



## Measurement of b-jet shapes at CDF

Alison Lister  
*UC Davis*  
(Dated: November 23, 2006)

For the first time at hadron colliders, the shape of b-jets is measured. This analysis uses about  $300 \text{ pb}^{-1}$  of CDF Run II data and covers a transverse momentum range from 52 to 300 GeV/c. The shapes of b-jets are expected to be significantly different for jets that contain a single b-quark inside the jet cone, i.e. b-quarks produced mainly by flavour creation, than for jets that contain two b-quarks. Jets with two b-quarks are mainly jets that are produced from a gluon splitting into a  $b\bar{b}$  pair. A sample of tagged jets as well as a sample of inclusive jets are used to extract, statistically, the shape of b-jets. The analysis methodology is presented here along with the preliminary results and systematic errors. The measured b-jet shapes are significantly broader than expected for Pythia Tune A and Herwig. This is likely to be coming from an underestimation in Leading Order Monte Carlos of the fraction of b-jets origination from gluon splitting. Decreasing the single b-quark jet fraction by 0.2 in Pythia Tune A and Herwig leads to a reasonable description of the measured b-jet shapes.

## I. INTRODUCTION

The internal structure of jets is dictated principally by multi-gluon emission from the primary parton. This process is described by fragmentation models. Multi-gluon emission involves high order QCD processes which are hard to calculate, so different models are used to implement the fragmentation in the simulation. On top of this, a good understanding of the hadronisation process is needed in order to compare simulation results with what is observed at detector level. Jet shapes are thus relevant quantities to study the overall decay structure leading to the observed jets. Moreover, the underlying event, an important component of any hadronic collision, also plays a role in the overall jet shapes.

Jet shapes are sensitive to whether the initial hard-scattered parton was a quark or a gluon. The inclusive jet shapes have been measured at CDF [1] and show that the ratio between the quark- and gluon-jet production cross sections is well reproduced by Pythia Tune A Monte Carlo simulation [2]; Pythia Tune A is tuned to the CDF Run I underlying event [3] [4]. It is also expected, but has never been measured at CDF, that jet shapes are sensitive to the quark flavour. Moreover, the shapes of b-jets are expected to be sensitive to the relative fraction of gluon splitting and flavour creation events. In the former case, the  $b$  and the  $\bar{b}$  quarks are expected to be most of the time inside the same jet [5], leading to significantly broader jet shapes than for the latter case. The fraction of gluon splitting events is an important parameter for the tuning of any Monte Carlo simulation. One of the aims of this analysis is to check if the fraction of b-jets originating from gluon splitting, as well as its evolution with  $p_T$ , is well described in the Monte Carlo models. This is particularly important for extrapolations to LHC energies.

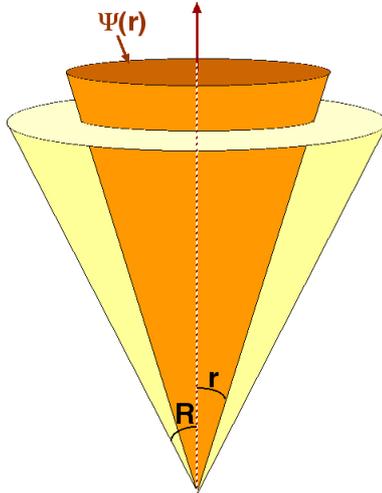


FIG. 1: Schematic drawing of the integrated jet shapes.

Jet shapes are defined as fractional transverse momentum ( $p_T$ ) distributions inside jets as a function of the distance away from the jet axis. The integrated jet shape is the fractional  $p_T$  inside a cone around the jet axis as illustrated in figure 1. This quantity can be computed at parton or hadron level in the Monte Carlo (MC) simulations and by using the calorimetric towers with  $p_T > 0.1$  GeV/c. Calorimeter towers with  $p_T > 0.5$  GeV/c and tracks with  $p_T > 0.5$  GeV/c are also considered in the study of the systematic uncertainties. The integrated shape is expressed as

$$\Psi(r/R) = \left\langle \frac{p_T(0, r)}{p_T(0, R)} \right\rangle \quad (1)$$

where  $R$  is the jet cone radius and  $p_T(0, r)$  is the sum of the transverse momentum of all objects inside a sub-cone of radius  $r$  around the jet axis. The integrated shapes are by definition normalised such that  $\Psi(r/R = 1) = 1$ . In Monte Carlo simulations, b-jets are defined as jets which have at least one b-quark inside the jet cone.

## II. DATA SAMPLE & EVENT SELECTION

This analysis is based on an integrated luminosity of about  $300 \text{ pb}^{-1}$  collected with the CDFII detector [6] between February 2002 and March 2004. This note presents results for central jets ( $|Y| \leq 0.7$ ) in a  $p_T$  range from 52 to 300 GeV/c. Four different datasets are used. The trigger paths for these datasets are similar, differing only by the cut thresholds and the nominal pre-scales. The four datasets select events with at least one jet with  $p_T > 20, 50, 70$  and  $100 \text{ GeV/c}$  respectively for each dataset. There are no other requirements on the events. These triggers are not fully efficient at the trigger threshold and the behaviour close to the trigger threshold might not be well understood. To avoid any trigger bias, each dataset is only considered when the trigger efficiency is above 99%.

The inclusive jet datasets are dominated by light flavoured and gluon jets. The b-jet content of the sample is enhanced by using the tight secondary vertex tagger (SecVtx). This tagging algorithm exploits the long life-time of the B-hadrons. Considering the large relativistic boost of the events, B-hadrons travel a few millimetres before decaying. The tagging algorithm is thus based on the reconstruction of secondary vertices inside the jets. The main limitation of this tagging method is the fact that hadrons that contain a c-quark can have decay lengths similar to those of B-hadrons in the CDF detector. Another limitation is the finite resolution of both the primary and secondary vertex locations.

In this analysis, jets are reconstructed using the MidPoint cone algorithm [7] with a cone size of 0.7 and a splitting/merging fraction of 75%. The following list summarises all the cuts applied to the jets in the samples

- One, and only one, primary vertex with  $|Z_{\text{vtx}}| \leq 50 \text{ cm}$  to ensure good secondary vertex reconstruction and to reject events with multiple  $p\bar{p}$  interactions;
- Missing  $E_T$  significance larger than a given threshold (3.5, 5.0, 6.0 and 7.0, respectively for the Jet20, Jet50, Jet70 and Jet100 datasets) to reject cosmic and beam halo backgrounds;
- $|Y_{\text{jet}}| \leq 0.7$  where the secondary vertex tagging algorithm and jet energy reconstruction are best understood;
- SecVtx tight tag on jets (for the tagged sample only).

An average jet correction is applied to correct the jet  $p_T$  to the particle level. It is calculated by matching, in the MC, hadron level and calorimeter level jets in the  $Y - \phi$  plane. The hadron level jet  $p_T$  is plotted as a function of the calorimeter jet  $p_T$ . The distribution obtained is fitted to a fourth order polynomial which is then used to obtain the corrected transverse momentum of the jets. The corrections are of the order of 20% at low  $p_T$ , down to 10% at higher  $p_T$ . The binning for this analysis and for all plots shown is done in corrected jet  $p_T$ .

## III. UNFOLDING METHOD

The sample of tagged jets used for this analysis does not contain only b-jets but also a background of jets that don't contain any b-quarks, the nonb-jets. The fraction of tagged jets that are b-jets is called the purity,  $p_b$ . The use of the SecVtx tagged jet sample biases the measured jet shapes due to the fact that the SecVtx algorithm requires jets with clean and well defined tracks. This bias is different for b-jets and nonb-jets. A bias term, dependent on the distance from the jet axis,  $r$ , must thus be added to correct for this effect. The bias terms,  $b_b(r/R)$  and  $b_{\text{nonb}}(r/R)$  for b- and nonb-jets respectively, are defined such that

$$\Psi_b^{\text{tag}}(r/R) = b_b(r/R)\Psi_b^{\text{det}}(r/R) , \quad \Psi_{\text{nonb}}^{\text{tag}}(r/R) = b_{\text{nonb}}(r/R)\Psi_{\text{nonb}}^{\text{det}}(r/R), \quad (2)$$

where the  $\Psi^{\text{det}}(r/R)$  terms represent the detector level shapes, before any tagging requirements. It is necessary to correct the b-jet shapes back to hadron level, i.e. to remove all influence of the tracker or calorimeters on the measurement. The corresponding correction factors,  $C^{\text{had}}(r/R)$ , are defined as

$$\Psi_b^{\text{had}}(r/R) = C_b(r/R)\Psi_b^{\text{det}}(r/R). \quad (3)$$

The final equation used to obtain the hadron level b-jet shape can be written as

$$\Psi_b^{\text{had}}(r/R) = C^{\text{had}}(r/R) \frac{\Psi^{\text{meas}}(r/R) - (1 - p_b)b_{\text{nonb}}(r/R)\Psi_{\text{nonb}}^{\text{det}}(r/R)}{p_b b_b(r/R)} \quad (4)$$

The different parameters of this equation are discussed in the following sections.

### A. Single b-quark jet content

Many of the distributions used for the unfolding are expected to be different for jets that contain one or two b-quarks. The parameters used in the unfolding are therefore somewhat sensitive to the fraction of b-jets that contain a single b-quark,  $f_{1b}$ . Because gluon splitting to  $b\bar{b}$  pairs only occurs as part of the fragmentation process and not in the matrix element in Leading Order (LO) MC, the fraction of gluon splitting is thought to be underestimated. For most jets where gluons split to a  $b\bar{b}$  pair, both b-quarks end up inside the same jet cone [5]. A comparison between the fraction of jets with more than one b-quark inside the same jet cone,  $1 - f_{1b}$ , predicted by Pythia Tune A and a Next to Leading Order (NLO) calculation is shown in figure 2. The maximum deviation between the Pythia Tune A and NLO prediction is of the order of 0.2. Before calculating any of the unfolding factors, the MC samples are re-weighted to decrease the  $f_{1b}$  fraction by 0.2. This lower  $f_{1b}$  fraction is used for the remainder of the analysis to obtain the secondary vertex mass templates as well as the tagging biases and hadron level corrections to the b-jet shapes.

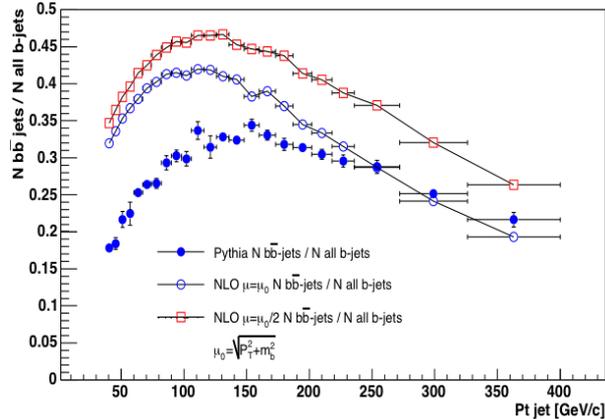


FIG. 2: Fraction of b-jets that contain more than one b-quark inside the same jet cone. The Pythia Tune A predictions are compared to NLO ones for two hadronisation and factorisation scales.

### B. Tagged Jet Shapes

The tagged jet shapes,  $\Psi^{\text{tag}}(r/R)$ , are defined as the average jet shapes, measured at calorimeter level, of all tagged jets in the samples.

### C. Inclusive Jet Shapes

Given the very low fraction of b-jets in inclusive jet production, less than 4%, it is possible to approximate the nonb-jet shapes to those of the inclusive jet shapes, before any tagging requirements. The assumption is that  $\Psi_{\text{nonb}}^{\text{det}}(r/R) \approx \Psi_{\text{incl}}^{\text{det}}(r/R)$ . The difference between these shapes, in Pythia Tune A MC, is negligible, with a maximum difference of less than 0.5%. No systematic uncertainty is therefore related to the use of this approximation.

### D. Purity

The fractions of b-jets in the tagged jet samples are extracted from a fit to the secondary vertex mass distributions for b- and nonb-jets. It is not possible to reconstruct the full hadron invariant mass mainly because of the presence of neutral particles in the B-hadron decays that are not detected in the tracking detectors. Nevertheless, the distribution of the invariant mass of the tracks used to find the secondary vertex, referred to as the secondary vertex mass, is significantly different for heavy flavoured jets and for light flavoured or gluon jets. Using the MC samples, distributions

of the secondary vertex masses for tagged jets are obtained for each  $p_T$  bin, separately for b- and nonb-jets. The secondary vertex mass distributions, as obtained from Pythia Tune A, for the second  $p_T$  bin are shown in the top of figure 3 (left). The measured distribution in data is fitted to the b- and nonb-templates, using a binned  $\chi^2$  minimisation method, to find the most likely fraction of jets that are b-jets. The bottom of figure 3 (left) shows the distribution in data along with the fitted distribution for the second  $p_T$  bin. The fit describes the data very well. Figure 3 (right) shows the extracted purity,  $p_b$ , as a function of the  $p_T$  of the jets fitting the data with templates from Pythia Tune A on the one hand and from Herwig [8] on the other hand.

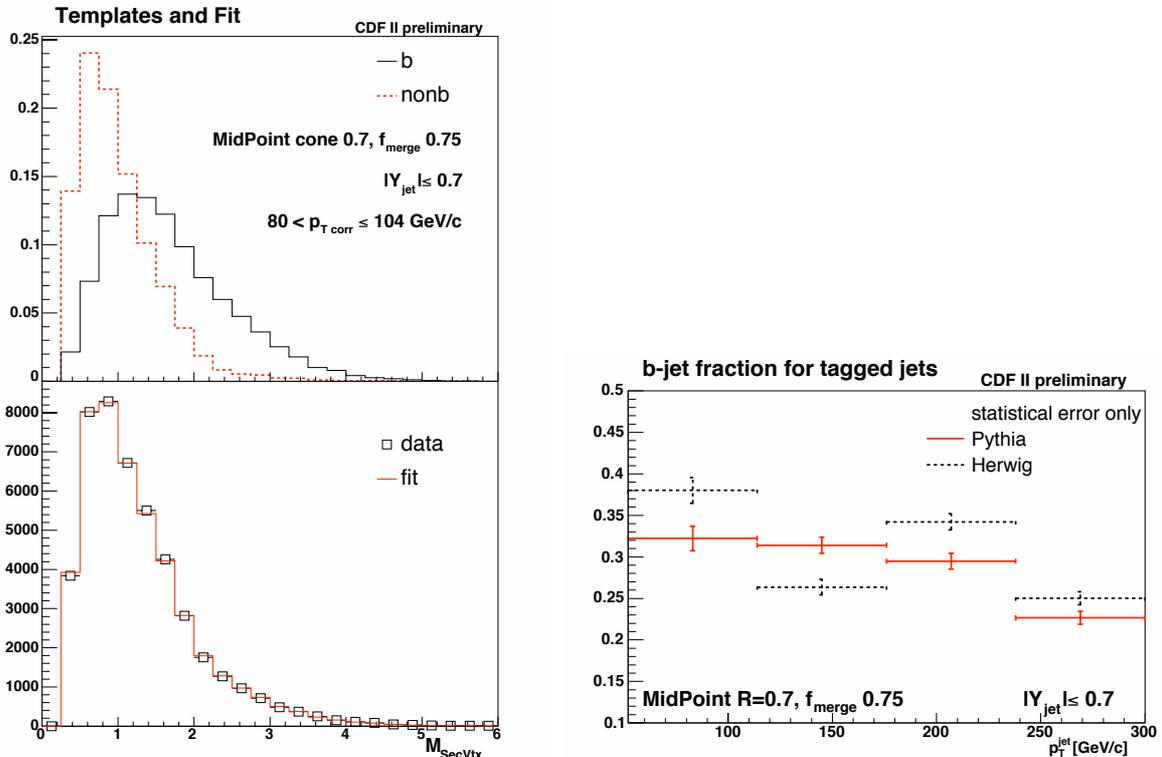


FIG. 3: (left) Normalised secondary vertex mass distributions for b- (black full line) and nonb-jets (red dashed line) as obtained from Pythia Tune A (top) and secondary vertex mass distribution in data (black open squares) compared to the fitted distribution (red line) (bottom) for the second  $p_T$  bin. (right) Extracted b-jet purity in data as a function of jet  $p_T$  using the templates obtained from Pythia Tune A (red full lines) and from Herwig (black dashed lines). The error bars indicate the statistical errors only.

### E. Biases Due to SecVtx Tagging

The requirement that the jets be tagged by the SecVtx tight algorithm introduces a bias in the measured jet shapes. These biases are different for each  $p_T$  bin, each bin in  $r$  and different for b- and nonb-jets. The bias terms are defined as the ratios, as obtained from MC, between the tagged and the inclusive jet shapes for b- and nonb-jets separately

$$b_b(r/R) = \frac{\Psi_{b \text{ MC}}^{\text{tag}}(r/R)}{\Psi_{b \text{ MC}}^{\text{incl}}(r/R)}, \quad b_{\text{nonb}}(r/R) = \frac{\Psi_{\text{nonb MC}}^{\text{tag}}(r/R)}{\Psi_{\text{nonb MC}}^{\text{incl}}(r/R)}. \quad (5)$$

Figure 4 shows the bias corrections obtained from Pythia Tune A for each  $p_T$  bin. The left hand plot shows the bias corrections for b-jets as well as single and double b-quark jets separately. The maximum bias for b-jets is of the order of 8%. The right hand plot shows the bias corrections for nonb-jets as well as the biases for c- and light+gluon-jets separately. The maximum bias for nonb-jets is of the order of 18%.

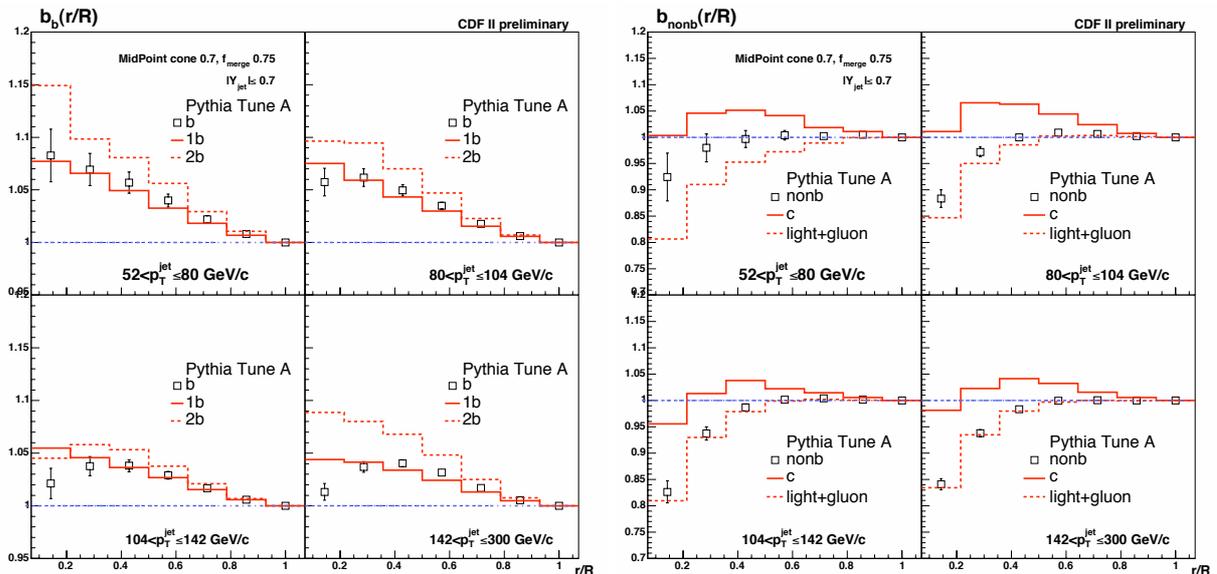


FIG. 4: (left) Biases due to the tagging on b-jets (black open squares) shown alongside the bias for single b-quark jets (full red lines) and double b-quark jets (dashed red lines). (right) Biases due to the tagging on nonb-jets (black open squares) shown alongside the bias for c- (full red line) and light+gluon-jets (red dashed line). The errors shown are the errors due to the MC statistics.

## F. Hadron Level Corrections

The hadron level correction factors,  $C^{\text{had}}(r/R)$  are evaluated from the MC for each bin in  $r$  and each bin in  $p_T$  and are defined as

$$C^{\text{had}}(r/R) = \frac{\Psi_{\text{bMC}}^{\text{had}}(r/R)}{\Psi_{\text{bMC}}^{\text{det}}(r/R)}, \quad (6)$$

where  $\Psi_{\text{bMC}}^{\text{det}}(r/R)$  are the MC b-jet shapes computed at calorimeter level, and  $\Psi_{\text{bMC}}^{\text{had}}(r/R)$  are the MC b-jet shapes computed at the hadron level. Figure 5 shows the correction factors for each  $p_T$  bin obtained using Pythia Tune A. The correction factors are of the order of 3% at most.

## IV. SYSTEMATIC UNCERTAINTIES

The different sources of systematic uncertainties for this analysis are described below.

To account for the sensitivity of the unfolding method to the variation of the  $f_{1\text{b}}$  fraction, it is decreased by 0.5. The difference in the measured b-jet shapes when using the default  $f_{1\text{b}} = 0.5$  instead of the normally used default  $f_{1\text{b}} = 0.2$  is taken as a systematic uncertainty.

The fraction of c-jets containing only one c-quark,  $f_{1\text{c}}$ , might also not be well described in LO MC. This fraction is decreased by 0.2, the same as the decrease to the  $f_{1\text{b}}$  fraction applied for the unfolding. The difference in the measured b-jet shapes when varying the  $f_{1\text{c}}$  fraction is taken as a systematic uncertainty.

The fraction of nonb-jets that are c-jets is assumed to be correct in the MC. From secondary vertex mass fits separating b-, c- and light-jets into three independent templates, a systematic error of 5% should be considered on the c-jet content of the nonb-jets.

In order to evaluate the effect of using a particular set of MC models for the fragmentation, hadronisation and underlying event, the whole unfolding is performed using Herwig MC samples instead of the Pythia Tune A samples.

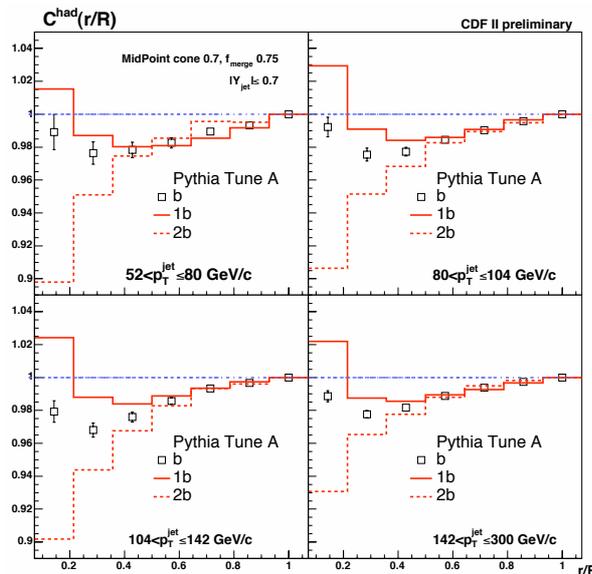


FIG. 5: Hadron level corrections to the b-jet shapes (black open squares). The errors shown are the MC statistical errors. The hadron level corrections for single b-quark jets (red full line) and double b-quark jets (red dashed line) are also shown.

The difference in the measured b-jet shapes obtained using these two MC samples is taken as a systematic uncertainty.

To investigate the calorimeter response to low  $p_T$  particles, the analysis is performed using only calorimeter towers with  $p_T > 0.5$  GeV/c and the difference with respect to the nominal measurement (which uses calorimeter towers with  $p_T > 0.1$  GeV/c) is taken as a systematic uncertainty.

To gauge any remaining detector effects, the whole analysis is performed using the tracks inside the jet cone instead of the calorimeter towers. Tracks with  $p_T > 0.5$  GeV/c are considered. The difference between these two methods is taken as a systematic uncertainty.

A 3% systematic uncertainty on the jet energy corrections is considered [9]. A variation of  $\pm 15\%$  on the Missing  $E_T$  significance is applied [1]. The cut on the location of the primary vertex is varied by  $\pm 5$  cm around the nominal cut at 50 cm. These variations are all found to have small effects on the final measurement. The dependence on the MC modelling of the SecVtx parameters was also investigated.

The total, statistical and systematic uncertainties are shown in figure 6 for each  $p_T$  bin and  $r$  bin. Also shown are the various contributions from the dominant effects.

The dominant sources of systematic errors vary as a function of the  $p_T$  bin. These are

- the difference in the b-jets shapes reconstructed using Pythia Tune A and Herwig,
- the difference in the b-jet shapes reconstructed from tracks instead of calorimeter towers,
- the  $f_{1b}$  variation from -0.2 to -0.5,
- the  $f_{1c}$  variation by -0.2,
- the jet energy scale.

## V. RESULTS

The measured integrated b-jet shapes are shown in figure 7. The results are compared to Pythia Tune A and Herwig predictions using the default  $f_{1b}$  fractions as well as the expected distributions if  $f_{1b}$  were decreased by 0.2. The left hand plot also shows the Pythia Tune A predictions for the inclusive jet shapes. The

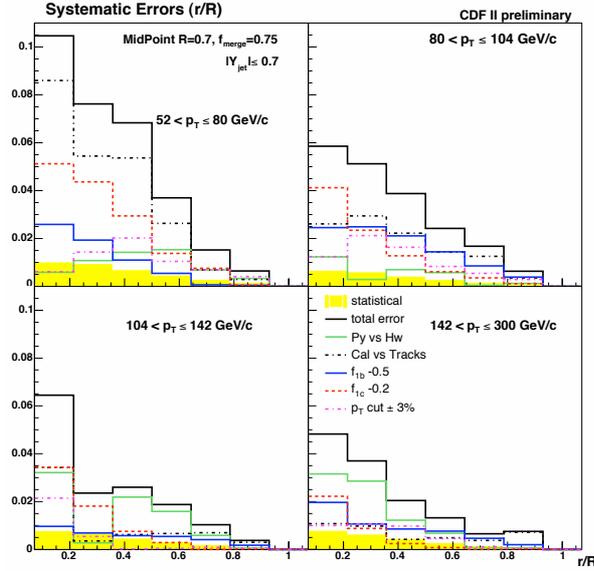


FIG. 6: Total error on the measurements for each jet  $p_T$  and  $r$  bin (black full lines). Also reported are the statistical errors (yellow bands) and the five dominant sources of systematic uncertainties.

Pythia Tune A and Herwig predictions for single and double b-quark jets are reported on the right hand plot. Figure 8 shows the ratio of the hadron level integrated jet shapes for various MC predictions over the measured values.

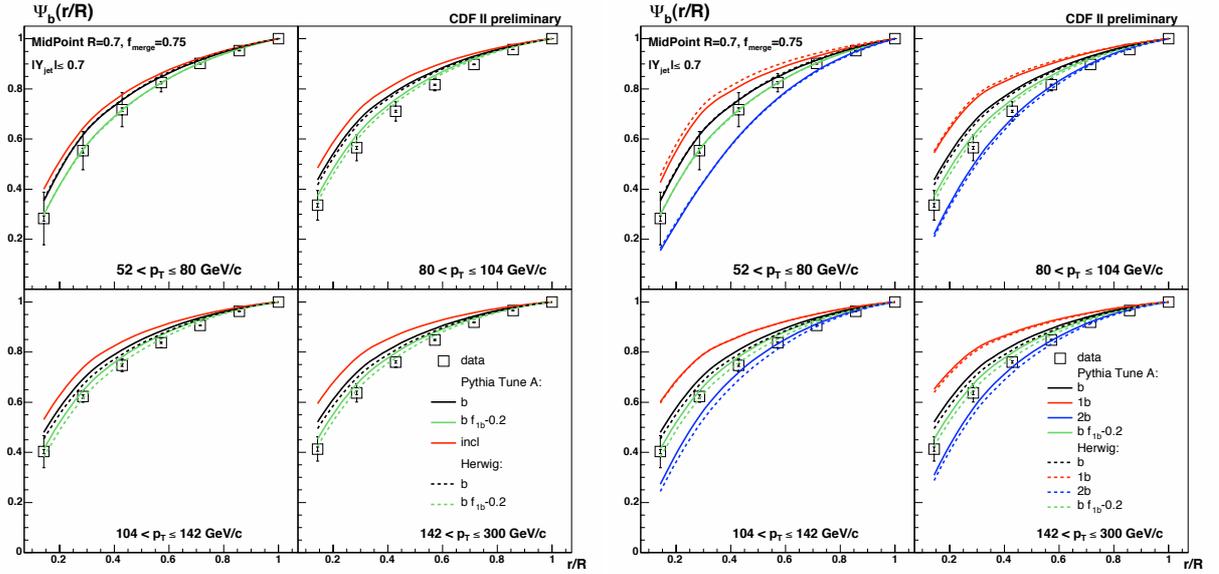


FIG. 7: Measured integrated b-jet shapes for each of the  $p_T$  bins considered. The results are shown as black open squares where the error bars represent the statistical and total errors. The statistical errors are smaller than the squares. The results are compared to Pythia Tune A (full lines) and Herwig (dashed lines) predictions using the default  $f_{1b}$  fractions (black lines) as well as the expected distributions if  $f_{1b}$  is decreased by 0.2 (green lines). (left) Also reported are the Pythia Tune A predictions for the inclusive jet shapes (red lines). (right) Also reported are the Pythia Tune A and Herwig predictions for single and double b-quark jets (red and blue lines, respectively).

Another way of looking at these results is to plot the fractional  $p_T$  outside a cone of fixed radius  $r$  as a function of the  $p_T$  of the jets. This gives an idea of the change in width of the jets as the energies increase. Jets of a particular

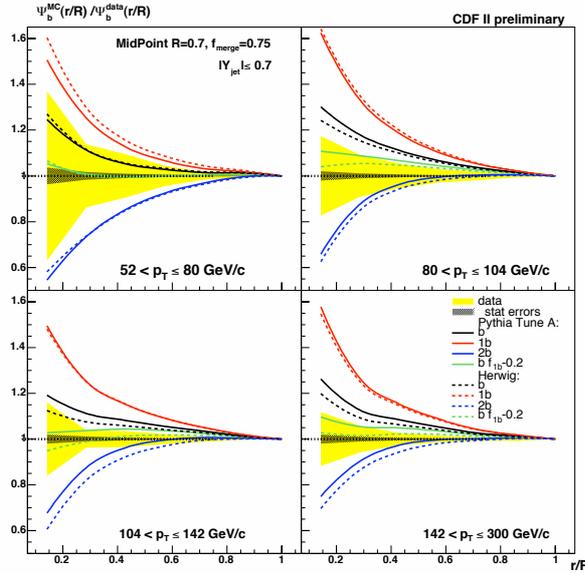


FIG. 8: Ratios of the hadron level integrated b-jet shapes for various MC predictions over the measured values. Pythia Tune A and Herwig predictions using the default  $f_{1b}$  fractions are shown as black lines (full lines for Pythia Tune A, dashed ones for Herwig). Pythia Tune A and Herwig predictions using  $f_{1b}$  fractions 0.2 below the default values are also reported (green lines) along with the predictions for single and double b-quark jet shapes (red and blue lines, respectively). The yellow bands represent the total errors on the measured b-jet shapes and the grey hashed lines are the statistical errors.

flavour are expected to become narrower as the  $p_T$  increases, mainly due to the running of the strong coupling constant,  $\alpha_s$ . There is also a small effect due to the boost of the jets. Figure 9 shows the evolution with jet  $p_T$  of the measured fractional  $p_T$  outside a cone of fixed radius  $r=0.2$ . The results are compared to Pythia Tune A and Herwig predictions using the default  $f_{1b}$  fractions as well as the expected distributions if  $f_{1b}$  is decreased by 0.2. The left hand plot also shows the Pythia Tune A predictions for the inclusive jet shapes as well as the previously published inclusive jet shape results which agree very well. The Pythia Tune A and Herwig predictions for single and double b-quark jets are reported on the right hand plot.

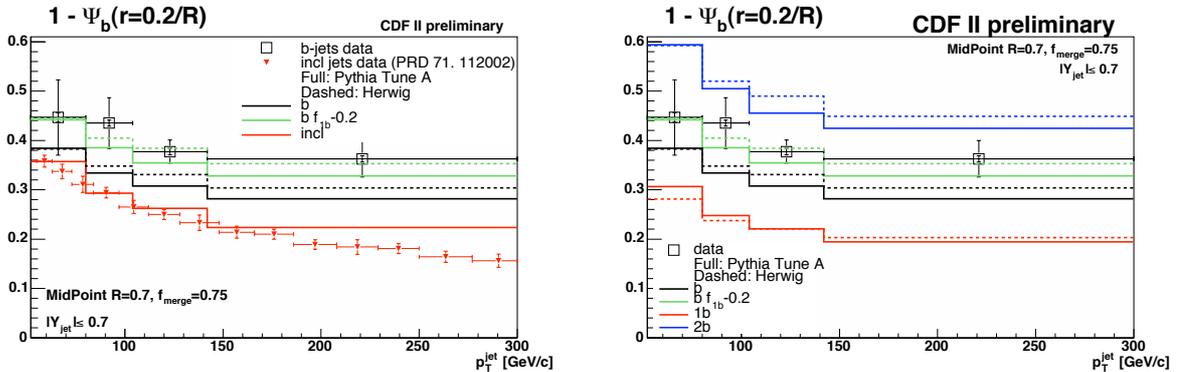


FIG. 9: Fractional  $p_T$  outside a cone of radius  $r = 0.2$  around the jet axis as a function of the  $p_T$  of the jet. The results for b-jet shapes are shown as black open squares and compared to different Pythia Tune A and Herwig predictions. The error bars on the plots represent the statistical and systematic errors. The statistical errors are smaller than the squares. The results are compared to Pythia Tune A (full lines) and Herwig (dashed lines) predictions using the default  $f_{1b}$  fractions (black lines) as well as the expected distributions if  $f_{1b}$  is decreased by 0.2 (green lines). (left) Also reported are the Pythia Tune A predictions for the inclusive jet shapes (red lines) as well as the previously published inclusive jet shape results (red triangles). (right) Also reported are the Pythia Tune A and Herwig predictions for single and double b-quark jets (red and blue lines, respectively).

The measured b-jet shapes are significantly broader than expected for Pythia Tune A and Herwig. This is likely

to be coming from an underestimation in Leading Order Monte Carlos of the fraction of b-jets origination from gluon splitting. Decreasing the single b-quark jet fraction by 0.2 in Pythia Tune A and Herwig leads to a reasonable description of the measured b-jet shapes.

### Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

- 
- [1] D. Acosta et al. [CDF Collaboration], Study of Jet Shapes in Inclusive Jet Production in p anti-p Collisions at  $\sqrt{s} = 1.96$  TeV, PRD **71**, 112002 (2005).
  - [2] T. Sjostrand et al., High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).
  - [3] R. Field and R. Craig Group, Pythia Tune A, Herwig, and Jimmy in Run 2 at CDF, arXiv: hep-ph/0510198.
  - [4] PYTHIA Tune A Monte Carlo samples are generated using the following tuned parameters in PYTHIA: PARP(67) = 4.0, MSTP(82) = 2.0, PARP(84) = 0.3, PARP(85) = 0.9, PARP(86) = 0.95, PARP(89) = 1800.0, PARP(90) = 0.25.
  - [5] S. Frixione et al, Heavy Quark Production, CERN-TH/97-16, arXiv: hep-ph/9702287 (1997).
  - [6] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, et al., Nucl. Instrum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, et al., Phys. Rev. D **52**, 4784 (1995); P. Azzi, et al., Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E.
  - [7] G. C. Blazey et al., arXiv: hep-ex/0005012 (2000).
  - [8] G. Corcella et al., JHEP **0101**, 010 (2001) [arXiv: hep-ph/0011363 (2000)]; arXiv: hep-ph/0210213 (2002).
  - [9] A. Bhatti et al. , Determination of the Jet Energy Scale at the Collider Detector at Fermilab, NIM A566 (2006) 375-412 [arXiv: hep-ex/0510047].