Measurements on top quark pairs in proton collisions recorded with the ATLAS detector
Mussche, I.

Citation for published version (APA):
Mussche, I. (2012). Measurements on top quark pairs in proton collisions recorded with the ATLAS detector

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
This thesis describes measurements performed with data sets recorded at different times, by the ATLAS detector. The data is divided into two distinct periods. The first data set, used to perform the top quark pair production cross section measurement in Chapter 5, is produced in 2010 and corresponds to an integrated luminosity of 35 pb$^{-1}$. In Chapter 6, we make use of 1.04 fb$^{-1}$ of data that is obtained in the period between April and July 2011 to measure the top quark charge asymmetry.

In this chapter, we introduce the data sets and scrutinize the characteristics of the objects we are interested in. This includes the standard event selection that is applied in all analyses, and the distributions of the observables we are interested in. Furthermore, we include a description of the background estimates that are obtained with data-driven techniques. Finally, in Section 4.7, we treat the reconstruction of the top quarks with the Kinematic Likelihood Fitter that will ultimately be used in the analyses in Chapters 5 and 6.

### 4.1 Introduction

In 2010 and 2011, proton beams ran with an energy of 3.5 TeV each. The maximum instantaneous luminosity was $7 \cdot 10^{27}$ cm$^{-2}$ s$^{-1}$ in the first week of 2010 data taking and the total integrated luminosity reached 90 µb$^{-1}$. At that time only one colliding bunch was present in each beam. Since then, the luminosity numbers increased strongly over time, due to the LHC improving the beam conditions. The beam focusing improved to form a smaller interaction area, the number of protons per bunch was increased and, on top of that, the number of bunches per beam was raised to 6 at the end of October (the
end of the 2010 LHC run). The resulting maximum instantaneous luminosity was almost five orders of magnitudes higher at the end of this data taking period: $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. The total recorded data we analyze, split up in periods following the ATLAS convention, is shown in Table 4.1. This includes the instantaneous luminosity and the average number of interactions per bunch crossing. The total integrated luminosity delivered by the LHC over the entire period, reached 50 pb$^{-1}$, of which ATLAS recorded 45 pb$^{-1}$. The periods we list form a consistent set of data, corresponding to 42.7 pb$^{-1}$. The data quality requirements reduce these numbers slightly, as is described later.

In 2011, the instantaneous luminosity increased even further, to $1.3 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. The average number of events per bunch crossing ranges from 6 to 8, producing extra interactions in the events we are interested in. In addition, the bunch spacing was reduced to 50 ns, resulting in additional activity in events. The phenomenon of extra interactions and activity that contaminate the events of interest is called ‘pile-up’ and is described in the following sections. The total recorded luminosity of the 2011 data set we analyze corresponds to 1.2 fb$^{-1}$ before quality requirements.

<table>
<thead>
<tr>
<th>Year</th>
<th>Period</th>
<th>Int. luminosity (pb$^{-1}$)</th>
<th>Peak luminosity ($\times 10^{32} \text{ cm}^{-2}$)</th>
<th>Peak evts per bunch crossing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>E</td>
<td>1.12</td>
<td>0.0391</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>1.96</td>
<td>0.102</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>8.81</td>
<td>0.701</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>8.81</td>
<td>1.47</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>23.0</td>
<td>2.03</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>42.7</td>
<td>2.03</td>
<td>3.71</td>
</tr>
<tr>
<td>2011</td>
<td>B</td>
<td>17.0</td>
<td>2.44</td>
<td>8.48</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>179</td>
<td>6.65</td>
<td>7.30</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>50.2</td>
<td>8.37</td>
<td>7.60</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>152</td>
<td>11.1</td>
<td>8.07</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>561</td>
<td>12.7</td>
<td>8.02</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>278</td>
<td>12.7</td>
<td>6.89</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1237</td>
<td>12.7</td>
<td>8.48</td>
</tr>
</tbody>
</table>
4.1. Introduction

4.1.1 Pile-up

We define as pile-up the occurrence of extra activity in the detector that results from proton scatterings in the same, or neighboring, bunch crossings that do not correspond to the interaction that triggered the event. The average number of events per bunch crossing (at peak luminosity) was 0.01 at the LHC start-up, and increased to almost 4 by the end of 2010. This means that any interaction that triggered ATLAS to record an event is most of the time polluted by extra vertices and signals. If additional protons from the same bunch that triggered the event collide, it is called in-time pile-up. If interactions from the preceding or succeeding bunches of protons leave behind signals in the collision of interest, they form out-of-time pile-up. In 2010, the bunch spacing of the proton beams was such that out-of-time pile-up was negligible.

The simulation samples we use are already corrected for pile-up effects. Additional interactions are simulated and added to the primary hard scatter. As a result, the distribution of the number of vertices with more than four good tracks in simulation is reasonably compatible with data in 2010. But, since in simulation only one setup value for pile-up is simulated for the full year and in reality pile-up conditions change during the year, small differences remain. This residual difference between data and simulation is considered as a systematic uncertainty.

In 2011 out-of-time pile-up becomes more relevant. The final data runs we use here are collected with a configuration of 1318 bunches in the beam (compared to ∼ 130 towards the end of 2010 and at the beginning of 2011). The time window between consecutive bunches (bunch spacing) is reduced to 50 ns. Out-of-time pile-up from neighboring bunches generates extra signals in the recorded event, more calorimeter activity, for instance, but not necessarily extra vertices. Simulated events are produced under the assumption of a certain distribution of the number of pile-up interactions, but in the presence of out-of-time pile-up, the simulation needs to be reweighted according to the pile-up conditions of the particular data set. Reweighting is performed using the average number of interactions per bunch crossing. This is an observable that is extracted from the data. Figure 4.1 shows the distributions of data (only period D of data) and simulation, before reweighting. The number of interactions in data is lower in this period than was anticipated in simulation. Events in the simulation are reweighted according to the data sets that are being studied. That means that in this particular example, weights in the range 0.0-4.0 occur.

4.1.2 Data quality

Not all events in the data meet identical and optimal detection quality requirements. The operation of the detector is monitored by a series of checks of standard distributions, of all subdetector parts, trigger figures, and of reconstructed physics objects. Collision recorded in a time window with approximately constant luminosity, form a ‘lumiblock’. A lumiblock is typically a minute long. The condition of all subdetector parts is stored per lumiblock, making it possible to select lumiblocks that satisfy the (sub)detector quality
Chapter 4. Selection and characteristics of data

Figure 4.1 – Average number of interactions per bunch crossing, before reweighting.

checks a specific analysis requires. In top quark analyses, where all subdetectors are of importance, a considerable fraction of lumiblocks is not accepted. This is mostly due to problems in the detector control or readout in one of the subsystems that occasionally occurs. After applying the quality criteria of all subdetector components, in total 35.3 pb\(^{-1}\) is left of the initial 42 pb\(^{-1}\) collected in 2010. In 2011, the data set is reduced from 1.24 fb\(^{-1}\) to 1.04 fb\(^{-1}\) due to the quality restrictions.

After the quality filter, the sample can still contain signals that do not originate from a hard interaction, as noise or cosmic ray events may have triggered the event. For this reason, a quality requirement on the track vertices is imposed. Vertices are reconstructed from all tracks that are present within a collision, and a vertex is considered of good quality when at least five tracks originate from it. Due to pile-up there can be multiple good vertices in one event. Events are selected if there is at least one ‘good’ vertex.

A hardware problem in the liquid argon calorimeter, caused by a readout device that broke down, affects the acceptance and energy resolution of electrons and jets. We account for this issue by discarding all events that contain jets that overlap with the troubled region of the detector, starting from run 180614 (April 30th, 2011). The electron quality requirements are also adapted such that unreliable measured values are avoided. We apply the same selection to a subset of the simulation to account for the acceptance and rescale the luminosity loss due to this effect.

4.2 Object definitions for analyses

In this section, we describe the identification and reconstruction of the objects that are relevant to the \(\bar{t}t\) analysis: electrons, muons, jets and missing transverse energy (\(\not{E}_T\)). Since we have to deal with ambiguities and overlap between these objects we discuss them in the order that overlap is resolved.
4.2. Object definitions for analyses

4.2.1 Electrons

An event is selected if it was triggered by one of the chosen single-electron triggers. We require a trigger that is efficient for the kinematic region above 20 GeV. For electrons we used $\text{EF}_\text{e15\_medium}$ in 2010, hence with an online transverse energy threshold of 15 GeV (see Section 2.2.6 for the naming conventions). For 2011, we switch to $\text{EF}_\text{e20\_medium}$ (electron), the lowest unprescaled trigger, which has an online threshold of 20 GeV for the transverse energy. Consequently, as described in Section 3.4.1, we apply a threshold of 20 (25) GeV for the transverse energy of reconstructed electrons in the 2010 (2011) analysis. The margin between the online and offline energy threshold is added to avoid inefficiencies around the threshold region (turn-on effects), see Section 3.4.1.

We require that the triggered object matches the reconstructed object, so that the electron we use in the analysis is the one that triggered the event. The absolute pseudorapidity (of the cluster) is required to be within 2.47, with the transition region of the detector, $1.37 < |\eta| < 1.52$, excluded. Finally, isolation criteria are applied: if the amount of energy in a cone of $\Delta R = 0.2$ around the electron candidate exceeds 4 GeV it is considered a jet, because then it is likely to be a misidentified jet, or an electron produced in the decay of the jet.

4.2.2 Jets

We use jets reconstructed with the anti-$k_t$ algorithm, as discussed in Section 3.4.2, with $\Delta R = 0.4$. The $p_T$ threshold of jets to be used in the analysis is 20 GeV, but in the event selection jets with a transverse momentum of at least 25 GeV are considered for the jet multiplicity cut. For 2010, we set the maximal absolute value of the pseudorapidity to 2.5. In 2011, the calibration of jets has developed, making it possible to use jets in a range up to $|\eta| < 4.5$. This means that in 2011 we can make use of very forward jets.

Events with jets that contain energy deposits falling in the unresponsive part of the calorimeter are discarded. Jets are removed if an accepted electron is present within $\Delta R = 0.2$.

4.2.3 Muons

The muon is required to have a transverse momentum above 20 GeV and $|\eta| < 2.5$, because of the trigger requirements (see Section 3.4.2). Muons too are required to be isolated, to avoid accepting fake muons, or muons produced in the core of a jet. Therefore, the calorimeter energy in a cone of $\Delta R = 0.3$ around the muon must below 4 GeV. In addition, the combined transverse momentum of all tracks within that cone cannot exceed 4 GeV either, to attempt to select only muons that originate from the hard scatter. And finally, muons that are within $\Delta R = 0.4$ from a jet are vetoed as well, to remove any potentially remaining overlap. If two back-to-back muons satisfy the signature of a cosmic muon, they are both removed.
4.2.4 Corrections to leptons

There are a number of corrections that are applied to leptons in top quark events. Firstly, the efficiency differences in simulation and data are accounted for: as mentioned in Section 3.4, we measure the trigger, reconstruction and identification efficiencies of leptons in both data and simulation. The relative differences between the respective efficiencies are corrected for by applying scale factors to the simulated events. We summarize the scale factors in Table 4.2, with the reconstruction and identification step convoluted. The methods to obtain the scale factors differ between the data sets (2010, 2011) and the lepton flavors.

For 2010, the scale factor of the electron trigger corresponds to a single number, whereas the reconstruction and identification scale factors are parametrized as a function of $\eta$ and $E_T$. For muons it is the other way around, the reconstruction and identification scale factors correspond to a single number, while the trigger scale factor is expressed in bins of $\eta$ and $\phi$. Most scale factors are close to, or compatible with 1.

In 2011 data, the binning is finer and in terms of more observables. Because of this, larger values for the scale factors may occur in some areas of the detector, due to statistical fluctuations. For example, the muon scale factors can run up to 1.6 for high momentum muons in certain parts of the endcaps. All events we select are weighted with the product of the trigger and reconstruction/ID scale factors.

<table>
<thead>
<tr>
<th>Year</th>
<th>Lepton flavor</th>
<th>Type</th>
<th>Value (range)</th>
<th>Dependent on</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>e</td>
<td>Trigger</td>
<td>0.995</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>Reco+ID</td>
<td>0.907-1.055</td>
<td>$\eta, E_T$</td>
</tr>
<tr>
<td></td>
<td>$\mu$</td>
<td>Trigger</td>
<td>0.657-1.046</td>
<td>$\eta, \phi$</td>
</tr>
<tr>
<td>2011</td>
<td>e</td>
<td>Trigger</td>
<td>0.966-0.997</td>
<td>$\eta$</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>Reco+ID</td>
<td>0.953-1.117</td>
<td>$\eta, E_T$</td>
</tr>
<tr>
<td></td>
<td>$\mu$</td>
<td>Trigger</td>
<td>0.85-1.57</td>
<td>$\eta, \phi, p_T$</td>
</tr>
<tr>
<td></td>
<td>$\mu$</td>
<td>Reco+ID</td>
<td>0.81-1.02</td>
<td>$\eta, \phi$</td>
</tr>
</tbody>
</table>

4.3 Multijet background

As explained in Section 3.3.2, the uncertainty on the cross section of multijet processes is large, and the numbers of events that need to be produced to have a reliable simulated sample is so large that simulation of multijet events is not feasible. Therefore the shape
and normalization of multijet events are estimated from data. The procedure we follow is somewhat different for the electron and muon channels.

We distinguish between fake leptons and non-prompt leptons. A fake lepton is an object, usually a jet, that is falsely identified as a lepton, because of a reconstruction inefficiency. A non-prompt lepton is a lepton that is a decay product of a pion or a $B$ meson, for example. Photon conversions also lead to non-prompt leptons. Typically fake leptons are electrons, because the identification of jets, photons and electrons are all largely based on calorimeter information. Fake muons are rare, because it is more difficult for other particles to traverse the other detectors and subsequently leave a track in the muon detector.

We discuss the methods for both channels in the following sections$^1$.

### 4.3.1 Muon channel

In the muon channel the fraction of genuine (but non-prompt) muons is large. Non-prompt muons from meson decays are typically non-isolated, as activity coming from the other decay products is usually present. A matrix method is deployed to estimate the amount of multijet events in the muon channel [69]. A ‘loose’ sample of muons is obtained by removing the isolation criteria. This is done such that the loose selection is a subset of the ‘tight’ or standard sample, which is the nominal muon selection. The number of events in the loose and standard sample can be expresses as a sum of the real muon (‘real’) and the fake and non-prompt component (‘fake’):

\[ N_{\text{loose}} = N_{\text{loose}}^{\text{real}} + N_{\text{loose}}^{\text{fake}} \]
\[ N_{\text{standard}} = \epsilon_{\text{real}} N_{\text{loose}}^{\text{real}} + \epsilon_{\text{fake}} N_{\text{loose}}^{\text{fake}}. \]

The efficiency $\epsilon_{\text{real}}$ ($\epsilon_{\text{fake}}$) is the fraction of events in the standard selection with respect to the loose sample, of the real (fake) component. This set of equations can be solved for $N_{\text{standard}}$, the number we are interested in, as a function of the two efficiencies and the total number of events in the loose and standard sample. The efficiencies are obtained from control regions. The efficiency of a prompt muon to end up in both the loose and standard selection can be extracted from $Z \rightarrow \mu\mu$ events. The efficiency of a non-prompt or fake muon to be found in both loose and standard is obtained from a multijet enriched sample, with $E_T < 10$ GeV. Including the efficiencies and measuring the number of events in the loose and standard selection leads to an estimate of the normalization of misidentified muons in the data sample.

### 4.3.2 Electron channel

In the electron channel non-prompt electrons, photon conversions and misidentified jets contaminate the electron selection in unknown amounts. Since no representative control

---

$^1$The final yields of the multijet background are listed in Tables 4.3 and 4.4.
region could be defined in the 2010 data set, a binned likelihood fit is applied to a data sample in the sideband area. The sideband is formed by the region with small missing transverse energy, $E_T < 35$ GeV (see selection criteria).

The fit is built up from templates for the $E_T$ distribution of $t\bar{t}$, $W$+jets, $Z$+jets. The multijet template is obtained from an ‘anti-electron sample’, a sample of data orthogonal to the standard selection, where one quality cut on the electrons is inverted. This selection results in a multijet-rich sample, from which a template of the multijet background can be obtained.

The fit, containing the simulated templates and the anti-electron template, is then applied to the data in the sideband region. The relative fractions of the other components can be extrapolated to the signal region using the shapes of the template in this range. Consequently, an estimate for the multijet component in the signal region follows from this. Since the fit is applied in different jet bins, it supplies an estimate for the multijet content as a function of the number of jets in the event. In this way the normalization of multijet events can be established. The shapes of other distributions can be extracted from the anti-electron sample, and are scaled to the normalization obtained in the fit.

As of 2011, a matrix method similar to the one described for the muons is deployed to assess the fraction of background in the electron channel. Loose and tight electrons are defined, where tight is the default electron and loose electrons are less stringent quality cuts and relaxed isolation criteria. The number of fake leptons is subsequently obtained from the efficiencies to go from loose to tight in combination with the values for loose and tight electrons as measured in data.

### 4.4 $W$+jets background normalization

In this section we address the normalization of the yield of $W$+jets. Ideally, the contribution of $W$+jets is obtained from data as there are substantial theoretical uncertainties on the normalization. As this is practically difficult, we make use of simulation for $W$+jets, especially for the shapes of kinematic observables, and try to reduce the dependence on the yield. The treatment of $W$+jets differs between the 2010 and 2011 analyses.

In the cross section measurement conducted with the 2010 data, we perform a template fit that returns an estimate for the total amount of background. Hence, we do not make use of the yield of $W$+jets events in the signal region. However, the yields of $W$+jets samples are used in supporting calculations that constrain the final result, for example for ratios of cross sections in bins with low jet multiplicities. Systematic uncertainties are assigned where the simulation is used.

For 2011 data, the number of $W$+jets events in the signal region is obtained from data-driven methods. It makes use of the property that $W^+$ events are produced in larger amounts than $W^-$ events in the LHC. This is a direct result of the abundance of $u$-quarks in the proton-proton collisions, compared to $d$-quarks. The ratio of $W^+/W^-$ events is theoretically more stable than the cross section itself [33]. The charge of the $W$ boson
is contained in the charge of the lepton that we measure. The total number of $W$+jets events after all selection can be expressed as:

\[ N_W(data) = \frac{r_{\text{sim}} + 1}{r_{\text{sim}} - 1} \left( N^+_\text{total}(data) - N^-_{\text{total}(data)} \right) . \]

The ratio $r_{\text{sim}}$ is the ratio of $W^+/W^-$ events in simulation, $N^+(N^-)$ is the total amount of data events with a lepton of positive (negative) charge. A small correction is applied to $N^+$ and $(N^-)$: single top and diboson events also have charge asymmetric production mechanisms at the LHC, but are produced in much lower amounts than $W$+jets. Nevertheless, we correct the data by subtracting the estimated number of single top and diboson events in that set. The ratio $r_{\text{sim}}$ is thus obtained from a sample of simulated $W$+jets that passes all selection requirements, and is measured to be $r_{\text{sim}} = 1.58 \pm 0.08$ in the electron channel and $1.72 \pm 0.07$ in the muon channel [70]. The sources of systematic uncertainty that this fraction is sensitive to are the lepton scale factors and the jet energy scale uncertainty.

Inserting the measured amount of data events (corrected for single top and diboson contamination) results in $N_W = 8244 \,(13374)$ events in the electron (muon) channel, corresponding to scale factors of 1.05 and 0.77 respectively. We assign a relative uncertainty of 15% and 12% to these numbers, obtained by propagating the uncertainty on data and on $r_{\text{sim}}$ to the number of $W$+jets events.

In the plots that will follow in this chapter, the uncertainty on the data-driven background estimates is included as a dashed area. For 2011 plots this includes, besides the multijet background, the uncertainty of obtaining $W$+jets from data.

### 4.5 Basic event selection

In this section we present our event selection for the 2010 and 2011 data sets\(^2\).

We summarize the offline selection criteria.

- **Good vertex.** At least one vertex with more than four tracks should be present in the event.

- **One lepton.** We require exactly one lepton (electron or muon), with $E_T/p_T > 20$ GeV, with $|\eta| < 2.5$. For 2011 data, we change the threshold for electrons to 25 GeV. The muon momentum and rapidity cuts are equal in 2010 and 2011.

- **Four or more jets.** An event needs to have at least four jets with $p_T$ above 25 GeV, with $|\eta| < 2.5$ for 2010 data and $|\eta| < 4.5$ for 2011 data.

\(^2\)Only a few exceptions in terms of the electron and jet properties have changed between 2010 and 2011, due to event trigger evolution and improvements in understanding of the detector. In the cross section analysis and in the charge asymmetry analysis we deviate from the basic selection, to include more events in control regions or purify the sample, but the basic strategy to select top quark events is similar.
Chapter 4. Selection and characteristics of data

- **Missing transverse energy and transverse W boson mass.** In the $e$+jets ($\mu$+jets) channel, we require $E_T \geq 35 \ (20)$ GeV. This cut is identical in 2010 and 2011 selection and supported by studies on the fake or non-prompt leptons in multijet background, that are of different origin in the two channels. In the electron channel, more multijet background is expected and stricter cuts are set to enhance the signal-to-background ratio. Besides $E_T$, the selection is also based on the $m_W$ observable, the transverse mass of the $W$ boson obtained from the lepton and $E_T$ components (we will present the exact definition in Section 4.6.4). In the $e$+jets channel, a threshold on $m_W$ is set to $m_W \geq 25$ GeV. To reduce background in the $\mu$+jets channel, a triangular cut is applied, $E_T + m_W \geq 60$ GeV, cutting away the majority of multijet events [69].

4.5.1 Event yields

The selection cuts are applied to the events in data and to the signal and background simulation, scaled to the estimated luminosities, 35.3 pb$^{-1}$ and 1.04 fb$^{-1}$, respectively. For the data set of 2010, the resulting numbers of events are listed in Table 4.3. The uncertainties listed reflect the statistical uncertainty only, except for the data-driven multijet estimate to which a 50% systematic uncertainty is assigned.

The number of signal events in the electron channel is estimated to be 190, the major background is $W+$ jets with 174 events. The total number of events in data is 397, which is in agreement with the total expected number of events. Due to the different selection criteria, the muon channel has a larger event yield: almost 50% more signal events are expected compared to the electron channel. The contribution of $W+$ jets is also relatively larger. The observed number of events, 646, is within the range that is expected from the combined signal and background estimates.

Note that we discuss the event yields of signal and backgrounds in the data set of 2010 although the measured number of $t\bar{t}$ events is the observable that is measured in the next chapter. The approach in that analysis uses more data-driven input and fits the contribution of signal and background to the measured number of events.

The equivalent table for 2011 data is shown in Table 4.4. The difference is that since the $W+$ jets estimation is data-driven now, we include the systematic uncertainty in the table, like we do for the multijet background. The contribution to the uncertainty that originates from the limited number of events in this background is shown in brackets for comparison to Table 4.3. The number of signal events is 5603 (7960) in the electron (muon) channel. The relative contribution of background increases, as a result of the relaxation of the threshold on the jet pseudorapidity, from $|\eta| < 2.5$ to $|\eta| < 4.5$. Top quark pair events are expected to produce more central activity, compared to the background that results from electroweak processes, $W+$ jets single top and $Z+$ jets and to the multijet background. The result of the increased $\eta$ range is that the purity of $t\bar{t}$ events reduces, but we allow this since the analysis that will be performed on this data set is mostly limited by the number of signal events and not on the systematical uncertainty on background estimates that is
4.6 Distributions in data: 2010 versus 2011

As we saw above, the event yields in data and simulation are compatible. In this section we will evaluate the distributions of reconstructed kinematic variables of selected events, and compare the 2010 and 2011 data sets.

4.6.1 Electron properties

The momentum of the electrons are shown in Figure 4.2, on a logarithmic scale. The momentum runs from 20 GeV for 2010 data (left) and from 25 GeV for 2011 data. It contains the cumulative contributions of signal ($tt$), $W$+jets, multijet background and the sum of all other backgrounds. The latter contains diboson, single top and $Z$+jets components. The hashed blocks reflect the total statistical uncertainty plus the systematic uncertainty as a result of the data-driven multijet approach. Finally, the data is plotted on top of the sum of all expected contributions. The shapes of the background events are comparable. Multijet events produce electrons of low momentum, mostly. For both data sets the expected distribution for the sum of signal and background matches the data well.
Table 4.4 – Expected and observed event yields in 2011 data. The numbers in brackets show the contribution of the statistical uncertainty on the W+jets background.

<table>
<thead>
<tr>
<th>Components</th>
<th>e+jets</th>
<th>µ+jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t} )</td>
<td>5603 ± 40</td>
<td>7960 ± 47</td>
</tr>
<tr>
<td>W+jets (data-driven)</td>
<td>8244 ± 1237 [90]</td>
<td>13374 ± 1605 [98]</td>
</tr>
<tr>
<td>Z+jets</td>
<td>723 ± 15</td>
<td>1262 ± 20</td>
</tr>
<tr>
<td>WW/ZZ/WZ</td>
<td>124 ± 4</td>
<td>192 ± 5</td>
</tr>
<tr>
<td>Single top</td>
<td>448 ± 7</td>
<td>609 ± 8</td>
</tr>
<tr>
<td>Multijet (data-driven)</td>
<td>1159 ± 579</td>
<td>2198 ± 1099</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum Backgrounds</td>
<td>10701 ± 1366 [587]</td>
<td>17637 ± 1945 [1104]</td>
</tr>
<tr>
<td>Total Expected</td>
<td>16304 ± 1366 [588]</td>
<td>25597 ± 1946 [1105]</td>
</tr>
<tr>
<td>Observed</td>
<td>16182 ± 127</td>
<td>26741 ± 164</td>
</tr>
</tbody>
</table>

Figure 4.2 – Electron momentum for data in 2010 (left) and in 2011 (right), after all cuts. The last bin shows the sum of the values beyond the plotted range.

The angular properties of the electrons in the sample are shown in Figure 4.3. The rapidity of the electrons (top) shows the acceptance gap at the crack region at \( 1.37 < |\eta_{\text{cluster}}| < 1.52 \), a detector region where no active material is present. (In 2010 data there is no visible empty bin, as the binning is not fine enough to show it.) Furthermore, the signal contribution is larger at the central region, as the \( t\bar{t} \) pair is expected to be produced
more centrally. A difference between the multijet estimates in 2010 and 2011 is visible in the high-\(\eta\) regions. This is due to the change from the anti-electron selection to the matrix method for obtaining the multijet estimate. The difference is not likely to affect our measurements in the following chapters, since we already assign large uncertainties to this contributions. The data is compatible with the expected shape.

The \(\phi\)-distributions, shown in 4.3, show more irregular behavior. In principal, the production of events is symmetric and electrons are expected to be produced isotropically in \(\phi\), resulting in a flat distribution. Due to acceptance irregularities, small deviations can arise. The liquid argon hardware problem in 2011, described earlier, does have a strong impact on this distribution. During part of the data taking period, the region \(\phi = (-1.2, -0.5)\) and \(\eta \geq 0\) might be affected by mismeasurements. The data events recorded during this period that contain jets and electrons in this area are removed from the selection, to avoid selecting unreliable events. Since there are some runs recorded from before the problem occurred, a small number of data points populate the gap. The simulation is corrected for this problem.

4.6.2 Muon properties

For the muons in the selected sample (in the muon channel), we examine the same distributions for 2010 and 2011 data. Figure 4.4 shows the momentum of muons for both periods. The momentum threshold as well as the angular requirements for muons to pass the selection are unaltered with respect to the previous chapter. One exception is added to correct for a software bug in 2011: scale factors for muons with a \(p_T\) larger than 150 GeV were not well defined. Therefore all events that contain muons above this threshold are removed\(^3\). There is no prominent difference between the distribution observed in data, and the expectation.

The angular distributions of data and simulation of the muons, shown in Figure 4.5, mostly agree within their uncertainty. The shape of \(\eta\) is different with respect to electrons. At \(\eta = 0\), the muons cannot be reconstructed as there is no detector material at that point. The feet of ATLAS, the support structure holding the detector, also produce areas that are insensitive to muon detection, at \(\phi = -2.0\) and \(\phi = -1.1\), for different values of \(\eta\). This is visible in both \(\phi\)-plots. It is, just as other non-flat behavior in the shape, well described in simulation. One data point in the \(\eta\) distribution lies significantly above the expectations (at \(\eta = 0.25\)), but we cannot associate this to a specific detector or physical effect. With regard to the overall match of the data and expectation in this distribution, we therefore assume it to be a fluctuation.

4.6.3 Jet properties

Jets are studied separately for the electron and muon channel, although the properties of jets should not depend on the flavor of the lepton in the event. We therefore show only

\(^3\)In later releases the software bug was corrected, restoring the full muon spectrum in the data set.
Chapter 4. Selection and characteristics of data

![Graphs showing event distributions for 2010 and 2011 data sets.](image)

**Figure 4.3** – Electron polar and azimuthal angles, for 2010 (left) and 2011 data (right), after all cuts.

The foremost change concerning jet requirements in 2011 is the extension of the allowed pseudorapidity range for jets from $|\eta| < 2.5$ to $|\eta| < 4.5$. This is visible in the $\eta$-distribution of the jet with the highest momentum in the event, shown in Figure 4.8.
4.6. Distributions in data: 2010 versus 2011

![Figure 4.4](image_url) – Muon momentum for data in 2010 (a) and 2011 (b), after all cuts. The last bin of (a) shows the sum of the values beyond the plotted range.

As mentioned before, $t\bar{t}$ produces more central jets, on average. The $\phi$-angle shows no irregularities, except for the bump in 2011 resulting from the problem in the calorimeter.

### 4.6.4 Event variables

Besides the leptons and jets, we are interested in event-wide variables, and especially the total transverse energy vector and consequently the missing transverse energy. The missing transverse energy is depicted in Figure 4.9. The magnitude of $\vec{E}_T$ in data is in agreement with simulation for both time periods. This is also true for the electron channel (not shown).

Contrary to $\eta$, the $\phi$-component of $\vec{E}_T$ is available, since it includes information only from the $x$- and $y$-components. The difference between the muon and $\vec{E}_T$, projected on the transverse plane, $\Delta\phi$, is shown in the same figure as well. It contains information on the angle between the neutrino and the lepton that result from a $W$ boson. The range of possible angles is large, but the shape of the $t\bar{t}$ contribution is somewhat different than the backgrounds: around $\Delta\phi = 1$, the contribution of $t\bar{t}$ seems relatively large. The $W$+jets background, for example, shows a steady slope with respect to this variable, with the most likely value being a back-to-back configuration for the lepton-neutrino pair. The $W$ boson in top quark decay is more likely to be boosted with respect to the $W$ boson in $W$+jets production, leading to slightly smaller angles between the lepton and the neutrino on average. The prediction overestimates the number of events in data for high values of $\Delta\phi$ (above $\Delta\phi = 1$), but the distributions are still compatible.

One important variable (used in the selection criteria) is the transverse $W$ boson mass ($m^W_T$), that is constructed from lepton and $\vec{E}_T$ information. When selecting events in the
single-lepton decay channel of $t\bar{t}$, one of the $W$ bosons is required to decay into a neutrino and a lepton. Whilst the lepton is actually measured by the detector, the kinematical properties of the neutrino are best estimated using $E_T$. As there is no $z$-component to $E_T$, a full mass reconstruction of this $W$ boson is not possible. The mass in the transverse plane, however, can be computed and is expressed as

$$M_W^T = \sqrt{(p_T + E_T)^2 - (p_x + E_x)^2 - (p_y + E_y)^2},$$

where the lepton and missing transverse momentum in the $x$ and $y$ direction are combined.

The distribution of $M_W^T$ peaks just below the $W$ boson mass for physics processes that indeed contain $W$ bosons. Figure 4.10 shows the reconstructed transverse mass, after all cuts. The signal ($t\bar{t}$) and $W$+jets background both peak at the expected position.
4.6. Distributions in data: 2010 versus 2011

**Figure 4.6** – Jet multiplicity in 2010 (a) and 2011 (b), after all other cuts. The jets need to have $p_T > 25$ GeV. The last bin shows the sum of the values beyond the plotted range.

**Figure 4.7** – Transverse momentum of the hardest jet in the event for 2010 (a) and 2011 (b), after all cuts. The last bin shows the sum of the values beyond the plotted range.

around 80 GeV. A genuine $W$ boson decaying into lepton-neutrino pair is present in $t\bar{t}$ and $W$+jets. Due to the differently applied cuts in the electron and muon channel, the shape is dissimilar between the two. In the electron channel (top plots) a straight cut at 25 GeV is imposed to cut away the majority of multijet background events. In the muon channel it suffices to apply only the triangular cut ($E_T + m_W^T \geq 60$ GeV), resulting in a different shape on the lower side of the spectrum. The muon channel has a residual slope that originates from the exponentially decreasing value of $m_W^T$ for multijet events, whereas in the electron channel, this is cut off. The agreement in both channels gives confidence
in the assumption that the multijet background is well estimated, also in terms of the shape of $m_T^W$.

### 4.6.5 Identification of $b$-jets

The identification of jets that originate from $b$-quarks is an important feature in our analyses, as it is a good differentiator between signal and background. In this section we will present the procedure of tagging $b$-jets. There are several methods to distinguish jets coming from $b$-quarks from light quark jets. We use the SV0 algorithm [71] to do the identification. The SV0 algorithm takes advantage of the fact that a $b$-quark hadronizes to form a $B$ hadron. The $B$ hadron has an average lifetime of about 1.5 ps, see for example [72], which implies it has a mean decay length $c\tau$ of 450 $\mu$m, before it decays.
Figure 4.9 – (a,b) Missing transverse energy, after selection for 2010 (left) and 2011 (right). The last bin shows the sum of the values beyond the plotted range. (c,d) Difference between the φ component of $E_T$ and the lepton.

into lighter particles. The boosted state in which a $B$ hadron is produced enhances the distance it travels to millimeter scale. At this scale it becomes possible to detect a 'secondary vertex’ in the collision, defined as a vertex with a distance $L$ from the primary vertex.

The reconstruction of a secondary vertex starts with selecting tracks that are matched to a jet; i.e., tracks that are within a distance $\Delta R$ from the axis of the jet. Quality cuts, like a minimum number of hits in the inner detector and a momentum threshold of 0.5 GeV, are applied to the tracks. Tracks with a transverse or longitudinal distance to the primary vertex larger than 2 mm are disregarded since they are too far away and are therefore likely to be misreconstructed or to originate from photon conversions. The remaining
tracks are input to the SV0 algorithm.

The algorithm reconstructs vertices from all remaining track pairs in the jet and selects those with a significant displacement from the primary vertex. The vertex mass is required to be inconsistent with \( K^0_s \) and \( \Lambda^0 \) decays and photon conversions. Finally, the vertex position is not allowed to coincide with material in the pixel detector.

In the final step one single secondary vertex is fitted to all of the two-track vertices. Iteratively, tracks that contribute to a low goodness-of-fit \( \chi^2 \) value are removed, until this value reaches a threshold value. The measure for the probability of a jet to originate from a \( b \)-quark is then expressed in the decay length significance, \( L/\sigma(L) \), where \( L \) is the distance to the primary vertex and \( \sigma(L) \) the uncertainty on this value. Figure 4.11 (left) shows the distribution of this weight variable for different source of jets; coming from light
quarks, $c$-quarks, tau leptons and $b$-quarks. The vector from the primary to the secondary vertex is projected on to the jet axis, allowing the weight to take negative values for cases where the secondary vertex is on the opposite side of the jet with respect to the primary vertex. Negative weights do occur, but indicate an unlikely $b$-jet candidate and result from the limited resolution. We consider a jet $b$-tagged if the SV0 algorithm assigns a weight larger than 5.85 to it (indicated by the line with arrow). The contamination from other sources (fake $b$-jets) in the form of $c$-jets is one order of magnitude lower, from the $\tau$ lepton, light quark, and gluon jets it is minimal. The working point is chosen based on optimization between efficiency, purity and light quark rejection.

![Figure 4.11](image.png)

**Figure 4.11** – (a) Distribution of SV0 weight for the different contributions of jet sources. (b) Efficiency curve for SV0 weight and rejection factor for non-$b$ sources.

The right plots in Figure 4.11 display the efficiency and rejection factor, as a function of the cut-off value of the SV0 weight. The weight cut at 5.85 corresponds to an efficiency of 50% calculated with simulations of signal from MC@NLO. The efficiency indicates that we can expect roughly only $\sim$25% of the signal events to yield two $b$-tagged jets. The rejection factor, which is the inverse of the mistag rate, at this point is 271 for light quarks, 9 for charm quarks and 38 for tau leptons.

As a result about 20-25% of the events with two $b$-quarks are recognized as such. About 50% of the two $b$-quark events end up to have only one $b$-tagged jet. The number of $b$-tagged jets in the selection for top quark events is plotted in Figure 4.12, and proves this statement. Signal events are mostly present in events with tagged jets, but still a significant part has zero $b$-tags found. Since in the bins with $b$-tagged jets the $t\bar{t}$ contribution is dominant, it proves that $b$-tagging is useful for differentiating top quarks from the background.
Figure 4.12 – Number of $b$-tagged jets in selected events in 2010 and 2011.

4.6.6 $W+\text{jets}$ background normalization after $b$-tagging

The amount of $W+\text{jets}$ events before $b$-tagging in 2011 data was estimated with a data-driven approach, as discussed in Section 4.4. The normalization after requiring at least one $b$-tagged jet, on top of the applied selection criteria, $N_{W}^{\geq 4,\text{tagged}}$, is obtained by correcting the total amount of events before $b$-tagging with correction factors obtained partly from data and partly from simulation in events with two jets. It follows the scheme in [73].

\[
N_{W}^{\geq 4,\text{tagged}} = N_{W}^{\geq 4}(\text{data}) \cdot f_{2,\text{tagged}}^{\text{data}} \cdot k_{2\rightarrow\geq4}^{\text{sim.}}.
\]

Here, $N_{W}^{\geq 4}$ is the amount of $W+\text{jets}$ events after all selection requirements, but before $b$-tagging. The factor $f_{2,\text{tagged}}^{\text{data}}$ is the fraction of $W+2$ jet events that have at least one $b$-tagged jet, and is obtained from data using the same charge asymmetry property as used for events with four or more jets. Finally, $k_{2\rightarrow\geq4}^{\text{sim.}}$ represents the ratio $f_{4,\text{tagged}}^{\text{sim.}} / f_{2,\text{tagged}}^{\text{data}}$ in our simulated samples.

For our selection, we obtain $f_{2,\text{tagged}}^{\text{data}} = 0.036 \pm 0.005 (0.043 \pm 0.005)$ in the electron (muon) channel. The correction ratio is evaluated to be $k_{2\rightarrow\geq4}^{\text{sim.}} = 1.91 \pm 0.21$ and $2.04 \pm 0.23$, for the two channels respectively. The total number of $W+\text{jets}$ event then amounts to $N_{W}^{\geq 4,\text{tagged}} = 567 (1165)$ in the electron (muon) channel, to which we conservatively assign a 30% uncertainty.

4.7 Top quark reconstruction

One straightforward algorithm for the reconstruction of top quarks from the hadronic decay products, $\sum p_{T}$ method, has already been discussed in Section 3.5. Another method we utilize is the ‘Kinematic Likelihood Fitter’ (KL-fitter) [74], which attempts to reconstruct both top quarks in the event. In the following section we present the performance of the two reconstruction algorithms.
4.7. Top quark reconstruction

4.7.1 Reconstruction with $\sum p_T$ method

In the $\sum p_T$ reconstruction algorithm the combination of three jets, from all possible 3-jet combinations, that together form the vector with the largest transverse momentum is considered the top quark candidate on the hadronic side. The jets in the range $20 < p_T < 25$ GeV are also used in this combination. The aforementioned additional requirement of having a 2-jet combination within the 3-jet combination that has a transverse mass compatible with the $W$ boson mass, we ignore for now. There is no distinction made between $b$-tagged and untagged jets in this procedure.

The reconstructed mass of the top quark is shown in Figure 4.13. There is a mass peak visible, on top of the background. The background is formed by the physics background processes introduced, but also by combinatorial background. Combinatorial background originates from cases where one or more of the chosen three jets do not originate from the top quark.

![Figure 4.13](image)

(a)

(b)

Figure 4.13 – (a, b) Reconstructed invariant mass of the top quark ($\sum p_T$), for 2010 and 2011.

We plot the invariant mass of the three 2-jet combinations that are possible of the three jets used for the top quark candidate in Figure 4.14. We sort the jets by their transverse momentum, jet 1 is the highest. Following this convention $m_{j1,j2}$ is the 2-jet combination obtained from the jets with the highest $p_T$, within the chosen three jets. And similarly we have $m_{j2,j3}$ and $m_{j1,j3}$. In all cases the $W$ boson mass peak is visible around 80 GeV, albeit of different height. First of all, there is combinatorial background present in these plots. By construction, the $W$ boson is reconstructed properly only in one of the three options. Secondly, the $b$-quarks in the top quark decay chain are expected to be of higher transverse momentum, on average. The $b$-quark is much lighter than the $W$ boson, in the rest frame of the top quark this deficit is compensated by a boosted $b$-quark. The result is that the likelihood of the combination $m_{j2,j3}$ being the right one, given that the top quark

\footnote{The algorithm was originally aimed at first data, under the assumption that $b$-tagging would not be sufficiently calibrated at the time LHC collisions started.}
was well reconstructed is the largest. This distribution also shows the clearest peak. The data matches this behavior reasonably well.

Figure 4.14 – (a,b) Reconstructed invariant mass of the $W$ boson candidate from (a) jet 1 and 2, (b), jet 1 and 3, (c) jet 2 and 3 (sorted in terms of transverse momentum).

In Section 3.5, we additionally applied a $W$ boson mass constraint, where events only pass if the invariant mass of one of the three $W$ boson candidates is close enough to 80.4 GeV. This was applied to 14 TeV collisions, but the procedure has also been used for simulated collision events of a CM-energy of 10 TeV in [75]. The results lead to a more distinct mass peak that could work as a measure for the number of well reconstructed top quarks and consequently a cross section measurement. A disadvantage is that it leads to
a reduction of signal events. We apply the $W$ boson mass constraint to the reconstructed top quark events in our data at 7 TeV, to investigate the feasibility of the algorithm at this CM-energy. Figure 4.15(a) shows all $W$ boson candidates (three candidates per event). We reconstruct the top quark only in case at least one jet pair has an invariant mass in a window of 20 GeV around 80 GeV, the limits set in the original paper. This leads to a 55% reduction of signal events in simulation. The distribution of the top quark mass under these conditions is shown in Figure 4.15(b). A shoulder is formed as a result of the $W$ boson mass constraint, in the region of 100 GeV, compared to Figure 4.13. The signal shows a broad ‘peak’ in the region 150-160 GeV, revealing the top quark as an excess on top of the combinatorial background in the mass spectrum. The low selection efficiency of this method, however, is problematic. Extending the mass window to a 30 or 40 GeV range around the central value leads to more selected events, but less pronounced peaks. This makes the extraction of the cross section via fits to the top quark mass peak difficult.

In conclusion, the method of reconstructing the mass on one side of the decay proves to work by producing a mass peak at the expected value. Adding the $W$ boson mass constraint enhances the peak, but at the cost of efficiency. For the cross section measurement a higher efficiency and a more pronounced mass peak are preferred. Using information from both sides of the decay in a more advanced algorithm will improve the distribution.

![Figure 4.15](a) Reconstructed invariant mass of all $W$ boson candidates (three per event). The dashed line indicates a window of 20 GeV around the expected mass, outside which events can be excluded. (b) Reconstructed invariant mass of the top quark, for events with one of the three $W$ boson candidates within a window of 20 GeV around the expected mass.

### 4.7.2 Reconstruction with Kinematic Likelihood Fitter

The Kinematic Likelihood Fitter package (KL-Fitter) is used to obtain a full reconstruction of semi-leptonic $t\bar{t}$ events using a maximum likelihood method. The goal is to associate detector objects to partons and leptons of the $t\bar{t}$ decay, in order to use the kinematic
properties of the top quarks to increase the resolution of the detector objects. The kinematic constraints are enforced by varying the energy and angles of the detector objects within their resolution, for each permutation of association of the objects to partons and leptons of the top quark decay. In this way, we assign a likelihood to each permutation. The maximization is performed by the Bayesian Analysis Toolkit [76].

The input to the KL-Fitter algorithm consist of the calibrated energy and angular information of the jets, lepton and \( \not{E}_T \) as measured by the ATLAS detector. Only the four jets with the highest transverse momentum are selected. The assumption is that these four jets come from the hard scattering process and originate from either the two b-quarks or the two light quarks in the top quark decay. Table 4.5 lists the 17 observables that are used. For jets and leptons the angles \( \tilde{\Omega}_i = (\tilde{\eta}_i, \tilde{\phi}_i) \) and energy \( \tilde{E} \) are input. Secondly, we use the measured \( x \)- and \( y \)-component of the missing transverse energy, \( \tilde{E}_x \) and \( \tilde{E}_y \).

<table>
<thead>
<tr>
<th>Detector object</th>
<th>Observables (.snapshot)</th>
<th>(#)</th>
<th>Particles</th>
<th>Fit parameters (snapshot)</th>
<th>(#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>jets (( \times 4 ))</td>
<td>((\tilde{\eta}_i, \tilde{\phi}_i))</td>
<td>(12)</td>
<td>light quarks (( \times 2 ))</td>
<td>(E, \eta, \phi)</td>
<td>(6)</td>
</tr>
<tr>
<td>lepton</td>
<td>((\tilde{\eta}_l, \tilde{\phi}_l), \tilde{E}_l)</td>
<td>(3)</td>
<td>(b)-quarks (( \times 2 ))</td>
<td>(E, \eta, \phi)</td>
<td>(6)</td>
</tr>
<tr>
<td>( \not{E}_T )</td>
<td>(\tilde{E}_x, \tilde{E}_y)</td>
<td>(2)</td>
<td>lepton</td>
<td>(E)</td>
<td>(1)</td>
</tr>
<tr>
<td>neutrino</td>
<td>(p_x, p_y, p_z)</td>
<td>(3)</td>
<td>top quark</td>
<td>(\hat{M}_t)</td>
<td>(1)</td>
</tr>
</tbody>
</table>

The likelihood is built up out of two elements: transfer functions and mass constraints.

**Transfer functions**

The transfer functions represent conditional probabilities. They define the probability of measuring the value for an observable quantity at detector level given its true value at parton level. In this way the resolution of the measurement is taken into account. The transfer functions for the angles \( W(\tilde{\Omega}_i | \Omega_i) \) are defined as a Gaussian, with \( \Omega_i \) representing the true angular coordinates \((\eta_i, \phi_i)\) and \( \tilde{\Omega}_i \) the observable values of these quantities. The Gaussian distribution accommodates the resolution on these variables. The angular transfer functions are only applied to transfer from quarks to jets; the lepton angular resolution is neglected and is approximated by a delta function.

The functions to transfer the energy from parton to detector level look somewhat different. The resolution effects on the lepton and jet energies cannot be represented by a plain Gaussian distribution due to detector losses of several kinds, calibration effects and initial and final state radiation. To accommodate this behavior, the transfer function for energy is parametrized as the sum of two Gaussian distributions:
4.7. Top quark reconstruction

\[ W(\tilde{E}_i | E_i) = \frac{1}{2\pi(c_2 + c_3c_5)} \left( e^{-\frac{(\Delta E - c_1)^2}{2c_2^2}} + c_3 e^{-\frac{(\Delta E - c_4)^2}{2c_5^2}} \right) , \]

with \( \tilde{E}_i \) and \( E_i \) being the measured and true value respectively, and \( \Delta E \) the difference between them. The five parameters \( c_i \) are determined from simulation for different bins of rapidity and depend on the true energy. Figure 4.16 shows an example of the fit that parametrizes the resolution (obtained from [77]), in an arbitrary rapidity-momentum range. The extra Gaussian (component 2) accommodates the tail effects of the resolution distribution.

![Figure 4.16](image)

**Figure 4.16** – Example of the energy resolution for light jets with \( 0.8 < |\eta| < 1.37 \) and \( 205 \text{ GeV} < p_T < 235 \text{ GeV} \).

For the missing transverse energy (\( \not{E}_T \)) we define two transfer functions, \( W(\tilde{E}_x | E^x_{\nu}) \) and \( W(\tilde{E}_y | E^y_{\nu}) \). In these functions, the \( x \)- and \( y \)-components of the parton momentum are mapped to the measured missing energy components.

**Implementations of the mass constraints**

The other terms in the likelihood function constrain the combined masses formed by the jets, leptons and \( \not{E}_T \) objects, with Breit-Wigner functions. The mass of the combination of the two jets assigned as light quarks (\( m_{jj} \)) is expected to be close to the \( W \) boson mass of 80.4 GeV. In the likelihood function it is constrained with a Breit-Wigner function around this value, with a width \( \Gamma_W \) of 2.1 GeV. In the likelihood this is expressed by the term \( f_{BW}(m_{jj}|M_W, \Gamma_W) \). The combined mass of the charged lepton and the neutrino, \( m_{l\nu} \), is restricted similarly to the \( W \) boson mass, \( f_{BW}(m_{l\nu}|M_W, \Gamma_W) \).
Chapter 4. Selection and characteristics of data

The masses of the hadronic and leptonic top quarks are required to be equal, but the absolute value of the pole mass is kept unconstrained. The top quark pole mass is a free parameter ($\hat{M}_t$), but we do impose a fixed width for $\Gamma_t$ of 1.5 GeV. The unconstrained mass value in the fit prevents a bias towards the expected value in the mass distribution of background events. The top mass constraints in the likelihood function are expressed as $f_{BW}(m_{\nu j}|\hat{M}_t, \Gamma_t)$ and $f_{BW}(m_{jj}|\hat{M}_t, \Gamma_t)$.

The likelihood is written as:

$$L(p_1, ..., p_{17}) = \prod_{i=1}^{4} (W(\tilde{E}_i|E_i) \cdot W(\tilde{\Omega}_i|\Omega_i)) \cdot W(\tilde{E}_l|E_l) \cdot W(\tilde{E}_x|E_x) \cdot W(\tilde{E}_y|E_y) \cdot f_{BW}(m_{jj}|M_W, \Gamma_W) \cdot f_{BW}(m_{\nu j}|\hat{M}_t, \Gamma_t) \cdot \delta_b,$$

where the index $i$ runs over the four jets. The delta function $\delta_b$ prevents $b$-tagged jets to be assigned to the position of a light quark: a zero probability is assigned to that permutation. The fit parameters are listed in Table 4.5.

The negative log of the likelihood is minimized, considering all permutations of parton-jet assignments. The reconstructed mass distributions for the best permutation are shown in Figure 4.17. It is important to notice that the constraint on the difference between the two top quark masses does not bias the background shape.

![KL-Fitter reconstructed mass for simulation and data, for the leptonically (a) and hadronically decaying top quark (b).](image)

**Figure 4.17** – KL-Fitter reconstructed mass for simulation and data, for the leptonically (a) and hadronically decaying top quark (b).
4.8 Comparison of the KL-fitter and $\sum p_T$ algorithms

We compare the properties of the top quarks on the hadronic side with both reconstruction algorithms, to each other and to the true properties as simulated by the generator. We apply an extra cut, with respect to the basic selection: at least one jet should be tagged as a $b$-jet. As we have seen, this leads to a reasonably pure sample of signal events. (Note that the $b$-tag veto in the likelihood of the KL-fitter only leads to a permutation that is not allowed, it does not discard events.) Figure 4.18 shows the reconstructed invariant mass of the top quarks on the hadronic side, for both reconstruction algorithms, after requiring at least one $b$-tagged jet. First of all, the reconstruction with the two algorithms results in reasonably different distributions of the top quark mass. The kinematic fit is superior in terms of width, showing a narrower peak. The KL-fitter mass also has a higher turn-on point, the distribution starts above 90 GeV. This is in contrast to the mass obtained with the $\sum p_T$ method, which can take values as low as 40 GeV. In all cases the data matches with the expectations.

![Diagram](a)

![Diagram](b)

![Diagram](c)

![Diagram](d)

**Figure 4.18** – Hadronic top quark mass in 2010 and 2011 data, reconstructed with the $\sum p_T$ method (a,c) and KL-Fitter (b,d).

We compare the two algorithms in simulation at reconstruction level with the simulation...
at parton level, to show the difference in performance. At the parton level, we take the top quark at the point where it is created in the hard scatter in simulation. In the simulation, the top quark mass is set to a fixed value, 172.5 GeV, making a straight comparison of reconstructed and simulated masses impossible. We therefore examine the momentum and angular distributions. We define the difference in momentum of the top quark as \( \Delta p_T = p_T(\text{reco}) - p_T(\text{parton}) \), where \( p_T(\text{reco}) \) is the \( p_T \) as reconstructed with one of the algorithms. This quantity is shown in Figure 4.19. The KL-fitter algorithm shows a distribution of \( \Delta p_T \) centered around zero, with tails to both sides. The tail towards the positive side is slightly higher, indicating that the reconstruction has a tendency to overestimate the transverse momentum more often than the other way around. The \( \sum p_T \) algorithm is much more biased, the distribution of \( \Delta p_T \) peaks around 50 GeV. This is expected, as the algorithm explicitly maximizes the momentum. The equivalent variables for the angles, \( \Delta \eta \) and \( \Delta \phi \) are displayed as well, and show equivalent spectra. For both reconstruction algorithms, there is a peak at 0, with symmetric tails. The KL-fitter is superior to a small extent. The spatial difference between the reconstructed top quark at reconstruction and parton level is expressed in \( \Delta R \). This distribution shows that a large part of the events has a top quark reconstructed in the vicinity of the original top quark. The KL-fitter performs better as well, with a larger peak close to zero. It is not possible to deploy an unambiguous matching algorithm to match the reconstructed and parton level quark. Even misreconstructed quarks can coincidentally get very close to the parton level top quark direction. We measure a fraction of 32\% of events within \( \Delta R = 0.4 \) of the original top quark for the KL-fitter algorithm and 26\% for the \( \sum p_T \) procedure. The actual number of well reconstructed events will be somewhat lower, due to the effect of coincidental matches.
4.9. Summary

We introduced the data that will be used in the analyses of the coming chapters. For the cross section chapter this amounts to 35 pb$^{-1}$, recorded in 2010. The charge asymmetry measurement will be performed on a part of the 2011 data set, 1.04 fb$^{-1}$. There are small differences in the conditions of the data in terms of pile-up and detector status, but we apply a common event selection and compare the distribution of the important objects in our analysis. Comparing the data with the expected shape and normalization of simulated and data-driven background, we showed that overall the agreement is very good. The procedure of tagging jets when they are likely to originate from $b$-quarks, $b$-tagging, works very well, and gives a handle to purify the sample and suppress backgrounds. We discussed two top reconstruction algorithms. The first is straightforward and makes an optimal combination of jets to form the top quark on the hadronic side of the decay. The second algorithm is a kinematic fitter that attempts to reconstruct both sides of the decay. It contains input from jets, the lepton and the missing transverse energy. Both algorithms show the invariant mass peak, but the KL-fitter is superior in all terms. It has
the advantage of reconstructing a top quark on the leptonic side of the decay as well, and thus gives information on the complete $t\bar{t}$-system. We make use of the KL-fitter algorithm where possible. We will, however, make use of the $\sum p_T$ algorithm in case of incomplete information, that is, when we take control regions with only three jets available.