CDF/ANAL/TOP/CDFR/7701
Version 3.00, January 25, 2006

Validation of the MadEvent Single-Top Samples with ZTOP NLO Calculations

Sarah Budd[b], Catalin Ciobanu[b], Thomas Junk[b], Yves Kemp[a], Jan Lück[a], Thomas Müller[a], Svenja Richter[a], Jeannine Wagner[a], Wolfgang Wagner[a], Thorsten Walter[a]

[a] *Universität Karlsruhe*
[b] *University of Illinois, Urbana-Champaign*

Abstract

We validate the s- and t-channel single-top-quark production Monte Carlo samples used in the CDF Run II single-top search by comparing to next-to-leading-order calculations. We see good agreement in the rates and in the shapes of transverse momentum (p_T) and pseudorapidity (η) distributions of leading and 2^{nd}-leading jets. In addition, we estimate the systematic uncertainty on the Monte Carlo model.

1 Introduction

Besides the top quark pair production via the strong interaction, discovered at the Fermilab Tevatron Collider in 1995 by the CDF and DØ collaborations [1, 2], the Standard Model (SM) also predicts the electroweak production of single-top-quarks, which has not yet been observed. The discovery of single-top production and the subsequent precise measurement of its cross section are main goals of the collaborations at the Fermilab Tevatron Collider. In Run I (1992-95) of the Tevatron several upper limits on single-top production cross sections were set by the CDF [3, 4] and DØ [5, 6] experiments. The Run I results have been recently surpassed by the latest Run II measurements [7, 8]. A comprehensive review of top quark physics can be found in reference [9].

In proton-antiproton ($p\bar{p}$) collisions at the Tevatron, two production modes for single-top are dominating: the t-channel process (Figure 1 (a)) with a total NLO cross section of $1.98^{+0.28}_{-0.22}$ pb and the s-channel process (Figure 1 (b)), also called W^* production (0.88 ± 0.11 pb at NLO) [10, 11]. In $p\bar{p}$ collisions the third electroweak production mode, the associated production (Figure 1 (c)), known as Wt production, has by comparison a small cross section of $0.094^{+0.015}_{-0.012}$ pb (no full NLO calculation) [12]. All indicated cross sections are calculated with a top mass of 175 GeV/c^2.

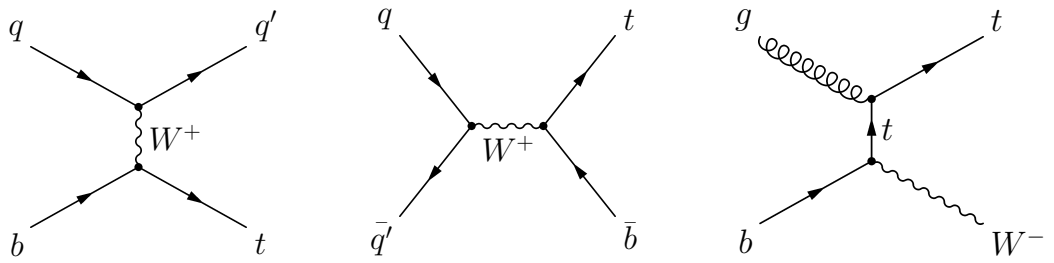


Figure 1: Leading-order (LO) Feynman diagrams of single-top production modes: t-channel (a), s-channel (b) and associated production (c).

A good understanding of the signal characteristics of the relevant single-top production modes is indispensable for further analyses. For the s-channel production mode, the available Monte Carlo (MC) generators are in good agreement with the next-to-leading-order (NLO) calculations. The t-channel production mode has one important NLO contribution which is not modeled by the present leading-order (LO) MC generators. This leads to the fact that there is yet no generator which produces an appropriate signal sample. The Tevatron Run I analyses used the PYTHIA MC generator [13], knowing that it is not accurate in some aspects. As a workaround a matching procedure of two matrix element MC samples was proposed [14] to compensate for the inaccuracy. In the recent Run II single-top analysis the CDF Collaboration used besides the LO sample an additional sample, which models the important NLO contribution. In the following this will be called NLO sample. Both are produced by the MadEvent MC generator [15, 16], showered by PYTHIA and matched in such a way as to achieve an improved modeling [17].

In this note, the matching of the t-channel samples is optimized. This is done by comparing the distribution, which has to be matched, to the prediction of a next-to-leading-order calculation, provided by the ZTOP software [11]. The occurring differences are minimized by changing the matching parameters, especially the fraction of the NLO sample. The s- and newly matched t-channel signal samples are then validated by means of the ZTOP NLO calculations. The resulting deviations are used to estimate the systematic uncertainty on the Monte Carlo model.
The note is organized as follows. In Section 2 the underlying theory of leading- and next-to-leading-order calculations of t-channel single-top production is briefly discussed. The software, which produced the NLO distributions, namely ZTOP, along with the relevant parameters is described in Section 3. In Section 4 we give a short description of the MadEvent MC samples and we illustrate the matching of the t-channel signal samples. The validation of the samples, presented in Section 5, is followed by Section 6, which covers the estimate of the systematic uncertainty on the signal modeling. The note concludes in Section 7 with suggestions for improvements of the single-top MC production.

2 Next-to-Leading-Order Calculations

NLO calculations are essential to compare theory with experimental results. The most important NLO correction to the t-channel leading-order process shown in Figure 1 (a) is the 2→3 process, which is known as W-gluon fusion, where an initial gluon splits into a $b\bar{b}$ pair (Figure 2 (a)).

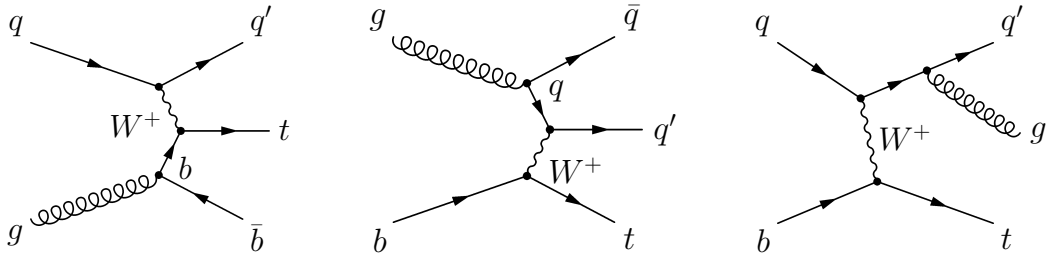


Figure 2: Some NLO Feynman diagrams of t-channel single-top production: W-gluon fusion (a), initial state gluon splitting (b) and gluon radiation (c)

If the b quark is considered massless in the computation of the 2→3 matrix element, the gluon splits into a real $b\bar{b}$ pair with the final state $\bar{b}$ quark (in the following called 2^{nd}-b quark in order to distinguish from the b quark coming from the decay of the top quark) being collinear with the incoming gluon. Given that the internal b quark is on-shell, its propagator is infinite and the Feynman diagram becomes singular. As in reality the b quark is not massless, the mass m_b regulates the collinear singularity, which is described by terms of $ln[(Q^2 + m_t^2)/m_b^2]$, where Q^2 is the virtuality of the W boson. The W-gluon fusion cross section contains these logarithmic terms of order $ln^n[(Q^2 + m_t^2)/m_b^2]/n!$ at every order n of the perturbative expansion in the strong coupling due to the collinear emission of gluons from the internal b quark propagator. This leads to the fact that, since the logarithms are large, the perturbation series does not converge quickly.

However, by introducing a b quark distribution function, one can sum up all collinear logarithms and bypass the convergence difficulty. Due to this method the cross section of the process in Figure 1 (a) is of order $ln[(Q^2 + m_t^2)/m_b^2]$, the diagram in Figure 2 (a) also contains these terms, even though they are already summed into the b quark distribution function. So, to avoid double counting, one needs to remove these terms. This subtraction method, carried out in reference [18], achieves the same results within the errors as the phase space slicing method [10], upon which the ZTOP software is based. This analytic form of NLO cross section calculation is fully differential, therefore experimental cuts can be implemented.

3 ZTOP Software

As mentioned above, the ZTOP software, a program to calculate NLO s-channel and t-channel single-top-quark production distributions, is based on the phase space slicing method, namely on a spin-averaged version. Since it provides the possibility to calculate transverse momentum and pseudorapidity distributions within the geometrical acceptance of a given detector, it is very appropriate to verify event samples of MC generators. The method of ZTOP is fully differential and produces final state jet and not parton distributions. This is a consequence of the formalism of the NLO calculations. Instead of flavor-ordered partons one has to compare p_T-ordered jets. Since b quark jet distributions differ from those of light quark jets, ZTOP distinguishes between jets with and without b quark content. This b quark is always the already mentioned 2^{nd}-b quark, because ZTOP does not cover the decay of the top quark. Further contributions to light quark jets may come from initial- or final-state gluon radiation, for instance Figure 2 (c). As a result of all these matters, ZTOP produces the following relevant distributions: p_T and η of the top quark at parton level, as well as of the p_T-ordered leading jet and of the 2^{nd}-leading jet. There are separate distributions for jets with or without b quark content.

To produce the NLO distributions, we used ZTOP version 1.0 with the default parameters, except for minor adjustments: the W boson mass was set to 80.42 GeV/c^2, the top quark mass was set to 175 GeV/c^2 and the k_T cone size was set to 0.54, which corresponds to a fixed cone size of 0.4 used by CDF. To get better statistics, the initial number of Vegas iterations and points were set to the values recommended for graphing. The parton density functions were set to CTEQ5M1. Only top (no antitop) quark production was performed. Thus, the cross section was doubled to match the predictions for top and antitop production. Since the antitop distributions in $p\bar{p}$ collisions differ solely by the signs of η, we plot $Q_l \cdot \eta$ distributions when comparing with MadEvent, which contains top and antitop events. Here Q_l is the charge of the lepton coming indirectly from the top quark decay (+1 for top, -1 for antitop quark production in units of the elementary charge).

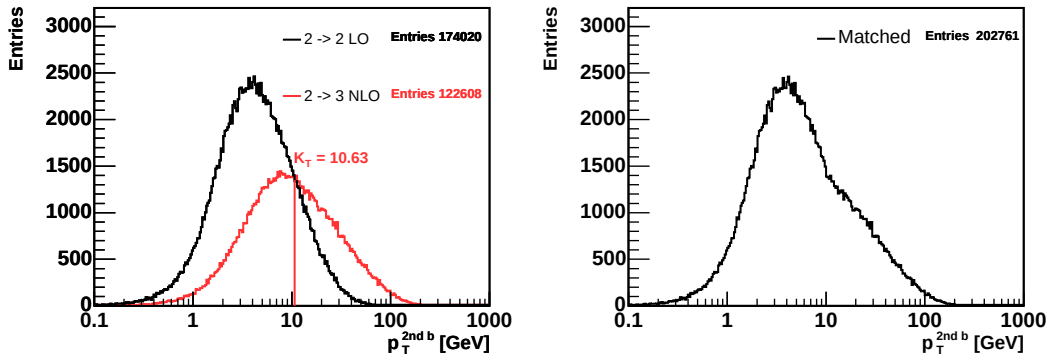
4 MadEvent Signal Samples

The CDF Run I single-top analysis used PYTHIA to generate s- and t-channel MC events. For the t-channel it is known that PYTHIA generates too soft and too far forward distributed 2^{nd}-b quarks. The reason for that is that PYTHIA starts with the LO 2→2 diagram (see Figure 1 (a)), i.e. with a b quark PDF and then creates the initial state through backward evolution (DGLAP scheme [19, 20, 21]). Using this method, only the soft region of the transverse momentum of the 2^{nd}-b quark is well modeled, while the hard region is underestimated. To bypass this problem, the CDF Run II single-top analysis chose MadEvent as MC generator, which brings along two advantages:
First, it provides the opportunity to generate two independent t-channel samples, the 2→2 LO process with a b quark PDF to cover the soft p_T range and another 2→3 NLO process (see Figure 2 (a)) with an initial state gluon splitting into a $b\bar{b}$ pair for the hard range. Both processes differ in the number of final state partons: the 2→2 matrix element includes the light quark and the decay products of the top quark, namely a lepton, a neutrino and the so-called 1^{st}-b quark in the final state. The 2→3 matrix element includes the same final state partons plus an additional $\bar{b}$ quark, the already mentioned 2^{nd}-b quark.
Second, MadEvent fully incorporates the spin of the top quark in contrast to the PYTHIA generator. One interesting feature of electroweak top quark production is that the top quarks are produced 100% polarized along the direction of the down-type quark (q' in Figure 1 (a), (b) and 2 (a)) in the top quark rest frame [22, 23, 24]. It is important to include this feature in the Monte Carlo description since it can be used to discriminate single-top-quark events from background.
Because MadEvent is designed to produce events at parton level, one needs a parton showering software to compute all desired final state particles and hadrons. For this purpose PYXTRA was used, a software interface that passes on the MadEvent output to PYTHIA, where a strong-angular-ordered showering is done through emission of QCD radiation.

4.1 Cross Section based Matching

For the upcoming CDF single-top analysis, the s- and t-channel MadEvent MC sample production at the end of 2004 was conducted with a top quark mass set to 178 GeV/c^2, corresponding to the then valid world average. As already addressed, two independent t-channel samples were generated.
A matching of the two MC samples was conducted by the single-top group in such a manner, that an event was accepted, if simultaneously both the p_T of the 2^{nd}-b quark was lower (higher) than a fixed threshold K_T and the event was from the 2→2 LO (2→3 NLO) process. By this procedure, as described in [14], double counting of events from the same phase space is avoided. One obtains a combined t-channel single-top MC sample with a p_T-spectrum of the 2^{nd}-b quark valid in both the soft and the hard region and with a continuous transition between them. The threshold K_T is defined as the intersection of the LO and NLO p_T distributions. To calculate this intersection point, the rates of the LO and NLO distributions were normalized to their respective cross sections obtained from MadEvent [17].

According to the MadEvent calculations, the ratio of the two cross sections was $R = \sigma_{LO}/\sigma_{NLO} = 1.42$. Using this value the intersection point K_T was about 10 GeV (see Figure 3 (a)). 144445 events out of the 2→2 sample and 58316 events of the 2→3 sample were merged into the matched MadEvent sample, whose resulting p_T distribution is shown in Figure 3 (b).



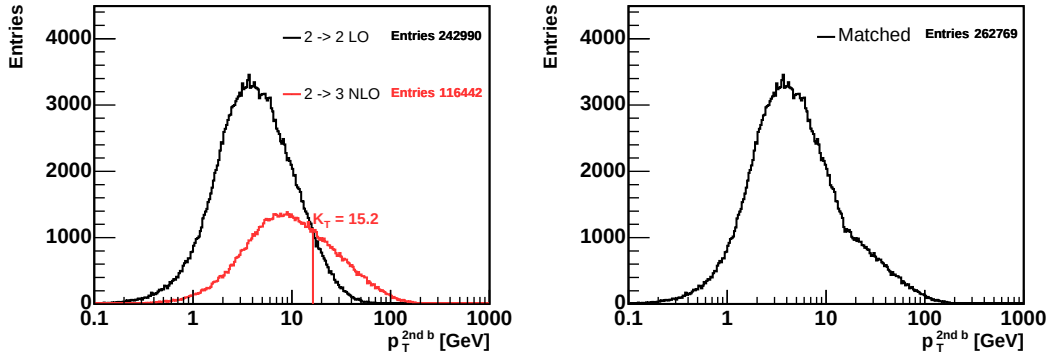
(a) (b)

Figure 3: Matching procedure of the $m_t = 178$ GeV/c^2 sample in log_{10} scale: 2^{nd}-b quark p_T of unmatched MadEvent 2→2 LO and 2→3 NLO samples with intersection at about 10 GeV (a), matched sample (b).

4.2 Slope based Matching

During 2005, new top-mass analyses with improved precision were conducted at CDF and DØ. They made clear, that the next, not yet published, world average of the top-mass will move towards a lower value. Thus, the CDF Collaboration decided to use a top mass of 175 GeV/c^2 for all upcoming analyses. Due to this, new samples had to be produced from the CDF single-top group. In the 178 GeV/c^2 t-channel sample, the rate of the 2^{nd}-b quark jet was found to be too low when comparing to NLO calculations. Since the b quarks contributing to these jets are all from the 2→3 MadEvent sample, which was, compared to the 2→2 sample, scaled down by the cross section ratio $R = \sigma_{LO}/\sigma_{NLO}$ and merged into the matched signal sample, one could improve the fraction between the 2→2 and 2→3 events by introducing a new matching procedure for the new generation. Such a new method is proposed in reference [17].

Instead of using the cross section ratio R, whose physical meaning is ambiguous, the single-top group used the minimum of the differences of the slopes of the two 2^{nd}-b quark p_T distributions at their intersection point K_T in dependence of the ratio R, to find the best mixing ratio, which leads to a smooth transition. The resulting distribution with an intersection at $K_T = 15.2$ GeV and a cross section ratio of $R = \sigma_{LO}/\sigma_{NLO} = 2.09$ is shown in Figure 4 (a), the matched sample in (b).



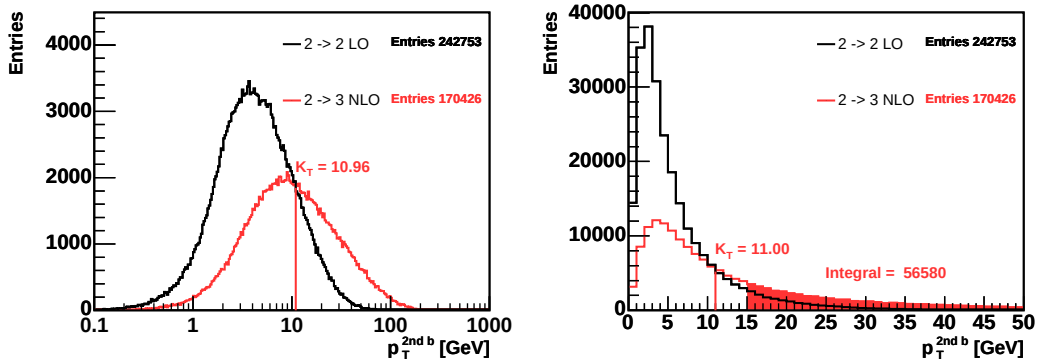
(a) (b)

Figure 4: Matching with slope minimization of the $m_t = 175$ GeV/c^2 sample in log_{10} scale: 2^{nd}-b quark p_T of unmatched MadEvent 2→2 LO and 2→3 NLO samples with intersection at 15.2 GeV (a), matched sample (b).

4.3 Fraction based Matching

It is obvious from Figure 4, that the new fraction of 2^{nd}-b quarks from the hard region decreases compared to the 178 GeV/c^2 sample, where we found that the rate of the hard 2^{nd}-b quark jets was too low. Therefore a third matching procedure, proposed in reference [11], is conducted. The underlying concept is to vary the ratio $R = \sigma_{LO}/\sigma_{NLO}$, more precisely the fraction of 2→3 events, until the rate of 2^{nd}-b quark jets from the hard region comply to the NLO predictions of the ZTOP software.

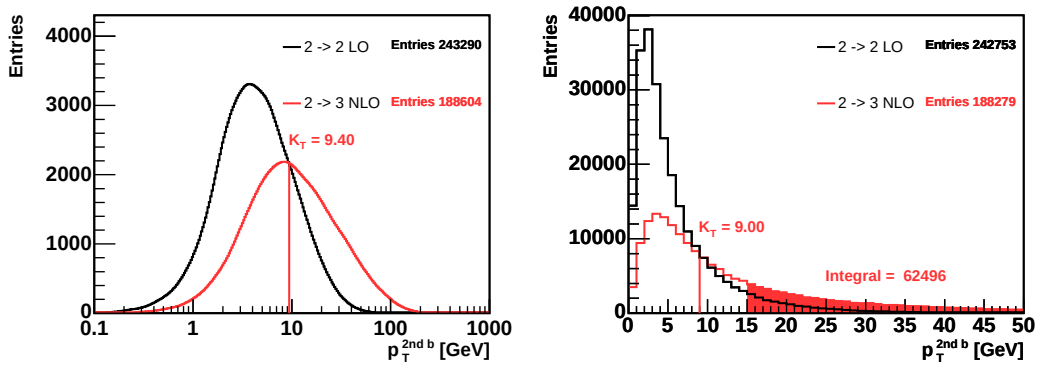
In a first pass, we choose a ratio of $R = 1.4$ (about the same as for the 178 GeV/c^2 sample) to be the working point for further estimates. Figure 5 shows the associated intersection point at $K_T = 11$ GeV.



(a) (b)

Figure 5: Intermediate step of the matching of the $m_t = 175$ GeV/c^2 sample: 2^{nd}-b quark p_T of unmatched MadEvent 2→2 LO and 2→3 NLO samples (in log_{10} (a) and linear (b) scale) with intersection at 11 GeV.

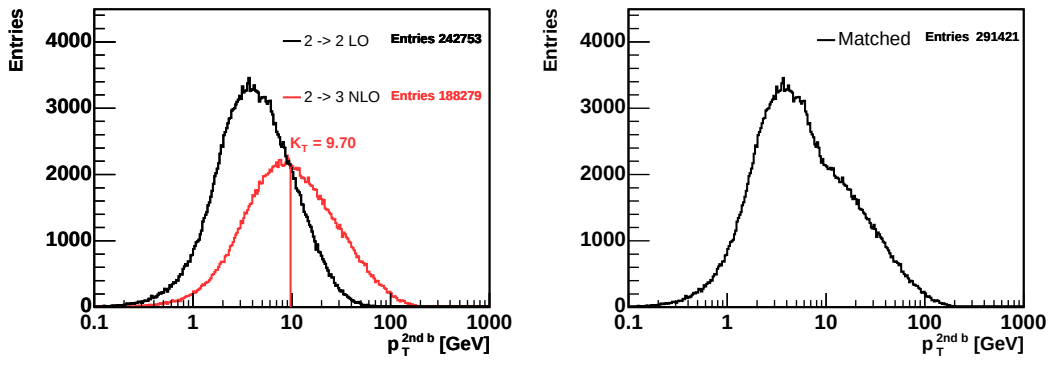
In the second pass, a new R ratio is estimated, which should meet the following prediction of the NLO calculation: within the detector cuts, a 2^{nd}-b quark jet appears in 21.1% of all events. The detector cuts, motivated by the single-top analysis, are $p_T > 15$ GeV and $|\eta| < 2.8$. To fulfill this NLO prediction, the MadEvent $R = 1.4$ sample should have about 59800 visible 2^{nd}-b quarks out of the total of 283410 events. As highlighted in Figure 5 (b), 56580 events include a 2^{nd}-b quark with $p_T > 15$ GeV, whereof about 98.5% pass the $|\eta| < 2.8$ cut.



(a) (b)

Figure 6: 2^{nd}-b quark p_T of unmatched MadEvent 2→2 LO and 2→3 NLO samples: intersection in smoothed log_{10} distribution (a), in linear scale (b).

By increasing the fraction of 2→3 events from 170426 to 188279 and thereby decreasing the ratio from $R = 1.4$ to $R = 1.29$, appropriate matching parameters are found. The value of $K_T = 9$ GeV, the nearest integer at the intersection of the smoothed distributions of the sub-samples (Figure 6 (a)), results in 62496 2^{nd}-b quarks with $p_T > 15$ (highlighted in Figure 6 (b)). 61521 of them pass the $|\eta| < 2.8$ cut, which exactly matches the NLO predictions of 21.1% of the total of 291738 events. The resulting sub-samples and the matched t-channel sample are shown in Figure 7.



(a) (b)

Figure 7: Matching of the $m_t = 175$ GeV/c^2 sample: 2^{nd}-b quark p_T of unmatched MadEvent 2→2 LO and 2→3 NLO samples (a), matched sample (b).

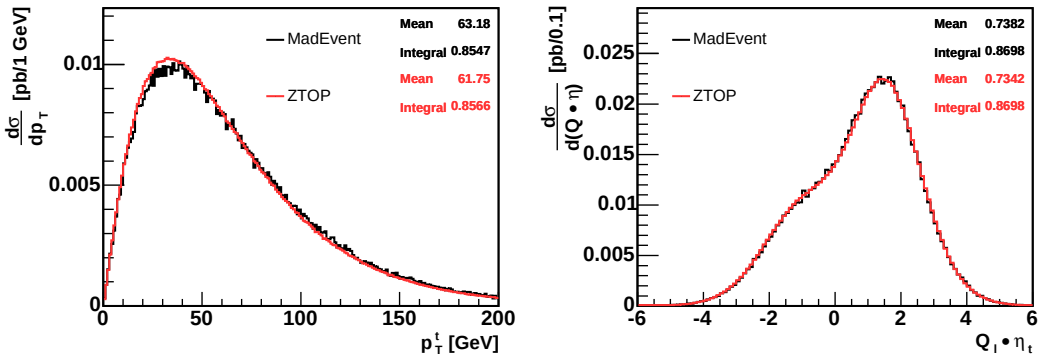
5 Validation

To validate the MadEvent s- and t-channel MC samples by means of the ZTOP calculations, acceptance cuts of $p_T > 15$ GeV and $|\eta| < 2.8$ are applied to the jet distributions of light quark and b quark. These cuts are motivated by the jet definition used in the single-top analysis. No cuts are applied to the top quark, since it further decays before building a jet. In the MadEvent event listing, the top quark is not included, so one has to build its fourvector out of the decay products at parton level, namely the W boson and the 1^{st}-b quark. The MadEvent s- and t-channel top quark distributions are normalized to the total ZTOP cross section of 0.870 pb or 1.985 pb, respectively. These normalizations yield the scales that are used for the normalization of all following s- and t-channel distributions. That means that the jet distributions are also normalized to the inclusive cross section. There is no additional normalization of these distributions.

For all but the top quark distributions, one has to compare p_T-ordered jets, in t-channel with the additional distinction between jets with or without b quark content. The potential 1^{st}-b quark jet from the top decay is not considered in the remainder of this thesis, since ZTOP does not cover this decay. We therefore will assume that the 1^{st}-b quark is well modeled in MadEvent. The differences between the integral of the p_T and $Q_l \cdot \eta$ distributions are due to overflow entries beyond the histogram boundaries.

5.1 s-channel Sample

In Figure 8 the top quark p_T and $Q_l \cdot \eta$ distributions of MadEvent and ZTOP are shown. The MadEvent top quark p_T is slightly harder than the ZTOP spectrum, in $Q_l \cdot \eta$ the distributions match perfectly.

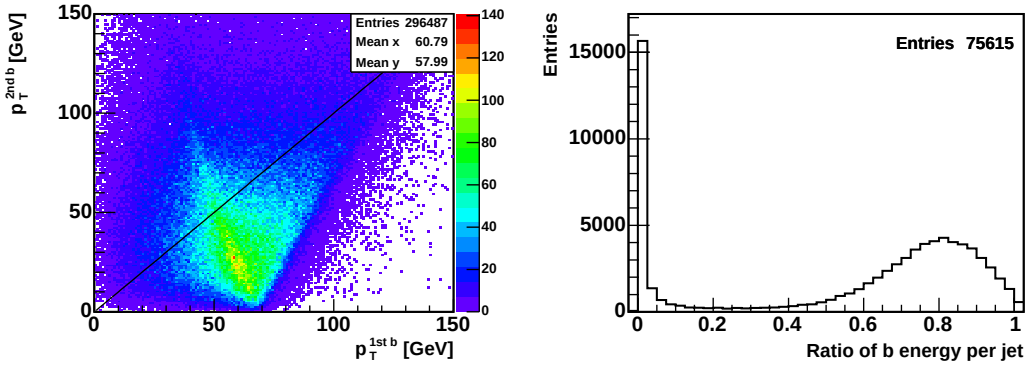


(a) (b)

Figure 8: s-channel top quark transverse momentum p_T (a) and lepton charge times top quark pseudorapidity $Q_l \cdot \eta$ (b).

The ZTOP software calculates, that at least one jet (apart from the 1^{st}-b quark jet) should be visible within the detectors acceptance in about 92.8% of all s-channel events. In the majority of events (84.3%) this p_T-leading jet is a 2^{nd}-b quark jet, which is in MadEvent represented by a 2^{nd}-b quark at parton level. Only in 8.5% ZTOP expects a light quark jet, which in the s-channel comes from initial or final state gluon radiation. Since the s-channel MadEvent sample is LO and does not include matrix elements of gluon radiations, this is modeled by the PYTHIA showering and does therefore not appear at parton level. To take this into account we have to leave the simple parton model which serves well for the 2^{nd}-b quark jets and investigate the events at hadron level after the showering. To do that one has to apply a jet clustering to the stable particles to include these contributions. The parameters of the k_T cluster algorithm, used for this purpose, were set to a cone size of 0.54 (equal to the ZTOP cone size) and to a minimum jet energy of 15 GeV, adjusted to the selection cut used on data. Since the clustering was done with all stable particles constructed by PYTHIA except for the lepton and the neutrino from the top decay, the 1^{st}-b quark from the top decay cannot be excluded and will contribute to the number of jets found by the cluster algorithm.

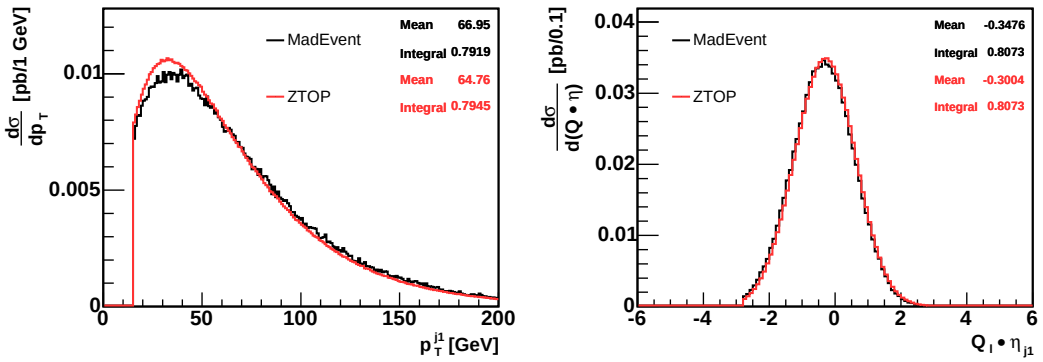
Due to the QCD color conservation implemented in the PYTHIA hadronization process, one cannot unambiguously trace back the jets to the original quarks, so another method to assign the jets to the partons is needed: at least 3 jets per event after the acceptance cuts were required, one for the 1^{st}-b quark, the 2^{nd}-b quark and the light quark jet. With the jets ordered in transverse momentum, we assume that the probability for the gluon radiated light jet to end up in neither the first nor the second jet is high, since both b quarks are rather hard (Figure 9 (a)). To get the light quark jet, one has to study the third ordered jet. The distinction, whether or not a jet is a light quark jet, was done by a cut of the ratio of the jet energy coming from B hadrons, thus from b quarks. We consider jets to be light if less than 30% of the jet energy is from particles coming from B hadrons, see Figure 9 (b).



(a) (b)

Figure 9: s-channel: p_T of 1^{st}-b quark versus p_T of 2^{nd}-b quark (a) with the line at equal values, ratio of jet energy of clustered particles coming from B hadrons shown for all third ordered jets with any B hadron content (b).

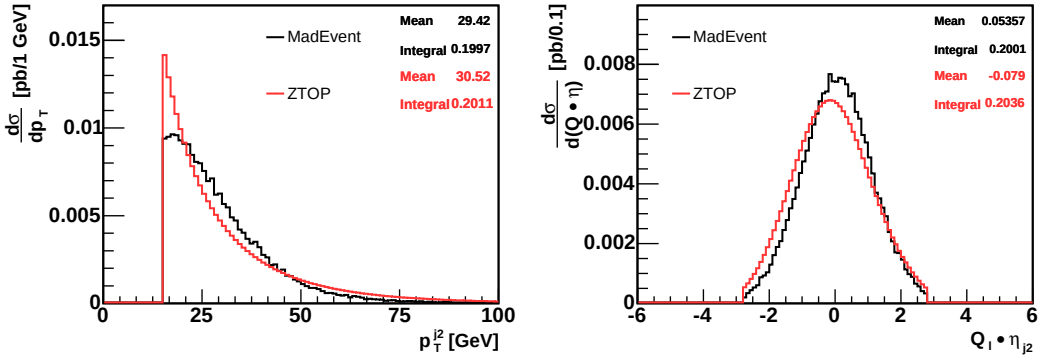
The MadEvent p_T-leading jet (j1) distributions, containing 2^{nd}-b quark jets and a small fraction of clustered light jets, are in good agreement with the ZTOP predictions, as seen in Figure 10. There are no differences in the rates, the MadEvent p_T spectrum is slightly too hard. The mean value of p_T is about 3% higher for MadEvent than for ZTOP.



(a) (b)

Figure 10: s-channel leading jet (j1) p_T (a) and $Q_l \cdot \eta$ (b). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

The p_T-ordered 2^{nd}-leading jet (j2), appearing in about 23.4% of all events according to ZTOP, is dominated by gluon radiated light jets, in contrast to the leading jets. Only in about 5.4% a 2^{nd}-b quark jet is the 2^{nd}-leading jet. Because of this the MadEvent p_T and $Q_l \cdot \eta$ distributions in Figure 11 do not match properly with the ones of ZTOP. The jet clustering of the stable particles does not result in a steeply falling p_T spectrum as predicted by ZTOP. Also in $Q_l \cdot \eta$ MadEvent differs from the theoretical calculations, solely the rates of the 2^{nd}-leading jets comply. Table 1 summarizes the relevant s-channel cross sections and fractions.



(a) (b)

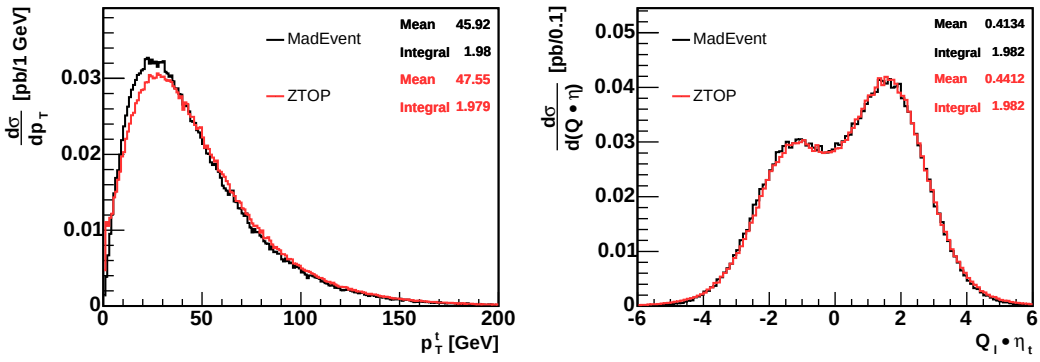
Figure 11: s-channel 2^{nd}-leading jet (j2) p_T (a) and $Q_l \cdot \eta$ (b). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

jet type	**ZTOP** cross section [pb] and fraction	**MadEvent** cross section [pb] and fraction
top quark (total cross section)	0.870 100%	0.870 N=296487 100%
leading b quark jet (b1)	0.733 84.3%	0.799 91.9%
leading light quark jet (q1)	0.074 8.5%	0.008 0.9%
leading jet seen (j1 = b1 + q1)	0.807 92.8%	0.807 92.8%
2^{nd}-leading b quark jet (b2)	0.047 5.4%	0.005 0.6%
2^{nd}-leading light jet (q2)	0.156 18.0%	0.195 22.4%
2^{nd}-leading jet seen (j2 = b2 + q2)	0.204 23.4%	0.200 23.0%

Table 1: Summary of all relevant s-channel cross sections and resultant fractions (defined as ratio of the cross sections, with 0.870 pb as denominator). The MadEvent s-channel sample contains a total of 296487 events. The 1^{st}-b quark from the top decay is excluded from the jet ordering.

5.2 t-channel Sample

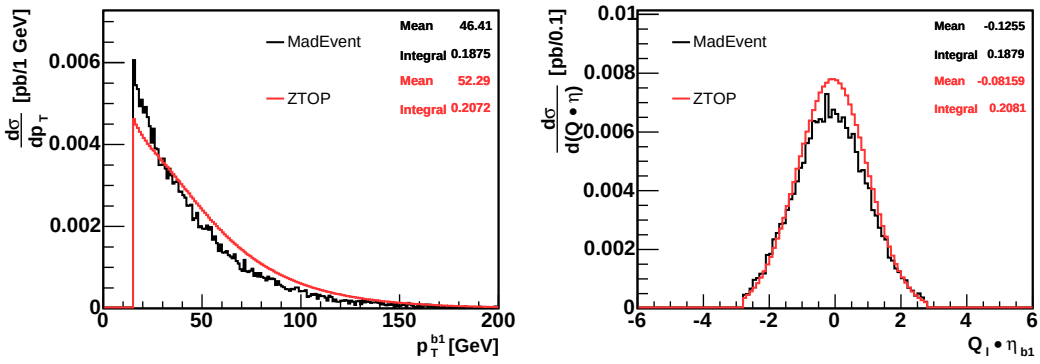
Figure 12 shows the comparison of top quark p_T and $Q_l \cdot \eta$ distributions between MadEvent and ZTOP. The t-channel MadEvent top quark distributions are in good agreement with the ZTOP output. The deviation in the mean of the transverse momentum is minor, the MadEvent $Q_l \cdot \eta$ mean value is slightly lower than the one from ZTOP.



(a) (b)

Figure 12: t-channel top quark transverse momentum p_T (a) and lepton charge times top quark pseudorapidity $Q_l \cdot \eta$ (b).

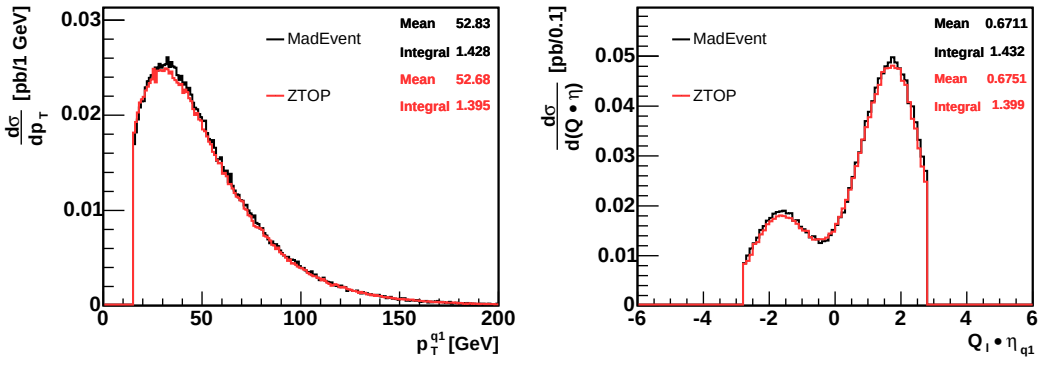
In the case of p_T-leading jets, MadEvent t-channel jets are represented by the corresponding quarks at parton level. The relevant cross sections and event fractions are summarized in Table 2. As computed by ZTOP, there should be at least one jet in 81.0% of all events within the detector acceptance cuts. When the 2^{nd}-b quark jet is the p_T-leading jet (b1), which is the case in 10.5% of all events according to ZTOP, MadEvent underestimates this rate by a factor of about 0.90, as shown in Figure 13. The falling p_T spectrum and the central $Q_l \cdot \eta$ distribution are reasonably well simulated by MadEvent.



(a) (b)

Figure 13: p_T (a) and $Q_l \cdot \eta$ (b) of the t-channel leading b quark jet (b1). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

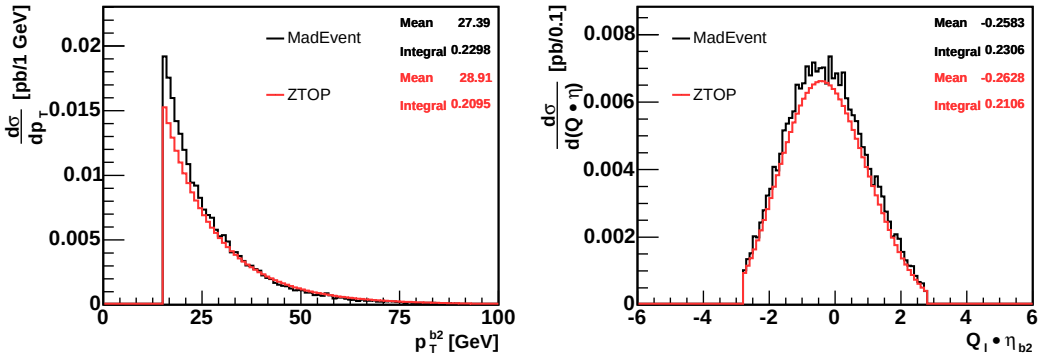
In the majority of events (70.5% according to ZTOP) the p_T-leading jet will be a light quark jet (q1). In Figure 14 one can see that MadEvent reproduces the corresponding distributions very well, the deviations in the mean of p_T as well as of $Q_l \cdot \eta$ are negligible, only the corresponding rate is slightly overestimated by a factor about 1.02.



(a) (b)

Figure 14: p_T (a) and $Q_l \cdot \eta$ (b) of the t-channel leading light quark jet (q1). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

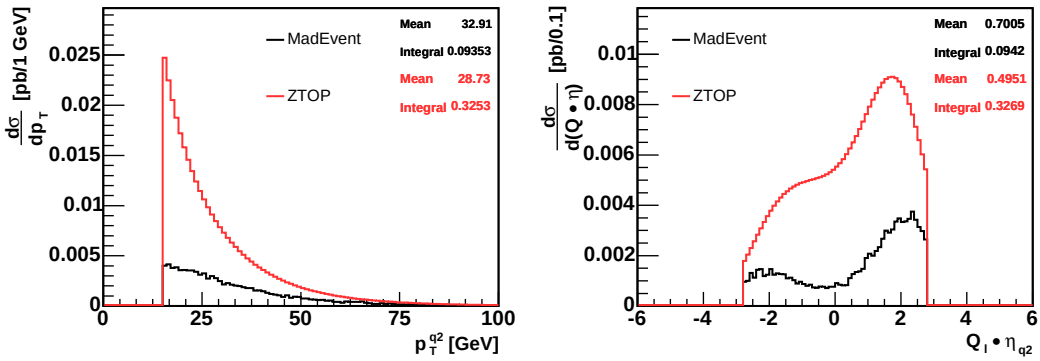
In 27.1% of all events a second jet (not considering the jet from the top decay as discussed earlier) should be visible within the detector's acceptance. If this p_T-2^{nd}-leading jet is a b quark jet (b2), which implies that in the same event the light quark jet is the p_T-leading jet, one obtains the following distributions of Figure 15. Both the falling p_T-spectrum and the slightly asymmetric shape of $Q_l \cdot \eta$ are well modeled by MadEvent, only the rate, i.e. the cross section, is slightly too high by a factor of about 1.09.



(a) (b)

Figure 15: p_T (a) and $Q_l \cdot \eta$ (b) of the t-channel 2^{nd}-leading b quark jet (b2). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

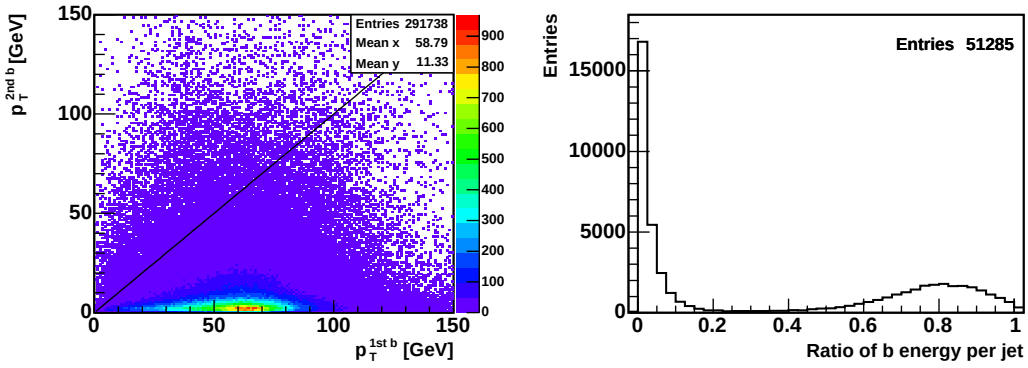
If the p_T-2^{nd}-leading jet is a light quark jet (q2), the MadEvent jet picture represented by the parton level light quark does not agree with the ZTOP predictions. In Figure 16 one can observe a big discrepancy in the rate, which is too low by a factor of about 0.29, as well as in the $Q_l \cdot \eta$ shape, which is too far forward.



(a) (b)

Figure 16: p_T (a) and $Q_l \cdot \eta$ (b) of the t-channel 2^{nd}-leading light quark jet (q2). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

The reason for the bad agreement could be that the matched MadEvent sample does not include all relevant 2→3 NLO matrix elements. We assume that initial state gluon splitting (Figure 2 (b)) as well as initial and final state gluon radiations (Figure 2 (c)) may have major contributions to light quark jets, especially to softer ones as considered in the p_T-2^{nd}-leading jets. The contributions to b quark jets should be significantly lower due to the small probability of gluons splitting in $b\bar{b}$ quark pairs.
The same circumstance is discussed in Section 5.1, where the s-channel light jets are produced via gluon radiations. Again, a jet clustering of stable particles will be necessary to collect the contributions generated by the PYTHIA showering. In the t-channel case, a different method to assign the jets to the partons is needed: at least 3 jets per event after the acceptance cuts were required, one for the p_T-leading, 2^{nd}-leading and for the 1^{st}-b quark jet respectively. With the jets ordered in transverse momentum, the probability for the 1^{st}-b quark parton to end up in the third jet can be assumed to be low, since the 1^{st}-b quark parton is mostly higher in p_T than both the light quark parton and the 2^{nd}-b quark parton as shown in Figure 17 (a).

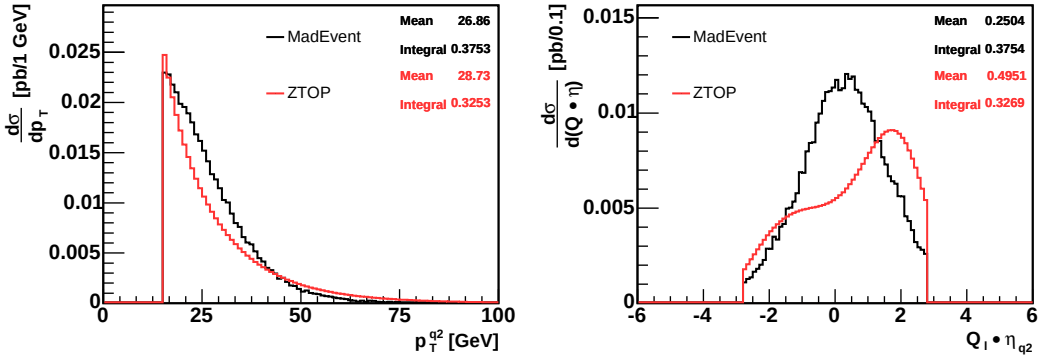


(a) (b)

Figure 17: t-channel: p_T of 1^{st}-b quark versus p_T of 2^{nd}-b quark (a) with the line at equal values, ratio of jet energy of clustered particles coming from B hadrons shown for all third ordered jets with any B hadron content (b).

To get the 2^{nd}-leading light quark jet, one has to study the third ordered jet, when there is no 2^{nd}-leading b quark parton in the event. The same distinction as in Section 5.1, whether or not a jet is a light quark jet, was used. In Figure 17 (b) the corresponding ratio for all jets with any b content is shown. The peak at very small values of the ratio is due to single particles coming from B hadrons, which are clustered into original light quark jets. If the majority of the clustered particles originates from B hadrons, the ratio of b energy per jet accumulates at high values. As already mentioned in Section 5.1, we consider jets to be light if less than 30% of the jet energy is from particles coming from B hadrons.

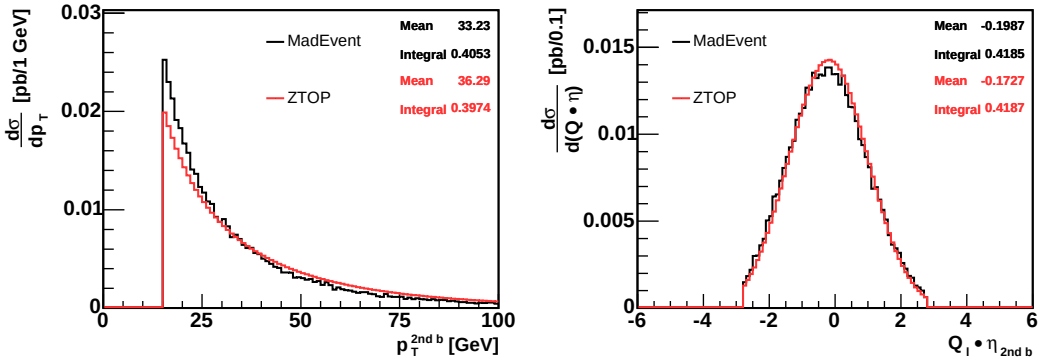
With these jet definitions, the distributions change heavily and a more reasonably agreement of the rates is reached as shown in Figure 18. The steeply falling p_T spectrum is not well modeled and the $Q_l \cdot \eta$ distribution has only poor agreement with the ZTOP predictions, since it is far too central. The rate is too high by about a factor of 1.15, but given the fact, that the matching procedure was performed with only two of the four relevant matrix elements, differences in these distributions are to be expected.



(a) (b)

Figure 18: p_T (a) and $Q_l \cdot \eta$ (b) of the clustered 2^{nd}-leading light jet (q2) in the t-channel. The 1^{st}-b quark from the top decay is excluded from the jet ordering.

Modeling the distributions of the 2^{nd}-b quark is the main reason for the MadEvent t-channel matching, performed in Section 4.3. This was done by adjusting the rate of the 2^{nd}-b quark jet to the NLO prediction. The corresponding p_T and $Q_l \cdot \eta$ distributions are shown in Figure 19. They are built by the sum of the leading b quark jet (b1) and 2^{nd}-leading b quark jet (b2) distribution. One can certainly state that the matching procedure of the t-channel MadEvent samples works very well.



(a) (b)

Figure 19: p_T (a) and $Q_l \cdot \eta$ (b) of the matched t-channel 2^{nd}-b jet (2^{nd} b=b1+b2).

jet type	**ZTOP** cross section [pb] and fraction	**MadEvent** cross section [pb] and fraction
top quark (total cross section)	1.985 100%	1.985 N=291738 100%
leading b quark jet (b1)	0.208 10.5%	0.188 9.5%
leading light quark jet (q1)	1.399 70.5%	1.432 72.2%
leading jet seen (j1 = b1 + q1)	1.607 81.0%	1.620 81.6%
2^{nd}-leading b quark jet (b2)	0.211 10.6%	0.231 11.6%
2^{nd}-leading light jet (q2)	0.327 16.5%	0.375 18.9%
2^{nd}-leading jet seen (j2 = b2 + q2)	0.538 27.1%	0.606 30.5%
2^{nd}-b quark jet (2^{nd}-b = b1 + b2)	0.419 21.1%	0.419 21.1%

Table 2: Summary of all relevant t-channel cross sections and resultant fractions (defined as ratio of the cross sections, with 1.985 pb as denominator). The 1^{st}-b quark from the top decay is excluded from the jet ordering.

6 Estimate of systematic Error on the MC Model

The underlying concept of estimating the systematic uncertainty is to use the deviations between the ZTOP and the MadEvent variables. By building a ratio-function of ZTOP over MadEvent bin by bin for each variable, one can see in which region MadEvent under- and overestimates the signal. Each event passing the single-top selection cuts gets a scale corresponding to the ratio of its bin. The sum of all scales approximates the corrected acceptance. Half of the arising difference between the corrected and the original acceptance is taken as symmetric systematic uncertainty. Since one has several variables and therefore several ratio-functions per event, an appropriate procedure has to be found to compute an overall scale per event. The applied proceeding is similar to the least-squares-method of weighted average. At this level the correlations between the ratio-functions $r_i(bin) = ZTOP_i(bin)/MadEvent_i(bin)$ are taken into account. The index i stands for one of the N_{var} used variables. A covariance-matrix V_{ij} of all produced ratio-functions $r_i(bin)$ is constructed by running over all N_{ev} events of the full MadEvent sample of s- or t-channel, respectively. For each event n the ratios $r_i(bin_{in})$ of the considered variables i are calculated by taking the content of bin_{in}, the event n is located in:

$$V_{ij} = \frac{S_{ij} - S_i \cdot S_j / N_{ev}}{N_{ev} - 1} \qquad \text{with } S_i = \sum_{n=1}^{N_{ev}} (r_i(bin_{in}) - 1) \tag{1}$$

$$\text{and } S_{ij} = \sum_{n=1}^{N_{ev}} (r_i(bin_{in}) - 1) \cdot (r_j(bin_{jn}) - 1) \qquad \text{for } i, j = 1, 2, ...N_{var}.$$

The weight-matrix M_{ij} is obtained by inverting the covariance-matrix: $M_{ij} = V_{ij}^{-1}$
For each event m of all N_{pre} events passing the preselection, a weighted scale s_m is calculated as follows:

$$s_m = \sum_{i=1}^{N_{var}} w_i \cdot r_i(bin_{im}), \text{ where the weight is given by } w_i = \sum_{j=1}^{N_{var}} M_{ij} / \sum_{k,l=1}^{N_{var}} M_{kl} \quad (2)$$

One obtains the corrected acceptance A by summing up all weighted scales: $A = \sum_{m=1}^{N_{pre}} s_m$
Since the single-top analysis investigates both the 2- and 3-jet-bin, this method is done for the 2-, 3- and 2+3-jet-bin separately. To take into account the statistical fluctuations, several actions are implemented by creating the ratio-functions. They are exemplarily shown in Figure 20: A smoothing of the original (a) t-channel p_T distribution of the leading b jet is conducted (b). To get the long tails of the p_T distributions under control, overflow bins are established in the ratio-functions for p_T higher than 75 GeV or 150 GeV, respectively (c). Under- and overflow bins are also used in the low statistic regions $|\eta| > 4.0$ of the top quark $Q_l \cdot \eta$ distributions. In the case of big unphysical oscillations in the ratio-functions, the affected region is approximated by a fit of a polynomial of low order (d).

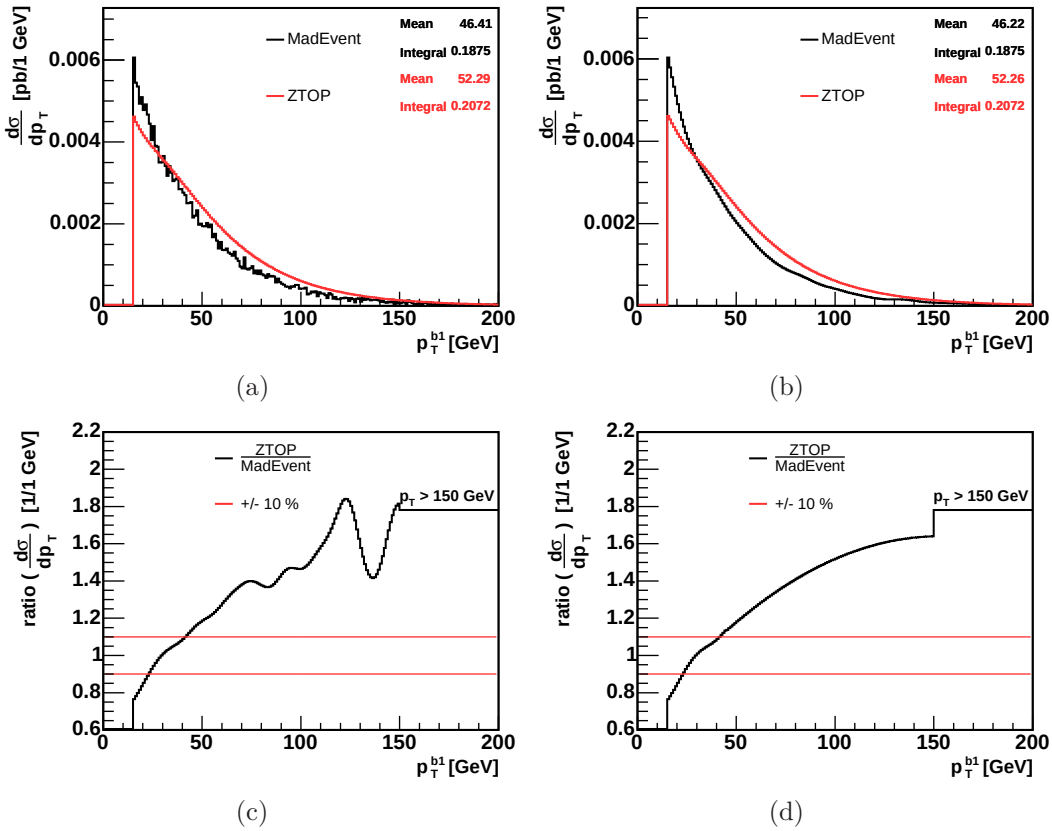


Figure 20: Intermediate steps by building the ratio-functions: the original (a) distributions are smoothed (b). Under- and overflow bins are established in the low statistic regions, here $p_T > 150$ GeV, of the ratio-functions (c). Appearing unphysical oscillations are approximated by a fit (d).

6.1 s-channel Estimate

$N_{var} = 6$ variables, validated in Section 5.1, are used to estimate the s-channel systematic uncertainty: p_T and $Q_l \cdot \eta$ of the top quark, of the p_T-ordered leading jet (j1) and of the 2^{nd}-leading jet (j2). Their ratio-functions are shown in Figure 21.

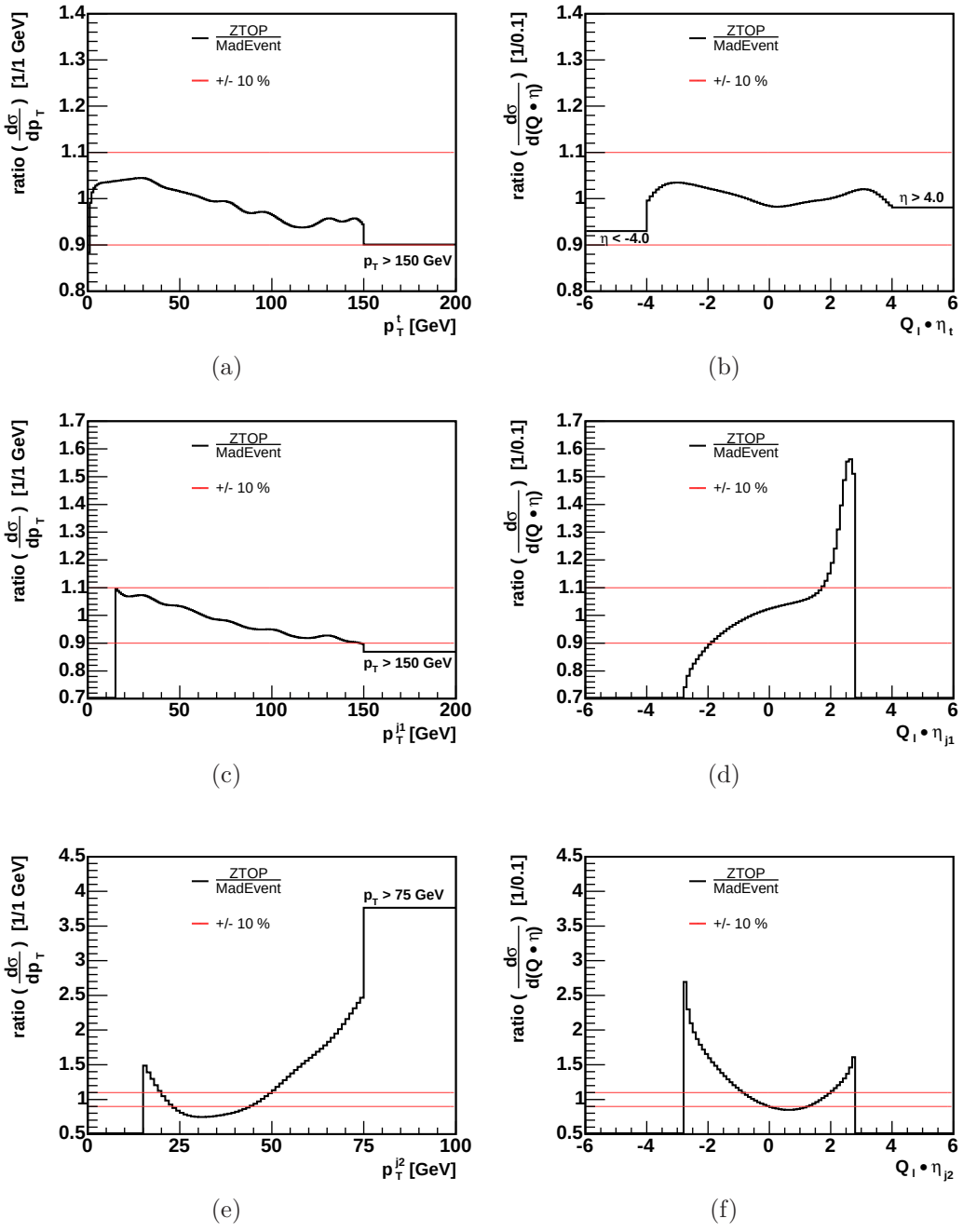
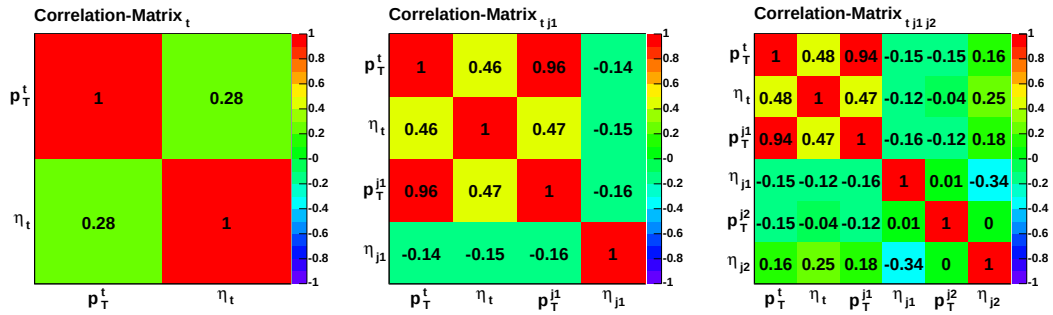


Figure 21: Ratio-functions of the s-channel top quark ((a) and (b)), leading jet j1 ((c) and (d)) and 2^{nd}-leading jet j2 ((e) and (f)).

One has to distinguish between three kinds of events at generator level: Those with no jets within the detectors acceptance (labeled: t), those with one jet (tj1) and those with two jets visible (tj1j2). The top quark variables are always included. For each of these cases a separate weight-matrix of the ratio-functions has to be computed. The corresponding correlation-matrices are shown in Figure 22.



(a) (b) (c)

Figure 22: s-channel correlation-matrices of the ratio-functions, separate for the subsamples t (a), tj1 (b) and tj1j2 (c).

Depending on the kind of event passing the preselection, the corresponding weight-matrix is used to calculate the weighted scale s_m. As an example, all scales of the preselected events of the 2+3-jet-bin are plotted in Figure 23. Table 3 specifies the corrected acceptance and its estimated uncertainty for each jet-bin.

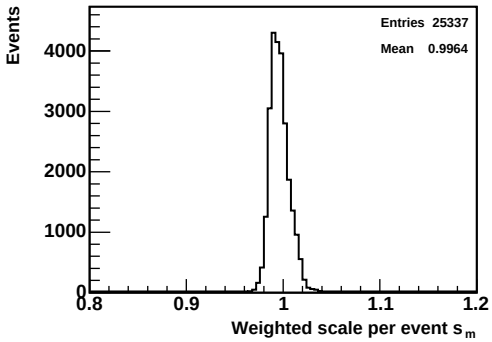


Figure 23: Scales of the s-channel events passing the 2+3-jet-bin preselection, the deviation of the mean from 1.0 is only about 0.4%.

	2-Jet-Bin	**3-Jet-Bin**	**2+3-Jet-Bin**
original acceptance	20895.0	4442.0	25337.0
corrected acceptance	20825.7	4419.4	25245.0
correction	-0.33%	-0.51%	-0.36%
estimated uncertainty	$\pm 0.2\%$	$\pm 0.3\%$	$\pm 0.2\%$

Table 3: s-channel: the estimated uncertainty is half the absolute value of the difference between original and corrected acceptance. The original acceptance is given as the number of MC events passing the single-top selection.

6.2 t-channel Estimate

For the t-channel, all $N_{var} = 10$ variables validated in Chapter 5.2 are considered. In contrast to the s-channel, it is distinguished whether or not a jet has a b quark content to treat the matched 2^{nd}-b quarks separately. The ratio-functions of p_T and $Q_l \cdot \eta$ of the top quark, of the leading b jet b1, of the leading light jet q1, of the 2^{nd}-leading b jet b2 and of the 2^{nd}-leading light jet q2 are shown in Figure 24 and 25.

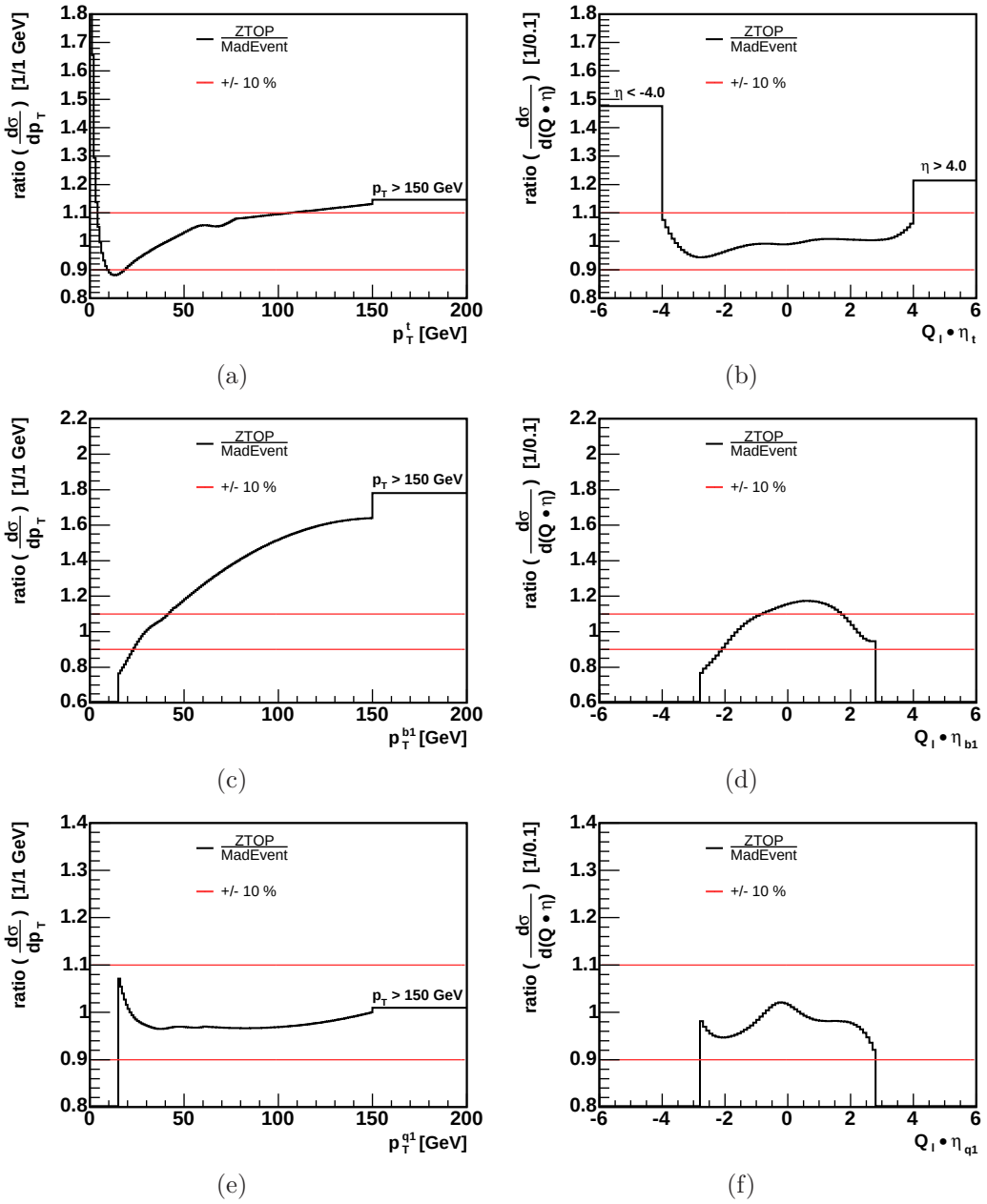
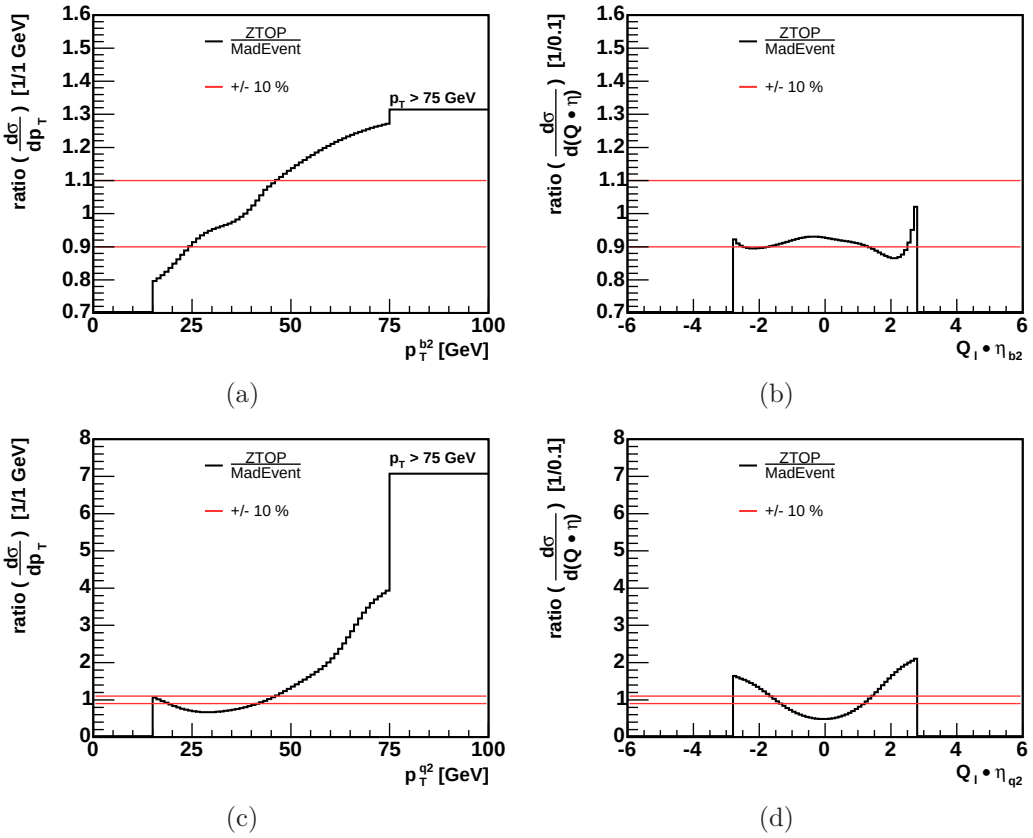


Figure 24: Ratio-functions of the t-channel top quark t ((a) and (b)), leading b jet b1 ((c) and (d)) and leading light jet q1 ((e) and (f)).

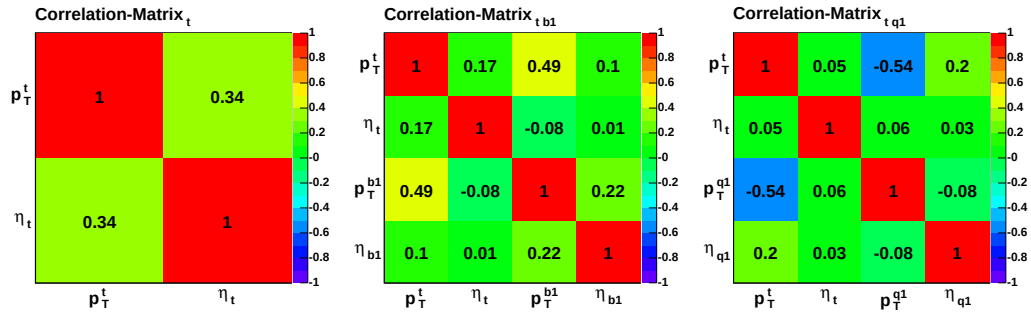


(a) (b)

(c) (d)

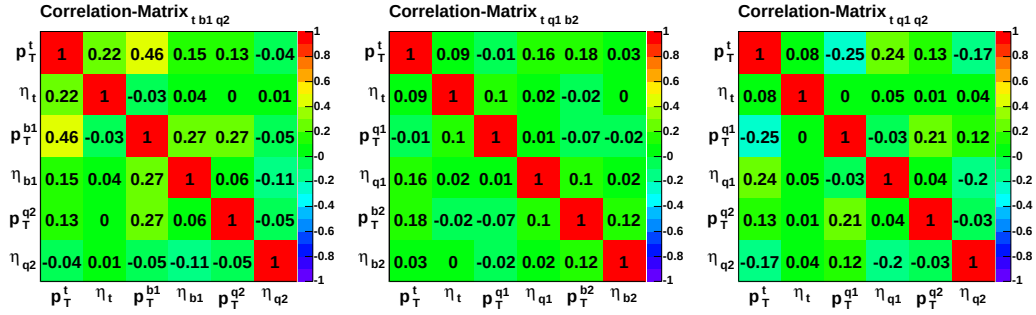
Figure 25: Ratio-functions of the t-channel 2^{nd}-leading b jet b2 ((a) and (b)) and 2^{nd}-leading light jet q2 ((c) and (d)).

Due to the flavor discrimination used in the t-channel, one has to deal with 6 different kinds of events: only top (labeled t), top plus b jet (tb1), top plus light jet (tq1), top plus leading b jet plus 2^{nd}-leading light jet (tb1q2), top plus leading light jet plus 2^{nd}-leading b jet (tq1b2) and top plus leading light jet plus 2^{nd}-leading light jet (tq1q2). As discussed for the s-channel, separate weight-matrices are used for each case. The corresponding correlation-matrices are shown in Figure 26 and 27.



(a) (b) (c)

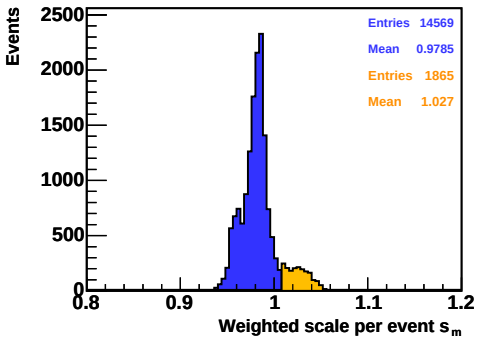
Figure 26: t-channel correlation-matrices: t (a), tb1 (b) and tq1 (c).



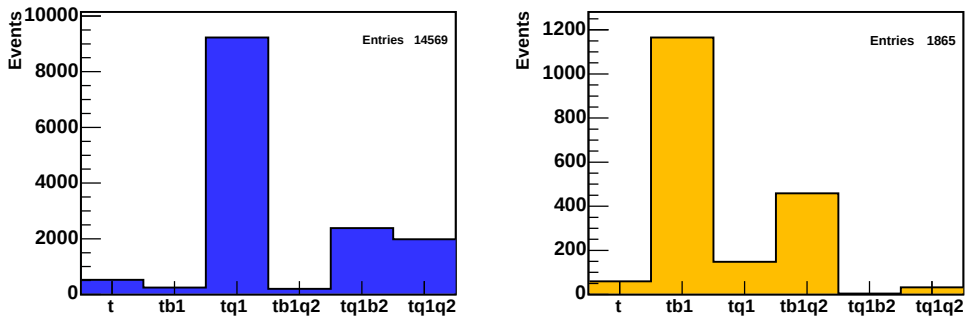
(a) (b) (c)

Figure 27: t-channel correlation-matrices: tb1q2 (a), tq1b2 (b) and tq1q2 (c).

In Figure 28 (a) the scales of the events passing the 2+3-jet-bin selection are shown. The distribution peaks at about 0.98, a smaller accumulation is located around 1.03. Figure 28 (b) and (c) show the frequency of the different kinds of events with a scale smaller and higher than about 1.01. It is clearly visible, that the scale is higher, when a leading b jet is found in the event. This arises from the high values in the ratio-functions (Figure 24 (c) and (d)). Table 4 lists the corrected acceptance and its estimated uncertainty for each jet-bin.



(a)



(b) (c)

Figure 28: Scales of events passing the 2+3-jet-bin preselection (a), lower scales are dominated by leading light jets q1 (b), higher scales are dominated by leading b jets b1 (c).

	2-Jet-Bin	**3-Jet-Bin**	**2+3-Jet-Bin**
original acceptance	13204.0	3230.0	16434.0
corrected acceptance	12971.6	3153.4	16128.9
correction	−1.76%	−2.37%	−1.86%
estimated uncertainty	±0.9%	±1.2%	±0.9%

Table 4: Summary of the t-channel results. The estimated uncertainty is half the absolute value of the difference between original and corrected acceptance. The original acceptance is given as the number of MC events passing the single-top selection.

7 Conclusion and Outlook

In this note, the t-channel matching procedure of two single-top signal Monte Carlo samples is optimized. The s- and matched t-channel samples, generated by MadEvent, are validated by comparing to ZTOP next-to-leading-order calculations. We find good agreement for all kinematic distributions we investigate, except for softer light quark jets due to gluon radiation. Since this has only minor impact on the s-channel, the corresponding MadEvent sample performs its task as expected. For the t-channel, we can conclude that the applied matching procedure leads to a MadEvent sample that successfully describes the kinematic distributions and rates of the 2^{nd}-b quark. However, small differences remain. The discrepancy in the p_T-ordered 2^{nd}-leading light jets is mainly due to the absence of initial state gluon splitting and initial and final state gluon radiation matrix elements in the MadEvent sample production. The subsequent PYTHIA showering of the partons is apparently inappropriate for modeling those contributions and not intended for this purpose.

The proper way would be to produce all relevant NLO matrix elements and match them as proposed in reference [11]. At present, an NLO-MC-generator for single-top is in preparation [25]. Probably it will be available for future iterations of single-top analyses and will redundantize further matching procedures.

We estimate the systematic uncertainty on the single-top acceptance due to the Monte Carlo modeling and find an uncertainty of about 1% on the t-channel acceptance. We obtain a negligible uncertainty well below 1% on the s-channel acceptance. These acceptance uncertainties are very well acceptable for the single-top analyses that are currently under way.

References

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

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

[3] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **65**, 091102 (2002).

[4] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **69**, 052003 (2004).

[5] DØ Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. B **517**, 282 (2001).

[6] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. D **63**, 031101 (2001).

[7] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **71**, 012005 (2005).

[8] DØ Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. B **622**, 265 (2005).

[9] W. Wagner, Rep. Prog. Phys. **68**, 2409 (2005).

[10] B. W. Harris *et al.*, Phys. Rev. D **66**, 054024 (2002).

[11] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).

[12] T. M. P. Tait, Phys. Rev. D **61**, 034001 (2000).

[13] T. Sjöstrand *et al.*, Comp. Phys. Commun. **135**, 238 (2001).

[14] E. E. Boos, L. V. Dudko and V. I. Savrin, CMS Note 2000/065 (2000).

[15] T. Stelzer and W. F. Long, Comp. Phys. Commun. **81**, 357 (1994).

[16] F. Maltoni and T. Stelzer, hep-ph/0208156 (2002).

[17] C. Ciobanu *et al.*, CDF Note 7020 (2004).

[18] T. Stelzer, Z. Sullivan and S. Willenbrock, Phys. Rev. D **56**, 5919 (1997).

[19] V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972).

[20] G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).

[21] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[22] D. O. Carlson and C. P. Yuan, Phys. Lett. B **306**, 386 (1993).

[23] G. Mahlon and S. Parke, Phys. Rev. D **55**, 7249 (1997).

[24] A. P. Heinson, A. S. Belyaev and E. E. Boos, Phys. Rev. D **56**, 3114 (1997).

[25] S. Frixione *et al.*, hep-ph/0512250 (2005).