CF/PHYS/TOP/ANAL/7252
November 3, 2004
Version 1.4

b-jet Energy Scale Uncertainty From Existing Experimental Constraints

Jean-Francois Arguin and Pekka K. Sinervo
University of Toronto

Abstract

The b-jets energy scale is a crucial uncertainty in the top quark mass measurement. So far in CDF, the jet energy scale uncertainty is estimated for generic jets regardless of the jet flavor and environment, and the properties specific to b-jets are considered to contribute negligibly to the overall uncertainty. We check this assumption in this note. The fragmentation, colour flow and semileptonic decays are identified as the main differences between b-jets and generic jets. The top quark mass uncertainties from the b-jets fragmentation and semileptonic decays are respectively estimated to be $\approx 0.2\ \mathrm{GeV}/c^2$ and $\approx 0.4\ \mathrm{GeV}/c^2$ using current experimental constraints on these topics. The colour flow uncertainty is at least partly covered by the fragmentation uncertainty. These uncertainties are small with respect to the total statistical and systematic top quark mass uncertainty at this stage of Run II.

1 Introduction

As more data are accumulated in Run II, the statistical uncertainty on the top quark mass measurement is shrinking and systematic uncertainties become increasingly important. Currently, the dominating systematic uncertainty is by far the jet energy scale (see recent measurements: [1][2][3][4]). Therefore, a better understanding of the jet energy scale uncertainty is of prime importance for the future of the top quark mass measurement.

The jet energy scale and is associated uncertainty is currently estimated by the jet corrections group [5] for "generic" jets, i.e. regardless of the flavor of the jets or the environment in which they are measured. In fact, the samples employed are dijets, W+jets, minimum bias and photon+jets events, all with different flavor composition and colour flow. For the top quark mass measurement, this estimate is problematic because $t\bar{t}$ events events contain b-jets and light quark jets that are different than generic jets in many regards. In the b-jets case: their fragmentation is harder, their colour flow is defined by the $t\bar{t}$ event nature, they have a large semileptonic decay fraction and they are measured in $t\bar{t}$ events that are more busy than other types of events. The uncertainties on the specific properties of b-jets have been so far neglected in the calculation of the top quark mass uncertainty. The necessity to verify this assumption has been raised recently in the top mass group.

A different approach to determine the jet energy scale is to use the hadronic W mass peak reconstructed from the light quark jets (W-jets) in $t\bar{t}$ lepton+jets events [6]. This measurement has the advantage of being done fully *in situ*, it is thus free of the potential ambiguities arising when applying the generic jet energy scale extracted from very different environments than $t\bar{t}$ to the W-jets. However, it is only sensitive to the W-jets energy scale and does not provide direct information on the b-jets. Consequently, the additional uncertainties due to the different flavor and colour flow of b-jets in $t\bar{t}$ events has to be estimated separately for this approach.

We emphasize the importance of estimating properly the b-jets energy scale by pointing out that the top quark mass is more sensitive to the b-jets energy scale than the W-jets energy scale. This is because

1. the b-jets originate directly from the top quark decay, and
2. most top quark mass analyses use the knowledge of the W boson mass to limit the dependence on the W-jets energy scale uncertainty.

One method to estimate the b-jets energy scale is to study $\mathrm{Z} \rightarrow b\bar{b}$ events triggered using the displaced vertices of b-jets [7]. It provides direct information but needs significantly more data than is currently available to meaningfully contrain the b-jets energy scale.

The approach we take in this note is to quantify the uncertainties of the properties specific to b-jets using current experimental constraints. The specificities of b-jets with respect to generic or W-jets are discussed in Sec. 2. We identify differences in the fragmentation, decay and colour. We describe the existing experimental contrains on these characterics and quantify their effect on the top quark mass measurement. This is described in Sec. 3 for the fragmentation and colour flow and Sec. 4 for the decay.

We use the template top mass analysis with b-tags in the lepton+jets channel [1] to estimate the impact of the b-jets uncertainty on the top quark mass measurement. We expect the general conclusions drawn from this analysis to be applicable to any top quark mass analysis. For this analysis, we employed the same event selection requirements and analysis technique as used previously [1]. In particular, we performed the analysis with version 4.11.1 of the CDF software.

2 Why Are b-jets Special?

As mentioned previously, it is possible to constrain the jet energy scale for the top quark mass measurement using the jet corrections group approach [5] or the *in situ* $W \to jj$ decays [6]. In both cases, we expect a large fraction of the generic jets or W-jets uncertainties to be correlated with the b-jets (e.g. single track response, underlying event, relative response, etc). The additional sources of uncertainties on b-jets will come from differences between generic or W-jets and b-jets in $t\bar{t}$ events. We identify four differences:

1. **Fragmentation**: The b-hadron resulting from b-quark fragmentation carries a larger fraction of the parent quark momentum than for the light quark fragmentation. This is because the b-quark is much heavier, and is thus only slightly decelerated when combined with a light quark to form a b-hadron.

2. **Color flow**: W-jets come from the decay of a colour singlet, while the b-jets are colour-connected with the initial state partons.

3. **Decay**: A large fraction of the decay products of b-hadrons are leptons and neutrinos. These particles interact very differently with the calorimeter than the more common hadronic particles. Therefore, b-jets will have a different response on average than W-jets because of the large b-hadrons semileptonic decay fraction.

4. $t\bar{t}$ **Environment**: The event characteristics (numbers of extra jets and possibly initial state and final state radiation effects) may have an effect on our understanding of the jet energy scales. However, we note that these effects are already tested and constrained by the $W \to jj$ measurement and the modelling of the $t\bar{t}$ decay through our Monte Carlo calculations. We don't believe such "environmental" effects introduce any additional uncertainties on the relative b-jets energy scale relative to that of light quark jets.

The following sections discuss the experimental constraints on heavy quark fragmentation and semileptonic decay and their impact on the top quark mass measurement. The uncertainty on the colour flow is partly covered by the experimental constraints on the fragmentation as discussed in Sec. 3.2. An independent study of the effect of colour flow on the b-jets energy scale in $t\bar{t}$ events has been completed [8].

3 b-quark Fragmentation Uncertainty

3.1 Theoritical and Experimental Background

The fragmentation of b-jets is governed by long distance QCD dynamics, and is thus very difficult to calculate from first principles. Consequently, the fragmentation of heavy quarks is described by phenomenological models with parameters that need to be determined from experimental input. Two of the most popular models are the Peterson [9] and Bowler [10] models that we are going to study in more detail. The heavy quark fragmentation models express the probability to observe a given ratio of the b-hadron and b-quark energy and momentum following the quark fragmentation. More rigorously, this is expressed in terms of the variable z, defined as

$$z \equiv \frac{(E+p_{||})_{hadron}}{(E+p)_{quark}}, \tag{1}$$

where the numerator is the sum of the hadron energy and momentum parallel to the quark and the denominator is the sum of the quark energy and momentum. The fragmentation function $f(z)$ parametrizes the probability density function of the variable z. The Peterson [9] fragmentation function is given by:

$$f(z) = N\frac{1}{z}(1-\frac{1}{z}-\frac{\varepsilon_b}{1-z})^{-2}. \tag{2}$$

This function has one parameter, ε_b, to be determined experimentally. N is just a normalization constant. The Bowler fragmentation function is given by:

$$f(z) = N\frac{1}{z^{1+bm_\perp^2}}(1-z)^a \exp(-\frac{bm_\perp^2}{z}), \tag{3}$$

where $m_\perp$ is the transverse mass of the hadron and a and b are the Bowler parameters to be determined experimentally.

An excellent laboratory for the study of b-quark fragmentation arises in $Z \to b\bar{b}$ decays, which have been studied at e^+e^- colliders by the SLD and the LEP experiments. Using their large $Z \to b\bar{b}$ datasets, those experiments have measured a variable similar to z defined above (but different because z is not accessible experimentally):

$$x_B^{wd} \equiv \frac{E_{had}}{E_{beam}}. \tag{4}$$

The variable E_{had} is the energy of the hadrons containing the b-quark and E_{beam} is the energy of the colliding beam. For experimental convenience, only the energy of the weak decaying b-hadrons are generally measured, thus the superscript "^{wd}". The measured b-hadrons spectrum is then compared to various theoritical models with their parameters fitted to yield the best representation of the data.

The most recent measurements have been made by the ALEPH [11], SLD [12] and OPAL [13] experiments. They use complementary techniques to measure x_B^{wd}. ALEPH reconstructs the kinematics of semi-exclusive decays $B \to l\nu D^{(*)}$, yielding precise energy measurement but with small data samples. The most precise measurements so far have used the kinematics of the tracks consistent with the secondary vertex of the b-hadrons decays. This is the technique employed for the SLD and OPAL measurements. The measured average values of x_B^{wd} ($< x_B^{wd} >$) are given in Table 1. The uncertainties on these measurements are $\approx \pm 0.5\%$ and they agree with each other at the ≈ 2 standard deviation level.

Table 1: Measured value of $< x_B^{wd} >$ for various experiments. The best fitted value of the Peterson parameters are also given for each measurement (It is not clear whether the value given by SLD is a fitted value, since no uncertainties are quoted [12]).

Measurements	$< x_B^{wd} >$	ε_b ($\times 10^{-4}$)
ALEPH [11]	0.716 $\pm$ 0.006(stat.) $\pm$ 0.006(syst.)	31 $\pm$ 3 $\pm$ 5
SLD [12]	0.709 $\pm$ 0.003(stat.) $\pm$ 0.003(syst.) $\pm$ 0.002(model)	55
OPAL [13]	0.7193 $\pm$ 0.0016(stat.) $^{+0.0038}_{-0.0033}$(syst.)	41.2 $\pm$ 0.7$^{+3.6}_{-3.5}$

We also give in Table 1 the fitted values of the Peterson parameter, ε_b, for the various measurements. The value published by SLD has no uncertainties, and it is unclear if it comes from a fit to the data. The authors have been contacted but have not replied yet. We show in Fig. 1 the data distribution of x_B^{wd} compared with the best fit of the Peterson model for the OPAL measurement [13]. The χ^2 over number of degrees of freedom (χ^2/ndf) for that fit is 159/45. The mediocre fit quality is partly understandable because the Peterson model has only one parameter to be fitted. The Bowler model has two parameters and yield better fit to the data (χ^2/ndf = 67/44), as shown in Fig. 2 for the OPAL measurement.

3.2 Impact on the Top Quark Mass Measurement

We have generated $t\bar{t}$ Monte Carlo events with variations of the b-quark fragmentation that are representative of the experimental knowledge presented in Sec. 3.1. We have used the PYTHIA Monte Carlo in which the Bowler and Peterson fragmentation functions are available (HERWIG has no obvious parameters directly related to the b-quark fragmentation). However, it is not possible to vary the Bowler parameters solely for the b-quark fragmentation in PYTHIA. Therefore, we have chosen to generate $t\bar{t}$ events with different values of the Peterson parameter.

Three samples of 20,000 $t\bar{t}$ events have been generated with the values of ε_b given in Table 2. The chosen range of the Peterson parameter is conservative:

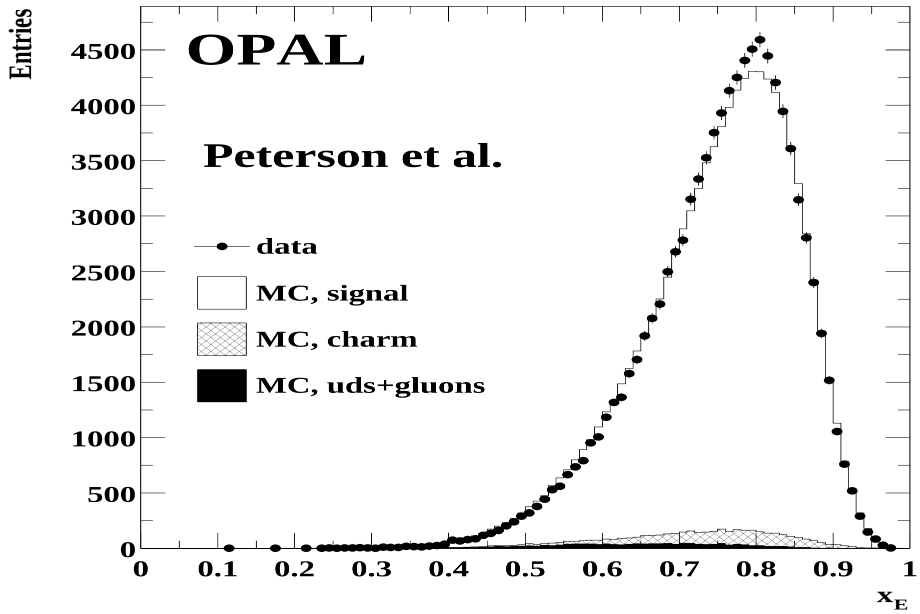

Figure 1: Data distribution of x_B^{wd} and the best fit of the Peterson model for the OPAL measurement [13].

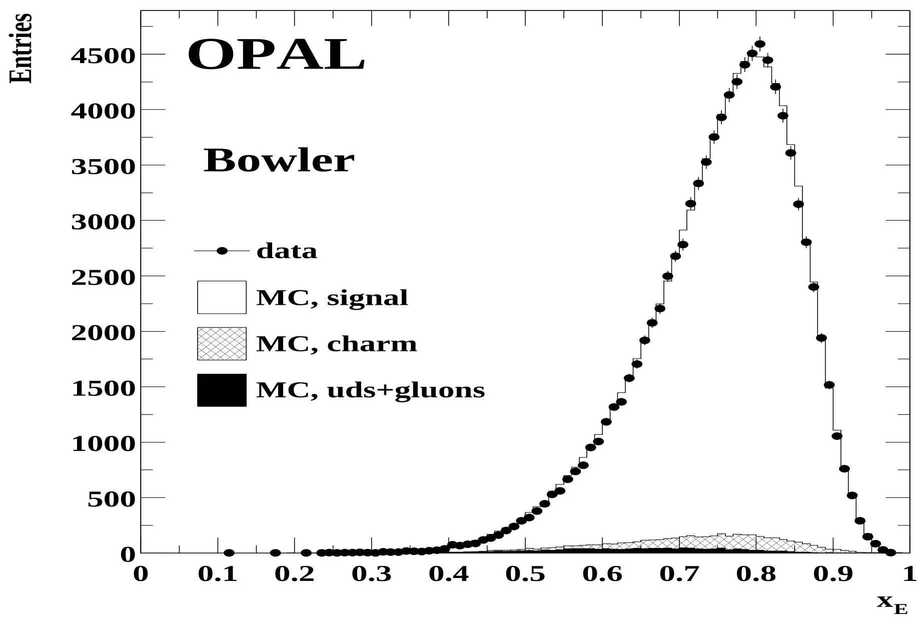

Figure 2: Data distribution of x_B^{wd} and the best fit of the Bowler model for the OPAL measurement [13].

ε_b = 25×10^{-4} corresponds to the lower bound of the ALEPH measurement, ε_b = 41×10^{-4} corresponds to the central value of the OPAL measurement and ε_b = 60×10^{-4} would correspond to the higher bound of the SLD measurement assuming it is a fitted value with uncertainties comparable to the OPAL measurement.

We verify that changing ε_b has a visible effect on the b-quark fragmentation in the generated PYTHIA $t\bar{t}$ samples. A variable analogous to x_B^{wd} is measured at the HEPG level in the different samples:

$$x' \equiv \frac{E_{had}^{1st}}{E_{part}}, \tag{5}$$

where E_{had}^{1st} is the energy of the b-hadron directly resulting from the b-quark fragmentation and E_{part} is the energy of the b-quark. The distributions of x' for each sample is shown in Fig. 3 and the average values are given in Table 2. The variations of $< x' >$ are significant between the samples: they are comparable in size with the variations of $< x_B^{wd} >$ from the lower to the higher experimental bound as given in Table 1. This gives us confidence that these samples yield reasonable variation of the b-quark fragmentation with respect to our experimental knowledge.

Table 2: Values of the Peterson parameter for the generated PYTHIA $t\bar{t}$ samples. The right-column shows the average value of x' (defined in Eqn. 5) for the various samples.

$\varepsilon_b(\times10^{-4})$	$< x' >$
25	0.705 $\pm$ 0.002
41	0.693 $\pm$ 0.002
60	0.679 $\pm$ 0.002

The generated events are required to pass the analysis selections [1]. The distributions of the event-by-event top quark mass are shown in Fig. 4. Very small differences can be observed between the mass templates with different values of ε_b.

Pseudo-experiments (as described in [1]) have been generated from these templates. The median of the pseudo-experiments mass distributions are given in Table 3. Also given is the result for the default PYTHIA sample (ttopei). This is a good cross-check since that sample uses a different fragmentation model: Bowler with $a = 0.3$ and $b = 0.58$ GeV^{-1}. As shown in Table 3, the top quark mass varies by $\lesssim 0.5\ \mathrm{GeV}/c^2$ when we vary the b-quark fragmentation.

We conclude from this exercise that the b-quark fragmentation has been measured by the LEP and SLD experiments to a level of precision corresponding to an uncertainty on the top quark mass of $\approx 0.2\ \mathrm{GeV}/c^2$ (half of the largest shift observed in Tab. 3). The size of this effect is small compared to the current statistical and generic jet energy scale uncertainty in the top quark mass measurement.

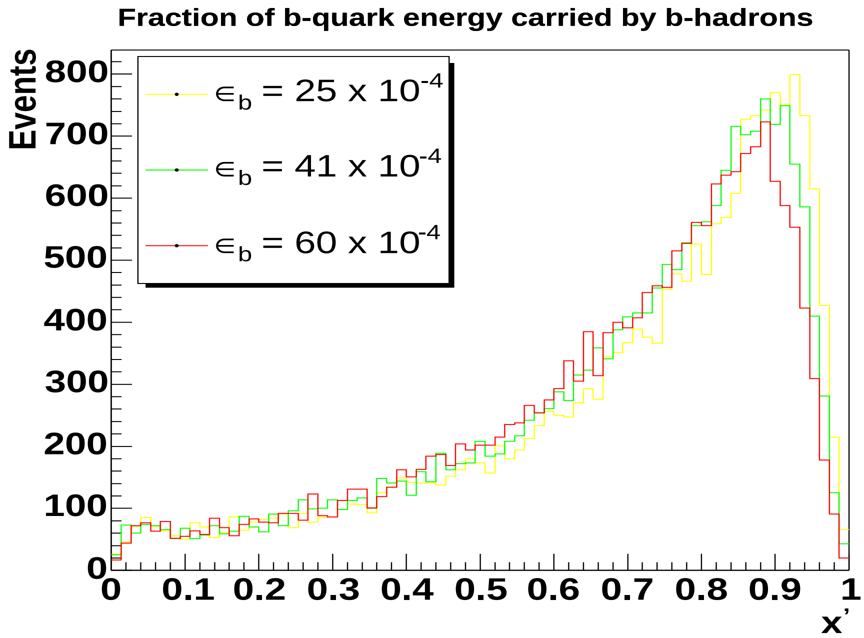

Figure 3: x' (see Eqn. 5) for various values of ε_b in PYTHIA $t\bar{t}$ events.

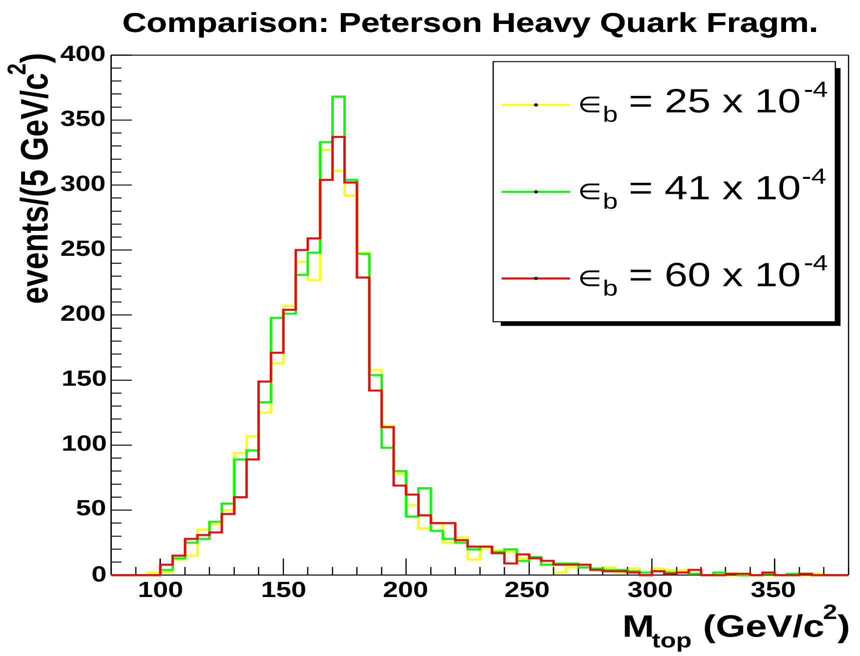

Figure 4: Top Mass templates for tagged lepton+jets events [1] for different values of ε_b.

As mentioned earlier, this study covers at least partly any possible uncertainty on the modelling of the colour flow in b-quarks decays observed in Z boson production. This is because a change in the colour flow affecting the b-quark would be observed as a variation of the b-hadron energy spectrum with respect to the b-quark. However, a independent study of this effect has been completed [8].

Table 3: Median of fitted mass from pseudo-experiments for $t\bar{t}$ events generated with various b-quark fragmentation.

Sample	M_{top} (GeV/c^2)
ε_b= 25$\times 10^{-4}$	174.66 $\pm$ 0.53
ε_b= 41$\times 10^{-4}$	174.22 $\pm$ 0.51
ε_b= 60$\times 10^{-4}$	174.23 $\pm$ 0.53
Default PYTHIA (Bowler: $a = 0.3$, $b = 0.58$ GeV^{-1})	174.45 $\pm$ 0.20

4 Uncertainty in the Decay

The b-hadrons decay differently than light quark hadrons. For instance, because of their higher mass, they decay to more particles in the final state on average and have a larger choice of decay reactions. However, the most striking characteristic of b-hadrons decay that can affect the jet energy scale is the abundance of semileptonic decays. The inclusive semileptonic branching ratio is (30.6 ± 2.7)% and (31.5 ± 2.4)% for B^+ and B^0 decays, respectively [14], while they constitute a much smaller portion in light hadron decays. The resulting muons, electrons and taus have a very different calorimeter response compared to the more common hadronic particles. Furthermore, these leptons are always accompanied by neutrinos that pass through the experimental apparatus undetected. Consequently, the average b-jets response is significantly lower than for the light quark jets coming from $\mathrm{W} \to jj$ decay in $t\bar{t}$ events.

We illustrate this in Fig. 5 where we show the response of b-jets (upper plot) and W-jets (bottom plot) as a function of the jet p_T for HERWIG $t\bar{t}$ events with M_{top} = $175\ \mathrm{GeV}/c^2$. The response is defined as

$$\frac{p_T(parton) - p_T(jet)}{p_T(jet)}, \tag{6}$$

where the jets are corrected up-to level 5. Furthermore, we can compare the response of generic b-jets with jets containing a semileptonic muon decay identified by the SLT tagger [15]. This is shown in Fig. 6 for SLT-tagged b-jets (upper plot) and non-SLT-tagged jets (bottom plot). The median response of W-jets, generic

b-jets and SLT-tagged jets are given in Table 4. W-jets have the largest response (9.5%) followed respectively by b-jets (13.6%) and SLT-tagged b-jets (40.6%).

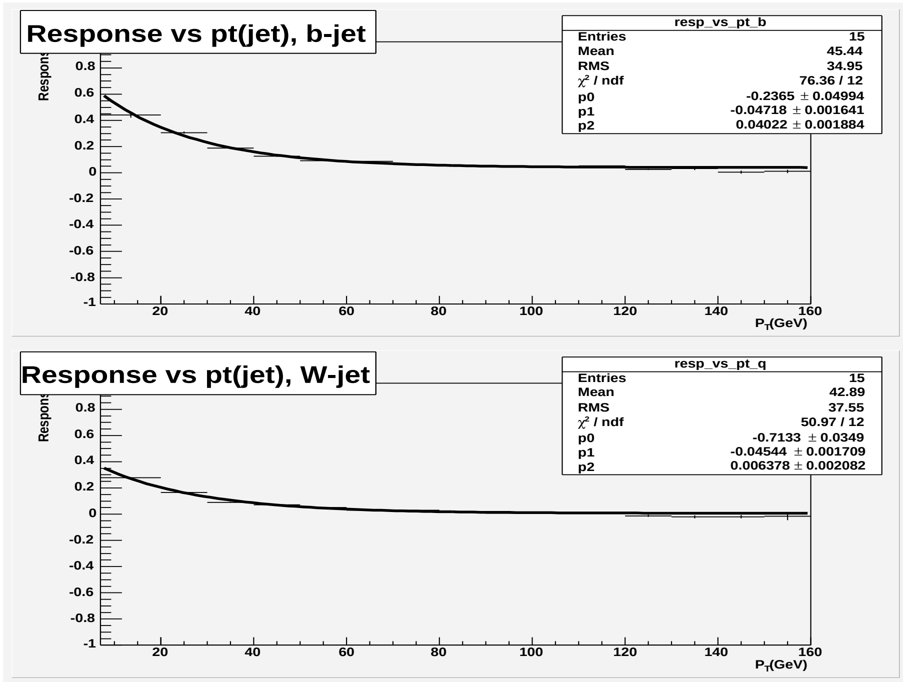

Figure 5: Response (defined as $(p_T(parton) - p_T(jet))/p_T(jet)$ for b-jet (upper plot) or W-jets (bottom plot) versus $p_T(jets)$.

Table 4: Median response ($\frac{p_T(parton)-p_T(jet)}{p_T(jet)}$ with jets corrected up-to level 5) of jets in HERWIG $t\bar{t}$ events with M_{top} = 175 GeV$/c^2$.

Jet type	Median Response (%)
W-jets	9.5
b-jets	13.6
SLT b-jets	40.6

We use these data to assign a conservative estimate of the top quark mass uncertainty due to the semileptonic b-hadrons branching ratio uncertainty. We start this calculation with the equation stating that the average b-jets response r_{bj} is a combination of the semileptonic jets response, r_{sl}, and jets without semileptonic decays (hadronic jets), r_{had}:

$$r_{bj} = f_{sl} r_{sl} + (1 - f_{sl}) r_{had}, \qquad (7)$$

where f_{sl} is the fraction of semileptonic jets. We want to calculate the uncertainty on the overall b-jets response due to the uncertainty on f_{sl}. The following variables are known:

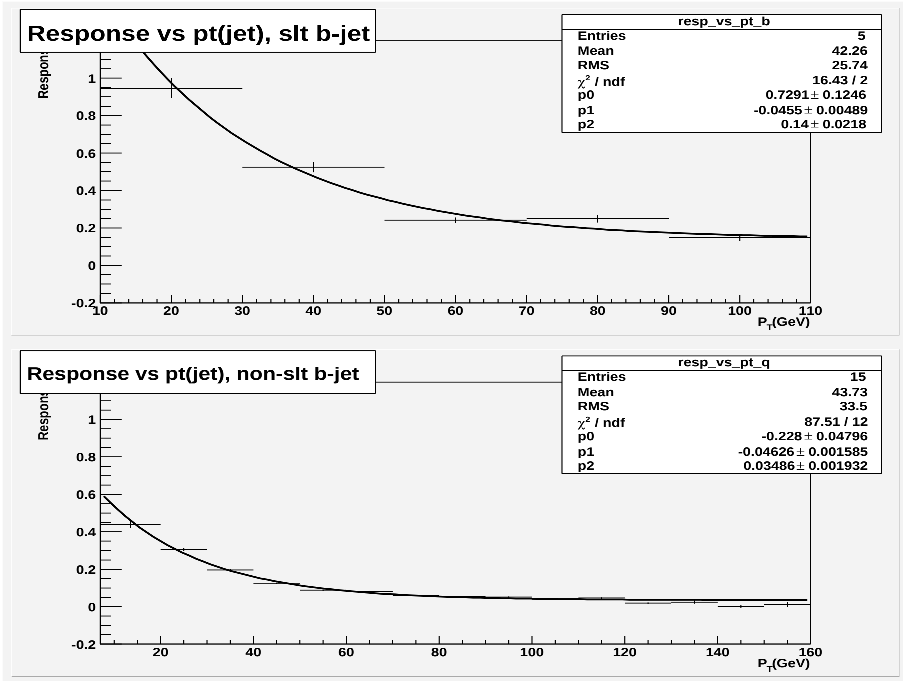

Figure 6: Response (defined as $\frac{p_T(parton)-p_T(jet)}{p_T(jet)}$ for b-jet (upper plot) or W-jets (bottom plot) versus $p_T(jets)$.

1. f_{sl} = 30 $\pm$ 3% as approximately given by the current world average on B-mesons inclusive semileptonic branching ratios [14].

2. $r_{had} = 9.5\%$ (as given in Table 4) assuming the response of hadronic b-jets is approximately equal to W-jets since the latter are mostly free of semileptonic decays.

3. $r_{bj} = 13.6\%$ as given in Table 4.

The only unkown is r_{sl} that we need to determine before calculating the uncertainty on r_{bj}. One estimate comes from the SLT-tagged b-jets response as given in Table 4: $r_{SLT} = 40.6\%$, but it is largely underestimated because SLT-jets contain only muons with $p_T \gtrsim 3$ GeV/c in order to pass the SLT requirements [15] and muons have a lower calorimeter response than electrons and taus on average. Alternatively, we can solve Eqn. 7 for r_{sl}:

$$r_{sl} = \frac{r_{bj} - (1 - f_{sl})r_{sl}}{f_{sl}} = 23.2\% \tag{8}$$

This calculation gives us a median response for semileptonic jets of $r_{sl} = 23.2\%$. This result confirms that the SLT-tagged b-jets response ($r_{SLT} = 40.6\%$) is a clear underestimation of the semileptonic jets response. We now use Eqn. 7 and compute the variation of r_{bj} by changing the semileptonic jets fraction f_{sl} by 3% corresponding approximatively to the current world uncertainty. We find an uncertainty of 0.4% on the total b-jets median response due to the semileptonic branching ratio uncertainty.

From the calculation of the jet energy scale uncertainty in the top mass analysis [1], we know that an average uncertainty of 1% in the jet energy scale correspond to an uncertainty of $\approx 1\ \mathrm{GeV}/c^2$ on the top quark mass. Since our calculation apply only to b-jets (and not to W-jets), it is conservative to apply this rule to get an uncertainty on the top quark mass of $0.4\ \mathrm{GeV}/c^2$.

In conclusion, we have calculated an uncertainty of 0.4% on the average b-jets energy scale from the current uncertainty on the semileptonic decay fraction and using the b-jets and W-jets response measured in $t\bar{t}$ Monte Carlo events. We extrapolate this uncertainty to the top quark mass uncertainty to get a conservative estimate of $0.4\ \mathrm{GeV}/c^2$.

5 Conclusion

The b-jets energy scale is a crucial uncertainty to the top quark mass measurement. It has been assumed so far that it is given by the generic jets energy scale uncertainty. In this note, we do a more careful examination by identifying special characteristics of b-jets and estimate the uncertainty from these properties using current experimental constraints. These specific properties of b-jets compared to generic or W-jets are the fragmentation, colour flow and decay.

We have estimated an uncertainty from b-quark fragmentation of $\approx 0.2\ \mathrm{GeV}/c^2$ using experimental constraints on b-quark fragmentation functions. This covers at least partly the uncertainty from the colour flow, although an independent study on the uncertainty arising from the modelling of this effect has been completed [8]. The main characteristic of b-jets decay subject to affect the jet energy scale is identified as the abundance of semileptonic decays. We calculate an uncertainty on the top quark mass of $\approx 0.4\ \mathrm{GeV}/c^2$ due to this effect.

In conclusion, the uncertainties specific to the b-jets energy scale are estimated to be small with respect to the total statistical and systematic uncertainty on the top quark mass measurement at this stage of Run II. We place upper conservative uncertainties on the top quark mass measurement arising from b-jet fragmentation and b-quark decay uncertainties of $0.2\ \mathrm{GeV}/c^2$ and $0.4\ \mathrm{GeV}/c^2$, respectively.

References

[1] E.M. Brubaker *et al.* “Updated Measurement of the Top Quark Mass in the Lepton+Jets Channel”, CDF Note 6845 (2004).

[2] K. Yorita *et al.* “Top Quark Mass Measurement by Dynamical Likelihood Method in the Lepton+Jets channel with SECVTX btags”, CDF Note 6829 (2004).

[3] I. Volobouev *et al.*. “Top Mass measurement in the Lepton+Jets Channel using a MultivariateTemplate Method”, CDF Note 6970 (2004).

[4] J.-F. Arguin *et al.* “Measurement of the Top Quark Mass Using the Template Method by Subdividing the Lepton+Jets Sample”, CDF Note 7139 (2004).

[5] A. Bhatti *et al.*, “Jet Corrections for summer 2003 analyses (offline version 4.10.4)”, CDF Note 6565, (2003).

[6] J.-F. Arguin and P.K. Sinervo, “Constraining the Jet Energy Scale Using W->jj Decays in ttbar Events”, CDF Note 7032, (2004).

[7] G.Busetto *et al.*, “A preliminary look at data collected by the Zbb trigger and extrapolations to 2 fb-1”, CDF Note 6036 (2002).

[8] S. Xie and P.K. Sinervo, “The Effects of Colour Flow on B Jet Energies in Top Quark Events,” CDF Note 7238 (September 2004).

[9] C. Peterson *et al.*, Phys. Rev. D27 (1983) 105.

[10] M.G. Bowler *et al.*, Z. Phys. C11 (1981) 169.

[11] ALEPH Collaboration, Phys. Lett. B512 (2001) 30.

[12] SLD Collaboration, Phys. Rev. D65 (2002) 092006.

[13] OPAL Collaboration, hep-ex/0210031 (2002), Submitted to Eur.Phys.J.C.

[14] S. Eidelman et al., Phys. Lett. B 592, 1 (2004)

[15] L. Cerrito, A. Taffard, “A Soft Muon Tagger for Run II“, CDF Note 6305 (2003).