First top quark physics with ATLAS : a prospect

van der Kraaij, E.E.

Citation for published version (APA):

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.
Chapter 6

Top quark and $W$ boson production cross section measurement

Most studies look into top quark pair events with at least four jets \[116\], as a semi-leptonic top quark pair decay leads to at least four partons. For the determination of the $t\bar{t}$ cross section we may gain more information by also measuring the number of top events with less than four reconstructed jets, since jets can be missed due to losses through the beam pipe or detector cracks, or due to mis-reconstruction, $p_T$ requirements, etc... In this chapter we discuss such a method, making use of b-tagging algorithms, see Section 4.4. An accurate measurement of the $t\bar{t}$ cross section is important for the searches for new physics beyond the Standard Model, such as supersymmetry.

First, we explain the setup of the analysis and discuss all significant systematic uncertainties. Second, we apply the analysis to a sample with unknown cross sections of the different channels, and to a small sample of events with integrated luminosity of $50 \text{ pb}^{-1}$, to determine the possibilities with the first data to be taken with ATLAS. Finally we discuss the implication of supersymmetric scenarios and their effect on the analysis results.

6.1 The analysis setup

The purpose of this analysis is to measure the contribution of top quark pair and $W$ boson production to events with ‘low’ jet multiplicities, that is events with four or less jets. Defining the $t\bar{t}$ channel as the signal channel, the main backgrounds are QCD multi-jets, $W+$jets, and single top events.\(^1\) With a b-tagger we can perform a counting measurement in which we count the number of events with one, two or three jets tagged. Since the top quark pair events have a different b-jet multiplicity than the background channels, the combination of the number of events tagged one, two or three times contains information on the amount of top quark pair events in the sample. We set as a goal to perform the analysis on first data with $50 \text{ pb}^{-1}$ of integrated luminosity. It is widely assumed that due to imperfect alignment and calibration of the detector the b-tagging efficiency cannot be predicted from MC samples, see \[101\], and thus also needs to be determined.

For the events that contain exactly two jets, which we shall refer to as the two-jet sample, the number of events with one or two jets tagged can be determined. The number of zero tagged 

\(^1\)See Sections 1.1.3 and 1.1.4. QCD multi-jet events are hard scatterings of $pp \rightarrow Xj$, with $X = 2, 3, \ldots$ and are also simply called QCD events.
jets is completely correlated with these two measurements. For events that contain three or four jets we can determine the number of events with one, two or three tags. In total we thus perform eight measurements on our entire sample. There are not many events containing four jets in a data sample of \( L = 50 \text{ pb}^{-1} \), and we therefore also perform this analysis on only the total of the two- and three-jet samples.

With the eight numbers extracted from the data we can solve the five unknowns in our analysis, i.e. the number of \( t\bar{t} \) events in the two-, three- and four-jet samples, and the two different b-tagging efficiencies: the efficiency for b-jets and the efficiency for light quark jets. For this last variable we must be careful as to what we call a light quark jet. In this analysis only the b-jets in top decays are categorized as b-jets. All other jets are categorized as light quark jets. This implies, however, that the light quark jets from \( W+\text{jets} \) events do not have the same composition as the light quark jets from top quark pair events: the latter have a substantial amount of s- and c-quarks coming from the hadronically decaying \( W \). As the jets in the \( W+\text{jets} \) events are all initial state radiation (ISR) and/or final state radiation (FSR), see Section 1.1.4, these mostly originate from gluons or u- and d-quarks. We will measure the b-tagging efficiency on light quark jets in top events, but not in \( W \) events. The small contribution from \( W+b\bar{b} \) events complicates this definition; we come back to this in section 6.1.4.

The standard b-tagging algorithm in ATLAS is the IP3D+SV1, a combined weight method mentioned in section 4.4. For the first data this is not suitable, for the same reason as mentioned before: the detector performance will not exactly be known, and a reliable calibration of the IP3D+SV1 b-tagger will take time. For the first data the JetProb algorithm is recommended, see Section 4.4.3. We have therefore used the IP3D+SV1 b-tagger for our study on the reliability of our analysis, and JetProb to show the feasibility of the analysis on first data.

### 6.1.1 Likelihood functions

To solve the five unknowns, i.e. the b-tagging efficiency on the b-jets (\( \varepsilon_b \)), the b-tagging efficiency on the light quark jets content of the top sample (\( \varepsilon_{\text{top}} \)), and the number of \( t\bar{t} \) events in the two-, three- and four-jet samples, called \( N_{2j}^{\text{top}} \), \( N_{3j}^{\text{top}} \) and \( N_{4j}^{\text{top}} \), we use RooFit [119] for the minimization of a negative log-likelihood (NLL), defined as follows:

\[
-\ln L = -\ln \prod_{x,n} P(N_{nj}^{\text{tag,meas}}, N_{nj}^{\text{tag}}) = -\ln \prod_{x,n} \frac{e^{-N_{nj}^{\text{tag}}}}{N_{nj}^{\text{tag,meas}}} \cdot \left( \frac{N_{nj}^{\text{tag}}}{N_{nj}^{\text{tag,meas}}} \right)^{N_{nj}^{\text{tag,meas}}},
\]

where \( P(N_{nj}^{\text{tag,meas}}, N_{nj}^{\text{tag}}) \) is the Poisson Likelihood of measuring \( N_{nj}^{\text{tag,meas}} \) events with \( x \) jets b-tagged in the \( n \)-jet sample, when \( N_{nj}^{\text{tag}} \) events are expected.

For the expected numbers of tagged jets in the NLL we have to make some assumptions to keep the number of unknowns reasonable. We will explain these assumptions by discussing the expected numbers of tagged jets in the two-jet sample as an example: with \( N_{2j}^{\text{tag}} \) and \( N_{2j}^{\text{tag,meas}} \) denoting the expected numbers of events with one and two jets tagged, we can write:

\[
N_{2j}^{\text{tag}} = \alpha_1 W_{2j} + \alpha_{\text{top}} W_{2j}^{\text{top}}
\]

\[
N_{2j}^{\text{tag,meas}} = \alpha_2 W_{2j} + \alpha_{\text{top}} W_{2j}^{\text{top}}
\]

with

\[
\alpha_1 W = 2\alpha_1 (1 - \alpha_2 W) \]

\[
\alpha_2 = (\alpha_1 W)^2.
\]
In the equations above we assume there are only two ‘types of events’: \(W\) events and top events. The assumption here is that single top events will be \(t\bar{t}\)-like in their b-jet content, hence we sum up the contributions of both the single top and the top quark pair channels in events with exactly two jets to \(N_{2j}^{\text{top}}\). The contribution to the number of \(W\) events with two jets, \(N_{2j}^{W}\), comes of course from the \(W+jets\) events containing two jets. Other background channels, such as QCD multi-jets, might however also contribute to this number. The assumption is that QCD events have few b-jets and will be ‘\(W+jets\)’ like; we note that we do not analyze QCD events in this chapter and we will not check this last assumption.

We previously stated that we define all jets which are not b-jets as light quark jets. In eq. (6.4) and (6.5) the probability for having one or two jets tagged in a \(W+jets\) event, \(e_1^W\) and \(e_2^W\), only depends on one variable \(e_1^W\). This b-tagging efficiency on the light quark jets content of the \(W+jets\) events is not a free variable in the fit and is to be measured elsewhere, see section 6.1.4.

For the contribution from the top events in eq. (6.2) and (6.3) we define the probabilities \(e_1^{\text{top}}\) and \(e_2^{\text{top}}\)

\[
e_1^{\text{top}} = 2e_1^{\text{top}}(1-e_1^{\text{top}})R_0^{b\text{Jet}} + (e_b(1-e_1^{\text{top}}) + e_1^{\text{top}}(1-e_b)) R_1^{b\text{Jet}} + 2e_b(1-e_b)R_2^{b\text{Jet}}
\]

\[
e_2^{\text{top}} = (e_1^{\text{top}})^2 R_0^{b\text{Jet}} + e_b e_1^{\text{top}} R_1^{b\text{Jet}} + e_b^2 R_2^{b\text{Jet}}.
\]

The b-tagging efficiency on the light quark jets content of the top sample \(e_1^{\text{top}}\) is a free variable in the fit, together with the b-tagging efficiency on b-jets, \(e_b\). Due to c-jets in the top sample, \(e_1^{\text{top}} \neq e_1^{W}\). A two-jet event can contain zero, one or two reconstructed b-jets, even before the b-tagging procedure. The probabilities for these three scenarios are given by \(R_i^{b\text{Jet}}\), with \(i = 0, 1, 2\). These three probabilities are the fractions of events in the total \(t\bar{t}\) two-jet sample with truly zero, one or two of the b-jets reconstructed and have to be determined by MC studies.

The exact sets of functions giving the expected numbers of tagged jets in the three- and four-jet sample are given in Appendix A. Only one more assumption is made for these last two samples: there are no events with three or more b-jets. In other words, we assume that the sum of the three fractions \(R_{nj}^{b\text{Jet}}\) with \(i = 0, 1, 2\) is equal to one in any jet multiplicity \(n\). ISR and FSR can lead to more b-quarks in the event, yet the high uncertainty on their production rate has led us to keep our model simply based on the presence of two b-quarks in a \(t\bar{t}\) event at Leading Order.

### 6.1.2 Trigger and Event Selection

For the event selection we have performed the standard event selection foreseen for early top quark pair analysis in the ATLAS collaboration [116]. Using the object definitions as mentioned in Section 4.6 the requirements are:

- **Missing transverse energy** \(E_T > 20\) GeV.
- **Exactly one isolated lepton**, be it electron or muon, with \(p_T > 20\) GeV. At the same time, the Event Filter trigger \(\text{EF}_e\text{25i}_\text{medium1}\) must have passed if the lepton is an electron, or the trigger \(\text{EF}_\mu\text{med}20\) must have passed if it is a muon.
- **Jets** must have a transverse momentum of \(p_T > 40\) GeV. Events with two, three or four jets are accepted.
Chapter 6. Top quark and $W$ boson production cross section measurement

Figure 6.1: Cumulative jet multiplicity after selecting one lepton (electron or muon) with $p_T > 20$ GeV and missing transverse energy of $E_T > 20$ GeV. A good jet is defined by $p_T > 40$ GeV. The distribution is normalized to an integrated luminosity of 500 pb$^{-1}$, and for the $t\bar{t}$ sample the events generated with MC@NLO are used.

MC samples used

We have analyzed MC samples for $pp$ collisions at a center-of-mass energy of 10 TeV. Throughout this chapter we refer to this data set as the MC08 data set.\footnote{The naming refers to the year 2008 in which all samples were generated and reconstructed.} In Fig. 6.1 the jet multiplicity after the other selection requirements is depicted for the samples used in this chapter, that is:

- $t\bar{t}$ generated with MC@NLO, in combination with the Herwig showering algorithm, or $t\bar{t}$ generated with AcerMC, in combination with the Pythia showering algorithm. Both samples only contain the di-leptonic and semi-leptonic decays of the top quark pair, i.e. there are no all-hadronic $t\bar{t}$ decays. The cross sections, including the branching ratios, are $\sigma_{MC@NLO} = 217$ pb and $\sigma_{AcerMC} = 218$ pb. In Fig. 6.1 the MC@NLO sample is used.

- $W$+jets, generated with the Alpgen program, with zero up to five partons. The total cross section amounts to 48.5 nb.

- $W$+$b\bar{b}$, generated with the Alpgen program, with zero up to three additional partons. The total cross section amounts to 17.9 pb.

- Single top, generated with AcerMC. The $Wt$- (no all-hadronic decays) and the $t$-channel (only leptonic decays) are used. No $s$-channel sample was produced. The cross sections, including the branching ratios, are respectively $\sigma_{Wt} = 14.3$ pb and $\sigma_t = 43.2$ pb.

The quoted cross sections are to near-NNLO\footnote{The near-NNLO is based on the NLO completed with several corrections, see [120]. It is an approximation of the NNLO.}. That is, first a cross section is given by the generator: at NLO for the $t\bar{t}$ events generated with MC@NLO, and at Leading Order for the
single top, the $W$ boson and the AcerMC $t\bar{t}$ samples. Then, for all cross sections a correction factor (the K-factor) is subsequently applied to correct them to near-NNLO calculations. See also [27,121].

From Fig. 6.1 it is clear that the $W+$jets events dominate in the two-jet sample, while in the four-jet sample the largest contribution is from $t\bar{t}$ events. The ratio of $(W+bb)/(W+\text{light jets})$ is approximately 0.01 for all jet multiplicities.

6.1.3 Fraction of b-jets in top quark pair events

The fractions $R^i_{nj}\,^\text{bjet}$, with $i = 0,1,2$ and $n = 2,3,4$ are the expected fractions of top quark pair events with truly zero, one or two of the b-jets reconstructed (but before b-tagging is applied), in the two-, three- and four-jet samples. For the determination of these fractions from the MC samples the reconstructed jets have to be identified as coming from either one of the b-quarks. We emphasize that we do not take into account b-quarks originating from ISR or FSR.

The approach we have adopted for the matching is the following: first the momentum vector of the two b-quarks is extracted from the MC generator information. Second, the direction of each b-quark is compared to the directions of all reconstructed jets passing the selection requirements. A reconstructed jet is defined to be originating from a b-quark, if the opening angle $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2}$ between the jet and the b-quark is less than 0.2. Finally the number of matched b-quarks in each event is counted. In this final step it can happen that the two b-quarks in the event are matched to the same reconstructed jet. In this case, the event is labeled as having one b-quark reconstructed.

In Table 6.1 we list the fractions extracted from the total semi-leptonic $t\bar{t}$ sample available (MC@NLO sample), with an equivalent of 1665 pb$^{-1}$ of integrated luminosity. We also list the fractions for samples with mis-calibrated jet energy scale ($JES$), i.e. the $JES$ with an offset of either + or $-10\%$.

<table>
<thead>
<tr>
<th>$t\bar{t}$</th>
<th>$R^0_{2j},^\text{bjet}$</th>
<th>$R^1_{2j},^\text{bjet}$</th>
<th>$R^2_{2j},^\text{bjet}$</th>
<th>$R^0_{3j},^\text{bjet}$</th>
<th>$R^1_{3j},^\text{bjet}$</th>
<th>$R^2_{3j},^\text{bjet}$</th>
<th>$R^0_{4j},^\text{bjet}$</th>
<th>$R^1_{4j},^\text{bjet}$</th>
<th>$R^2_{4j},^\text{bjet}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$JES = 0.9$</td>
<td>15.6</td>
<td>59.0</td>
<td>25.3</td>
<td>6.80</td>
<td>43.7</td>
<td>49.5</td>
<td>3.75</td>
<td>35.0</td>
<td>61.2</td>
</tr>
<tr>
<td>$JES = 1.0$</td>
<td>15.4</td>
<td>59.2</td>
<td>25.4</td>
<td>6.75</td>
<td>43.1</td>
<td>50.2</td>
<td>3.65</td>
<td>33.7</td>
<td>62.6</td>
</tr>
<tr>
<td>$JES = 1.1$</td>
<td>15.3</td>
<td>58.8</td>
<td>25.9</td>
<td>6.57</td>
<td>42.9</td>
<td>50.5</td>
<td>3.75</td>
<td>33.0</td>
<td>63.2</td>
</tr>
<tr>
<td>Stat.error</td>
<td>±0.2</td>
<td>±0.4</td>
<td>±0.3</td>
<td>±0.1</td>
<td>±0.3</td>
<td>±0.4</td>
<td>±0.2</td>
<td>±0.5</td>
<td>±0.6</td>
</tr>
</tbody>
</table>

Table 6.1: The $R^i_{nj}\,^\text{bjet}$ fractions with correct and mis-calibrated $JES$, extracted from the semi-leptonic $t\bar{t}$ sample, given in percentages. A factor $JES = 1.1$ means that the jet energy is scaled 10% too high and thus more jets pass the selection cut.

There is a small influence of the $JES$ on the number of matched b-jets. In Section 6.2.1 we study the systematic uncertainty as a consequence of different $R^i_{nj}\,^\text{bjet}$ fractions; in Section 6.2.2 we study the consequences of mis-calibrated energy scales.

A last difficulty, which we however cannot quantify, is the possibility that a jet is not matched to a b-quark, even though it follows the hadronization of one of the initial b-quarks. In the approach adopted in the calculation of the fractions $R^i_{nj}\,^\text{bjet}$ this jet is counted as a light quark jet. Although this could affect the analysis, we note that as the jet is poorly reconstructed (its direction is off compared to the initial b-quark's direction), the performance of the b-tagging algorithm will probably also be far from ideal.
Fraction of b-jets in AcerMC top quark pair events

For the systematic uncertainty study we will also analyze the effect of using a different $t\bar{t}$ generator: AcerMC. In Table 6.2 the fractions for both samples are summarized and we observe that in the MC@NLO sample on average the events have more b-jets reconstructed. The difference between the two samples is definitely non-negligible. In [93, 122] the different R$_i$ bJet$_n$ fractions to study the effective systematic uncertainty in Section 6.2.1.

<table>
<thead>
<tr>
<th>$R^0_{2j}$</th>
<th>$R^1_{2j}$</th>
<th>$R^2_{2j}$</th>
<th>$R^0_{3j}$</th>
<th>$R^1_{3j}$</th>
<th>$R^2_{3j}$</th>
<th>$R^0_{4j}$</th>
<th>$R^1_{4j}$</th>
<th>$R^2_{4j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC@NLO</td>
<td>15.4</td>
<td>59.2</td>
<td>25.4</td>
<td>6.75</td>
<td>43.1</td>
<td>50.2</td>
<td>3.65</td>
<td>33.7</td>
</tr>
<tr>
<td>AcerMC</td>
<td>17.5</td>
<td>60.4</td>
<td>22.1</td>
<td>8.16</td>
<td>45.4</td>
<td>46.5</td>
<td>6.00</td>
<td>37.8</td>
</tr>
</tbody>
</table>

Table 6.2: The $R^i_{n_j}$ fractions for MC@NLO and AcerMC samples for $n = 2, 3, 4$ and $i = 0, 1, 2$. The differences between the two samples are larger than the statistical errors, which are of the order of the errors listed in Table 6.1.

showing programs used for the two samples are shown to cause a discrepancy in the event jet multiplicity. We will use the difference between the sets of $R^i_{n_j}$ fractions to study the effective systematic uncertainty in Section 6.2.1.

Fraction of b-jets in single top events

The NLO cross sections for the different single top production channels at the LHC are summarized in Table 1.2 on page 11. The s-channel results at LO in two b-quarks, the two other channels can only result in two b-quarks at NLO. By analyzing the MC generator information in the events, we can extract the $R^i_{n_j}$ fractions for the selected single top events. As we stated in Section 6.1.2 there are no s-channel simulations from AcerMC available at center-of-mass energy of 10 TeV.

From the available single top samples we extract the fractions as summarized in Table 6.3. It is clear that less b-jets are reconstructed in single top events than in $t\bar{t}$ events. There are however far less single top events than $t\bar{t}$ events, making the differences between the $R^i_{n_j}$ for the single top sample and the $R^i_{n_j}$ for the $t\bar{t}$ sample less relevant. We will quantify this in Section 6.1.5.

<table>
<thead>
<tr>
<th>single top</th>
<th>$R^0_{2j}$</th>
<th>$R^1_{2j}$</th>
<th>$R^2_{2j}$</th>
<th>$R^0_{3j}$</th>
<th>$R^1_{3j}$</th>
<th>$R^2_{3j}$</th>
<th>$R^0_{4j}$</th>
<th>$R^1_{4j}$</th>
<th>$R^2_{4j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-channel</td>
<td>15.4</td>
<td>72.7</td>
<td>11.9</td>
<td>10.5</td>
<td>56.6</td>
<td>32.9</td>
<td>10.0</td>
<td>44.8</td>
<td>45.2</td>
</tr>
<tr>
<td>Wt-channel</td>
<td>25.1</td>
<td>69.9</td>
<td>5.00</td>
<td>14.1</td>
<td>68.9</td>
<td>17.0</td>
<td>12.3</td>
<td>53.5</td>
<td>34.2</td>
</tr>
<tr>
<td>Total</td>
<td>18.9</td>
<td>71.6</td>
<td>9.50</td>
<td>11.7</td>
<td>60.0</td>
<td>28.3</td>
<td>10.6</td>
<td>47.2</td>
<td>42.2</td>
</tr>
</tbody>
</table>

Table 6.3: The $R^i_{n_j}$ fractions extracted from the available single top samples for $n = 2, 3, 4$ and $i = 0, 1, 2$, given in percentages.
6.1.4 Extracting the b-tagging efficiency for $W+$jets from $Z+$jets events

Figure 6.2 depicts the $\varepsilon_i^W$ distribution for the $W+$jets samples with two, three and four jets. Above is the result of the IP3D+SV1 b-tagger, below of the JetProb b-tagger. The efficiencies for the three samples are similar in most of the b-tagger weight ranges. In our analysis we will use a weighted average of the $\varepsilon_i^W$ for the two- and three jet sample. This choice was made for two reasons: first, to have the same $\varepsilon_i^W$-distribution in the analyses including and excluding the four-jet sample. Second, the average was taken to prevent ending up with an efficiency on the three jet sample based on too few events.

The small differences visible in each plot of Fig. 6.2 between the efficiencies for the three samples are negligible: the differences are of the order of the statistical uncertainty on the

![Graphs](image)

Figure 6.2: The $\varepsilon_i^W$ for the two-, three- and four-jet sample as extracted from the MC generator information of the $W+$jets samples, after the event selection requirements of Section 6.1.2. All available $W+$jets events are used, including $W+b\bar{b}$, corresponding to approximately $L = 150 \text{ pb}^{-1}$. Above are the distributions for the IP3D+SV1 b-tagger, below for the JetProb b-tagger.
efficiencies (not depicted in the figures).

The b-tagging efficiency $\varepsilon^W_l$ is an input to the analysis in this chapter; it must be measured on the data before the analysis can be performed. As stated before, we have labeled all jets in the $W$+jets sample as ‘light quark jets’, including those from $W+b\bar{b}$ events. For this latter contamination of our light quark jet sample not to be a problem in the determination of $\varepsilon^W_l$, we must measure the b-tagging efficiency on a sample of events with a jet content resembling that of the expected $W$+jets sample. The approach we have adopted is to measure $\varepsilon^W_l$ from $Z$+jets events, i.e. events with exactly two leptons.

With a leptonic decay of either bosons the jets present in a $W$- or $Z$+jets event are all ISR and/or FSR. If the contamination from additional b-jets is also similar in the two samples, a clean sample of $Z$+jets events could thus be used to measure the b-tagging efficiency $\varepsilon^W_l$. From Section 1.1.4 we already know that this is not exactly the case: the ratio of $Z+b\bar{b}$ to $Z$+jets is approximately four times larger than the ratio $W+b\bar{b}$ to $W$+jets.

We now explain how the efficiency can be extracted from the first $50 \text{ pb}^{-1}$ of data, selecting $Z$+jets events with two leptons.

**Two-lepton event selection**

To show that the b-tagging efficiency is similar on $W$- and $Z$-events, we analyze the 10 TeV center-of-mass energy $Z$+light quark jets samples. We emphasize that we do not study any $Z+b\bar{b}$ events. To select the two-lepton $Z$-events to be as similar as possible to the one-lepton $W$-events in our analysis, the $E_T$ requirement is simply replaced by an extra lepton. That is, events are selected with exactly two electrons or two muons, with $p_T > 20$ GeV; no selection on charges, and no $E_T$ requirement. To prevent biases introduced by the trigger, the same trigger is used on the $Z$-events.

With a two-lepton selection, the sample consists mostly of $Z$+jets and $t\bar{t}$ events, see [123].

---

4) This one-to-one replacement is not entirely correct. First, the acceptances of the leptons and of the neutrino are not the same. Second, the $p_T$ distribution of the neutrino is not the same as that of the electron or muon in a $W$ decay.

![Figure 6.3: Jet multiplicity for events with exactly two leptons in a sample of 500 pb$^{-1}$. See text for details on event selection.](image_url)
6.1. The analysis setup

In Fig. 6.3(a) the cumulative jet multiplicity is depicted after our two-lepton requirements for the Z+jets, $t\bar{t}$ and single top samples of 500 pb$^{-1}$. The contribution from top events can be reduced with an extra requirement of $|M_{ll} - M_Z| < 5.0$ GeV, where $M_{ll}$ is the invariant mass of the two leptons. Figure 6.3(b) depicts the jet multiplicity distributions after this extra requirement: most of the top events are removed, while the amount of Z-events is reduced by only $\sim 1/3$. Extrapolating this to a first data sample of 50 pb$^{-1}$ would give us $\sim 500$ Z-events to measure the b-tagging efficiency.

**Extracting $\varepsilon^W_l$ from $\varepsilon^Z_l$**

For the measurement of the b-tagging efficiency we only use the average of the two- and three-jet samples. In Fig. 6.4(a) and 6.4(b) the results are shown for respectively the combined IP3D+SV1 b-tagger and the JetProb b-tagger. The black distribution is the efficiency on the selected $W$-events (that is with the one lepton and $E_T$ cut). This is the efficiency needed in our analysis. The ‘stars’-distribution is the efficiency for all events with exactly two leptons.\(^5\) For both b-taggers, and especially IP3D+SV1, the b-tagging efficiency is higher on the two-lepton sample. This is caused by the many top events present, as shown with Fig. 6.3(a).

With the extra requirement of $|M_{ll} - M_Z| < 5.0$ GeV on the two-lepton sample most of the top events are removed. In Fig. 6.4(a) and 6.4(b) the gray distributions (with error-bars) show the results after this extra requirement on samples with integrated luminosity of 50 pb$^{-1}$; the uncertainties shown are statistical. The conclusion is that even with not more than approximately 500 Z+jets events with two leptons a meaningful b-tagging efficiency can be extrapolated. In Fig. 6.5 the exact b-tagging efficiency for only the Z+jets sample is again shown and we see that it is almost a perfect match with the $\varepsilon^W_l$.

**Uncertainty on the b-tagging efficiency**

One problem remains, that is the possibility of contamination of the Z+jets events with Z+$b\bar{b}$. In Section 1.1.4 we listed the cross sections of $W$- and Z+$b\bar{b}$ with leptonic decay. From the same section we can deduce that the ratio of Z+$b\bar{b}$ to Z+jets is approximately four times larger than the ratio $W+b\bar{b}$ to W+jets. Unfortunately ATLAS has not produced a Z+$b\bar{b}$ sample at 10 TeV at the time of writing of this thesis; to measure the possible offset for $\varepsilon^W_l$ we have therefore calculated the b-tagging efficiency on a sample of W+jets including a sample of $W+b\bar{b}$ four times too large or four times too small. The result is depicted in Fig. 6.6, where the offset of the gray distributions compared to the upper/lower error-bars is the offset caused by too many/few $W+b\bar{b}$ events. These two outer distributions will be used in Section 6.4.2 to determine the uncertainty on the fit results as a consequence of the uncertainty on $\varepsilon^W_l$ due to b-jets in the Z+jets sample.

At a later stage the ratios of $W+b\bar{b}$ to W+jets and Z+$b\bar{b}$ to Z+jets will probably be measured by ATLAS. These measurements should be used to validate the analysis presented in this chapter.

The study presented in this section has shown that $\varepsilon^W_l$ can be extracted from two-lepton events. In the remaining studies in this chapter the true number of tagged jets in the W+jets events is used to define $\varepsilon^W_l$.

\(^{5}\)Both the black and the stars-distributions are measured on all available events to increase the number of selected events.
Figure 6.4: B-tagging efficiencies with the IP3D+SV1 b-tagger (a) and the JetProb b-tagger (b). Black: IP3D+SV1 efficiency on the $W$ + light quark jets, i.e. one-lepton events. Stars: IP3D+SV1 efficiency on all events with two leptons. Gray with error-bars: same as stars, with extra requirement $|M_\ell - M_Z| < 5.0$ GeV, sample of 50 pb$^{-1}$. The figures show how in the two-lepton channel a simple requirement on the di-lepton invariant mass reduces the b-tagging efficiency on the present jets to the b-tagging efficiency on the one-lepton $W+\text{jets}$ events.
6.1. The analysis setup

Figure 6.5: Black: IP3D+SV1 efficiency on the $W +$ light quark jets events with one lepton (the errors are too small to see). Gray: IP3D+SV1 efficiency on $Z +$ light quark jets events with two leptons. Both are calculated on a sample of 150 pb$^{-1}$. No background channels are included. The results show how the b-tagging efficiency on the pure $W$- and $Z$+jets samples almost coincide; the difference is less than the statistical uncertainty at 50 pb$^{-1}$ in Fig. 6.4(a).

Figure 6.6: B-tagging efficiency on the $W+\text{jets}$ sample with the JetProb b-tagger. The black distribution is the efficiency on $W+\text{jets}$ events (incl. $W+b\bar{b}$ ). The uncertainty of this black distribution is calculated from a sample of $5 \cdot 10^2$ events to estimate the statistical error to be expected from an analysis on 50 pb$^{-1}$ $Z+\text{jets}$ sample with two leptons. The upper/lower gray distribution is again the efficiency on the $W+\text{jets}$ sample. This is calculated with the $W+b\bar{b}$ contribution scaled up/down by a factor of 4, and also includes the upper/lower fluctuation of the statistical uncertainty shown in black. These two extreme distributions will be used later on in this chapter.
Chapter 6. Top quark and $W$ boson production cross section measurement

6.1.5 Results on a large sample

In this section we show how a set of solutions is extracted. A large sample of events with $L = 500 \text{ pb}^{-1}$ is used to reduce the statistical uncertainty, except for the $W+\text{jets}$ sample, for which the results have been scaled up from approximately $150 \text{ pb}^{-1}$. Performing a scan over a large range of the IP3D+SV1 b-tagger weights, the results are the five distributions as shown in Fig. 6.7.

**Definition:** The $\#top$ is defined as the sum of $t\bar{t}$ and single top events and will be used as such throughout this chapter.

In the results for the $\#top/\#total$ fits in Fig. 6.7 we see a clear plateau over a large range in the b-tag weight, corresponding to a large range in $\varepsilon_b$. As the same sample is used in each of the 100 fits at different weights, the results in each of the five distributions are correlated. To extract one value from these fits, we must therefore chose one specific weight. Throughout this chapter we state the results at a IP3D+SV1 weight of 4.0, as this is usually the point halfway the plateau.

At this weight the $\varepsilon_l^{\text{top}}$ is 0.05, while $\varepsilon_b$ is still at 0.64. Higher weights reduce the $\varepsilon_b$ faster than the $\varepsilon_l^{\text{top}}$, reducing the discriminating power of the b-tagger and effectively decreasing the degree of freedom in the maximization of the likelihood. This effect is visible for weights above $\sim 10$, for which the results are less accurate. At lower weights the plateau can extend down to a weight of $\sim -2.0$, below which the $\varepsilon_l^{\text{top}}$ increases rapidly. With more and more jets tagged, the number of events with one tag reaches zero, effectively again decreasing the degree of freedom in the analysis. Below a weight of $\sim -2.0$ the analysis therefore does not result in sensible measurements.

The conclusion is that over a large range of the b-tagger weight the analysis finds a stable solution and the analysis is thus almost independent of the choice of b-tagging efficiency.

Table 6.4 lists the number of events of each of the sub-samples selected and the $\#top/\#total$ ratios that should be expected. The fitted $\#top/\#total$ stated in the Table are extracted from Fig. 6.7 at a IP3D+SV1 weight of 4.0. Here we see that the single top events are fitted as tops: the fitted $\#tops$ corresponds to the true $\#tops$, defined as the sum of $t\bar{t}$ plus single top events. The conclusion can thus be drawn that the differences between the $R_i^{3\text{jet}}$ fractions for the single top sample and those for the $t\bar{t}$ sample, as mentioned in Section 6.1.3, are of little importance.

<table>
<thead>
<tr>
<th></th>
<th>$t\bar{t}$ (MC@NLO)</th>
<th>Single top (incl. $W+bb$)</th>
<th>$W+\text{jets}$</th>
<th>True $#top/#total$</th>
<th>Fitted $#top/#total$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>$1.20 \cdot 10^4$</td>
<td>$3.05 \cdot 10^3$</td>
<td>$7.07 \cdot 10^4$</td>
<td>0.17</td>
<td>0.17 ± 0.005</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>$1.10 \cdot 10^4$</td>
<td>$1.31 \cdot 10^3$</td>
<td>$1.23 \cdot 10^4$</td>
<td>0.50</td>
<td>0.51 ± 0.01</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>$4.81 \cdot 10^3$</td>
<td>313</td>
<td>$2.25 \cdot 10^3$</td>
<td>0.70</td>
<td>0.70 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.4: The number of events with 2, 3 or 4 jets passing the selection cuts. An integrated luminosity of $500 \text{ pb}^{-1}$ is used for each sub-sample. The fitted $\#top/\#total$ are extracted from Fig. 6.7 at a IP3D+SV1 weight of 4.0.
Table 6.5: The $b$-tagging efficiencies on the $b$-jets in the $t\bar{t}$ sample, on the light quark jets in the $t\bar{t}$ sample and on the jets in the $W+$jets sample. The latter is shown for comparison only. The fitted efficiencies are extracted from Fig. 6.7 at a IP3D+SV1 weight of 4.0. The fitted $\varepsilon_b$ agrees with the true value within the statistical uncertainty, the fitted $\varepsilon_{\text{top}}$ is too low. See text for more details.

<table>
<thead>
<tr>
<th></th>
<th>True efficiency</th>
<th>Fitted efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_b$</td>
<td>0.630 ± 0.003</td>
<td>0.638 ± 0.011</td>
</tr>
<tr>
<td>$\varepsilon_{\text{top}}$</td>
<td>0.053 ± 0.001</td>
<td>0.037 ± 0.004</td>
</tr>
<tr>
<td>$\varepsilon_W$</td>
<td>0.027 ± 0.0004</td>
<td>-</td>
</tr>
</tbody>
</table>

In Table 6.5 the fit results for $\varepsilon_b$ and $\varepsilon_{\text{top}}$ at a IP3D+SV1 weight of 4.0 can be compared to the true values for the efficiencies. The fitted $\varepsilon_b$ agrees with the true value within the statistical uncertainty. As can also be seen in Fig. 6.7(a), the fitted $\varepsilon_{\text{top}}$ is significantly lower than the true value. During this research we have not found any specific observable which has an influence on the $\varepsilon_{\text{top}}$ only, nor have we found any sign that the offset propagates in a mis-measurement of the $\#\text{top}/\#\text{total}$ ratios.

As the other measurements do correspond with the expected values, the small offset of $\varepsilon_{\text{top}}$ might imply that the likelihood functions as constructed in Section 6.1.1 are not completely accurate: in Section 4.4.3 it was discussed how the $b$-tagging efficiencies depend on the jet transverse momentum, yet in the likelihood functions this is not taken into account. Although this negligence might be the reason for the offset of $\varepsilon_{\text{top}}$, it is not important. The offset is small, and all other measurements work well.

The overall conclusion is that, within the statistical uncertainty on the fit, the results are a non-biased match of the values for $\#\text{top}/\#\text{total}$ with the true ratios and an accurate measurement of the $b$-tagging efficiency on $b$-jets, over a large range of the efficiencies.
Chapter 6. Top quark and $W$ boson production cross section measurement

Figure 6.7: Fit result for a sample with integrated luminosity of 500 pb$^{-1}$, where IP3D+SV1 is used as b-tagger. All results correspond with the expected values, except for a small offset in the fit of $\epsilon_t^{top}$. The correct values for the ratio fits are depicted by dotted lines.
6.2 Systematic uncertainties

In this section we discuss the systematic uncertainties present in our analysis. The best performing b-tagger is used in order to get a good estimate of the size of the uncertainties.

6.2.1 The expected fraction of b-jets

In Section 6.1.3 we explained the $R^{bjet}_{nj}$ fractions and showed their values in top samples from different MC generators. For a study of the systematic uncertainty due to variations in these values we have chosen to apply the fractions obtained from the AcerMC sample on a 500 pb$^{-1}$ sample containing $t\bar{t}$ events from MC@NLO, and vice-versa.

<table>
<thead>
<tr>
<th></th>
<th>True #top/#total</th>
<th>Fitted #top/#total $R^{bjet}_{nj}$ from AcerMC</th>
<th>Fitted #top/#total $R^{bjet}_{nj}$ from MC@NLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>0.17</td>
<td>0.17 ± 0.005</td>
<td>0.17 ± 0.005</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>0.50</td>
<td>0.50 ± 0.01</td>
<td>0.51 ± 0.01</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>0.70</td>
<td>0.71 ± 0.02</td>
<td>0.70 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.6: The fitted #top/#total ratios with the $R^{bjet}_{nj}$ fractions from two different $t\bar{t}$ samples applied on a sample containing $t\bar{t}$ events from MC@NLO.

<table>
<thead>
<tr>
<th></th>
<th>True #top/#total</th>
<th>Fitted #top/#total $R^{bjet}_{nj}$ from AcerMC</th>
<th>Fitted #top/#total $R^{bjet}_{nj}$ from MC@NLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>0.17</td>
<td>0.17 ± 0.005</td>
<td>0.17 ± 0.005</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>0.52</td>
<td>0.53 ± 0.01</td>
<td>0.54 ± 0.01</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>0.74</td>
<td>0.75 ± 0.02</td>
<td>0.75 ± 0.02</td>
</tr>
</tbody>
</table>

Table 6.7: The fitted #top/#total ratios with the $R^{bjet}_{nj}$ fractions from two different $t\bar{t}$ samples applied on a sample containing $t\bar{t}$ events from AcerMC.

The results are summarized in Table 6.6 and 6.7, where we observe that the effect is minimal: within each table the differences between AcerMC and MC@NLO are not more than 1% in the three- and four-jet samples. No effect is visible for the results in the two-jet sample. The differences between the two tables listed are another systematic effect, which is discussed in Section 6.2.3.

The results for the fitted values of the b-tagging efficiencies $\epsilon_b$ and $\epsilon_{top}$ show no systematic effect and are in agreement with the results stated in section 6.1.5.

We note that the systematic uncertainty due to variations in values of the $R^{bjet}_{nj}$ fractions does not have to be identical when a different b-tagging algorithm is used. Yet we will use these results when applying the JetProb algorithm, for the simple reason that a similar study to obtain the uncertainty with the JetProb b-tagger remains inconclusive. The performance of JetProb is worse and the small uncertainty is not measurable.
6.2.2 Energy scale variations

We have studied independently two energy scales subject to mis-calibration: the Jet Energy Scale (JES) and the $E_T$ scale. The two are however correlated. If for example we select an event with only one jet, with the direction of the $E_T$ opposite to the jet direction, and the jet energy is measured with an offset of +10%, the missing transverse energy is also measured with an offset of approximately +10%. The correlation is not exactly 100%, as there are some out-of-cluster effects, see Section 4.5.1, but to first approximation the correlation is very high.

In our studies of the JES we have adjusted the jet energies with +10% and -10%, while at the same time adjusting the $E_T$ such that the vectorial sum of all jets+$E_T$ stays constant. These steps are of course taken before the event selection. In our studies of the $E_T$ scale we only adjust the $E_T$, again with +10% and -10%. The changes have no effect on the stability of the fits. Although the $R_{nj}^{bJet}$ fractions used are slightly off in the mis-calibrated samples, see Table 6.1, we know from Section 6.2.1 that this effect is negligible.

In Table 6.8 we state the effect of JES offsets on the fitted #tops and #W, in Table 6.9 the results are listed with $E_T$ scale offsets. For clarity: with a JES = 1.1 we mean that the jet energies have been scaled up. As the selection of jets stays at $p_T > 40$ GeV, this results in more jets passing the cut. A similar reasoning goes for $E_T$ scale $= 1.1$, etc... As the energy scale offsets alter the absolute number of events selected, the tables list the absolute number of tops and $W$ fitted; the total number of events is of course the sum of both. The changes compared to the fitted number of events in Table 6.4 are given in percentages. The fitted number of events agree well with the true values of the selected events in each channel.

<table>
<thead>
<tr>
<th></th>
<th>Fitted #tops (JES = 1.1)</th>
<th>Fitted #W (JES = 1.1)</th>
<th>Fitted #tops (JES = 0.9)</th>
<th>Fitted #W (JES = 0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>14.1(-4.3%)</td>
<td>83.8(+18.1%)</td>
<td>15.3(+3.6%)</td>
<td>58.9(-17.0%)</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>13.8(+10.1%)</td>
<td>14.9(+24.2%)</td>
<td>10.9(-12.7%)</td>
<td>8.93(-25.7%)</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>6.45(+25.0%)</td>
<td>2.76(+25.6%)</td>
<td>3.96(-23.2%)</td>
<td>1.54(-30.0%)</td>
</tr>
</tbody>
</table>

Table 6.8: Results for the fitted #tops and #W ($\times 10^3$) with jet energy over-calibrated (JES=1.1) and under-calibrated (JES=0.9). The changes compared to the fitted number of events in Table 6.4 are given in percentages.

<table>
<thead>
<tr>
<th></th>
<th>Fitted #tops ($E_T$ scale = 1.1)</th>
<th>Fitted #W ($E_T$ scale = 1.1)</th>
<th>Fitted #tops ($E_T$ scale = 0.9)</th>
<th>Fitted #W ($E_T$ scale = 0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>15.0(+1.4%)</td>
<td>73.0(+2.8%)</td>
<td>14.4(-2.1%)</td>
<td>68.1(-4.1%)</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>12.8(+1.9%)</td>
<td>12.3(+2.0%)</td>
<td>12.1(-3.0%)</td>
<td>11.6(-3.0%)</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>5.32(+3.0%)</td>
<td>2.22(+1.0%)</td>
<td>5.12(-1.0%)</td>
<td>2.09(-5.1%)</td>
</tr>
</tbody>
</table>

Table 6.9: Results for the fitted #tops and #W ($\times 10^3$) with $E_T$ over-calibrated ($E_T$ scale=1.1) and under-calibrated ($E_T$ scale=0.9).

---

6) This is in fact an overestimation, as the $E_T$ scale is correlated to the JES.
In Table 6.10 the results are listed for the b-tagging efficiencies with \( JES \)- and \( \mathbb{E}_T \) scale offsets. Although the absolute number of events selected showed large changes with the energy scale offsets, the efficiencies are almost not affected. Within the statistical uncertainties the results correspond with those in Table 6.5.

<table>
<thead>
<tr>
<th></th>
<th>True efficiency</th>
<th>Fitted efficiency</th>
<th>True efficiency</th>
<th>Fitted efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((JES = 1.1))</td>
<td>((JES = 1.1))</td>
<td>((JES = 0.9))</td>
<td>((JES = 0.9))</td>
</tr>
<tr>
<td>(\varepsilon_b)</td>
<td>0.626 ± 0.002</td>
<td>0.627 ± 0.011</td>
<td>0.635 ± 0.003</td>
<td>0.626 ± 0.013</td>
</tr>
<tr>
<td>(\varepsilon_{t\bar{t}}^{\text{top}})</td>
<td>0.052 ± 0.001</td>
<td>0.039 ± 0.004</td>
<td>0.056 ± 0.002</td>
<td>0.040 ± 0.005</td>
</tr>
<tr>
<td>(\varepsilon_{t\bar{t}}^{W})</td>
<td>0.026 ± 0.0003</td>
<td>-</td>
<td>0.028 ± 0.0004</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>True efficiency</th>
<th>Fitted efficiency</th>
<th>True efficiency</th>
<th>Fitted efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((E_T \text{ scale} = 1.1))</td>
<td>((E_T \text{ scale} = 1.1))</td>
<td>((E_T \text{ scale} = 0.9))</td>
<td>((E_T \text{ scale} = 0.9))</td>
</tr>
<tr>
<td>(\varepsilon_b)</td>
<td>0.630 ± 0.003</td>
<td>0.637 ± 0.011</td>
<td>0.629 ± 0.003</td>
<td>0.635 ± 0.011</td>
</tr>
<tr>
<td>(\varepsilon_{t\bar{t}}^{\text{top}})</td>
<td>0.053 ± 0.001</td>
<td>0.037 ± 0.004</td>
<td>0.054 ± 0.001</td>
<td>0.038 ± 0.004</td>
</tr>
<tr>
<td>(\varepsilon_{t\bar{t}}^{W})</td>
<td>0.027 ± 0.0004</td>
<td>-</td>
<td>0.027 ± 0.0004</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6.10: Results for the b-tagging efficiencies with jet energy over-calibrated \((JES=1.1)\) and under-calibrated \((JES=0.9)\), and with \(\mathbb{E}_T\) over-calibrated \((\mathbb{E}_T \text{ scale}=1.1)\) and under-calibrated \((\mathbb{E}_T \text{ scale}=0.9)\).

### 6.2.3 The expected total number of events

In Table 6.4 results are given with the \(t\bar{t}\) events generated by MC@NLO. The change which occurs in the number of \(t\bar{t}\) events selected when an AcerMC sample is used, is shown in Table 6.11. In fact it is not the difference in generators which results in the discrepancies, it is the showering procedure after the generation which is the cause: the same effect was measured in [93]. For more details we refer to [122], with the comment that this work was performed on MC samples generated in an earlier version of Athena (version 12). Similar studies on newer MC samples are in progress. For the final measurement of the number of top and the number of \(W\) events in Section 6.4.3 we will use the average of the measurements in Table 6.11, and assign half of the differences as systematic uncertainties.

The uncertainty expected on the total number of events also depends on the uncertainty on the measurement of the luminosity. In [29] it is explained that the luminosity initially can be measured to an accuracy of \(\sim 20\%\), which ultimately should improve to \(\sim 5\%\).

<table>
<thead>
<tr>
<th></th>
<th>AcerMC</th>
<th>MC@NLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>11.8</td>
<td>12.0</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>12.2</td>
<td>11.0</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>6.08</td>
<td>4.81</td>
</tr>
</tbody>
</table>

Table 6.11: The number of \(t\bar{t}\) events \((\times 10^3)\) selected from an AcerMC and from a MC@NLO sample, both normalized to an integrated luminosity of 500 pb\(^{-1}\).
6.2.4 Discussion

We have studied most of the systematic uncertainties important for our analysis. One uncertainty we could not quantify: the uncertainty on the expected number of $W$+jets events, which has been calculated so far only by looking at Alpgen samples. Unfortunately we do not have immediate access to enough events from other generators to check systematic effects on $W$+jets events, but the effect could be as high as 50% on the expected number of events in the four-jet sample, see [116]. The uncertainty on the expected number of single top events is also not accounted for, although its low value compared to the expected number of $t\bar{t}$ events makes this uncertainty less important.

With a good overview of the performance of the analysis we now first show how it can be applied to measure the cross sections for $t\bar{t}$, single top and $W$+jets in a sample with known sub-samples, yet unknown cross sections. In Section 6.4 we continue with a small sample of 50 pb$^{-1}$ and we implement the systematic uncertainties.

6.3 Analysis of a ‘blind’ mixed sample

As to check that the analysis presented does not run into problems if for example the yield of the $W$+jets events is varied without changing the $t\bar{t}$ normalization, we apply the analysis to a ‘blind’ mixed sample.

This Mixed sample was generated into a Egamma-, Muon- and a Jet-stream for the ATLAS collaboration, see also 2.6.3. In this section only the Muon-stream is used, i.e. events triggered by the muon trigger. The sample has an integrated luminosity of 146 pb$^{-1}$. Before starting the exercise, the list of sub-samples in the mix is:

- $t\bar{t}$ generated with AcerMC, in combination with the Pythia showering algorithm.
- $W$+jets generated with Alpgen, with between two and five additional partons.
- Single top generated with AcerMC; only the Wt- and $t$-channel.

This resembles closely the samples mentioned in Section 6.1.2, except that no $W$+$b\bar{b}$ is used. We have applied our analysis to this Mixed sample, as if it were the real unknown data from ATLAS. To make a sensible comparison with MC-results, we cannot compare these to the results in the previous sections, as these did contain the $W$+$b\bar{b}$. We therefore have rerun the analysis on a new MC-sample, with the same sub-samples as present in the Mixed sample. Throughout the rest of this section we refer to this sample as the MC08 sample. As a test, we do compare the results on MC08 containing $t\bar{t}$ from MC@NLO, and on MC08 containing $t\bar{t}$ from AcerMC.

6.3.1 Mixed sample results

Figure 6.8 shows the fit results of the analysis on the Mixed muon-stream using the IP3D+SV1 b-tagger and using the $R_{b,j}^{b\bar{b}}$ fractions from MC@NLO in the maximization of the likelihood functions. As in Fig. 6.7(a) the $\varepsilon_{\ell}^{\text{top}}$ has a visible offset. From the discussion on the results in Section 6.1.5 we conclude that this is no reason for concern and that the measured $\#\text{top}/\#\text{total}$ ratios are unbiased.

Measuring the ratios again at a weight of 4.0, we obtain for the number of tops and the number of $W$’s the results as listed in Table 6.12. Performing the same steps on MC08, with the
6.3. Analysis of a ‘blind’ mixed sample

<table>
<thead>
<tr>
<th>Total # events</th>
<th>Fitted # tops</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>17.2</td>
<td>2.49 ± 0.17</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>4.78</td>
<td>2.32 ± 0.12</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>1.52</td>
<td>1.07 ± 0.08</td>
</tr>
<tr>
<td>TOTAL</td>
<td>23.5</td>
<td>5.87 ± 0.22</td>
</tr>
</tbody>
</table>

Table 6.12: The results ($\times 10^3$) for the fitted #tops and #W on the selected events in the Mixed sample muon-stream. The $R_{nj}^{bjet}$ fractions from MC@NLO are used in the likelihood functions.

<table>
<thead>
<tr>
<th>Total # events</th>
<th>Fitted # tops</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>15.0</td>
<td>2.53 ± 0.08</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>4.23</td>
<td>2.14 ± 0.04</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>1.28</td>
<td>0.90 ± 0.03</td>
</tr>
<tr>
<td>TOTAL</td>
<td>20.5</td>
<td>5.57 ± 0.09</td>
</tr>
</tbody>
</table>

Table 6.13: The results ($\times 10^3$) for the fitted #tops and #W on MC08, containing $t\bar{t}$ generated by MC@NLO. These values are obtained by using all available events and scaling the samples to 146 pb$^{-1}$; the statistical uncertainties are thus smaller compared to those in Table 6.12. The $R_{nj}^{bjet}$ fractions from MC@NLO are used in the likelihood functions.

$t\bar{t}$ events generated by MC@NLO, the results in Table 6.13 are obtained. For an extraction of the cross section of the $W+$jets we can now simply compare the total number of measured $W$ events in the Mixed sample to the number of $W$ events in MC08, resulting in

$$\sigma_W^{\text{Mix}} = \frac{(1.76 \pm 0.02) \cdot 10^4}{1.49 \cdot 10^4} \cdot \sigma_W^{\text{MC}} = (1.18 \pm 0.02) \cdot \sigma_W^{\text{MC}}$$ (6.8)

For the calculation of the cross sections of the $t\bar{t}$ and of the single top events we use the measurements of the fitted number of tops in the two- and three-jet samples. Defining the factors

$$f_{t\bar{t}} = \frac{\sigma_{\text{Mix}}^{t\bar{t}}}{\sigma_{\text{MC}}^{t\bar{t}}}$$
$$f_s^{\text{singleTop}} = \frac{\sigma_{\text{Mix}}^{\text{singleTop}}}{\sigma_{\text{MC}}^{\text{singleTop}}}$$ (6.9)

we can solve these two factors from the two equations

$$f_{t\bar{t}}(N_{\text{MC2j}}^{\text{tops}} - N_{2j}^{\text{singleTop}}) + f_s^{\text{singleTop}}(N_{\text{MC2j}}^{\text{tops}} - N_{2j}^{\text{t\bar{t}}}) = N_{\text{Mix,2j}}^{\text{tops}}$$ (6.10)
$$f_{t\bar{t}}(N_{\text{MC3j}}^{\text{tops}} - N_{3j}^{\text{singleTop}}) + f_s^{\text{singleTop}}(N_{\text{MC2j}}^{\text{tops}} - N_{2j}^{\text{t\bar{t}}}) = N_{\text{Mix,3j}}^{\text{tops}}.$$ (6.11)

The variables $N_{\text{Mix,XX}}^{\text{tops}}$ are the fitted number of tops in the Mixed sample, in the two- or three jet sample. $N_{XX}^{t\bar{t}}$ is the true number of $t\bar{t}$ in MC08 (the same holds for $N_{XX}^{\text{singleTop}}$), while $N_{\text{MC,XX}}^{\text{tops}}$ is the fitted number of tops in MC08. Although no bias is measured in the fitted ratios so far, this setup ensures that any possible bias is accounted for when measuring on real data. The end result is:

$$\sigma_{\text{Mix}}^{t\bar{t}} = (1.24 \pm 0.13) \cdot \sigma_{\text{MC}}^{t\bar{t}}$$
$$\sigma_{\text{Mix}}^{\text{singleTop}} = (0.00 \pm 0.60) \cdot \sigma_{\text{MC}}^{\text{singleTop}}.$$ (6.12)
Chapter 6. Top quark and W boson production cross section measurement

Figure 6.8: Fit results for the Mixed sample (muon stream) with $L = 146 \text{ pb}^{-1}$. The IP3D+SV1 is used as b-tagger, and the $R_{nj}^{\ell, kjet}$ fractions are extracted from the MC@NLO sample.
6.3. Analysis of a ‘blind’ mixed sample

<table>
<thead>
<tr>
<th></th>
<th>Total # events</th>
<th>Fitted # tops</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>14.9</td>
<td>2.52 ± 0.08</td>
<td>12.4 ± 0.08</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>4.42</td>
<td>2.37 ± 0.04</td>
<td>2.06 ± 0.04</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>1.49</td>
<td>1.10 ± 0.03</td>
<td>0.39 ± 0.03</td>
</tr>
<tr>
<td>TOTAL</td>
<td>20.8</td>
<td>5.99 ± 0.10</td>
<td>14.8 ± 0.10</td>
</tr>
</tbody>
</table>

Table 6.14: The results ($\times 10^3$) for the fitted # tops and #W on MC08, containing $t\bar{t}$ generated by AcerMC. These values are obtained by using all available events and scaling the samples to 146 pb$^{-1}$; the statistical uncertainties are thus negligible compared to those in Table 6.12. The $R_{nj}^{b\text{jet}}$ fractions from MC@NLO are used in the likelihood functions.

For a determination of the cross sections in the Mixed sample we do not have to take into account any of the systematic uncertainties mentioned in Section 6.2, as we have used the same sub-samples in the Mixed sample and in MC08, except for the $t\bar{t}$ events: we used the $t\bar{t}$ from MC@NLO instead of the $t\bar{t}$ from AcerMC.

If we apply the analysis on MC08 with $t\bar{t}$ events from AcerMC, while keeping the $R_{nj}^{b\text{jet}}$ fractions from the MC@NLO sample, we can immediately compare the new results (summarized in Table 6.14) with Table 6.12. The end result is

$$
\sigma_{\text{Mix}}^{W} = (1.19 \pm 0.02) \cdot \sigma_{\text{MC}}^{W} \\
\sigma_{\text{Mix}}^{t\bar{t}} = (1.00 \pm 0.10) \cdot \sigma_{\text{MC}}^{t\bar{t}} \\
\sigma_{\text{Mix}}^{\text{singleTops}} = (0.90 \pm 0.47) \cdot \sigma_{\text{MC}}^{\text{singleTops}}.
$$

(6.13)

The b-tagging efficiencies

The fit results for the b-tagging efficiencies in Fig. 6.8 are as expected: the $\epsilon_b$ is fitted as 0.63 ± 0.02 and corresponds with the true value of 0.63 obtained from MC@NLO $t\bar{t}$ events, see Table 6.5. The $\epsilon_{t\bar{t}}^{\text{top}}$ is fitted as 0.02 ± 0.01 and is lower than the true value of 0.05 obtained from MC@NLO $t\bar{t}$ events. This last discrepancy is not the result of using different MC generators, but a bias we observed earlier.

6.3.2 Discussion

To emphasize the importance of accurate MC predictions, we performed the analysis with MC@NLO $t\bar{t}$ events and with Alpgen $t\bar{t}$ events. The results in eq. (6.12) and eq. (6.13) show a large difference, in effect resulting in a significant systematic uncertainty in the analysis. In the next section we will quantify this more accurately.

The correct values for the cross sections in this ‘blind’ sample study are obtained with the Alpgen sample, that is the values in eq. (6.13). The study shows that a variation of the cross sections can be measured; the W cross section is measured with the highest accuracy.
6.4 Early data analysis

In this section we study the possible results on a first data sample of 50 pb$^{-1}$. For clarity, the selection of events with electrons is again applied and the $W + b\bar{b}$ sample is again included in our MC set of samples.

JetProb is the recommended b-tagger for first data. It is a much simpler and more robust algorithm, with the drawback that its performance is worse compared to a (calibrated) IP3D+SV1, see Fig. 4.12. We observe this again in Fig. 6.9 which shows the results of the analysis performed on a sample of 500 pb$^{-1}$. On the horizontal axis $-\log_{10}P$ is plotted, where $P$ is the probability output of the b-tagger. The probability range is chosen such that the same $\epsilon_b$ range is covered as was the case with the IP3D+SV1 b-tagger. Compared to Fig. 6.7 where the latter is used, the results are worse: the efficiencies show more fluctuation and the distributions for the ratios have less clear plateaus.

Looking at the results for the four-jet events, which have the largest uncertainties, we conclude that with the JetProb b-tagger the best weight to use is in the range $0.4 - 0.7$. Within the given fit uncertainties the results are in agreement with the expected values from Table 6.4 (depicted with the dotted lines); only the results in the four jet sample in Fig. 6.9(d) are more than ‘1-sigma’ off. Figure 6.10 depicts the results on a sample of no more than 50 pb$^{-1}$. These results do show a correct fit for the ratio in the four jet sample.

6.4.1 Excluding the four-jet sample

In Fig. 6.9(d) the results in the four jet sample show a more than ‘1 sigma’ offset compared with the expected value from Table 6.4, and in Fig. 6.10(d) the results in the four jet sample show a large uncertainty; it is worth checking whether the measurements in the four jet sample improve the analysis, or worse, whether the measurements do not bias the results in the two and three jet sample. As the latter are more important, the analysis can be repeated excluding the measurements on the four jet sample. This implies only five measurements are used in the likelihood discussed in Section 6.1.1 to solve four unknowns: the two efficiencies, and the number of top events in the two- and three jet samples.

We performed such an analysis on a 50 pb$^{-1}$ sample and the conclusion is that no difference is visible with the previous results; using larger samples or the IP3D+SV1 b-tagger does not alter this conclusion. The three extra measurements in the four jet sample thus give too little information to influence the fit results in the two and three jet sample, yet can be used to estimate the number of tops and $W$’s in the four jet sample.
6.4. Early data analysis

Figure 6.9: Fit result for a sample with $L = 500 \text{ pb}^{-1}$. JetProb is used as $b$-tagger. Compared to the results in Fig. 6.7 with IP3D+SV1 as $b$-tagger, these results show more fluctuations. Also, the fitted ratios show larger deviations from the true values (from Table 6.4), depicted by the dotted lines.

(a) Left: results for $\epsilon_{b}$. Right: results for $\epsilon_{t}$ (top). The distribution with error-bars is the fit result, the distribution without is the true efficiency for the used sample.

(b) Ratio of $top/total$ for events with 2 jets.

(c) Ratio of $top/total$ for events with 3 jets.

(d) Ratio of $top/total$ for events with 4 jets.
Chapter 6. Top quark and $W$ boson production cross section measurement

Figure 6.10: Fit result for a sample with integrated luminosity of 50 pb$^{-1}$. JetProb is used as $b$-tagger. The correct values for the ratio fits are depicted by dotted lines.
6.4.2 Uncertainty on the b-tagging efficiency on W+jets

Before extracting the cross sections from the results on the 50 pb⁻¹ sample, the uncertainty on the b-tagging efficiency on W+jets still has to be taken into account. In Section 6.1.4 an upper and lower limit on \( \varepsilon^W_b \) was estimated by taking into account the difference between the W- and Z+jets cross sections, plus the statistical fluctuations expected from a 50 pb⁻¹ sample of Z+jets events. The results for the upper and lower limits were shown in Fig. 6.6.

Repeating the analysis on a first data sample of 50 pb⁻¹ with these extreme distributions for \( \varepsilon^W_b \) the fit results at a JetProb weight of 0.5 in Table 6.15 are obtained. We observe that compared to the ‘nominal’ results in the distributions in Fig. 6.10 a too low value for \( \varepsilon^W_b \) increases the fitted number of top events, while at the same time decreases the \( \varepsilon^W_b \). A too high value for \( \varepsilon^W_b \) has the opposite effect.

We conclude that the uncertainty on \( \varepsilon^W_b \) results in errors on the fitted \#top/#total with sizes of 9%, 12% and 13%, respectively for the two, three and four jet sample. For the b-tagging efficiencies the systematic uncertainty is ±0.05 for \( \varepsilon^W_b \) and ±0.01 for \( \varepsilon^W_{l\text{top}} \).

<table>
<thead>
<tr>
<th>Fitted #top/#total:</th>
<th>With lower limit ( \varepsilon^W_b )</th>
<th>With correct ( \varepsilon^W_b )</th>
<th>With upper limit ( \varepsilon^W_b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 jet sample</td>
<td>0.25 ± 0.13</td>
<td>0.17 ± 0.10</td>
<td>0.09 ± 0.05</td>
</tr>
<tr>
<td>3 jet sample</td>
<td>0.55 ± 0.20</td>
<td>0.44 ± 0.18</td>
<td>0.32 ± 0.15</td>
</tr>
<tr>
<td>4 jet sample</td>
<td>0.84 ± 0.16</td>
<td>0.71 ± 0.24</td>
<td>0.59 ± 0.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fitted efficiencies:</th>
<th>With lower limit ( \varepsilon^W_b )</th>
<th>With correct ( \varepsilon^W_b )</th>
<th>With upper limit ( \varepsilon^W_b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon^W_b ) (true = 0.36)</td>
<td>0.30 ± 0.15</td>
<td>0.36 ± 0.15</td>
<td>0.40 ± 0.15</td>
</tr>
<tr>
<td>( \varepsilon^W_{b\text{top}} ) (true = 0.07)</td>
<td>0.07 ± 0.10</td>
<td>0.06 ± 0.06</td>
<td>0.05 ± 0.06</td>
</tr>
</tbody>
</table>

Table 6.15: The fit results from the analyses with different distributions of the efficiency \( \varepsilon^W_b \) on a sample of 50 pb⁻¹, at a JetProb weight of 0.5.

6.4.3 Measurement of the cross sections

In Section 6.3 the cross section for W+jets, single tops and \( t\bar{t} \) is calculated with the use of six values: the measured number of top events in the two- and three jet sample, \( N^{\text{top}}_{MC,2j} \) and \( N^{\text{top}}_{MC,3j} \), plus the total number of W events measured, \( N^W \). These are compared with the three expected values \( N^{\text{top}}_{MC,2j} \) and \( N^{\text{top}}_{MC,3j} \) from a MC sample.

From Section 6.2.3 we conclude that the measurements for \( N^{\text{top}}_{MC,2j} \) and \( N^{\text{top}}_{MC,3j} \) can be made to respectively 1% and 5%. To stay on the safe side, we also assume an uncertainty of 5% on \( N^{W}_{MC} \). Adding the uncertainties on the lepton trigger and reconstruction⁷, we summarize the uncertainty on the MC predictions as in Table 6.16. As for the results obtained from the fits, Table 6.17 gives an overview of the uncertainties measured, and their uncorrelated total sum per variable. The uncertainty on the luminosity is listed with its lower and upper limit.

With the maximum uncertainties on \( N^{W}_{MC} \) and \( N^W \) in Table 6.17 we conclude that the cross section for the W+jets can be measured up to 32% with first data. The total cross section for

⁷In Section 4.6.1 we stated these uncertainties, which were gained from studies performed on a sample of 100 pb⁻¹ 14 TeV collisions. As we study the performance on a smaller sample at lower center-of-mass energy, these quoted uncertainties are not entirely correct. Yet these uncertainties are almost negligible compared to the total; we will therefore leave them unchanged.
top quark pairs plus single tops can be measured to 50% combining the results in the two and three jet sample; if we were to follow similar steps as taken in Section 6.3 to obtain the cross sections in eq.(6.13) for the individual channels, the uncertainties on the cross sections for $t\bar{t}$ and for single top would be 110% and 580% respectively.

### 6.4.4 Conclusion

We have shown that for a first data sample of 50 pb$^{-1}$ a measurement up to 32% accuracy can be made for the $W$+jets background cross section to the top channels in the lower jet multiplicities. An individual cross section measurement for the top quark pair and for the single top events results in large uncertainties, respectively 110% and 580%, yet a total cross section for the top quark pair plus single top events can be determined up to 50% accuracy.

Compared to a dedicated $W$ or top cross section analysis the advantage is that the presented analysis is independent of the b-tagging efficiency, which is in fact determined simultaneously with the cross sections. The JetProb efficiency $\varepsilon_b$ on b-jets is measured with a statistical uncertainty of ±0.15 and a systematic uncertainty of ±0.05, both of which are correlated to the uncertainties on the cross section measurements.

The uncertainty on the measurement of the b-tagging efficiency on $W$+jets events was estimated from the higher expected value for the cross section ratio of $Z+b\bar{b}$ to $Z$+jets compared to that of $W+b\bar{b}$ to $W$+jets. If the first ratio is truly higher, we can expect a positive offset in the measurement of the efficiency $\varepsilon^W_b$, biasing our end results. Further research using simulated $Z+b\bar{b}$ events is needed to investigate the exact offset and to explore possibilities to compensate for the bias. The same conclusion can be drawn on the subject of QCD events. With only a small fraction of $b\bar{b}$ in QCD events, we expect these events to be $W$+jets like. Using simulated multi-jet events the expected number of QCD events passing the selection requirements, and their effect on the analysis presented, should be investigated.
6.5 The effect of supersymmetry on the top/$W$ ratio

In this section we study the effect of the presence of supersymmetric particles on the analysis set up in this chapter. As in section 5.9 we study the SU3 (bulk) and SU4 (low mass) models.

6.5.1 Event selection

We first study the effect of the presence of SU4 with the standard event selection listed in Section 6.1.2. We then perform the analysis with higher requirements on the $E_T$ and transverse momentum of the jets as to increase the signal to background ratio, where the signal in this case is the events with supersymmetric particles. We have not performed any research to define the best requirements, yet based on the work done in [60] we have chosen for:

- Missing transverse energy $E_T > 100$ GeV.
- Exactly one isolated lepton, be it electron or muon, with $p_T > 20$ GeV. At the same time, the Event Filter trigger $\text{EF}_e25\_\text{medium1}$ must have passed if the lepton is an electron, or the trigger $\text{EF}_\mu20$ must have passed if it is a muon.
- At least one jet with $p_T > 100$ GeV. Any additional jet must have $p_T > 50$ GeV.

With these new criteria we analyze the effect of two supersymmetric models. The samples analyzed are reconstructed with a center-of-mass energy of 10 TeV and are the SU3 and SU4 samples generated with Isajet [118] plus HERWIG, see Section 4.2. We note that for the SU4 sample we only had access to a sample corresponding to 169 pb$^{-1}$ and we normalized the results in this section to 500 pb$^{-1}$.

6.5.2 Results

The requirements mentioned above on the missing transverse energy and the jet $p_T$ could affect the results of the fit procedure and introduce biases in the events selected. Before including the supersymmetric channels we checked the results of the scan procedure on the Standard Model events with these requirements. The fit results are still correct: the distributions for the fitted number of top events are flat and at the correct values. Also the fitted efficiencies correspond to the ‘true’ values.

The SU4 model is one of the models with the lowest masses for the supersymmetric particles and its cross section is relatively high: 402 pb [60]. Table 6.18 lists the results for the fitted number of top and $W$ events using the standard selection requirements from Section 6.1.2 on a sample of 500 pb$^{-1}$ including the SU4 supersymmetric events. The columns to the left in the Table list the number of events for each channel passing the selection; the number of events for the Standard Model channels are the same as in Table 6.4. The values for the fitted number of tops and number of $W$'s are extracted from the results in Fig. 6.11 at a IP3D+SV1 weight of 4.0.

Within the statistical uncertainty the measured number of $W$ events corresponds to the actual number of $W$ events present; only in the two jet sample is the measured number of $W$ events slightly too high. The measured number of tops actually equals the sum of top and supersymmetry events. In Section 6.4.2 a too high number of $W$'s was correlated with a too high fit result for $\varepsilon_b$. This is not the case here: the fitted b-tagging efficiency $\varepsilon_b$ is too low.
With the flat distributions for the ratios in Fig. 6.11 and the efficiencies having only a small offset from the true distributions we can conclude that the SU4 events resemble the top events in their b-jet contents. This is as we might expect from the phenomenology of SU4 events, see [124]:

- 63% of the SU4 events produced contain at least one gluino. The gluino is dominantly decaying to the third generation $\tilde{g}$, with
  - $\tilde{g} \rightarrow \tilde{b}_1 b$ (branching ratio = 47%),
  - $\tilde{g} \rightarrow \tilde{t}_1 t$ (branching ratio = 42%),
  - $\tilde{g} \rightarrow \tilde{b}_2 b$ (branching ratio = 4%).

The squarks $\tilde{b}_1$ and $\tilde{b}_2$ decay to $\tilde{t}_1 W$ with a probability of about 50%. The lightest scalar top $\tilde{t}_1$ is only 30 GeV heavier than its Standard Model partner ($m_{\tilde{t}_1} = 206$ GeV) and decays to final states similar to those of the top quark:

- $\tilde{t}_1 \rightarrow \tilde{\chi}^{\pm}_1 b$ (branching ratio = 100%). The $\tilde{\chi}^{\pm}_1$ decays through a virtual $W^\pm$ and a $\tilde{\chi}^0_1$.

- 11% of the SU4 events produced in the hard scattering result in a scalar top pair $\tilde{t}_1 \tilde{t}_1$.

All in all, a large fraction of SU4 events (53%) is characterized by a decay through at least two (virtual) $W$-bosons and two b-quarks, with as result a $t\bar{t}$-like final state.

In Table 6.19 we summarize the results after the selection requirements mentioned in Section 6.5.1. The results for all five distributions are shown in Fig. 6.12. We can almost draw the same conclusion as before: for the three- and four-jet samples the measured number of $W$ events is correct and all supersymmetry events are measured as top events. In the two-jet sample the number of $W$'s is too high however.

Repeating the last analysis with the high energetic requirements on a sample containing SU3 instead of SU4 we obtain the results from Table 6.20 and Fig. 6.13. These measurements are inconclusive: the small amount of supersymmetry events does not make it possible to see any significant effect on the results.

### 6.5.3 Conclusion

Combining the results of this section with those of Section 5.9 from the previous chapter we can conclude that the SU4 events resemble the top events in their b-jet contents, while at the same time they do not satisfy a complete semi-leptonic $t\bar{t}$ hypothesis. In the analysis in Section 5.9 on SU3 events an excess of events was visible after the tight selection cuts. With the low jet multiplicity demanded in this section the excess is not clear.

With the large parameter space available, supersymmetric models with other b-jet contents than those in the SU3 and SU4 models analyzed in this section are possible; we cannot assume that all models will be easily identified as $t\bar{t}$ like, as was the case for the SU4 model, or $W+$jets like. We can only conclude that if supersymmetry exists, with some luck, the $b$-tagging analysis presented in this chapter might give a hint of the composition of the supersymmetric events.
6.5. The effect of supersymmetry on the top/W ratio

<table>
<thead>
<tr>
<th></th>
<th>$t\bar{t}$ + $W+$jets</th>
<th>SU4 Fitted</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 jet sample</strong></td>
<td>14.9</td>
<td>70.7</td>
<td>3.98</td>
</tr>
<tr>
<td><strong>3 jet sample</strong></td>
<td>12.3</td>
<td>12.3</td>
<td>4.74</td>
</tr>
<tr>
<td><strong>4 jet sample</strong></td>
<td>5.23</td>
<td>2.23</td>
<td>4.05</td>
</tr>
</tbody>
</table>

Table 6.18: Results for the fitted number of top and W events ($\times 10^3$) using the standard selection requirements from Section 6.1.2 on a sample of 500 pb$^{-1}$ including the supersymmetric SU4 model. To the left the number of events are listed for the different samples present. Within the statistical uncertainty the fitted number of W events corresponds to the actual number of W events present; the fitted number of tops equals the sum of top and supersymmetry events.

<table>
<thead>
<tr>
<th></th>
<th>$t\bar{t}$ + $W+$jets</th>
<th>SU4 Fitted</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 jet sample</strong></td>
<td>1.60</td>
<td>4.43</td>
<td>3.21</td>
</tr>
<tr>
<td><strong>3 jet sample</strong></td>
<td>1.23</td>
<td>1.19</td>
<td>3.11</td>
</tr>
<tr>
<td><strong>4 jet sample</strong></td>
<td>0.47</td>
<td>0.21</td>
<td>1.86</td>
</tr>
</tbody>
</table>

Table 6.19: Results ($\times 10^3$) using the supersymmetry selection requirements on a sample of 500 pb$^{-1}$ including SU4 events. Almost the same conclusion can be drawn as in Table 6.18: the fitted number of W events is correct in the three- and four-jet sample and all supersymmetry events are measured as top events. Yet in the two-jet sample the fitted number of W events is too high.

<table>
<thead>
<tr>
<th></th>
<th>$t\bar{t}$ + $W+$jets</th>
<th>SU3 Fitted</th>
<th>Fitted #W</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 jet sample</strong></td>
<td>1.60</td>
<td>4.43</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>3 jet sample</strong></td>
<td>1.23</td>
<td>1.19</td>
<td>0.27</td>
</tr>
<tr>
<td><strong>4 jet sample</strong></td>
<td>0.47</td>
<td>0.21</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Table 6.20: Results ($\times 10^3$) using the supersymmetry selection requirements on a sample of 500 pb$^{-1}$ including SU3 events. Compared to the results with SU4 events, it is unclear whether the supersymmetry events are fitted as top or as W events. The small amount of supersymmetry events does not make it possible to see any significant effect on the results.
Figure 6.11: Fit result for a sample containing the supersymmetry events SU4, with luminosity of 500 pb$^{-1}$. The standard event selections from Section 6.1.2 are used, while IP3D+SV1 is used as b-tagger.
6.5. The effect of supersymmetry on the top/$W$ ratio

Figure 6.12: Fit result for a sample containing the supersymmetry events SU4, with luminosity of 500 pb$^{-1}$. The supersymmetry requirements mentioned in Section 6.5.1 are used, while IP3D+SV1 is used as b-tagger. The $R_i^{bj}$ fractions are extracted from the MC@NLO $t\bar{t}$ events, after the event selection.
(a) Left: results for $\varepsilon_\text{b}$. Right: results for $\varepsilon_\text{t}^{\text{top}}$. The distribution with error-bars is the fit result, the distribution without is the true efficiency when using $t\bar{t}$ events generated by MC@NLO.

(b) Ratio of $\text{top}/\text{total}$ for events with 2 jets.

(c) Ratio of $\text{top}/\text{total}$ for events with 3 jets.

(d) Ratio of $\text{top}/\text{total}$ for events with 4 jets.

Figure 6.13: Fit result for a sample containing the supersymmetry events SU3, with luminosity of 500 pb$^{-1}$. The susy requirements mentioned in Section 6.5.1 are used, while IP3D+SV1 is used as b-tagger. The $R_n^{\text{jet}}$ fractions are extracted from the MC@NLO $t\bar{t}$ events, after the event selection.