other hadrons such as pions, kaons or protons can be produced. Some of them (e.g. neutral pions) decay electromagnetically, causing small electromagnetic showers inside the hadronic showers. This makes hadronic calorimetry more difficult than electromagnetic calorimetry and is one cause of the uncertainties on jet energies.
Muons are heavier than electrons and do not interact strongly. Most muons produced at the LHC will have energies between 1 GeV and 1 TeV. They are so-called minimum ionizing particles, which means they deposit little energy in the calorimeters (up to a few GeV per muon). Muons are the only particles reaching and interacting with the outermost detector layers, the muon chambers, which are tracking chambers similar to those in the inner detector. If they are surrounded by a magnetic field, they can be used to measure muon momentum.
Neutrinos only interact weakly with matter. They leave no tracks in the inner detector and do not deposit any energy in the calorimeters. However, they can be detected indirectly: The magnitude of the vector sum of the transverse energies of all "visible" particles is defined as the *missing transverse energy* $\not{E}_T$. In hadron collisions, the momentum of the interacting partons in the initial state is unknown. However, it can be assumed that their transverse momentum is negligibly small. If all particles in the final state are detected, the transverse energy should be balanced and the missing energy is zero (or close to zero, accounting for detector accuracy). If the missing energy is not zero, one knows that there was at least one undetected particle, e.g. a neutrino, that caused this imbalance. As the longitudinal momentum of the partons in the initial state is not known and the longitudinal momentum of the beam remnants in the final state cannot be measured, the longitudinal missing energy cannot be determined. Because the whole system may be boosted along the z-axis, projections to the transverse plane are often used as they are invariant under such boosts.

3.2. Important Physical Quantities

3.2.1. Luminosity, Counting Rate and Cross Section

The luminosity L describes the flux density of particles in the beam. For synchrotrons it is given by

$$L = \frac{n_B \cdot f \cdot n_1 \cdot n_2}{A}, \tag{3.1}$$

where f is synchrotron frequency, n_B is the number of bunches per beam, n_i is the number of particles per bunch in beam i and A is the effective bunch cross section at the interaction point [2]. However, determining the actual luminosity is not a trivial task, especially for hadron colliders. At lepton colliders, the luminosity can be measured using eq. 3.2 and a process with a well known cross section (BHABHA scattering). At the LHC, the luminosity will be determined from the total elastic and inelastic cross section using the optical theorem [15]. The integrated luminosity $L_{int} = \int L \mathrm{d}t$ is defined as the integral of the luminosity over a certain amount of time.

For a reaction with a given final state, the exclusive cross section $\sigma(\sqrt{s})$ gives the probability of that reaction happening. It can be calculated from Feynman rules and the available phase space. However, cross sections for QCD events like top quark pair production are difficult to calculate, leading to large uncertainties on the predicted cross sections.

Luminosity, cross section, and interaction rate $\frac{\mathrm{d}P}{\mathrm{d}t}$ (for a given process) are related by the formula

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \sigma \cdot L. \tag{3.2}$$

The counting rate $\frac{\mathrm{d}N}{\mathrm{d}t}$ is again related to the interaction rate by a factor ϵ_{eff}, which depends on the acceptance of the detector as well as the efficiencies of the trigger, the reconstruction algorithms and the cuts used (see section 5.1).

3.2.2. Detector Coordinates

Because of the layout of the ATLAS detector, cylindrical coordinates are used to describe the positions and momenta of particles inside it [16]. The z-axis follows the beam direction, so the $x - y$ plane (transverse plane) is perpendicular to the beam. The origin is located at the center of the detector at the planned interaction point. The x-axis points to the center of the LHC ring, the y-axis points upward.

The azimuthal angle φ is measured in the transverse plane with respect to the x-axis, the polar angle θ is measured with respect to the z-axis. Instead of the polar angle, the pseudorapidity $\eta := -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$ is used. In the relativistic limit (i.e. for $\frac{m}{E} \ll 1$), this quantity is equal to the rapidity $y = \ln\left(\frac{E+p_z}{E-p_z}\right)$, which is invariant not under Lorentz boosts along the z-axis, but recieves an additive constant. The distance $\Delta\eta$ between two objects is Lorentz-invariant (for boosts along the z-axis). The angular distance ΔR between two objects is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$.

3.3. Experimental Setup

3.3.1. The LHC

The LHC (Large Hadron Collider) is a proton-proton collider that is expected to start operating in winter 2009. It is expected to start operating at a center of mass energy of 10 TeV and will eventually reach an energy of 14 TeV at a design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [17]. The beam tunnel has a circumference of 27 km and was previously used for the Large Electron-Positron Collider (LEP). The four main experiments are ATLAS, CMS, LHCb and ALICE. In addition to proton-proton collisions, lead ion collisions will also be studied.
As it will be the particle collider with the highest center of mass energy, the LHC will be focused on searching for new physics, for example the predicted Higgs boson or hypothetical supersymmetric partners to known matter. There will also be precision measurements of the top quark properties, e.g. its production cross section. In addition to proton-proton collisions, lead ion collisions will be used to study the properties of quark-gluon plasma.

3.3.2. The ATLAS Detector

The ATLAS (A Toroidal LHC Apparatus) detector (cf. fig. 3.1) is one of the two multipurpose detectors at LHC [16]. It is composed of an inner detector inside a 2 T solenoidal magnetic field, electromagnetic and hadronic calorimeters and muon chambers.
The inner detector consists of the Pixel detector, the Silicon Strip Detector (SCT) and the Transition Radiation Tracker (TRT). In the Pixel Detector, the approximately 80 million pixels are arranged in three barrel layers and six endcap disks. The closest barrel layer is only about 50 mm from the beam pipe. Following the Pixel

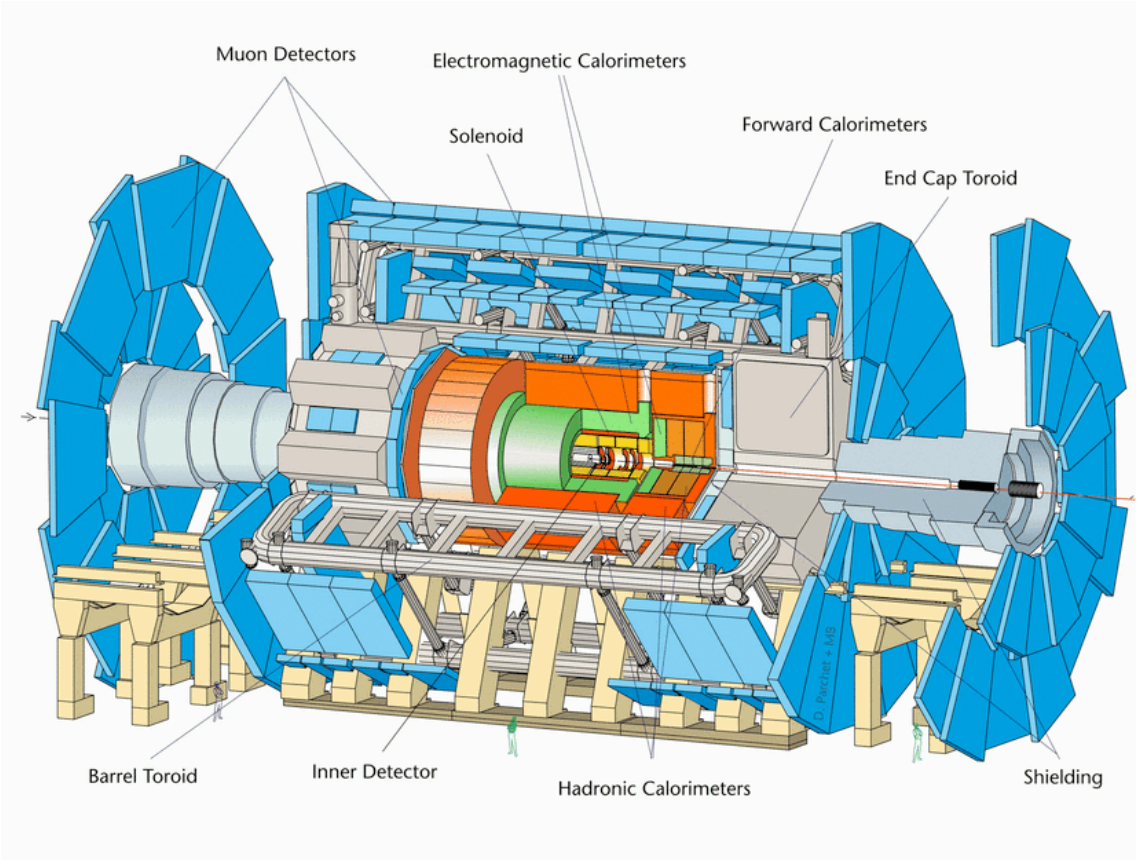

Figure 3.1.: The ATLAS experiment.

Detector, the Silicon Strip Detectors are arranged so that each track crosses eight strip layers. The TRT, made up of about 350,000 straw tubes, measures radiation that is produced when a charged particle crosses between two regions with different permittivities. This allows for the determination of the particle's velocity. It can therefore be used to distinguish between e.g. electrons and charged pions.
The calorimeters can again be divided into barrel parts and endcap parts. The liquid argon electromagnetic calorimeter is followed by the hadronic calorimeters. In the endcap regions, the latter are filled with liquid argon, while in the barrel region the hadronic calorimeter is made from scintillating fiber tiles. There are also two liquid argon forward calorimeters. There are approximately 200,000 readout channels from all of the calorimeters.
The muon system is immersed in a toroidal magnetic field to be able to reconstruct the muon momentum. There are four different kinds of muon chambers: Monitored Drift Tubes (MDTs) and Resistive Plate Chambers (RPCs) in the barrel region, as well as Cathode Strip Chambers (CSCs) and Thin Gap Chambers (TGCs) in the

endcap regions. The RPCs and TGCs have fast readout times and are mainly used for triggering (see below) while the MDTs and CSCs are used for precision tracking. The muon chambers are arranged so that each muon passes through several chambers, allowing for the determination of the muon's momentum and sign.

Reading out the full ATLAS detector at each bunch crossing would lead to data rates that are too high to be able to save or analyse all of the data. Apart from that, most events are expected to be "simple" QCD scattering events. Interesting physics makes up only a fraction of the observed reactions. Because of these two facts, a trigger system is needed. ATLAS employs a three level trigger system (L1, L2, and Event Filter). Its task is to reduce the recorded data rate to a manageable amount, preventing buffer overflows, and to filter out uninteresting events.

The L1 triggers look for high-p_T objects in certain regions of the detector, using muon chambers and calorimeter information. From their decisions, so-called *Regions of Interest* are identified which the L2 trigger systems investigate further. The event filters have access to information from all parts of the detectors. Their output rate is about 200 Hz, which is low enough to be written to tape for further offline analysis.

4. Monte Carlo Samples

For the following analyses, ROOT files containing simulated Monte Carlo data were used. The $t\bar{t}$ samples were obtained using the Monte Carlo generator MC@NLO [18, 19]. This is a next-to-leading order Monte Carlo event generator. In order for some of the second order corrections to converge, negative weights were introduced. After "unweighting", events can now have weights of ± 1. Events with negative weights have to be subtracted whenever one is counting events, for example when filling a histogram or counting the number of events that passed a certain cut. The W+jets and $b\bar{b}$ background samples were produced using ALPGEN [20]. The showering and hadronisation was provided by JIMMY [21] in all cases.
The reconstructed data was obtained using GEANT4 [22] to simulate the ATLAS detector and digitize the data. It was then reconstructed using the ATHENA framework [23].
The following samples (cf. tab. 4.1) were used:

1. $\boldsymbol{t\bar{t} \rightarrow e}$**+jets**: $t\bar{t}$ events with one hadronically and one leptonically decaying W boson. The latter must decay into an electron or into a τ lepton decaying into an electron.

2. $\boldsymbol{t\bar{t} \rightarrow \mu}$**+jets**: $t\bar{t}$ events with one hadronically and one leptonically decaying W boson. The latter must decay into a muon or into a τ lepton decaying into a muon.

3. $\boldsymbol{t\bar{t}}$ **dilepton** channel: $t\bar{t}$ events with two leptonically decaying W bosons where at least one of the W bosons must decay into an electron or into a τ lepton decaying into an electron.

4. $\boldsymbol{W \rightarrow e\nu}$ with associated jets (for example from initial/final state radiation and/or processes like $gq \rightarrow Wq$). In the available samples, the production of 3, 4, or 5 additional partons had been specified. ALPGEN uses MLM matching to avoid double counting of events in certain regions of the phase space. The jet

multiplicity does not always equal the number of partons because of showering effects and reconstruction efficiencies. Samples with two or fewer partons were not used because the majority of them will not pass preselection cuts. Samples with more than five partons were not available. Their absence is not expected to influence the results as the cross section decreases rapidly with increasing parton multiplicity.

5. $\boldsymbol{b\bar{b}}$**+jets**: Events where bottom quark pairs were produced along with light quark or gluon jets. At least one bottom quark must decay weakly into a muon.

In addition to the requirements on truth (generator) level, only events passing the following preselection criteria on reconstruction level were written to files:

- At least one (reconstructed) electron with $p_T > 15\,\text{GeV}$, $\eta < 2.47$ and outside the crack region (only e+jets, dilepton, $W \to e\nu$).
- At least one (reconstructed) muon with $p_T > 15\,\text{GeV}$, $\eta < 2.5$ (only μ+jets, $b\bar{b}$+jets).
- At least one jet[1] with $p_T > 15\,\text{GeV}$[2], $\eta < 2.5$ after overlap removal between jets and isolated electrons.

The "crack region" ($1.37 \leq \eta \leq 1.52$) is the border between the barrel and forward calorimeters. Electrons in this region are excluded as they cannot be reliably reconstructed.
The overlap removal is necessary because the jet algorithm reconstructs most electrons as jets. During this procedure, jet candidates with $p_T > 20\,\text{GeV}$ that have an angular distance $\Delta R < 0.2$ from an isolated[3] (`Etcone20` $< 6\,\text{GeV}$) electron with $p_T > 20\,\text{GeV}$ are removed.
These preselection cuts are very loose and serve to make the sample files smaller and easier to handle without losing interesting events. To compare contributions from signal and background events, all histograms were scaled to $L_{int} = 200\,\text{pb}^{-1}$. The necessary scaling factor is the ratio between the number of expected events $N_{exp} = \sigma \cdot L_{int}$ and the number of available Monte Carlo events N_{MC}.

[1] Jets were reconstructed using "Cone4H1TowerJets", which is a cone algorithm with a cone size (radius in $\eta - \phi$ space) of 0.4.
[2] The quantity called p_T is determined from calorimeter measurements, not from tracker information. The name p_T is used to stay consistent with ATLAS nomenclature.
[3] For a definition of isolation, especially the `Etcone20` variable, see section 6.8.1.

channel	σ [pb]	BR	N_{exp}	N_{MC}	scale factor
total $t\bar{t}$	401.6	1	80320		
$t\bar{t} \rightarrow e$+jets	70.1	0.1746	14023	401547	0.0349
$t\bar{t} \rightarrow \mu$+jets	69.8	0.1739	13965	403309	0.0346
$t\bar{t}$ dilepton channel	28.2	0.0701	5634	161128	0.0350
$W \rightarrow e\nu + 3$ partons	248.0		49605	224074	0.221
$W \rightarrow e\nu + 4$ partons	68.4		13688	58872	0.233
$W \rightarrow e\nu + 5$ partons	20.3		4050	17492	0.232
$b\bar{b} \rightarrow \mu$+jets	25855		5171000	249802	20.7

Table 4.1.: Some properties of signal and background samples used. Expected cross sections taken from [14]. N_{exp} refers to the expected number of events in $200\,\mathrm{pb}^{-1}$ of data, N_{MC} to the number of Monte Carlo events.

5. Methods

5.1. The Cut and Count Method for Cross Section Determination

If one could directly measure the production rate of top quark pairs inside the detector, the $t\bar{t}$ production cross section could be calculated by inverting eq. 3.2:

$$\sigma_{t\bar{t}} = \frac{\mathrm{d}S}{L \cdot \mathrm{d}t} = \frac{S}{L_{int}}, \tag{5.1}$$

where S is the number of $t\bar{t}$ events and L_{int} is the integrated luminosity. However, events cannot be counted this easily. Detector, trigger, and reconstruction are not perfect, so single objects or even whole events might be lost. There are also lots of other kinds of events that take place, some with similar signatures to the ones we are looking for, that have to be accounted for.
The cut and count method is an easy, but not very accurate way of finding the cross section of a given reaction. It works best for reactions with low, well-understood backgrounds. The strategy is the following:

1. Apply a series of cuts to your sample, i.e. only keep events that pass certain requirements
2. Count N, the number of events that pass all cuts
3. Find the efficiency ε of these cuts for signal events
4. Calculate the expected number of background events B that pass all cuts
5. The total cross section is then given by

$$\sigma = \frac{N - B}{\varepsilon \cdot L_{int} \cdot BR} \tag{5.2}$$

where BR is the branching rate of the decay channel(s) used for the analysis.

There is no straight-forward procedure for finding the best set of cuts, neither is there finding the efficiencies and estimating the remaining background. The optimal cuts depend on the expected amount of signal and background as well as the uncertainties on the amount of background: If the background is large and the cross section is not known well, it is important to suppress most of the background, even at the loss of statistics. If the background is well understood and described, and the amount of signal is very small, it is important to choose cuts that are efficient, even if they let through a lot of background. In principle, it is not possible to get both.
The cut efficiencies both for signal events and the background can be calculated using Monte Carlo simulations. These have large systematic uncertainties, especially for the LHC as no one has ever taken data at $\sqrt{s} = 10\,\text{TeV}$ before. There are ways to estimate cut efficiencies and remaining background from data. In this thesis, Monte Carlo generated samples were used to test various sets of cuts and estimate their efficiencies.

5.1.1. Cut Efficiencies and their Uncertainties

The cut efficiencies calculated using Monte Carlo generated samples will have both systematic and statistic uncertainties. The systematic uncertainties are due to the fact that the response of the detector is not known very well yet and the Monte Carlo predictions (e.g. for kinematic distributions) could not be tuned to data at $\sqrt{s} = 10\,\text{TeV}$ or to the specific properties of the ATLAS detector. They were not considered in this thesis.
The statistical uncertainties arise because the samples used are of finite size. They can be calculated precisely. For these studies, a Bayesian approach [24] was used. It assumes that the prior distribution is flat in the interval $[0, 1]$ and calculates the probability density function for the efficiency, given the observed number of events that passed/failed the cut. It is now possible to find intervals so that the probability for the true efficiency to be inside that interval is at least 68.3%, corresponding to the 1σ interval of a Gaussian distribution. There are of course infinitely many such intervals, but usually one of the following is used:

1. The symmetric interval (symmetric around the mode),

2. the central interval, where the probabilities for the efficiency to be below the lower limit and to be above the upper limit are equal,

3. the shortest interval covering a probability of 68.3%.

In the case of a symmetric distribution (e.g. a Gaussian distribution) all three intervals are equal. That is generally not the case for an asymmetric distribution like the one that is used here. In this thesis, the shortest interval was used. After rounding the uncertainties to one significant figure, the intervals will appear symmetric in most cases.

5.2. The Matrix Method for QCD Background Determination

The matrix method [25, 26] can be used to estimate the amount of signal and (instrumental) background events in a given sample using only information gathered from data.
We define a loose and a tight sample, where the latter is a subset of the former, obtained by applying an additional cut or set of cuts (for example a tighter lepton isolation requirement). We assume both samples contain a certain amount of signal events N_s as well as some amount of background events N_b. We can then write:

$$N^l = N_s^l + N_b^l, \tag{5.3}$$

$$N^t = N_s^t + N_b^t = \varepsilon_s \cdot N_s^l + \varepsilon_b \cdot N_b^l, \tag{5.4}$$

Where N^l and N^t are the amount of events in the loose and the tight sample, ε_s and ε_b are the efficiencies of the additional cut applied to a pure signal or pure background sample. N^l and N^t can be easily measured in a counting experiment. If ε_s and ε_b are known and differ significantly from each other, one can estimate the amount of background/signal remaining in the sample:

$$N_b^t = \frac{\varepsilon_b}{\varepsilon_s - \varepsilon_b}\left(\varepsilon_s N^l - N^t\right), \qquad N_s^t = \frac{\varepsilon_s}{\varepsilon_b - \varepsilon_s}\left(\varepsilon_b N^l - N^t\right). \tag{5.5}$$

The two efficiencies should be determined from data if possible.
Later on, it will be shown that the matrix method can be used to estimate the rate of fake isolated muons from QCD background. This will be done in the $t\bar{t} \to \mu$+jets channel; the background will be simulated $b\bar{b}$ events with muons from B-meson decays. The additional cut will be a tighter muon isolation requirement. In a complete study, one would not need to rely on Monte Carlo simulations for the

efficiencies. ε_b can be determined from data in regions where there is little to no signal, in our case for example from events with low $\not\!\!E_T$. To determine ε_s from data, $Z \to \mu\mu$ events can be used. They contain isolated muons with a high p_T (similar to the muons from top quark decays) and it is possible to obtain relatively pure $Z \to \mu\mu$ samples from data. As ε_s depends on η or p_T, it needs to be corrected for the different kinematics of the muons from $t\bar{t}$ events.

5.2.1. Uncertainties

For estimating the uncertainties on N_s^t and N_b^t, the approach described in [27] was followed. There are four sources of uncertainties: both ε_s and ε_b have statistical and systematical uncertainties that need to be determined separately. We assume that these are neither correlated to each other nor to N^l or N^t as they are determined from different data samples. N^l and N^t also have statistical uncertainties as they are determined by a counting experiment, but they are correlated. Instead, we can use $N_1 := N^l - N^t$ and $N_2 := N_t$, which are uncorrelated. For sufficiently high statistics we can assume Gaussian distributions with $\Delta N_i = \sqrt{N_i}$. For the sake of simplicity, we also assume Gaussian errors on ε_s and ε_b. The uncertainty on N_s^t is then given by:

$$\Delta N_s^t = \sqrt{\left(\frac{\partial N_s^t}{\partial \varepsilon_s} \cdot \Delta\varepsilon_s\right)^2 + \left(\frac{\partial N_s^t}{\partial \varepsilon_b} \cdot \Delta\varepsilon_b\right)^2 + \left(\frac{\partial N_s^t}{\partial N_1} \cdot \Delta N_1\right)^2 + \left(\frac{\partial N_s^t}{\partial N_2} \cdot \Delta N_2\right)^2}. \tag{5.6}$$

ΔN_b^t can be calculated the same way.

6. Results

6.1. Properties of $t\bar{t}$ Events and Dominant Backgrounds

It makes sense to look at some kinematic distributions of the samples before applying any cuts. As cuts on electron p_T, $\not{E}_T$, and jet p_T will be introduced later, these distributions will be examined here in detail. All histograms have been scaled to $200\,\mathrm{pb}^{-1}$.

In fig. 6.1, you can find the p_T spectrum of the reconstructed electrons. Electrons with a transverse momentum of less than 15 GeV are not saved in the ROOT file. The distributions look similar for signal and background events. As events with no electrons were removed during the preselection before generating the ROOT file, further cuts on electron p_T are not expected to suppress a lot of background events. They are necessary because of the electron triggers that will be used, though. Triggers were not considered for this study, but they will be necessary during data taking. It is likely that a trigger requiring an electron with $p_T > 15\,\mathrm{GeV}$ will be used for studies in the e+jets channel. As there are large uncertainties on the trigger efficiency for electron momenta close to the trigger threshold, those events need to be removed.

Figure 6.2 shows the distribution of missing energy $\not{E}_T$ for both signal and background events. Again, the spectra are quite similar. That can be expected as the W+jets events contain neutrinos in the final state, leading to real missing energy. However, cuts on $\not{E}_T$ will hopefully help suppress QCD background with fake isolated leptons. For these processes, there is no real $\not{E}_T$. There can be fake missing energy due to, for example, objects not being reconstructed or malfunctions in the calorimeter. As can be seen in fig. 6.2, cutting on $\not{E}_T$ will not reduce statistics too much.

In fig. 6.3, the jet multiplicity for signal and background events is shown. Only jets with $p_T > 20\,\mathrm{GeV}$ are counted. Although four quarks are produced during the semi-leptonic decay of a $t\bar{t}$ pair, there are many signal events with more or fewer than

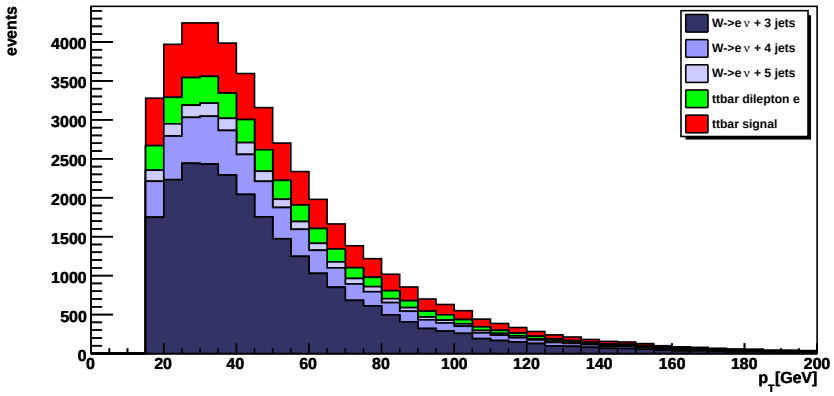

Figure 6.1.: Transverse momentum p_T of reconstructed electrons for signal and background events.

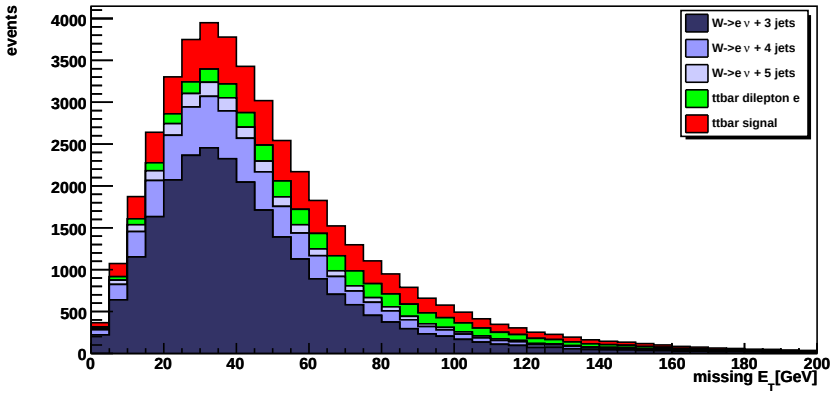

Figure 6.2.: Missing transverse energy $\not{E}_T$ for signal and background events.

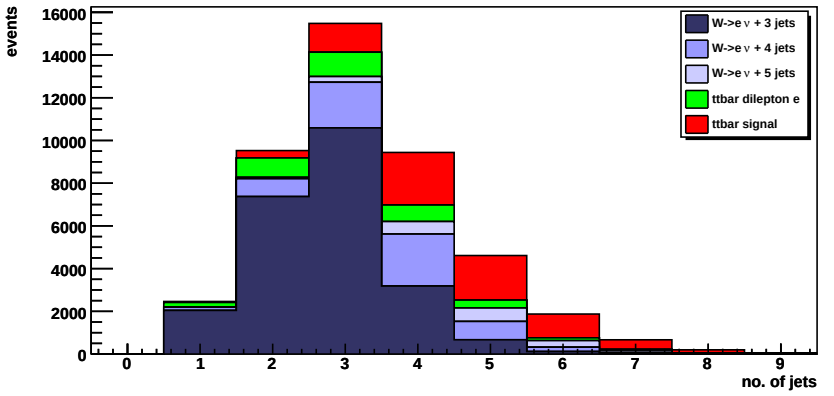

Figure 6.3.: Number of jets with $p_T > 20\,\mathrm{GeV}$ for signal and background events.

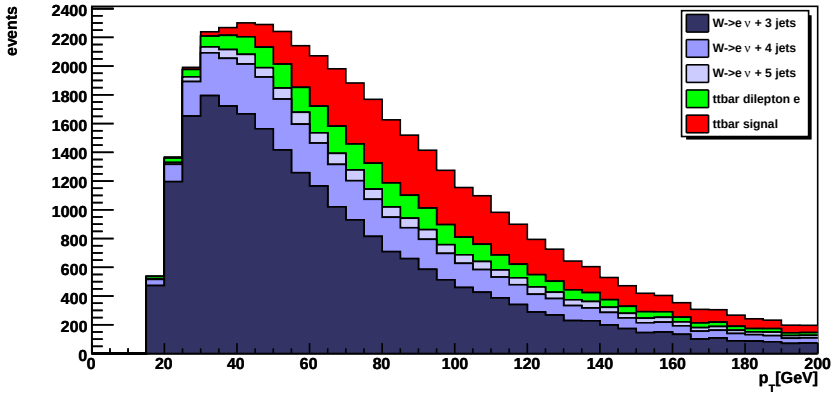

Figure 6.4.: Transverse momentum p_T of leading jets for signal and background events.

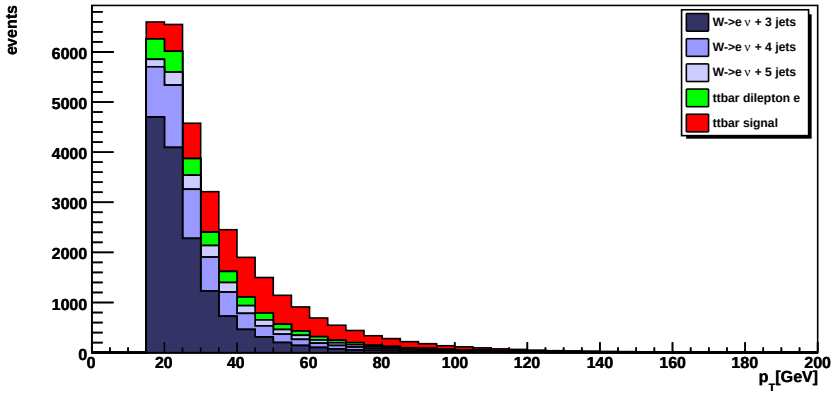

Figure 6.5.: Transverse momentum p_T of 3rd leading jets for signal and background events.

four jets. This is not unexpected: Some of the jets produced during the decay might be outside of the detector's acceptance region or have a p_T below the reconstruction threshold. There can also be additional jets, e.g. due to final state radiation.
For three or fewer jets, background events dominate, but in bins with four or more jets, there is a large contribution from signal events. Cuts on jet multiplicity are therefore helpful in discriminating against W+jets background.
In fig. 6.4 and 6.5, the transverse momentum spectra for leading jets (i.e. the jets with the highest p_T) and third leading jets are shown. For the former, the distribution peaks at around $30 - 40\,\text{GeV}$ for background events, but at about $80\,\text{GeV}$ for signal events. A similar behaviour can be seen in the p_T spectra of the third leading jets. It peaks at around $15\,\text{GeV}$ for background events but at about $40\,\text{GeV}$ for signal events. Because of that, jet p_T seems like a useful quantity for background

discrimination.

6.2. Preselection Cuts

The ATLAS collaboration conducted a series of analyses on Monte Carlo data to study the expected performance of the ATLAS detector. The results have been published in what is often referred to as the "ATLAS CSC[1] book" [1]. The CSC chapter on top quark physics uses four preselection cuts for the electron+jets channel (see below). They will be referred to as "CSC cuts" in the following sections.

1. 4-jet requirement: There need to be at least four jets with $p_T > 20\,\text{GeV}$.

2. 3-jet requirement: Of those jets, at least three need to have $p_T > 40\,\text{GeV}$.

3. Electron requirement: There needs to be exactly one electron, which needs to have $p_T > 20\,\text{GeV}$, and no isolated (`Etcone20`$< 6\,\text{GeV}$) muon.

4. Missing E_T requirement: The total missing transverse energy has to be $\not{E}_T > 20\,\text{GeV}$.

The cut flows and efficiencies for the preselection cuts applied to the $t\bar{t}$ sample and to the W+jets samples can be found in tables 6.1–6.5.

	applying cuts in order			applying cuts separately	
requirement	events	events/total	events/last cut	events	events/preselection
total events	401547	100.00		401547	
preselection	230310	57.36 ± 0.08	57.36 ± 0.08	230310	100.00
4 jet requirement	152381	37.95 ± 0.08	66.16 ± 0.10	152381	66.16 ± 0.10
3 jet requirement	105842	26.36 ± 0.07	69.46 ± 0.12	126791	55.05 ± 0.10
electron p_T	95612	23.81 ± 0.07	90.33 ± 0.09	209940	91.16 ± 0.06
missing E_T	84881	21.14 ± 0.06	88.78 ± 0.10	206291	89.57 ± 0.06
W mass c.	43457	10.82 ± 0.05	51.20 ± 0.17	102440	44.48 ± 0.10
top mass c.	13388	3.33 ± 0.03	30.81 ± 0.22	41020	17.81 ± 0.08

Table 6.1.: Cut efficiencies [%] for $t\bar{t} \to e + j$ events.

[1] "Computing System Commissioning"

	applying cuts in order			applying cuts separately	
requirement	events	events/total	events/last cut	events	events/preselection
total events	161128	100.00		161128	
preselection	102612	63.68 ± 0.12	63.7 ± 0.1	161128	100.00
4 jet requirement	26154	16.23 ± 0.09	25.5 ± 0.1	102612	25.49 ± 0.14
3 jet requirement	14831	9.20 ± 0.07	56.7 ± 0.3	26154	22.99 ± 0.13
electron p_T	8905	5.53 ± 0.06	60.0 ± 0.4	14831	50.55 ± 0.16
missing E_T	8419	5.23 ± 0.06	94.5 ± 0.2	8905	93.80 ± 0.08
W mass c.	2877	1.79 ± 0.03	34.2 ± 0.5	8419	20.03 ± 0.13
top mass c.	498	0.31 ± 0.01	17.3 ± 0.7	2877	7.03 ± 0.08

Table 6.2.: Cut efficiencies [%] for dilepton events.

	applying cuts in order			applying cuts separately	
requirement	events	events/total	events/last cut	events	events/preselection
total events	224074	100.00		224074	
preselection	108476	48.41 ± 0.11	48.4 ± 0.1	108476	100.00
4 jet requirement	7170	3.20 ± 0.04	6.6 ± 0.1	7170	6.61 ± 0.08
3 jet requirement	1931	0.86 ± 0.02	26.9 ± 0.5	7107	6.55 ± 0.08
electron p_T	1589	0.71 ± 0.02	82.3 ± 0.9	100394	92.55 ± 0.08
missing E_T	1307	0.58 ± 0.02	82.3 ± 0.9	92013	84.82 ± 0.11
W mass c.	430	0.19 ± 0.01	32.9 ± 1.3	16574	15.28 ± 0.11
top mass c.	61	0.027 ± 0.004	14.2 ± 1.7	5563	5.13 ± 0.07

Table 6.3.: Cut efficiencies [%] for $W \to e + \nu + 3$ partons.

	applying cuts in order			applying cuts separately	
requirement	events	events/total	events/last cut	events	events/preselection
total events	58872	100.00		58872	
preselection	28703	48.75 ± 0.21	48.8 ± 0.2	28703	100.0
4 jet requirement	9501	16.14 ± 0.15	33.1 ± 0.3	9501	33.1 ± 0.3
3 jet requirement	3911	6.64 ± 0.10	41.2 ± 0.5	5612	19.6 ± 0.2
electron p_T	3587	6.09 ± 0.10	91.7 ± 0.4	26661	92.9 ± 0.2
missing E_T	3111	5.28 ± 0.09	86.7 ± 0.6	24493	85.3 ± 0.2
W mass c.	1045	1.78 ± 0.05	33.6 ± 0.9	7195	25.1 ± 0.3
top mass c.	170	0.29 ± 0.02	16.3 ± 1.2	2411	8.4 ± 0.2

Table 6.4.: Cut efficiencies [%] for $W \to e + \nu + 4$ partons.

	applying cuts in order			applying cuts separately	
requirement	events	events/total	events/last cut	events	events/preselection
total events	17492	100.0		17492	
preselection	8733	49.9 ± 0.4	49.9 ± 0.4	8733	100.0
4 jet requirement	5782	33.1 ± 0.4	66.2 ± 0.5	5782	66.2 ± 0.5
3 jet requirement	3237	18.5 ± 0.3	56.0 ± 0.7	3608	41.3 ± 0.5
electron p_T	3019	17.3 ± 0.3	93.3 ± 0.4	8103	92.8 ± 0.3
missing E_T	2622	15.0 ± 0.3	86.9 ± 0.6	7572	86.7 ± 0.4
W mass c.	793	4.5 ± 0.2	30.2 ± 0.9	2570	29.4 ± 0.5
top mass c.	117	0.7 ± 0.1	14.8 ± 1.3	883	10.1 ± 0.3

Table 6.5.: Cut efficiencies [%] for $W \rightarrow e + \nu + 5$ partons.

6.3. W Boson and Top Quark Mass Constraints

The mass of the top quark and the W boson are well known quantities (compared to the energy resolution expected in the early stages of ATLAS data taking). As the hadronically decaying top quark can be fully reconstructed, it makes sense to require the mass of the reconstructed top quark and W boson to be close to the values measured in previous experiments. The width of this mass window is finite due to the finite width of top quark and W boson as well as inaccuracies in the measurements. If cuts like these are used for cross section determination, the result will depend on the mass of the top quark. Independent mass measurements are needed to extract the cross section.

The CSC note defines the top quark decay candidate as the three jets with the highest p_T sum, chosen from all jets with $p_T > 20\,\text{GeV}$. They then require at least one of the three possible dijet combinations that can be chosen from the top quark decay candidate to have an invariant mass close to the mass of the W boson $m_W \approx 80\,\text{GeV}$.

For the studies described here, the W boson decay candidate was defined as the dijet combination chosen from the top quark decay candidate with the invariant mass closest to m_W (cf. section 6.6). Cuts were applied to the invariant mass m of this W boson decay candidate as well as the invariant mass M of the top quark decay candidate. This approach is equivalent to the CSC approach.

The following mass constraints were used:

- loose W mass constraint: $60\,\text{GeV} < m < 100\,\text{GeV}$.
- tight W mass constraint: $70\,\text{GeV} < m < 90\,\text{GeV}$.

- top mass constraint: $157\,\text{GeV} < M < 187\,\text{GeV}$.

The efficiencies for the mass constraints can be found in table 6.1–6.5.
The invariant mass distribution for the W boson and top quark decay candidates after the preselection cuts can be seen in figs. 6.6 and 6.7. The mass of the W boson candidate peaks at about 80 GeV as it should. There are tails to both sides from combinatorics. The top quark candidate mass peaks at about $160 - 170\,\text{GeV}$. There are again large tails due to combinatorics. The tails can be reduced by applying a tight W mass constraint (cf. fig. 6.8). However, one observes a pronounced shoulder at about 100 GeV.

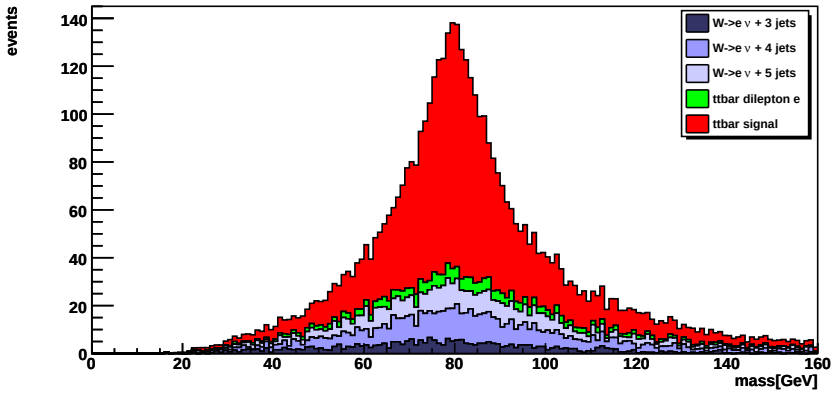

Figure 6.6.: Mass of the W boson decay candidate after CSC cuts.

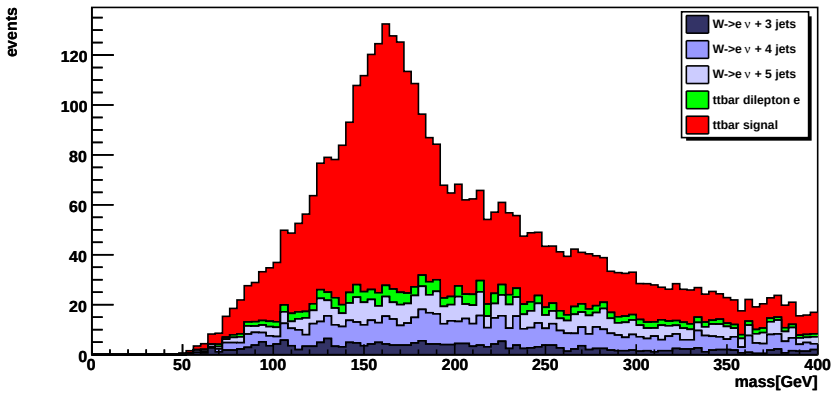

Figure 6.7.: Mass of the top quark decay candidate after CSC cuts.

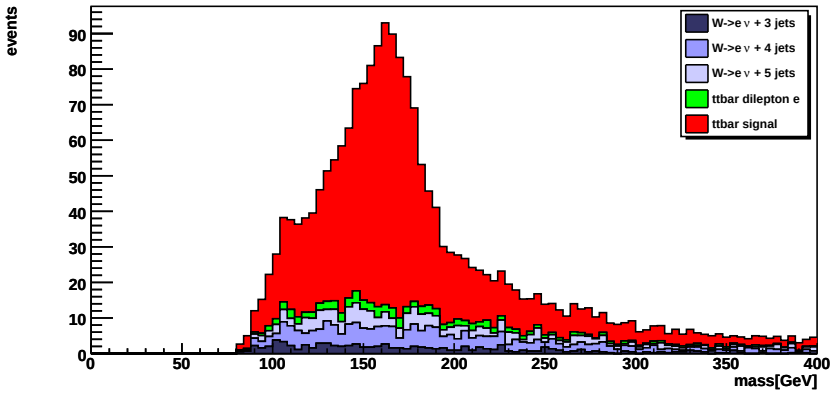

Figure 6.8.: Mass of the top quark decay candidate after CSC cuts and tight W mass constraints.

6.4. Jet-Parton Matching

To test the efficiencies of the top quark and W boson decay candidate algorithms, a matching algorithm was developed that matches truth quarks (from the Monte Carlo generator) to reconstructed jets in $\eta - \varphi$-space. Only jets with $p_T > 20\,\text{GeV}$ are considered for the matching. This matching is done in such a way that the outcome does not depend on the order of the objects to be matched:

1. Calculate the distances in $\eta - \varphi$-space between all possible quark-jet pairs, save them in a matrix.

2. Find the quark-jet pair with the minimum distance. Call this pair a "match" and record it.

3. Remove the corresponding row and column from the matrix so the jet and the quark that were just matched will not be considered for further matches.

4. Continue with step 2 until you run out of jets or quarks.

It is not always possible to find a jet coming from a given quark, as the jet might be outside the detector's acceptance region or it might not have enough energy to be reconstructed properly. There are also additional jets from e.g. gluon radiation that might be close to one of the quarks in $\eta - \varphi$-space. Because of that, it is necessary to introduce a cutoff and only use matches where the distance ΔR between quark and jet is smaller than some maximal distance ΔR_m. I chose $\Delta R_m = 0.5$ as the distribution peaks well below that value. Of course, there might still be some

wrongly matched pairs, but the algorithm performs correctly to first order. The dependence on ΔR_m was tested later on.
The distribution of the distances between matched pairs of quarks and jets can be seen in fig. 6.9.

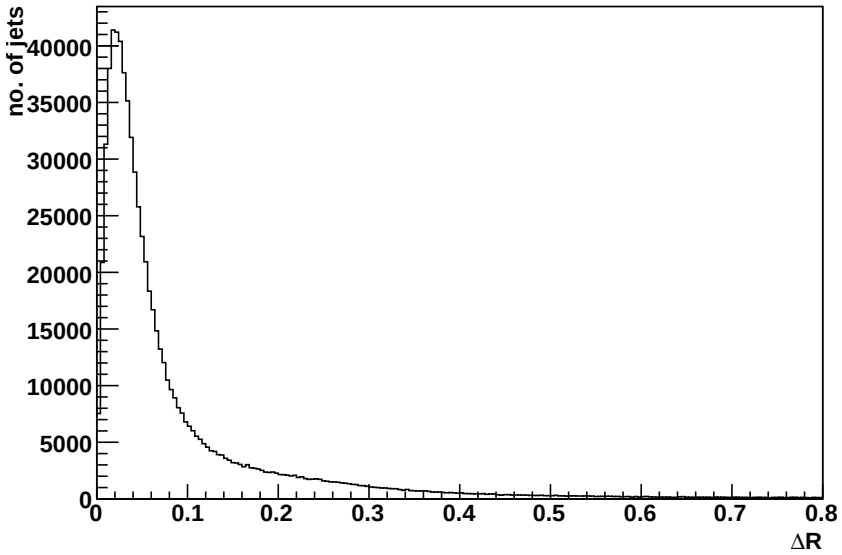

Figure 6.9.: Angular distance ΔR between jets and quarks.

6.5. Top Quark Candidate Algorithm

The performance of a simple algorithm that finds the decay products of the hadronically decaying top quark was tested and compared to the expected performance of randomly choosing three jets. The algorithm was taken from [1]. It takes the 3-jet combination from all jets with $p_T > 20\,\text{GeV}$ with the highest p_T sum, i.e. $i, k, j < N_{jet}$ so that $\sqrt{(p_{x,i} + p_{x,j} + p_{x,k})^2 + (p_{y,i} + p_{y,j} + p_{y,k})^2}$ is maximal and $i \neq j$, $i \neq k$, $j \neq k$ where i, j, k are the indices of the jets.
This algorithm was tested after applying the CSC preselection cuts, with a loose or tight W mass constraint and after applying a top mass constraint. Using information about truth quarks, the jets coming from the bottom quarks and the W boson decay products are identified using the matching algorithm described in 6.4. The number of successes s is defined as the number of jets in the top quark decay candidate matched to a truth quark from the hadronically decaying top quark with

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.050 ± 0.001	0.034 ± 0.001	0.029 ± 0.001
1	0.240 ± 0.001	0.188 ± 0.002	0.160 ± 0.002
2	0.500 ± 0.002	0.502 ± 0.002	0.486 ± 0.002
3	0.211 ± 0.001	0.276 ± 0.002	0.325 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.021 ± 0.001	0.016 ± 0.001	0.013 ± 0.001
1	0.122 ± 0.002	0.102 ± 0.002	0.081 ± 0.002
2	0.354 ± 0.003	0.339 ± 0.004	0.301 ± 0.004
3	0.503 ± 0.004	0.543 ± 0.004	0.606 ± 0.004

Table 6.6.: Performance of the top quark decay candidate algorithm.

$\Delta R \leq 0.5$. One can then count events with a given number of successes and calculate the relative frequencies for $s \in \{0, 1, 2, 3\}$. The relative frequency of getting $s = 3$ is a good measure for the performance of the algorithm. As can be seen in table 6.6, the CSC algorithm performs quite well. After applying tight mass constraints, the right 3-jet combination is found in about 61% of events.

To compare against random choice, the number of successes was calculated for the different jet multiplicities (cf. table A.1). If 3 out of N jets are chosen randomly, the probability of getting r jets coming from the top quark decay products is given by

$$p(r) = \frac{\binom{3}{r} \cdot \binom{N-3}{3-r}}{\binom{N}{3}}$$

for positive r with $N - 6 < r < 3$, and $p(r) = 0$ otherwise. The results are displayed in table A.1. The algorithm works better than random guessing even without any mass constraints for all jet multiplicities (results for 8 or more jets per event are not displayed due to lack of statistics).

One source of systematic uncertainties is the performance of the jet-parton matching, especially the parameter ΔR_m (maximal distance between matched quark-jet pairs). To see how the results depend on the value of ΔR_m, the analysis was also performed for $\Delta R_m = 0.4$ and $\Delta R_m = 0.6$ (cf. table A.2 and A.3). One can see that there is a slight dependence, which was to be expected, but the results are mostly stable against variations of ΔR_m. Still, this dependence leads to a systematic uncertainty. Understanding the performance of the top quark decay candidate algorithm leads to

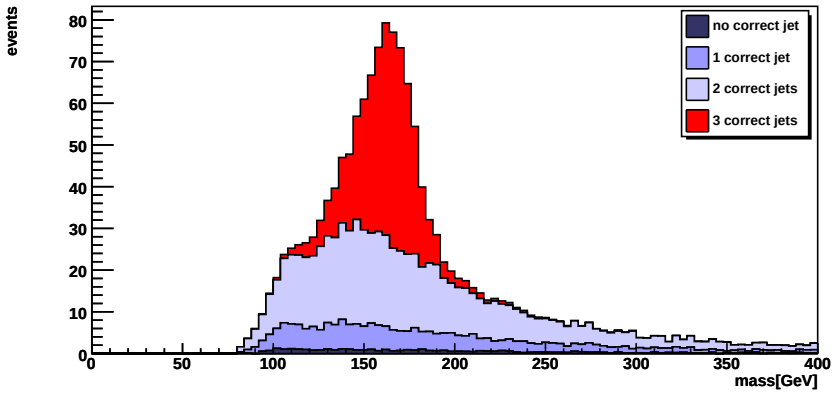

Figure 6.10.: Mass of top quark decay candidate for different numbers of correctly chosen jets.

a better understanding of the shape of the top quark candidate mass distribution. In figure 6.10 this distribution is plotted for the right combination of jets (in red) and for the combinatoric background (in blue). There is an almost-Gaussian peak on top of the flat background. Due to the previously applied W mass constraints, the background drops off for masses below 100 GeV, leading to the shoulder that can also be seen in figure 6.8.

6.6. W Boson Candidate Algorithm

The W boson decay candidate was simply chosen as the dijet combination from the top quark decay candidate with the invariant mass closest to $m_W = 80\,\text{GeV}$.

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.263 ± 0.002	0.217 ± 0.002	0.184 ± 0.002
1	0.472 ± 0.002	0.442 ± 0.002	0.418 ± 0.002
2	0.265 ± 0.002	0.342 ± 0.002	0.398 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.166 ± 0.003	0.142 ± 0.003	0.112 ± 0.003
1	0.361 ± 0.003	0.347 ± 0.004	0.321 ± 0.004
2	0.473 ± 0.004	0.511 ± 0.004	0.567 ± 0.004

Table 6.7.: Performance of the W boson decay candidate algorithm.

This is equivalent to the approach used by [1]. The performance of the algorithm was again evaluated using information from the matching between truth quarks and jets. It can be seen in table 6.7. The algorithm was compared to the expected results of randomly guessing the two jets. The results can be seen in table A.4. The performance was also tested for different values of the cut-off parameter ΔR_m for jet-parton matching (cf. tables A.5 and A.6). There is a slight dependence on ΔR_m that is a source of systematic uncertainty on the performance.

6.7. Re-evaluation of CSC preselection cuts

The preselection cuts in [1] were chosen and examined for a collision energy of $\sqrt{s} = 14\,\text{TeV}$. As the LHC will start at $\sqrt{s} = 10\,\text{TeV}$, it makes sense to reexamine those cuts as their efficiencies are related to the collision energy. For this part of the analysis, the four preselection cuts from [1] were used, but with varying cut parameters `m_4jet`, `m_3jet`, `m_e_pt`, and `m_MET`:

1. 4-jet requirement: There need to be at least four jets with $p_T >$ `m_4jet`.
2. 3-jet requirement: Of those jets, at least three need to have $p_T >$ `m_3jet`.
3. electron p_T requirement: There needs to be exactly one electron, which needs to have $p_T >$ `m_e_pt`, and no isolated muon.
4. Missing E_T requirement: The total missing transverse energy has to be $\not{E}_T >$ `m_MET`.

Jets with a transverse momentum of less than 20 GeV were not used as the reconstruction is not accurate for low-p_T jets. Moreover, it is likely that an electron with $p_T \geq 20\,\text{GeV}$ will be required for triggering, so the limits on the 4-jet requirement and on the electron p_T requirement should not go below 20 GeV. The two parameters for the jet p_T cuts were varied between 20 GeV and 80 GeV, `m_e_pt` between 20 GeV and 60 GeV, and `m_MET` between 0 GeV and 60 GeV.
There is no obvious way to define the "quality" of a cut or a set of cuts. In the context of a full analysis, one would like to choose the cuts so that the expected uncertainty (statistical and systematical) on the cross section (or whatever is being measured) is minimal. However, the expected systematical uncertainty on the cross section depends on the methods that are used to estimate/eliminate background. The optimal set of cuts therefore depends on the analysis and cannot be calculated

here.
In the context of this thesis, two numbers are worth looking at:

- The expected number of signal events S that pass all cuts, which can be calculated for a given integrated luminosity L_{int} using $S = L_{int} \cdot \sigma_{t\bar{t}} \cdot BR \cdot \varepsilon$, where ε is the efficiency of the set of cuts when used on a pure signal sample.

- The ratio $\frac{S}{B}$ of the expected signal over the expected background. The expected number of background events can be calculated using a similar formula: $B = \sum_b L_{int} \cdot \sigma_b \cdot BR_b \cdot \varepsilon_b$, where b runs over all relevant backgrounds and ε_b is the efficiency of the set of cuts used on a pure background sample of type b.

When the cross sections for the background processes are not well known, it is often more accurate to estimate the amount of background from the measured data, for example by looking in regions where one would expect no or little signal and extrapolating to the signal region. For the scope of this thesis, however, that cannot be done as there are no measured data available yet. The cross sections and branching ratios used can be found in table 4.1.
To get precise results, both S and $\frac{S}{B}$ should be large. Generally, one can find sets of cuts that maximize one or the other, not both at the same time. The exact figure of merit that needs to be maximized depends on the analysis (see above). However, if a set of cuts gives both a better signal and a better $\frac{S}{B}$ ratio, it can be considered "better" than the cut it was compared to.
For these studies, three scenarios were considered:

1. Applying the four preselection cuts (4-jet requirement, 3-jet requirement, electron p_T requirement and $\not\!\!E_T$ requirement) on their own,

2. Applying the four preselection cuts and a tight constraint on the W boson candidate mass ($70\,\text{GeV} < m < 90\,\text{GeV}$),

3. Applying the above cuts as well as a constraint on the mass of the top quark candidate ($157\,\text{GeV} < M < 187\,\text{GeV}$).

The possible values of signal events S and purity $\frac{S}{B}$ are displayed in figures 6.11 – 6.13. As expected, one can get good background suppression at the price of low signal and vice versa. The original CSC cuts (marked "-1") are somewhere in the middle, erring on the side of "more signal". Most cuts (especially those with tight jet requirements) can be found in regions with very low signal. Employing top mass

constraints (scenario 3), very high purity can be achieved, but in a region where only 100 or fewer signal events are expected. As this would lead to a relative statistic uncertainty on the $t\bar{t}$ production cross section of more than $\frac{1}{\sqrt{100}} = 10\%$ (for $200\,\mathrm{pb}^{-1}$ of data), using these cuts in the context of a full analysis of early data is not feasible.

The numbers in the graphs correspond to some specific examples of sets of cut parameters that can be found in table 6.8. There are some sets of cut parameters that appear better than the original CSC cuts. It seems that selections with loose or no cuts on missing transverse energy are favored. However, cuts on missing energy might be needed to filter out QCD background, which was not considered for this part of the study. With this in mind, it seems that the CSC preselection cuts still perform quite well at $\sqrt{s} = 10\,\mathrm{TeV}$. Another promising set of cuts is the one marked "5". It requires four jets with a p_T of at least 25 GeV and three jets with a p_T of at least 35 GeV. Electron p_T and $\not{E}_T$ requirements are left at 20 GeV. This set of cuts performs slightly better than the original CSC preselection cuts in scenarios 1 and 2.

It can be seen here that a simple cut-and-count based method, using the four preselection cuts and some mass constraints, will not be enough to determine $\sigma_{t\bar{t}}$ with acceptable precision during the first $200\,\mathrm{pb}^{-1}$ of data. In scenario 3 (using both W and top mass constraints), the lower bound on the expected relative statistic uncertainty $\frac{\Delta\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}}$ is given by $\min(\frac{1}{\sqrt{S}}) \approx \frac{1}{\sqrt{700}} \approx 4\%$.

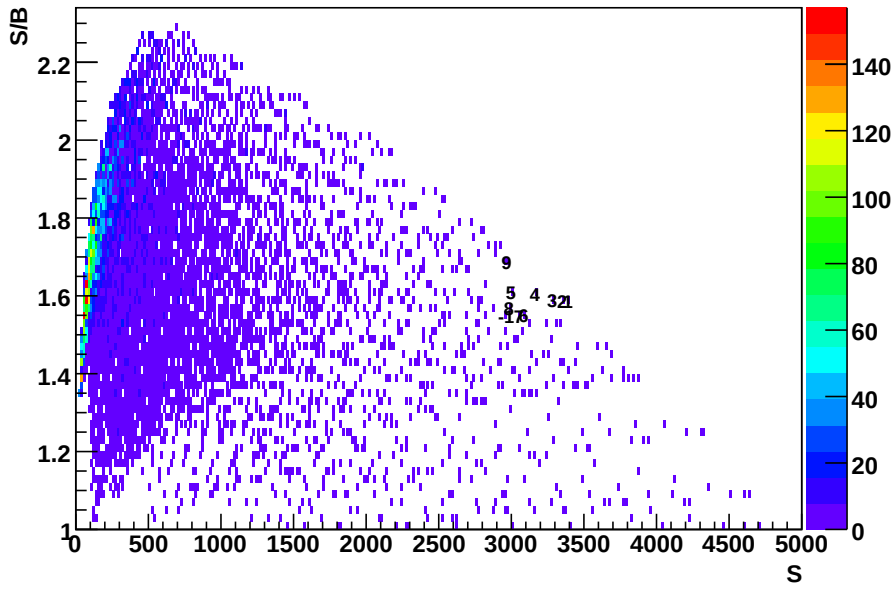

Figure 6.11.: Possible combinations of S and $\frac{S}{B}$ after CSC-like preselection cuts. The colours correspond to the number of sets of cut parameters that lead to a given combination of S and $\frac{S}{B}$. The numbers refer to specific sets of cut parameters (cf. tab. 6.8).

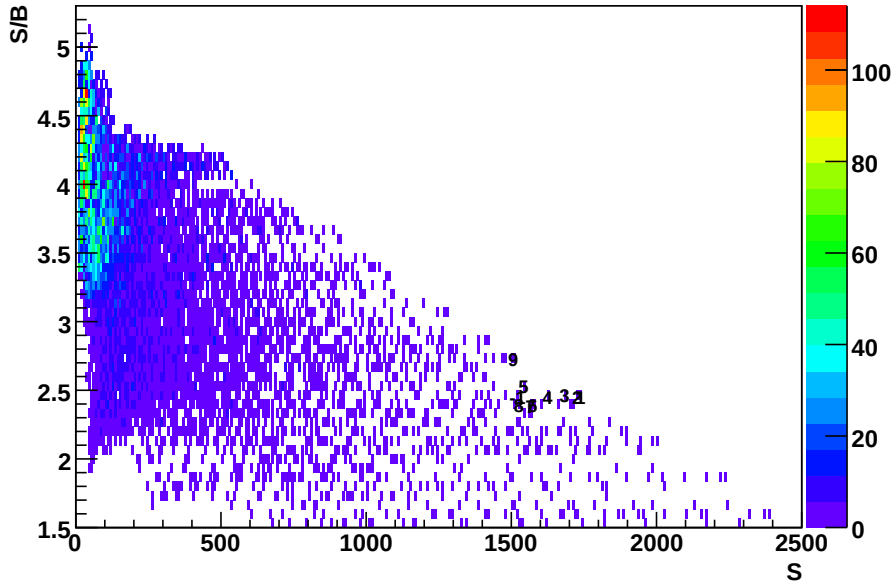

Figure 6.12.: Possible combinations of S and $\frac{S}{B}$ after CSC-like preselection cuts and W mass constraint. The colours correspond to the number of sets of cut parameters that lead to a given combination of S and $\frac{S}{B}$. The numbers refer to specific sets of cut parameters (cf. tab. 6.8).

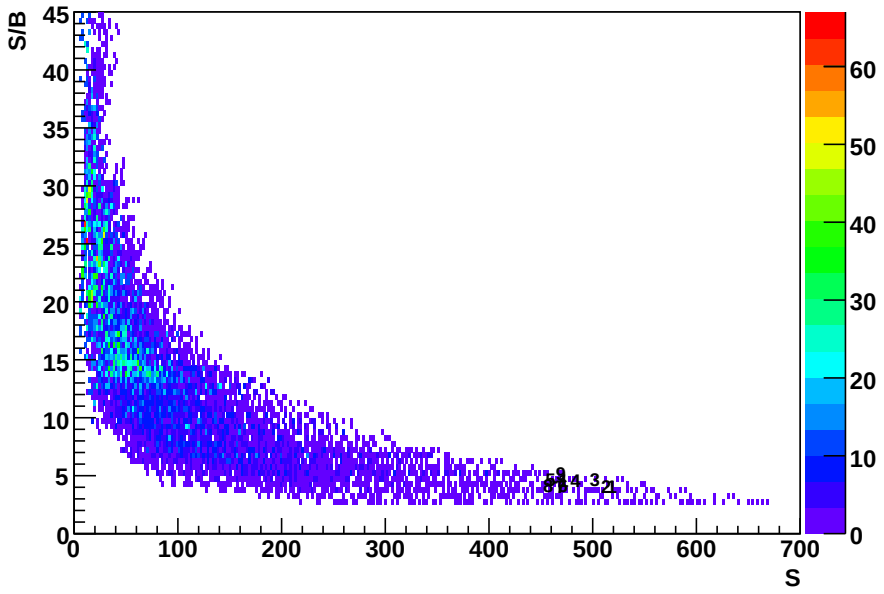

Figure 6.13.: Possible combinations of S and $\frac{S}{B}$ after CSC-like preselection cuts, top and W mass constraints. The colours correspond to the number of sets of cut parameters that lead to a given combination of S and $\frac{S}{B}$. The numbers refer to specific sets of cut parameters (cf. tab. 6.8).

	m_4jet [GeV]	m_3jet [GeV]	m_e_pt [GeV]	m_MET [GeV]	purity $\frac{S}{B}$		
					scenario 1	scenario 2	scenario 3
CSC	20	40	20	20	$\frac{2964.3}{1914.2} = 1.549$	$\frac{1517.6}{622.4} = 2.438$	$\frac{467.55}{97.5} = 4.794$
1	25	35	20	0	$\frac{3387.8}{2139.9} = 1.583$	$\frac{1738}{708.27} = 2.454$	$\frac{518.88}{115.7} = 4.486$
2	25	35	20	5	$\frac{3364.2}{2121.3} = 1.586$	$\frac{1725.5}{703.99} = 2.451$	$\frac{514.94}{115.1} = 4.475$
3	25	35	20	10	$\frac{3292.3}{2064.1} = 1.595$	$\frac{1687.8}{680.88} = 2.479$	$\frac{503.52}{110.0} = 4.579$
4	25	35	20	15	$\frac{3172.4}{1981.7} = 1.601$	$\frac{1626.6}{651.38} = 2.497$	$\frac{485.15}{105.5} = 4.598$
5	25	35	20	20	$\frac{3011}{1867.4} = 1.612$	$\frac{1541.9}{614.71} = 2.508$	$\frac{460.21}{99.7} = 4.614$
6	25	35	25	0	$\frac{3085.9}{1988.3} = 1.552$	$\frac{1580.5}{657.37} = 2.404$	$\frac{473.17}{108.3} = 4.369$
7	25	35	25	5	$\frac{3063.4}{1970} = 1.555$	$\frac{1568.6}{653.13} = 2.402$	$\frac{469.43}{107.7} = 4.358$
8	25	35	25	10	$\frac{2995.3}{1915} = 1.564$	$\frac{1533.2}{631.09} = 2.430$	$\frac{458.54}{102.7} = 4.466$
9	25	40	20	0	$\frac{2975}{1756.7} = 1.694$	$\frac{1511.5}{558.35} = 2.707$	$\frac{470.94}{88.1} = 5.346$

Table 6.8.: Examples of parameters for the preselection cuts and the resulting values for S and $\frac{S}{B}$. For scenario 1, only the four preselection cuts were applied. Scenario 2 has a tight W mass constraint in addition to the preselection cuts and in scenario 3, an additional top mass constraint was applied.

6.7.1. Uncertainties

The dominant uncertainties are those on the cross sections of the W+jets background processes. It can be assumed that the relative scaling between the samples with 3, 4, and 5 partons is approximately correct while the absolute cross section is not known very precisely. To test the dependence of the results on variations of the W+jets cross section, the same analysis was performed with the W+jets cross sections scaled up/down by 50%. Of course the purity $\frac{S}{B}$ changes, but this does not influence the relative quality of the cuts.
There are statistical uncertainties as the available samples are finite. These uncertainties are again highly correlated between the different sets of parameters as the analysis was performed on one set of samples. Statistical variations should therefore influence all results in approximately the same way.

6.8. Evaluation of QCD Background Using the Matrix Method

6.8.1. Muon Isolation Criteria

As the matrix method can be used to estimate the rate of fake isolated leptons, a brief overview of what is meant by "isolated leptons" will be given here. The focus will be on muon isolation because that is what was studied, but the same isolation criteria can be applied electrons.
There are several ways that muons can be produced in hadron collisions. One is usually interested in muons from (real) W boson decays, for example $t \to b\,W \to b\,\mu\,\nu_\mu$. Muons can also be produced inside jets, e.g. in weak decays of pions, kaons or B-mesons. These muons can be helpful for flavor-tagging (b-jets are more likely to contain muons than light quark jets) but should not be confused with the first kind of muons. The origin of every muon in the detector cannot be determined exactly. Instead, isolation criteria are used in the hope that the muons that pass these criteria are the ones that are wanted.

Angular Distance between Muon and Jets

If a muon is produced inside a jet, it will generally be boosted in the same direction as the jet. The angular distance $\Delta R(\mu, j)$ between a muon and the closest jet is

therefore a good measure of isolation. There are some problems with this approach, though: The muon will not go in the exact same direction as the jet it was produced in, so one needs to choose the minimum distance ΔR_m for isolated muons accordingly. The more jets there are in the detector, the more likely it is for an isolated muon to be found nearby an unrelated jet. So ΔR_m will have to be a compromise between "loosing" some isolated muons that are produced close to jets and "gaining" muons that were produced inside a jet but with a slight angle between them and the rest of the jet. The probabilities for those outcomes and the best choice for ΔR_m need to be determined from data.
There is another issue, though: The efficiency of this criterion depends on the jet reconstruction algorithm. If a muon was produced inside a jet that was not reconstructed properly or had a p_T of less than 20 GeV, that muon will appear isolated. Because of this, the rate of fake isolated muons depends on the jet algorithm and the p_T cutoff used.
In fig. B.1 and B.2, the $\Delta R(\mu, j)$ spectrum can be seen for the $b\bar{b}$ sample and the $t\bar{t} \to \mu$+jets sample. In the former, there should only be non-isolated muons. The $\Delta R(\mu, j)$ spectrum is expected to peak around $\Delta R(\mu, j) = 0$. This behaviour can be observed: most muons have $\Delta R(\mu, j) < 0.5$. There are some events where larger distances are observed. These probably are events where the muon was produced inside a jet that was not reconstructed properly. The spectrum for $t\bar{t} \to \mu$+jets events shows a small peak at $\Delta R(\mu, j) = 0.05$ but most muons have $\Delta R(\mu, j) > 0.5$ as expected. The peak at small distances can be explained by the fact that there are some high-p_T non-isolated muons produced inside b-jets in this sample. No sharp border between the two types of muons can be observed, but comparing the two spectra, $\Delta R_m = 0.5$ seems like a good choice.

Muon Isolation in the Calorimeter

Another set of isolation criteria focuses on calorimeter information. Most muons produced at the LHC will be minimum ionizing particles and will not deposit much energy in the calorimeters. If there was a lot of energy deposited near the path of the muon, it can be assumed that it was produced inside a jet.
To quantify the energy deposited "near" the muon, two cones (in $\eta - \varphi$-space) are defined: An inner cone, encompassing every point with angular distance ΔR to the muon of 0.05 or less, and an outer cone containing everything that is closer to the muon than 0.2. The quantity `Etcone_20` is then defined as the total amount of

transverse energy deposited inside the outer cone, but outside the inner cone. Other outer cone sizes (e.g. 0.3, 0.4) can also be used, the corresponding quantities are called `Etcone_30` and `Etcone_40`.
In figures B.3 and B.4, the `Etcone_20` spectra of non-isolated and isolated muons are shown. For muons from b-jets, `Etcone_20` peaks at about 5 GeV and falls of slowly. Most muons from top quark decays have `Etcone_20` $< 10\,\text{GeV}$. There is again no obvious value of `Etcone_20` that can be used to distinguish between isolated and non-isolated muons. In the context of this studies, `Etcone_20` $< 10\,\text{GeV}$ was used to define isolated muons, which is a rather loose cut. For studies on real data, the optimal cutoff value needs to be determined from data as it depends on detector response, noise etc. `Etcone_20` is not a perfect parameter to determine muon isolation, either: There could be energy from objects unrelated to, but close to the muon, or some deposited energy could be missed because of dead calorimeter cells.

Other Isolation Criteria

There are other criteria based on tracker information (e.g. number of tracks or total p_T in a hollow cone around the muon). They were not used for these studies.

6.8.2. Composition of the Loose and Tight Sample

As no LHC data is available yet, the loose and tight samples were created using Monte Carlo data. The events in the loose sample have to pass the following cuts:

1. preselection (cf. section 4): At least one jet and one muon with $p_T > 15\,\text{GeV}$.
2. 4-jet requirement: There need to be at least four jets with $p_T > 20\,\text{GeV}$.
3. 3-jet requirement: Of those jets, at least three need to have $p_T > 40\,\text{GeV}$.
4. muon requirement: There needs to be exactly one muon.

The tight sample contains those events that pass the loose cuts, with the additional constraint that the muon has to be isolated (`Etcone_20` $< 10\,\text{GeV}$). This is a loose isolation cut. I would have liked to use a tighter cut here and introduce a loose isolation requirement to the loose sample. However, there were not enough $b\bar{b}$ events over after these cuts as there were only about $10\,\text{pb}^{-1}$ of $b\bar{b}$ events with muons available, so the cuts had to be loosened to get acceptable statistics.

The true efficiencies ε^{tr} for signal/background events to pass the isolation cut were determined:

$$\varepsilon_s^{tr} = 0.9436 \pm 0.0007, \qquad \varepsilon_b^{tr} = 0.15 \pm 0.01.$$

The efficiencies of the loose cuts for signal/background are

$$\varepsilon_s^{loose} = 0.2883 \pm 0.0007, \qquad \varepsilon_b^{loose} = 0.0050 \pm 0.0001.$$

From that and the cross sections from tab. 4.1, the expected number of signal events in the loose and the tight sample can be computed. As the cross sections for QCD events are not very reliable, the analysis was performed with a varying amount of background in the loose sample (the background ratio $\frac{N_b^l}{N_s^l}$ was varied in steps of 0.1 from 0.1 to 0.9). The resulting composition can be seen in table 6.9.

background ratio	N^l	N_s^l	N_b^l	N^t	N_s^t	N_b^t
0.1	4473.25	4025.93	447.33	3867.64	3798.71	68.93
0.2	5032.41	4025.93	1006.48	3953.8	3798.71	155.09
0.3	5751.33	4025.93	1725.4	4064.58	3798.71	265.87
0.4	6709.88	4025.93	2683.95	4212.28	3798.71	413.57
0.5	8051.86	4025.93	4025.93	4419.07	3798.71	620.36
0.6	10064.82	4025.93	6038.89	4729.25	3798.71	930.54
0.7	13419.76	4025.93	9393.83	5246.21	3798.71	1447.5
0.8	20129.64	4025.93	16103.71	6280.14	3798.71	2481.43
0.9	40259.28	4025.93	36233.36	9381.93	3798.71	5583.22

Table 6.9.: True composition of loose/tight sample for varying amount of background.

6.8.3. Determination of Isolation Efficiencies

In the context of this study, the isolation efficiency for signal events was determined from Monte Carlo, so $\varepsilon_s = \varepsilon_s^{tr} = 0.9436 \pm 0.0007$. For a real analysis, one could determine this efficiency from data. It is possible to get highly pure $Z \to \mu\mu$ samples (which should have only isolated muons) from data and use them to determine ε_s. It is however not possible to get a pure QCD sample from data in the interesting regions. ε_b is usually determined in regions with low $\not\!\!E_T$ as we do not expect many signal events there. It then has to be assumed that the isolation efficiency can be extrapolated to regions with high $\not\!\!E_T$. To test the validity of this assumption, ε_b was calculated for different values of $\not\!\!E_T$. The results are presented in figure B.5. The

distribution is mostly flat, but ε_b is slightly higher than average in the first three bins.

Two scenarios were considered. For the first one, ε_b was calculated from Monte Carlo, using background events that passed the loose cuts minus the $\not{E}_T$ requirement and had $\not{E}_T < 15\,\text{GeV}$. For the second scenario, signal events passing the same cuts were also considered when calculating ε_b.

6.8.4. Results

Using the values from table 6.9 and equation 5.5, the amount of signal/background in the tight sample was estimated for both scenarios. The results are displayed in tables 6.10 and 6.11. To estimate the uncertainties, equation 5.6 was used. Only statistical uncertainties on ε_s and ε_b were considered. They were calculated using the approximation $\Delta\varepsilon = \sqrt{\frac{\varepsilon\cdot(1-\varepsilon)}{N}}$. In the case of $\Delta\varepsilon_s$ and $\Delta\varepsilon_b$ in scenario 1, N was taken to be the number of events in the Monte Carlo sample used to calculate the true efficiency. For $\Delta\varepsilon_b$ in scenario 2, N represents the expected number of events used to calculate the efficiency. Because of that, $\Delta\varepsilon_b$ goes down as the expected amount of background goes up.

In extrapolating ε_b to regions with higher $\not{E}_T$, the amount of background in the sample is overestimated while the amount of signal is underestimated. Also, the uncertainties are quite large, even though systematic uncertainties on ε_s and ε_b were not considered. For now, the uncertainties for larger background ratios are dominated by uncertainties on ε_b and (especially in scenario 2) by the statistical uncertainties on N^t.

background ratio	N^l	N^t	N_s^t	N_b^t
0.1	4473 ± 67	3868 ± 62	3765 ± 64	102 ± 10
0.2	5032 ± 71	3954 ± 63	3724 ± 67	230 ± 19
0.3	5751 ± 76	4065 ± 64	3670 ± 71	394 ± 30
0.4	6710 ± 82	4212 ± 65	3599 ± 80	613 ± 45
0.5	8052 ± 90	4419 ± 66	3499 ± 95	920 ± 67
0.6	10065 ± 100	4729 ± 69	3350 ± 122	1380 ± 99
0.7	13420 ± 116	5246 ± 72	3100 ± 170	2146 ± 153
0.8	20130 ± 142	6280 ± 79	2601 ± 274	3679 ± 262
0.9	40259 ± 201	9382 ± 97	1105 ± 594	8277 ± 586

Table 6.10.: Reconstructed composition of tight sample for varying amount of background (scenario 1).

background ratio	N^l	N^t	ε_b	N_s^t	N_b^t
0.1	4473 ± 67	3868 ± 62	0.473 ± 0.019	3513 ± 75	354 ± 37
0.2	5032 ± 71	3954 ± 63	0.356 ± 0.014	3473 ± 74	481 ± 35
0.3	5751 ± 76	4065 ± 64	0.303 ± 0.010	3420 ± 75	645 ± 38
0.4	6710 ± 82	4212 ± 65	0.273 ± 0.008	3349 ± 78	864 ± 42
0.5	8052 ± 90	4419 ± 66	0.254 ± 0.007	3249 ± 83	1170 ± 47
0.6	10065 ± 100	4729 ± 69	0.240 ± 0.005	3099 ± 89	1630 ± 55
0.7	13420 ± 116	5246 ± 72	0.230 ± 0.004	2850 ± 99	2396 ± 66
0.8	20130 ± 142	6280 ± 79	0.223 ± 0.003	2351 ± 116	3929 ± 83
0.9	40259 ± 201	9382 ± 97	0.217 ± 0.002	854 ± 156	8527 ± 121

Table 6.11.: Reconstructed composition of tight sample for varying amount of background (scenario 2).

7. Summary and Outlook

In this thesis, I have studied two methods used to determine the top quark pair production cross section, and a simple algorithm for reconstructing the hadronic top quark. I have shown that these simple techniques can in principle be used to determine the cross section using the first $200\,\mathrm{pb}^{-1}$ of data. The top quark mass peak should be clearly visible using this simple algorithm. However, applying a simple cut-and-count method will lead to large uncertainties on the cross section. For this method to be useful, the background processes have to be better understood first. Data-based approaches like the matrix method can be used to estimate QCD background, but might lead to large systematic uncertainties.

For a serious measurement, other methods will be needed, for example multivariate analyses that will lead to better background suppression while retaining more signal events.

I have also shown that the CSC preselection cuts, which had been optimized for a center-of-mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$, are still useful at $\sqrt{s} = 10\,\mathrm{TeV}$, at which the first physics run of the LHC will probably take place. In the context of further analyses, it might be advisable to change the jet p_T requirements slightly. The consequences of lowering or removing the cuts on missing E_T should also be studied, especially the effects on QCD background suppression.

A. Tables

	$N_{jet} = 4$	$N_{jet} = 5$	$N_{jet} = 6$	$N_{jet} = 7$
s	random guessing (assuming perfect matching)			
0	0	0	0.050	0.114
1	0	0.300	0.450	0.514
2	0.750	0.600	0.450	0.343
3	0.250	0.100	0.050	0.029
s	top quark algorithm without any mass constraints			
0	0.013 ± 0.001	0.046 ± 0.001	0.112 ± 0.003	0.177 ± 0.006
1	0.174 ± 0.002	0.271 ± 0.003	0.326 ± 0.004	0.342 ± 0.008
2	0.537 ± 0.003	0.507 ± 0.003	0.430 ± 0.004	0.364 ± 0.008
3	0.276 ± 0.002	0.176 ± 0.002	0.132 ± 0.003	0.116 ± 0.005
s	top quark algorithm with tight W mass cut			
0	0.007 ± 0.001	0.028 ± 0.001	0.069 ± 0.003	0.109 ± 0.008
1	0.106 ± 0.002	0.188 ± 0.003	0.237 ± 0.006	0.275 ± 0.011
2	0.484 ± 0.003	0.509 ± 0.004	0.470 ± 0.007	0.410 ± 0.012
3	0.403 ± 0.003	0.275 ± 0.004	0.225 ± 0.005	0.206 ± 0.010
s	top quark algorithm with tight W mass cut and top mass cut			
0	0.003 ± 0.001	0.012 ± 0.002	0.030 ± 0.004	0.096 ± 0.015
1	0.051 ± 0.003	0.106 ± 0.005	0.135 ± 0.009	0.124 ± 0.017
2	0.298 ± 0.005	0.312 ± 0.007	0.298 ± 0.012	0.269 ± 0.023
3	0.649 ± 0.006	0.570 ± 0.008	0.537 ± 0.013	0.510 ± 0.025

Table A.1.: Performance of the top quark decay candidate algorithm for different jet multiplicities. s refers to the number of correctly identified jets.

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.053 ± 0.001	0.037 ± 0.001	0.031 ± 0.001
1	0.246 ± 0.001	0.195 ± 0.002	0.166 ± 0.002
2	0.497 ± 0.002	0.500 ± 0.002	0.485 ± 0.002
3	0.204 ± 0.001	0.268 ± 0.002	0.318 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.023 ± 0.001	0.018 ± 0.001	0.014 ± 0.001
1	0.127 ± 0.002	0.106 ± 0.002	0.084 ± 0.002
2	0.354 ± 0.003	0.340 ± 0.004	0.302 ± 0.004
3	0.496 ± 0.004	0.536 ± 0.004	0.600 ± 0.004

Table A.2.: Performance of the top quark decay candidate algorithm using $\Delta R_m = 0.4$. s refers to the number of correctly identified jets.

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.048 ± 0.001	0.033 ± 0.001	0.027 ± 0.001
1	0.237 ± 0.001	0.185 ± 0.002	0.157 ± 0.002
2	0.501 ± 0.002	0.502 ± 0.002	0.486 ± 0.002
3	0.215 ± 0.001	0.280 ± 0.002	0.330 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.023 ± 0.001	0.015 ± 0.001	0.012 ± 0.001
1	0.127 ± 0.002	0.100 ± 0.002	0.079 ± 0.002
2	0.354 ± 0.003	0.337 ± 0.004	0.299 ± 0.004
3	0.496 ± 0.004	0.548 ± 0.004	0.610 ± 0.004

Table A.3.: Performance of the top quark decay candidate algorithm using $\Delta R_m = 0.6$. s refers to the number of correctly identified jets.

	$N_{jet} = 4$	$N_{jet} = 5$	$N_{jet} = 6$	$N_{jet} = 7$
s	random guessing (assuming perfect matching)			
0	0.17	0.3	0.400	0.476
1	0.67	0.6	0.533	0.476
2	0.17	0.1	0.067	0.048
s	W boson algorithm without any mass constraints			
0	0.199 ± 0.002	0.274 ± 0.003	0.360 ± 0.004	0.428 ± 0.008
1	0.492 ± 0.003	0.473 ± 0.003	0.438 ± 0.004	0.398 ± 0.008
2	0.308 ± 0.002	0.252 ± 0.003	0.202 ± 0.004	0.174 ± 0.006
s	W boson algorithm with tight W mass cut			
0	0.138 ± 0.002	0.197 ± 0.003	0.262 ± 0.006	0.306 ± 0.011
1	0.419 ± 0.003	0.425 ± 0.004	0.409 ± 0.006	0.400 ± 0.012
2	0.444 ± 0.003	0.378 ± 0.004	0.329 ± 0.006	0.295 ± 0.011
s	W boson algorithm with tight W mass cut and top mass cut			
0	0.089 ± 0.003	0.121 ± 0.005	0.167 ± 0.010	0.194 ± 0.020
1	0.313 ± 0.005	0.333 ± 0.007	0.322 ± 0.012	0.350 ± 0.024
2	0.598 ± 0.006	0.546 ± 0.008	0.512 ± 0.013	0.456 ± 0.025

Table A.4.: Performance of the W boson decay candidate algorithm for different jet multiplicities. s refers to the number of correctly identified jets.

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.271 ± 0.002	0.225 ± 0.002	0.191 ± 0.002
1	0.467 ± 0.002	0.437 ± 0.002	0.413 ± 0.002
2	0.262 ± 0.002	0.339 ± 0.002	0.396 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.172 ± 0.003	0.147 ± 0.003	0.117 ± 0.003
1	0.359 ± 0.003	0.345 ± 0.004	0.319 ± 0.004
2	0.469 ± 0.004	0.508 ± 0.004	0.564 ± 0.004

Table A.5.: Performance of the W boson decay candidate algorithm using $\Delta R_m = 0.4$. s refers to the number of correctly identified jets.

	without top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.259 ± 0.002	0.213 ± 0.002	0.181 ± 0.002
1	0.474 ± 0.002	0.444 ± 0.002	0.420 ± 0.002
2	0.267 ± 0.002	0.343 ± 0.002	0.400 ± 0.002
	with top mass cut		
s	no W mass cut	loose W mass cut	tight W mass cut
0	0.163 ± 0.003	0.139 ± 0.003	0.109 ± 0.003
1	0.362 ± 0.003	0.348 ± 0.004	0.322 ± 0.004
2	0.475 ± 0.004	0.513 ± 0.004	0.569 ± 0.004

Table A.6.: Performance of the W boson decay candidate algorithm using $\Delta R_m = 0.6$. s refers to the number of correctly identified jets.

B. Plots

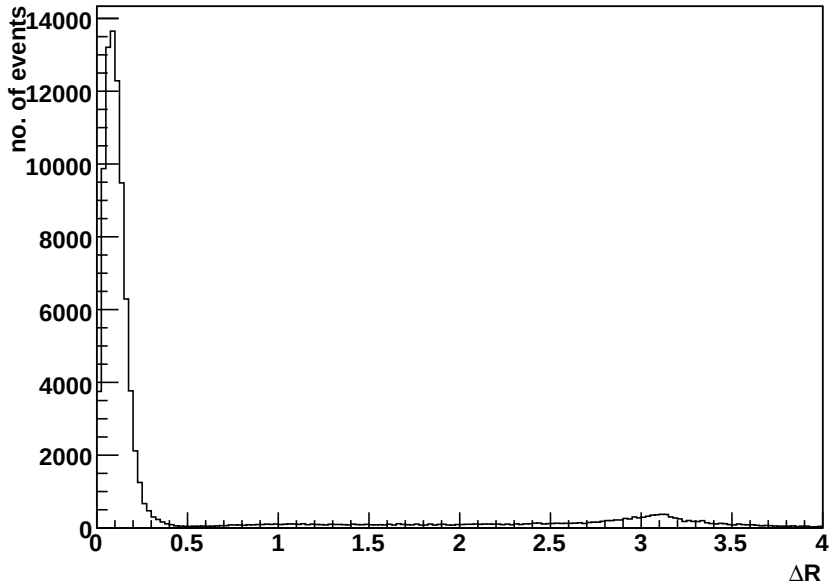

Figure B.1.: Angular distance ΔR between muons and closest jet in $b\bar{b}$ sample (non-isolated muons).

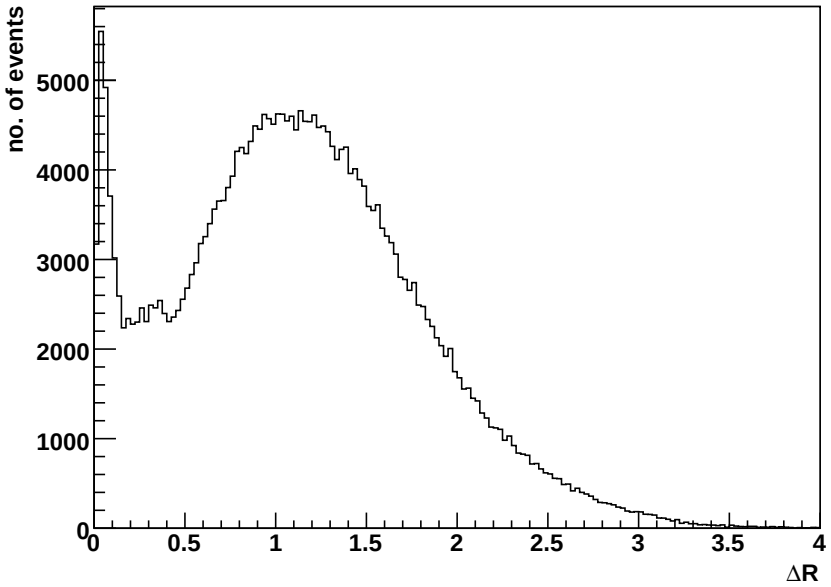

Figure B.2.: Angular distance ΔR between leading muon and closest jet in $t\bar{t}$ sample (mostly isolated muons).

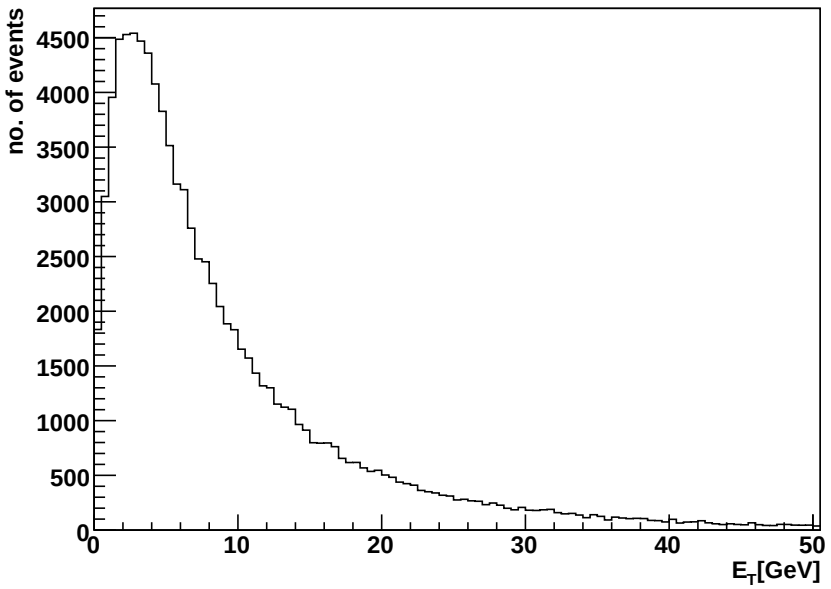

Figure B.3.: Transverse energy in hollow cone around muons in $b\bar{b}$ sample (non-isolated muons).

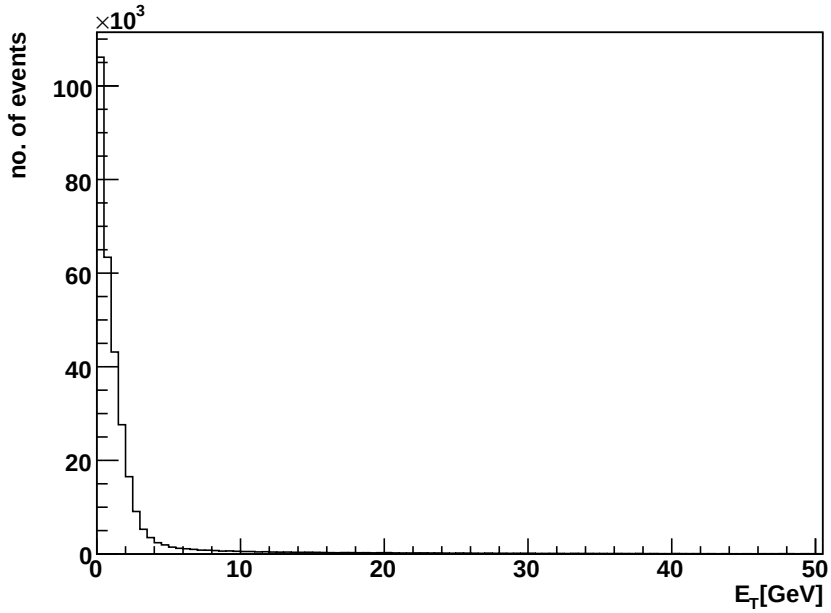

Figure B.4.: Transverse energy in hollow cone around leading muon in $t\bar{t}$ sample (mostly isolated muons).

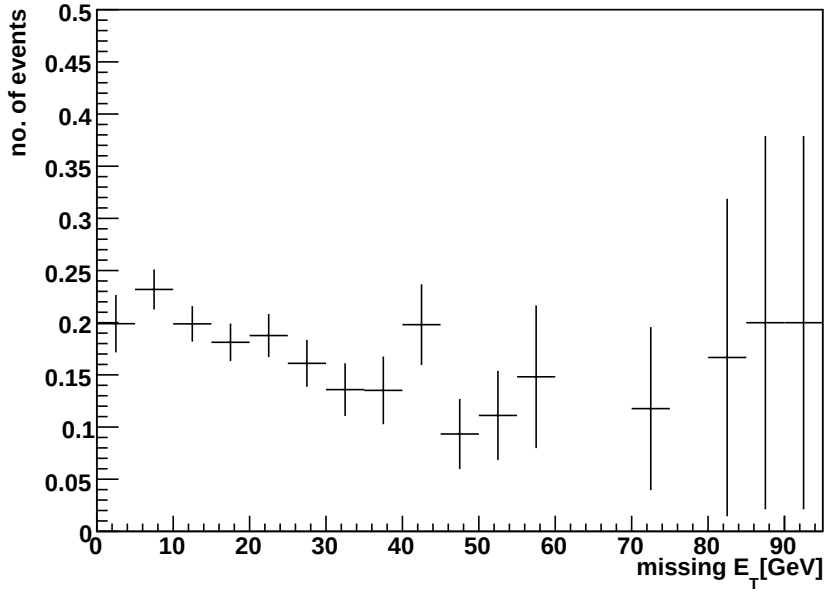

Figure B.5.: Isolation efficiency for background events for varying amounts of $\not\!\!E_T$.

Bibliography

[1] The ATLAS Collaboration, G. Aad, et al. *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics*. CERN, Geneva, 2008.

[2] C. Berger. *Elementarteilchenphysik*. Springer, 2nd edition, 2006.

[3] Tevatron Electroweak Working Group. Combination of CDF and D0 Results on the Mass of the Top Quark. 2009, arXiv:0903.2503 [hep-ex]. FERMILAB-TM-2427-E, TEVEWWG-TOP-2009-03, CDF-NOTE-9717, DØ-NOTE-5899.

[4] The DØ Collaboration, S. Abachi, et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[5] The CDF Collaboration, F. Abe, et al. Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[6] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[7] The DØ Collaboration, V. M. Abazov, et al. Experimental discrimination between charge 2e/3 top quark and charge 4e/3 exotic quark production scenarios. *Physical Review Letters*, 98(4):041801, 2007.

[8] The CDF collaboration, Z. Gunay Unalan, et al. First CDF measurement of the top quark charge using the top decay products. *Nucl. Phys. Proc. Suppl.*, 177-178:297–299, 2008.

[9] C. Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[10] S. Moch and P. Uwer. Heavy-quark pair production at two loops in QCD. *Nucl. Phys. Proc. Suppl.*, 183:75–80, 2008.

[11] The DØ Collaboration, V. M. Abazov, et al. Combination and interpretations of $t\bar{t}$ cross section measurements with the D0 detector. 2009, arXiv:0903.5525 [hep-ex]. Submitted to Phys. Lett. D.

[12] The DØ Collaboration, V. M. Abazov, et al. Observation of Single Top Quark Production. 2009, 0903.0850. Submitted to Phys. Rev. Lett.

[13] The CDF Collaboration, T. Aaltonen, et al. First Observation of Electroweak Single Top Quark Production. 2009, arXiv:0903.0885v3 [hep-ex]. Submitted to Phys. Rev. Lett.

[14] M. Bosman et al. Understanding Monte Carlo Generators for Top Physics. ATL-COM-PHYS-2009-334.

[15] Peter Jenni, Markus Nordberg, Marzio Nessi, and Kerstin Jon-And. *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Technical Design Report. Geneva, 2008.

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

[17] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock. *LHC Design Report*. CERN, Geneva, 2004.

[18] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.

[19] S. Frixione, P. Nason, and B. R. Webber. Matching NLO QCD and parton showers in heavy flavour production. *JHEP*, 08:007, 2003.

[20] M. L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[21] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour. Multiparton Interactions in Photoproduction at HERA. *Z. Phys.*, C72:637–646, 1996.

[22] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, February 2006.

[23] P. Calafiura et al. The athena control framework in production, new developments and lessons learned. In *Computing in High Energy Physics and Nuclear Physics*, page 456, Interlaken, Switzerland, 2004.

[24] F. Kohn. A software package for event reweighting according to trigger efficiencies. ATLAS Internal Note, publication in preparation.

[25] The DØ Collaboration, V. M. Abazov, et al. Measurement of the t anti-t production cross section in p anti-p collisions at s**(1/2) = 1.96-TeV using secondary vertex b tagging. *Phys. Rev.*, D74:112004, 2006.

[26] T. Golling. *Measurements of the Top Quark Pair Production Cross Section in Lepton+Jets Final States using a Topological Multivariate Technique as well as Lifetime b-Tagging in Proton-Antiproton Collisions at* $\sqrt{s} = 1.96$ *TeV with the D0 Detector at the Tevatron.* PhD thesis, Universität Bonn, 2005.

[27] E. Barberis, T. Golling, I. Iashevili, A. Quadt, and P. Schieferdecker. The Matrix Method and its Error Calculation. DØ-Note 4564.

Thanks

Writing this thesis would not have been possible without the support of a lot of people.
First and foremost, I want to thank Prof. A. Quadt for giving me the opportunity to work in his department and being my first referee. I also want to thank Prof. A. Frey and Prof. Bosman for reviewing this thesis.
I want to thank my advisors J. Meyer and A. Roe for answering all my questions and helping me during the writing process. I had a great time working with you!
I also want to thank everyone from the high-energy physics group who helped me and answered my many questions, especially D. Schiepel.
Thanks to my fellow B.Sc. candidates E. Fuchs, J. Hofestädt, and A. Runde, who provided great inspiration and helped edit this document.
Last but not least I want to thank my family and friends for supporting me and providing some entertainment outside of particle physics.
I hope I have not forgotten to mention anyone!

Erklärung nach §13(8)der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, 17.7.2009

(Henrike Fleischhack)