CDF/PHYS/JET/PUBLIC/8143
June 6, 2006

Measurement of b-quark Jet Shapes at CDF: Inclusive Method

Alison Lister [1]
University of Geneva (ETHZ), Switzerland

For the first time at hadron colliders, the shape of b-quark jets is measured. This analysis uses about $300\,\mathrm{pb}^{-1}$ of CDF Run II data and covers a transverse momentum range from 52 to 300 GeV. The shapes of b-quark 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. The measurement the b-quark jet shapes is presented here. A sample of tagged jets as well as a sample of inclusive jets are used to extract, statistically, the shape of b-quark jets. The analysis methodology is presented here along with the results and systematic errors. This measurement can be used to extract the fraction of gluon splitting events in b-jet production at the Tevatron. The agreement between the data and Pythia Tune A MC of the hadron level b-quark jet shapes, using the fitted values for the single b-quark jet fraction is found to be reasonable. The difference with respect to the previous note, CDF note 7973, is that not attempt is made to measure separately the shape of single and double b-quark jet shapes but a simpler, inclusive method is used. This is motivated by the fact that the total systematic errors for the previous method were very large, especially for double b-quark jets.

[1] alister@fnal.gov

Contents

Chapter 1

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 higher 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 at parton level 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, plays a non-negligible role in the overall jet shapes.
It is also expected, but has never been measured at hadron colliders, that the jet shapes are sensitive to the flavour of the jets. In particular it is interesting to look at the shapes of b-jets where the difference with respect to the inclusive jets is expect to be maximal. A comparison of the b-jet shapes in data with those predicted in Monte Carlo models for 1b and 2b jets allows to constrain the gluon splitting fraction in data, something which is useful for the tuning of the Monte Carlo models. This is particular important for extrapolations to LHC energies.
There was an unpublished analysis at HERA which looked at the jet shapes of c-quark jets. The results of this analysis are therefore a first measurement of the shapes of b-quark jets.
This note presents results for central jets ($|Y| \leq 0.7$) in a p_T range from 52 to 300 GeV on about $300\,\mathrm{pb}^{-1}$ of CDF Run II data. The analysis uses both calorimeter tower information and tracks.

1.1 Jet Shapes

The internal structure of jets is dictated principally by multi-gluon emission from a primary outgoing parton. This process is described by fragmentation models. Multi-gluon emission involves higher order QCD processes which are hard to calculate. Different models are thus used to implement the fragmentation process in the simulation. The implementations must often be tuned to reproduce the experimental data as accurately as possible. On top of this, a good understanding of the hadronisation process is needed in order to compare simulation results at parton level with what is observed in the detector. Moreover, the underlying event, an important component of any hadronic collision, plays a non-negligible role in the internal structure of jets. A relevant quantity for studying the internal structure of jets is the jet shape as will be discussed shortly but first a definition of what is meant by jet shapes is necessary.

Jet shapes look at the fractional transverse momentum (p_T) distribution inside the jets as a function of the distance away from the jet axis. This can be expressed either as a differential or integrated quantity. The differential jet shape is the rate of change of p_T with increasing distance away from the jet axis, whereas the integrated shape is the fractional p_T inside a cone around the jet axis as illustrated in figure 1.1. This quantity can be computed at parton or hadron level

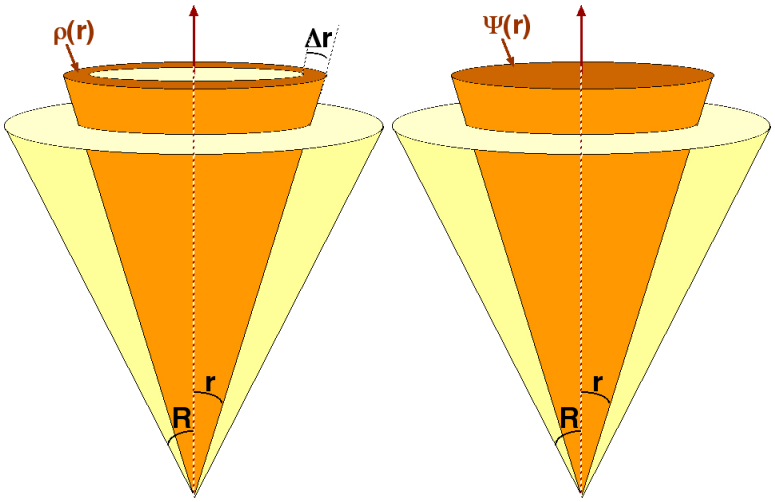

FIGURE 1.1: Schematic drawing of the differential (left) and integrated (right) jet shapes.

in the Monte Carlo (MC) simulations and by using either the tracks or the calorimetric towers at detector level. Formally, the distance away from the jet axis is defined as

$$r = \sqrt{(\Delta\phi)^2 + (\Delta Y)^2} \tag{1.1}$$

where $\Delta\phi$ and ΔY are the angular distances in the (ϕ,Y)-plane between the objects and the jet direction. Y is the rapidity which is given by

$$Y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{1.2}$$

and ϕ is the direction in the plane orthogonal to the beam-direction relative to the vertical. The integrated shape is expressed as

$$\Psi(r) = \left\langle \frac{p_T(0,r)}{p_T(0,R)} \right\rangle \tag{1.3}$$

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. Similarly, the differential jet shape is defined as

$$\rho(r) = \frac{\partial\Psi(r)}{\partial r} = \left\langle \lim_{\Delta r\to 0}\left(\frac{p_T(0,r+\Delta r) - p_T(0,r)}{p_T(0,R)\Delta r}\right)\right\rangle \tag{1.4}$$

The integrated shapes are normalised such that $\Psi(r = R) = 1$, i.e. the fractional transverse momentum of the objects inside a cone of radius equal to the jet cone radius around the jet axis is unity. Similarly, the differential shapes are normalised such that $\int_0^R \rho(r)dr = 1$. This requirement again comes from the fact that the fractional transverse momentum of the objects inside the jet cone must be equal to unity.

Jet shapes are sensitive to whether the initial hard-scattered parton was a quark or a gluon. The flavour inclusive jet shapes have been measured at CDF and show that the ratio between the quark- and gluon-jet production cross sections is well reproduced by the Pythia Tune A Monte Carlo simulation [1] [2].

[1] The parameters of Pythia Tune A were tuned to the CDF Run I underlying event. [1]

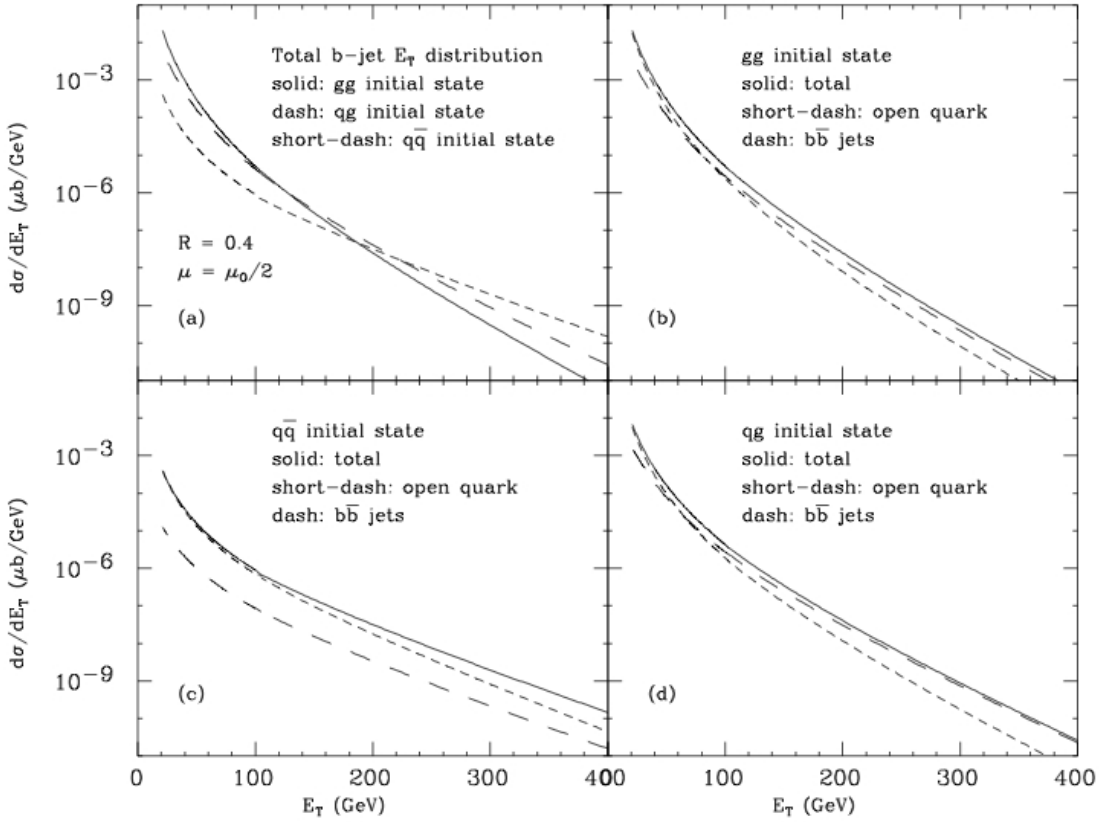

FIGURE 1.2: Differential b-jet cross section for different intial state configurations for a renormalisation/factorisation scale $\mu = \mu_0/2$ and jet cone radius of 0.4 (a). Differential components of the production process: $gg \rightarrow$ b-jet (b), $q\bar{q} \rightarrow$ b-jet (c) and $qg \rightarrow$ b-jet (d).

It is also expected, but has never been measured at CDF, that jet shapes are sensitive to the flavour of the quark jets. In particular it is interesting to look at the shape of b-quark jets where the difference with respect to the inclusive shape is expected to be maximal. Moreover, the shapes of b-quark 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 [3], 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. The b-jet cross section p_T dependence is sensitive to the relative fraction of gluon splitting to flavour creation b-jets, which is directly linked to the relative fractions of initial state production processes which are gg, qg or $q\bar{q}$ [3]. Figure 1.2, reproduced from [3], shows in the top left hand plot the relative contributions of the different initial state configurations to the differential b-jet cross section. The other three plots (b)-(d) show the differential b-jet cross section for each of the three initial state configurations along with the relative contribution due to jets which contain two b-quarks (called $b\bar{b}$ jets in these plots). These plots were made for a scale $\mu = \mu_0/2$ where $\mu_0 = \sqrt{p_T^2 + m_b^2}$ and for a jet cone algorithm with a cone of size 0.4. It is interesting to note that even though the probability that a gluon jet will split into a $b\bar{b}$ pair grows with jet energy, the fraction of primary gluons in the final state becomes smaller. One of the aims of this analysis is to check if the fraction of b-jets originating from gluon splitting, as well as it's evolution with p_T, is well described in the Monte Carlo models.

It is very important to have a good understanding of b-quark jets because they appear in a

number of searches for new physics both at the Tevatron and at future accelerators such as the Large Hadron Collider (LHC). The tuning of the Monte Carlo simulations at the Tevatron is very important for any extrapolation to the LHC energies. Such extrapolations provide a good basis for both qualitative and quantitative tests of the future sensitivity of many new physics channels at the LHC.

In Monte Carlo simulations, b-quark jets are defined as jets which have at least one b-quark inside the jet cone. The measurement presented in this thesis therefore combines jets originating from flavour creation with those where the b-quarks come from gluon splitting. The condition for a jet to be a b-quark jet can be expressed by the condition

$$\Delta R_{\mathrm{b-quark} \rightarrow \mathrm{jet\ axis}} \leq R_{\mathrm{jet\ cone}} \tag{1.5}$$

As mentioned above, the shapes of jets containing one or two b-quarks inside the jet cone are expected to be significantly different. Figure 1.3 shows the hadron level predictions using Pythia Tune A for the integrated b-quark jet shapes in four different p_T bins (see chapter 2 for the definition of the binning). Also shown in these plots are the predictions for inclusive jets as well as single and double b-quark jets. By measuring the shape of the b-quark jets and comparing them to the Monte Carlo predictions, it is possible to verify if the relative fraction of flavour creation versus gluon splitting jets is correctly implemented in the simulation, at least for heavy flavoured jets.

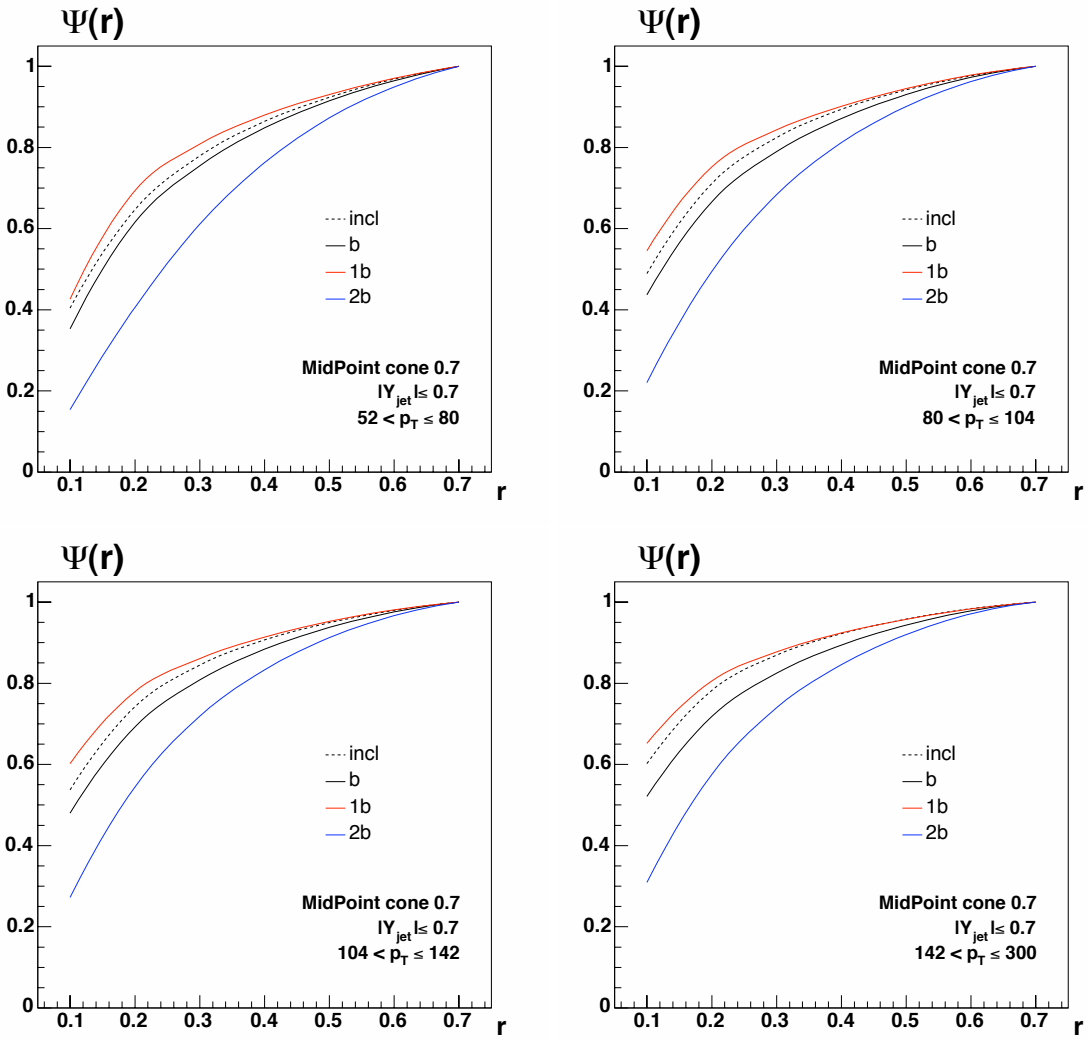

FIGURE 1.3: Hadron level predictions for the integrated jet shapes for b-quark jets and inclusive jets. Also shown are the predictions for single and double b-quark jets.

Chapter 2

Event Selection

In this chapter the description of the different datasets and MC samples is presented along with the event selection. First, in section 2.1 a description of the triggers is given for the different datasets used. Next, the different Monte Carlo samples used for this analysis will be described and in section 2.3 the selection cuts will be discussed. Section 2.4 discusses the average jet correction which corrects the p_T of the jets for detector and resolution effects.

2.1 Datasets

A sample of inclusive jet events is selected from $\sim 300\,\mathrm{pb}^{-1}$ of CDF Run II data collected from February 2002 through to September 2004. Four different datasets are used. These are called Jet20, Jet50, Jet70 and Jet100. The trigger paths for these datasets are similar, differing only by the cut thresholds and the nominal pre-scales. They are summarised in table 2.1. The basic idea of these triggers is to select events which have at least one jet which passes the trigger thresholds. There are no requirements on anything other than the jets.

At the trigger level 1, events are selected that have at least one calorimeter tower with the transverse component of the energy deposit above a nominal threshold (5 or 10 GeV). These trigger paths are called ST05 and ST10, respectively, where ST stands for single tower. A cut on the transverse energy of calorimetric clusters is then applied at level 2 (cut at 15, 40, 60 and 90 GeV for the Jet20, Jet50, Jet70 and Jet100 trigger paths, respectively). These triggers are called CL15, CL40, CL60 and CL90, where CL stands for cluster. A calorimetric cluster is defined as follows: starting from the seed towers, all towers whose p_T is above a certain threshold and that touch the seed tower are attached to the cluster. Then all towers, again passing the p_T cut, that touch the cluster are attached to the cluster. This iteration continues until the there are no more neighbouring calorimeter towers which have energy deposits above the p_T threshold. These are the stable calorimetric clusters.

Finally, at level 3, full jets are reconstructed using the JetClu cone algorithm with a cone of size 0.7. The jet algorithm is the same as the offline JetClu algorithm but at trigger level the primary vertex of the event is assumed to be at the centre of the detector, i.e. at (0,0,0) in the (Y, ϕ, z) coordinates of the detector. Cuts are applied on the transverse energy of the jets (20, 50, 70 and 100 GeV for the different datasets). The level 3 triggers for the different datasets are called J20, J50, J70 and J100, where J stands for jets.

Dataset	Level 1	Level 2	Level 3
Jet20	ST05	CL15	J20
Jet50	ST05	CL40	J50
Jet70	ST10	CL60	J70
Jet100	ST10	CL90	J100

TABLE 2.1: Summary of the trigger tables for the jet datasets.

2.2 Monte Carlo Samples

A number of different Monte Carlo (MC) samples are used both in the unfolding procedure of this analysis, presented in section 3, and as a final comparison with the data. The samples used are leading order MC simulations of the QCD $2 \to 2$ process that include initial and final state gluon radiation through parton showering.

The main samples used for the unfolding procedure of this analysis are generated using Pythia Tune A [1]. They use CTEQ5L PDFs and the default Lund fragmentation model [4][5]. Different samples were generated with different requirements on the minimum transverse momentum of the hard scattering; these are: Pt18, Pt40, Pt60, Pt90, Pt120, Pt150 and Pt200, where the number is the applied hard-scattering p_T cut. A p_T cut is needed on the hard-scattered process to avoid the effect of the divergence of the matrix element as $p_T \to 0$. The different cuts are used in order to have high statistics over a wide range of transverse momentum.

The fraction of b-jets in the inclusive jet samples is only of the order of 4%. In order to increase the b-jet statistics, samples were generated with the same transverse momentum cuts as above but which have an additional filter (HEPG filter) which selects only events which have at least one b-quark in the event. These samples are referred to as the b-filtered samples.

For comparison purposes and for systematic studies, samples generated using Herwig [6], with the same transverse momentum cuts as above, are also used. Only inclusive samples are available, not b-filtered ones.

To investigate the influence of the PDFs on the unfolding procedure as well as on the final shapes, a sample is generated using Pythia Tune A but changing the PDFs to CTEQ6L [7]. This sample was generated using a transverse momentum cut at 60 GeV and applying the additional b-quark filter to enhance the b-jet content.

Further samples are generated using Pythia Tune A but changing the fragmentation model used for heavy quarks to the Peterson model with the ϵ_b parameter set to 0.006 [8]. In the Peterson model, the fragmentation function for heavy quarks is described by

$$D_{\mathrm{Q}}^{\mathrm{b,c}}(z) = \frac{N}{z(1-\frac{1}{z}-\frac{\epsilon_{\mathrm{b,c}}}{(1-z)})}. \tag{2.1}$$

The most likely value for the ϵ_b parameter is 0.006 ± 0.002 at the Tevatron [9]. Two samples of this type were generated with transverse momentum cuts at 18 and 120 GeV and with the additional b-quark filter. Table 2.2 shows a summary of all the MC samples used.

[1] The different parameters of Pythia are tuned to the CDF Run I underlying event.

Pythia Tune A inclusive	Pythia Tune A HEPG b-filter	Herwig inclusive	Pythia Tune A CTEQ6L HEPG b-filter	Pythia Tune A Peterson frag. HEPG b-filter
Pt18	Pt18	Pt18		Pt18
Pt40	Pt40	Pt40		
Pt60	Pt60	Pt60	Pt60	
Pt90	Pt90	Pt90		
Pt120	Pt120	Pt120		Pt120
Pt150	Pt150	Pt150		
Pt200	Pt200	Pt200		

TABLE 2.2: Summary of the MC samples used.

2.3 Event Selection

In this analysis, jets are reconstructed using the MidPoint cone algorithm with a cone size of 0.7 and a splitting/merging fraction of 75%. Events are selected which have at least one jet in the central rapidity region ($|Y_{\text{jet}}| \leq 0.7$).

For this analysis, it is necessary to ensure that, when the data were taken, not only the calorimeters were functioning properly but also that the silicon detectors were on and functioning properly. At CDF, each run is assigned something like a "quality stamp". This has quality bits for each potential type of analysis requirement. In this analysis, the requirement is that the silicon detector and the QCD bits are set [2].

A cut is applied on the total missing E_T significance of the event to remove a large fraction of the cosmic background [3]. It must be below a certain threshold that is dependent on the dataset used. These thresholds are: 3.5, 5.0, 6.0 and 7.0, respectively, for Jet20, Jet50, Jet70 and Jet100, the same as those used for most other QCD analyses [10][11].

This analysis considers all jets which are in the central rapidity region where the secondary vertex tagging algorithm is the best understood. A cut on $|Y_{\text{jet}}| \leq 0.7$ is thus applied.

To ensure good secondary vertex reconstruction, a cut is required on the z-component of the primary vertex which must be within 50 cm of the centre of the detector. To remove any potential effects due to multiple parton interactions on the final jets, all events with multiple primary vertices were removed.

As mentioned previously, the b-jet content of the jet samples is relatively low, below 4%. It would therefore be impossible to make any precision measurement of b-jet properties with such a low purity. It is thus necessary to require that the jets be positively tagged by the SecVtx tight tagger. This increases the b-jet fraction to $20 - 40\%$ for high and low p_T jets, respectively, as shown in section 3.3. The requirement that the jets be tagged introduces a bias in the selected sample. Nevertheless, this can be accounted for in this analysis, as will be shown in chapter 3. It should be noted that the sub-cone inside which tracks are considered by the SecVtx algorithm is maintained at the default value of 0.4 [12].

[2] Version 7 of the good run list is used.

[3] The missing E_T significance is defined as $\frac{E_T^{\text{mis}}}{\sqrt{\sum E_T}}$, where E_T^{mis} and $\sum E_T$ are the missing and total transverse energy of the event, respectively.

The following list summarises all the cuts applied to the jets in the samples, after requiring the runs to pass the quality cuts

- One (and only one) primary vertex with $|Z_{\text{vtx}}| \leq 50$ cm,
- Missing E_T significance larger than a given threshold (dependent on the dataset),
- $|Y_{\text{jet}}| \leq 0.7$,
- SecVtx tight tag on jets (for the tagged sample only).

2.4 Average p_T Correction

Jets, therefore also jet shapes, are computed at the level of calorimeter towers but it is important to obtain results that are detector independent. Jet properties are affected by the detector response. The detector resolution is not perfect and is not necessarily the same over the whole ϕ and η region.

For this reason it is necessary to correct the p_T of the jets for all detector effects. This correction is called the hadron level correction. It corrects the p_T of the jets back to what they are expected to have been before the particles entered the detector. This correction factor is calculated in this analysis by matching, in MC, hadron level and calorimeter jets based on an angular selection in the $Y - \phi$ plane as detailed in [10]. For each tagged calorimeter jet in the range $|Y_{\text{jet}}| < 0.7$, the angular difference, ΔR, between that jet and each of the hadron level jets in the event is computed. The hadron level jet associated to the calorimetric jet is taken to be the one with the smallest ΔR (as long as this difference is smaller than 0.7, the jet cone radius). A scatter plot of calorimetric jet p_T versus hadron level jet p_T was produced for the whole p_T range covered by the jet datasets and is reproduced in figure 2.1. The scatter plot is fitted to a fourth order polynomial. The fit results give a correction function

$$p_T^{\text{corr}} = 0.5 + 1.24\,p_T - 0.0012\,p_T^2 + 4 \cdot 10^{-6}\,p_T^3 - 5 \cdot 10^{-9}\,p_T^4 \tag{2.2}$$

where p_T is the measured calorimetric transverse momentum and p_T^{corr} is the corrected transverse momentum of the jet. This correction function is shown on the right hand side of figure 2.1. The corrections are of the order of 20% at low p_T, down to 10% at higher p_T. The last two terms of this correction function are only important for very large p_T jets.

Except where explicitly mentioned, the binning for this analysis and for all plots shown is done in corrected p_T, from now simply referred to as p_T, to enable a better comparison with other experimental measurements or theoretical models.

2.5 p_T Thresholds

Because of its nature, the trigger is not fully efficient at the trigger threshold and the behaviour close to the trigger threshold might not be well understood. A similar effect is seen in the MC samples where there is a turn-on effect.

It is therefore necessary to consider each dataset only in the p_T region where the trigger effects are negligible. This is assumed to be the case when the trigger efficiency is above 99%. The dataset thresholds are thus taken to be the same as those used in [10]. The trigger efficiency curves as a function of raw, calorimeter level, p_T are shown in figure 2.2. These plots are

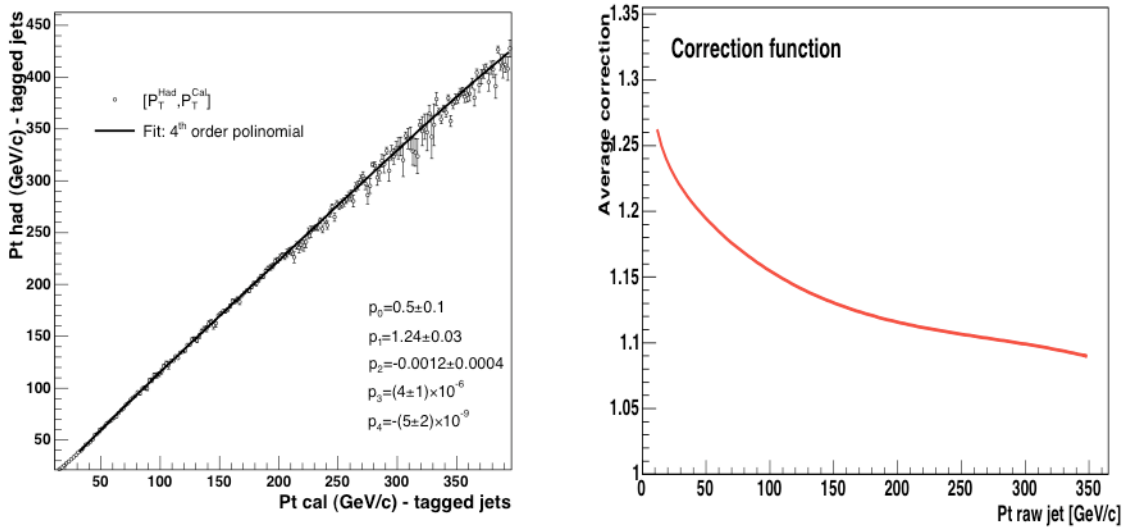

FIGURE 2.1: Scatter plot of the hadron level vs calorimetric jet p_T for a wide range of p_T (left). The correction function applied is shown in the right hand plot. These plots are taken from [10].

Dataset	p_T^{raw} limit [GeV/c]	p_T^{corr} limit [GeV/c]
Jet20	43	52
Jet50	67	80
Jet70	89	104
Jet100	123	142

TABLE 2.3: p_T threshold for the different datasets in raw and corrected p_T.

taken from [10]. Table 2.3 shows the thresholds in raw and corrected p_T for the different datasets.

In MC, there is a similar problem to that of the trigger efficiency. It is due to the cut on the minimum p_T of the hard-scattered partons. The jet energy will not be the same as the energy of the hard-scattered partons; it is therefore necessary to make sure that the jets are not biased by this effect. The final binning of this analysis is more conservative than the one used for the b-jet cross section measurement [10]. Only one MC sample is used for each data bin. The ranges in p_T of the four final bins, along with the datasets and MC samples used, are shown in table 2.4. Also shown in this table are the total number of inclusive and tagged jets in data that pass all the selection cuts.

p_T limits [GeV/c]	Dataset	MC sample	N_{jets}	$N_{\text{tagged jets}}$
52-80	Jet20	Pt18	161'524	4'677
80-104	Jet50	Pt40	354'922	13'367
104-142	Jet70	Pt60	134'907	5'874
142-300	Jet100	Pt90	377'650	18'673

TABLE 2.4: Final binning in jet p_T, showing the datasets and MC samples used for each p_T bin. Also shown are the total number of jets and the total number of tagged jets in data which pass all the selection cuts.

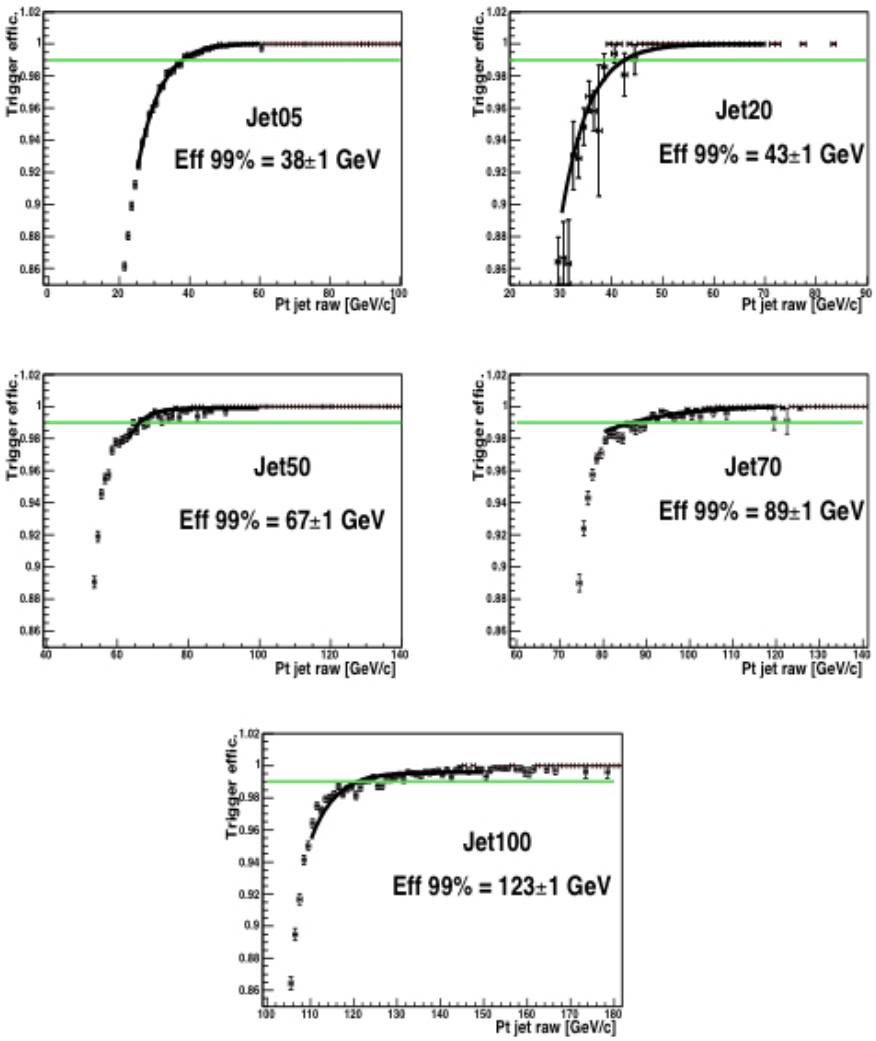

FIGURE 2.2: Trigger efficiency as a function of raw jet p_T for the different jet datasets, showing as a horizontal green line the 99% threshold used. The figures are taken from [10].

Chapter 3

b-quark Jet Shapes

The subject of this analysis is the study of b-quark jet shapes. As described in section 1.1, a b-jet is defined as a jet which contains at least one b-quark. These jets are mainly flavour creation jets (for the cases where there is one b-quark inside the jet) and jets from gluon splitting processes where the gluon splits into a $b\bar{b}$ pair. For the latter case, the two b-quarks are most of the time expected to be inside the same jet cone, as was shown in section 1.1. The fraction of b-quark jets from gluon splitting has not been measured at CDF Run II. It is important measure this fraction experimentally in order to check if the MC models describe the interplay between the initial state gluon production, where the gluon splits into a $b\bar{b}$ pair, and the initial state heavy flavour production. In section 4.3, a plot comparing the Pythia Tune A predictions to a next-to-leading order calculation shows that the gluon splitting rate is most likely underestimated in Pythia Tune A. The shapes of b-quark jets, in MC, show a significant dependence on the relative fraction of b-quark jets containing one or two b-quarks inside the same jet cone as was shown towards the end of chapter 1. The measurement of b-quark jet shapes is thus a good test to verify the modelling of the gluon splitting process in MC. This is particularly useful for extrapolations of such processes to LHC energies, where a good understanding of b-jets is necessary for many searches for new particles.

In chapter 1.1 jet shapes were introduced. In practise, the average jet shapes are measured, where the average is taken over all jets in the samples. Equation 1.3 for the integrated shapes becomes

$$\Psi(r) = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \frac{p_T(0,r)}{p_T(0,R)}. \tag{3.1}$$

For the differential shape, equation 1.4 becomes

$$\rho(r - \Delta r/2, r + \Delta r/2) = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \frac{p_T(0, r + \Delta r/2) - p_T(0, r + \Delta r/2)}{p_T(0,R)\,\Delta r}, \tag{3.2}$$

where Δr is the bin size in r. It is not possible to obtain, in data, even with good tagging, a sample which consists of only b-jets. It is not even possible to obtain a sample containing only heavy flavoured jets. For this reason, it is necessary to extract the jet shapes of b-quark jets from a background of light-, gluon- and c-jets. This can only be done statistically.

For the remainder of this thesis, the term "raw" will be used for quantities measured at the level of the detector, using calorimetric or track information, before any correction factors are applied. The terms "corrected" or "hadron level" are used when the effects of the detector have been taken into account or, in MC, when the quantities are calculated directly using the final set of particles, after the fragmentation and hadronisation processes. Hadron level or corrected

quantities are therefore detector independent.

In the next section, the unfolding procedure is presented and all its different parameters are explained. The raw shapes for both inclusive and tagged jets are then shown. Section 3.3 details the method used to calculate the purity of the sample. Section 3.4 describres the bias corrections used to correct for the effect of applying the SecVtx tight tagger. In section 3.5 the hadron level corrections to the jet shapes are shown. These are the correction factors applied to the b-jet shapes in order to correct the shapes for detector effects. Finally, section 3.6 presents the hadron level b-quark jet shapes showing only the statistical errors. Chapter 4 presents the systematic studies carried out for this analysis. Chapter 5 presents the final results for the b-quark jet shapes.

3.1 Unfolding Method

As mentioned above, the sample of tagged jets used for this analysis does not contain only b-quark jets but also a background of jets that don't contain any b-quarks; these are called nonb-jets. The measured shape, after tagging, will thus be a combination of the b-jets and nonb-jets which can be written as

$$\Psi_{\text{meas}}(r) = p_{\text{b}}\Psi^{\text{b}}_{\text{meas}}(r) + (1 - p_{\text{b}})\Psi^{\text{nonb}}_{\text{meas}}(r), \tag{3.3}$$

where p_{b} is the fraction of jets which contain at least one b-quark; this is called the purity. The superscripts b and nonb represent b- and nonb-jets, respectively, and the shapes are the raw, measured shapes after tagging. Although the jet algorithm is only run at calorimeter level, the jet shapes can be measured using either the calorimetric towers or the tracks that are inside the jet cone.

The fact that a SecVtx tagged jet sample is used most probably biases the measured shapes. These biases could be due to the fact that the SecVtx algorithm requires jets with cleaner, better defined tracks which could lead to cleaner jet shapes with fewer soft particles. This bias is most likely different for b-jets than for nonb-jets. A bias term, which depends on the distance from the jet axis, r, must thus be added to correct for this effect in order to extract the unbiased detector level shapes for b- and nonb-jets. These bias terms, $b^{\text{b}}(r)$ and $b^{\text{nonb}}(r)$ for b- and nonb-jets respectively, are defined such that

$$\Psi^{\text{b}}_{\text{meas}}(r) = b^{\text{b}}(r)\Psi^{\text{b}}_{\text{det}}(r), \tag{3.4}$$

and

$$\Psi^{\text{nonb}}_{\text{meas}}(r) = b^{\text{nonb}}(r)\Psi^{\text{nonb}}_{\text{det}}(r), \tag{3.5}$$

where the $\Psi_{\text{det}}(r)$ terms represent the detector level shapes, before any tagging requirements. Combing equations 3.3 through 3.5 gives for the measured shape, after tagging,

$$\Psi_{\text{meas}}(r) = p_{\text{b}}b^{\text{b}}(r)\Psi^{\text{b}}_{\text{det}}(r) + (1 - p_{\text{b}})b^{\text{nonb}}(r)\Psi^{\text{nonb}}_{\text{det}}(r). \tag{3.6}$$

This equation can be re-written, extracting the b-jet shape, as

$$\Psi^{\text{b}}_{\text{det}}(r) = \frac{\Psi_{\text{meas}}(r) - (1 - p_{\text{b}})b^{\text{nonb}}(r)\Psi^{\text{nonb}}_{\text{det}}(r)}{p_{\text{b}}b^{\text{b}}(r)}. \tag{3.7}$$

It is important that no measured quantity is influenced by the particular experimental setup. This is particularly important when comparing results with theoretical models or to other experimental results. For this reason, it is necessary to correct the b-jet shapes back to hadron

level, i.e. to remove all influence on the measurement of the tracker or calorimeters. This is in part done by an average correction to the p_T of the jets, discussed in section 2.4. It is also important to correct the actual jet shapes back to hadron level. This removes, among other things, the influence of the magnetic field and the calorimeter response on the jet shapes. Correction factor, $C_{\rm had}(r)$, are thus applied to extract the b-jet shapes at hadron level. Equation 3.7 can be written in it's final form as

$$\Psi^{\rm b}_{\rm had}(r) = C_{\rm had}(r)\frac{\Psi_{\rm meas}(r) - (1-p_{\rm b})b^{\rm nonb}(r)\Psi^{\rm nonb}_{\rm det}(r)}{p_{\rm b}b^{\rm b}(r)} \tag{3.8}$$

This last equation is the equation used to unfold the hadron level b-jet shapes from the raw tagged and inclusive jet shapes. Appendix A contains a discussion into the reasons why it was not possible, in this analysis, to measure the differential b-quark jet shapes using a similar unfolding method.

It is apparent from the above discussion that in order to measure the shapes of b-quark jets it is necessary to measure a number of other parameters. These will be discussed individually in the following sections. The different terms are

- $\Psi_{\rm meas}(r)$: the measured shapes in data, after tagging
- $\Psi^{\rm nonb}_{\rm det}(r)$: the detector level shapes of nonb-jets, without any tagging
- $p_{\rm b}$: the b-jet purities, after tagging
- $b^{\rm b}(r)$: the biases on the b-jets due to the tagging requirement
- $b^{\rm nonb}(r)$: the biases on the nonb-jets due to the tagging requirement
- $C_{\rm had}(r)$: the hadron level corrections to the b-jet shapes

3.2 Raw Shapes

3.2.1 Tagged Jet Shapes

It is necessary for this analysis to measure, in data, the shapes of the tagged jets as well as the shapes of the nonb-jets. The shapes are measured at calorimeter level. The shapes measured using tracks are used in an independent measurement of the hadron level b-quark jet shapes, to get an idea of the systematic effect linked to the response of the calorimeters to low momentum particles.

The tagged jet shapes are defined as the average shapes for all the tagged jets in the samples, measured at calorimeter level. Figure 3.1 shows the integrated jet shapes for tagged jets, as measured in data (black points) and compared to the Pythia Tune A MC predictions for each of the four p_T bins. The data and Pythia Tune A MC jet shapes are not expected to be the same because of the different fractions of b-quark jets in data and MC, as will be shown in section 3.3, and because of the different fraction of b-jets which contain more than one b-quark. The latter issue is discussed in the next paragraphs.

Jets that contain two b-quarks inside the same jet cone are mainly jets which originate from a gluon splitting into a $b\bar{b}$ pair. Such jets are expected to be significantly wider than jets from flavour creation, containing most of the time only a single b-quark inside the cone. The fraction of gluon splitting events has not been measured to a high precision at hadron colliders. The total fraction of inclusive jets at CDF that are gluon jets has been shown, by the inclusive

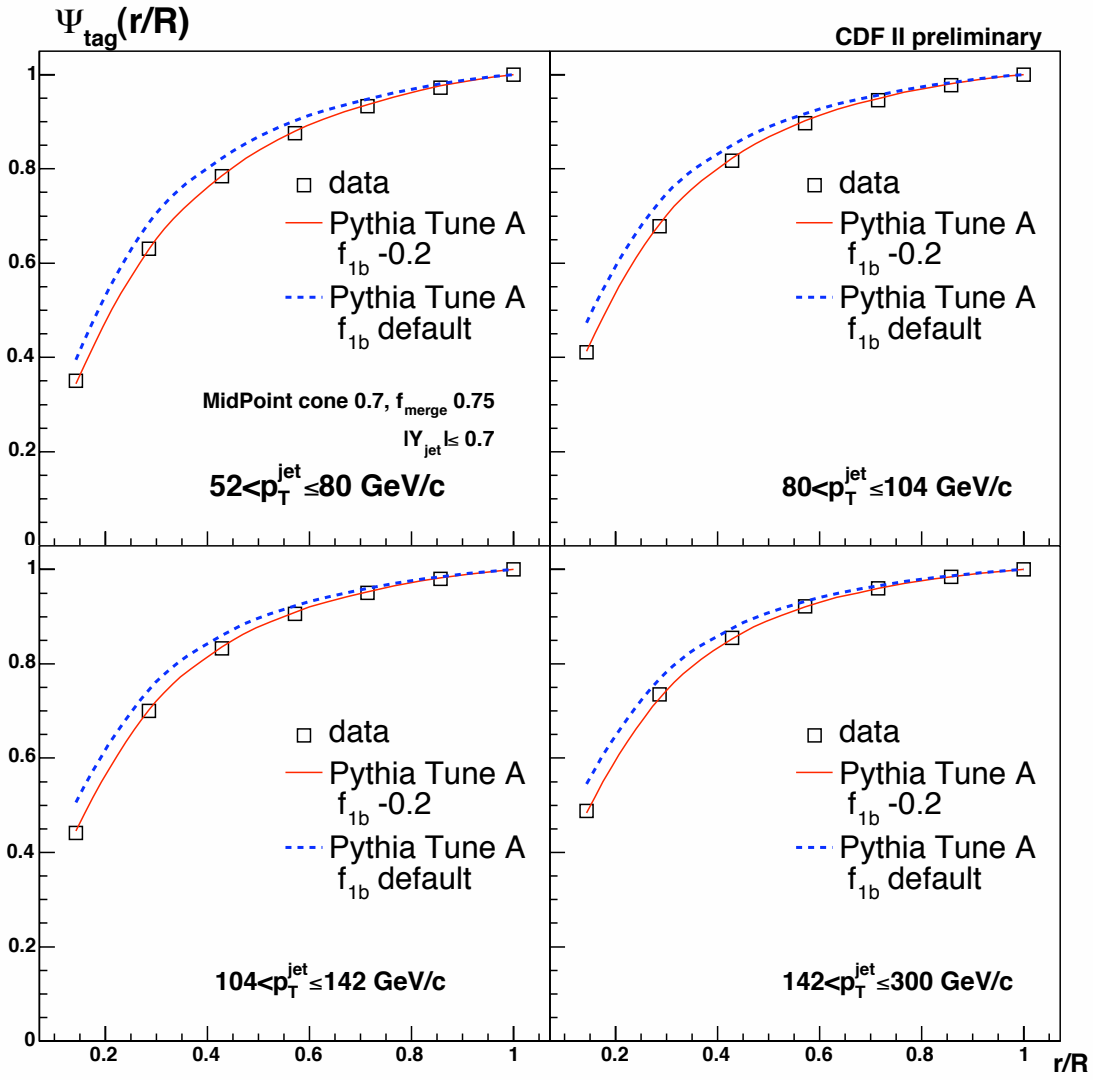

FIGURE 3.1: Raw integrated shapes for tagged jets. The measured shapes in data (black points) are compared to the Pythia Tune A MC predictions using the default f_{1b} fraction (blue dashed line) and using the f_{1b} fraction which best agrees with the data (red curve). Only the statistical errors are shown. These are smaller than the data points in these plots.

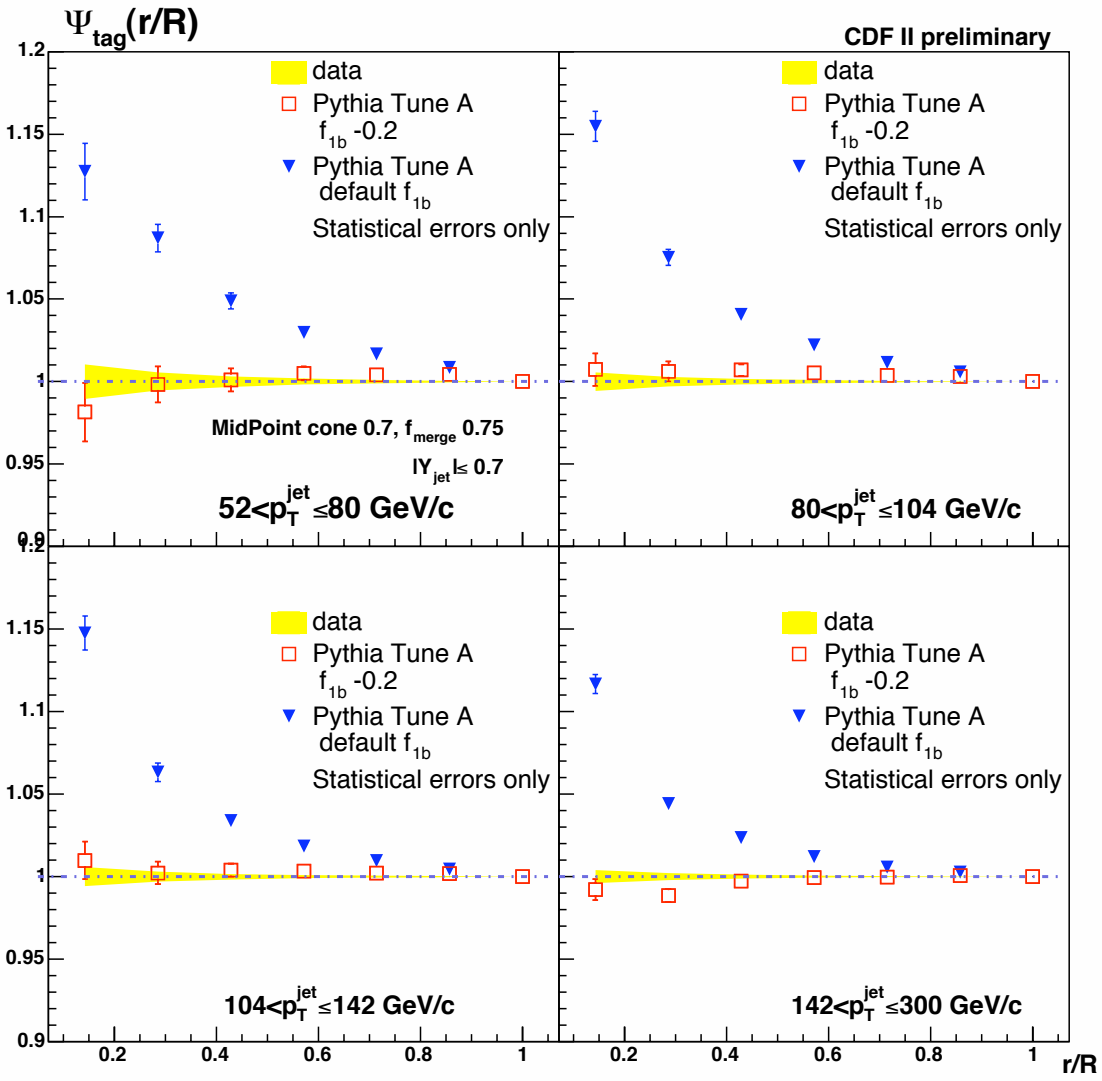

FIGURE 3.2: Ratio of the raw integrated shapes for tagged jets for Pythia Tune A over that measured in data. The statistical errors on the measured shapes in data are shown as yellow bands. The Pythia Tune A MC predictions using the default f_{1b} fraction(blue triangles) and using the f_{1b} fraction which best agrees with the data (red open squares). Only the statistical errors of the Pythia Tune A MC are shown.

jet shape analysis, to agree well with the data [2]. One of the aims of this analysis is to determine how well the fraction of gluon splitting events in b-jet production is modelled by the MC.

Many of the distributions used for the unfolding, such as the secondary vertex mass templates, the bias corrections and the hadron level corrections, 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 have a single b-quark, f_{1b}. This fraction is defined as

$$f_{1b} = \frac{N^{1b}_{jets}}{N^{1b}_{jets} + N^{2b}_{jets}}, \tag{3.9}$$

where N^{1b}_{jets} and N^{2b}_{jets} are the number of b-jets which contain one and two b-quarks, respectively. Figure 3.1 shows that the tagged jet shape is not well modelled by the Pythia Tune A MC. It is important, before calculating any of the unfolding parameters, to vary the f_{1b} fraction in such a way that the agreement between data and Pythia Tune A MC is best. An increase in the fraction of b-jets which contain more than one b-quark inside the jet cone by 20% (absolute) seems to make the agreement much better than the default fraction. This can be seen both in figures 3.1 and 3.2. The former shows as a blue dashed curve the Pythia Tune A predictions for the tagged jet shapes at detector level for the default f_{1b} fraction and the prediction with a f_{1b} fraction 20% higher then predicted are shown as a red curve. The second of these plots shows the ratio between the Pythia Tune A predictions and the measured tagged jet shapes for both the default and the decreased f_{1b} fractions. The agreement is seen to be much better with the lower f_{1b} fraction. This lower f_{1b} fraction is used for the remainder of the analysis to obtain the tagging biases and hadron level corrections to the b-jet shapes.

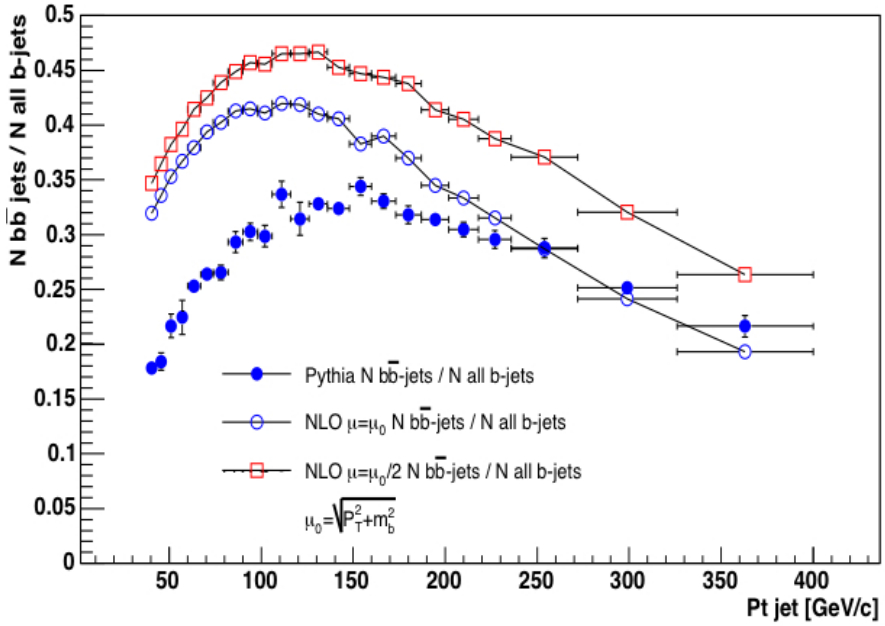

FIGURE 3.3: Fraction of b-jets containing two b-quarks inside the jet cone. Pythia Tune A is compared to NLO calculations using both a scale of $\mu = \mu_0$ and $\mu = \mu_0/2$. This plot is taken from [11].

The decrease of 20% in the f_{1b} fraction can be in part explained by the following. Reference [11] shows a comparison of the fraction of jets containing two b-quarks, which is equal to $1 - f_{1b}$, between Pythia Tune A, where the jet algorithm is run at hadron level, and Next to Leading Order (NLO) calculations. This comparison is reproduced in figure 3.3. The renormalisation and factorisation scale choice of the NLO calculation affects the $b\bar{b}$ fraction. The

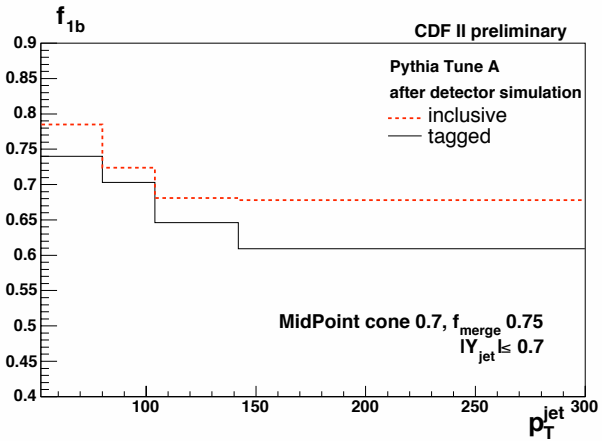

FIGURE 3.4: Fraction of b-jets containing a single b-quark inside the jet cone as predicted by Pythia Tune A MC. The f_{1b} fraction after tagging (black line) is compared to the fraction before tagging (red dashed line).

maximum increase of the double b-quark jet fraction with respect to the Pythia Tune A values is obtained with a scale of $\mu = \mu_0/2$, where $\mu_0 = \sqrt{p_T^2 + m_b^2}$. The maximum increase is of the order of 17% (absolute). The plotted fractions are for inclusive b-jets, before any tagging is applied. It is reasonable to expect that the tagging efficiency for double b-quark jets is higher than that for single b-quark jets. This can be verified by comparing, in Pythia Tune A MC, the f_{1b} fraction before and after tagging, as shown in figure 3.4. The double b-quark jet fraction increases by about 10% after the tagging requirement is applied. There is no reason to believe that the tagging fractions should increase differently for the Pythia Tune A and the NLO predictions but it is reasonable to increase the expected variation by a few percent which gives us the total decrease of the f_{1b} fraction of 20%. There might also be a difference between the true f_{1b} fraction and the one predicted by the NLO calculation but this effect is hard to evaluate.

3.2.2 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 that

$$\Psi_{\text{det}}^{\text{nonb}}(r) \approx \Psi_{\text{det}}^{\text{incl}}(r) \tag{3.10}$$

is used. The difference between these shapes, in Pythia Tune A MC, is negligible as can be seen from figure 3.5 and 3.6. The latter shows the ratio of the nonb over the inclusive shapes, which is compatible with unity with a maximum deviation of less than 0.5%. No systematic uncertainty is therefore related to the use of this approximation. Figure 3.7 shows the integrated shapes for inclusive jets, as measured in data (black points) and compared with the Pythia Tune A MC predictions for each of the four p_T bins. The agreement between data and Pythia Tune A MC is not perfect, especially for the lowest p_T bin and the first bins in r. No reasonable explanation for this difference has been found so far. The effect of the difference in the inclusive jet shapes between data and Pythia Tune A MC on the hadron level b-quark jet shapes will be further investigated in chapter 5. Figure 3.8 shows, as black open squares, the raw inclusive shapes and as red triangles the raw tagged shapes. It is the difference between the inclusive and the tagged jet shapes that is needed in order to be able to expect a reasonable precision in the final measurement.

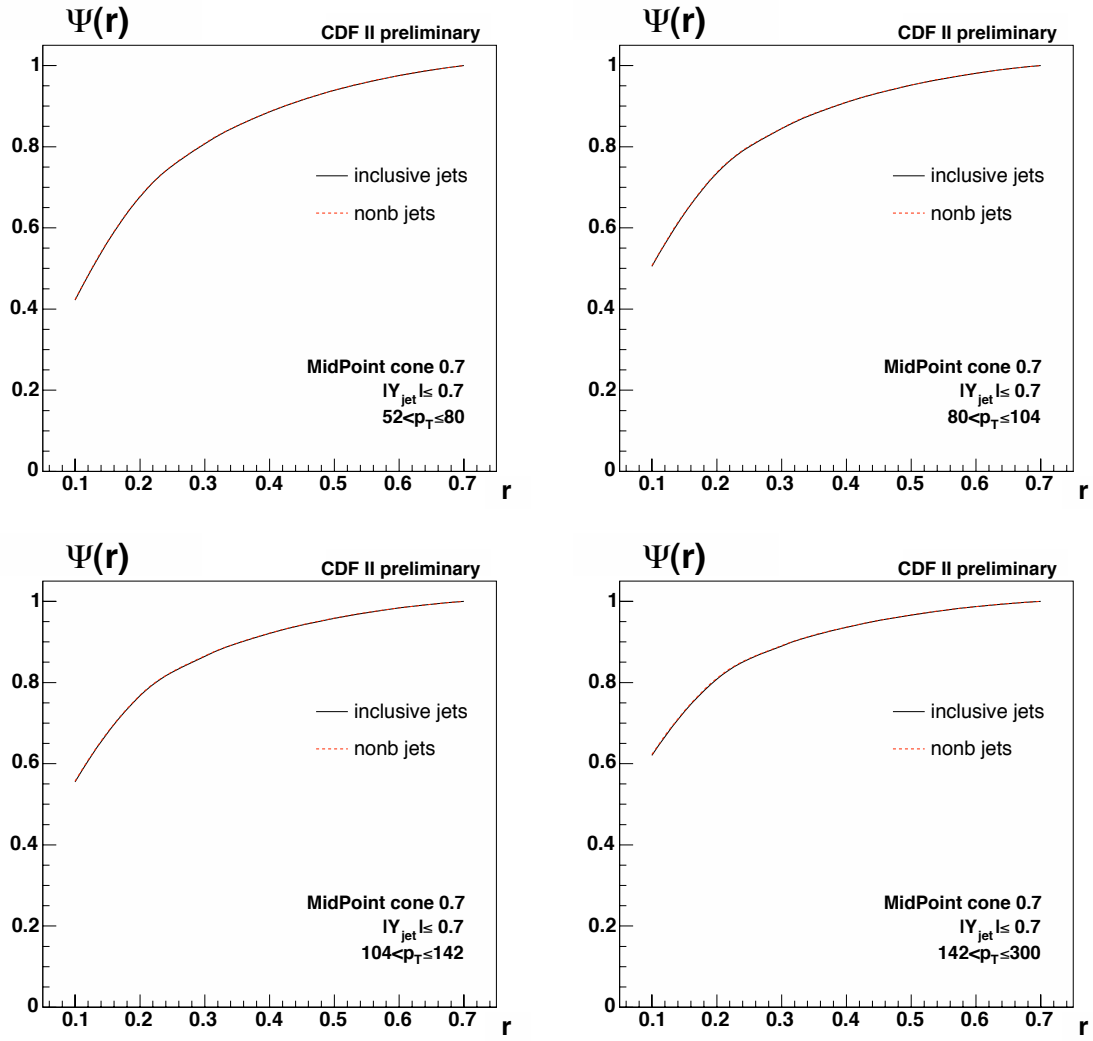

FIGURE 3.5: Raw integrated shapes, in Pythia Tune A, for inclusive (black full line) and nonb-jets (red dashed line).

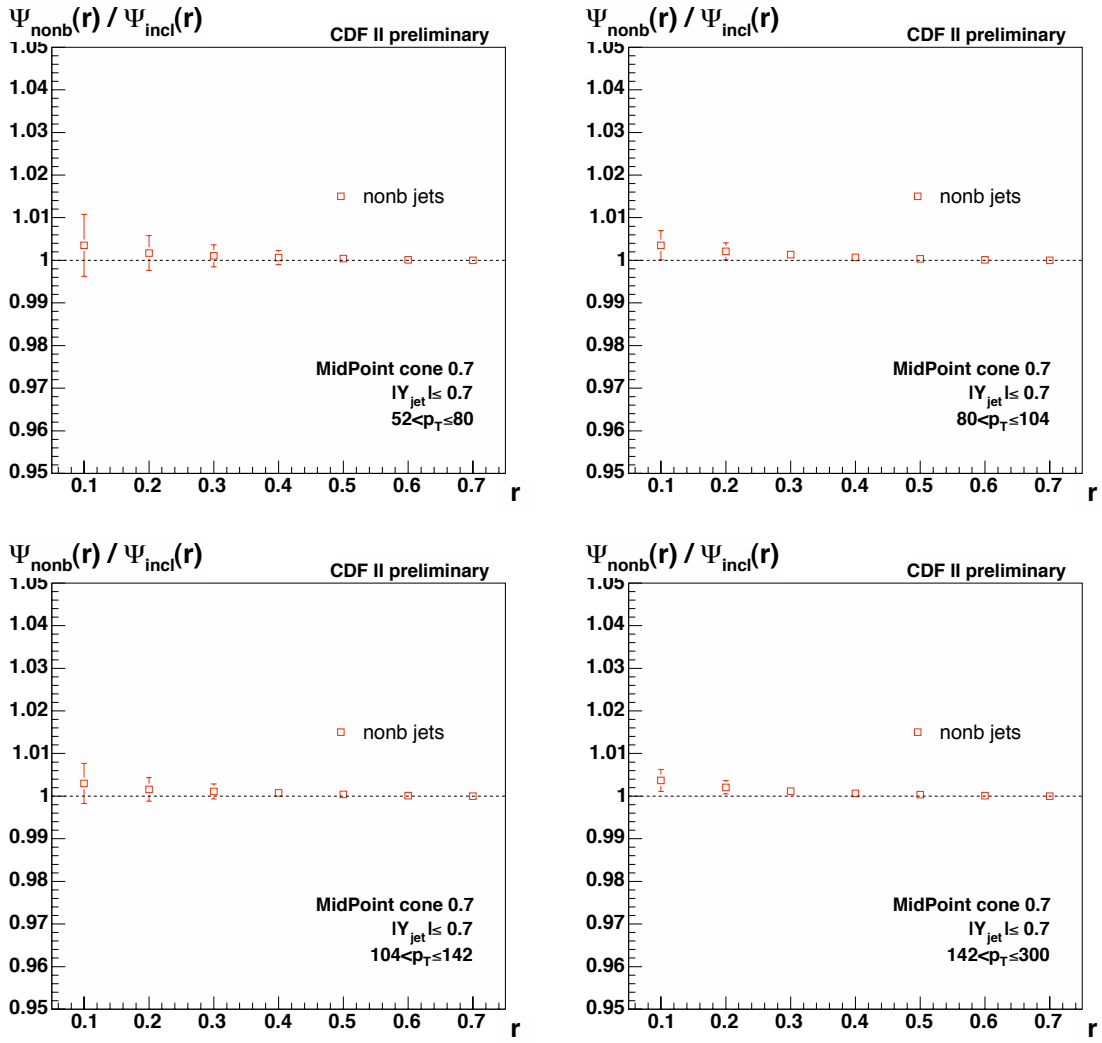

FIGURE 3.6: Ratio of the raw nonb-jet integrated shapes, in Pythia Tune A, over the inclusive integrated shapes (red points). Only the statistical errors are shown.

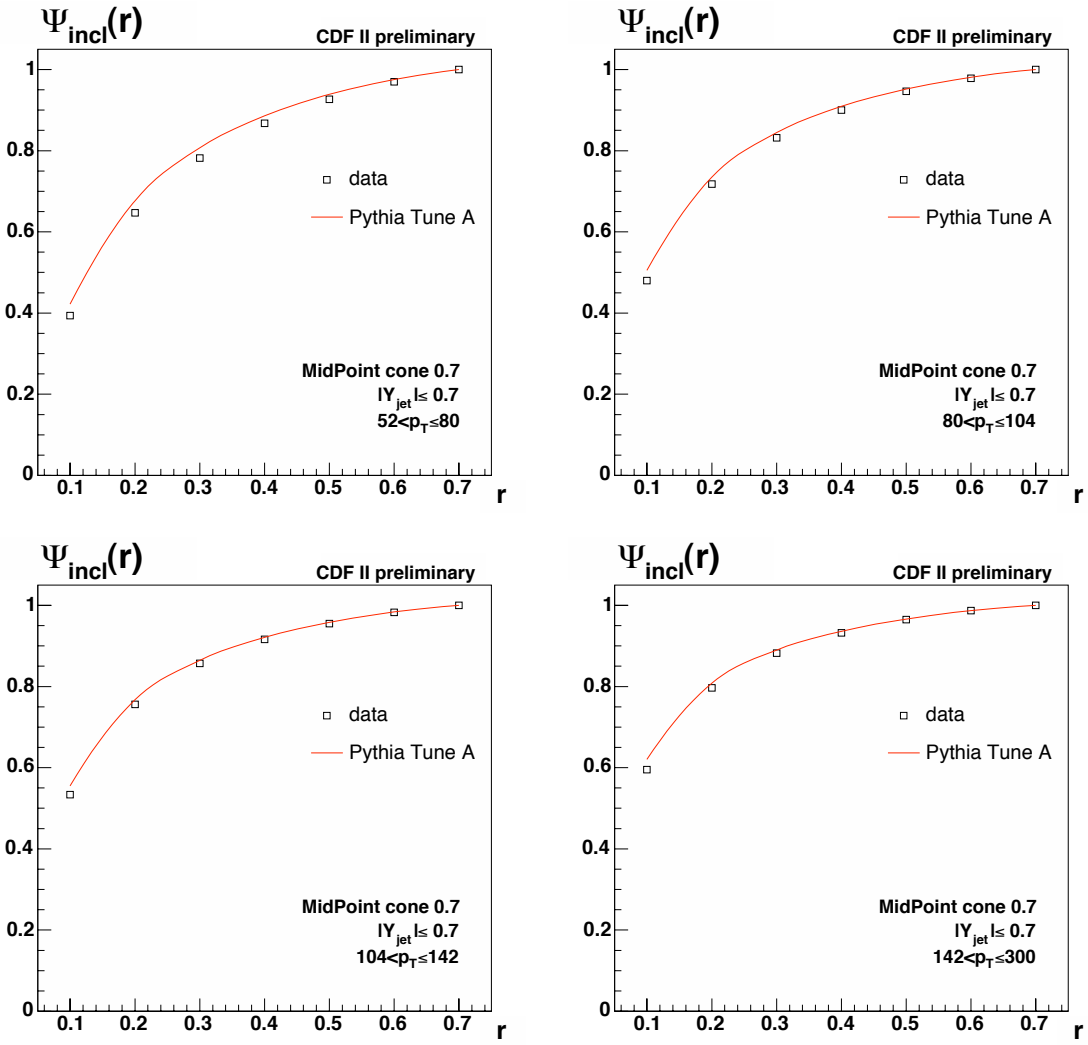

FIGURE 3.7: Raw integrated shapes for inclusive jets. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

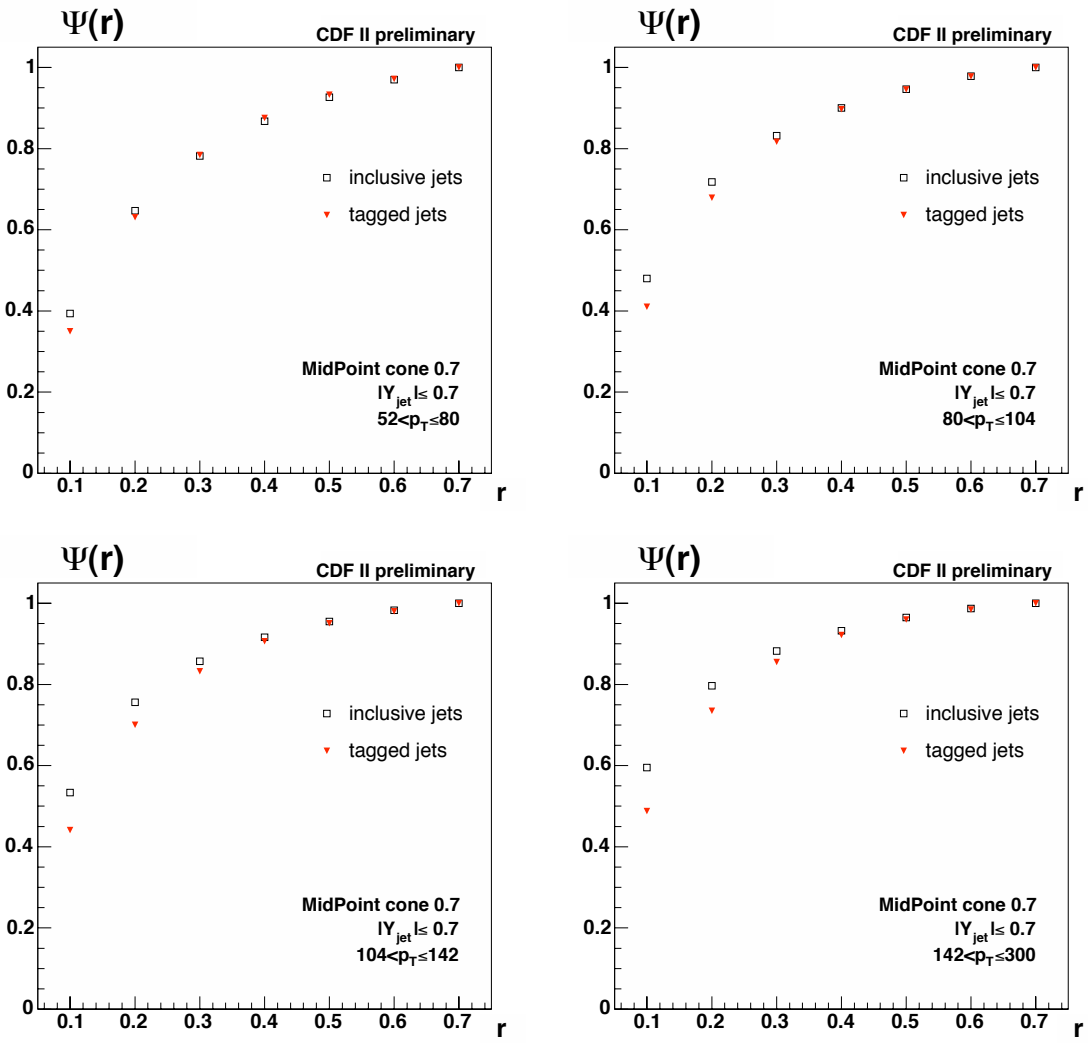

FIGURE 3.8: Raw integrated shapes for inclusive (black open squares) and tagged jets (red full triangles).

3.3 Purity

The fraction of tagged jets which are b-jets, the purity, is extracted from a fit to the secondary vertex mass distributions for b- and nonb-jets [10]. It is not possible to reconstruct the full hadron invariant mass due to the presence of neutral particles which are not detected in the tracking detectors and also due to the detector resolution. 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 and is also different for b-quark jets and c-quark 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. These distributions, known as templates, are renormalised to unit area. The errors on each bin of the templates correspond to the inverse of the square root of the number of entries in that bin. Templates are obtained from the Pythia Tune A MC samples for the b-jets from the b-filtered samples and for the nonb-jets from the inclusive samples. These templates are shown in figure 3.9. The measured distribution in data of the secondary vertex masses is fitted to the b- and nonb-templates, using a binned χ^2 minimisation method [1], to find the most likely fraction of jets that are b-jets. The stability of the fits is tested with respect to changes in the fit range and changes in the number of bins used to cover this fit range. The stability is found to be very good. Figure 3.10 shows the distributions in data along with the fitted distributions for each p_T bin. This shows that the fit describes the data well. Figure 3.11 shows the fit results for p_{b} as a function of the p_T of the jets along with the values in MC. The errors shown in this figure are the statistical errors on the purity (see section 4.1). Table 3.1 shows the results for the purity in both data and Pythia Tune A MC for all the bins.

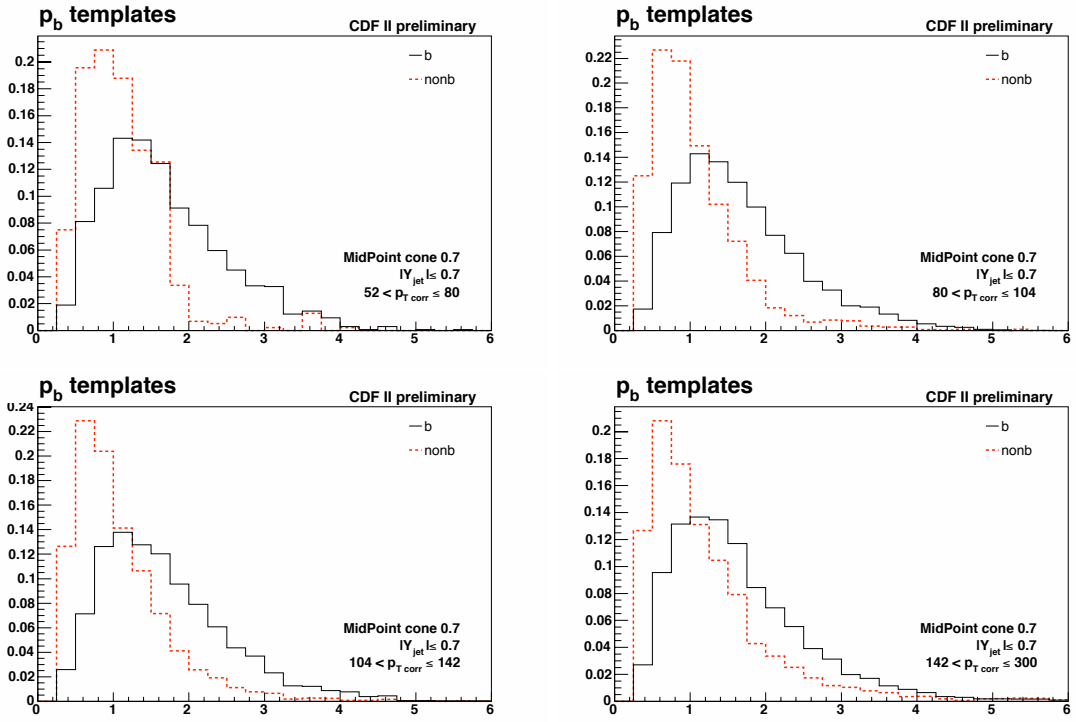

FIGURE 3.9: Normalised secondary vertex mass distributions for b- (black full line) and nonb-jets (red dashed line) for all the p_T bins.

[1] This is done using the ROOT routine TFractionFitter

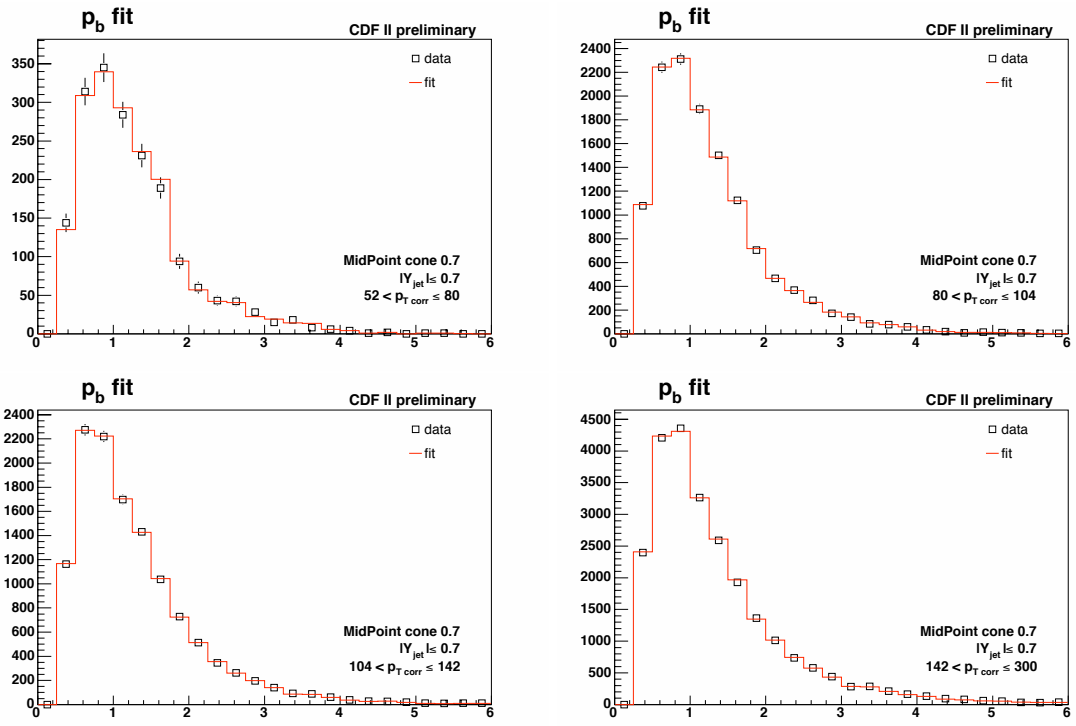

FIGURE 3.10: Secondary vertex mass distribution in data (black points) compared to the fitted distribution (red line) for all the p_T bins.

p_T range	p_{b}	p_{b} MC
52-80	0.320 ± 0.018	0.519
80-104	0.338 ± 0.010	0.448
104-142	0.300 ± 0.010	0.368
142-300	0.242 ± 0.009	0.304

TABLE 3.1: b-jet purity in data, as obtained from the fit results for each p_T bin, shown together with the Pythia Tune A MC values. The errors quoted correspond to the statistical errors only.

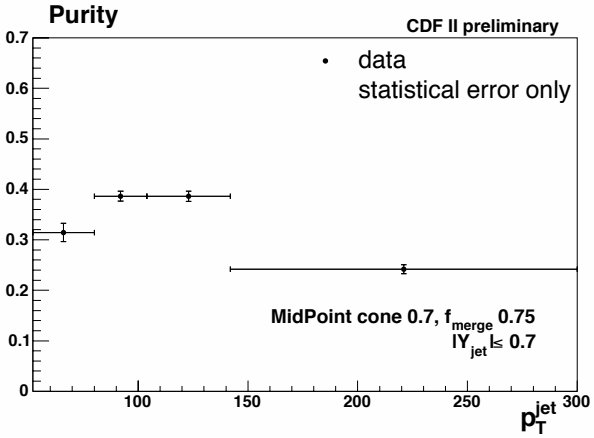

FIGURE 3.11: b-jet purity in data (points) as a function of jet p_T, shown along with the values obtained from MC (dashed red line). The error bars indicate the statistical errors only.

3.4 Biases Due to SecVtx Tagging

The requirement that the jets must be tagged by the SecVtx tight tagger introduces a bias in the measured shapes. These biases must be corrected for. They 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. They are thus defined as

$$b_{\mathrm{b}}(r) = \frac{\Psi^{\mathrm{tag}}_{\mathrm{b\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{b\,MC}}(r)} \tag{3.11}$$

for b-jets and

$$b_{\mathrm{nonb}}(r) = \frac{\Psi^{\mathrm{tag}}_{\mathrm{nonb\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{nonb\,MC}}(r)} \tag{3.12}$$

for nonb-jets. A bias term larger than unity for the low r bins implies that the shapes are narrower after tagging than before. Similarly, a bias term smaller than unity for the low r bins implies that the shapes get wider after the tagging requirement is applied.

Figure 3.12 shows the bias corrections, as a function of r, for b-jets (black open squares) for all the p_T bins considered. These bias corrections are obtained using Pythia Tune A with the $f_{1\mathrm{b}}$ fraction for which the tagged jet shapes best agree with the data. Only the errors due to the MC statistics are shown. Also shown, as red lines, are the biases for single and double b-quark jets (called 1b- and 2b-jets). The maximum bias for b-jets is of the order of 8%. The fact that the biases due to tagging on b-jets are sometimes somewhat smaller than both the biases due to tagging on the single or the double b-quark jet shapes can be understood by the fact that there are two competing effects. The first is that the tagging on b-quark jets leads to narrower jets, the tagging algorithm selects preferentially events where the heavy flavoured quarks are close to the jet axis. The second is that the tagging efficiencies are not the same for single and double b-quark jets. It is expected, from MC studies, that is is more likely to tag a double b-quark jet than a single b-quark jet. Thus the fraction of b-quark jets which have two b-quarks inside the same jet cone is significantly higher after tagging than before tagging. This was shown in figure 3.4 for Pythia Tune A using the default $f_{1\mathrm{b}}$ fraction. Double b-quark jet shapes being broader on average than single b-quark jets, the b-quark jets after tagging are expected to get broader under this effect. These two competing effects: selection of narrower jets by the algorithm but selection of a larger fraction of broader double b-quark jet shapes, tend to cancel each other out to a certain extent, leading to smaller biases.

Figure 3.13 shows the bias corrections, as a function of r, for nonb-jets (black open squares) for all the p_T bins considered. Only the errors due to MC statistics are shown. Also shown are the biases for c- and light+gluon-jets (as red lines). The maximum bias for nonb-jets is of the order of 18%. It is immediately apparent from this plot that the errors on the tagging bias corrections for nonb-jets due to the MC statistics are relatively large. These errors could only be reduced by a considerable increase in the MC statistics for each p_T bin, something which was not possible within the framework of this thesis.

3.5 Hadron Level Corrections

In order to compare the results to any future theoretical models or other experimental results, it is important to measure the jet shapes at hadron level. It is therefore necessary to correct for the detector effects. This is in part done by correcting the p_T of the jets, as described in section 2.4. On top of this, an additional correction is needed to correct the jet shapes.

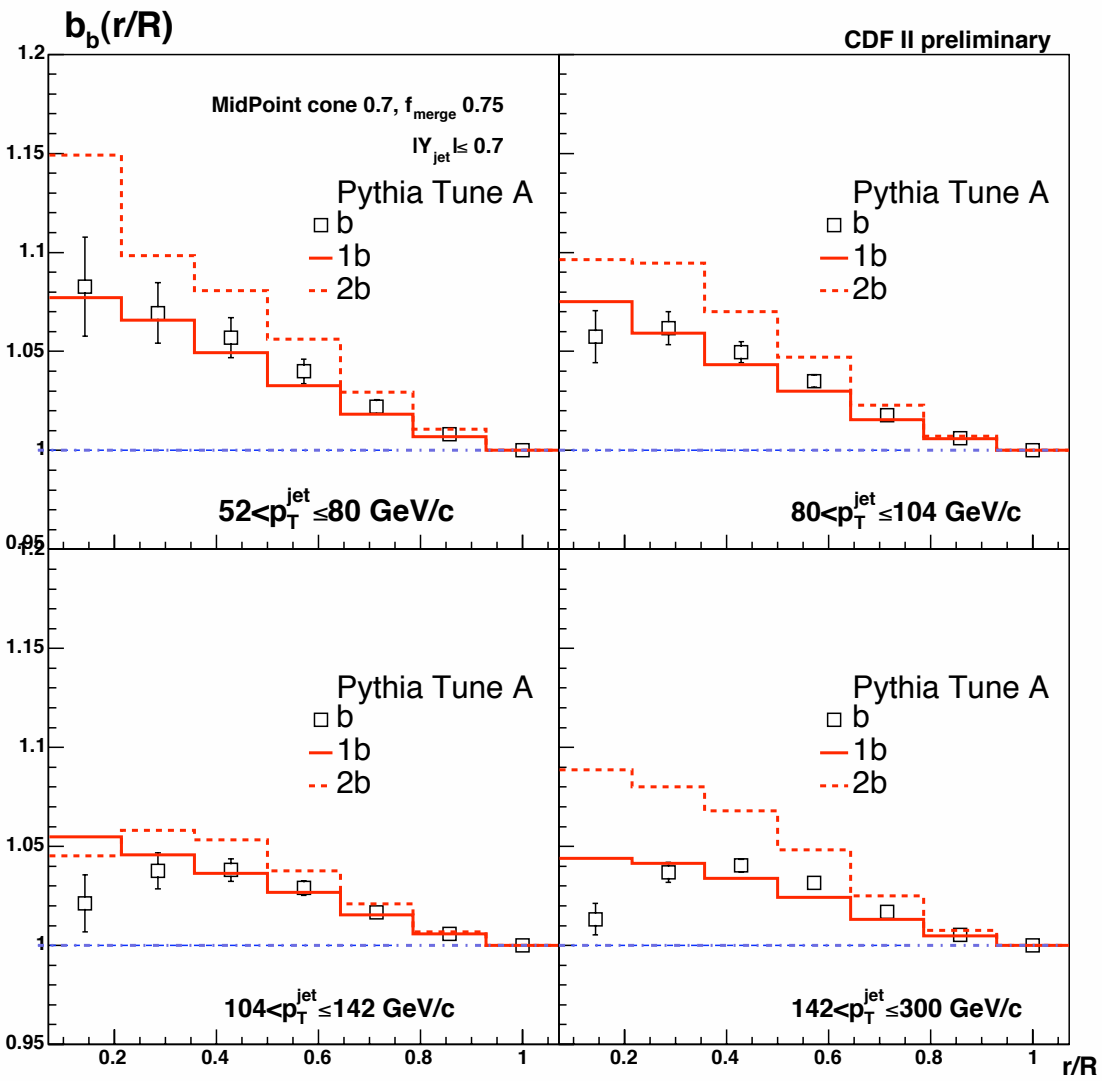

FIGURE 3.12: Bias due to tagging on b-jets (open black squares) shown alongside the bias for 1b- and 2b-jets (full and dashed red lines, respectively). The errors shown are the errors due to the MC statistics.

The correction factors are obtained in a similar way to the one used in reference [2]. Correction factors are computed for each bin in r and each bin in p_T, from MC. The b-quark jet shapes in MC are computed at detector level and at hadron level. The correction factors are defined as

$$C^{\text{had}}(r) = \frac{\Psi^{\text{had}}_{\text{MC}}(r)}{\Psi^{\text{det}}_{\text{MC}}(r)}, \tag{3.13}$$

where $\Psi^{\text{det}}_{\text{MC}}(r)$ are the shapes computed at calorimeter level or using tracks, and $\Psi^{\text{had}}_{\text{MC}}(r)$ are the MC truth shapes computed considering all stable particles, after the fragmentation and hadronisation processes. Figure 3.14 shows the correction factors for each p_T bin obtained using Pythia Tune A with the best $f_{1\text{b}}$ fraction. The correction factors are of the order of 3% at most.

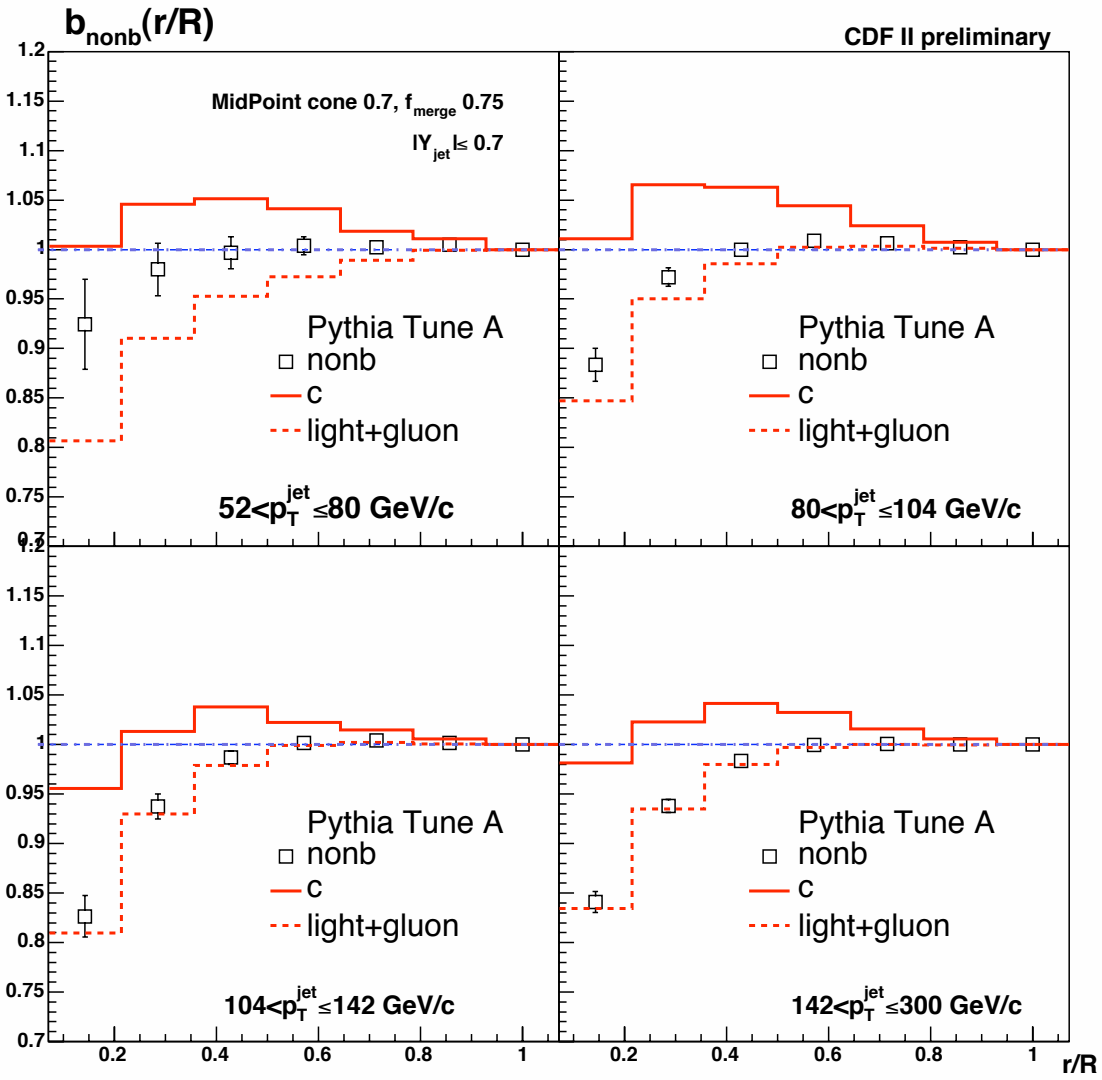

FIGURE 3.13: Bias due to tagging on nonb-jets (black open squares) shown alongside the bias for c- (full red line) and light+gluon-jets (red dashed line). The errors shown are the errors due to the MC statistics.

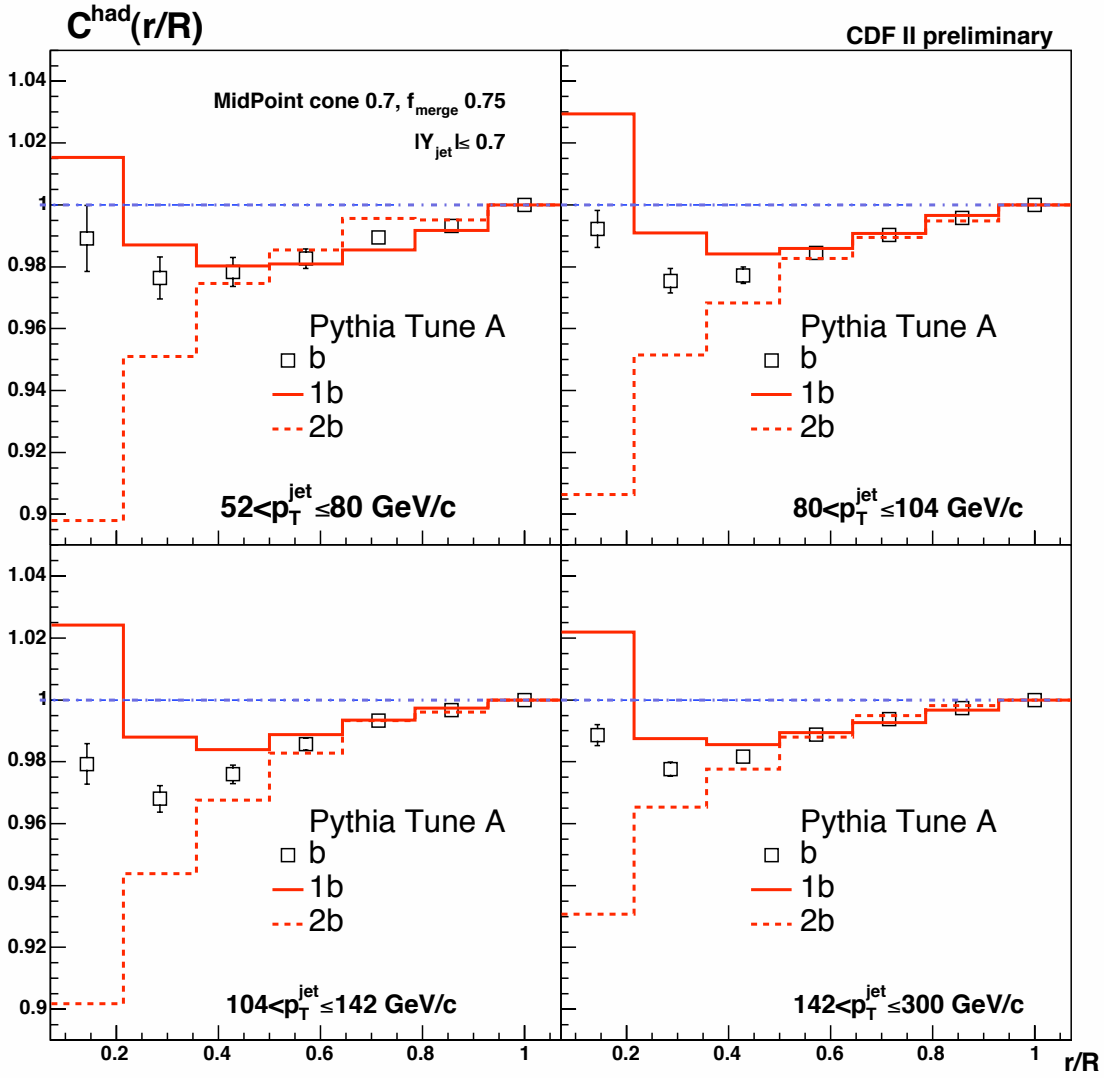

FIGURE 3.14: Hadron level corrections to the b-jet shapes (black points). The errors shown are the MC statistical errors. The hadron level corrections for single b-quark jets (red line) and double b-quark jets (red dashed line) are also shown.

3.6 Hadron Level b-quark Jet Shapes

Figure 3.15 shows the hadron level b-quark jet shapes obtained from the unfolding procedure discussed in the previous sections. Only the statistical errors, which are most of the time smaller than the points, are shown at this point. The final results will be shown in chapter 5. The data are compared to the Pythia Tune A MC hadron level predictions for inclusive and b-quark jet shapes.

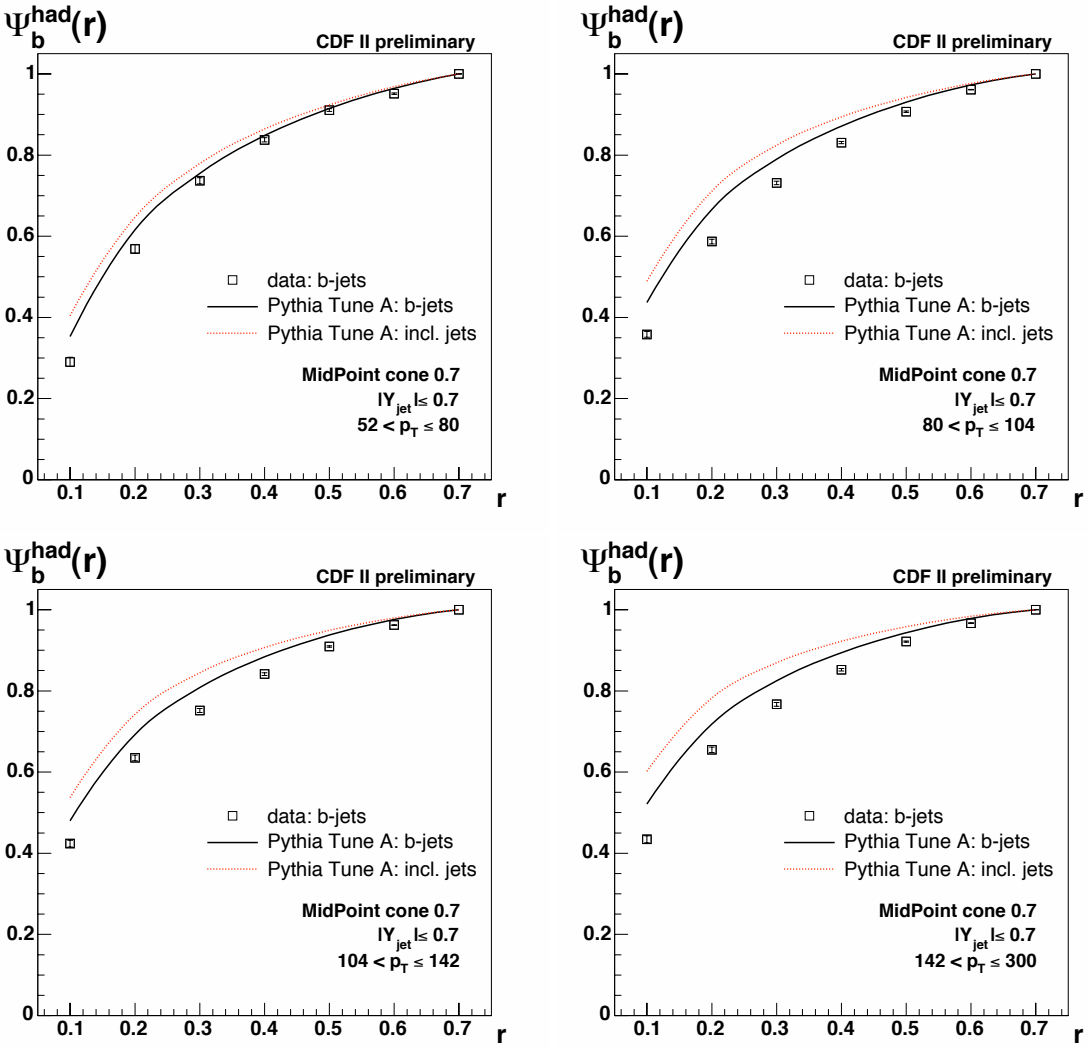

FIGURE 3.15: Hadron level b-quark jet shapes for all p_T bins considered (black open squares). Only the statistical errors are shown. Also shown are the Pythia Tune A MC predictions for inclusive (red line) and b-quark jet shapes (black line).

Chapter 4

Systematic Studies

In this chapter, the systematic studies carried out for this analysis are presented. First the various uncertainties are evaluated in order to determine what variation of each parameter is needed. The hadron level b-quark jet shapes are calculated for each parameter variation and the differences between the default hadron level b-quark jet shapes and those computed using the varied parameters are taken as the systematic uncertainties on the measurements due to that effect. Except where explicitly mentioned, the uncertainties on the different fractions discussed below are quoted as absolute errors, not relative ones.

First, it is necessary to extract the statistical error on the purity from the total error quoted by the fitter. The error quoted by the secondary vertex mass fitter is a combination of the statistical errors of the data and of the MC templates. This method is presented in the next section.

The relative fraction of c-jets in the nonb-jets is not guaranteed to be accurate in the Pythia Tune A MC. It is necessary to find a method to estimate how good the MC description of this fraction is. This is done in section 4.2. This fraction has an effect on the SecVtx mass templates for nonb-jets and on the biases on nonb-jets due to the tagging requirement.

A variation in the relative fraction of c-jets which contain only a single c-quark inside the jet cone needs to be considered. The estimate on the variation of the single b-quark jet content of b-jets is used to set a systematic variation on the single c-quark jet fraction. This is described in section 4.3.

There are a certain number of standard systematic studies which are carried out for most QCD analyses [2] [10]. These are the variations of the jet energy scale, the primary vertex location and the missing E_T significance.

The systematic uncertainty related to the differences between the hadron level b-quark jet shapes calculated using tracks instead of calorimeter towers is presented in section 4.8. This systematic study is carried out to verify the detector response to low momentum particles. Section 4.9 looks into the differences in the hadron level b-quark jet shapes obtained from shapes measured at calorimeter level but using only towers with $p_T > 0.5\,\text{GeV}$ and those obtained using the default cut at 0.1 GeV.

As mentioned in section 2.2, different samples were generated in order to investigate the effect on the unfolding method of using a specific MC sample with a specific parameter set. These studies are presented in section 4.10.

Section 4.11 presents a study to investigate the dependence on the MC modelling of the

SecVtx performance. Some distributions, related to the SecVtx algorithm, such as the number of Pass 1 tracks, are not well described by the Pythia Tune A MC. The distribution of the number of Pass 1 tracks in Pythia Tune A MC is re-weighted to agree with the data. The different biases and hadron level corrections used for the unfolding of the b-quark jet shapes are re-computed using the re-weighted Pythia Tune A MC and are found to be within the errors due to the MC statistics from the default unfolding paramters.

The final systematic uncertainties are shown in section 4.12 for each bin in r and each bin in p_T for each of the effects studied in this chapter. The dominant sources of systematic uncertainties are also discussed.

The final section of this chapter presents a study, which is not included as a source of systematic error, on the effect of changing the jet rapidity cut from $|Y| \leq 0.7$ to $0.1 \leq |Y| \leq 0.7$. The purpose of this study is to check if the previously published inclusive results [2], computed using the latter rapidity cuts, can be directly compared to the results of this analysis. It turns out that removing the central rapidity region does not affect the b-jet shapes or the inclusive shapes.

4.1 Extraction of the Statistical Error on the Purity

Before being able to investigate systematic uncertainties on the purity, it is important to extract the contribution of the statistical error to the total error on the purity quoted by the fitting algorithm. The fitting algorithm returns a total error which takes into account both the statistical errors of the templates and of the data. In order to extract the errors due to the statistics of the data, the statistics of the templates are artificially increased until the contribution to the total error of the fit is negligible. The error contribution due to the statistics of the templates goes roughly as $1/\sqrt{N}$; by increasing artificially the statistics of the Monte Carlo templates, one is effectively reducing the error due to the templates to zero. This method is similar to that carried out in [10]. In practise this is done by scaling the template histograms for b- and nonb-jets by a factor of 10^3, which reduces the effective error of the templates by a factor of $\sqrt{10^3}$. This factor is motivated by figure 4.1, taken from [10], which shows for one particular bin in p_T the total error of the fit as a function of the factor by which the templates are scaled. The blue line shows the asymptotic error. This is taken as the statistical error. This shows that a scaling of 10^3 is more than sufficient to consider that the error quoted by the fitter is the statistical error.

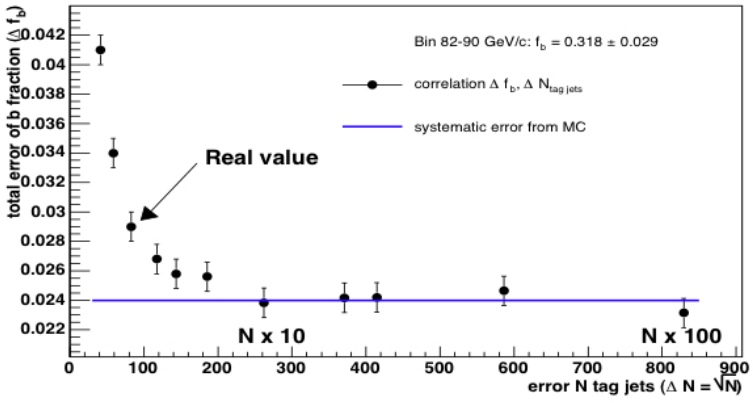

FIGURE 4.1: Error on the purity of the sample returned by the fitting algorithm as a function of the artificial scaling applied to the MC templates.

The error on the fit when using the scaled histograms is taken as the statistical error, $\Delta p_{\mathrm{b}}^{\mathrm{stat}}$.

The error due to the MC statistics is thus given by

$$\Delta p_{\mathrm{b}}^{\mathrm{MC\,stat}} = \sqrt{(\Delta p_{\mathrm{b}}^{\mathrm{fit}})^2 - (\Delta p_{\mathrm{b}}^{\mathrm{stat}})^2} \tag{4.1}$$

where $p_{\mathrm{b}}^{\mathrm{fit}}$ is the total error quoted by the fitting algorithm.

4.2 Estimate of the c-fraction in Data

An important source of systematic uncertainty is the fraction of nonb-jets which are c-jets. This will have an effect on the secondary vertex mass templates for nonb-jets which are used for the measurement of the purity of the sample. It will also have an effect on the biases due to tagging on the nonb-jets. Previous studies quote a conservative error of $20 - 30\%$ (relative error) [10]. It is important to see if this error can be reduced. This was done by carrying out a secondary vertex mass fit using three different templates instead of the standard two, as was presented in section 3.3. Different templates are obtained for b-jets, c-jets and light+gluon-jets.

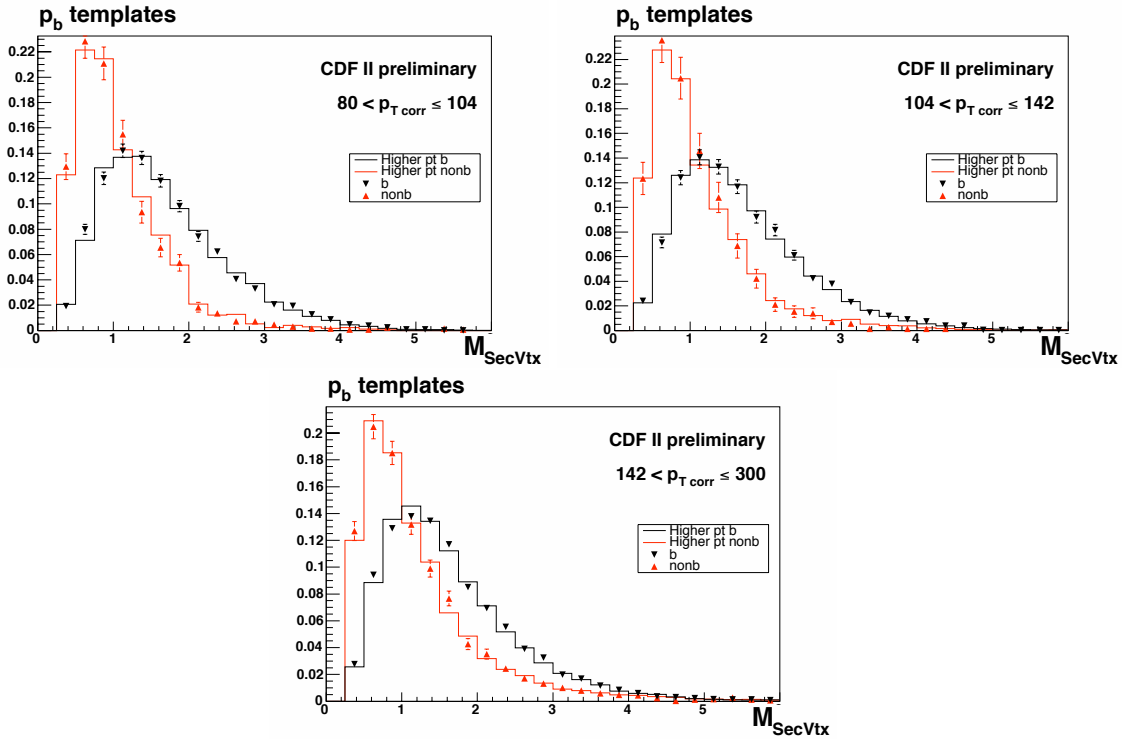

FIGURE 4.2: Secondary vertex mass distributions for b- (black) and nonb-jets (red) for all the p_T bins. The standard MC samples, shown as points with the MC statistical errors, are compatible with the higher p_T samples (lines).

The limitation of this method is that the statistics for the c- and light+gluon-jet templates is too low for the fits to converge. In order to overcome this statistical limitation, it is possible to use higher p_T MC samples without biasing the outcome of the fit. The higher p_T samples have a larger number of events in the p_T bins of interest because the cut on the hard-scattered p_T is higher. The assumption that the use of the higher p_T MC samples does not bias the outcome of the secondary vertex mass fit must be checked before results for the c-fraction can be obtained. As shown in section 2.2, the MC samples used for the four p_T bins of this analysis are: Pt18, Pt40, Pt60 and Pt90. In this section these are referred to as the default samples. The higher

p_T MC samples would thus be: Pt40, Pt60, Pt90 and Pt120, respectively. The comparisons in this section are not applicable to the lowest p_T bin (which uses Pt18 by default) because the p_T distribution of the Pt40 sample is not correct, due to turn on effects, up to about 60 GeV. The p_T distributions for Pt18 and Pt40 are therefore not the same over the whole range of the bin. For the other three p_T bins, a set of comparisons are made which show that the use of the higher p_T samples do not bias the fit results. The three tests carried out are the following

- The b- and nonb-jet templates for the default and higher p_T samples are within errors of each other as shown in figure 4.2.
- The results of the fit of the data to the standard b- and nonb-jet templates using the default and the higher p_T samples are within 4% of each other, well within the total error on the fit. The results of the fit are shown for both cases in table 4.1. The difference between the two fitted fractions is also shown in this table.
- The fit of the data to the b-, c- and light+gluon-templates using the higher p_T samples give results for the b-jet fraction that are within 2% of the results using the default, two-template fit. The results are shown in table 4.2.

These tests all show that the use of the higher p_T MC samples for the templates does not bias in any way the results of the fits. The values for the c-purity can therefore be trusted.

p_T range	$p_b \pm \Delta p_b$	$p_b \pm \Delta p_b$ higher p_T templates	Δp_b
80-104	0.339±0.023	0.299±0.016	0.040
104-142	0.300±0.025	0.275±0.015	0.025
142-300	0.242±0.026	0.242±0.015	0.000

TABLE 4.1: Results for the b-jet purity, p_b, using the standard MC templates and the ones obtained using the higher p_T templates. The errors shown are the errors returned by the fitter which include the statistical errors and the errors due to the MC statistics. Also shown is the difference between these two values for each bin.

p_T range	$p_b \pm \Delta p_b$	$p_c \pm \Delta p_c$	$f_c \pm \Delta f_c$	f_c in MC	Δf_c	Δp_b
80-104	0.327±0.022	0.307±0.047	0.456±0.071	0.418	0.038	0.012
104-142	0.284±0.018	0.245±0.039	0.342±0.055	0.309	0.033	0.016
142-300	0.240±0.023	0.190±0.044	0.250±0.060	0.254	0.004	0.002

TABLE 4.2: Results for p_b, p_c and f_c in data from the fits using the higher p_T templates. The f_c fraction in MC is also shown along with the difference between the fit and the MC values. The differences between the p_b values obtained with the standard fit and the 3-parameter fit are shown in the last column. The errors shown are the errors returned by the fitter which include the statistical error and the error due to the MC statistics.

The templates for b-, c- and light+gluon-jets using the higher p_T MC samples are shown in figure 4.3 for the three highest p_T bins. The distribution in data, along with the fit to the three templates is shown in figure 4.4. This shows that the fit is very good. The c-content of the nonb-jets, f_c, is defined as

$$f_c = \frac{N^c_{\text{jets}}}{N^{\text{nonb}}_{\text{jets}}} = \frac{N^c_{\text{jets}}}{N^{\text{tot}}_{\text{jets}} - N^b_{\text{jets}}} = \frac{p_c}{1 - p_b}, \tag{4.2}$$

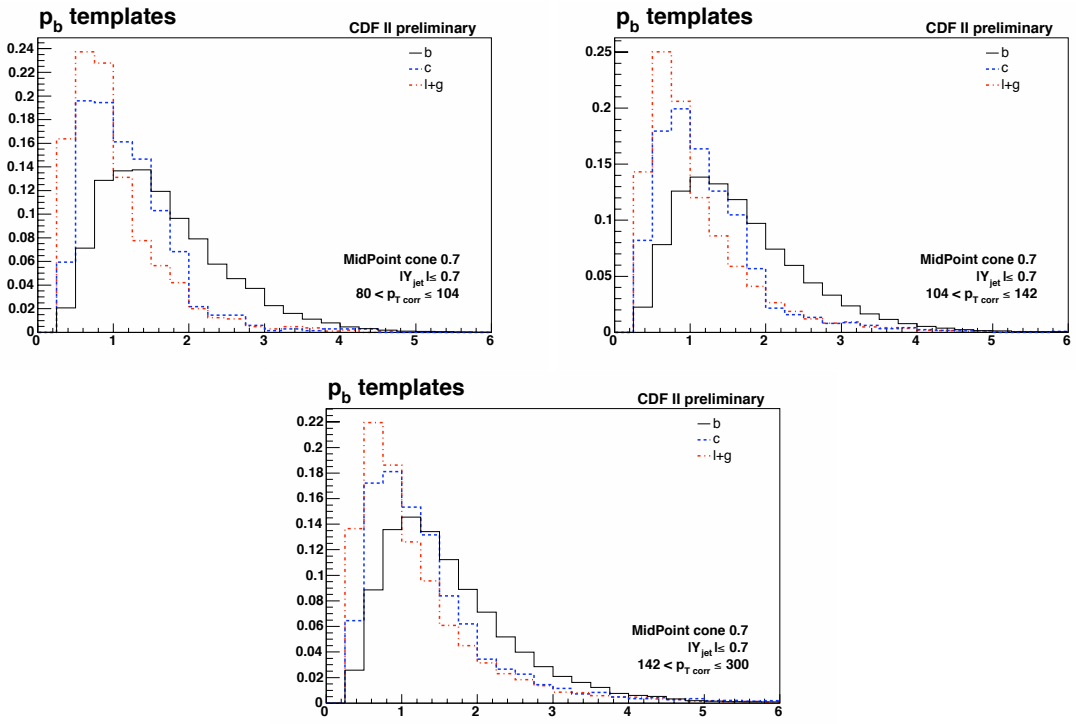

FIGURE 4.3: Secondary vertex mass distributions for b- (black solid line), c- (blue dashed line) and light+gluon-jets (red dotted line) for all the p_T bins.

where p_b and p_c are the fitted b- and c-fractions and $N^\mathrm{x}_\mathrm{jets}$ are the number of jets of flavour x in the sample. Figure 4.5 shows the results for f_c as a function of p_T for the three highest p_T bins. This shows that that the c-content of the nonb-jet fraction is well described in MC. The fit gives results, shown in table 4.2, that are within 4% of the MC values. For this analysis, the MC truth values for f_c are used because of the fact that it is not possible to extract from data the c-content for the lowest p_T bin. In order to account for the 4% difference in the fitted fractions, along with the uncertainty on the fit, about 3%, a systematic error of 5% is considered for the fraction of nonb-jets which are c-jets.

4.3 Estimate of the Uncertainty on $f_{1\mathrm{c}}$

In section 3.2.1 it was found that the fraction of b-jets which have more than one b-quark inside the jet cone was not well modelled by Pythia Tune A. In order to get the best agreement in the tagged jet shapes between the data and the MC predictions, it was necessary to increase this fraction by 20%. A similar effect could occur in c-quark jets where the fraction of c-jets which contain more than one c-quark inside the jet cone could be different in the data than in the Pythia Tune A MC prediction. This fraction has an effect on the nonb-jet secondary vertex templates used for the purity measurement as well as on the tagging biases on nonb-jets. The fraction of c-jets containing only one c-quark, $f_{1\mathrm{c}}$, is varied by $\pm 20\%$, the same as the increase to the $f_{1\mathrm{b}}$ fraction applied for the unfolding, in order to determine the effect on the final hadron level b-quark jet shapes.

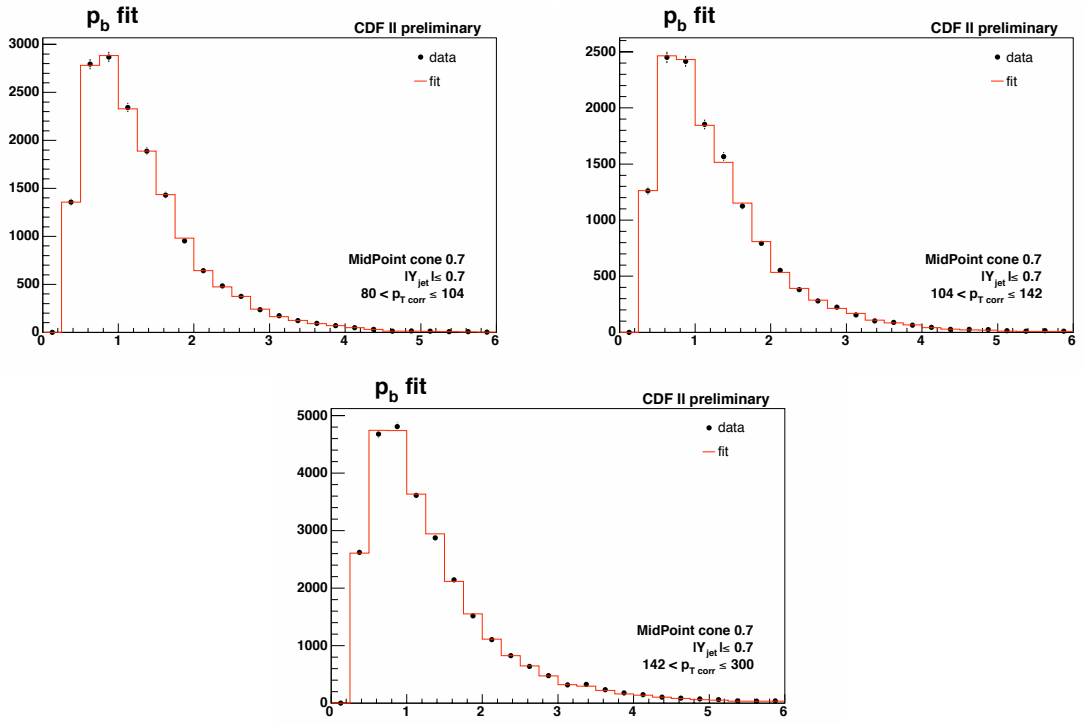

FIGURE 4.4: Secondary vertex mass distribution in data (black points), showing the statistical errors, compared to the fitted distribution using the three templates (red line) for the three highest p_T bins.

4.4 Jet Energy Scale

A systematic error must be associated to the uncertainty of the jet corrections as well as the uncertainty on the detector level jets. This systematic error is a measure of how well the jet energy corrections, presented in section 2.4, correct the measured jet p_T back to the hadron level one. The systematic uncertainty on the jet energy scale quoted for the official jet energy corrections is 3% (relative error) [13]. Another study specifically calculated the jet corrections for b-jets and found results that were in good agreement with the inclusive jet corrections [10]. The comparison between the official jet corrections and the b-jet specific corrections is shown in figure 4.6 and 4.7 for inclusive and tagged jets, respectively. Also shown in this figure are the corrected p_T distributions using the same correction factors as the inclusive jet shape analysis [2]. This shows that the official jet correction are very similar to the b-jet specific corrections and to the inclusive jet shape analysis corrections. The systematic error on the jet energy scale should combine both the error quoted for the official jet corrections and the uncertainty on the b-jet fragmentation which is 0.6%. Adding these two effects in quadrature leads to a total systematic uncertainty on the jet energy scale of 3% (relative error). For this analysis, the effect on the final hadron level b-quark jet shapes is evaluated by shifting the p_T of the bins by $\pm 3\%$ and re-computing the shapes.

4.5 Missing E_T Significance

The systematic effect due to the choice of the missing E_T significance cut is taken to be the difference in the hadron level b-quark jets between the default shapes and those computed using the standard CDF variation of $\pm 15\%$ (relative) [10] [2]. The cut on the missing E_T significance removes a large fraction of the cosmic events but the shapes are not expected to be sensitive to

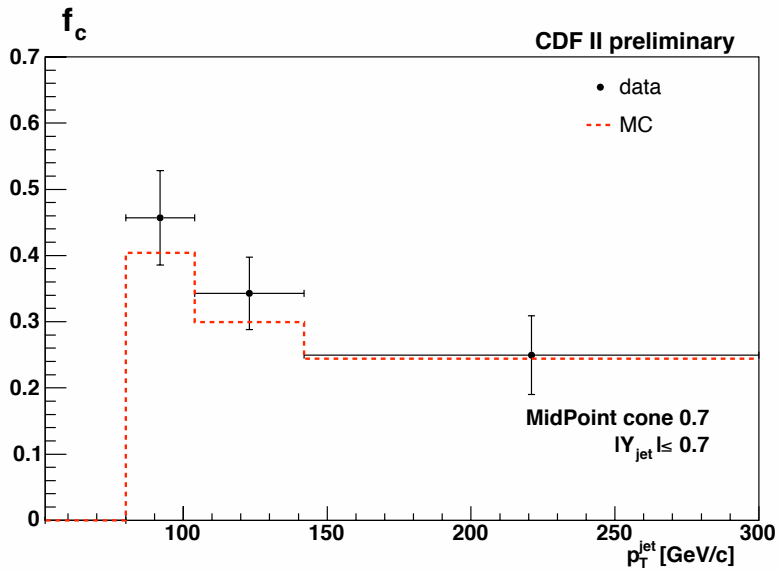

FIGURE 4.5: Fraction of nonb-jets which are c-jets as a function of jet p_T (black points) shown along with the values obtained from MC (red dashed line) which are used in this analysis. The errors shown are the errors returned by the fitter which include the statistical errors and the errors due to the MC statistics.

the values of the cuts because the number of events which are removed by this cut is very small. This systematic effect is thus expected to be small for all the parameters.

4.6 Primary Vertex z-position

The cut on the location of the primary vertex is varied by ± 5 cm around the event selection cut at 50 cm and the effect on the hadron level b-quark jet shapes is investigated. Again this effect is expected to be small because the corrections to the jets and secondary vertices due to the location of the primary vertex are well understood in this region.

4.7 Inclusive vs. nonb-jet Shapes

As was shown in section 3.2, the use of the inclusive jet shapes instead of the nonb-jet shapes in the unfolding equation does not introduce any significant systematic error. The differences between the raw inclusive and nonb-jet shapes, as obtained from the Pythia Tune A MC, are found to be smaller than the statistical errors.

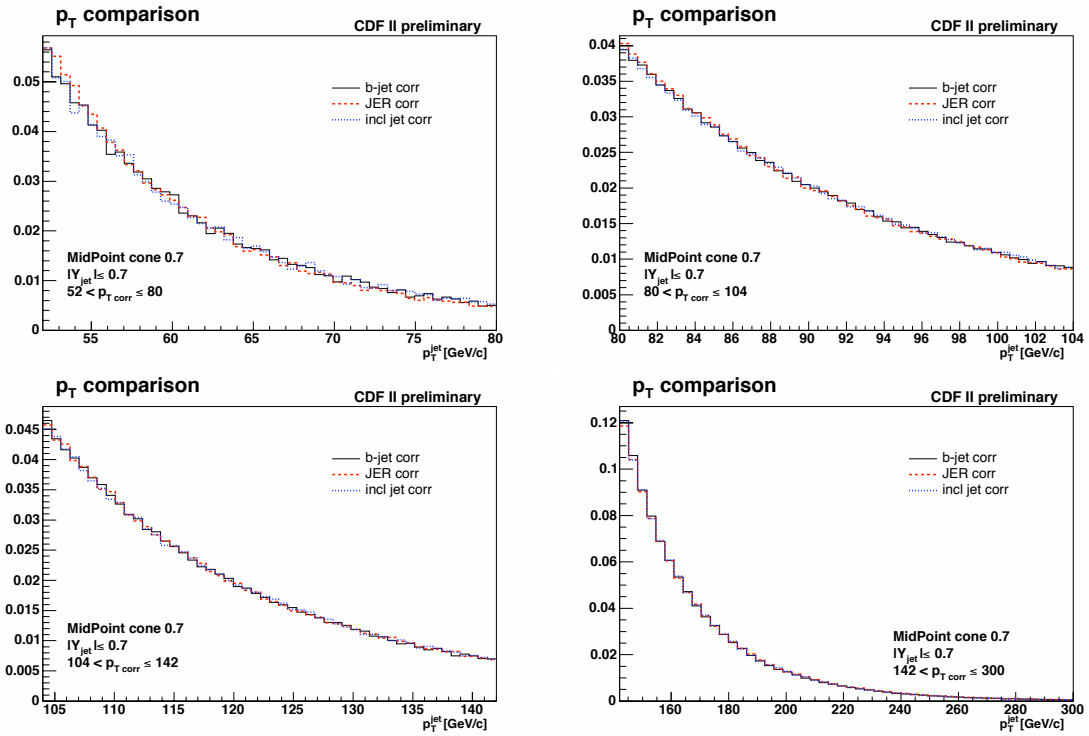

FIGURE 4.6: Comparison in data of the corrected p_T distributions of all jets for each dataset obtained with the official jet corrections (red dashed line), the b-jet specific corrections (black line) and the corrections used in the inclusive jet shape analysis (blue dotted line). The distributions have been normalised to unit area.

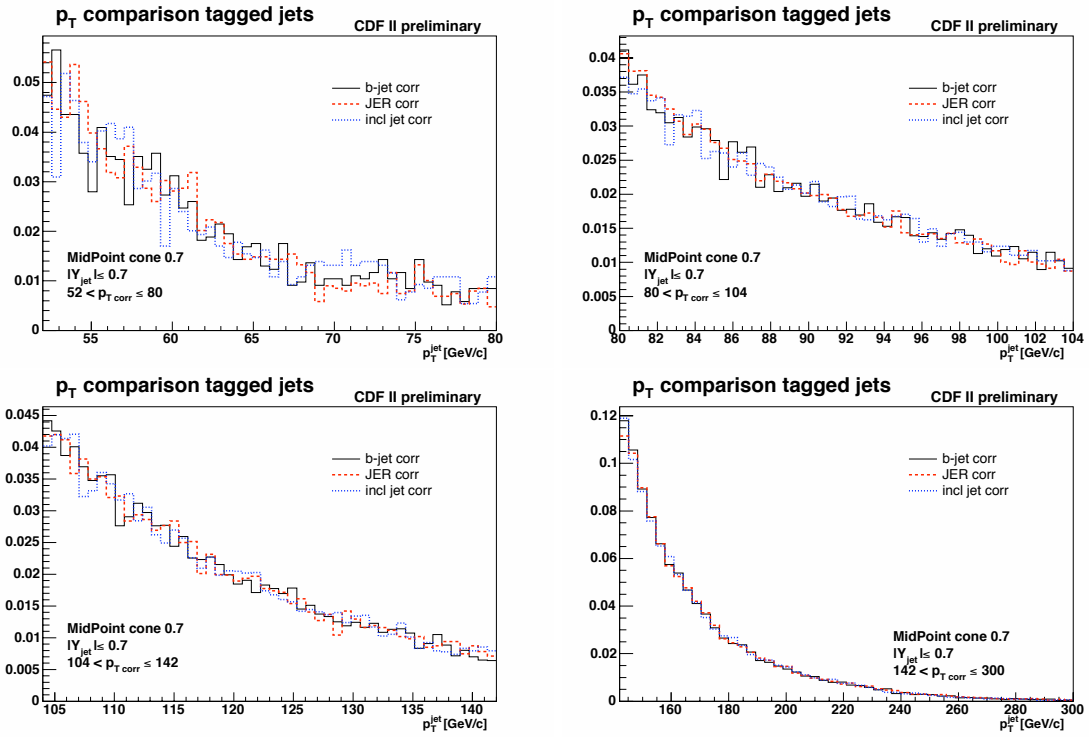

FIGURE 4.7: Comparison in data of the corrected p_T distributions of tagged jets for each dataset obtained with the official jet corrections (black line), the b-jet specific corrections (red dashed line) and the corrections used in the inclusive jet shape analysis (blue dotted line).

4.8 Hadron Level b-jet Shapes Obtained Using Tracks

The hadron level b-jet shapes obtained starting from the raw calorimetric shapes are compared to those obtained starting with tracks. The jet algorithm is still run using calorimeter towers. To be included inside a jet, the tracks must pass certain cuts

- $\Delta R_{\text{track}\to\text{jet}} \leq 0.7$, to ensure the tracks are inside the jet cone
- $0.5 < p_T^{\text{track}} < 100\,\text{GeV}$, to ensure the tracks are well reconstructed
- track $|Z_0| < 2\,\text{cm}$ with respect to the primary vertex, to ensure the tracks come from the same primary vertex as the jet

These cuts are the same as used for the inclusive jet shape analysis [2].

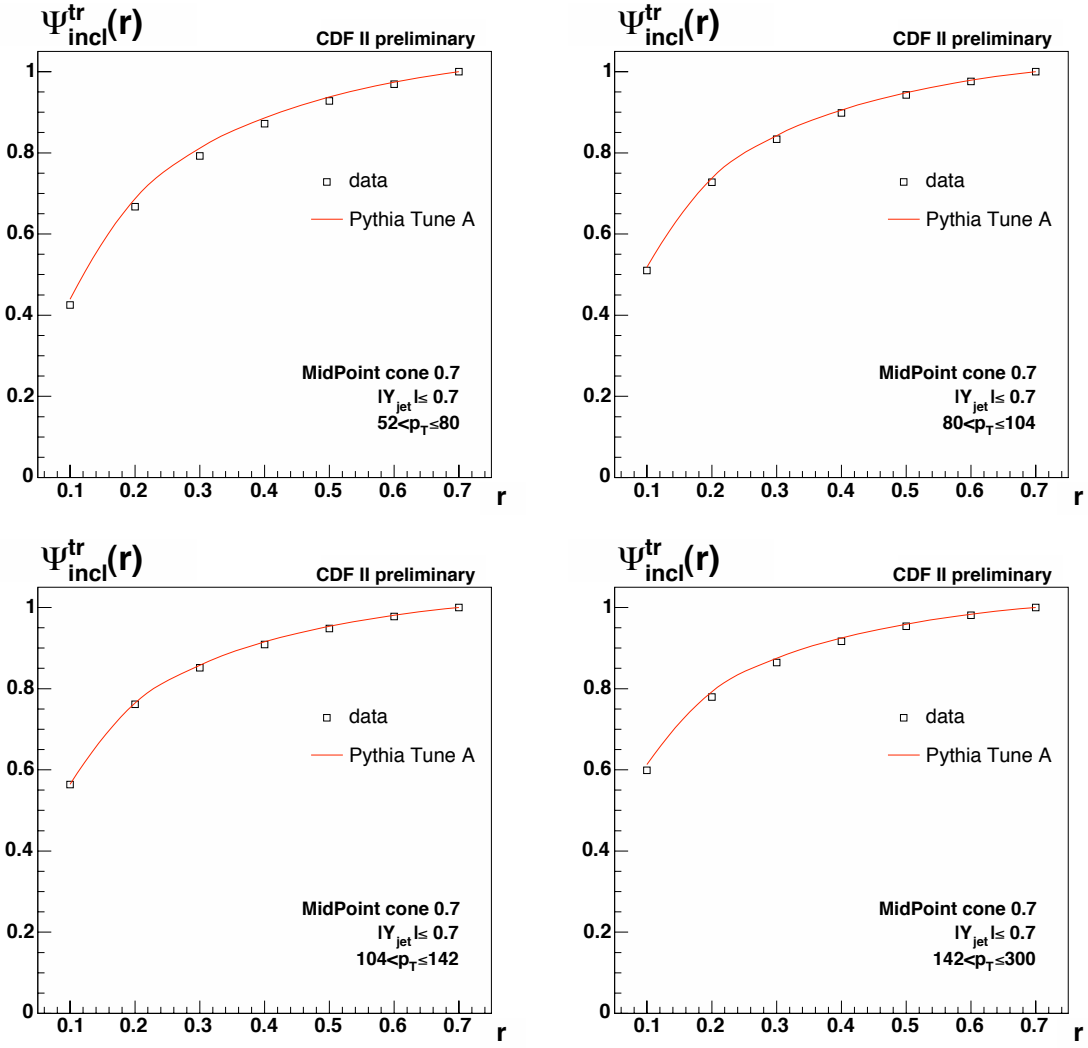

FIGURE 4.8: Raw integrated shapes for inclusive jets measured using tracks. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

Figure 4.8 shows a comparison between data and the Pythia Tune A MC of the raw inclusive shapes measured using tracks. Figure 4.9 shows the same comparison for tagged jets. As for the case of calorimetric jet shapes, the agreement between data and MC is not perfect. The data and MC shapes for the tagged jets are expected to be slightly different due to the different

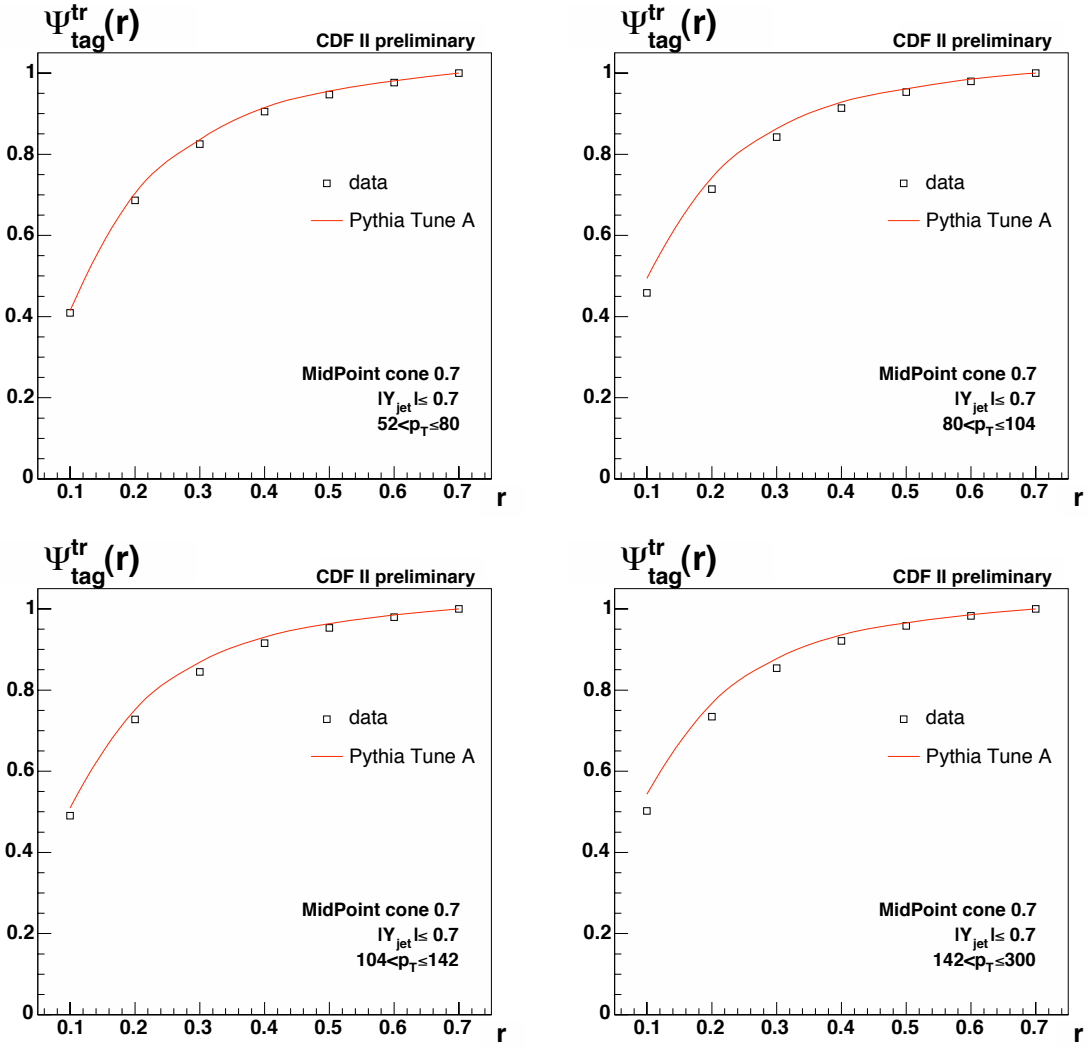

FIGURE 4.9: Raw integrated shapes for tagged jets measured using tracks. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

fractions of b-jets.

The bias corrections to correct for the tagging requirement and the hadron level corrections are re-computed for the shapes measured using the tracks. The hadron level b-quark jet shapes are shown in figure 4.10. The differences between the final hadron level b-jet shapes calculated using the default towers and using tracks are shown in figure 4.11. These differences are included as systematic errors. The largest difference between the hadron level b-quark jet shapes reconstructed from tracks and calorimeter towers is obtained for the lowest p_T bin for which the difference is of the order of 0.05.

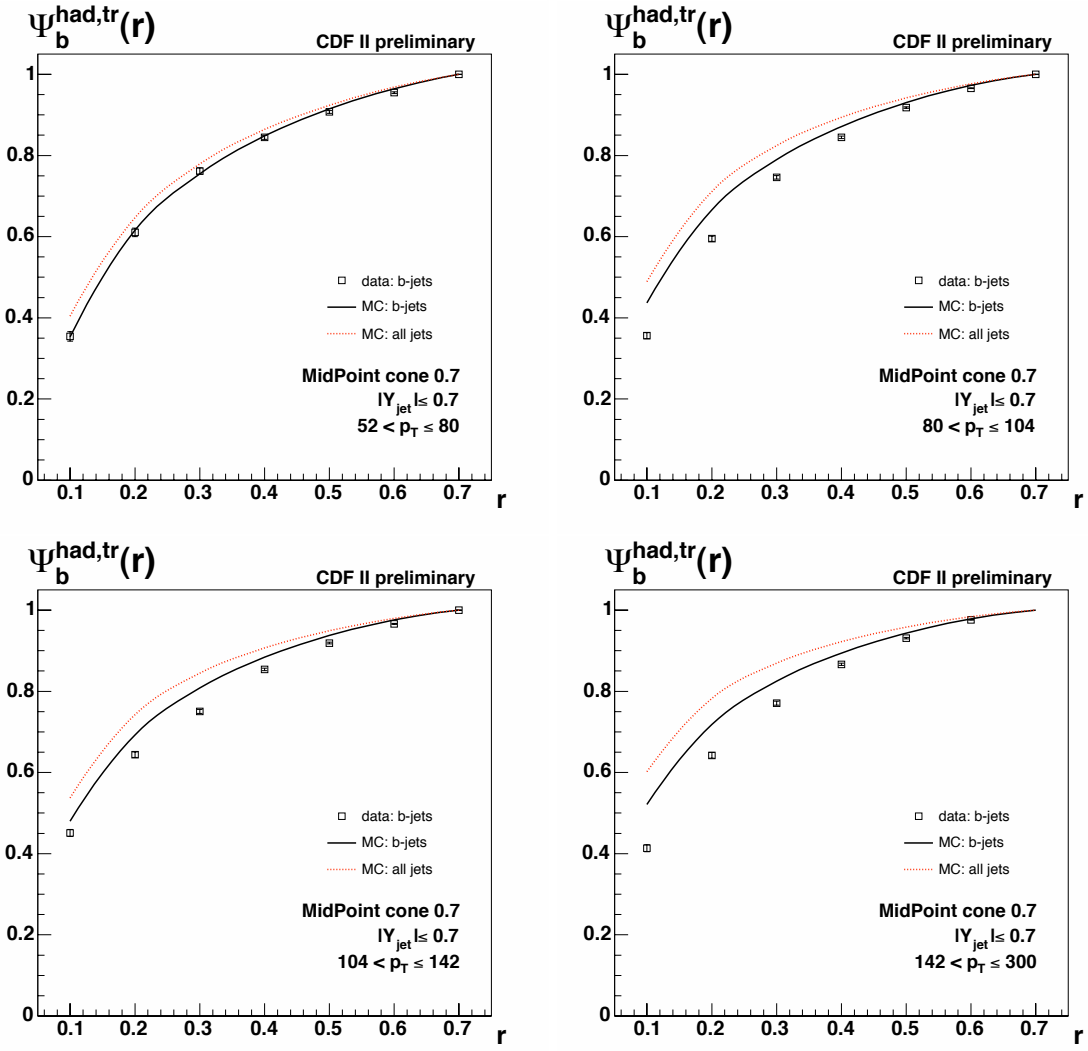

FIGURE 4.10: Hadron level integrated b-jet shapes for each of the p_T bins considered, starting from tracks. The results are shown as black points, only the statistical errors are shown. The MC predictions for inclusive jets and for b-jets are shown as red and black curves, respectively.

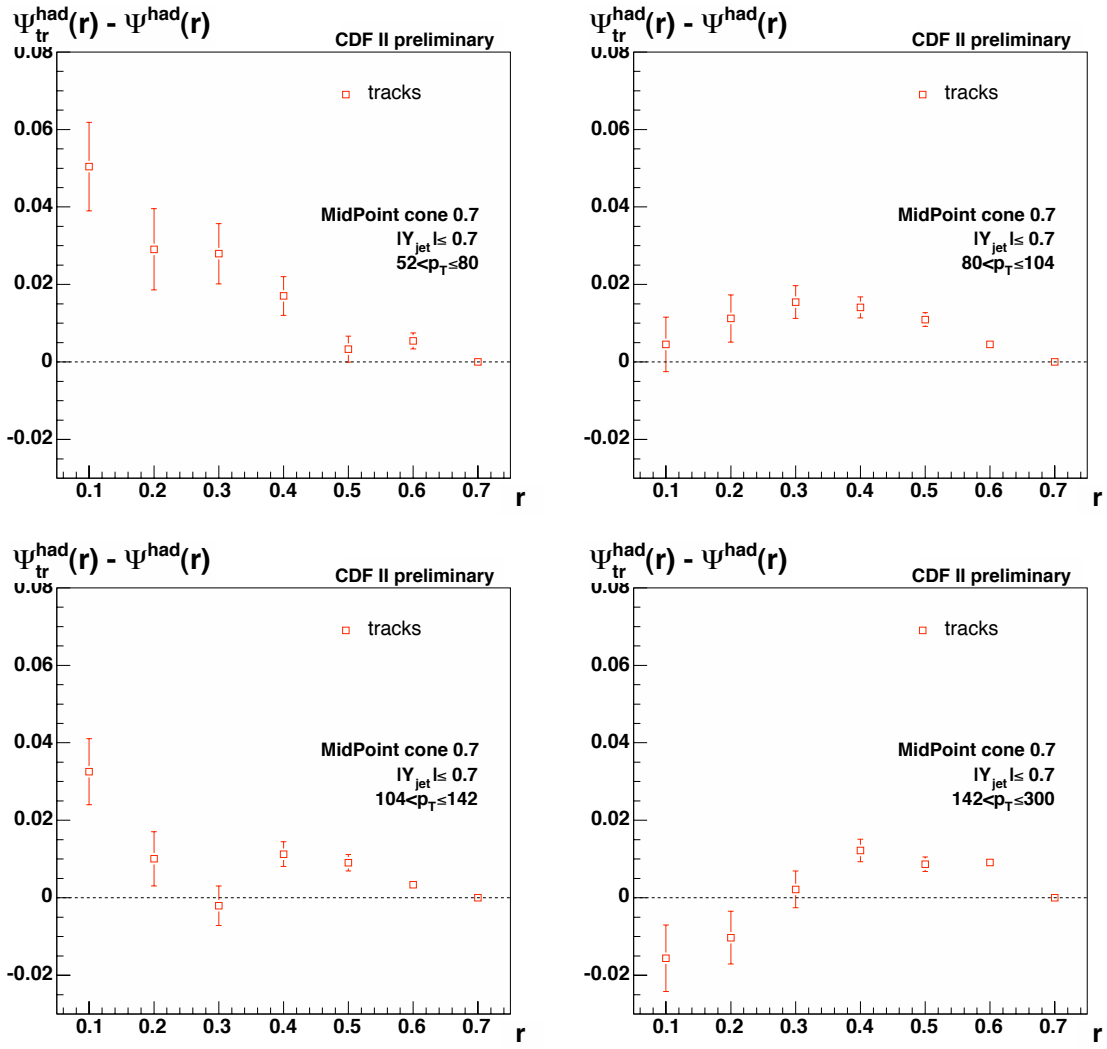

FIGURE 4.11: Difference between the hadron level integrated b-jet shapes starting from tracks and starting from calorimeter towers (red points). Only the statistical errors of the shapes reconstructed using tracks are shown.

4.9 Hadron Level b-jet Shapes Obtained Using Calorimeter Towers with $p_T > 0.5\,\mathrm{GeV}$

The hadron level b-jet shapes obtained starting from the raw calorimetric shapes with the default cut on the p_T of the towers at 0.1 GeV are compared to those obtained with a cut on the p_T of the towers at 0.5 GeV. This study is carried out in order to investigate how well the CDF detector simulation models low transverse momentum calorimeter energy deposits.

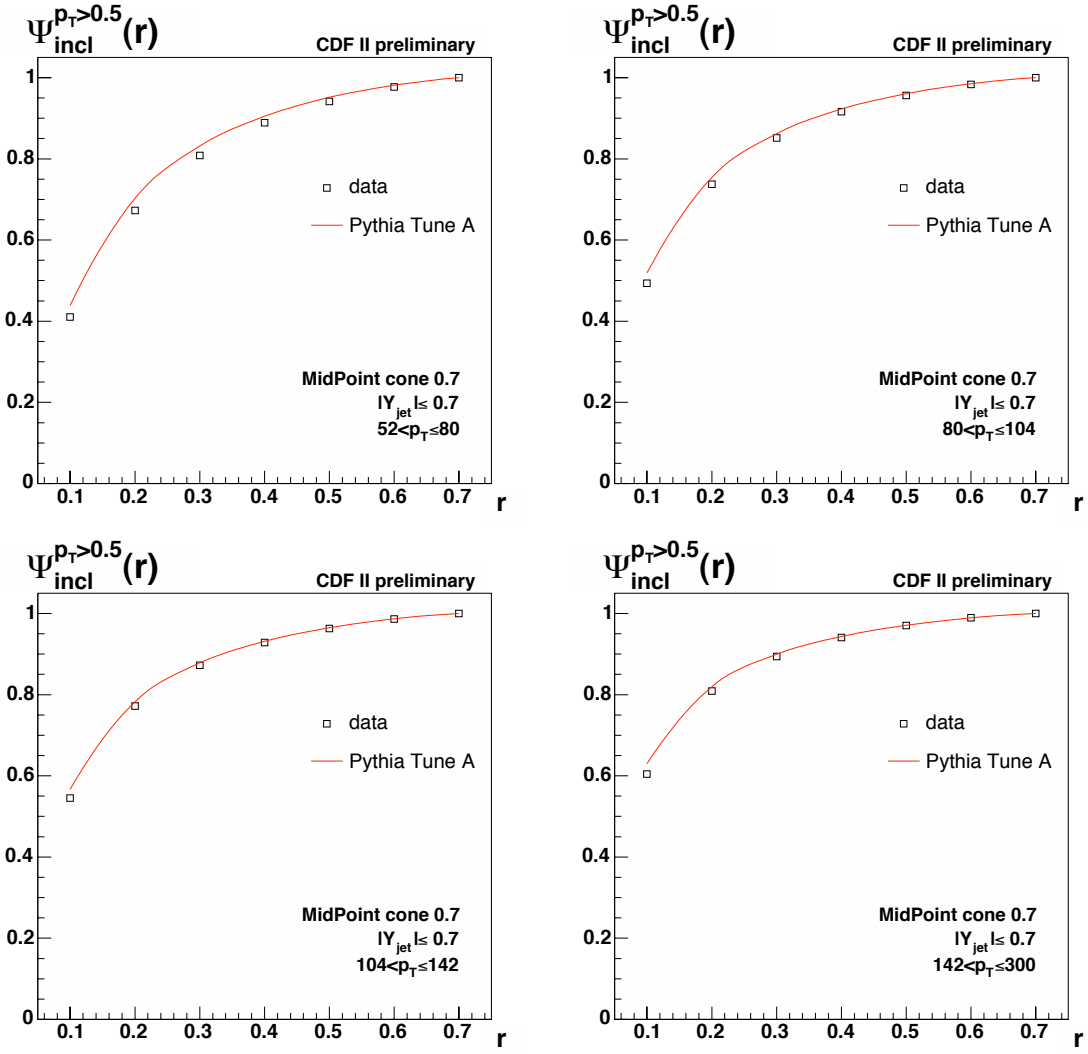

FIGURE 4.12: Raw integrated shapes for inclusive jets measured using calorimeter towers with a cut on the p_T of the towers at 0.5 GeV. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

Figure 4.12 shows a comparison between data and the Pythia Tune A MC of the raw inclusive shapes measured with the additional p_T cut. Figure 4.13 shows the same comparison for tagged jets. As for the default case, the agreement between data and MC is not perfect. The data and MC shapes for the tagged jets are expected to be different because of the different fractions of b-jets in data and MC.

The bias corrections to correct for the tagging requirement and the hadron level corrections are re-computed for the raw shapes measured with this additional p_T cut on the towers. The

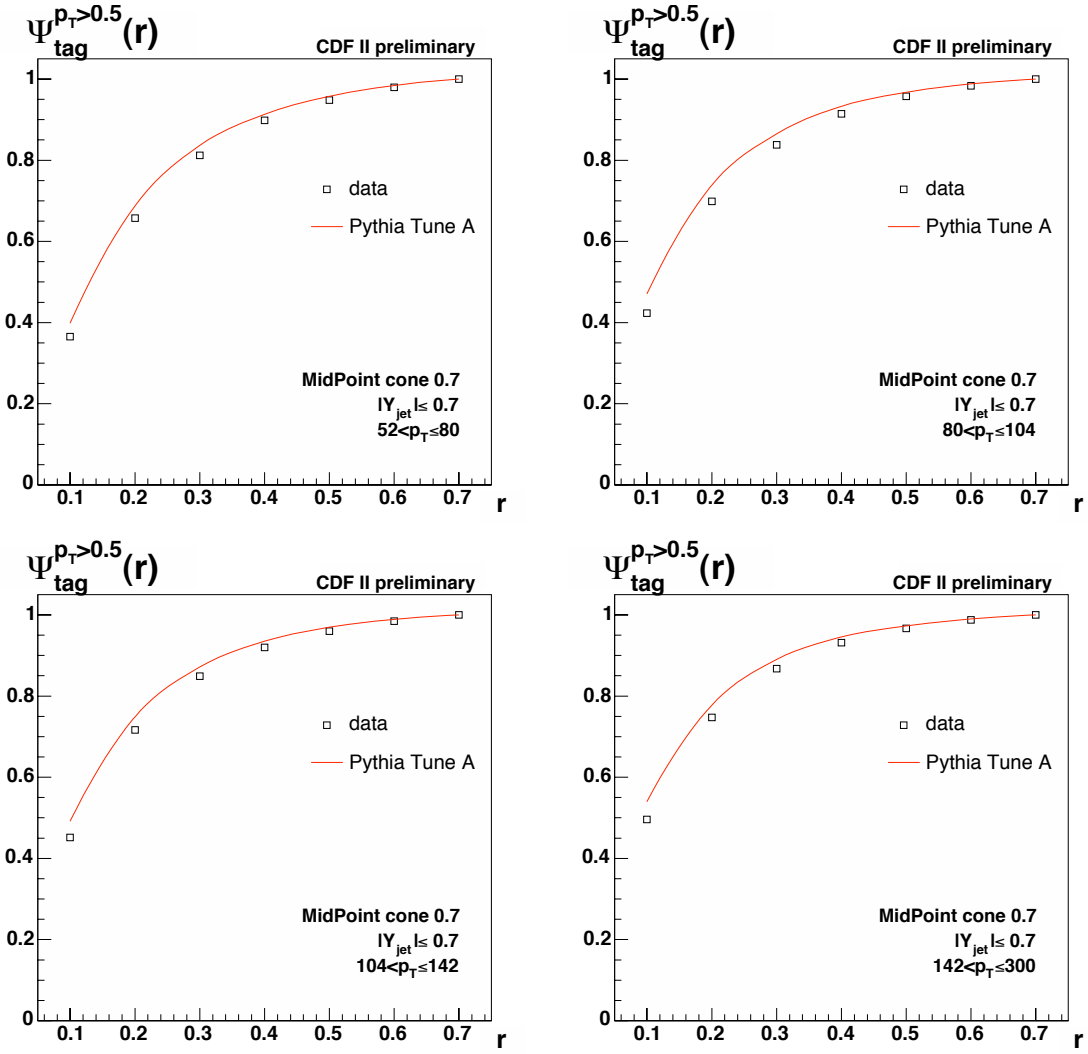

FIGURE 4.13: Raw integrated shapes for tagged jets measured using calorimeter towers with cut on the p_T of the towers at 0.5 GeV. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

hadron level b-jet shapes measured using the additional p_T cut on the towers are shown in figure 4.14. The differences between the hadron level b-jet shapes calculated using the default cut on the tower p_T and the cut at 0.5 GeV are shown in figure 4.15. This difference is included as a source of systematic error. It is found to be of similar size to the statistical errors and therefore not a dominant source of systematic uncertainty. The largest difference with respect to the default cut on the calorimeter tower p_T is about 0.005.

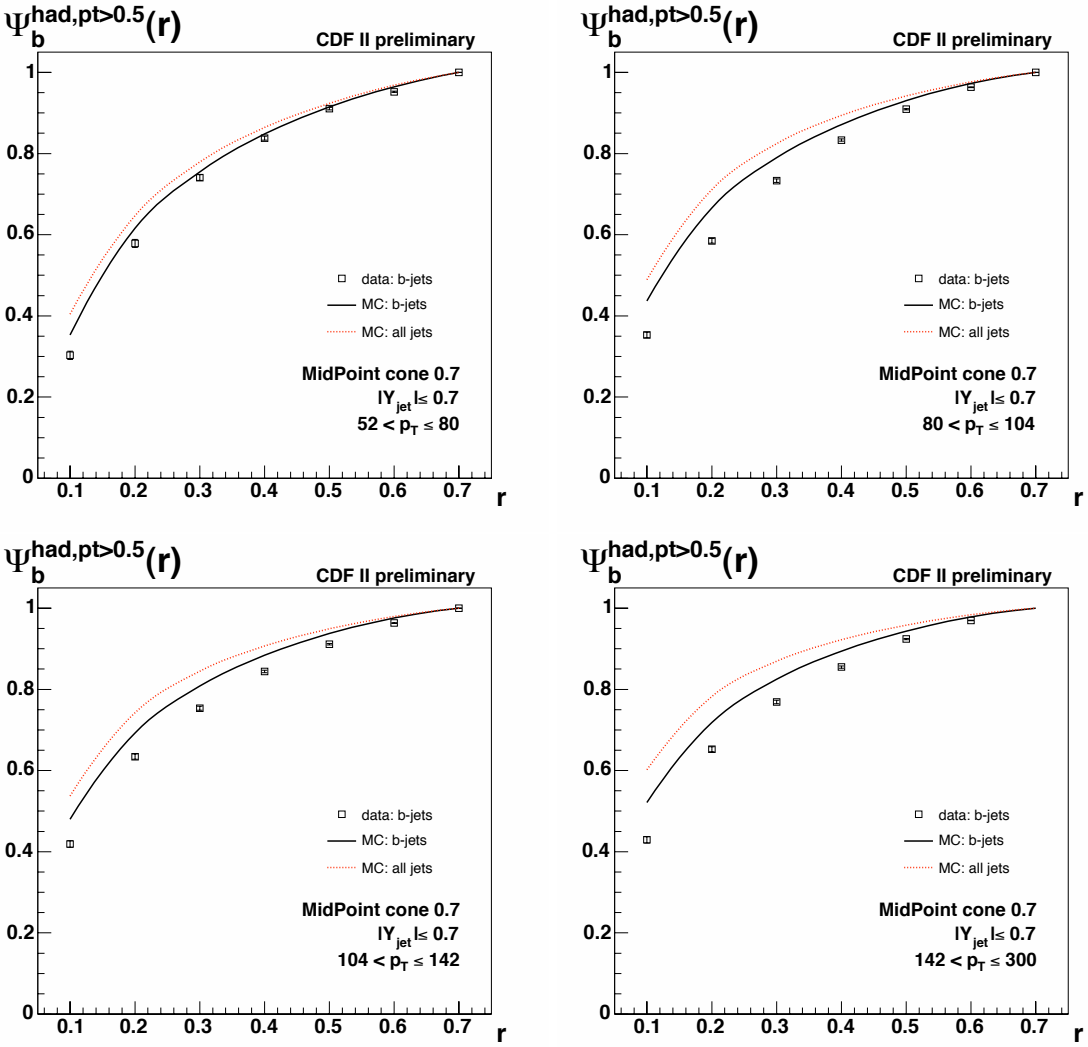

FIGURE 4.14: Hadron level integrated b-jet shapes for each of the p_T bins considered, starting from calorimeter towers with p_T larger than 0.5 GeV. The results are shown as black points, only the statistical errors are shown. The MC predictions for inclusive jets and for b-jets are shown as red and black curves, respectively.

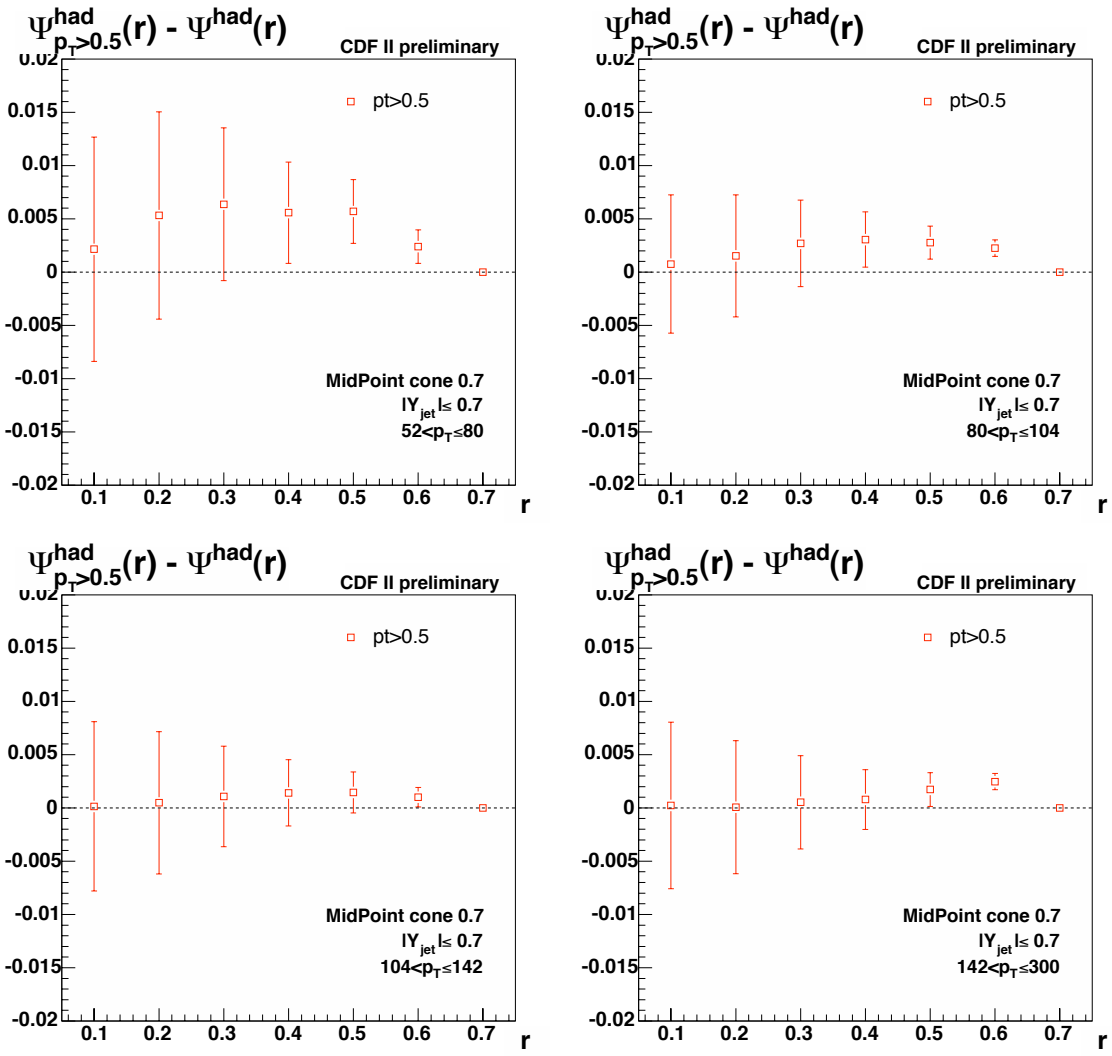

FIGURE 4.15: Difference between the hadron level integrated b-jet shapes starting from calorimeter towers with p_T larger than 0.5 GeV and the default cut at 0.1 GeV. The results are shown as black points, only the statistical errors of the shapes from towers with $p_T > 0.5$ GeV are shown.

4.10 MC Dependence of the Unfolding Parameters

On top of the effects mentioned in the previous sections, the MC dependence of the parameters of the unfolding procedure is investigated. These studies are detailed in the next sections.

4.10.1 Purity

The difference in the templates obtained with different MC samples is investigated. Because the statistics of the Herwig MC samples is not very large, it was necessary to carry out the same procedure as detailed in section 4.2. The higher p_T Herwig nonb-templates were used. As for the c-fraction extraction, the lowest p_T bin cannot be considered because of the different p_T distributions. The Herwig templates agree within statistical errors with those from Pythia Tune A. Figure 4.16 shows the comparison between the templates using Pythia Tune A (solid lines) and Herwig (points). The fit results using the higher p_T Herwig samples are within the quoted fit errors of the Pythia Tune A results, as can be seen in table 4.3.

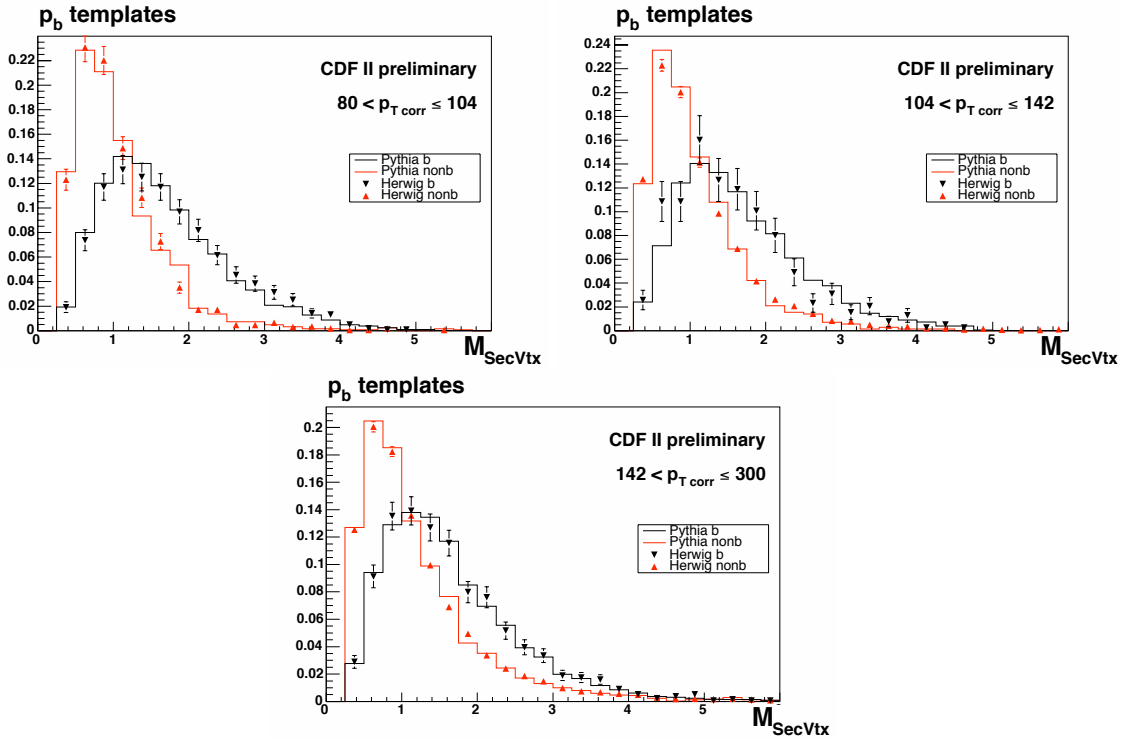

FIGURE 4.16: Secondary vertex mass distributions for b- (black) and nonb-jets (red) for all the p_T bins. The full lines represent the Pythia Tune A MC templates whereas the points represent the Herwig templates using the higher p_T MC samples.

A similar procedure can be applied to investigate the effect of the fragmentation function and PDFs on the secondary vertex mass templates. Because only b-filtered samples are available for the Pythia Tune A MC samples with the CTEQ6L PDFs as well as those with the Peterson fragmentation function, the comparison is only made for the b-jet templates. The b-jet templates for the default Pythia Tune A MC are shown in figures 4.17 and 4.18 along with the b-jet template obtained using Pythia Tune A with CTEQ6L PDFs for the former and with the Peterson fragmentation functions for the latter. The statistics for the comparison to Pythia Tune A with the Peterson fragmentation for the lowest bin are not very high. This shows that

	Herwig		
p_T range	p_{b}	fit error	Δp_{b}
80-104	0.342	0.024	0.019
104-142	0.327	0.032	0.027
142-300	0.253	0.020	0.011

TABLE 4.3: Fitted purities using the Herwig templates. The total fit errors quoted by the fitting algorithm are shown as well as the differences between the fit results and those obtained with Pythia Tune A.

at least for b-jets, where we expect the difference to be maximal, the templates are very similar.

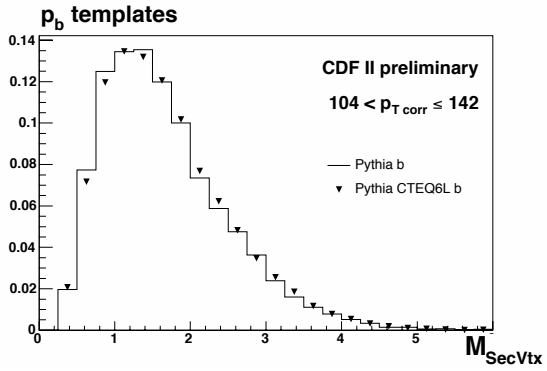

FIGURE 4.17: Secondary vertex mass distributions for b-jets. The line represents the default Pythia Tune A templates, whereas the points represent the Pythia Tune A predictions using the CTEQ6L PDFs.

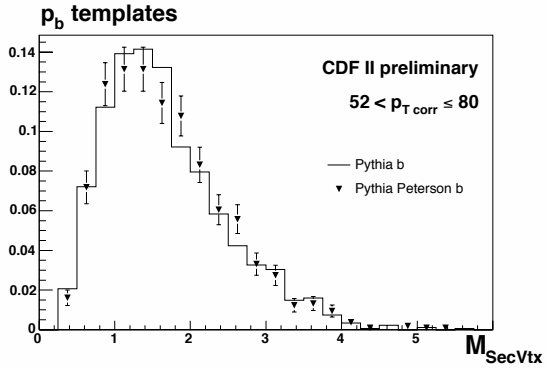

FIGURE 4.18: Secondary vertex mass distributions for b-jets. The line represents the default Pythia Tune A templates whereas the points represent the Pythia Tune A predictions using the Peterson fragmentation function.

There is not enough statistics in the samples produced with different PDFs or a different fragmentation function to carry out the secondary vertex mass fit. It is therefore not possible to compare fitted purities to verify that these samples do not change the outcome of the fit. Nevertheless it is possible to see, by comparing the templates for b-quark jets to the default Pythia Tune A MC ones, that the differences between the fitted purities should not be large.

4.10.2 Biases Due to SecVtx Tagging

The bias corrections are re-computed for all the different MC samples available: Pythia Tune A, Herwig, Pythia Tune A with CTEQ6L PDFs, Pythia Tune A with the Peterson fragmentation model. The biases obtained are compared to the default Pythia Tune A ones.

Biases Due to the Tagging on b-jets

Figure 4.19 shows the comparison of the tagging biases on b-jets between Pythia Tune A (red line) and Herwig (black full points). The biases obtained using Herwig are compatible with those obtained using Pythia Tune A. The errors on the biases for Herwig are relatively large due to the limited b-jet statistics.

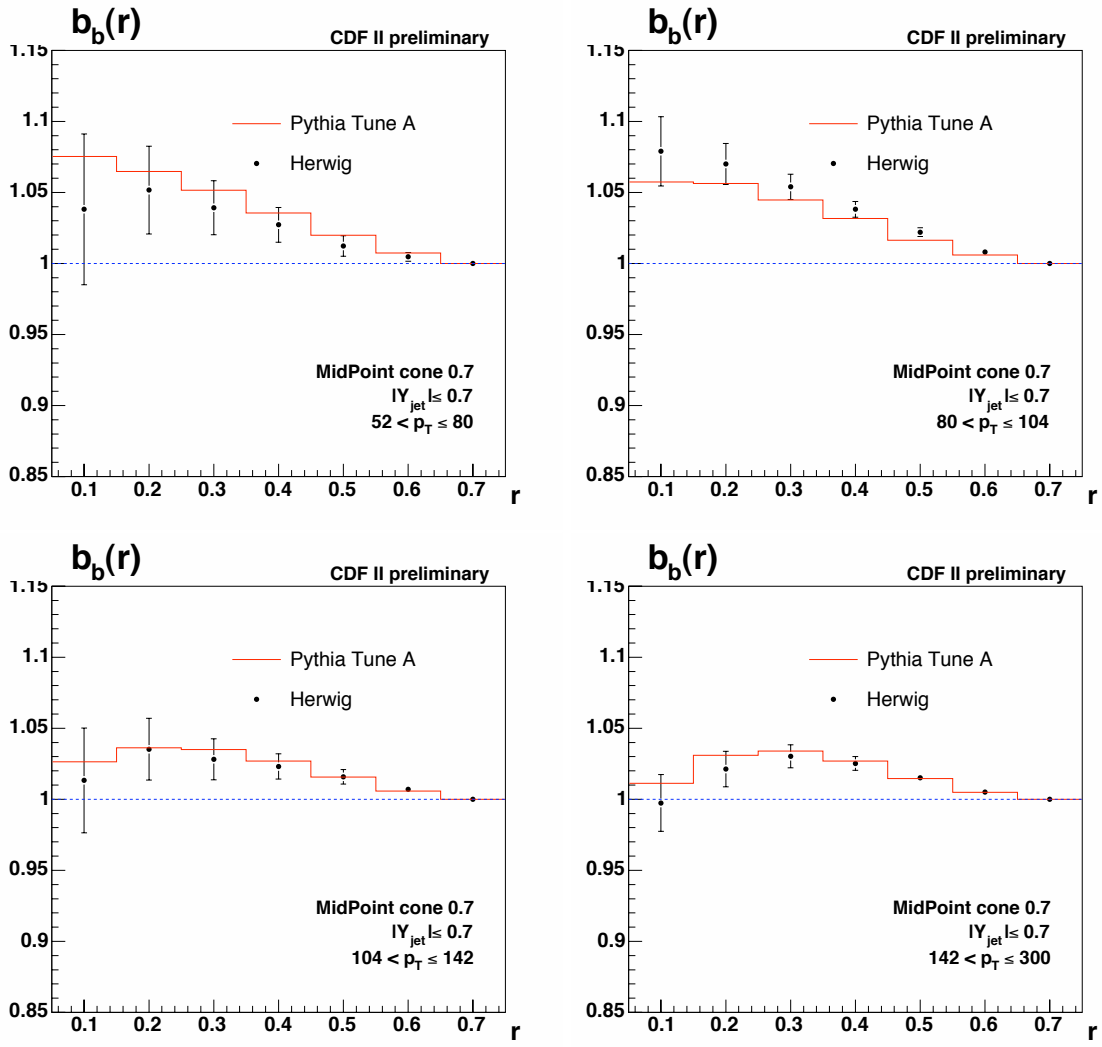

FIGURE 4.19: Comparison between Herwig (black full points) and Pythia Tune A (red line) MC of the biases due to the tagging on b-jets. The errors shown are the MC statistical errors.

Figure 4.20 shows a comparison of the tagging biases on b-jets between the default Pythia Tune A MC and Pythia Tune A with the Peterson fragmentation model. The difference in the bias for the first bin in r is relatively large but all other bins are similar to the default Pythia Tune A biases.

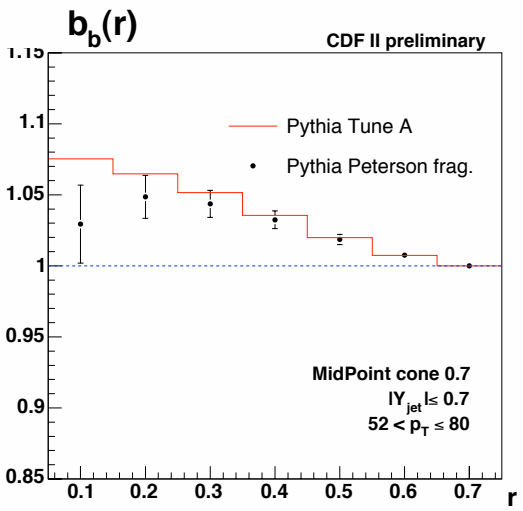

FIGURE 4.20: Comparison of the biases due to the tagging on b-jets between the default Pythia Tune A MC (red line) and Pythia Tune A with the Peterson fragmentation model (black full points). The errors shown are the MC statistical errors.

Figure 4.21 shows a comparison of the tagging biases to b-jets between the default Pythia Tune A MC and Pythia Tune A with CTEQ6L PDFs. The biases obtained for Pythia Tune A with CTEQ6L PDFs are similar the default Pythia Tune A biases.

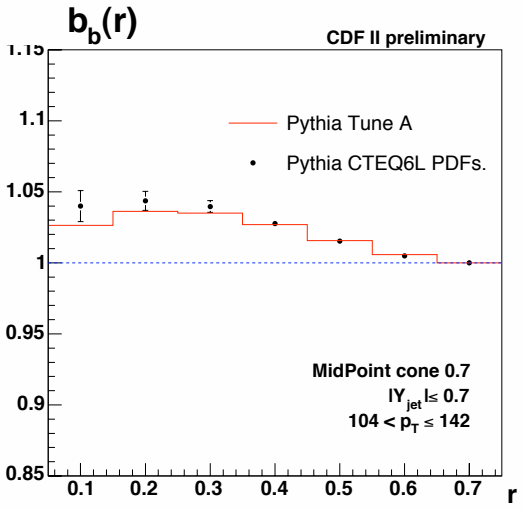

FIGURE 4.21: Comparison of the biases due to the tagging on b-jets between the default Pythia Tune A MC (red line) and Pythia Tune A with CTEQ6L PDFs (black full points). The errors shown are the MC statistical errors.

Biases Due to the Tagging on nonb-jets

The tagging biases on nonb-jets obtained using Herwig are compatible with those obtained using Pythia Tune A. The errors on the biases are relatively large due to the limited statistics in the tagged MC samples. Figure 4.22 shows the comparison between Pythia Tune A (red line) and Herwig (black full points). It is not possible to make a comparison between the default Pythia

Tune A and Pythia Tune A with Peterson fragmentation or CTEQ6L PDFs because there is not enough statistics for tagged nonb-jets.

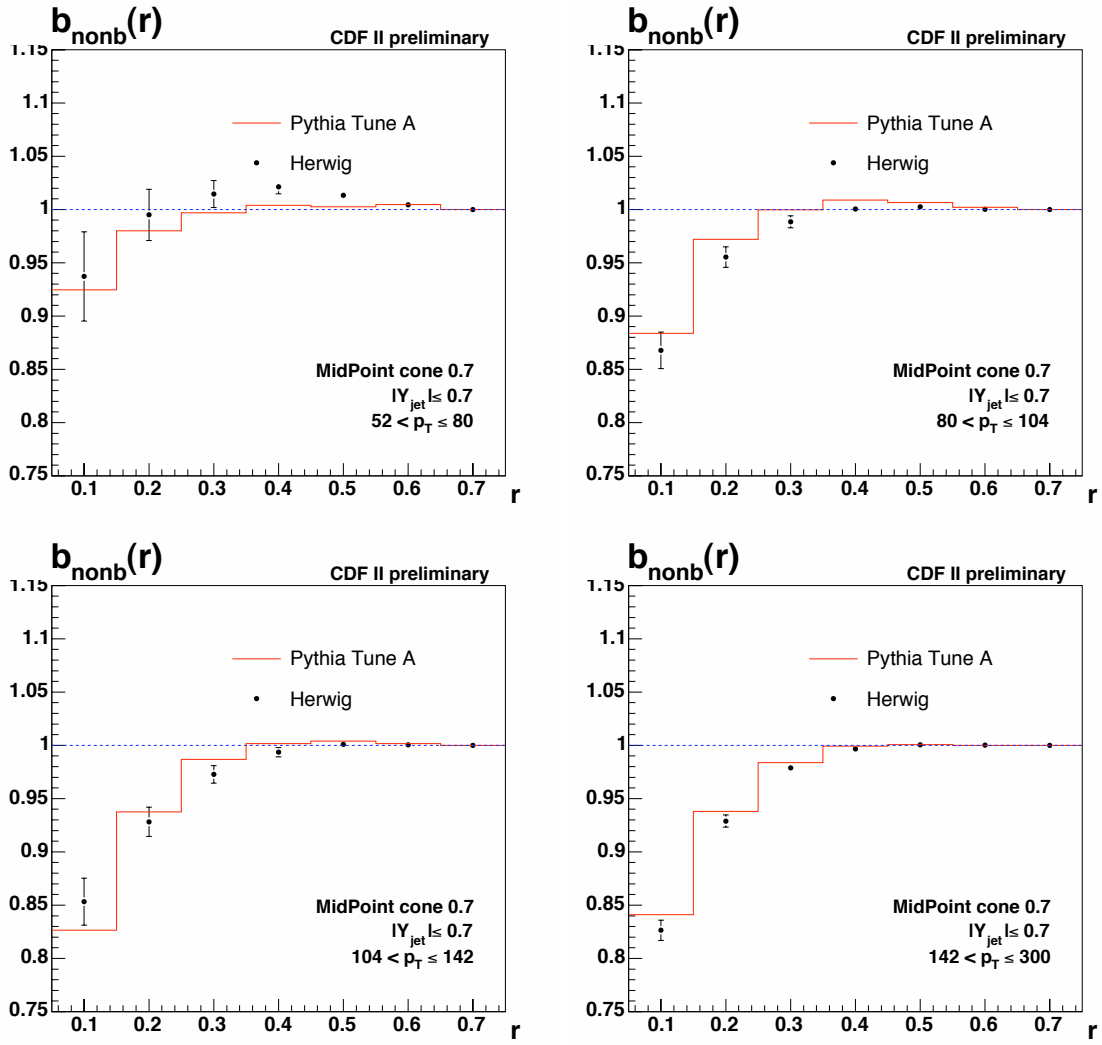

FIGURE 4.22: Comparison between Herwig (black full points) and Pythia Tune A (red line) MC of the biases due to the tagging on b-jets. The errors shown are the MC statistical errors.

4.10.3 Hadron Level Corrections to the Jet Shapes

The hadron level corrections are computed for all the different MC samples used. Figure 4.23 shows the comparison between the hadron level corrections obtained using Herwig and Pythia Tune A. The correction factors are not in good agreement for the 2^{nd} and 4^{th} jet p_T bins. The statistics for the 1^{st} bin does not allow any conclusions to be drawn. The hadron level b-jet shapes are shown in figure 4.24. The shapes in Herwig are slightly wider than in Pythia Tune A. Figure 4.25 shows the same comparison for detector level b-quark jet shapes where the agreement between Herwig and Pythia Tune A is reasonably good. It is not possible to draw any conclusions about any systematic effect due to the MC dependence of the hadron level corrections from these comparisons because of the low statistics of the Herwig samples. Section 4.10.4 compares the final hadron level b-quark jet shapes unfolded using the Pythia Tune A MC to those obtained using the Herwig MC. This was only possible for the three highest p_T bins

because of the lack of statistics for the lowest p_T bin.

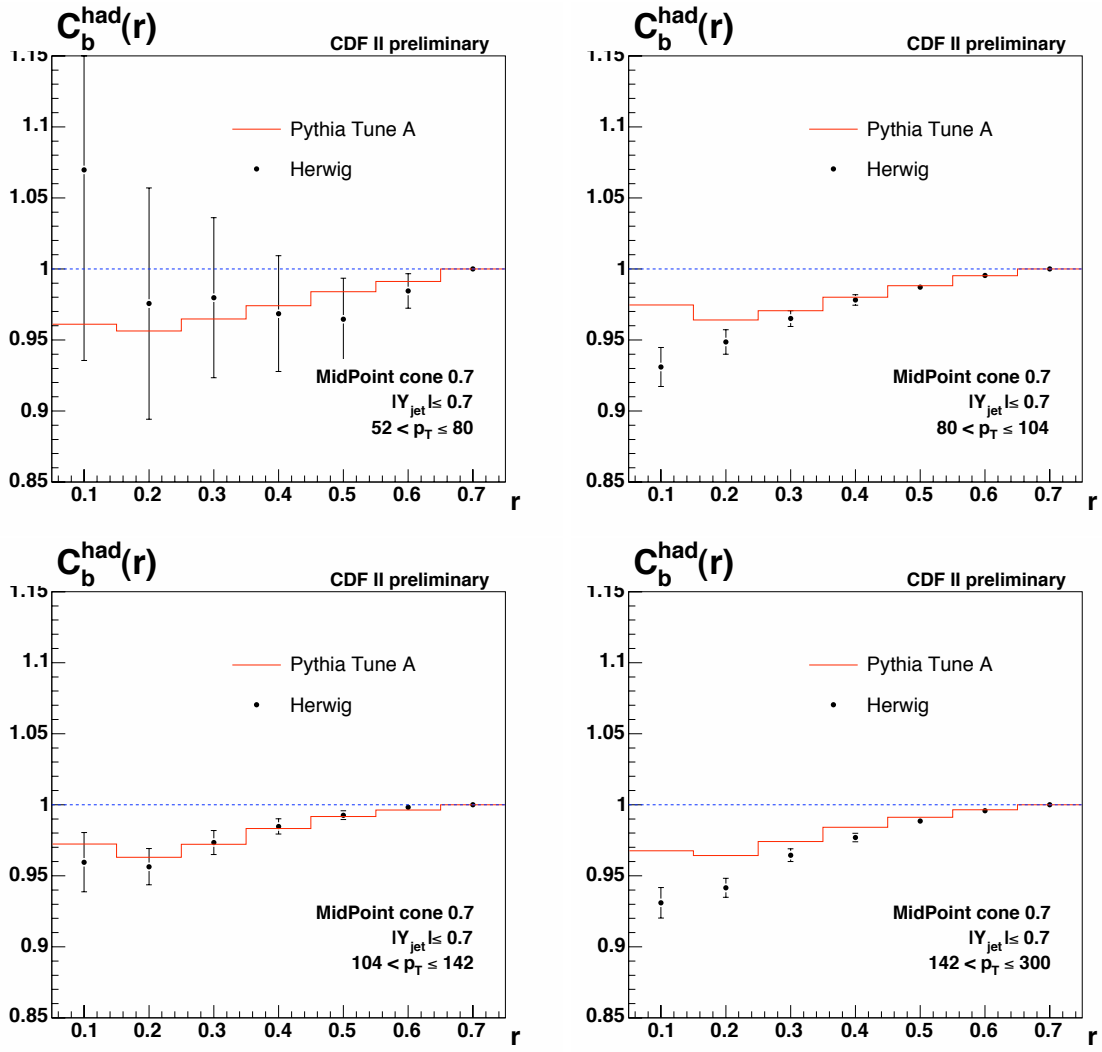

FIGURE 4.23: Comparison between Herwig (black full points) and Pythia Tune A (red line) MC of the hadron level corrections to b-jets. The errors shown are the MC statistical errors.

It is also possible to compare the default hadron level corrections to those obtained with Pythia Tune A using the Peterson fragmentation model. This is shown in figure 4.26. The hadron level corrections are compatible.

Finally, a comparison is made between the hadron level corrections to Pythia Tune A using the CTEQ6L PDFs and the default Pythia Tune A. This is shown in figure 4.27. The hadron level corrections are very similar for these two cases.

4.10.4 Hadron Level b-jet Shapes Using the Herwig MC for the Unfolding Parameters

Despite the fact that the statistics of the Herwig MC samples are not very large, it is possible to carry out the whole of the unfolding procedure using the tagging biases, hadron level corrections and secondary vertex mass templates from the Herwig MC samples instead of from the Pythia

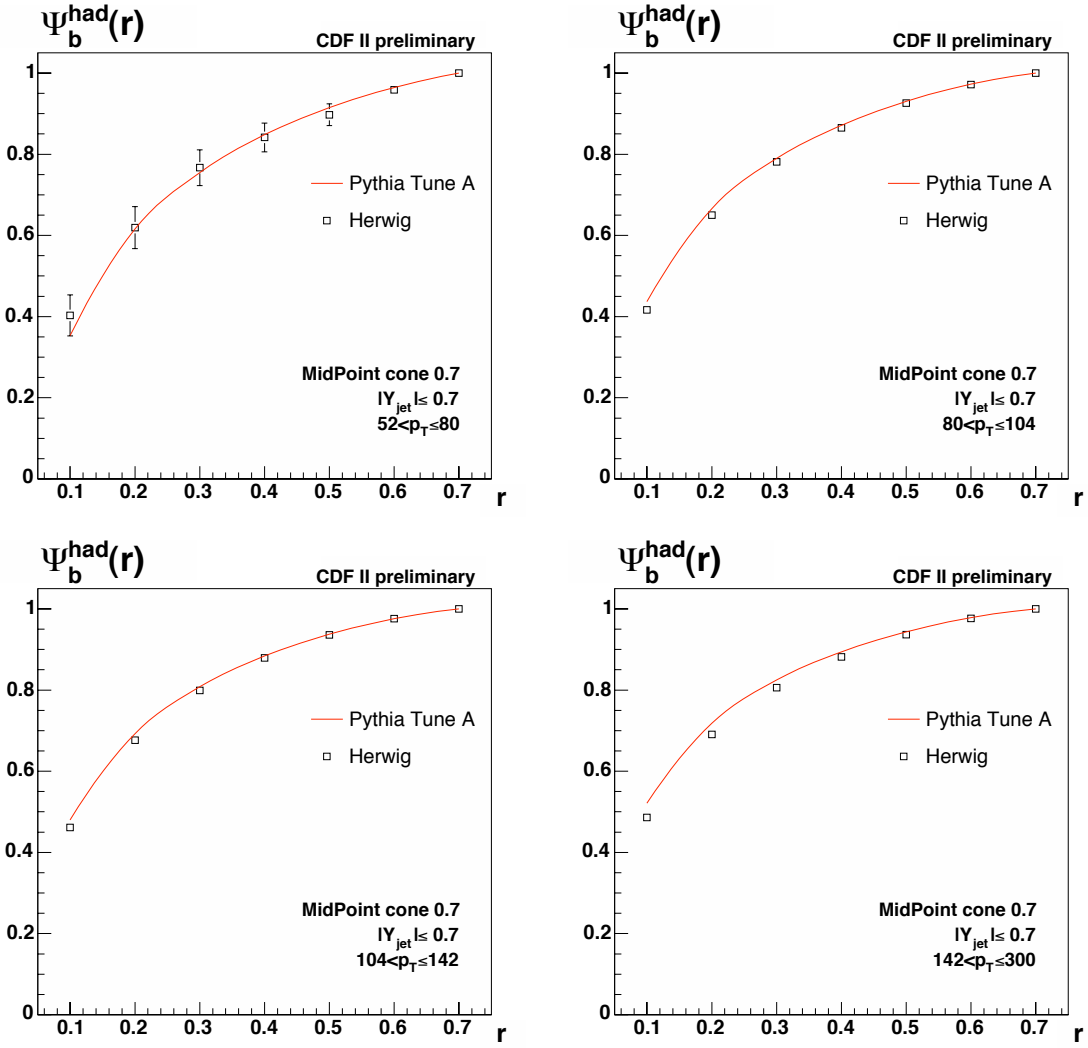

FIGURE 4.24: Comparison between Herwig (black full points) and Pythia Tune A MC (red line) of the hadron level b-jet shapes. The errors shown are the MC statistical errors.

Tune A samples. However, this was not possible for the lowest jet p_T bin because of the lack of statistics in that bin.

Figure 4.28 shows the hadron level b-quark jet shapes in data using the Herwig MC samples for the unfolding. Figure 4.29 shows the differences between the hadron level b-quark jet shapes using Herwig for the unfolding and using the default Pythia Tune A MC for the unfolding. The results are within the total systematic errors. Due to the lack of statistics in the lowest p_T bin, this effect can not be included as a systematic uncertainty. It is possible to conclude from this study that the hadron level b-quark jet shapes are not particularly sensitive to the use of a particular choice of MC for the unfolding parameters. In order to make a better comparison, more Herwig MC samples would have to be generated, particularly b-filtered samples, which was beyond the scope of this analysis.

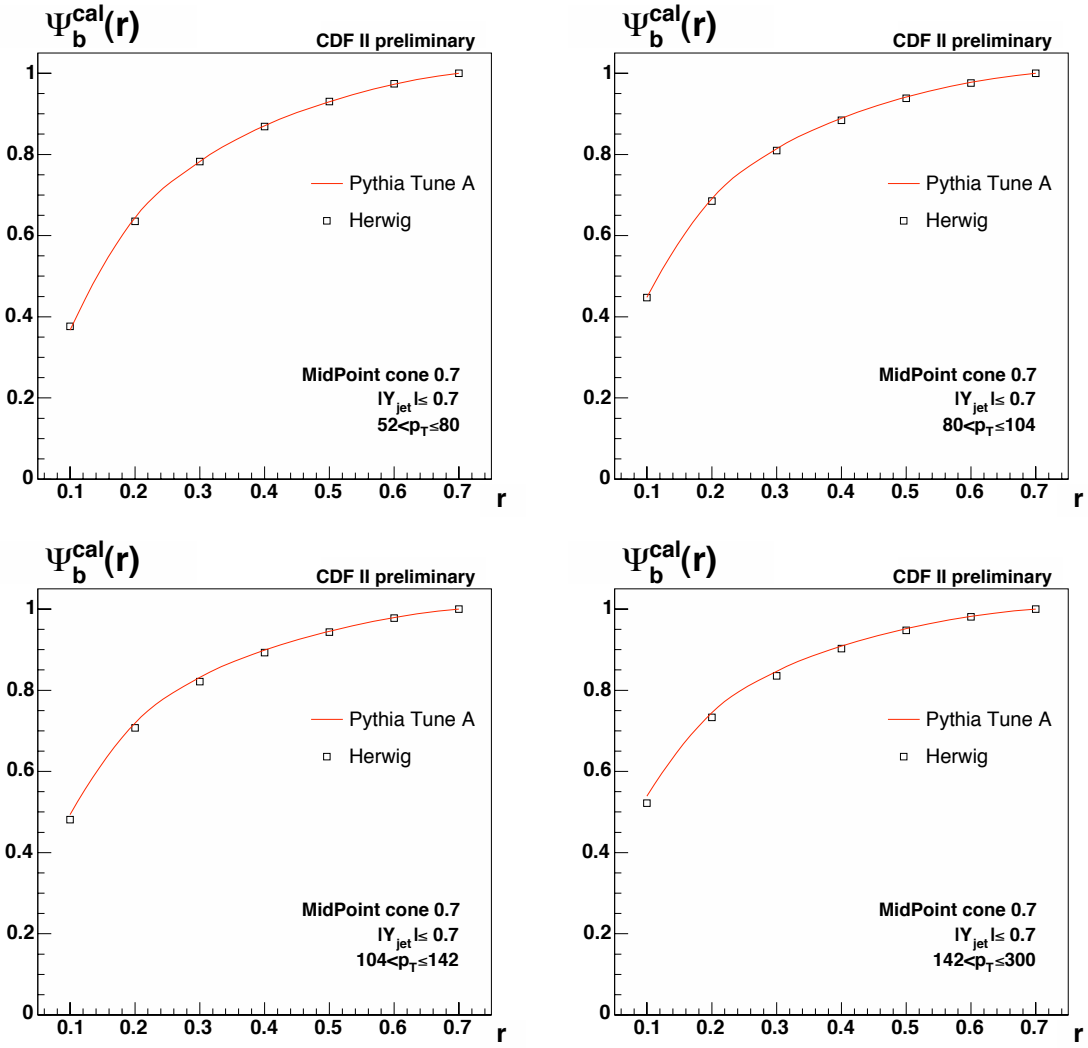

FIGURE 4.25: Comparison between Herwig (black full points) and Pythia Tune A MC (red line) of the detector level b-jet shapes. The errors shown are the MC statistical errors.

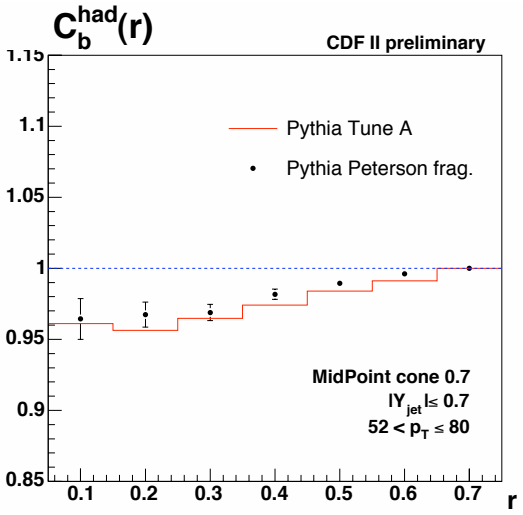

FIGURE 4.26: Comparison between Pythia Tune A using the Peterson fragmentation model (black full points) and the default Pythia Tune A MC (red line) of the hadron level corrections to b-jets. The errors shown are the MC statistical errors.

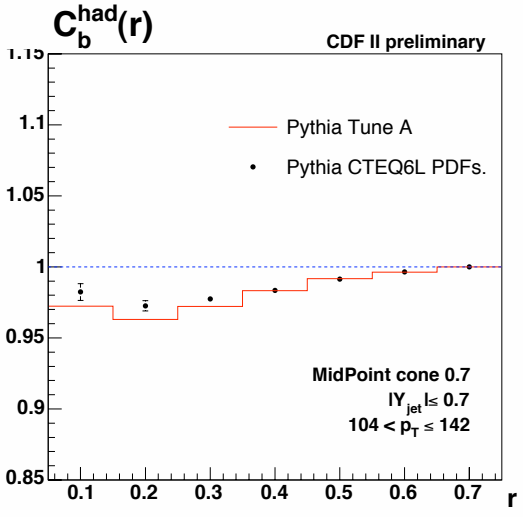

FIGURE 4.27: Comparison between Pythia Tune A using the CTEQ6L PDFs (black full points) and the default Pythia Tune A MC (red line) of the hadron level corrections to b-jets. The errors shown are the MC statistical errors.

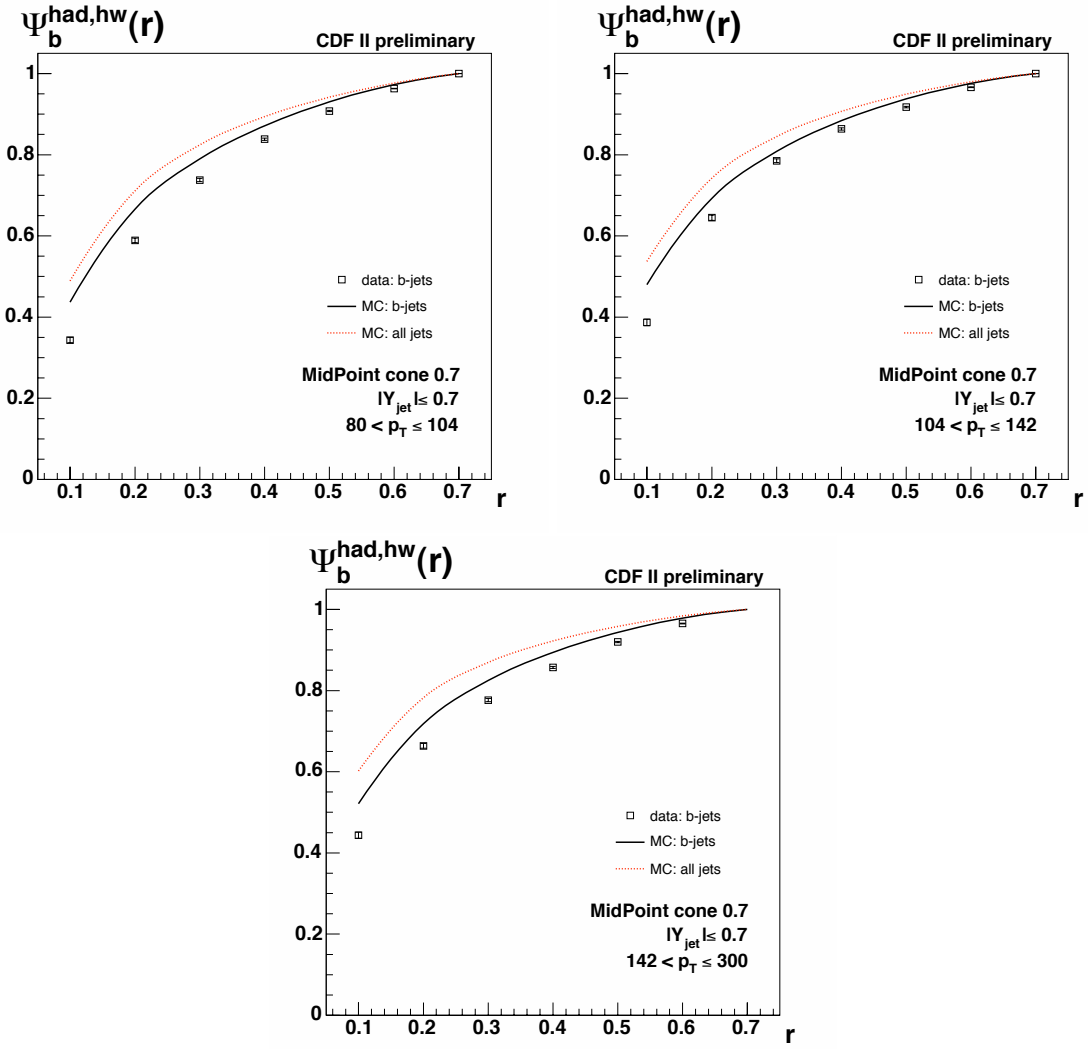

FIGURE 4.28: Hadron level b-quark jet shapes. Data (shown as black points) using the unfolding parameters obtained from the Herwig MC samples is compared to the Pythia Tune A MC predictions (red line). Only the statistical errors are shown.

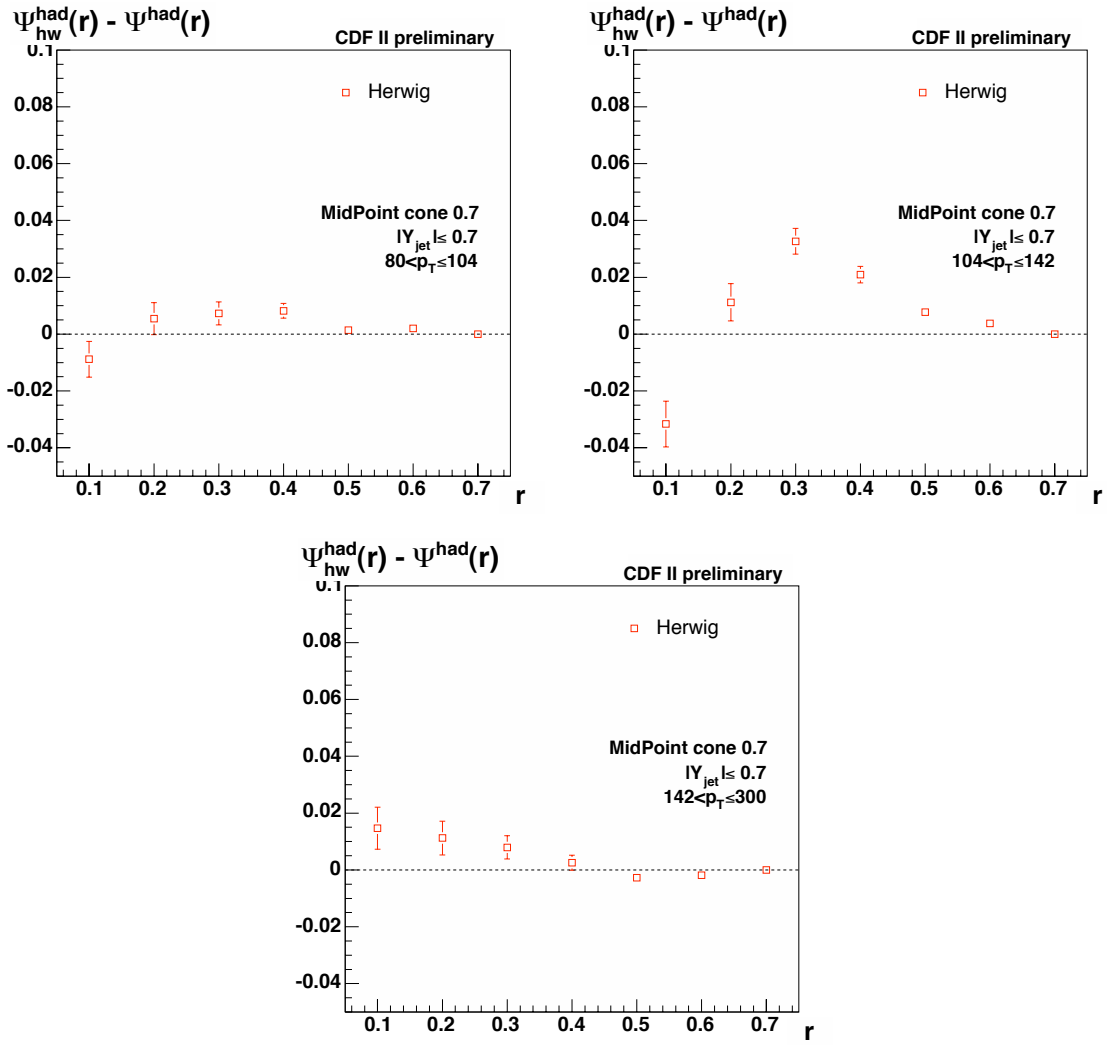

FIGURE 4.29: Difference between hadron level b-quark jet shapes using the unfolding parameters obtained from the Herwig MC samples and the ones obtained from Pythia Tune A. Only the statistical errors when using the Herwig unfolding factors are shown.

4.11 Dependence on the MC Modelling of the SecVtx Performance

Despite huge progress in the past two years in the understanding of the SecVtx performance and in particular in the differences in performance between data and MC, there are still a number of parameters, relevant to SecVtx, which do not agree perfectly between data and MC. In particular the number of Pass 1 tracks does not agree perfectly between data and MC, as can be seen in figure 4.30.

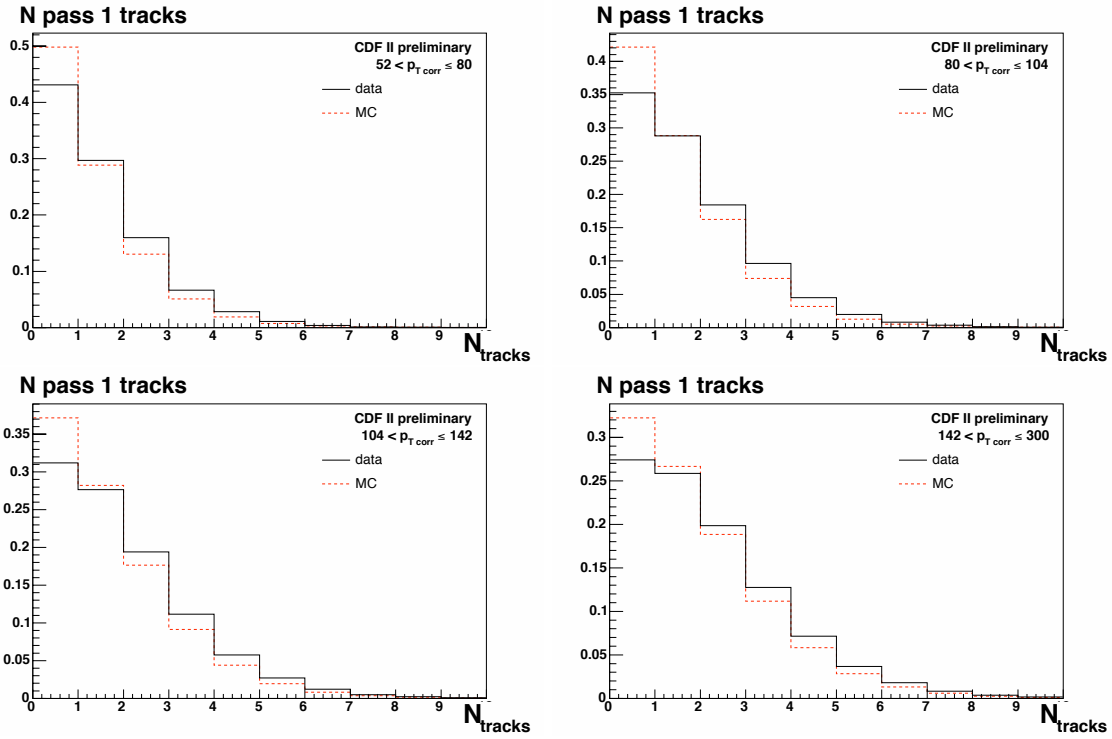

FIGURE 4.30: Plot of the distribution of the number of Pass 1 tracks inside the jet for all jets.

In order to investigate if this discrepancy between the data and the MC affects any of the unfolding parameters, the Pythia Tune A MC samples are re-weighted in such a way that the distributions of the number of Pass 1 tracks agree between data and MC. The weight function is defined for each value of the number of tracks as

$$W(n) = \frac{f_{\text{data}}(n)}{f_{\text{MC}}(n)} \tag{4.3}$$

where $f_{\text{data/MC}}(n)$ is defined as the fraction of inclusive jets which pass the selection cuts that have n Pass 1 tracks for data and MC, respectively. These functions are normalised to unit area. Having defined the weight function for each of the four p_T bins, the jet shapes in MC are re-weighted according to the number of Pass 1 tracks in the jet. The biases due to tagging on b-jets (figure 4.31) and on nonb-jets (figure 4.32) are re-computed using the new weighted shapes. The hadron level corrections are also re-computed (figure 4.33). These biases and hadron level corrections are compared to the default values obtained for the unweighted shapes. The biases are always found to be within the statistical errors of the default biases. The hadron level corrections are also compatible, within statistical errors, to those obtained using the default MC shapes. There is therefore no systematic uncertainty associated to the slight differences with

respect to data of the modelling of the SecVtx performance in Pythia Tune A MC.

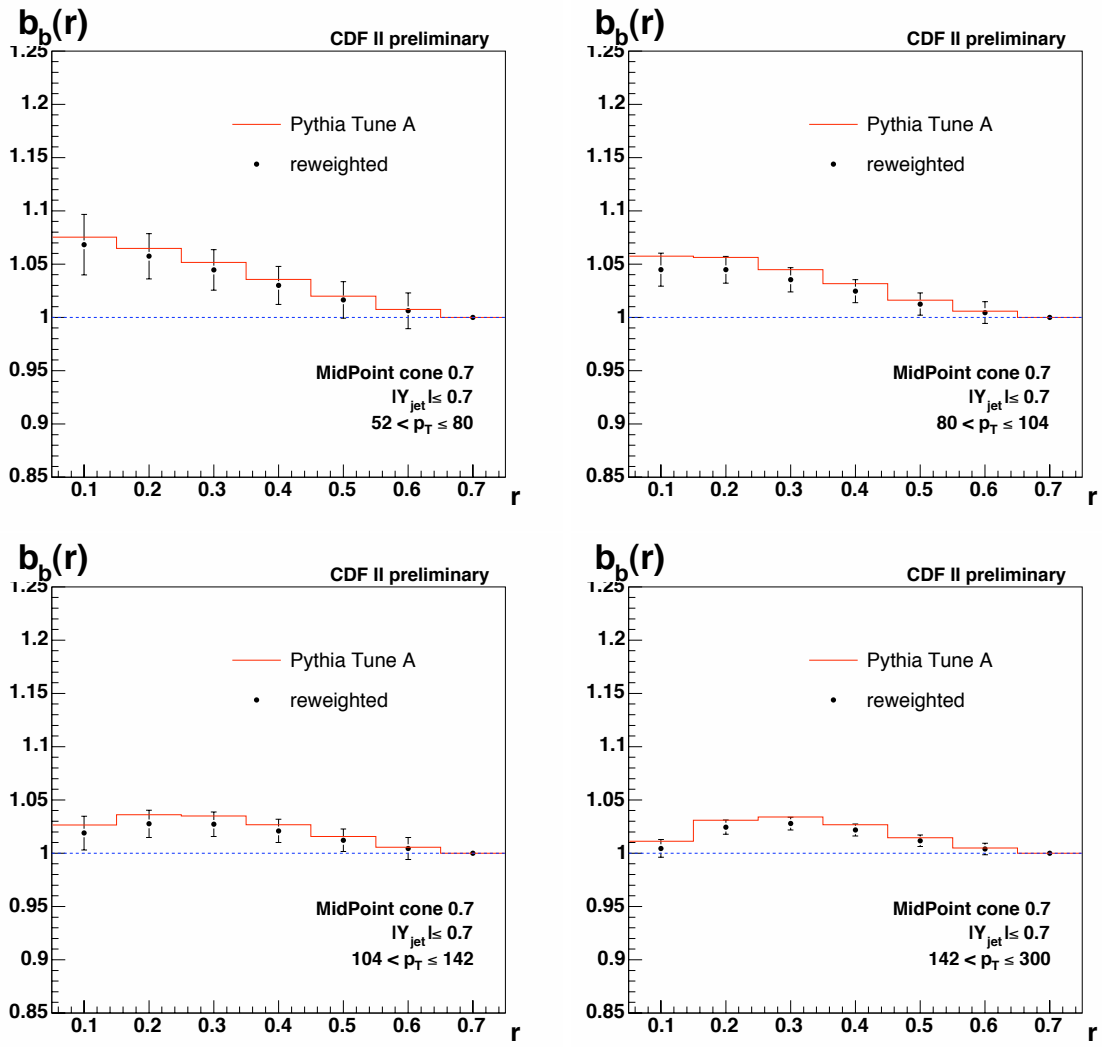

FIGURE 4.31: Comparison of the tagging biases on b-jets between the default Pythia Tune A MC shapes (red line) and the reweighted ones (black points). Only the MC statistical errors are shown.

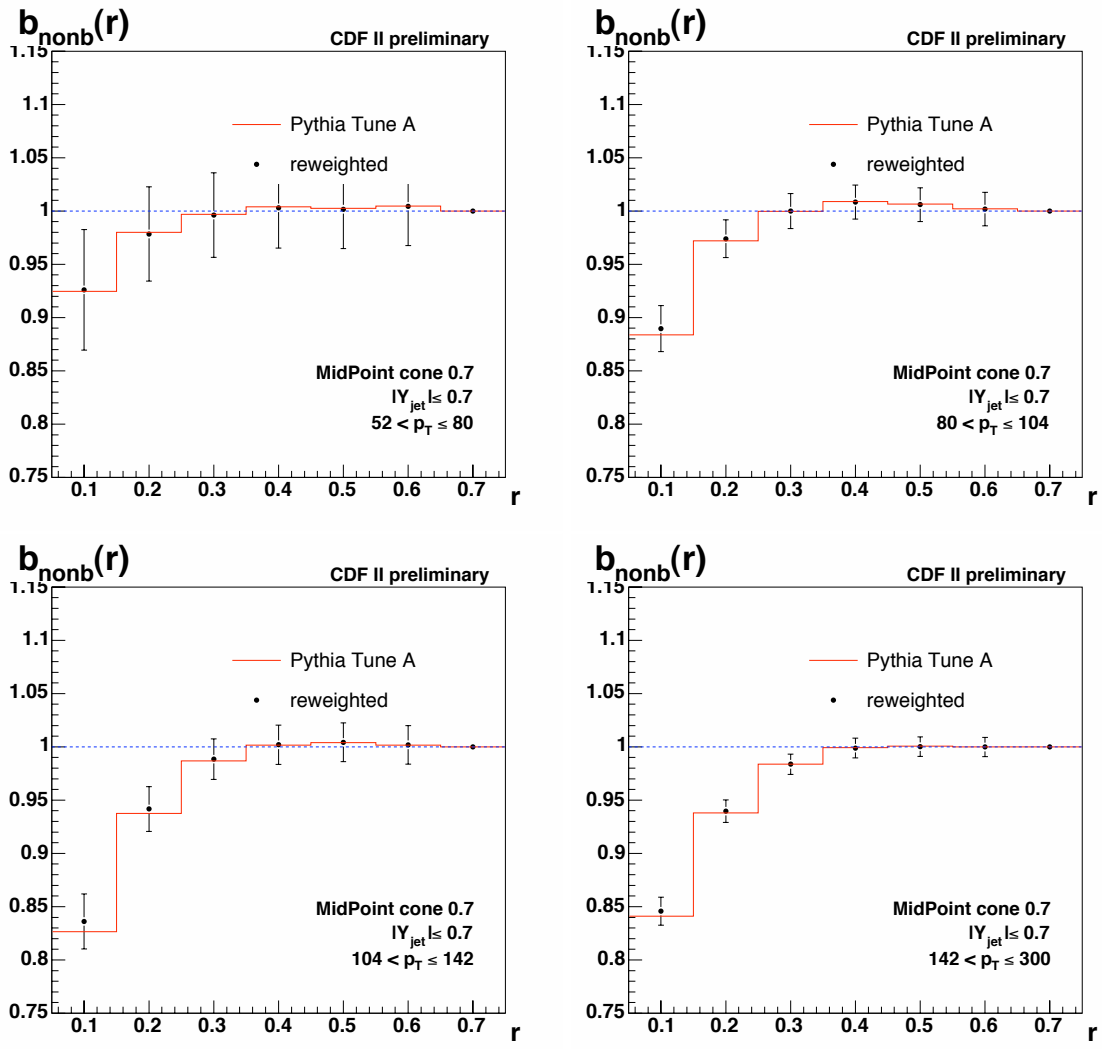

FIGURE 4.32: Comparison of the tagging biases on nonb-jets between the default Pythia Tune A MC shapes (red lines) and the reweighted ones (black points). Only the MC statistical errors are shown.

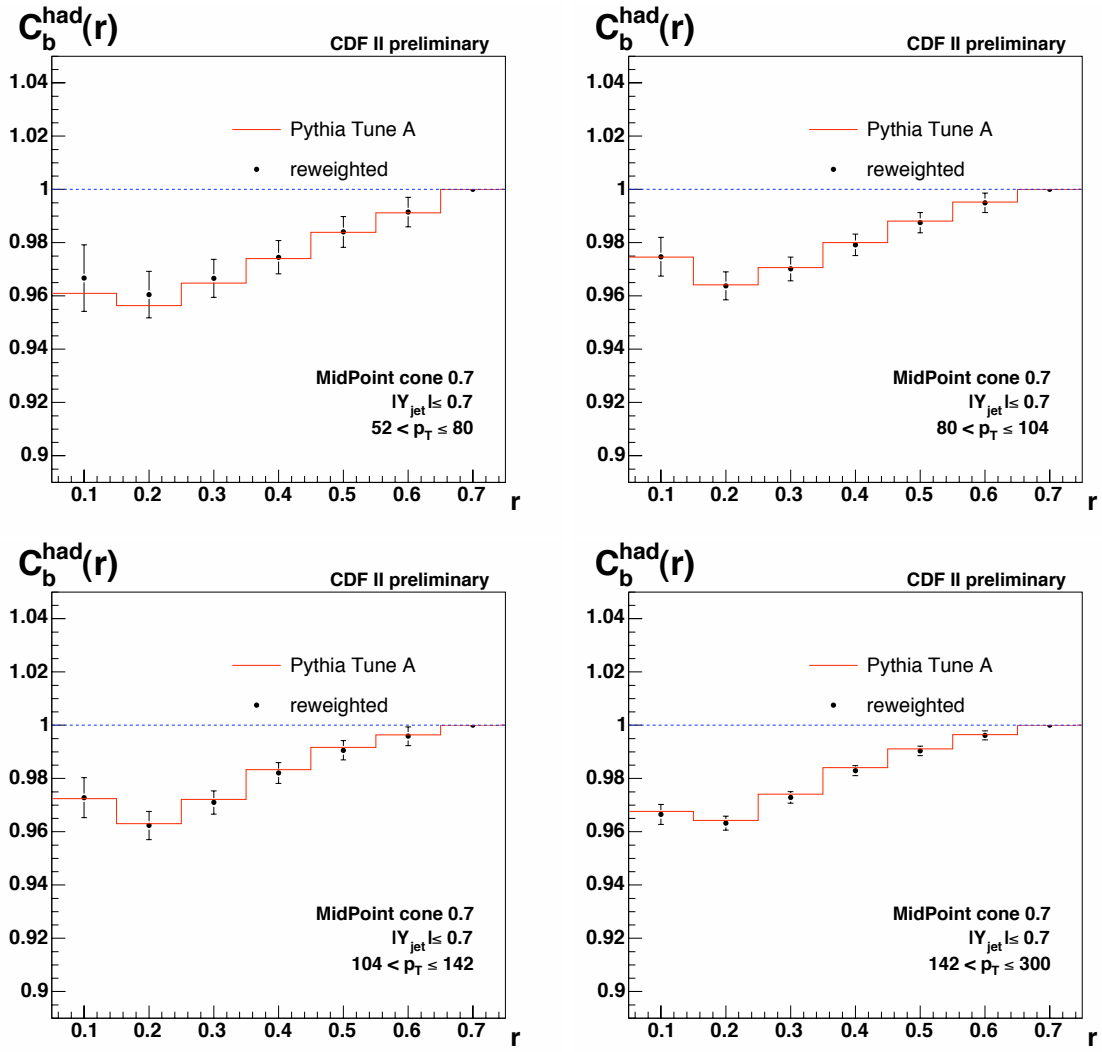

FIGURE 4.33: Comparison of the b-jet hadron level corrections between the default Pythia Tune A MC shapes (red lines) and the reweighted ones (black points). Only the MC statistical errors are shown.

4.12 Total Systematic Errors

The hadron level b-jet shapes are computed from equation 3.7 for the default case as well as for each of the systematic variations mentioned in the previous sections. The systematic variations considered for this analysis are

- jet p_T bins $\pm 3\%$
- missing E_T significance cut $\pm 15\%$
- $|Z_{\text{vtx}}|$ cut $\pm 5\,\text{cm}$
- $f_{1\text{c}} \pm 20\%$ (absolute)
- $f_{\text{c}} \pm 5\%$ (absolute)
- jet shapes measured using tracks
- jet shapes measured using calorimeter towers with $p_T > 0.5\,\text{GeV}$

The total, statistical and systematic uncertainties are shown in table 4.4 for each p_T bin and r bin. Also shown are the various contributions from the individual effects.

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

- $f_{1\text{c}}$ variation,
- the difference between the b-jet shapes reconstructed from tracks and from calorimeter towers,
- MC statistics.

Despite the fact that the limited amount of statistics in the MC samples leads to a significantly large errors on the hadron level b-quark jets, particularly for the nonb-jet tagging biases and the nonb-jet secondary vertex mass templates, this effect is never the largest systematic one. Reducing the error due to the MC statistics will not significantly reduce the total systematic error.

The systematic uncertainties were also computed in an alternative manner [14]. The total systematic and statistical errors were computed for each of the parameters individually. These errors were then combined to get the total systematic and statistical errors on the hadron level b-quark jet shapes. In this different approach, the systematic uncertainty calculation is assuming no correlations between the different parameters but is at risk of double-counting some of the effects. The total systematic errors obtained using this method were found to be very similar to the ones obtained using the default method. They turn out to be very slightly higher, by a maximum of 15%, than the ones obtained using the approach detailed above.

$52 - 80\,\mathrm{GeV}$											
				Individual systematic errors							
r bin	tot	stat tot	syst tot	MC stats	$p_T \pm 3\%$	$\lvert Z_{\mathrm{vtx}}\rvert \pm 5\,\mathrm{cm}$	$\not\!\!E_T$ sig $\pm 15\%$	$f_{1\mathrm{c}} \pm 20\%$	$f_{\mathrm{c}} \pm 5\%$	tracks	tower $p_T > 0.5\,\mathrm{GeV}$
0.1	0.071	0.011	0.070	0.035	0.011	0.001	0.000	0.032	0.008	0.050	0.002
0.2	0.054	0.010	0.053	0.034	0.001	0.002	0.000	0.027	0.007	0.029	0.005
0.3	0.044	0.007	0.043	0.025	0.006	0.001	0.001	0.018	0.005	0.028	0.006
0.4	0.026	0.005	0.026	0.016	0.002	0.001	0.000	0.008	0.004	0.017	0.006
0.5	0.013	0.003	0.013	0.010	0.000	0.000	0.000	0.004	0.002	0.003	0.006
0.6	0.008	0.002	0.007	0.004	0.001	0.000	0.000	0.000	0.001	0.005	0.002

$80 - 104\,\mathrm{GeV}$											
				Individual systematic errors							
r bin	tot	stat tot	syst tot	MC stats	$p_T \pm 3\%$	$\lvert Z_{\mathrm{vtx}}\rvert \pm 5\,\mathrm{cm}$	$\not\!\!E_T$ sig $\pm 15\%$	$f_{1\mathrm{c}} \pm 20\%$	$f_{\mathrm{c}} \pm 5\%$	tracks	tower $p_T > 0.5\,\mathrm{GeV}$
0.1	0.038	0.006	0.038	0.015	0.007	0.001	0.000	0.033	0.006	0.005	0.001
0.2	0.029	0.006	0.028	0.013	0.012	0.002	0.000	0.019	0.004	0.011	0.002
0.3	0.023	0.004	0.023	0.009	0.009	0.001	0.000	0.010	0.003	0.015	0.003
0.4	0.017	0.003	0.017	0.006	0.003	0.000	0.000	0.005	0.002	0.014	0.003
0.5	0.013	0.002	0.012	0.003	0.003	0.000	0.000	0.003	0.001	0.011	0.003
0.6	0.006	0.001	0.006	0.002	0.002	0.000	0.000	0.001	0.000	0.005	0.002

$104 - 142\,\mathrm{GeV}$											
				Individual systematic errors							
r bin	tot	stat tot	syst tot	MC stats	$p_T \pm 3\%$	$\lvert Z_{\mathrm{vtx}}\rvert \pm 5\,\mathrm{cm}$	$\not\!\!E_T$ sig $\pm 15\%$	$f_{1\mathrm{c}} \pm 20\%$	$f_{\mathrm{c}} \pm 5\%$	tracks	tower $p_T > 0.5\,\mathrm{GeV}$
0.1	0.054	0.008	0.053	0.025	0.014	0.003	0.001	0.030	0.005	0.033	0.000
0.2	0.029	0.007	0.028	0.021	0.000	0.001	0.001	0.015	0.003	0.010	0.000
0.3	0.017	0.005	0.016	0.014	0.004	0.000	0.000	0.006	0.003	0.002	0.001
0.4	0.015	0.003	0.015	0.009	0.003	0.001	0.000	0.002	0.002	0.011	0.001
0.5	0.011	0.002	0.011	0.005	0.001	0.000	0.000	0.001	0.001	0.009	0.001
0.6	0.004	0.001	0.004	0.002	0.001	0.000	0.000	0.000	0.001	0.003	0.001

$142 - 300\,\mathrm{GeV}$											
				Individual systematic errors							
r bin	tot	stat tot	syst tot	MC stats	$p_T \pm 3\%$	$\lvert Z_{\mathrm{vtx}}\rvert \pm 5\,\mathrm{cm}$	$\not\!\!E_T$ sig $\pm 15\%$	$f_{1\mathrm{c}} \pm 20\%$	$f_{\mathrm{c}} \pm 5\%$	tracks	tower $p_T > 0.5\,\mathrm{GeV}$
0.1	0.033	0.008	0.032	0.020	0.007	0.000	0.001	0.018	0.003	0.016	0.000
0.2	0.023	0.006	0.023	0.016	0.009	0.001	0.001	0.007	0.002	0.010	0.000
0.3	0.015	0.004	0.014	0.011	0.007	0.000	0.000	0.002	0.004	0.002	0.001
0.4	0.015	0.003	0.014	0.007	0.002	0.000	0.000	0.001	0.003	0.012	0.001
0.5	0.010	0.002	0.010	0.004	0.000	0.000	0.000	0.000	0.001	0.009	0.002
0.6	0.010	0.001	0.010	0.002	0.000	0.000	0.000	0.000	0.001	0.009	0.002

TABLE 4.4: Total, statistical and systematic errors on the final b-jet shapes along with the contributions from each parameter. The last bins in r always have by definition zero error so they are not displayed.

4.13 Changing the Rapidity Cut to $0.1 \leq |Y| \leq 0.7$

In order to compare the hadron level b-jet shapes to the previously published inclusive jet shapes, it is important to see if the different cut requirements on the rapidity of the jets have an effect on the measured shapes. Figure 4.34 shows a comparison, in data, between the inclusive shapes using a rapidity cut on the jet of $|Y| \leq 0.7$ (black curve) and those measured removing the central rapidity region, i.e. requiring $0.1 \leq |Y| \leq 0.7$ (red curve). Figure 4.35 shows the ratio of the shapes removing the central rapidity region over the shapes including the central rapidity region. Figure 4.36 shows a comparison between the shapes for tagged jets using the two different rapidity cuts. Figure 4.37 shows the ratio of the shapes removing the central rapidity region over the shapes including the central rapidity region. Both these sets of figures show that the measured shapes are very similar for both cases; they are within less than 1% of eachother. This allows the published inclusive jet shape results to be plotted on the same figures as the b-jet shape results.

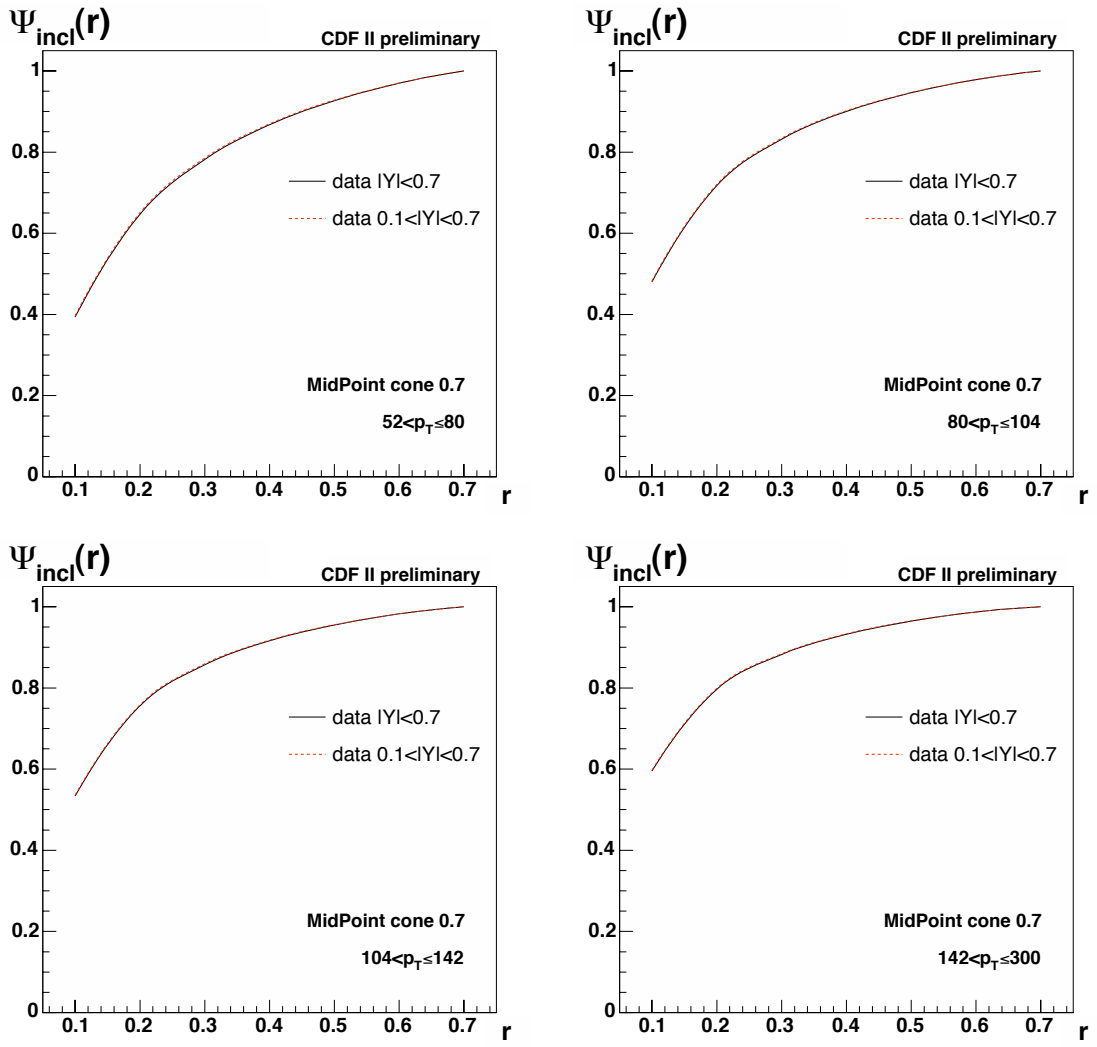

FIGURE 4.34: Comparison in data between the inclusive shapes with the rapidity cut $|Y| \leq 0.7$ (black curve) and the cuts used for the measurement of the inclusive jet shapes, $0.1 \leq |Y| \leq 0.7$ (red curve).

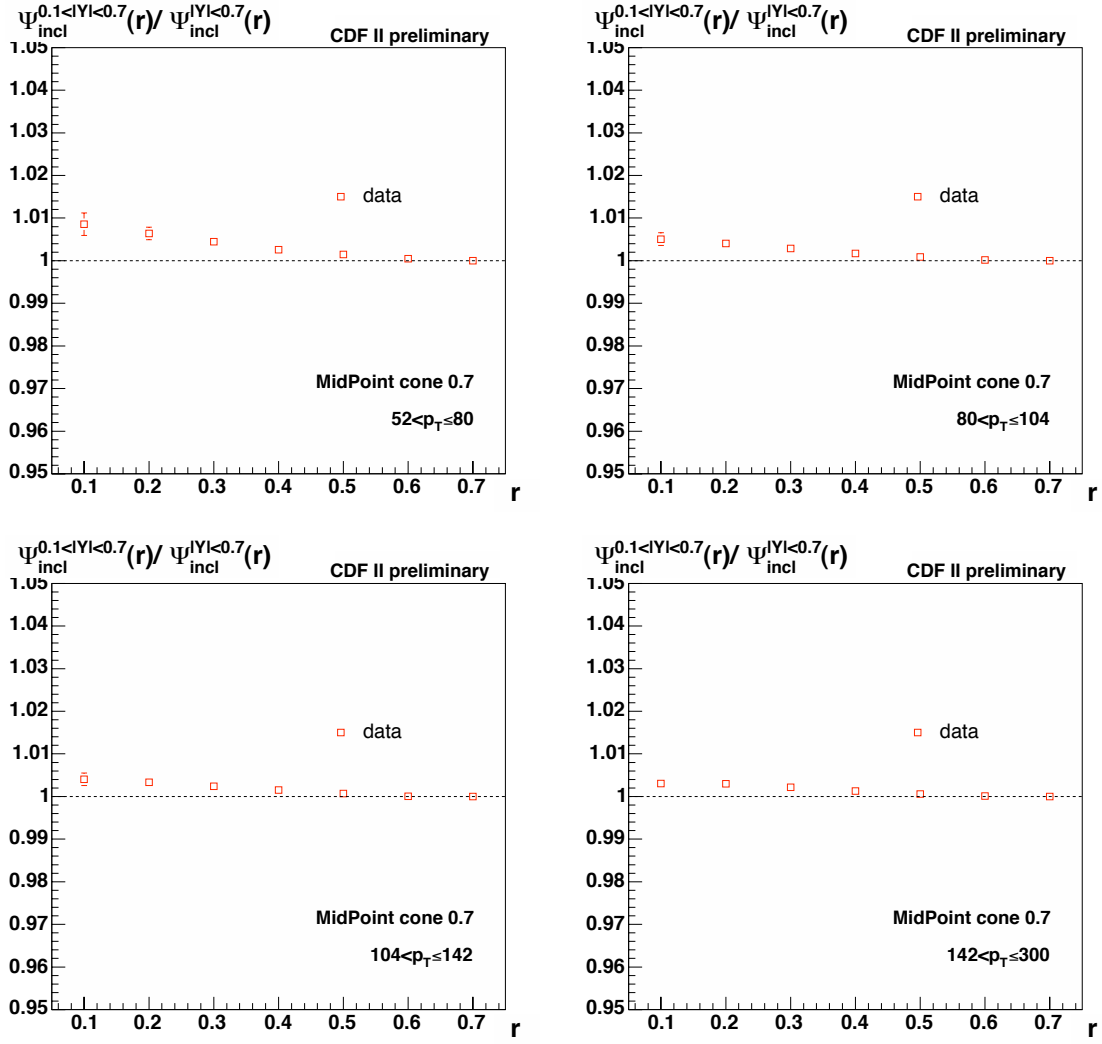

FIGURE 4.35: Ratio in data between the inclusive shapes with the rapidity cut $0.1 \leq |Y| \leq 0.7$ over the shapes with $|Y| \leq 0.7$ (red points). Only the statistical errors of the measured shapes with $0.1 \leq |Y| \leq 0.7$ are shown.

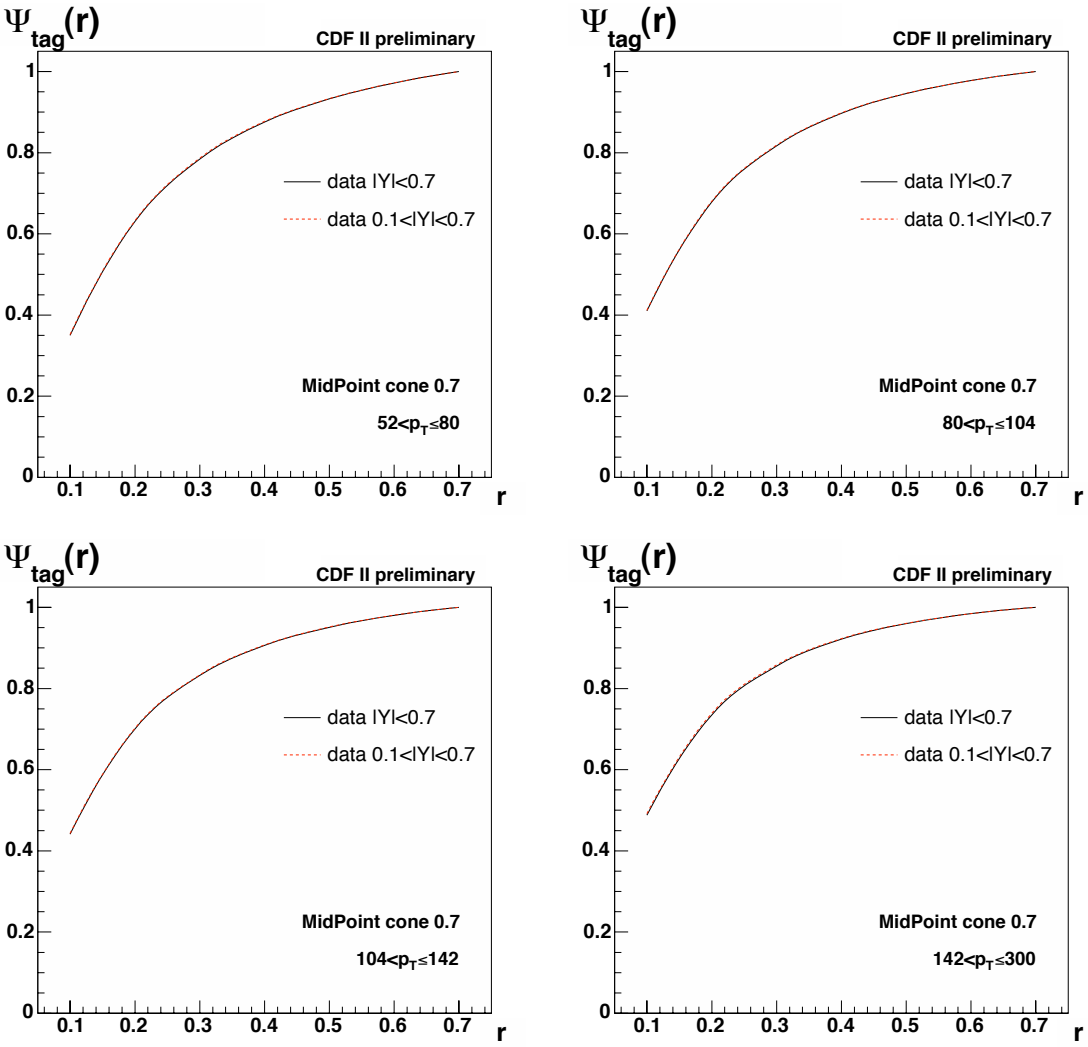

FIGURE 4.36: Comparison in data between tagged jet shapes with the rapidity cut $|Y| \leq 0.7$ (black curve) and the cuts used for the measurement of the inclusive jet shapes, $0.1 \leq |Y| \leq 0.7$ (red curve).

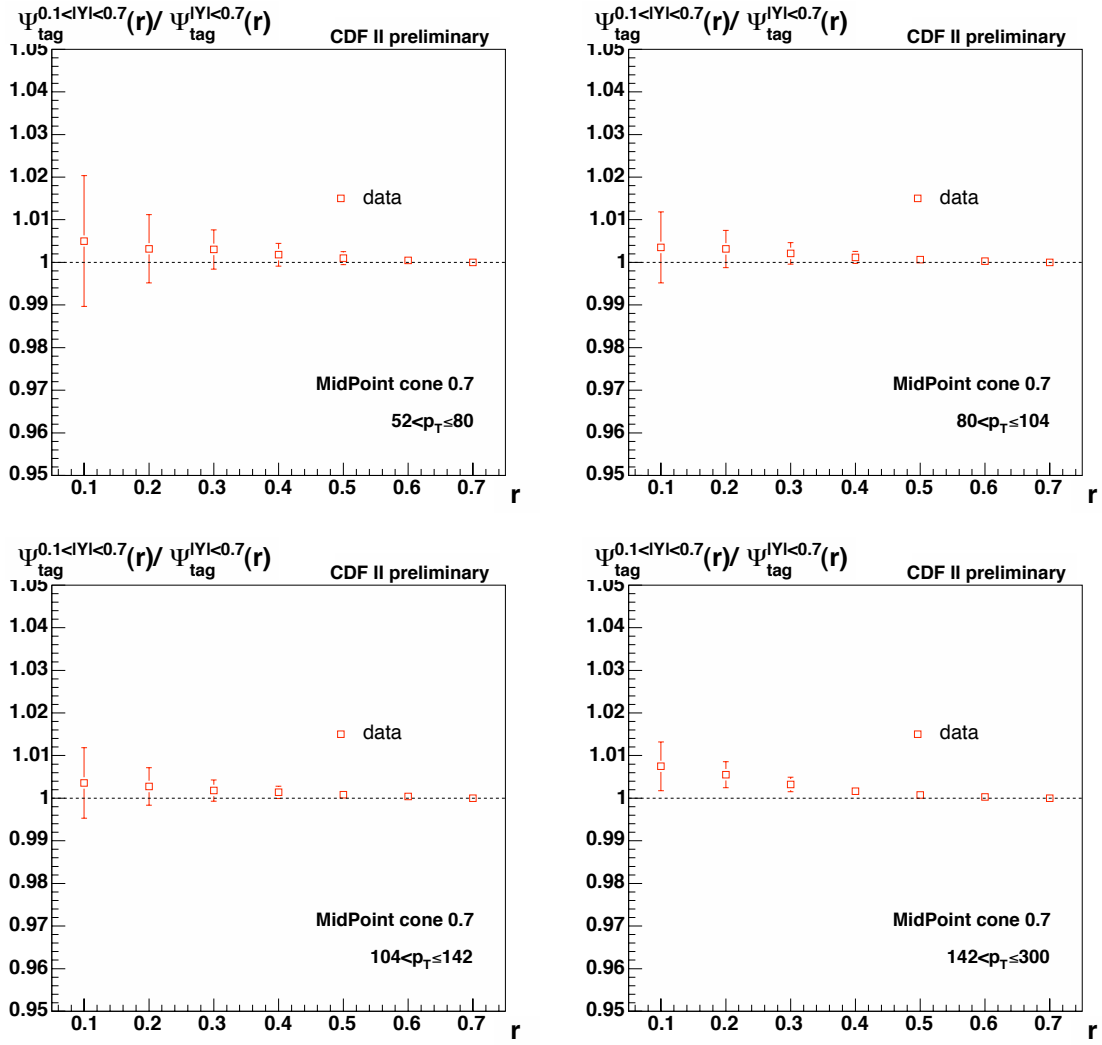

FIGURE 4.37: Ratio in data between the tagged jet shapes with the rapidity cut $0.1 \leq |Y| \leq 0.7$ over the shapes with $|Y| \leq 0.7$ (red points). Only the statistical errors of the measured shapes with $0.1 \leq |Y| \leq 0.7$ are shown.

Chapter 5

Results

Combining the measurements of all the different terms of equation 3.7, the b-jet shapes can be calculated. Section 5.1 shows that it is possible, starting from the Pythia Tune A MC predictions for the inclusive and tagged jet shapes, to reconstruct the detector level and hadron level b-quark jet shapes predicted by the same MC. This is done for shapes measured using calorimeter towers and for those measured using tracks. Section 5.2 presents the final results for the integrated jet shapes, including all the systematic errors. Then, in section 5.3, the same results are shown but this time the fractional energy outside a fixed cone radius is plotted as a function of the p_T of the jet. These results are compared to the inclusive jet shape results and the Pythia Tune A predictions for inclusive and b-jets. Also shown is a comparison with the shapes expected from Pythia Tune A for single b-quark jets and double b-quark jets. Finally, section 5.4 shows the results of a fit to the data in order to determine the most likely $f_{1\mathrm{b}}$ fraction. This shows that the Pythia Tune A MC predictions are between 10% and 20% higher (in absolute terms) than the values found in data.

5.1 MC Reconstruction Checks

In this section, the whole analysis is carried out starting from the raw calorimeter and track shapes as measured when taking Pythia Tune A MC predictions instead of the data. This is to test that the reconstruction method actually returns the correct detector level and hadron level shapes.

Figure 5.1 shows the comparison between the detector level b-quark jet shapes and the reconstructed shapes starting from inclusive and tagged jets. In this plot, the "true" MC purity is used instead of the fraction obtained from the secondary vertex mass fit to the inclusive MC distribution. Moreover, all the tagging bias terms, for b- and for nonb-jets, are taken from the b-filtered Pythia Tune A MC. The inclusive shapes are still taken from the inclusive Pythia Tune A MC but the tagged shapes are taken from the b-filtered Pythia Tune A MC. This means that the whole procedure, except for the values of the inclusive shapes, is taken from the same MC samples as used for the comparison. Figure 5.2 shows the same plots but starting from the shapes measured using tracks. Figures 5.3 and 5.4 show the shapes unfolded back to hadron level compared to the hadron level b-jet shapes when starting from calorimeter towers and tracks, respectively. All these sets of plots show basically perfect agreement between data and MC [1].

Figures 5.5, 5.7 and 5.9 show the comparison between the unfolded Pythia Tune A MC and the hadron level predictions of Pythia Tune A when starting from raw calorimetric towers, tracks

[1] The very small differences can be attributed to the rounding of the values for each individual parameter to floats.

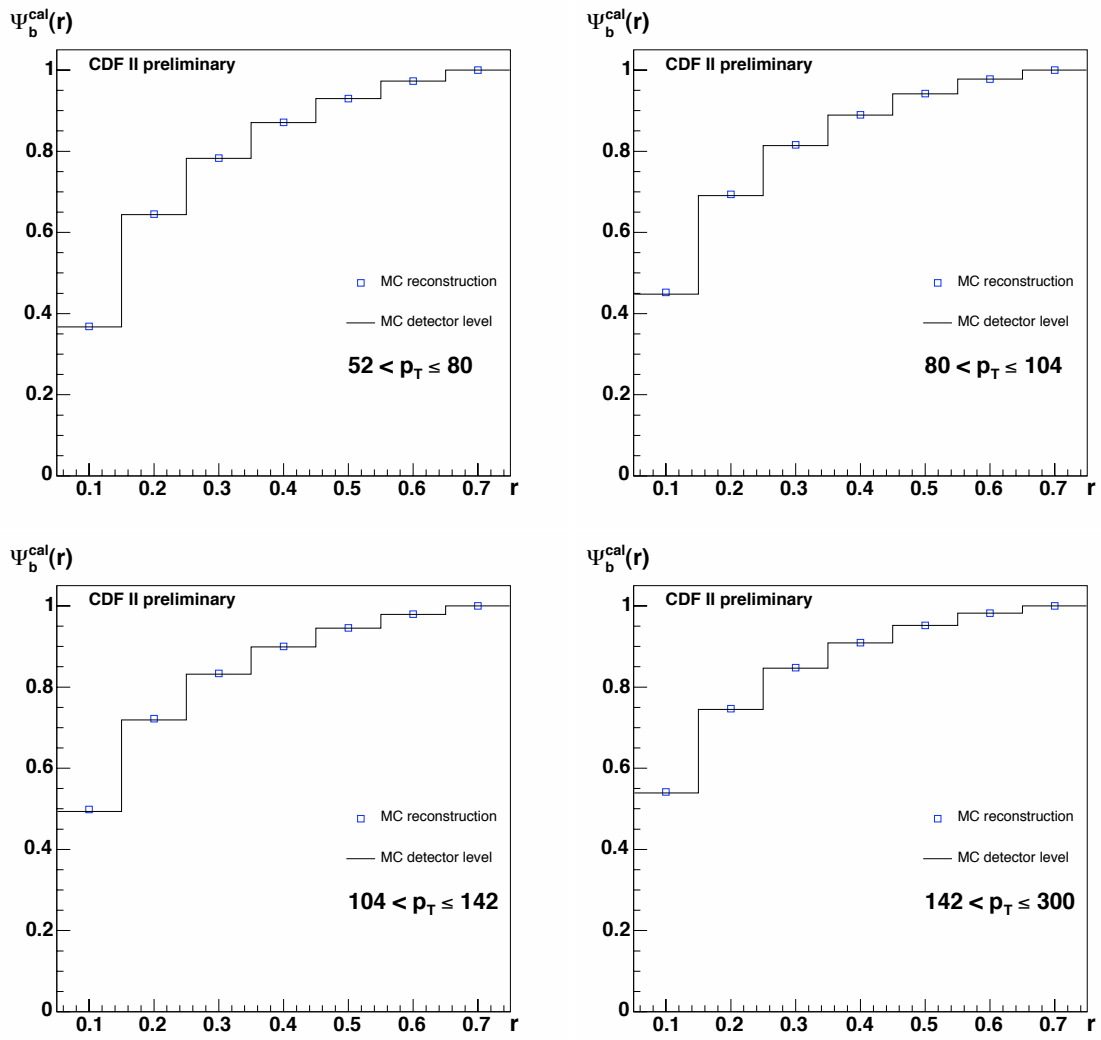

FIGURE 5.1: Detector level integrated b-jet shapes using calorimeter towers for each of the p_T bins considered. The reconstructed detector level MC shapes are shown as blue points. The Pythia Tune A MC predictions for b-jets at calorimeter level are shown as black lines.

and calorimeter towers with $p_T > 0.5\,\text{GeV}$, respectively. The differences with respect to the previous plots is that the parameters are taken from the relevant Pythia Tune A MC samples and not all from the b-filtered samples. The purities are also taken from the secondary vertex mass fit results when fitting the templates to the MC distributions for tagged jets. Figures 5.6, 5.8 and 5.10 show the ratio of the reconstructed over the "true" hadron level b-quark jet shapes for the shapes reconstructed using the calorimeter towers, the tracks and the calorimeter towers with $p_T > 0.5\,\text{GeV}$, respectively. The shapes are in good agreement with the hadron level predictions.

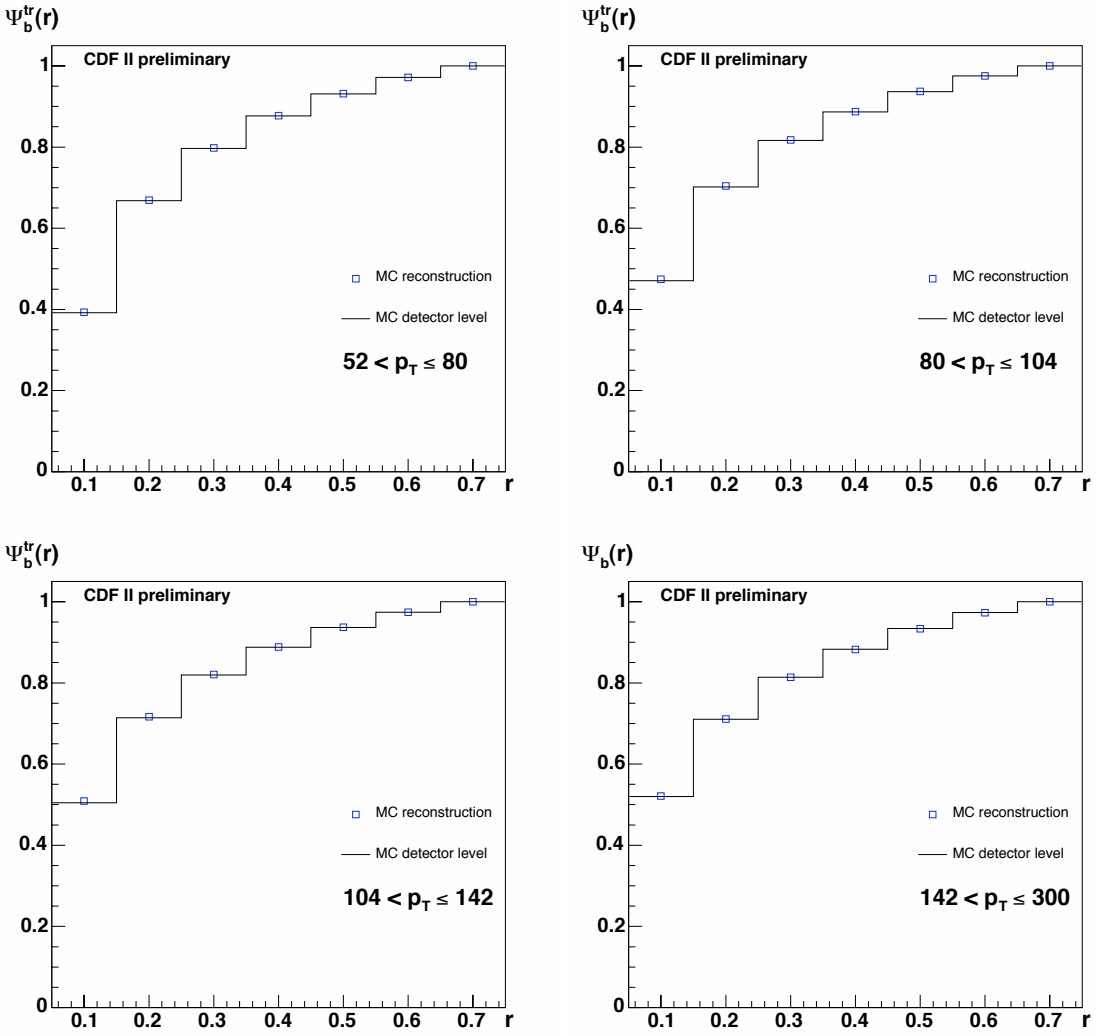

FIGURE 5.2: Detector level integrated b-jet shapes using tracks for each of the p_T bins considered. The reconstructed detector level MC shapes are shown as blue points. The MC predictions for b-jets at track level are shown as black lines.

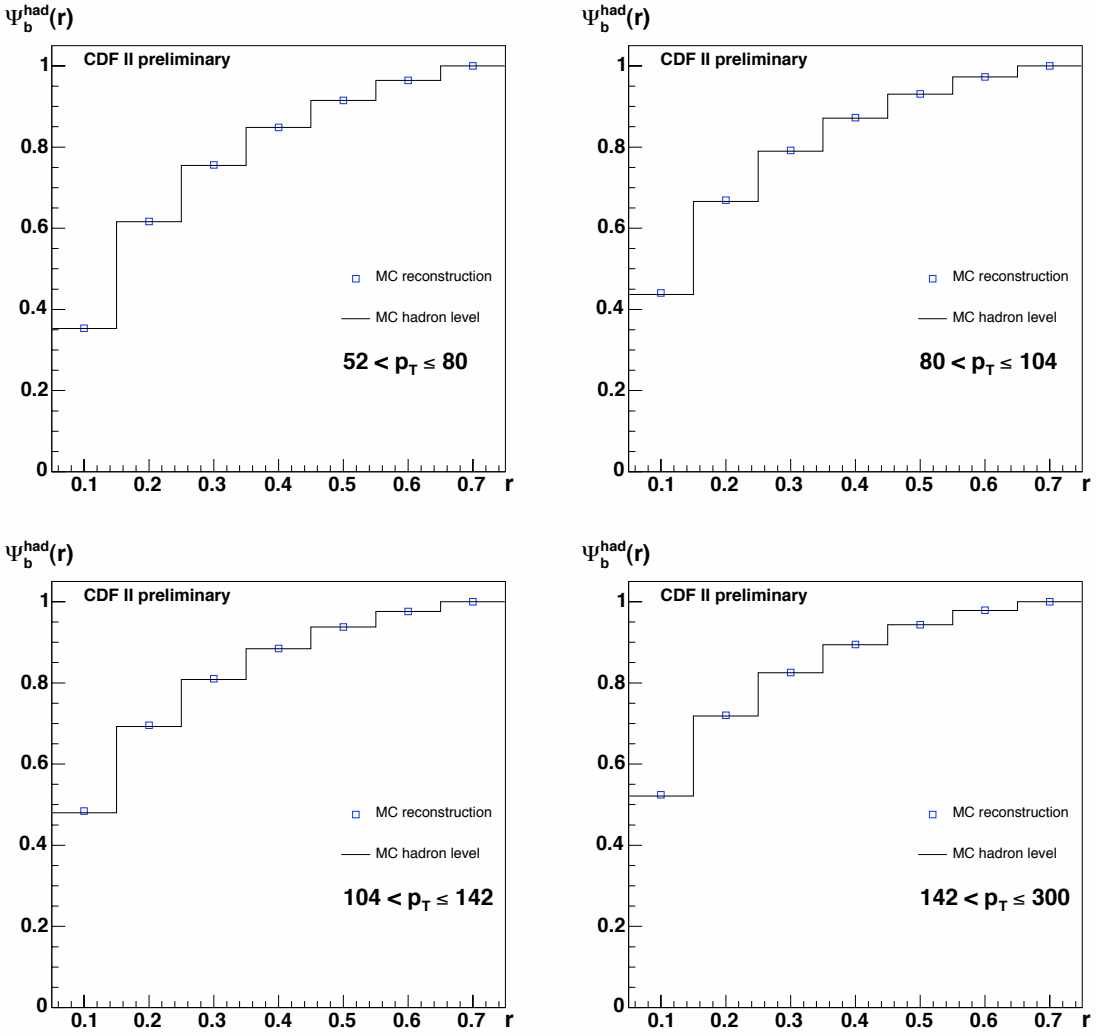

FIGURE 5.3: Hadron level integrated b-jet shapes for each of the p_T bins considered. The reconstructed hadron level MC shapes are shown as black points. The MC predictions for b-jets at hadron level are shown as black lines.

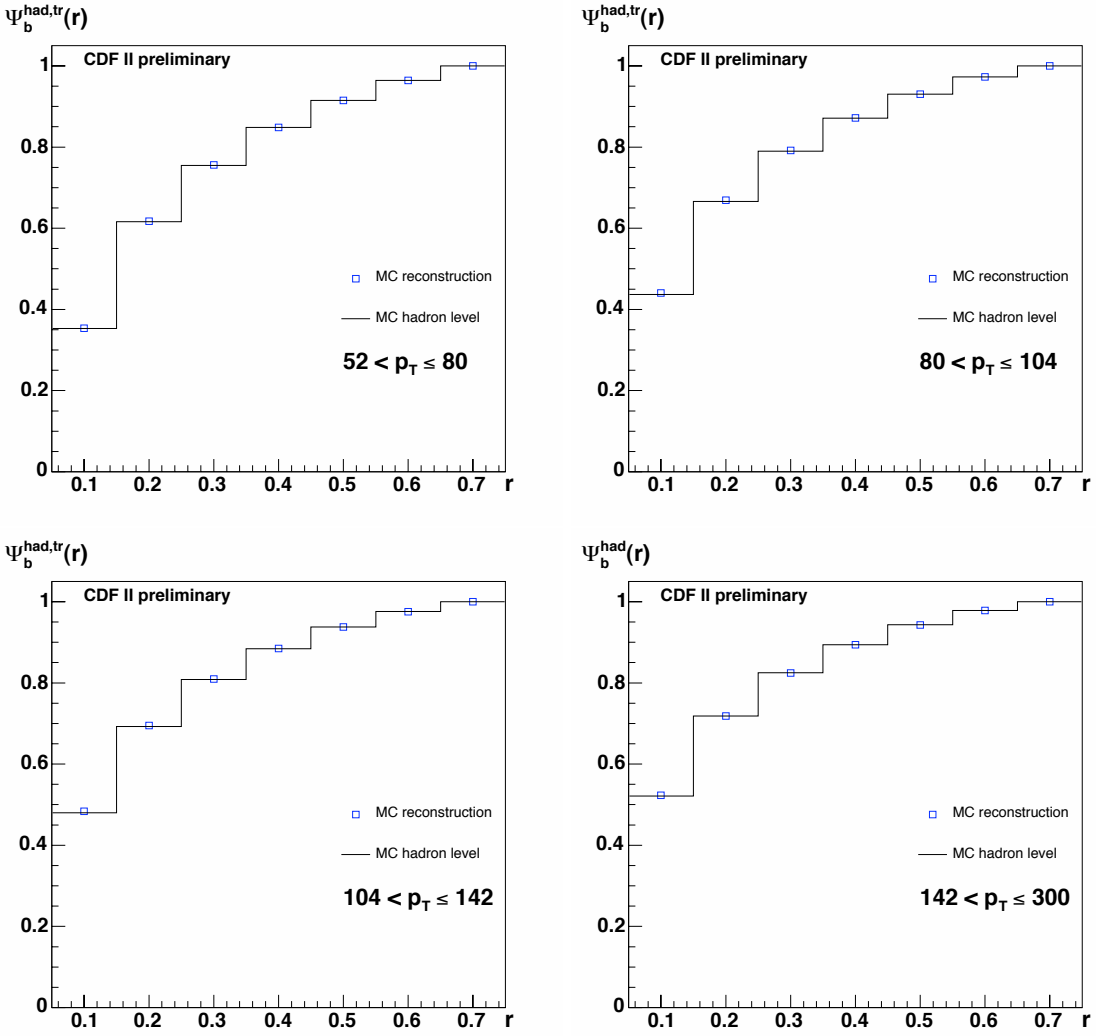

FIGURE 5.4: Hadron level integrated b-jet shapes for each of the p_T bins considered. The reconstructed hadron level MC shapes, starting from tracks, are shown as black points. The MC predictions for b-jets at hadron level are shown as black lines.

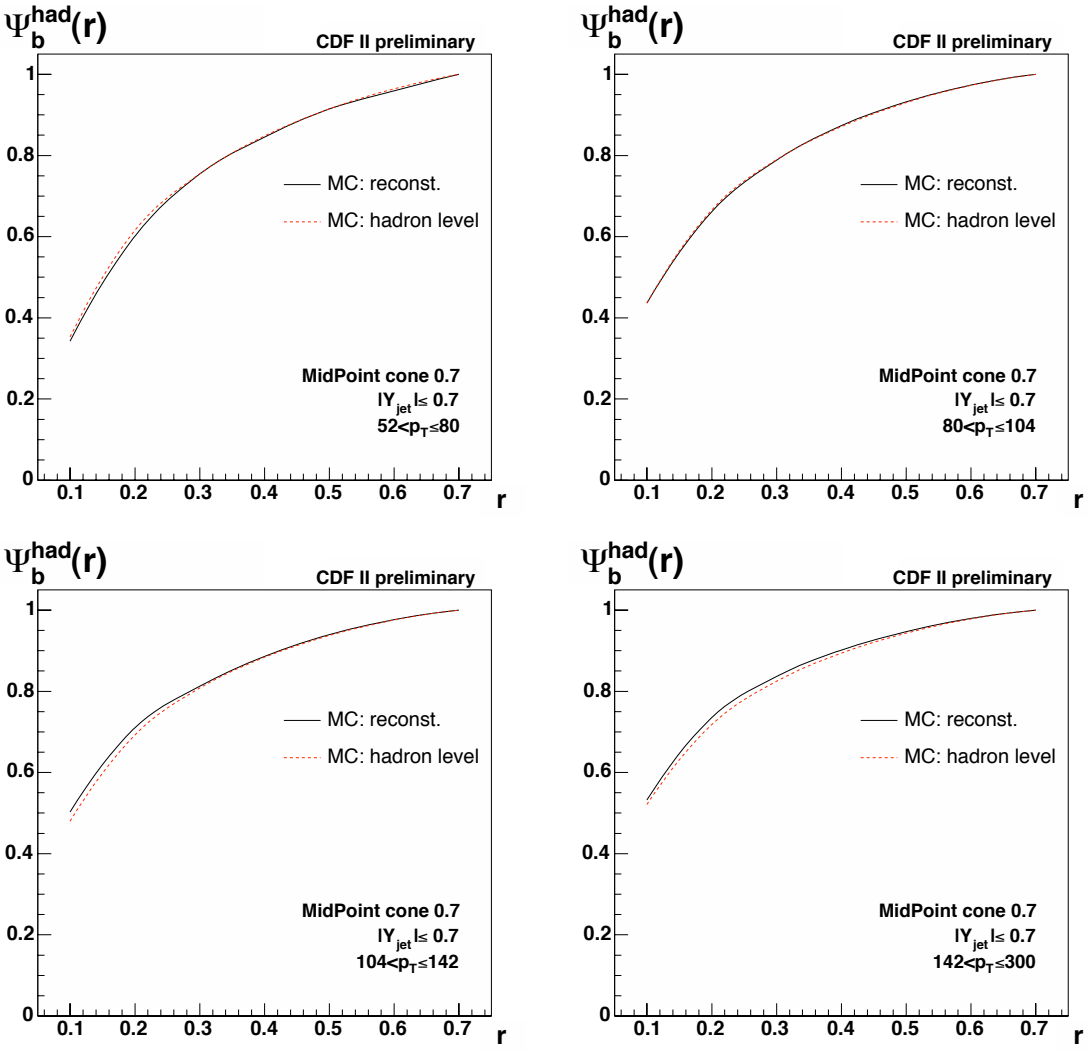

FIGURE 5.5: Hadron level integrated b-jet shapes in Pythia Tune A MC for each of the p_T bins considered. The reconstructed hadron level MC shapes are shown as black lines. The MC predictions for b-jets at hadron level are shown as red dashed curves.

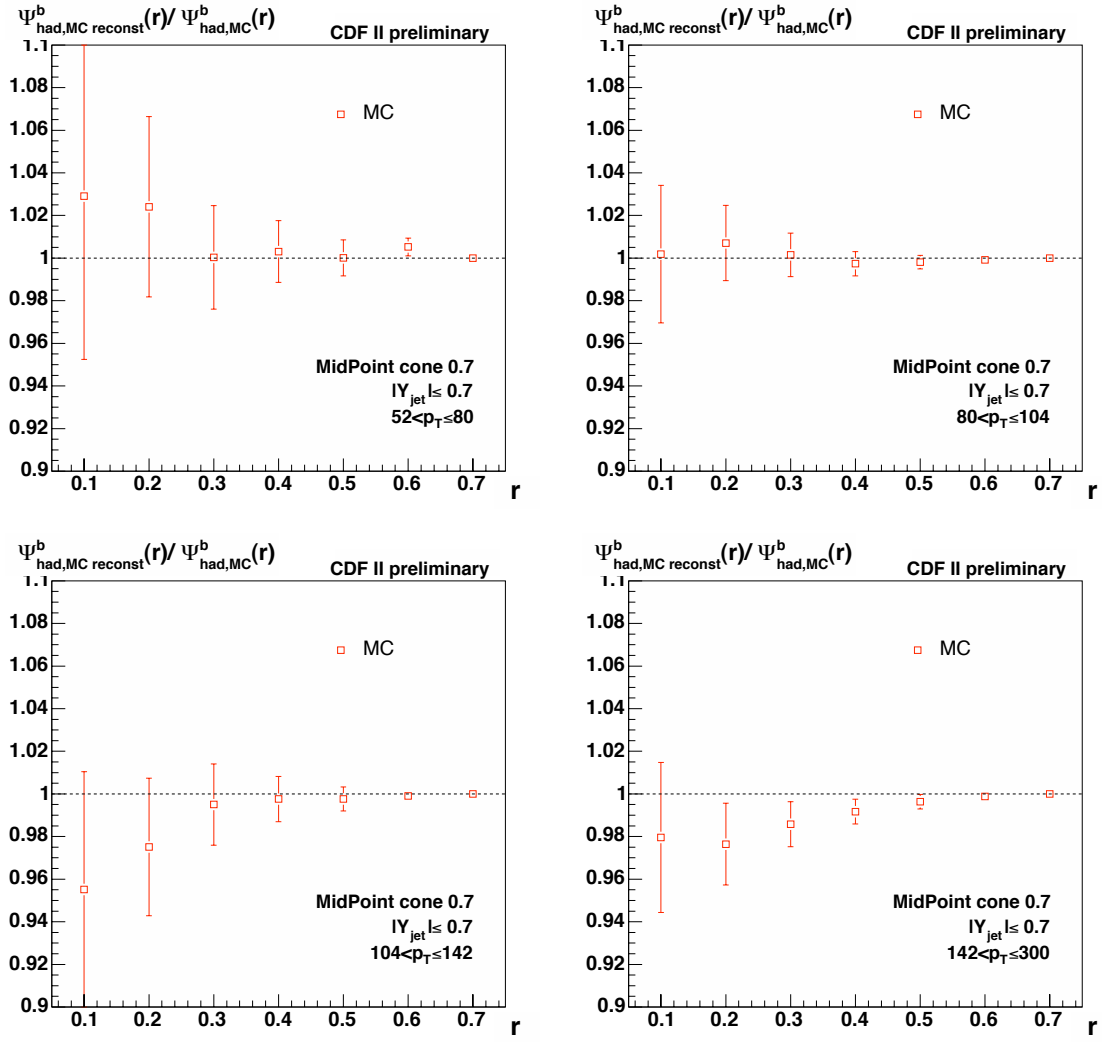

FIGURE 5.6: Ratio of the reconstructed hadron level Pythia Tune A MC shapes over the "true" hadron level Pythia Tune A MC shapes. The errors shown are only the statistical errors of the reconstructed shapes.

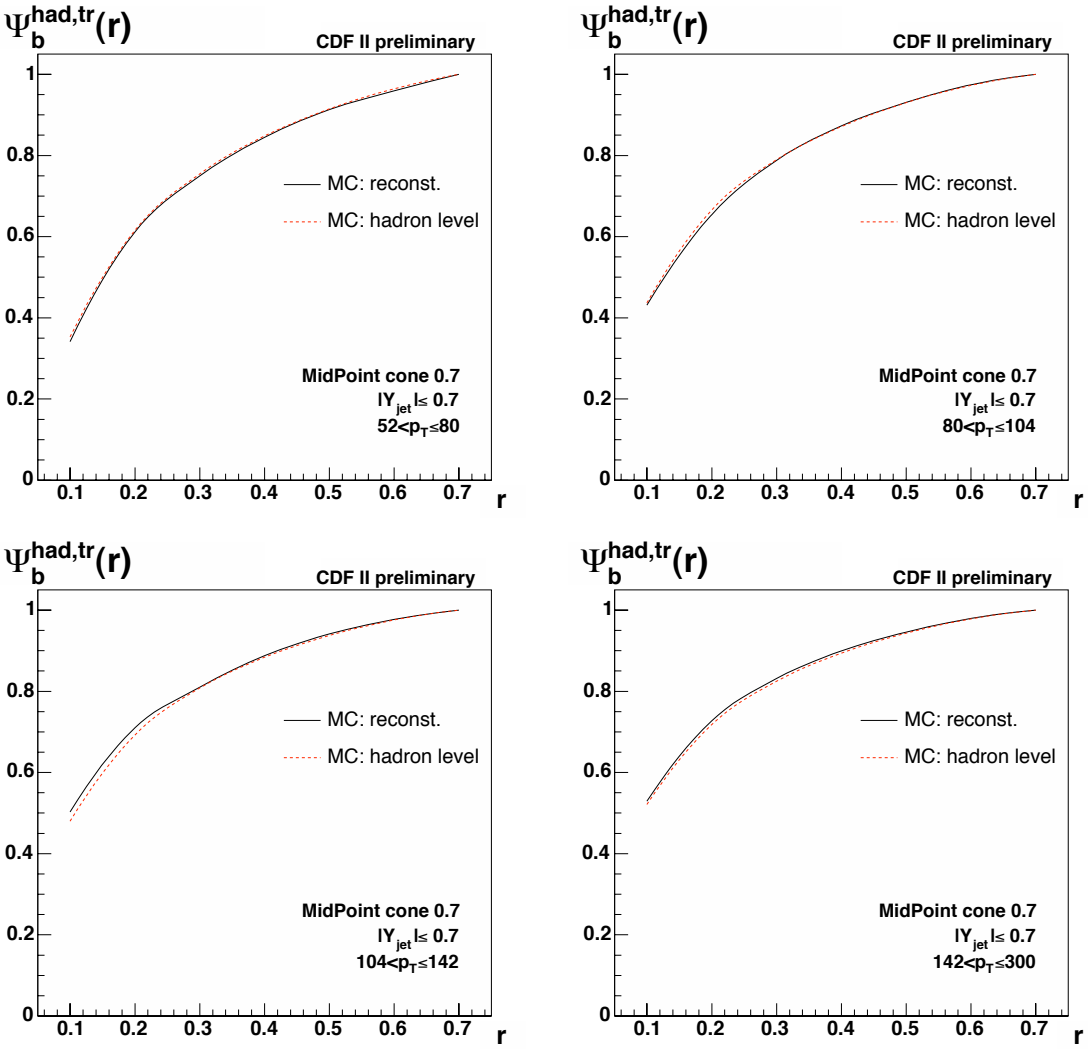

FIGURE 5.7: Hadron level integrated b-jet shapes in Pythia Tune A MC for each of the p_T bins considered. The reconstructed hadron level MC shapes, starting from tracks, are shown as black lines. The MC predictions for b-jets at hadron level are shown as red dashed curves.

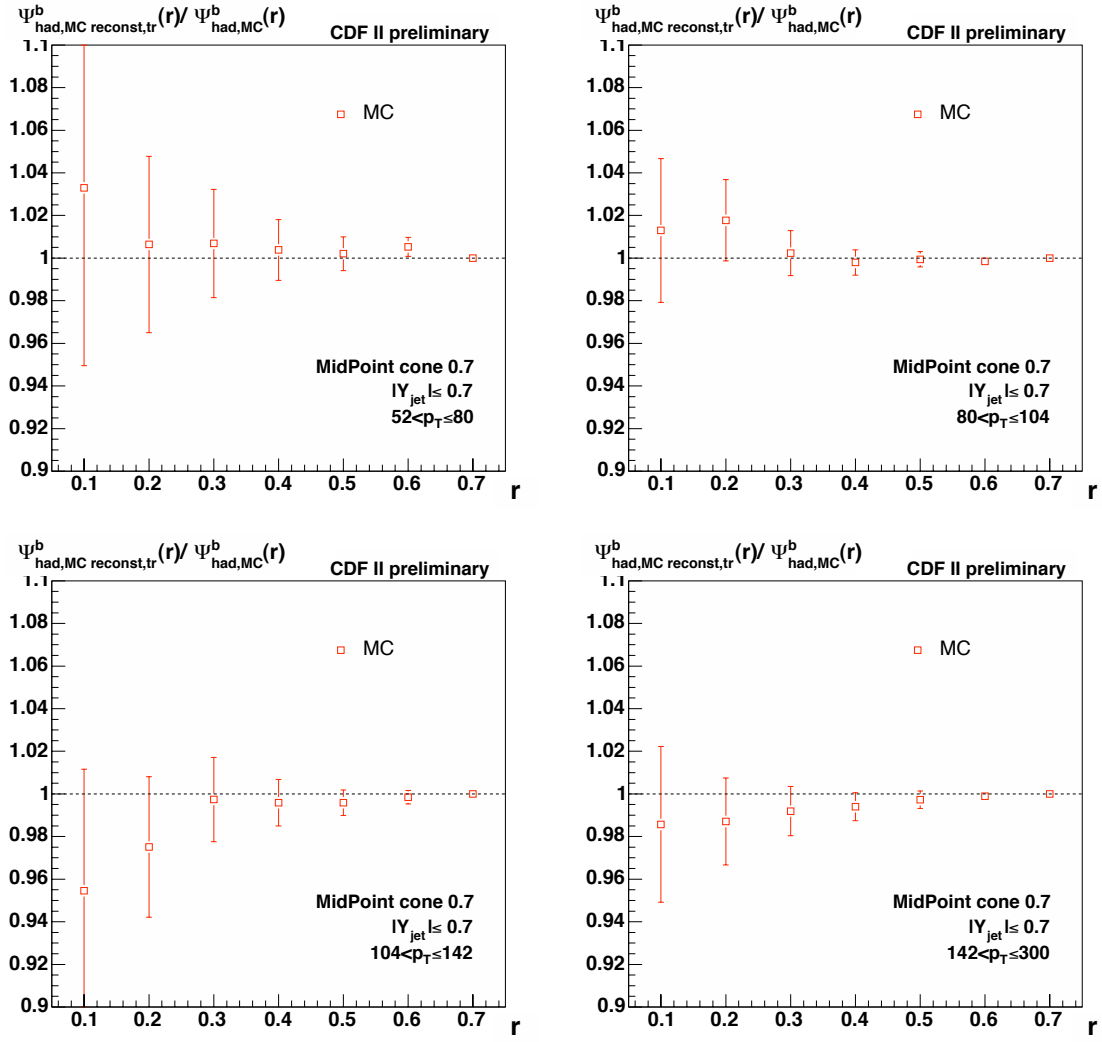

FIGURE 5.8: Ratio of the reconstructed hadron level Pythia Tune A MC shapes, starting from raw track shapes, over the "true" hadron level Pythia Tune A MC shapes. The errors shown are only the statistical errors of the reconstructed shapes.

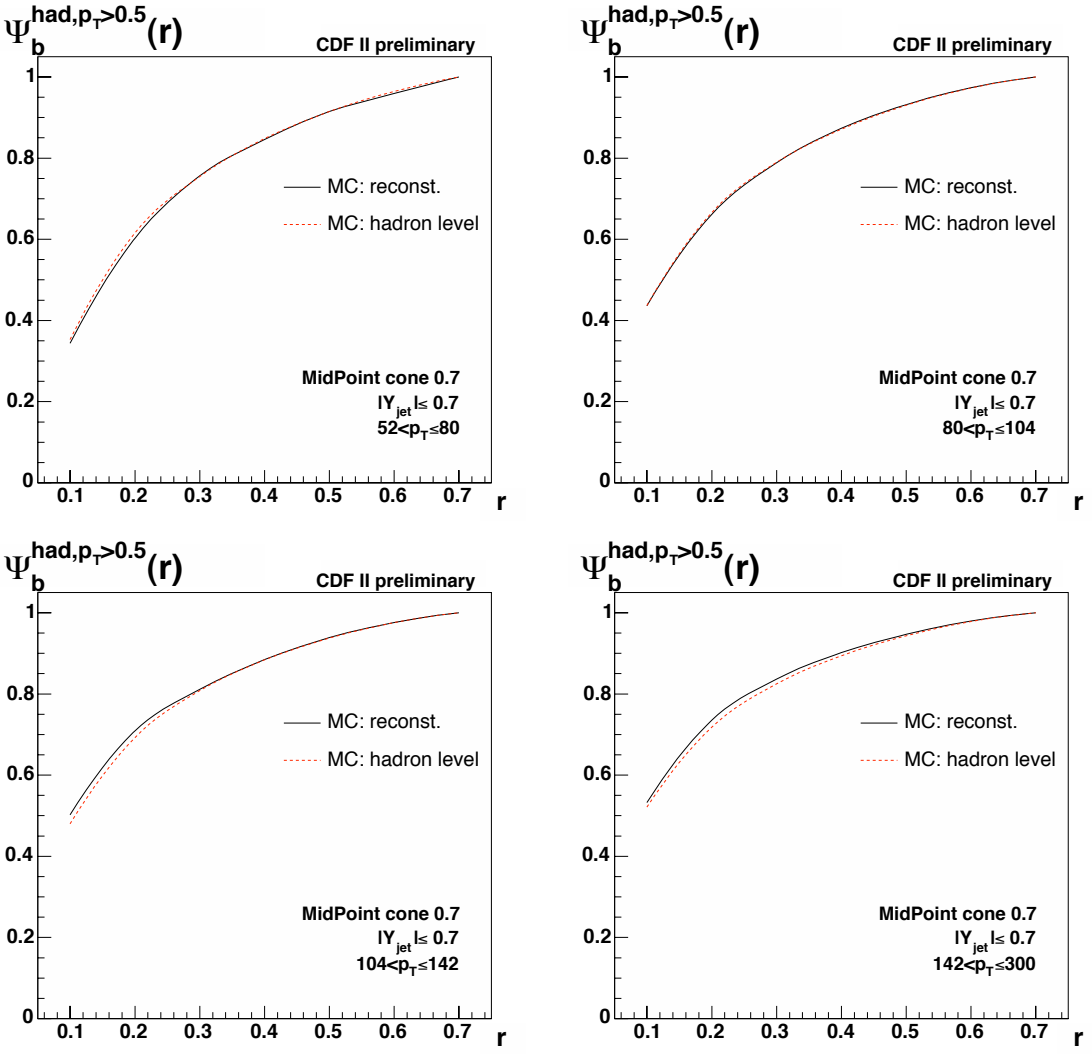

FIGURE 5.9: Hadron level integrated b-jet shapes in Pythia Tune A MC for each of the p_T bins considered. The reconstructed hadron level MC shapes, starting from calorimeter towers with p_T larger than 0.5 GeV, are shown as black lines. The MC predictions for b-jets at hadron level are shown as red dashed curves.

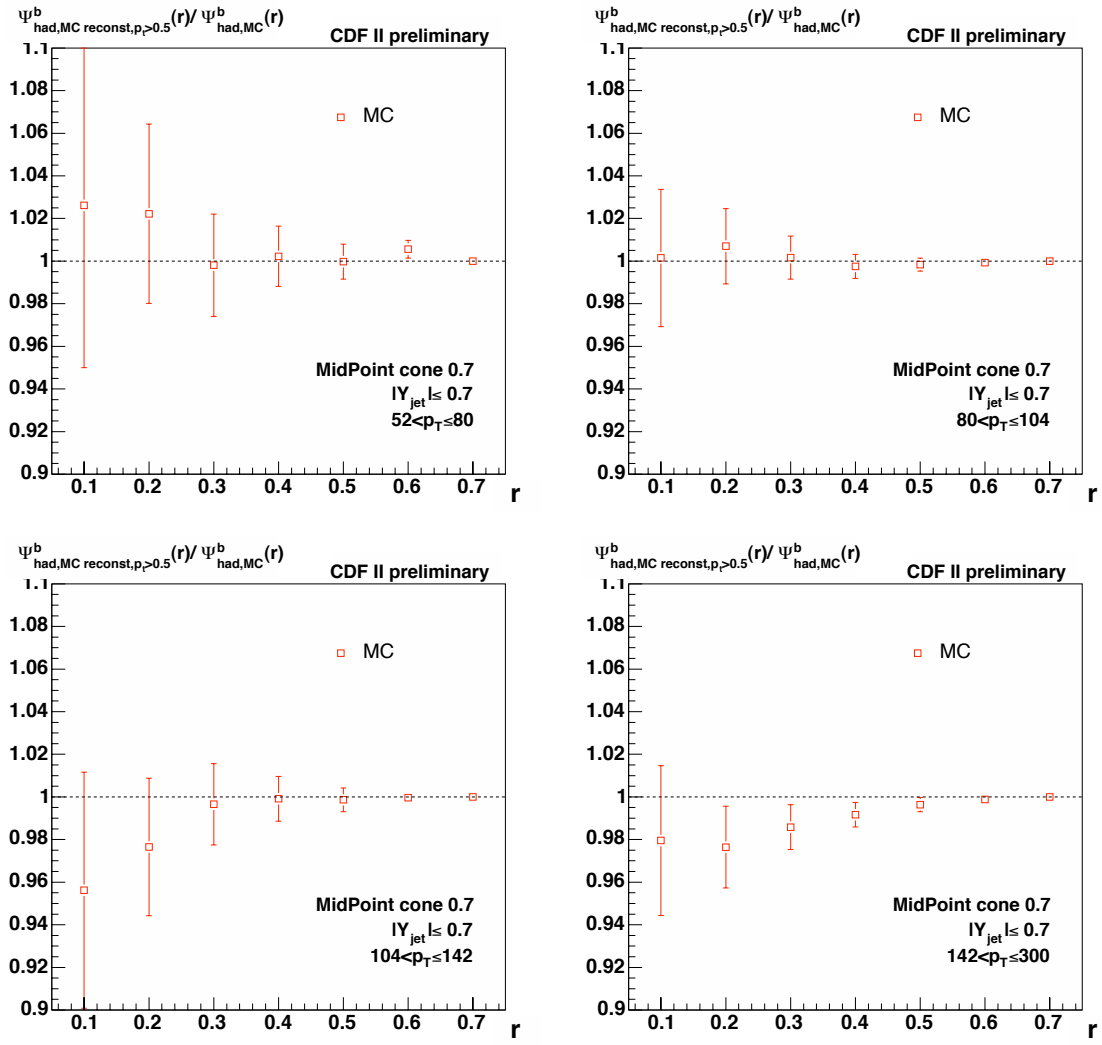

FIGURE 5.10: Ratio of the reconstructed hadron level Pythia Tune A MC shapes, starting from raw shapes reconstructed using calorimeter towers with $p_T > 0.5\,\text{GeV}$, over the "true" hadron level Pythia Tune A MC shapes. The errors shown are only the statistical errors of the reconstructed shapes.

5.2 Integrated b-jet Shapes

The hadron level integrated b-jet shapes are shown in figure 5.11, where both the statistical and the systematic errors are shown. The statistical errors are very small with respect to the systematic error, smaller than the points on the plots. Also shown in this plot are the MC Pythia Tune A predictions for the inclusive and the b-jet shapes. There are significant deviations from the MC predictions of the b-jet shapes for all but the first p_T bin.

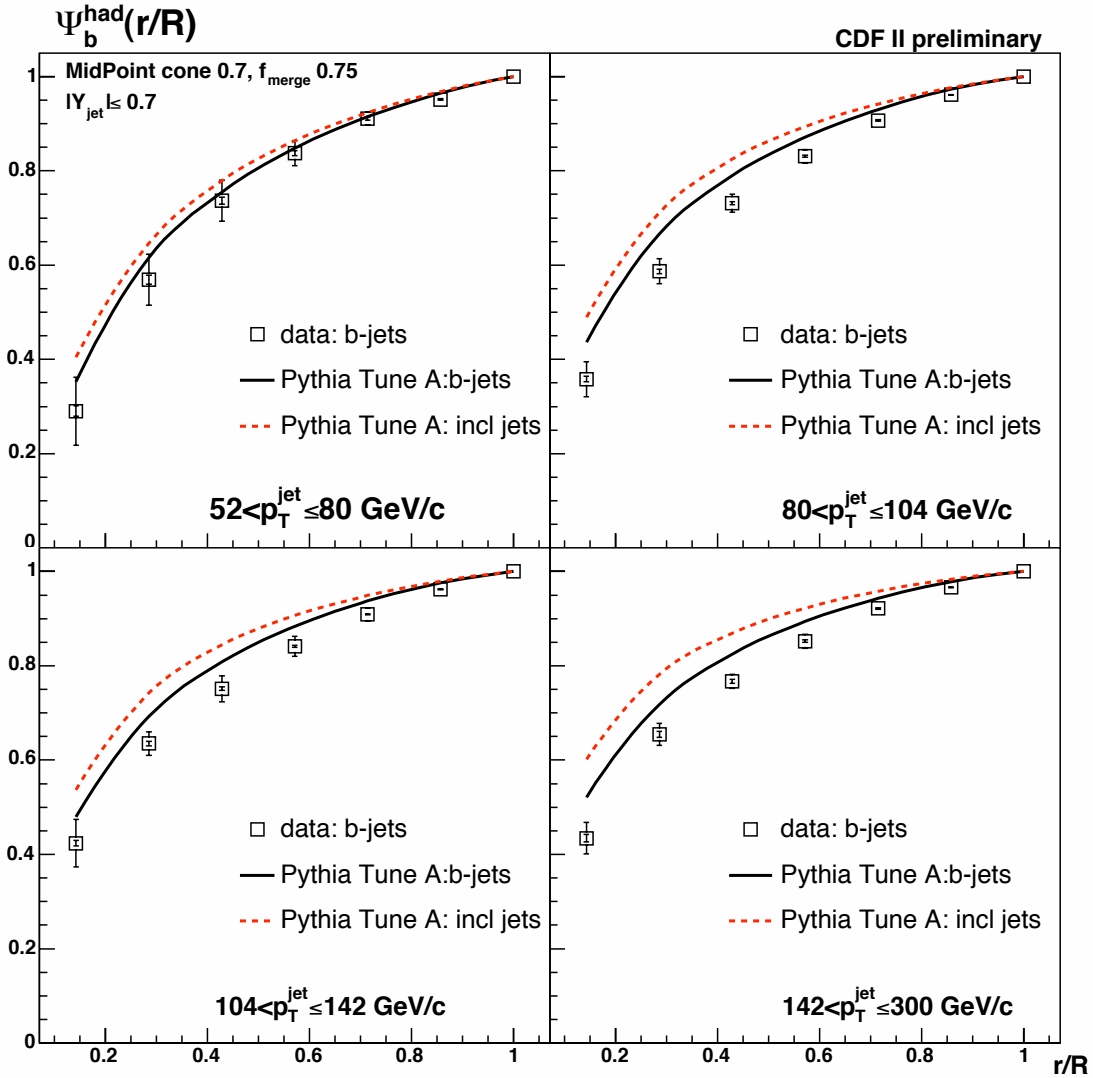

FIGURE 5.11: Hadron level integrated b-jet shapes for each of the p_T bins considered. The results are shown as black points where the error bars represent the total and statistical errors. The statistical errors are not visible as they are smaller than the points. The MC predictions for inclusive jets and for b-jets are shown as red and black curves, respectively.

Figure 5.12 shows the reconstructed hadron level b-quark jet shapes, where the inclusive jet shapes are taken from the Pythia Tune A MC predictions, whereas the tagged shapes are still measured in data. Only the statistical errors are shown. Figure 5.13 shows the ratio between the hadron level b-quark jet shapes reconstructed using the inclusive shapes from MC and the default hadron level b-quark jet shapes. Only the statistical errors of the numerator are shown on the points. The yellow bands show the total systematic errors of the hadron level b-quark jet shapes. This shows that the final results are somewhat sensitive to the small differences

in the measured inclusive jet shapes between the data and the Pythia Tune A MC. These differences are most of the time within the total systematic uncertainties on the final measurement.

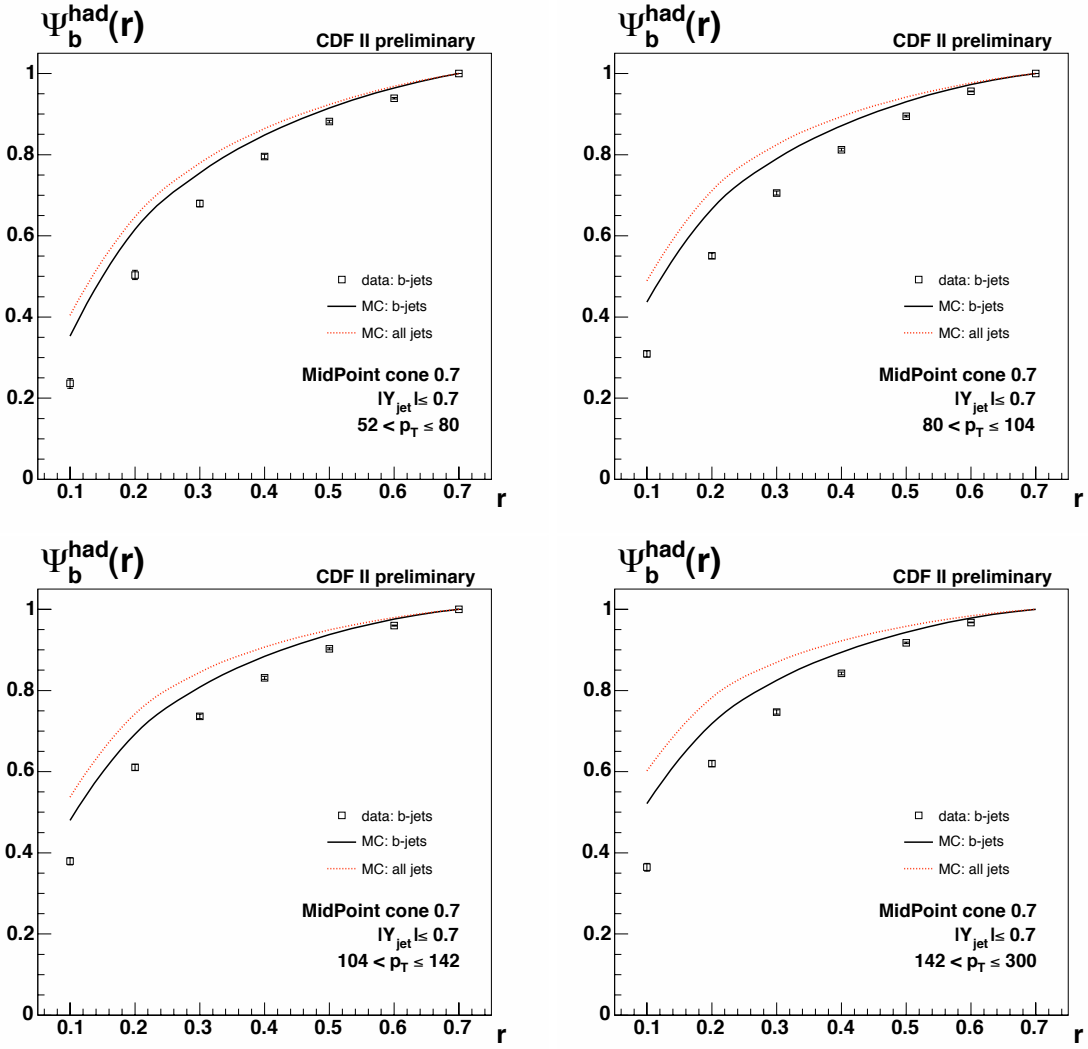

FIGURE 5.12: Hadron level integrated jet shapes in data but using the raw inclusive shapes obtained from Pythia Tune A MC. The results are shown as black points. Only the statistical errors are shown. The MC predictions for inclusive jets and for b-jets are shown as red and black curves, respectively.

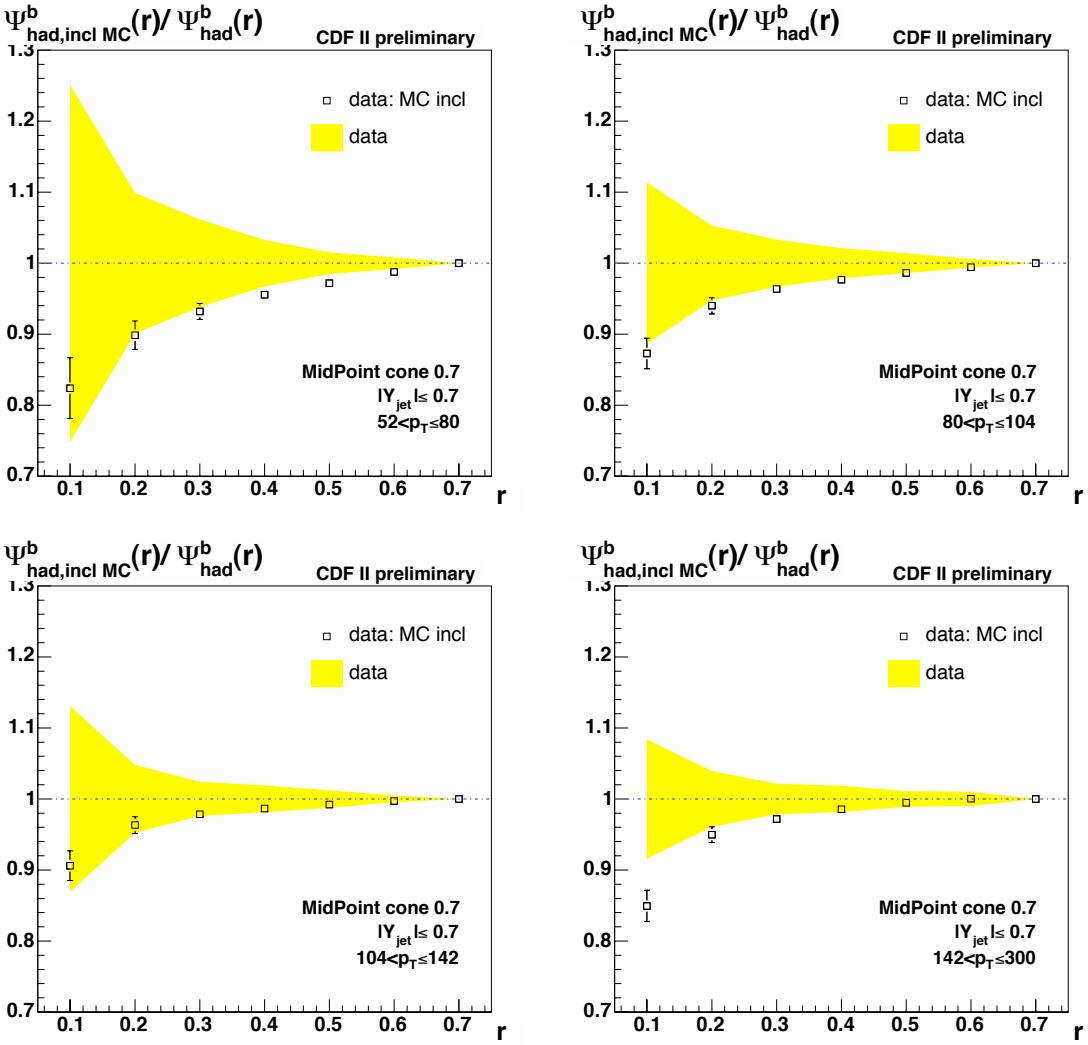

FIGURE 5.13: Ratio of the hadron level b-quark jet shapes reconstructed using the raw inclusive shapes taken from Pythia Tune A MC over those reconstructed from the raw inclusive shapes measured in data. The ratios are shown as black points, where the errors are the statistical errors on the numerator only. The yellow shaded areas represent the total errors on the hadron level b-quark jet shapes.

5.3 Variation with p_T

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 of the jets increase. Jets of a particular flavour are expected to become narrower as the p_T increases, mainly due to the running of the strong coupling constant, α_s. There is also a small effect due to the boost of the jets.

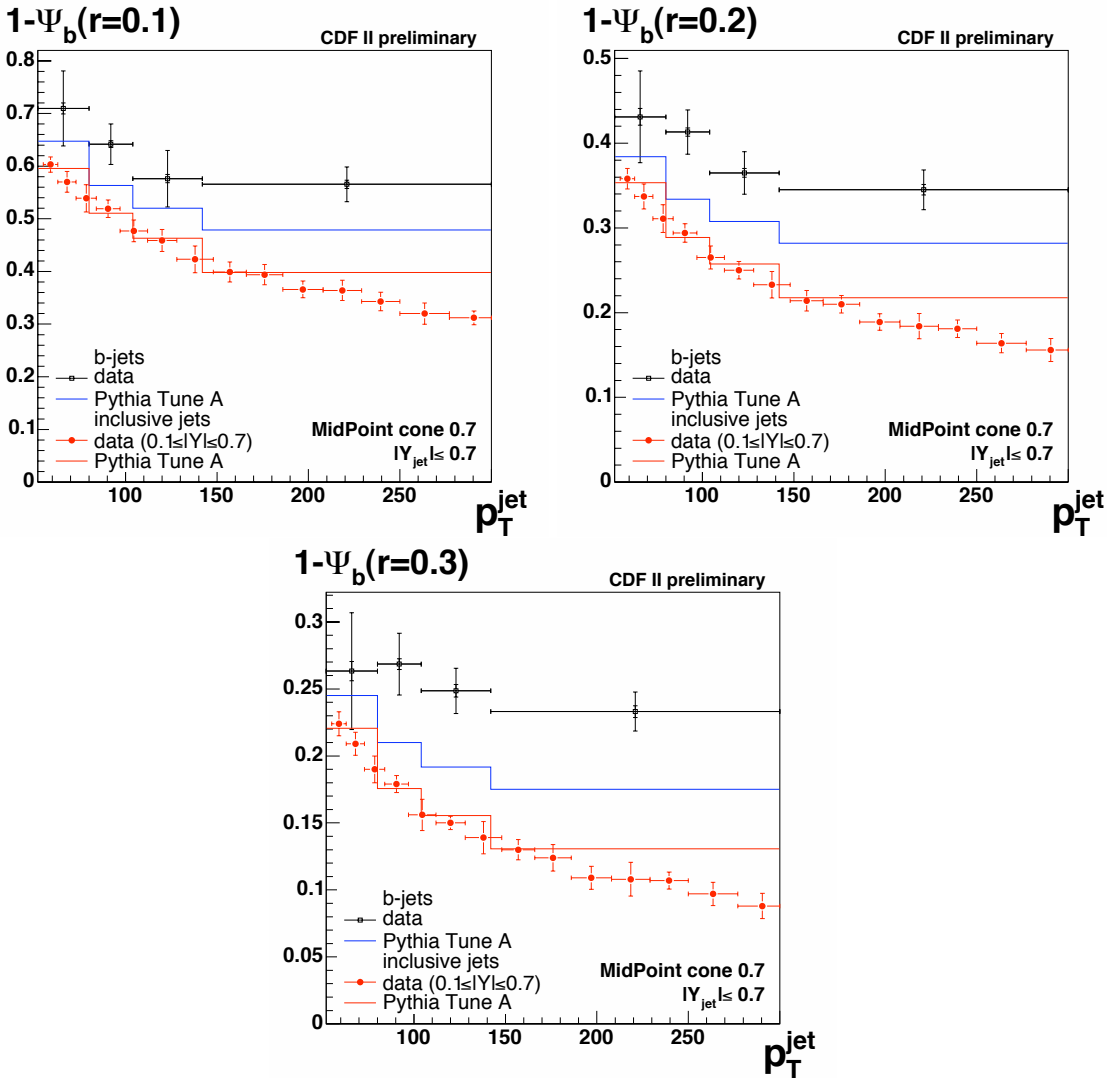

FIGURE 5.14: Fractional p_T outside a cone of radius $r = 0.1$, 0.2 and 0.3 around the jet axis as a function of the p_T of the jet. The results for b-quark jet shapes are shown as black points and the Pythia Tune A MC predictions are shown as blue lines. Also shown are the previously published inclusive jet shape results (red points) and the Pythia Tune A MC predictions (red lines) for inclusive jet shapes. The errors bars on the data points represent the total errors.

Figure 5.14 shows the evolution with jet p_T of the fractional p_T outside a cone of fixed radius (0.1, 0.2 and 0.3 are shown). Also shown on these plots are the predictions from the Pythia Tune A MC for both the b-jets and the inclusive jets. The previously published inclusive jet shapes results are shown on these plots as red points [2]. The latter agree very well with the MC predictions. The b-quark jet shapes measured in data do not agree with the Pythia Tune A MC predictions. The values shown represent the average value over the whole p_T range of the bin.

Because there are many more jets in the low p_T region of each bin, the average shapes for each bin are going to be dominated by the lower p_T jets. For this reason, when comparing the inclusive results and predictions, the two are expected to agree for the lower p_T regions of the bins.

5.4 Dependence on the Single b-quark Jet Fraction

The main reason why the Pythia Tune A MC predictions do not match the data very well is thought to be because the fraction of b-jets which come from gluon splitting is badly reproduced in Leading order MC. Figure 5.15 shows the same results as figure 5.14 but comparing the results to the Pythia Tune A MC predictions for b-quark jets as well as jets containing a single b-quark and jets with two b-quarks inside the same jet cone.

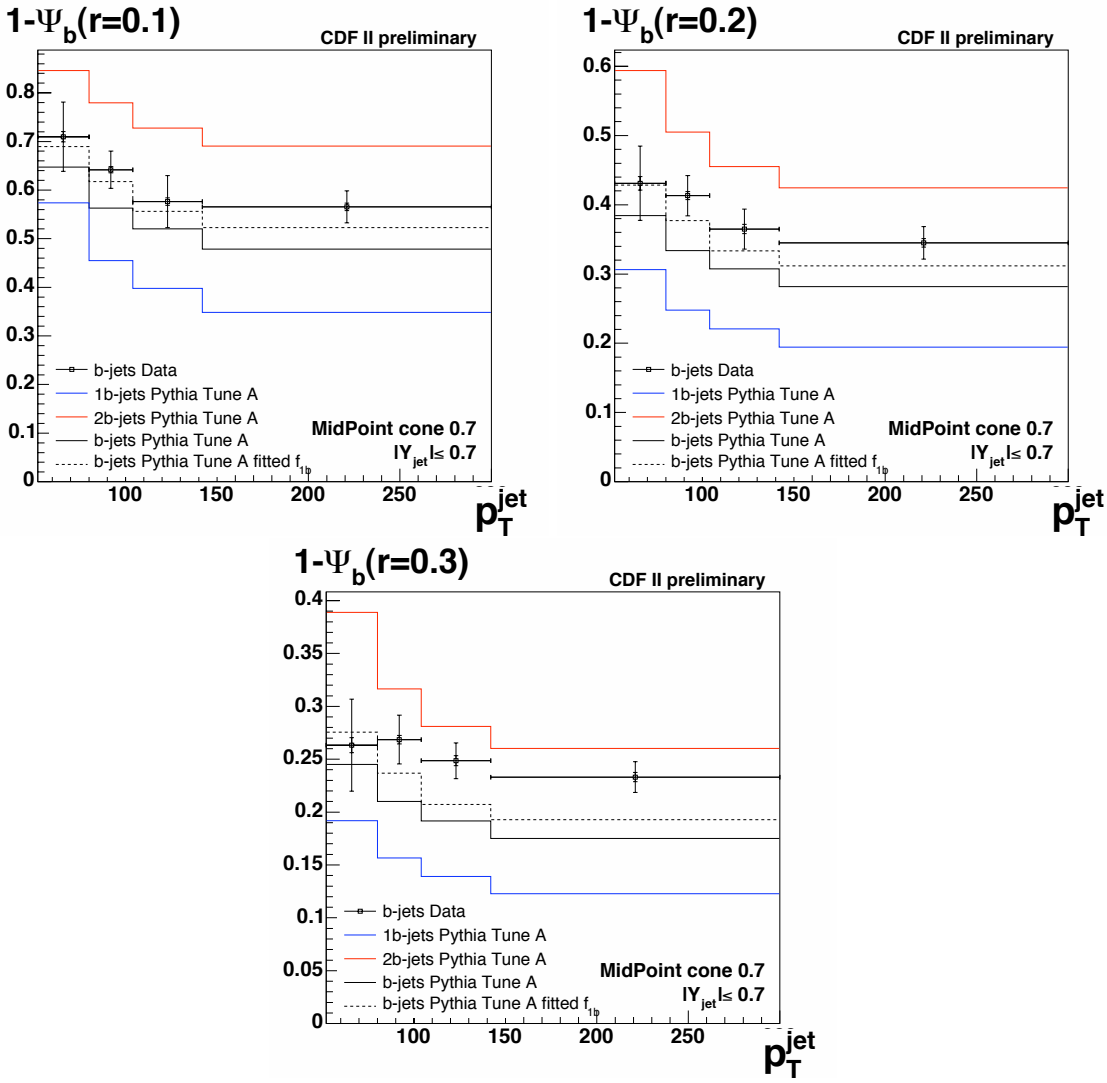

FIGURE 5.15: Fractional p_T outside a cone of radius $r = 0.1$, 0.2 and 0.3 around the jet axis as a function of the p_T of the jet. The results are shown as black points. The vertical error bars represent the total errors. The Pythia Tune A MC predictions for single and double b-quark jet shapes are shown as blue and red lines, respectively. Also shown are the Pythia Tune A MC b-quark jet shapes obtained from the fitted f_{1b} fractions (black dotted lines) and using the default f_{1b} fractions (black full lines).

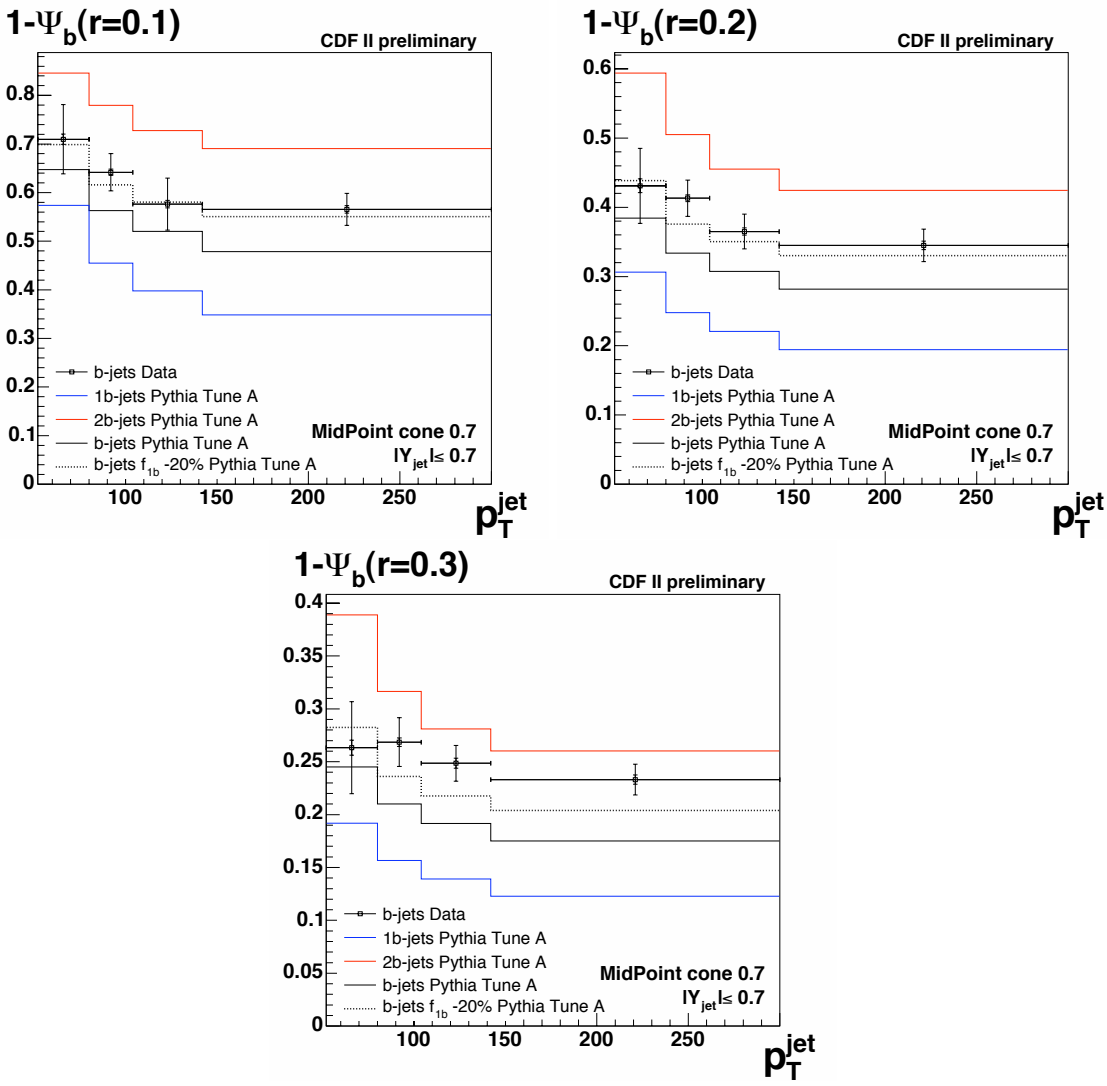

FIGURE 5.16: Fractional p_T outside a cone of radius $r = 0.1$, 0.2 and 0.3 around the jet axis as a function of the p_T of the jet. The results are shown as black points. The Pythia Tune A MC predictions for single and double b-quark jet shapes are shown as blue and red lines, respectively. Also shown are the Pythia Tune A MC predictions (black full lines) for b-quark jet shapes using the default f_{1b} fraction as well as using an f_{1b} fraction 10% and 20% lower, in absolute terms (black dashed and dotted lines, respectively).

p_T range	f_{1b}	f_{1b} MC	Δf_{1b}
52-80	0.575	0.774	-0.199
80-104	0.498	0.684	-0.186
104-142	0.519	0.627	-0.108
142-300	0.490	0.602	-0.112

TABLE 5.1: Fitted fraction of single b-quark jets in data, alongside the Pythia Tune A MC predictions and the difference between these two values.

It is possible to extract from data, for each p_T bin, the most probable f_{1b} fraction. This is done using the same binned χ^2 minimisation method as the one used for the SecVtx mass fit. The b-jet shapes obtained in data are fitted to the Pythia Tune A MC predictions for the single and double b-quark jets. The results obtained from the fit for each bin are shown in table 5.1. The fit errors are very large because of the large errors on the data and the fact that

the correlations between the bins have not been taken into account. For this reason it is not possible, at this stage, to state that these values truly constrain the f_{1b} fraction but they are an indication of the most probable value for this f_{1b} fraction. Also shown in that table are the f_{1b} fractions for each bin as predicted by Pythia Tune A and the difference between the predicted and fitted values. The difference is always within the 20% systematic uncertainty used for this analysis. It appears that the Pythia Tune A Leading Order MC systematically underestimates the gluon splitting rate in b-quark jet production. Figure 5.16 shows the variation with p_T of the fractional transverse momentum of the b-quark jets outside a cone of fixed radius compared to the values predicted by Pythia Tune A MC for b-quark jets, single b-quark jets, double b-quark jets as well as the predictions for b-quark jets if the f_{1b} fraction is decreased by 10% and 20% (absolute) with respect to the default Pythia Tune A MC values.

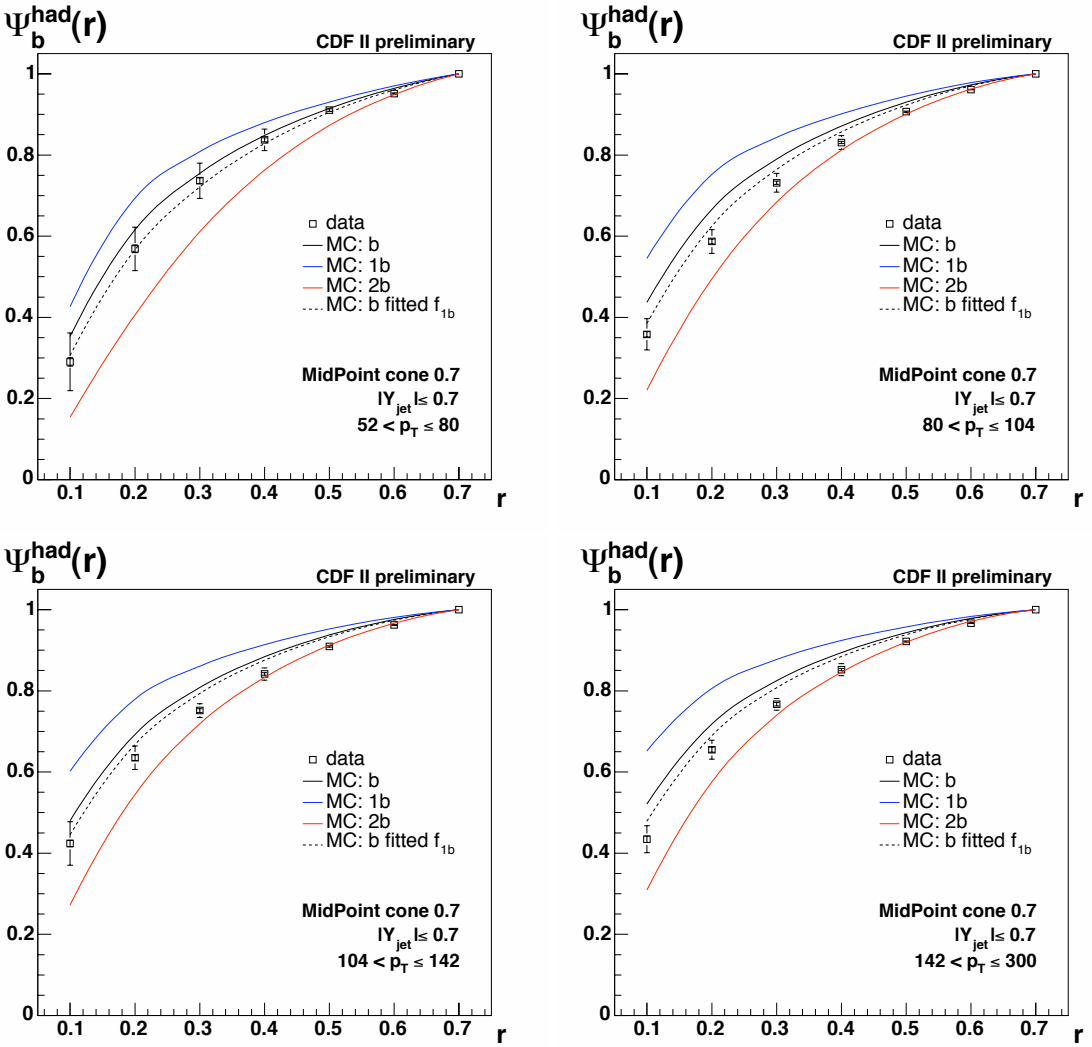

FIGURE 5.17: Hadron level integrated jet shapes for b-quark jets in data (black points). The total errors are shown. The Pythia Tune A MC predictions using the default and the fitted f_{1b} fractions are also shown (dotted and full black lines, respectively) along with the predictions for single and double b-quark jet shapes (red and blue lines, respectively).

Figure 5.17 shows a comparison of the hadron level integrated shapes of b-quark jets in data (black points) alongside the predictions from Pythia Tune A for b-quark jets, using the MC values for the f_{1b} fraction (dotted black line) and the fitted values for the f_{1b} fractions (full

black line). Also shown are the predictions for single and double b-quark jets (red and blue full lines, respectively). These plots show that the agreement between data and Pythia Tune A MC with the fitted f_{1b} fractions is better than with the default fraction. Figure 5.18 shows the ratio of the predicted Pythia Tune A MC jet shapes over the measured b-quark jet shapes. The total errors on the measured b-quark jet shapes are shown as yellow bands. This shows that the b-quark jet shapes in Pythia Tune A MC using the fitted f_{1b} fraction agree much better with the data than the default b-quark jet shape predictions.

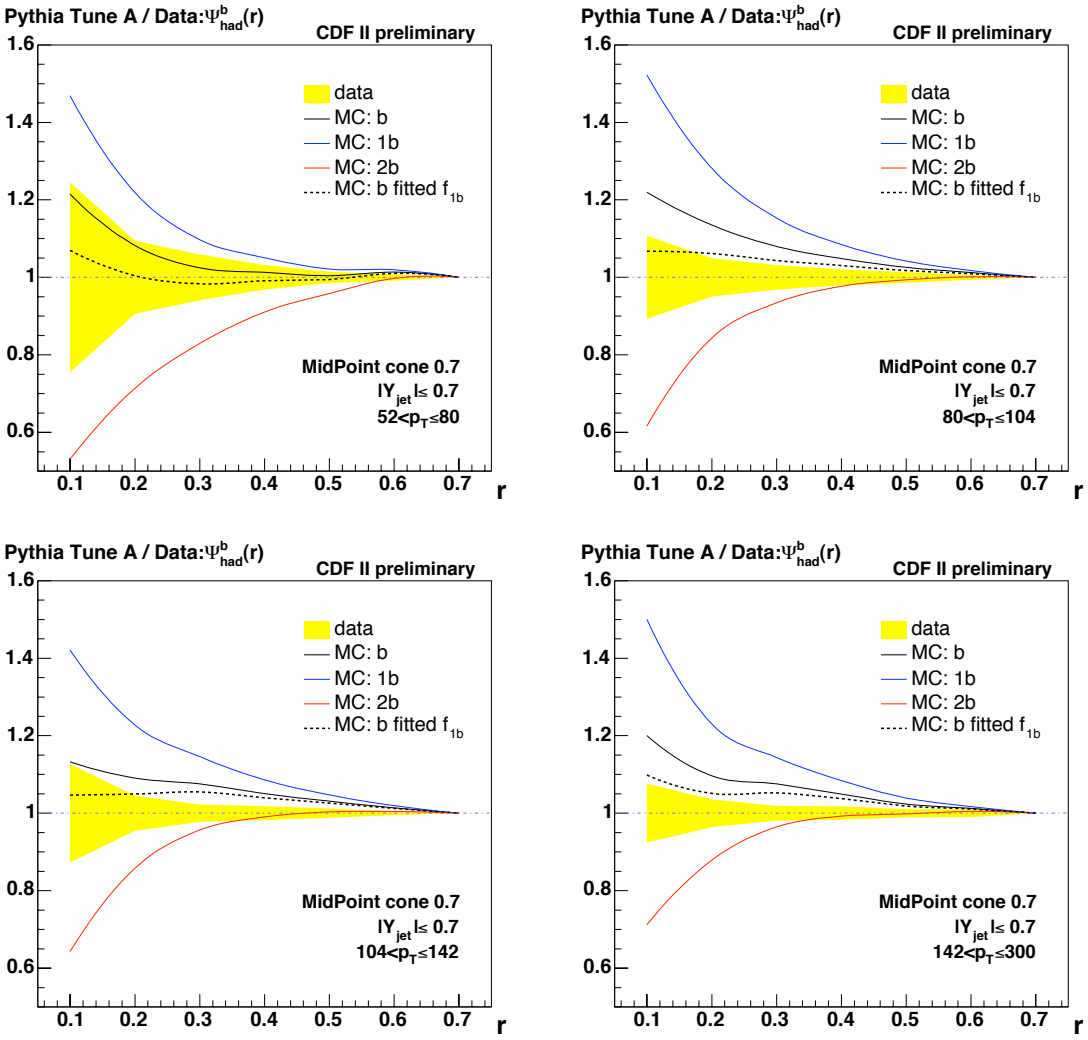

FIGURE 5.18: Ratio of the hadron level integrated jet shapes for b-quark jets for various theoretical predictions over the values obtained in data. The Pythia Tune A MC predictions using the default and the fitted f_{1b} fractions are shown (full and dotted black lines, respectively) along with the predictions 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.

Figure 5.19 shows the hadron level b-quark jet shapes in data compared to the Pythia Tune A predictions using the default values for the f_{1b} fractions (black lines). The Pythia Tune A predictions are also shown for the cases where the f_{1b} fraction is decreased by 10% and 20% (absolute) with respect to the default values (dotted and dashed black curves, respectively). This emphasises that the fraction of gluon splitting events is underestimated in Pythia Tune A.

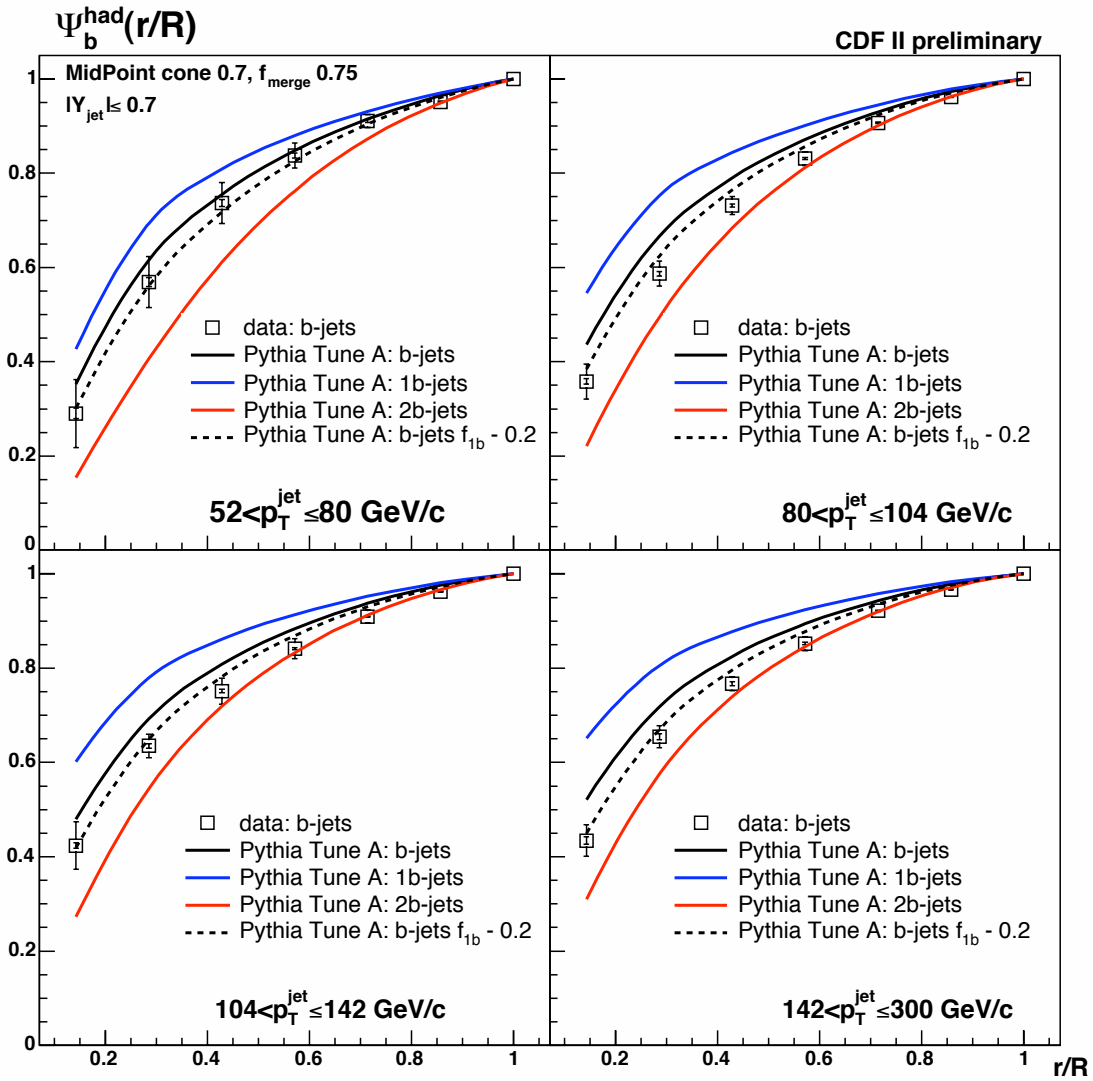

FIGURE 5.19: Hadron level integrated jet shapes for b-quark jets in data (black points). The total errors are shown. The Pythia Tune A MC predictions using the default f_{1b} fractions are shown as black lines. The Pythia Tune A MC predictions using f_{1b} fractions 10% and 20% (absolute) below the MC values are also shown (dotted and dashed black lines, respectively) along with the predictions for single and double b-quark jet shapes (red and blue lines, respectively).

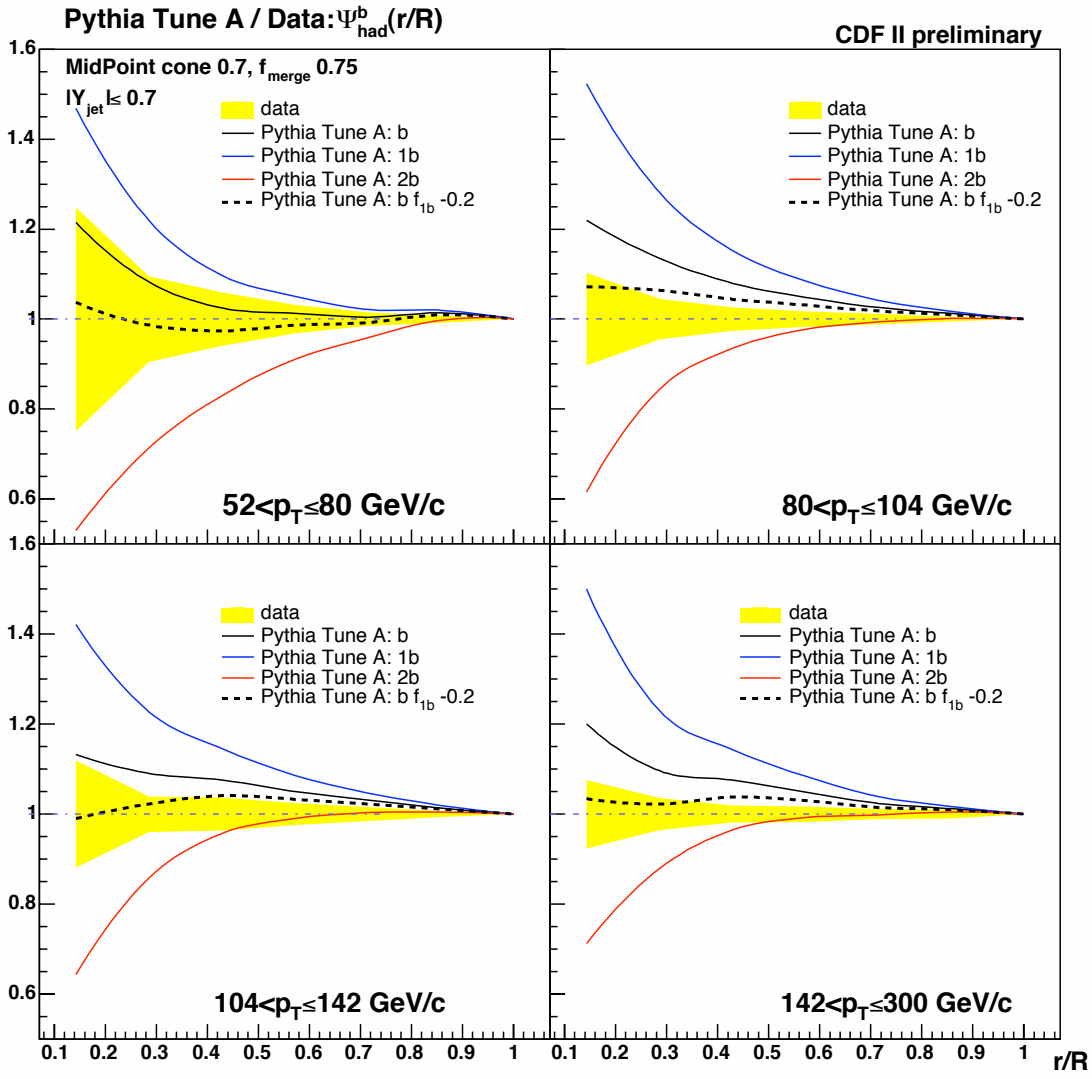

FIGURE 5.20: Ratio of the hadron level integrated jet shapes for b-quark jets for various theoretical predictions over the values obtained in data. The Pythia Tune A MC predictions using the default f_{1b} fractions are shown as black lines. The Pythia Tune A MC predictions using f_{1b} fractions 10% and 20% (absolute) below the MC values are also shown (dotted and dashed black lines, respectively) 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.

Chapter 6

Summary and Conclusions

For the first time at hadron colliders, the shapes of b-quark jets have been measured. This measurement uses approximately $300\,\mathrm{pb}^{-1}$ of data from the CDF detector. Because it is not possible to obtain a jet sample which contains only b-quark jets, a sophisticated unfolding method was needed in order to extract the b-quark jet shapes from the measured shapes for all jets and those measured for tagged jets. The tagging of the jets was achieved using the SecVtx tight tagger which increases the b-jet content from about 5% before any tagging requirement to $20-40\%$ after tagging, depending on the p_T of the jet. It was necessary to verify that this tagging algorithm could be applied not only to the default jet cone algorithm (JetClu with a cone size of 0.4) but also to other cone algorithms, in particular the one used in this analysis (MidPoint with a cone size of 0.7). It was found that the algorithm could be used in exactly the same way as for the default case, as long as the sub-cone inside which tracks used to reconstruct the secondary vertex can be found is kept at 0.4. Changing this sub-cone size to 0.7 significantly increases the tagging efficiency but the number of mistags also increases. A full re-optimisation of the algorithm would be needed before increasing this sub-cone to the full jet cone size of 0.7.

Other than the secondary vertex tagging, the analysis presented in this thesis involves the calculation of a number of different parameters before the b-quark shapes can be extracted statistically. These other parameters are the biases due to the tagging requirement, both on b- and nonb-jets, and also the hadron level corrections which correct the detector level jet shapes back to hadron level where they can more easily be compared to other experimental results or to theoretical models. The b-quark jet shape results show a disagreement with respect to the Pythia Tune A MC predictions. This is thought to be mainly due to the fact that the fraction of b-jets that contain one or two b-quarks is not correctly modelled in Pythia Tune A. The results, despite the relatively large systematic errors, can be used to extract the most probable value of this single b-quark jet fraction for each p_T bin considered. This fraction is found to be between 10% and 20%, in absolute terms, lower in data than in Pythia Tune A. The agreement between the data and the Pythia Tune A MC predictions using this new, fitted, fraction is better.

The systematic errors on the measurement are still relatively high. The statistical errors on the other hand are very small. The dominant sources of systematic errors are the fraction of b-quark jets which contain one or two b-quarks and similarly the fraction of c-quark jets, part of the background, which contain one or two c-quarks inside the jet cone. The difference between the hadron level shapes reconstructed from raw shapes using tracks and the default shapes, reconstructed using calorimeter towers, is also a dominant source of systematic error. Despite the fact that the Pythia Tune A MC statistics was limited, in particular for the calculation of the tagging biases on nonb-jets, generating more MC would not have significantly decreased the total systematic errors for any of the measurement points.

The measurement of b-quark jet shapes has been shown to be a good method for checking the MC modelling of the ratio between flavour creation and gluon splitting in b-quark jet production. The analysis methodology is somewhat sensitive to this ratio through the different parameters of the unfolding but this dependence is not very strong. It does not enter linearly into the unfolding equations and an uncertainty on the single b-quark jet fraction can be accounted for in the systematic error calculation. Despite this weak dependence on the single b-quark jet fraction, it is possible to obtain a relatively precise measurement of the b-quark jet shapes at CDF. This method should also be applicable to LHC studies, where b-jets will be involved in many searches for new particles. A good understanding of b-jets, in particular the relative amount of gluon splitting to flavour creation jets, is very important for such searches.

The b-quark jet shapes could also be used as parameters to apply cuts on in order to enhance or reduce the relative contributions of gluon splitting or flavour creation b-quark jets. For example, by requiring the fractional jet energy outside a fixed cone radius to be below a certain threshold, it is possible to enhance the flavour creation component of the b-quark jets.

A summary of the work presented in this thesis on the measurement of the b-quark jet shapes can be found in the CDF internal note number 8143. This note contains additional sections on comparisons of general jet related quantities as well as an appendix which compares the inclusive jet shape results using my version of the analysis code with the previously published inclusive jet shape results. I expect this analysis to be approved by the CDF collaboration in the very near future. These results would then become official CDF preliminary results.

Appendix A

Differential b-quark Jet Shapes

It is possible to write a similar equation to equation 3.8 but using the differential instead of the integrated shapes

$$\rho^{\mathrm{b}}_{\mathrm{had}}(r) = C_{\mathrm{had}}(r)\frac{\rho_{\mathrm{meas}}(r) - (1-p_{\mathrm{b}})b^{\mathrm{nonb}}(r)\rho^{\mathrm{nonb}}_{\mathrm{det}}(r)}{p_{\mathrm{b}}b^{\mathrm{b}}(r)} \tag{A.1}$$

This would be the simplest equation for the b-quark jet unfolding. The bias terms would be defined, in analogy to the integrated jet shape definition, as

$$b^{\mathrm{b}}(r) = \frac{\rho^{\mathrm{tag}}_{\mathrm{b\,MC}}(r)}{\rho^{\mathrm{incl}}_{\mathrm{b\,MC}}(r)} \tag{A.2}$$

for b-jets and

$$b^{\mathrm{nonb}}(r) = \frac{\rho^{\mathrm{tag}}_{\mathrm{nonb\,MC}}(r)}{\rho^{\mathrm{incl}}_{\mathrm{nonb\,MC}}(r)} \tag{A.3}$$

for nonb-jets. The hadron level corrections would become

$$C_{\mathrm{had}}(r) = \frac{\rho^{\mathrm{had}}_{\mathrm{MC}}(r)}{\rho^{\mathrm{det}}_{\mathrm{MC}}(r)}. \tag{A.4}$$

Combining all these terms into equation A.1 gives

$$\rho^{\mathrm{b}}_{\mathrm{had}}(r) = \frac{\rho^{\mathrm{had}}_{\mathrm{MC}}(r)}{\rho^{\mathrm{det}}_{\mathrm{MC}}(r)}\frac{\rho_{\mathrm{meas}}(r) - (1-p_{\mathrm{b}})\frac{\rho^{\mathrm{tag}}_{\mathrm{nonb\,MC}}(r)}{\rho^{\mathrm{incl}}_{\mathrm{nonb\,MC}}(r)}\rho^{\mathrm{nonb}}_{\mathrm{det}}(r)}{p_{\mathrm{b}}\frac{\rho^{\mathrm{tag}}_{\mathrm{b\,MC}}(r)}{\rho^{\mathrm{incl}}_{\mathrm{b\,MC}}(r)}} \tag{A.5}$$

It is apparent from this equation that replacing the measured quantities by the MC quantities, the equation unfolds correctly. The problem with the use of the differential shapes is that the normalisations are not easy to take into account because of the constraint on the integral of the jet shapes. This constraint must be true both for tagged shapes and the shapes corrected back to detector level, for both the data and the MC,

$$\int_0^R \rho^{\mathrm{b}}_{\mathrm{meas}}(r)dr = \int_0^R \rho^{\mathrm{b}}_{\mathrm{det}}(r)b^{\mathrm{b}}(r)dr = 1 \tag{A.6}$$

and

$$\int_0^R \rho^{\mathrm{b}}_{\mathrm{det}}(r)dr = 1 \tag{A.7}$$

It is not trivial to obtain bias terms which preserve this normalisation. In the numerator it would be possible to redefine the terms involving nonb-jets in order to maintain the normalisation. The $b^{\mathrm{nonb}}(r)$ terms could be replaced by

$$b^{\mathrm{nonb}}_{\mathrm{new}}(r) = \frac{b^{\mathrm{nonb}}(r)}{\int_0^R b^{\mathrm{nonb}}(r)\rho^{\mathrm{nonb}}_{\mathrm{det}}(r)dr}. \tag{A.8}$$

The normalisations would depend on the measured nonb-jet shapes (i.e. on the raw inclusive jet shapes using our approximation) but would not depend on the shapes we want to extract from this equation, the b-quark jet shapes.

The bias terms in the denominator, on the other hand, are where the problems lie. The tagging bias terms on b-quark jets must have additional renormalisations in order to obtain properly normalised b-quark jet shapes. The problem is that these renormalisations will depend on the detector level b-quark jet shapes which we want to measure. The $b^{\mathrm{b}}(r)$ terms should be replaced by something which would have the form

$$b^{\mathrm{b}}_{\mathrm{new}}(r) = \frac{b^{\mathrm{b}}(r)}{\int_0^R b^{\mathrm{b}}(r)\rho^{\mathrm{b}}_{\mathrm{det}}(r)dr}, \tag{A.9}$$

where $\rho^{\mathrm{b}}_{\mathrm{det}}(r)$ are the detector level b-quark jet shapes we want to measure. We would therefore end up with a dependence on the b-quark jet shapes on the left hand side of the unfolding equation but also on the right hand side of the equation. It is therefore impossible to obtain such a simple unfolding equation for the differential shape as the one used for the integrated jet shape, even requiring additional renormalisation of the shapes.

An alternative approach would be to obtain an unfolding equation for the differential b-jet shapes by differentiating the equation for the integrated shapes. This would lead to the correct normalisation of the jet shapes because the normalisation of the integrated shapes is taken care of in the integrated shape unfolding equation and will be maintained after differentiation. This equation turns out to be relatively complicated even for the detector level b-quark jet shapes (i.e. without considering the hadron level jet shape corrections)

$$\begin{aligned}\rho^{\mathrm{b}}_{\mathrm{det}}(r) = \frac{\partial \Psi^{\mathrm{b}}_{\mathrm{det}}(r)}{\partial r} &= \frac{1}{p_{\mathrm{b}} b^{\mathrm{b}}(r)}\Bigg[\frac{-\frac{\partial b^{\mathrm{b}}(r)}{\partial r}}{b_{\mathrm{b}}(r)}\Big(\Psi_{\mathrm{meas}}(r) - (1-p_{\mathrm{b}})b^{\mathrm{nonb}}(r)\Psi^{\mathrm{nonb}}_{\mathrm{det}}(r)\Big) + \rho_{\mathrm{meas}}(r) \\ &\quad -(1-p_{\mathrm{b}})b^{\mathrm{nonb}}(r)\rho^{\mathrm{nonb}}_{\mathrm{det}}(r) - (1-p_{\mathrm{b}})\frac{\partial b^{\mathrm{nonb}}(r)}{\partial r}\Psi^{\mathrm{nonb}}_{\mathrm{det}}(r)\Bigg],\end{aligned} \tag{A.10}$$

where

$$\frac{\partial b^{\mathrm{b}}(r)}{\partial r} = \frac{\rho^{\mathrm{tag}}_{\mathrm{b\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{b,MC}}(r)} - b^{\mathrm{b}}(r)\frac{\rho^{\mathrm{incl}}_{\mathrm{b\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{b\,MC}}(r)} \tag{A.11}$$

and

$$\frac{\partial b^{\mathrm{nonb}}(r)}{\partial r} = \frac{\rho^{\mathrm{tag}}_{\mathrm{nonb\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{nonb,MC}}(r)} - b^{\mathrm{nonb}}(r)\frac{\rho^{\mathrm{incl}}_{\mathrm{nonb\,MC}}(r)}{\Psi^{\mathrm{incl}}_{\mathrm{nonb\,MC}}(r)}. \tag{A.12}$$

This equation involves both differential and integrated terms for all the r dependent parameters of the unfolding equation and for the measured shapes. Moreover, the error calculation would be extremely complex. It was therefore deemed not feasible to carry out this analysis using the differential jet shapes.

For the integrated jet shape, these problems are not present because the bias terms and hadron level corrections tend to 1 for $r \to R$ without any additional requirements. Also, because of the nature of the unfolding equation, the integrated b-jet shapes tend to 1 as $r \to R$.

Appendix B

Comparison of some General quantities

In this appendix some general quantities are shown with the comparison between data and Pythia Tune A MC simulation. The quantities are plotted for the four p_T bins used in this analysis for jets which pass all of the selection cuts of my analysis. The plots shown are

- p_T distribution of all jets: data and MC agree very well
- p_T distribution of tagged jets
- ϕ distribution of all jets: data and MC agree very well
- ϕ distribution of tagged jets
- rapidity distribution of all jets: slight difference between data and MC
- rapidity distribution of tagged jets
- Zvtx of all events: data and MC agree
- Missing E_T significance of all events: data and MC agree
- Number of towers inside a cone of 0.7 around the jet axis for all jets: 1st bin: OK. Highest p_T bin MC shifted 2 towers lower
- Number of towers inside a cone of 0.7 around the jet axis for tagged jets
- Number of tracks inside a cone of 0.7 around the jet axis for all jets: data and MC agree
- Number of tracks inside a cone of 0.7 around the jet axis for tagged jets
- p_T distribution of the towers inside a cone of 0.7 around the jet axis for all jets: data and MC agree well but MC slightly higher at high p_T
- p_T distribution of the towers inside a cone of 0.7 around the jet axis for all jets showing the low p_T behaviour: data and MC agree very well but MC slightly higher in lowest p_T bin
- p_T distribution of the towers inside a cone of 0.7 around the jet axis for tagged jets
- p_T distribution of the towers inside a cone of 0.7 around the jet axis for tagged jets showing the low p_T behaviour
- dR distribution of the towers inside a cone of 0.7 around the jet axis for all jets

- p_T distribution of the tracks inside a cone of 0.7 around the jet axis for tagged jets
- p_T distribution of the tracks inside a cone of 0.7 around the jet axis for tagged jets showing the low p_T behaviour
- Z_0 distribution of the tracks inside a cone of 0.7 around the jet axis for all jets
- Z_0 distribution of the tracks inside a cone of 0.7 around the jet axis for tagged jets
- d_0 distribution of the tracks inside a cone of 0.7 around the jet axis for all jets
- d_0 distribution of the tracks inside a cone of 0.7 around the jet axis for tagged jets

Most variables show a good agreement between data and Pythia Tune A MC. The jet p_T and ϕ distributions are fine. The jet rapidity distributions show a significant difference between the data and the MC. This difference is especially visible in the high p_T bins where the MC distribution is shifted to higher rapidity values and is somewhat broader than the data. The Z_{vtx} distributions are reasonably well described by the MC. The missing E_T significance is also fine. The number of towers inside a cone of 0.7 around the jet axis is well described for the lowest p_T bin but is shifted down with respect to the data by 1-2 towers at high p_T. This shift is also visible in the tagged jet sample. The distribution of the number of tracks inside a cone of 0.7 is well reproduced in the MC both for inclusive and tagged jets except for the lowest p_T bin where the distribution in MC is slightly narrower. The distribution of the p_T of the tracks inside the jet cone is well described at low tower p_T in MC but tends to be slightly broader than the data at high tower p_T. The track p_T distribution is well modelled by the MC at high track p_T. At low track p_T the distribution in MC is slightly flatter than in data. The Z_0 distribution of the tracks inside a cone of 0.7 around the jet axis is relatively well described by the MC. The MC is slightly narrower than the data but this effect is not large.

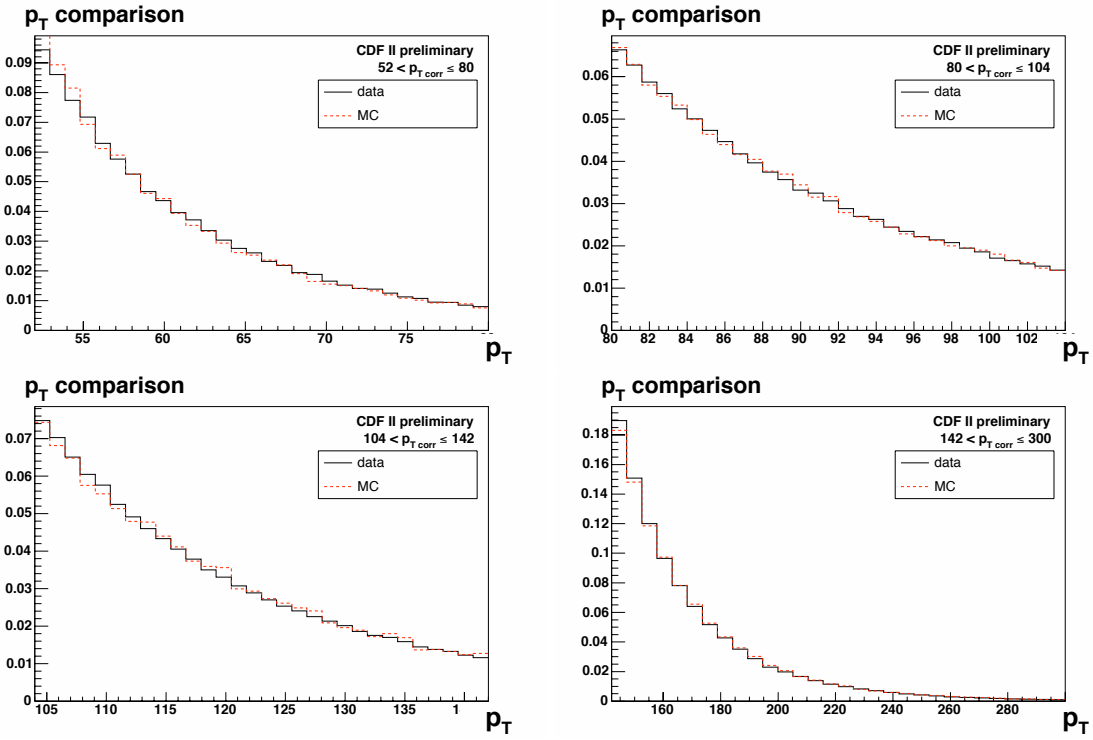

FIGURE B.1: p_T distribution of jets.

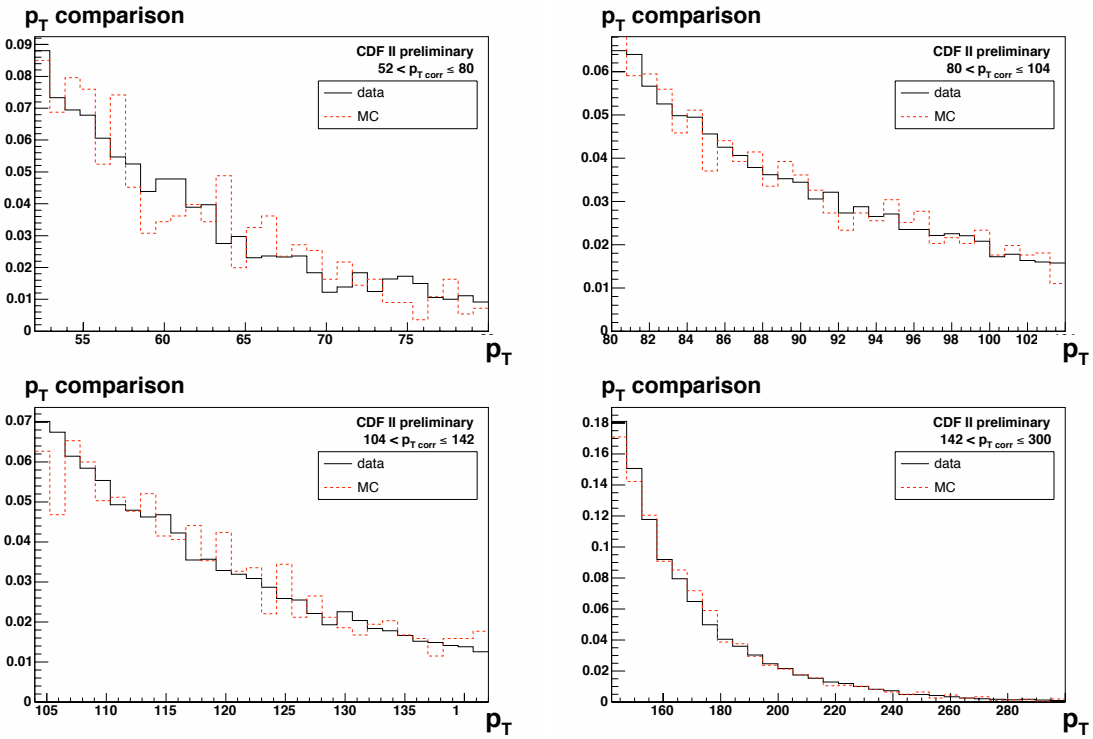

FIGURE B.2: p_T distribution of tagged jets.

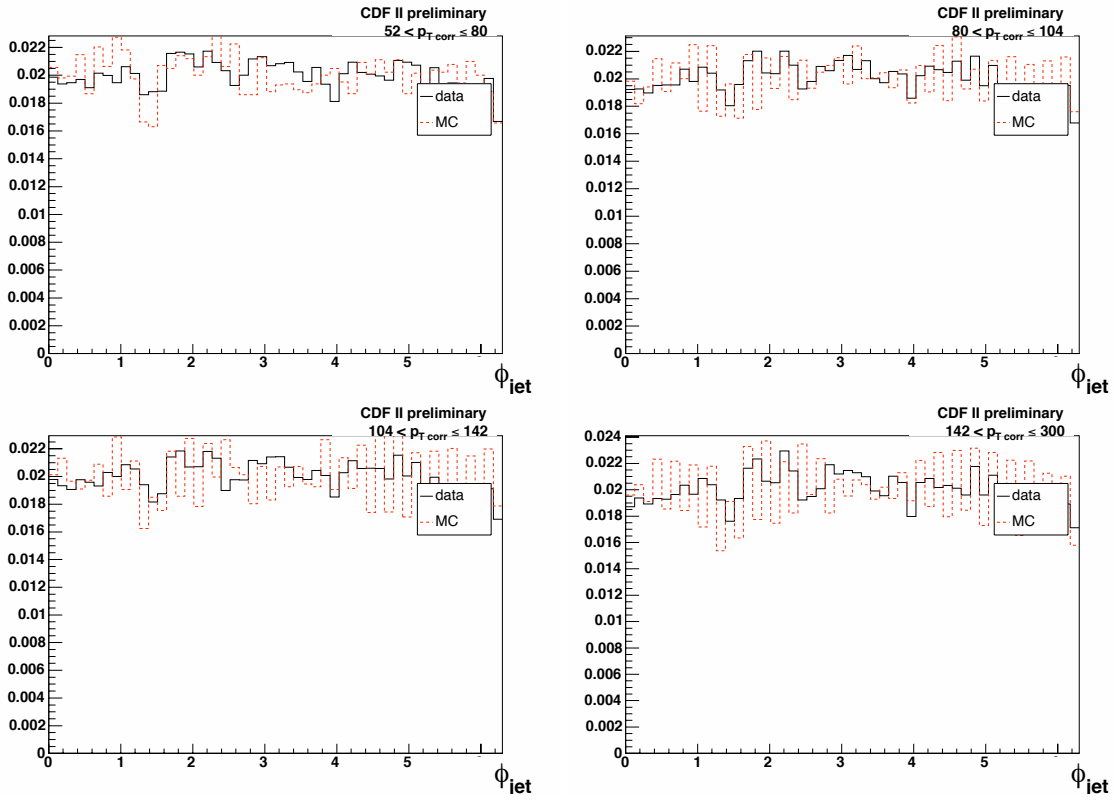

FIGURE B.3: ϕ distribution of jets.

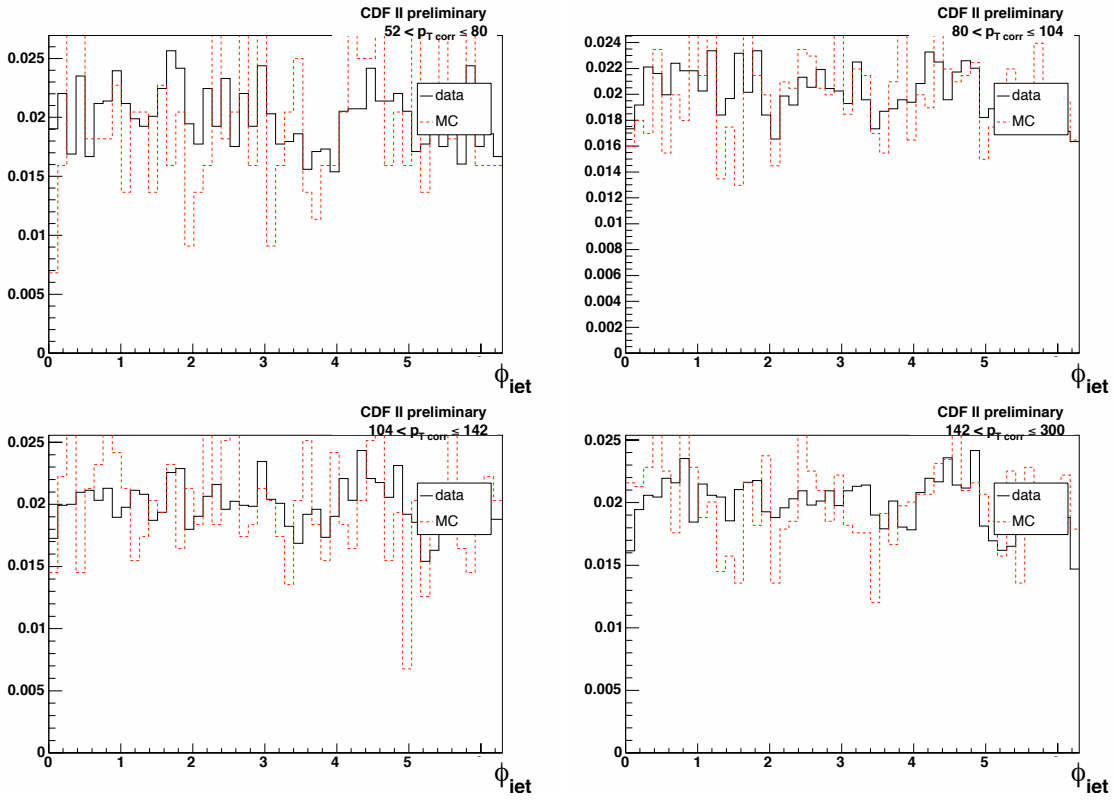

FIGURE B.4: ϕ distribution of tagged jets.

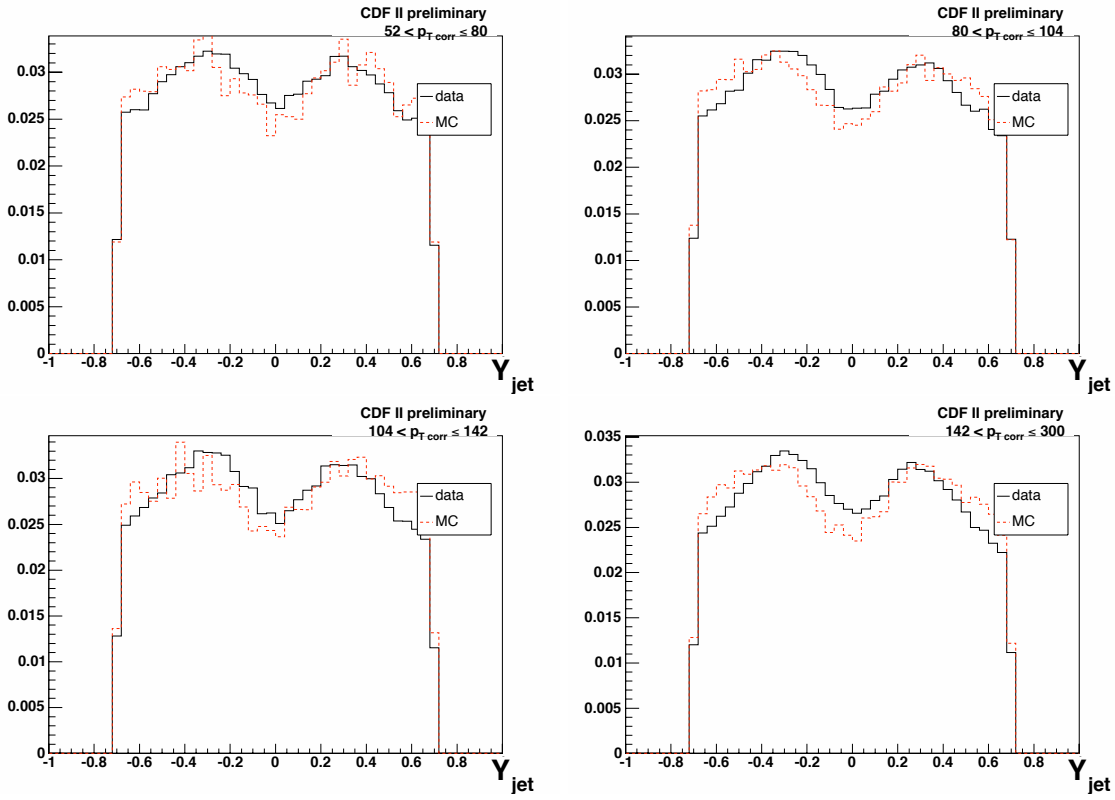

FIGURE B.5: Y distribution of jets.

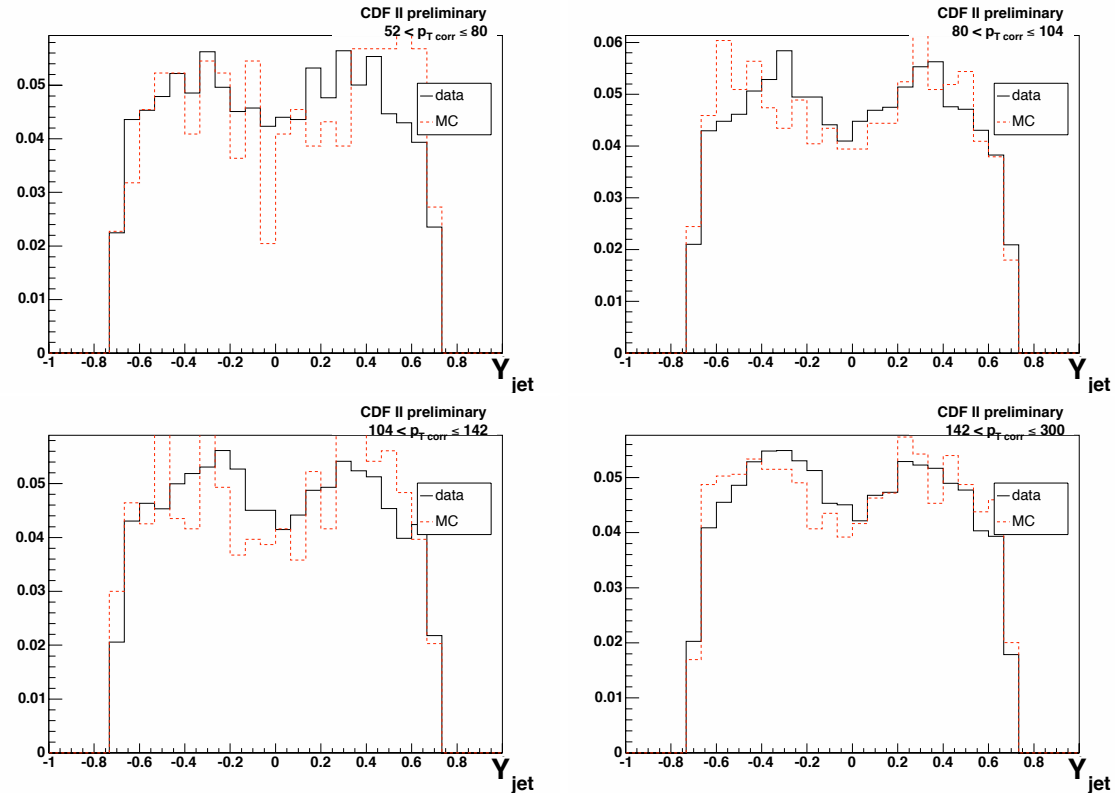

FIGURE B.6: Y distribution of tagged jets.

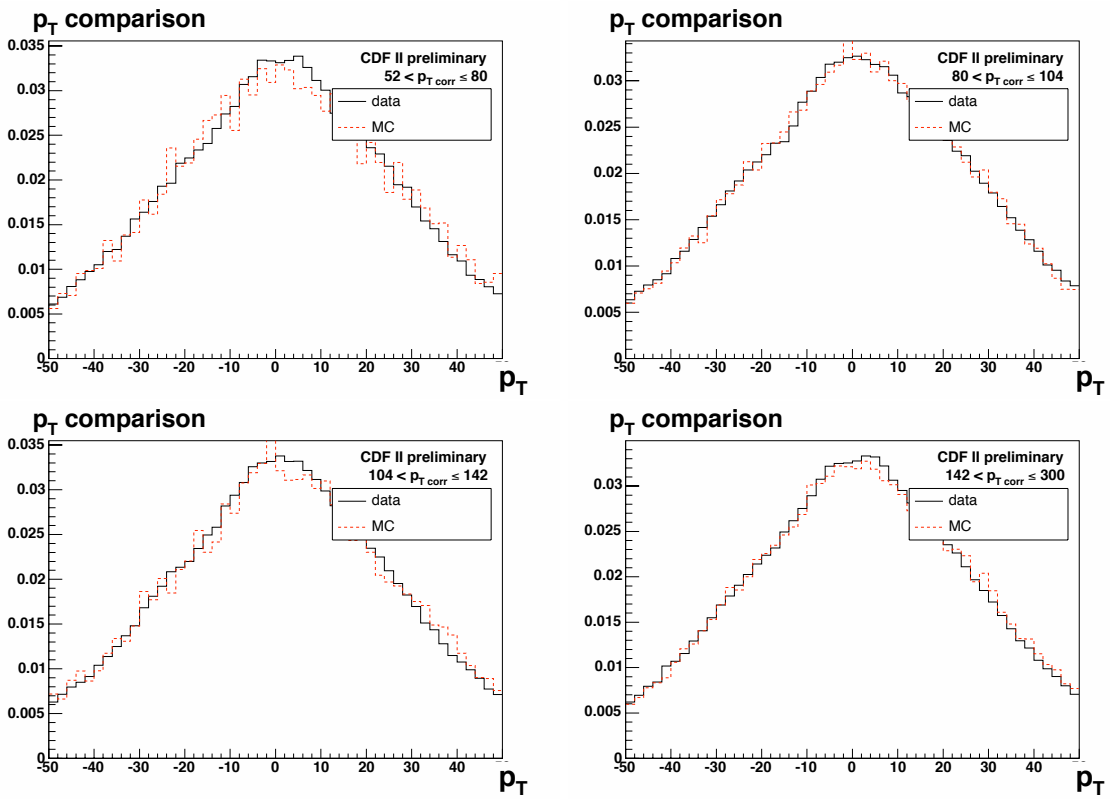

FIGURE B.7: Z_{vtx} distribution.

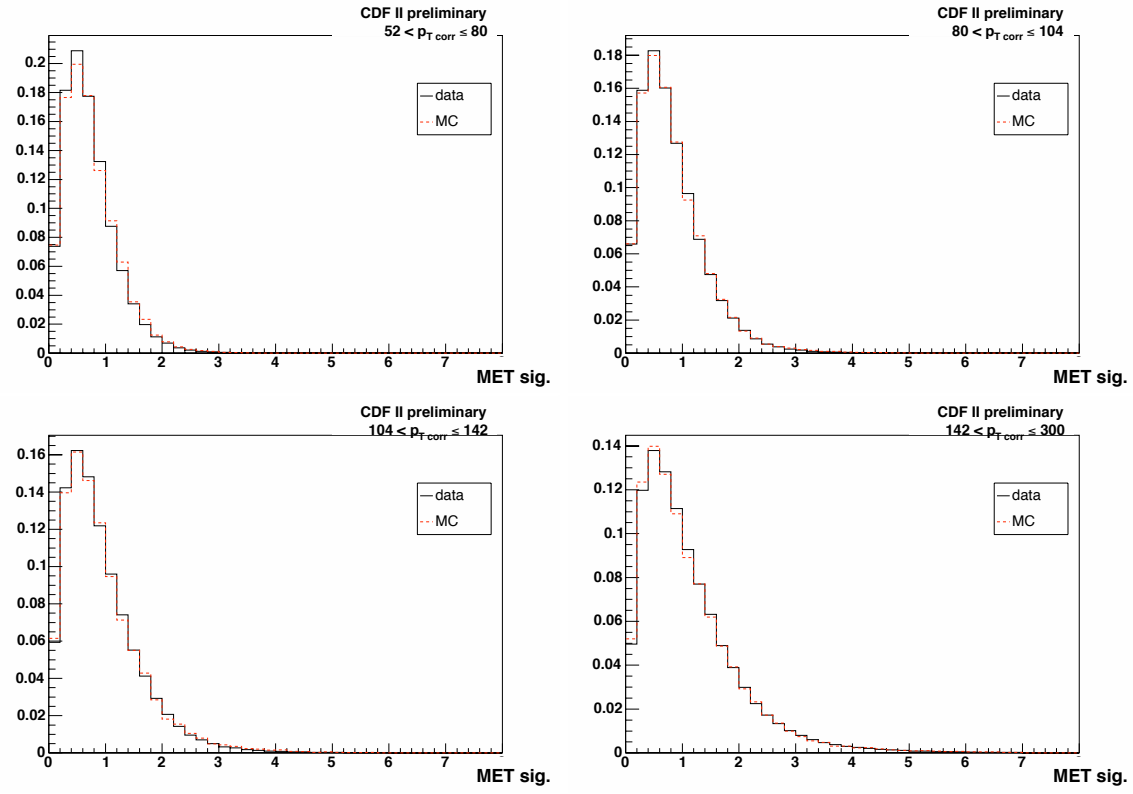

FIGURE B.8: Missing E_T significance distribution.

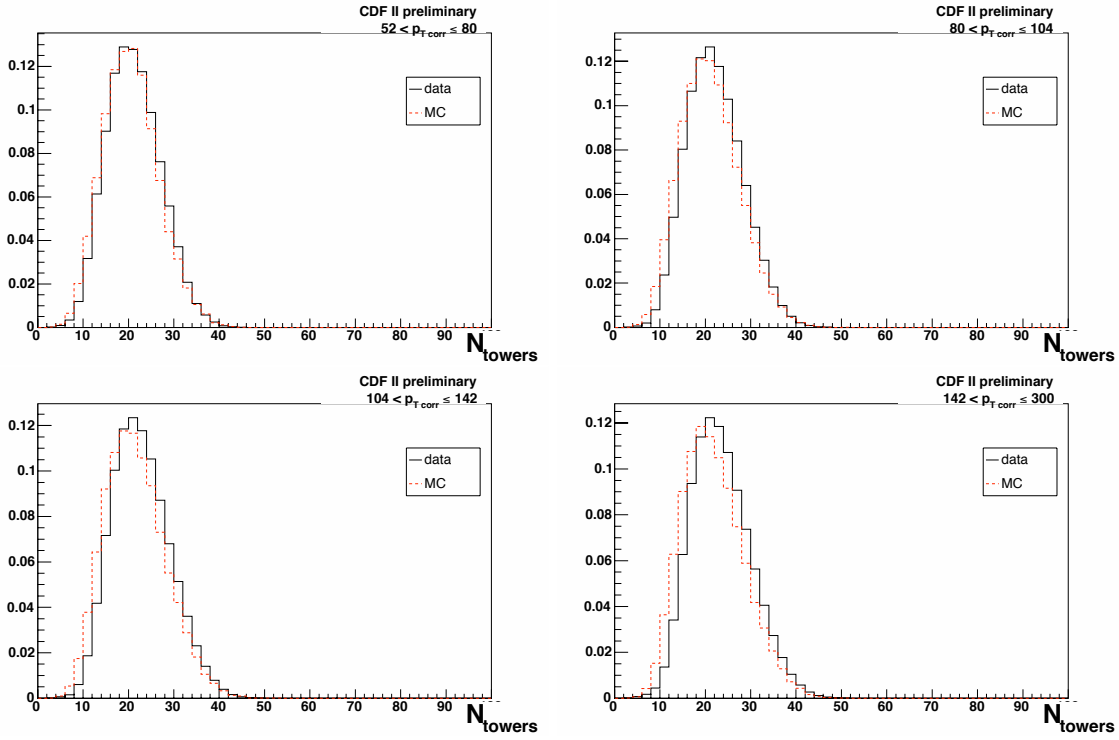

FIGURE B.9: Number of towers inside a cone of 0.7 around jet axis.

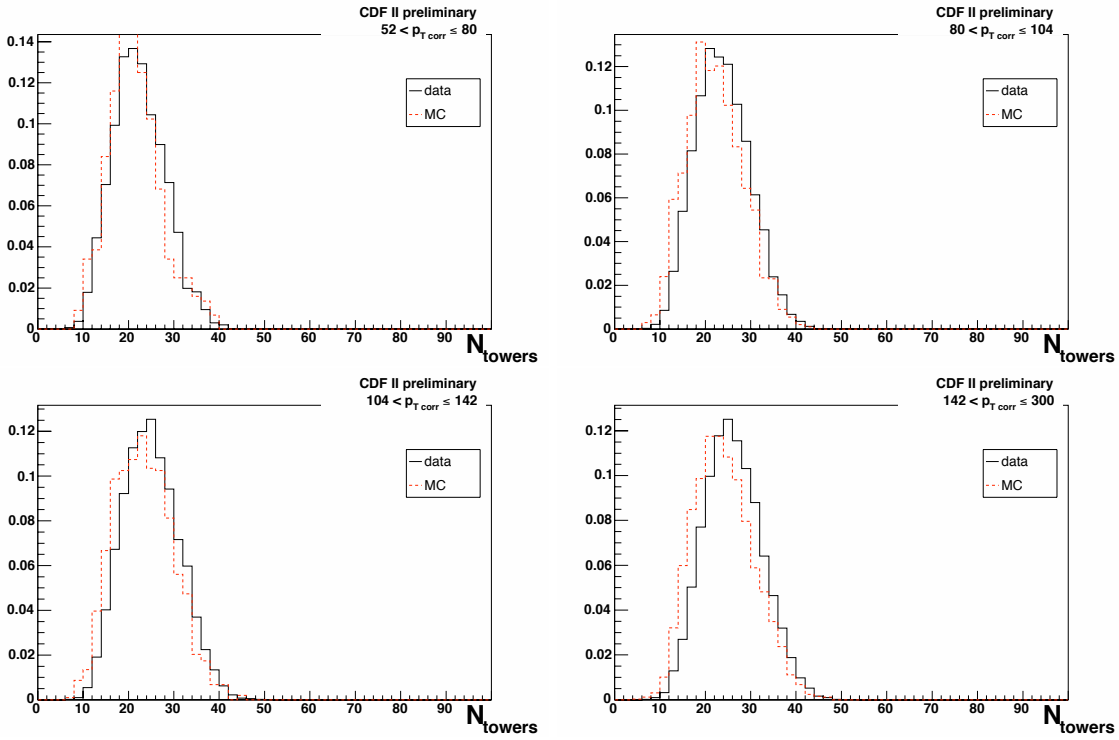

FIGURE B.10: Number of towers inside a cone of 0.7 around jet axis for tagged jets.

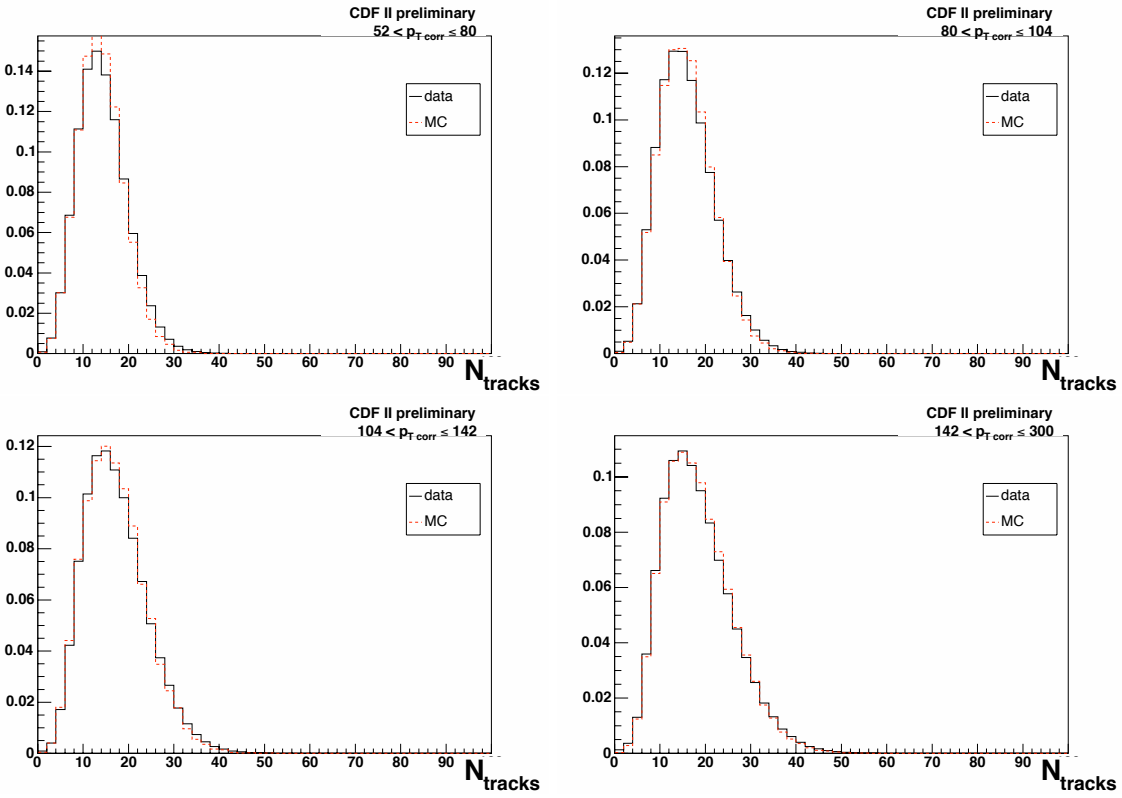

FIGURE B.11: Number of tracks inside a cone of 0.7 around jet axis.

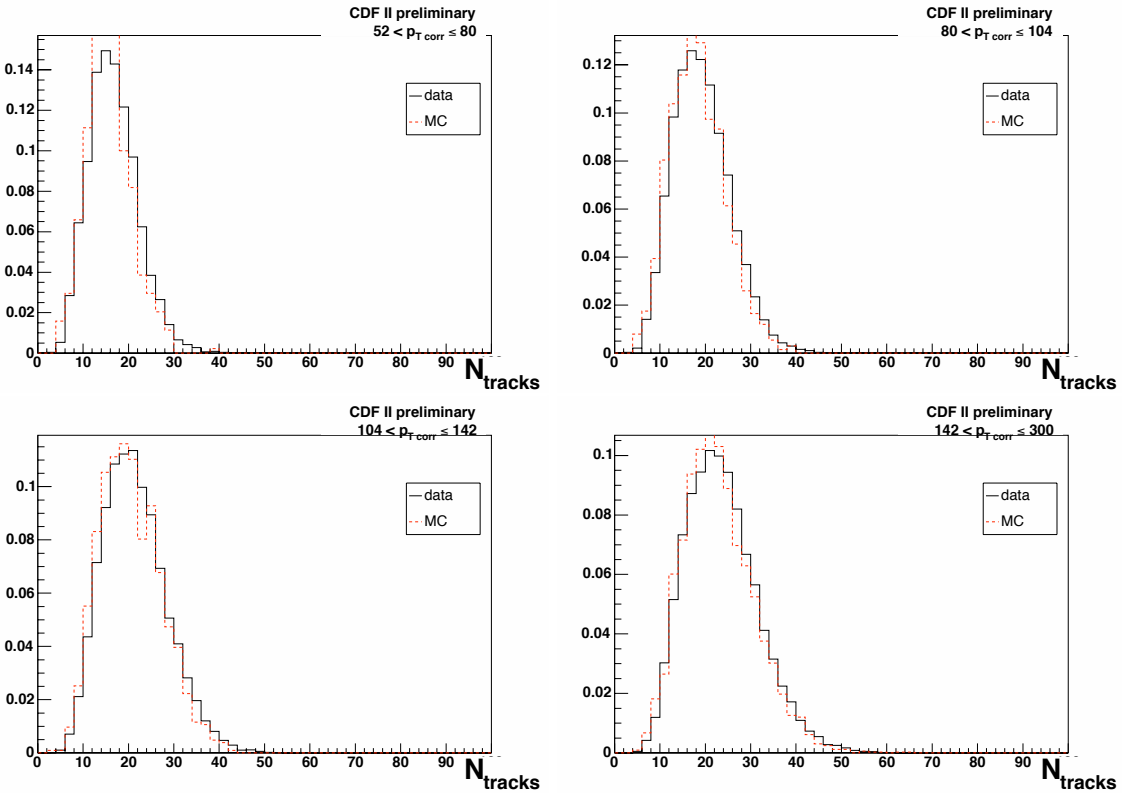

FIGURE B.12: Number of tracks inside a cone of 0.7 around jet axis for tagged jets.

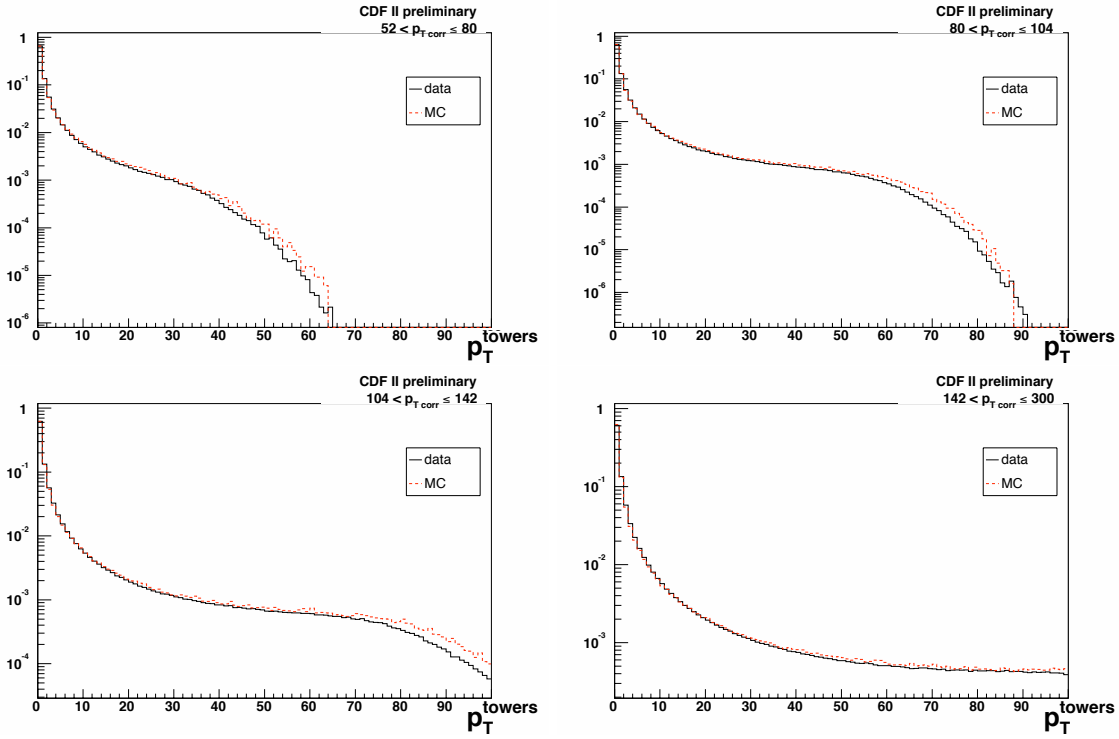

FIGURE B.13: Logarithmic plot of the p_T distribution of the towers inside a cone of 0.7 around jet axis for all jets.

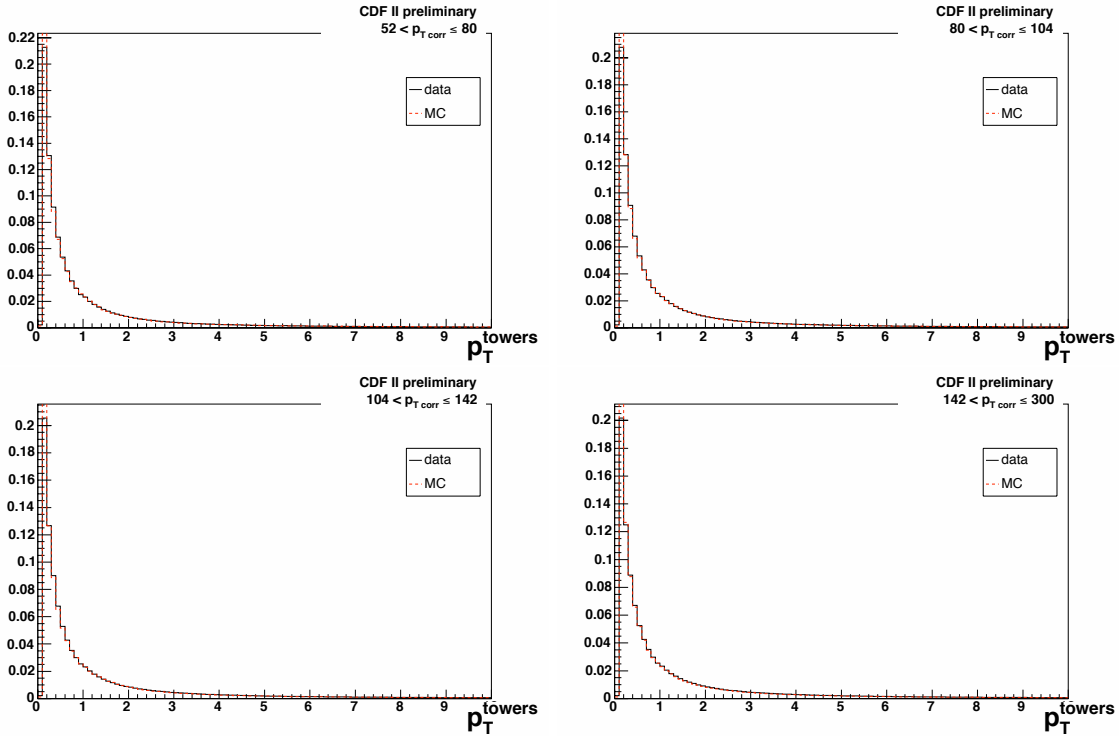

FIGURE B.14: Plot of the p_T distribution of the towers inside a cone of 0.7 around jet axis for all jets. Zoomed in to show the low p_T towers.

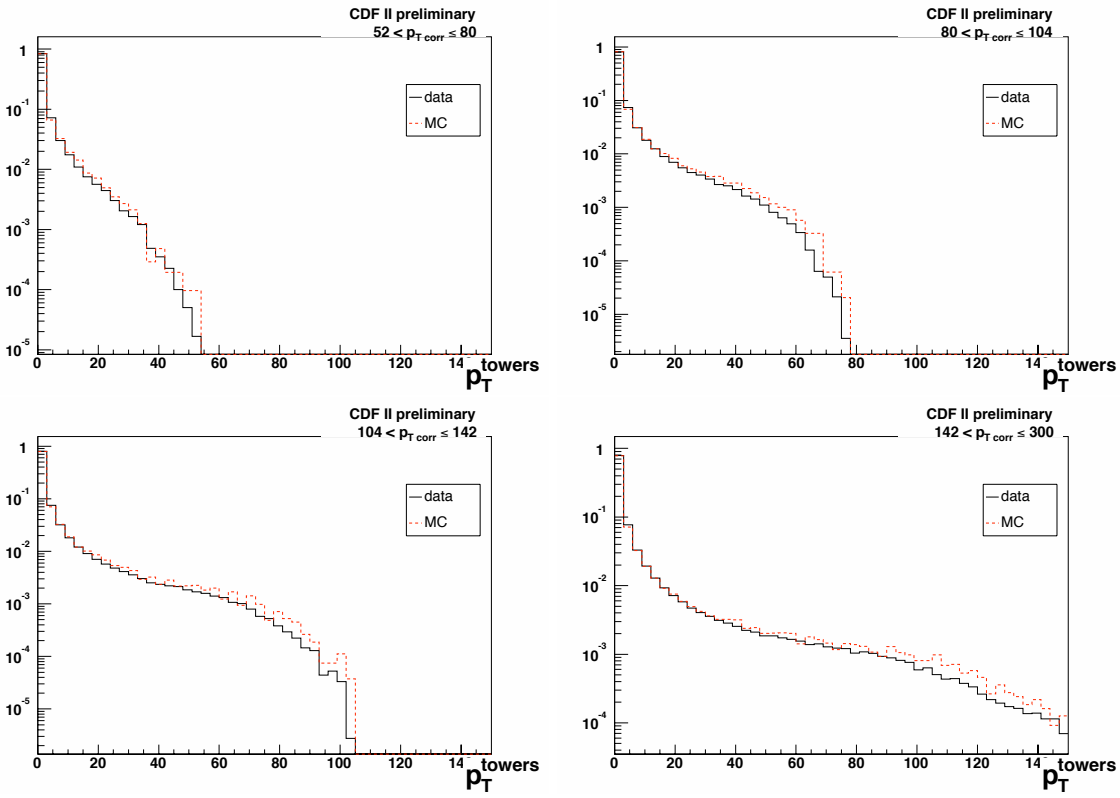

FIGURE B.15: Logarithmic plot of the p_T distribution of the towers inside a cone of 0.7 around jet axis for tagged jets.

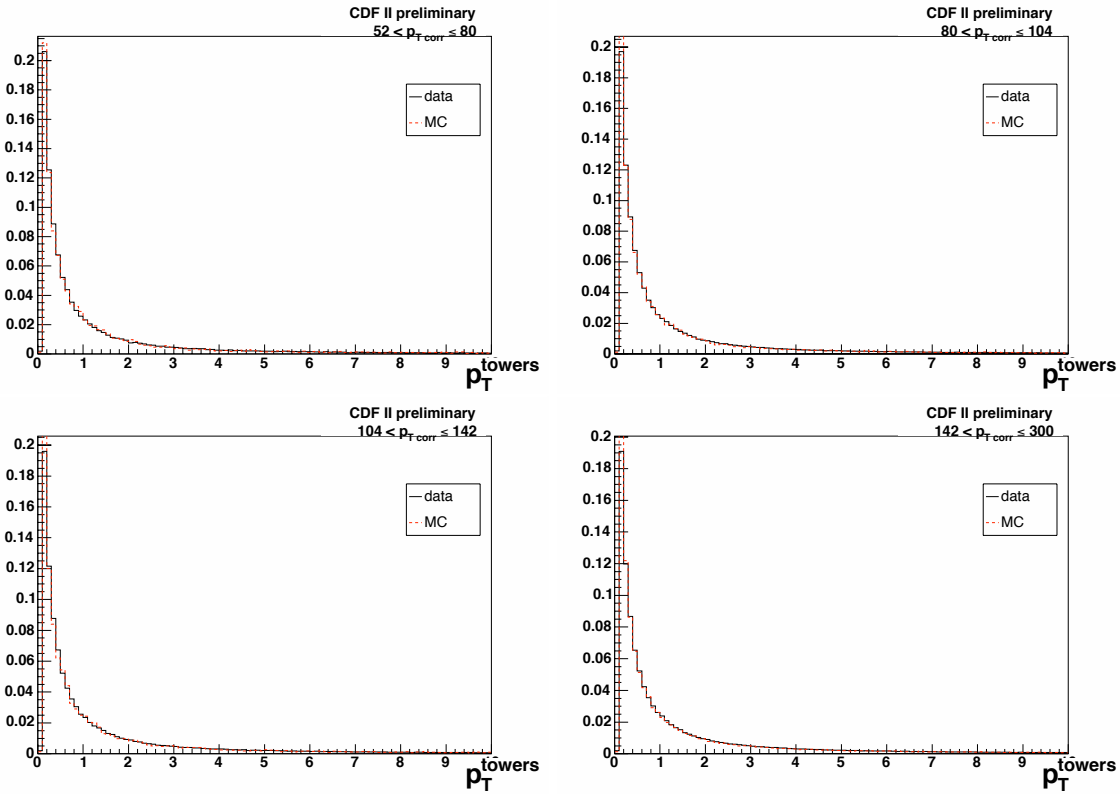

FIGURE B.16: Plot of the p_T distribution of the towers inside a cone of 0.7 around jet axis for tagged jets. Zoomed in to show the low p_T towers.

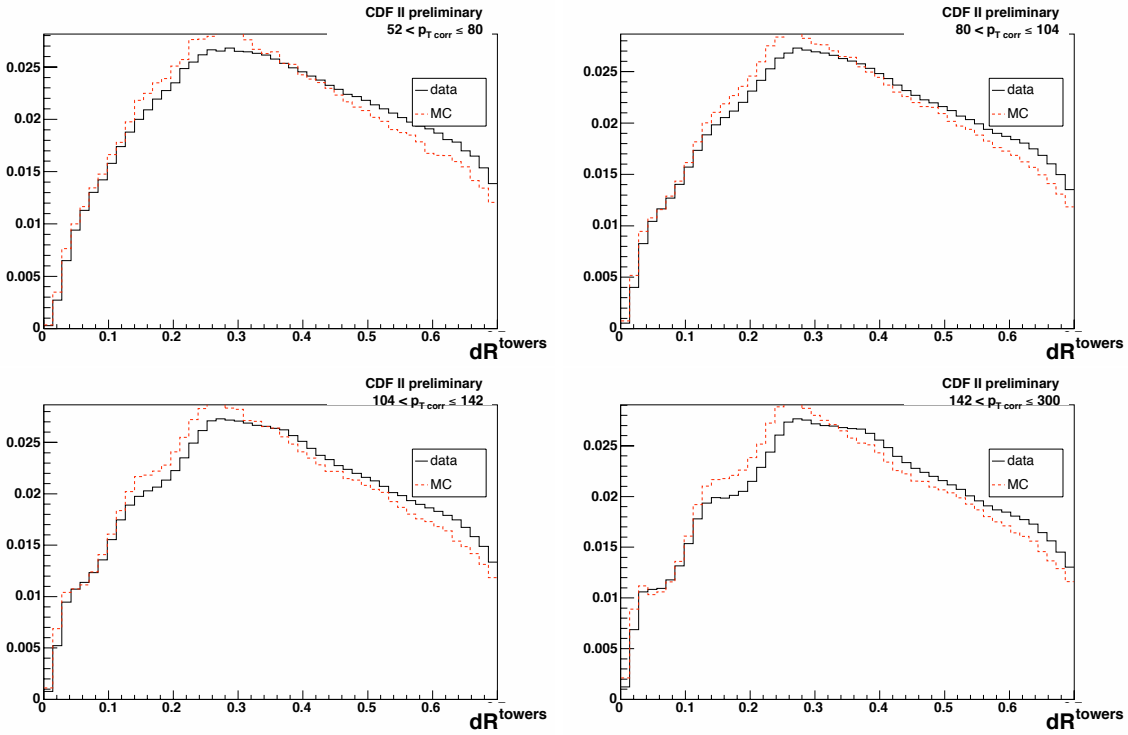

FIGURE B.17: Logarithmic plot of the dR distribution of the towers inside a cone of 0.7 around jet axis for all jets.

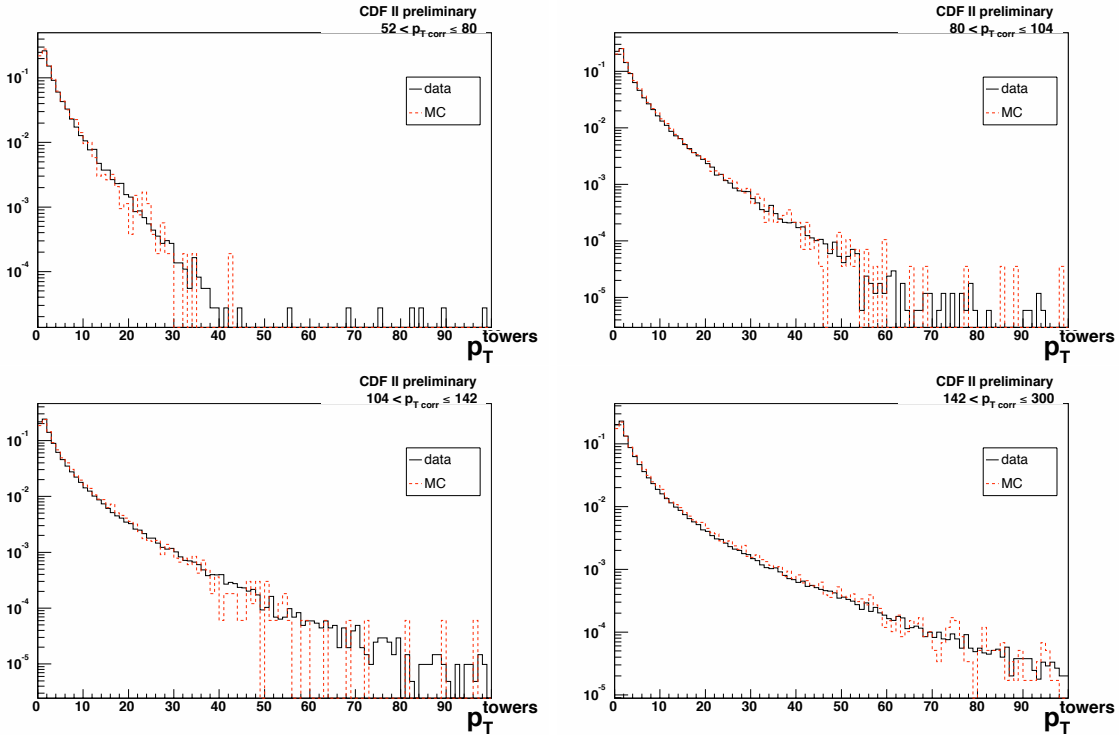

FIGURE B.18: Logarithmic plot of the p_T distribution of the tracks inside a cone of 0.7 around jet axis for all jets.

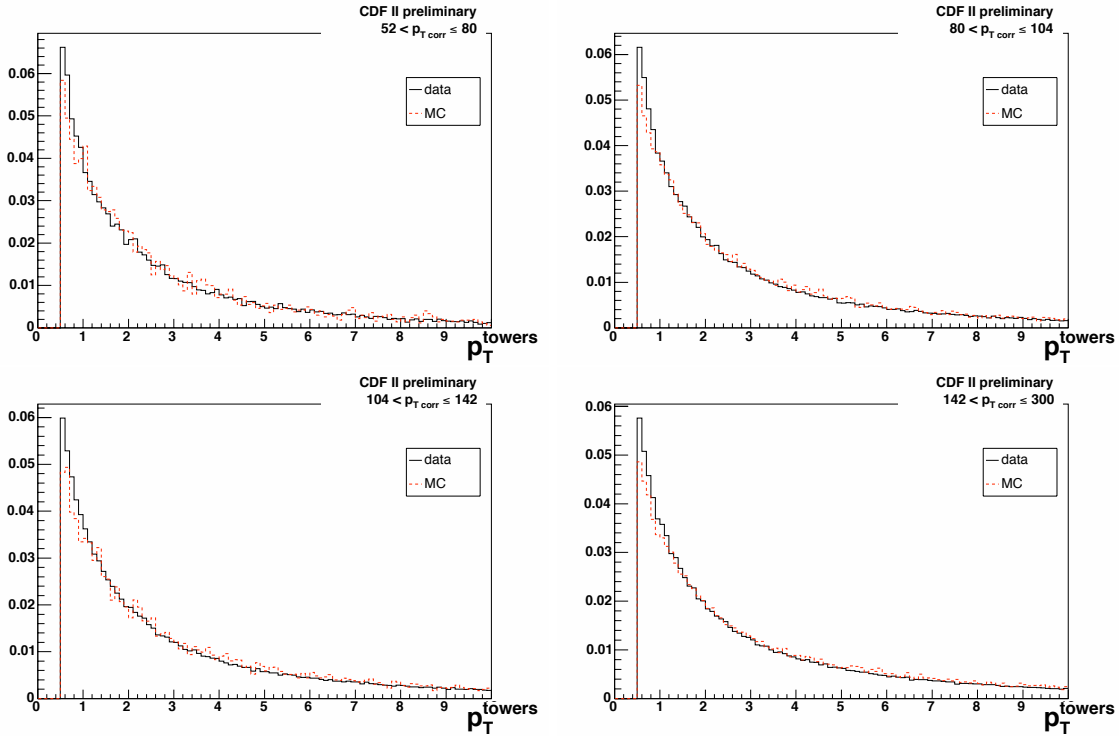

FIGURE B.19: Plot of the p_T distribution of the tracks inside a cone of 0.7 around jet axis for all jets. Zoomed in to show the low p_T tracks.

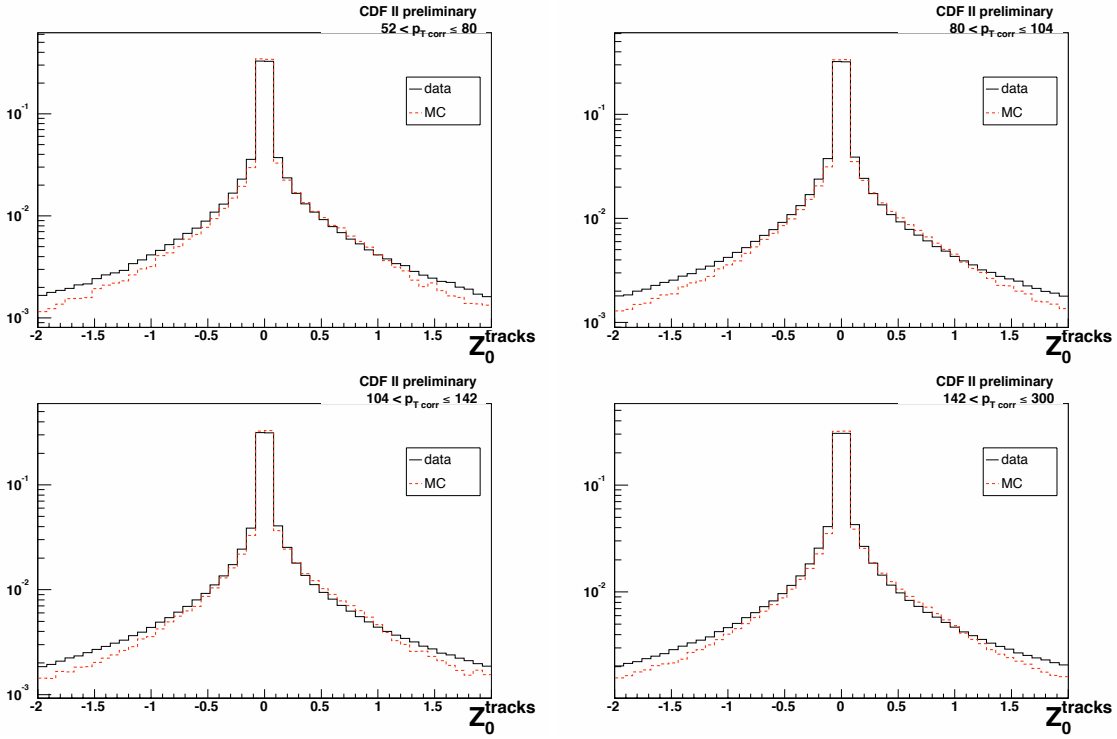

FIGURE B.20: Logarithmic plot of the Z_0 distribution of the tracks inside a cone of 0.7 around jet axis for all jets.

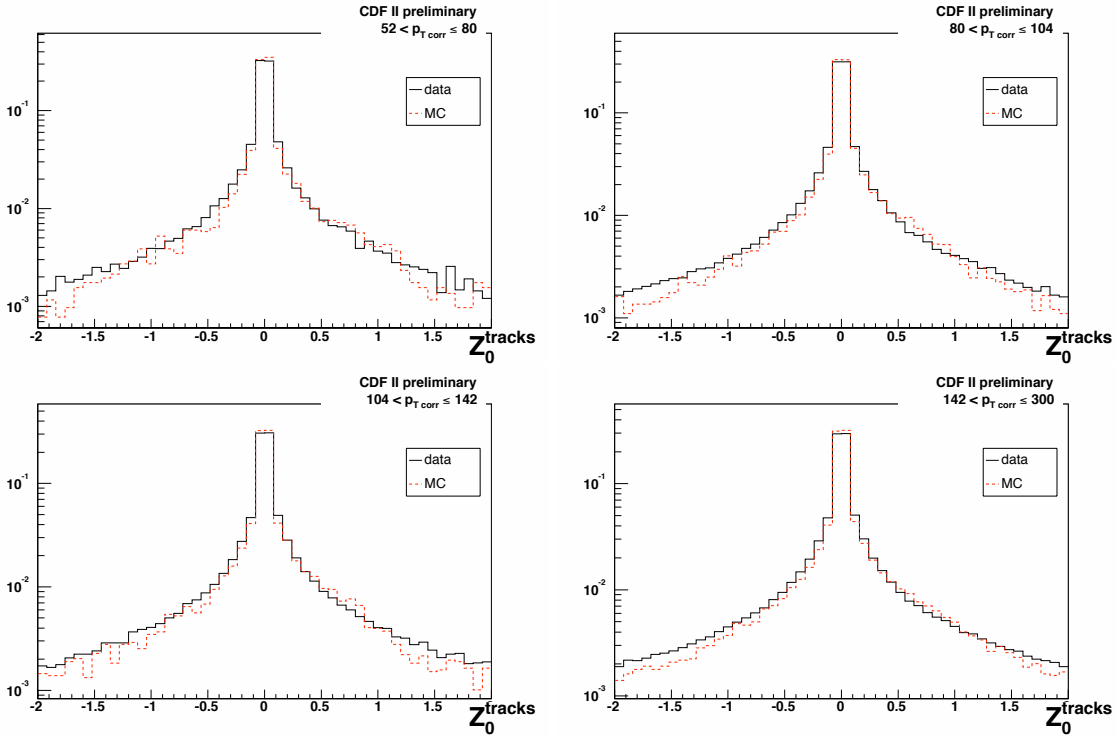

FIGURE B.21: Logarithmic plot of the Z_0 distribution of the tracks inside a cone of 0.7 around jet axis for tagged jets.

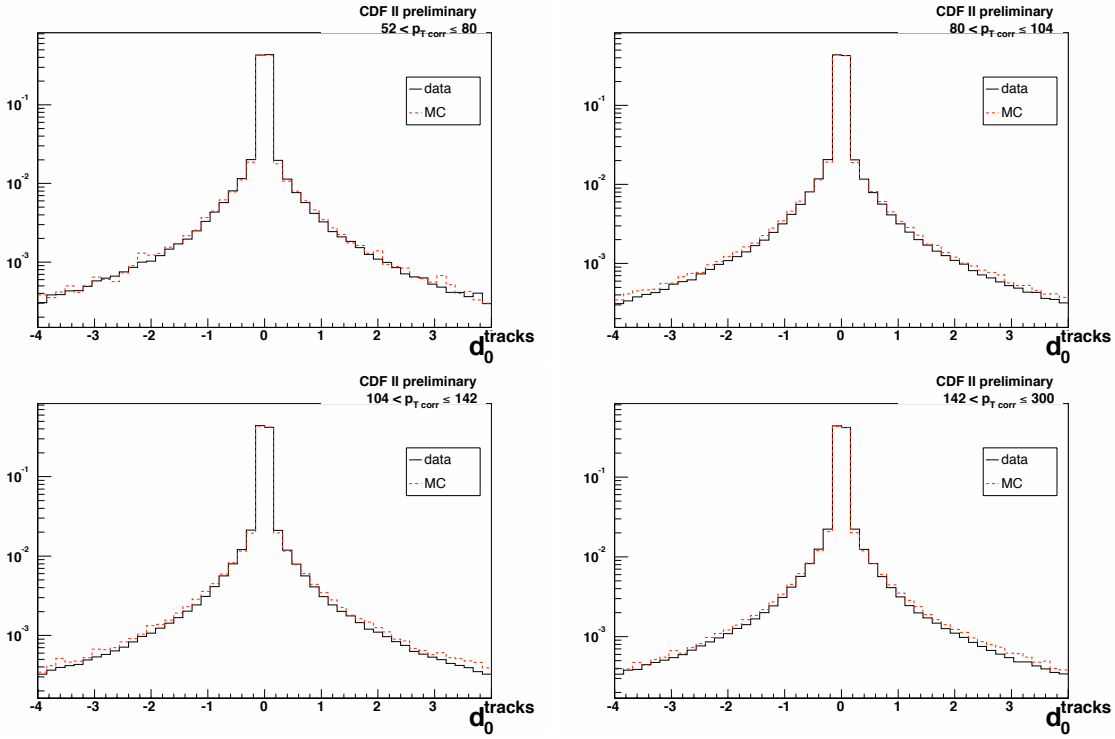

FIGURE B.22: Logarithmic plot of the d_0 distribution of the tracks inside a cone of 0.7 around jet axis for all jets.

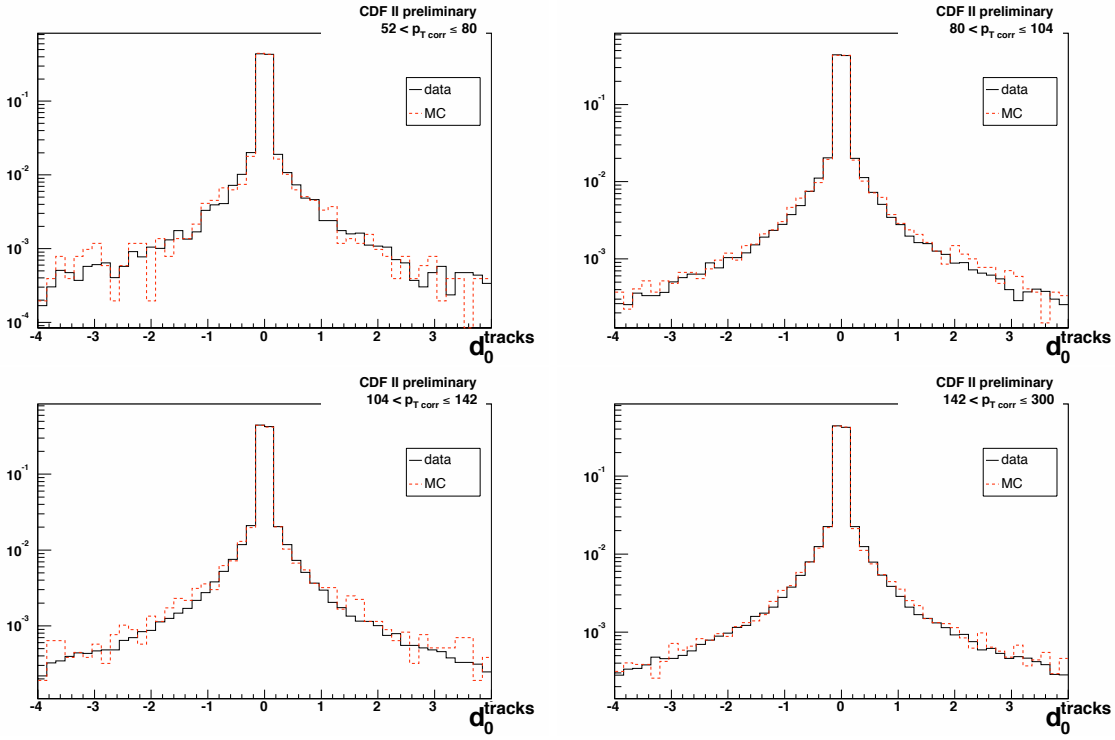

FIGURE B.23: Logarithmic plot of the d_0 distribution of the tracks inside a cone of 0.7 around jet axis for tagged jets.

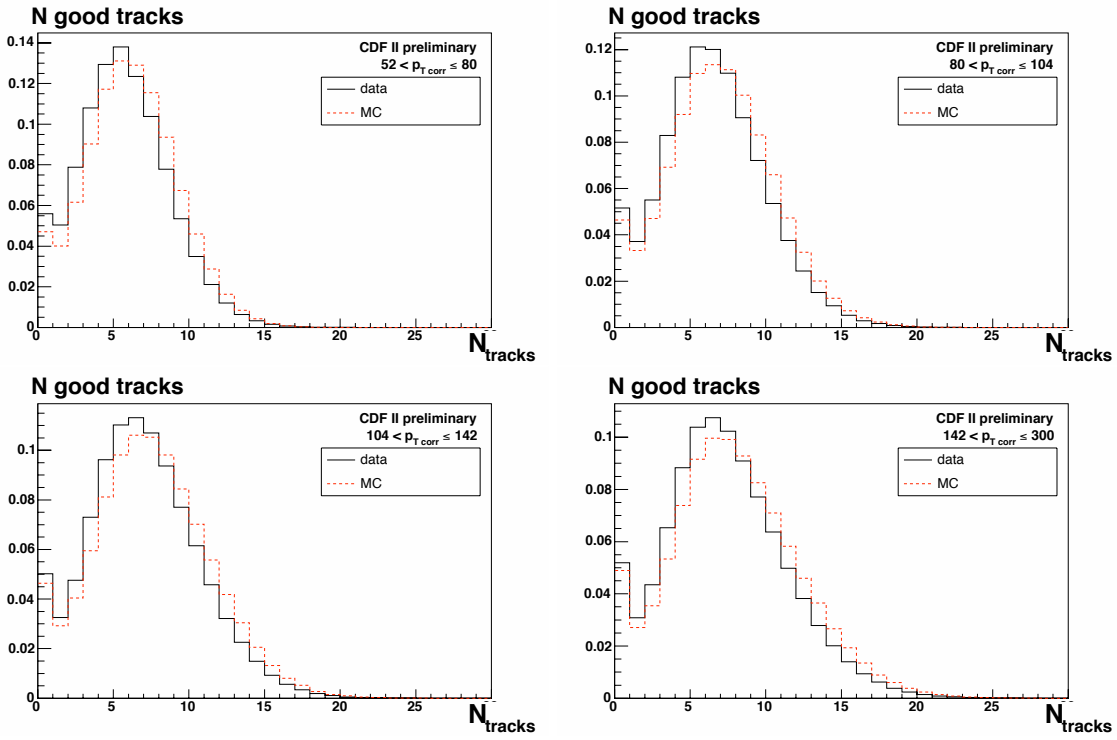

FIGURE B.24: Plot of the distribution of the number of good tracks inside the jet for all jets.

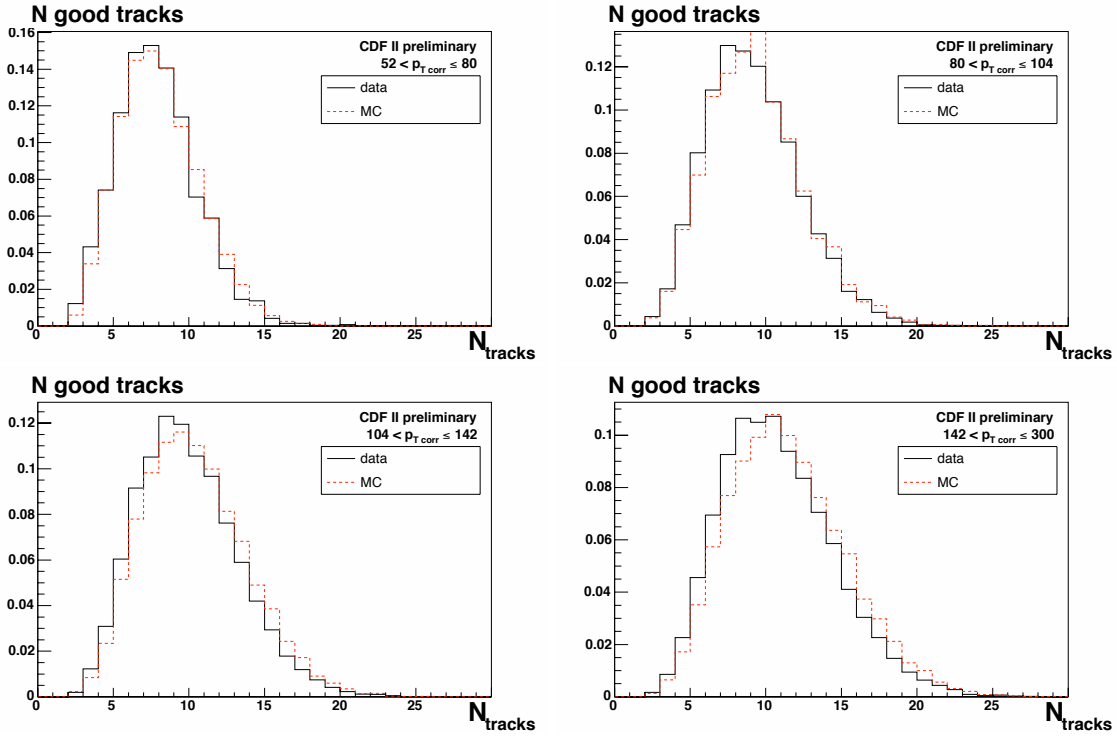

FIGURE B.25: Plot of the distribution of the number of good tracks inside the jet for tagged jets.

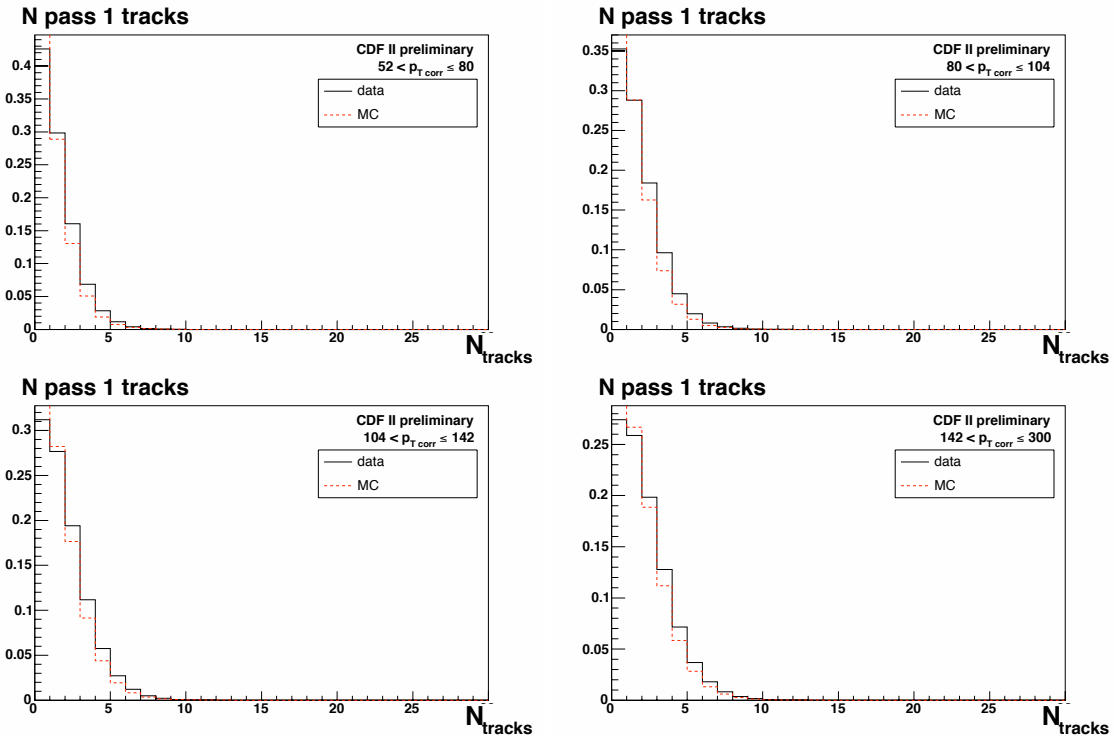

FIGURE B.26: Plot of the distribution of the number of pass1 tracks inside the jet for all jets.

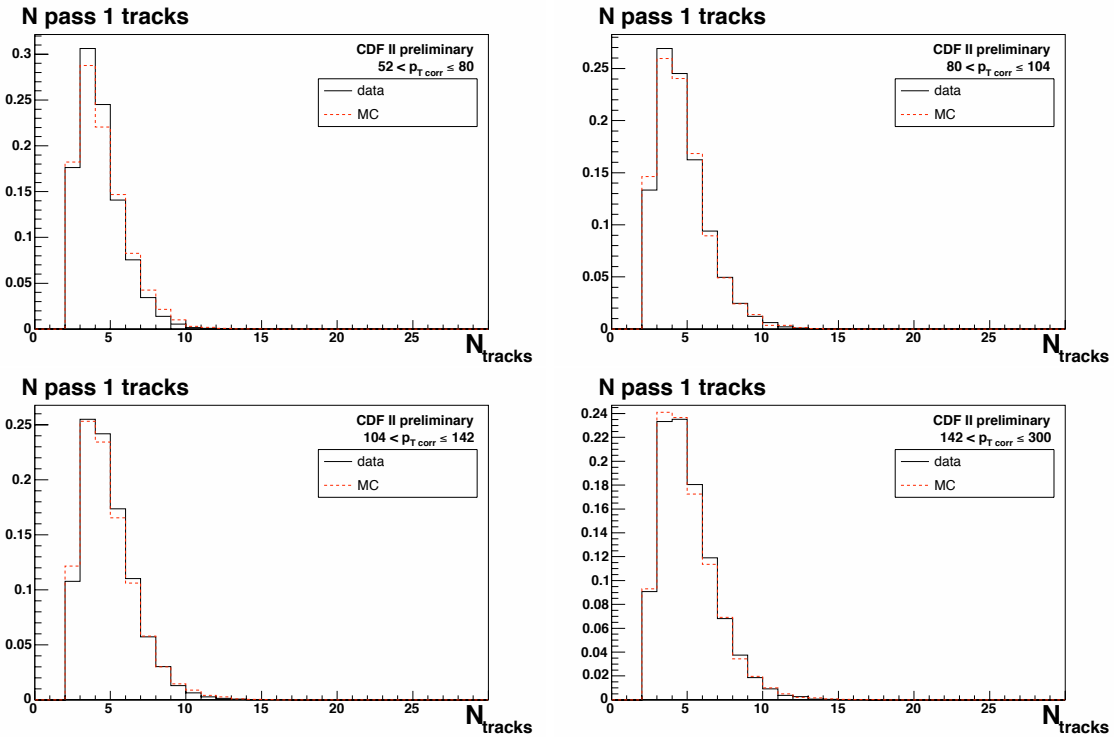

FIGURE B.27: Plot of the distribution of the number of pass1 tracks inside the jet for tagged jets.

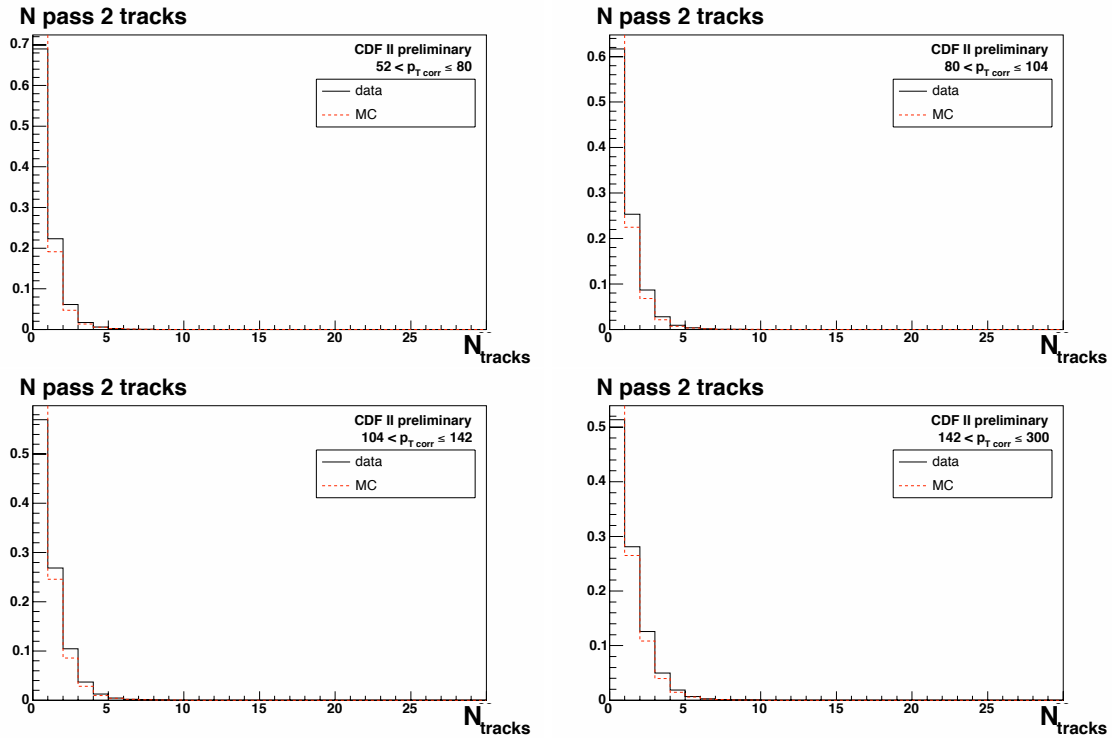

FIGURE B.28: Plot of the distribution of the number of pass 2 tracks inside the jet for all jets.

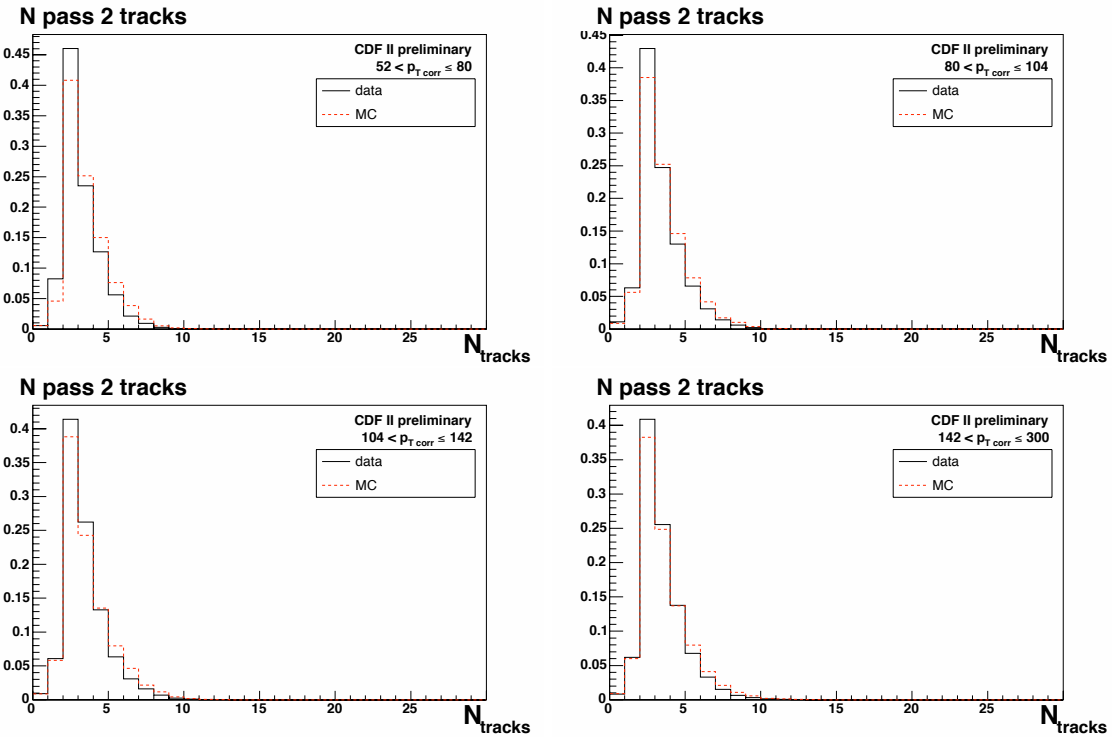

FIGURE B.29: Plot of the distribution of the number of pass 2 tracks inside the jet for tagged jets.

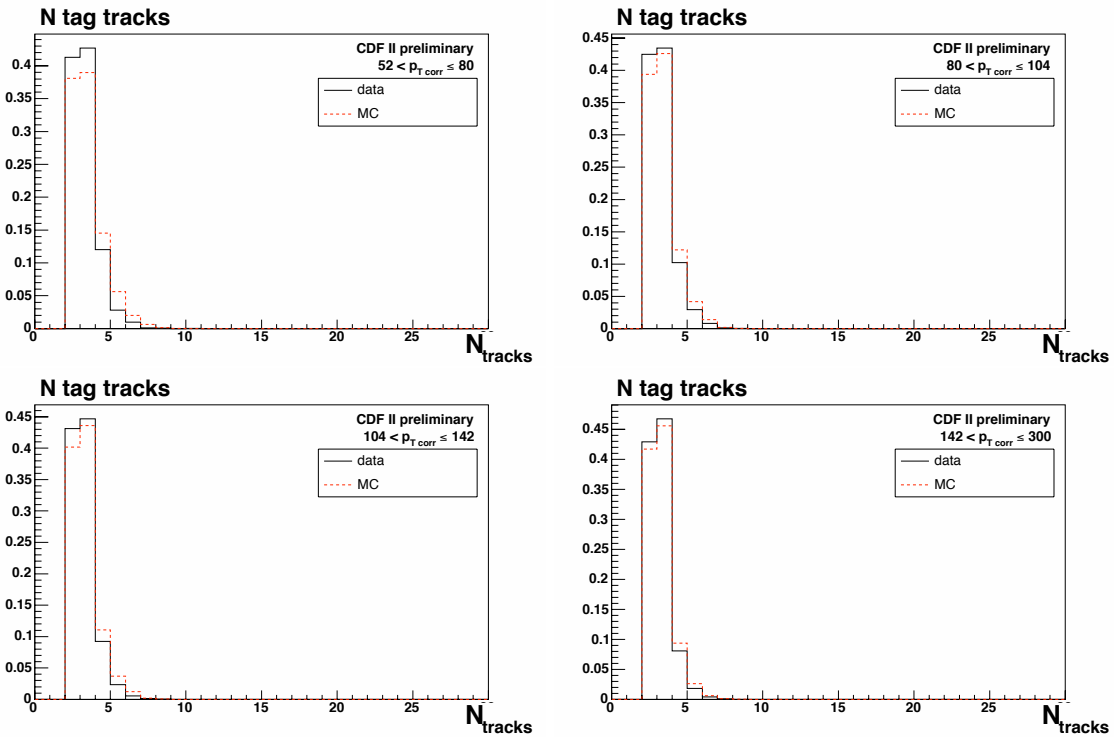

FIGURE B.30: Plot of the distribution of the number of tracks inside the jet used for the tag for tagged jets.

Appendix C

Inclusive jet shapes

This appendix compares the hadron level inclusive jet shapes with the previously published inclusive jet shapes [2]. The published results used about $170\,\mathrm{pb}^{-1}$ of CDF Run II data. All runs above and including 166328 where excluded. The other differences with respect to my analysis are

- Uses earlier datasets (5.1.1 instead of 5.3.1) and a different version of the off-line reconstruction software (5.3.1pre4 instead of 5.3.3_nt)
- Rapidity cut: $0.1 < |Y| < 0.7$. The central rapidity gap is excluded
- $|Z_{\mathrm{vtx}}| < 60\,\mathrm{cm}$
- jet p_T corrections: $p_T^{\mathrm{corr}} = 0.7509 + 1.148 p_T - 2.5 \cdot 10^{-4} p_T^2$
- missing E_T significance $< 3.5\,\mathrm{GeV}^{-1/2}$ for all datasets

I changed my analysis to respect these cuts and use the same binning as for [2]. The hadron level inclusive jet shapes is given by

$$\Psi_{\mathrm{had}}(r) = C_{\mathrm{had}}(r)\Psi_{\mathrm{det}}(r) \tag{C.1}$$

Where the hadron level correction C_{had} are calculated in the same way as my b-jet hadron level corrections

$$C_{\mathrm{had}}(r) = \frac{\Psi_{\mathrm{had}}^{\mathrm{MC}}(r)}{\Psi_{\mathrm{det}}^{\mathrm{MC}}(r)} \tag{C.2}$$

It is also possible to define an average fraction of the track multiplicity inside the jet. This is defined as

$$\pi(r) = \frac{1}{N_{\mathrm{jets}}} \frac{1}{\Delta r} \sum_{\mathrm{jets}} \frac{N_{\mathrm{tracks}}(r - \Delta r/2, r + \Delta r/2)}{N_{\mathrm{tracks}}(0, R)} \tag{C.3}$$

I compare the published results with my results applying all cuts as above.
The results are also compared to my MC simulation at hadron and detector level.

Figure C.1 and C.2 show the comparison at calorimeter and track level respectively of the raw, detector level, jet shapes as measured by the MC simulations. The agreement for calorimeter level shapes is very good between my MC and the published MC values. This shows that my reconstruction of the jet shapes at calorimeter level is done in the same way as the published analysis. The agreement for the track shapes is not so good for a few bins, in particular the three highest p_T bins. This could be investigated further but as for my analysis the systematic error on the b-jet shape due to the difference in the final shapes when starting from tracks or from

calorimeter towers is not a dominant error. I therefore decided not to investigate this any further.

Figure C.3 and C.4 show the comparison at calorimeter and track level respectively of the raw, detector level, jet shapes in data (red points) and MC (black line). Figure C.6 shows the calorimeter level raw shapes where an additional cut is applied on the p_T of the towers which must be above 0.5 GeV.

Figure C.7 and C.8 show the same comparisons, for calorimeter tower and tracks respectively, but comparing the differential jet shapes.

Figure C.10 and C.11 show the hadron level jet shapes starting from raw calorimetric and track jet shapes in Pythia Tune A MC respectively. These plots show that the hadron level corrections are well understood as the MC reconstructed hadron level shapes agree perfectly with the MC hadron level shapes.

Figure C.12 and C.13 show the hadron level jet shapes starting from calorimetric towers and tracks respectively.

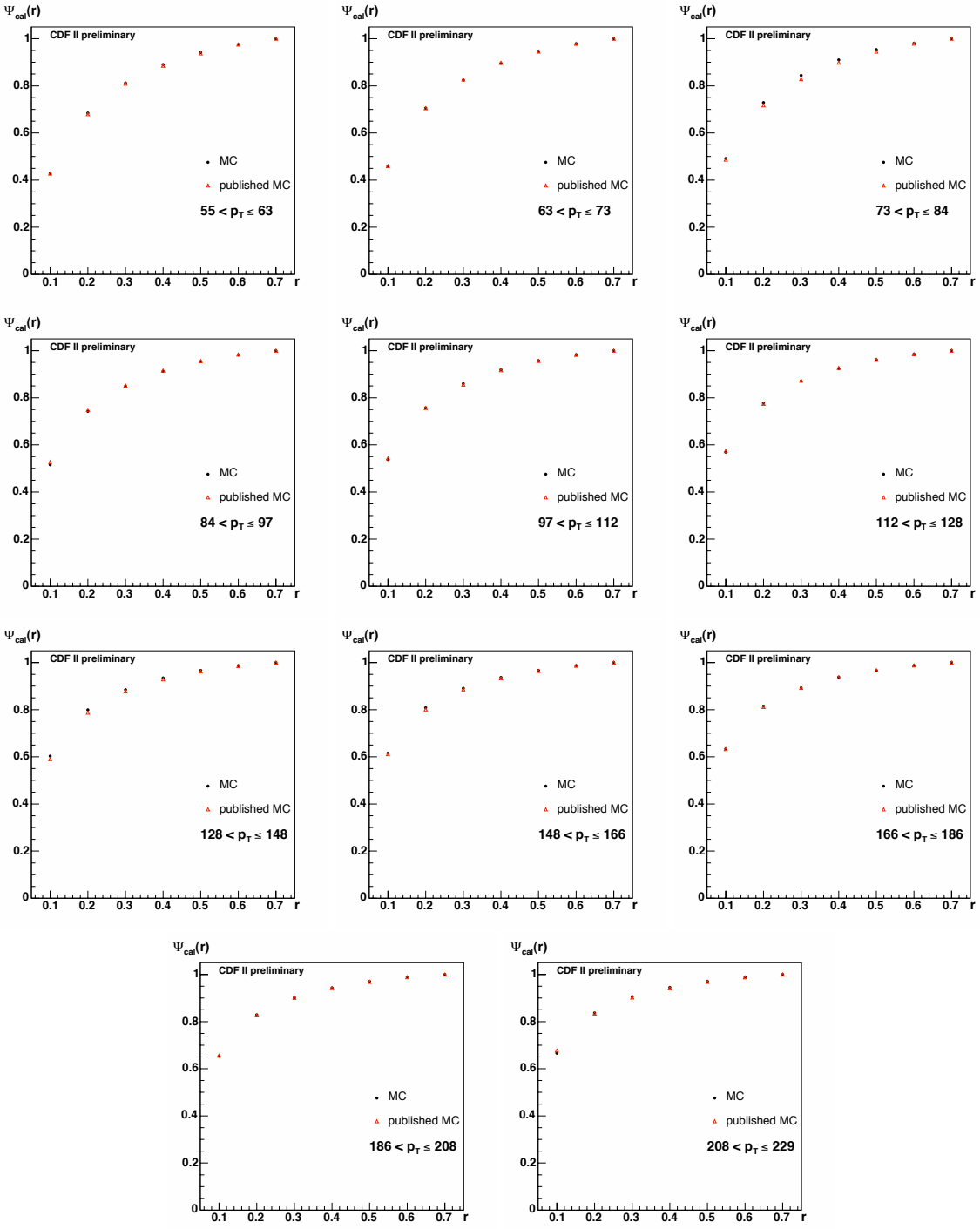

FIGURE C.1: Calorimeter level inclusive jet shapes. Comparing my detector level MC (black points) with the published ones (red points). No errors are shown.

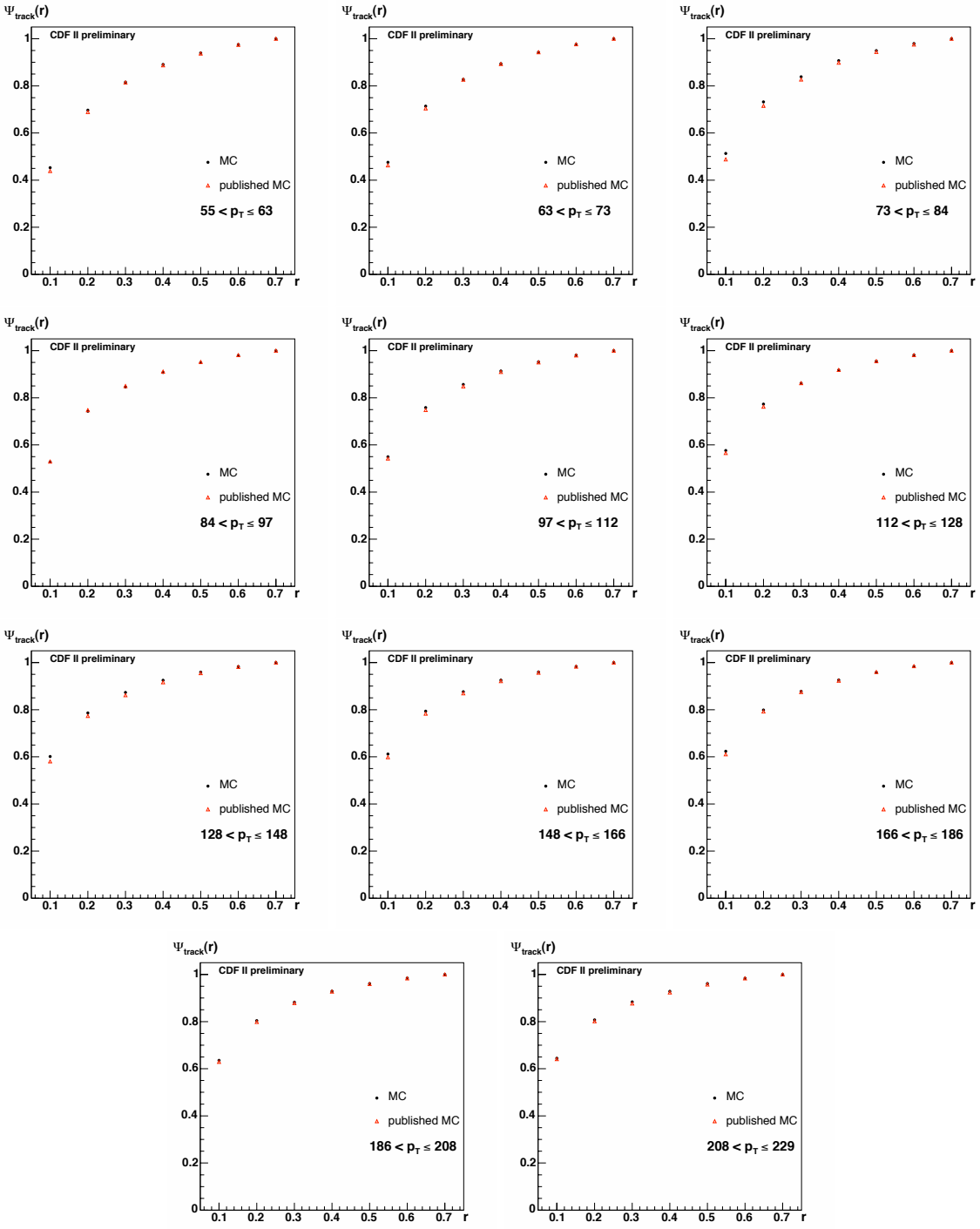

FIGURE C.2: Track level inclusive jet shapes. Comparing my detector level MC (black points) with the published ones (red points). No errors are shown.

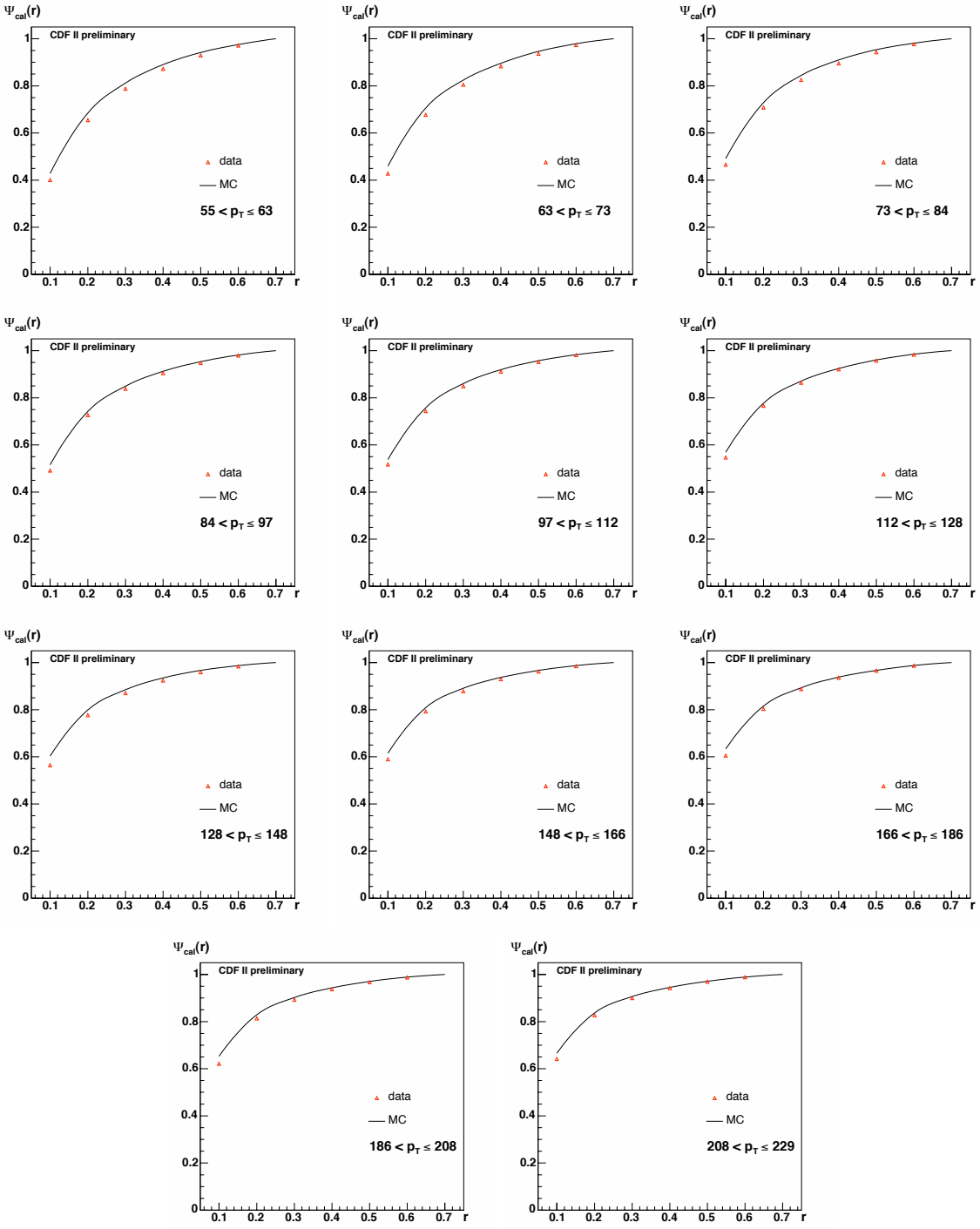

FIGURE C.3: Calorimeter level inclusive jet shapes. Comparing my data (red points) to the level MC (black line). Only the statistical errors are shown.

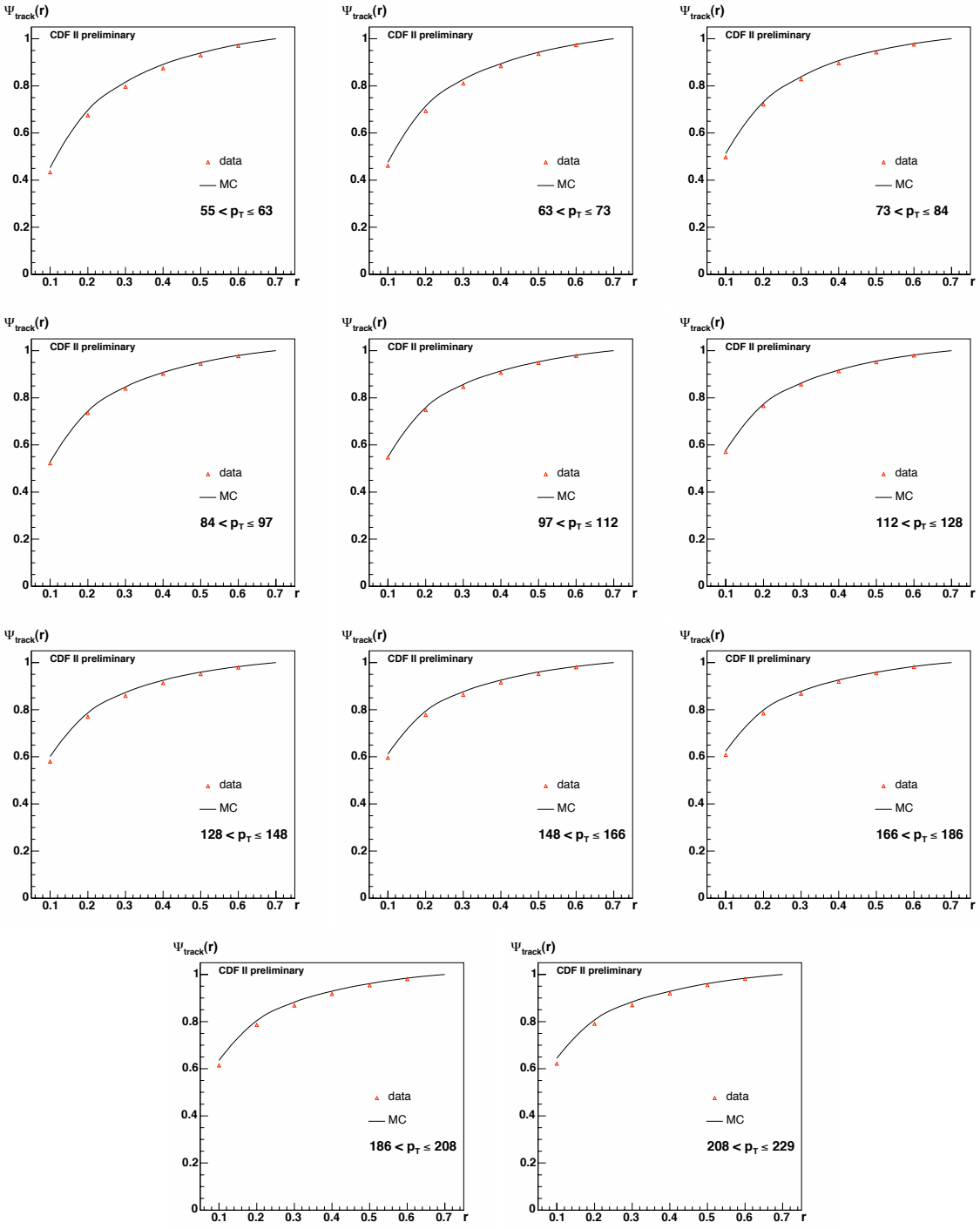

FIGURE C.4: Track level inclusive jet shapes. Comparing my data (red points) to the level MC (black line). Only the statistical errors are shown.

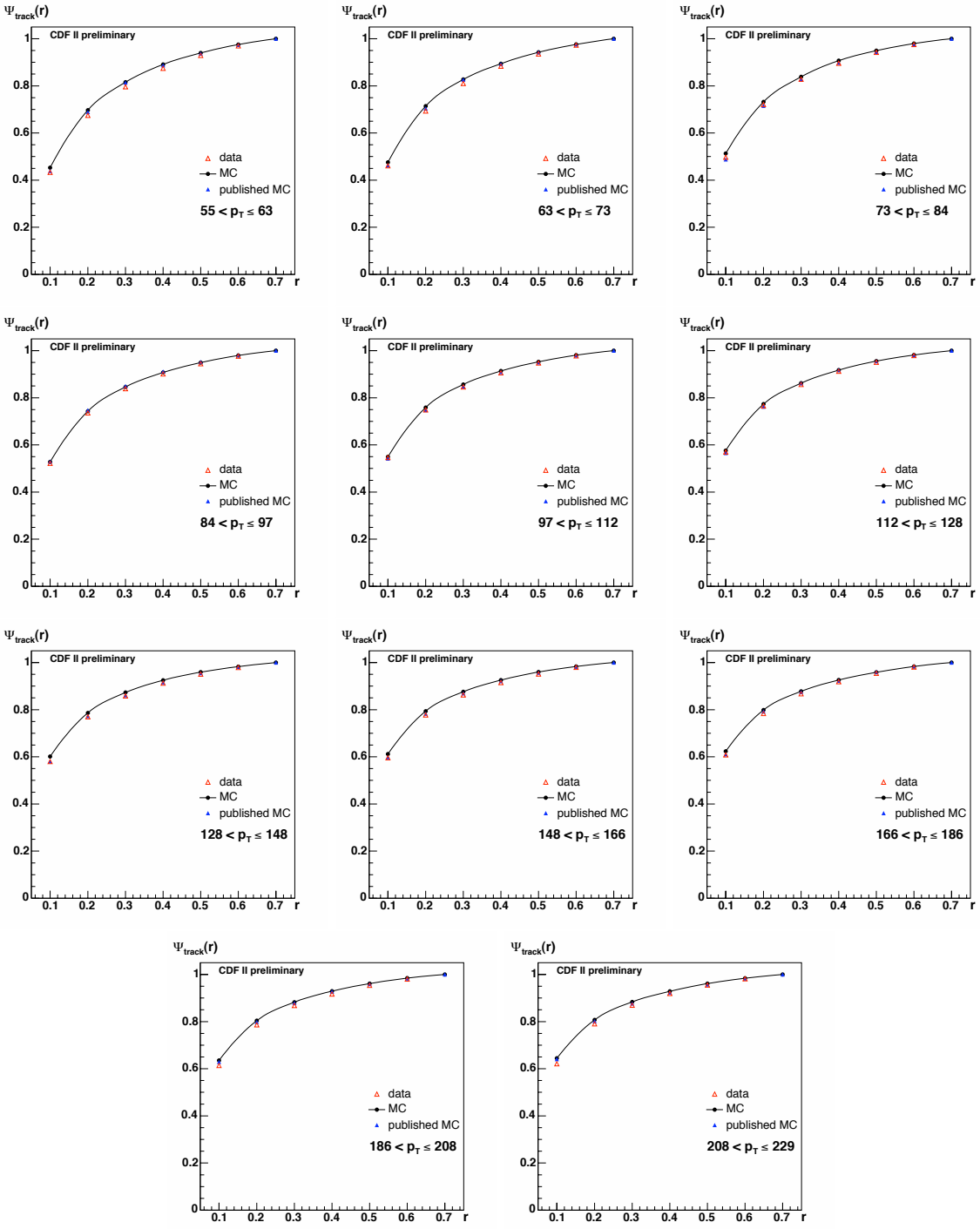

FIGURE C.5: Track level inclusive jet shapes. Comparing my data (red points) to the MC (black line) and the published MC (blue points). Only the statistical errors are shown.

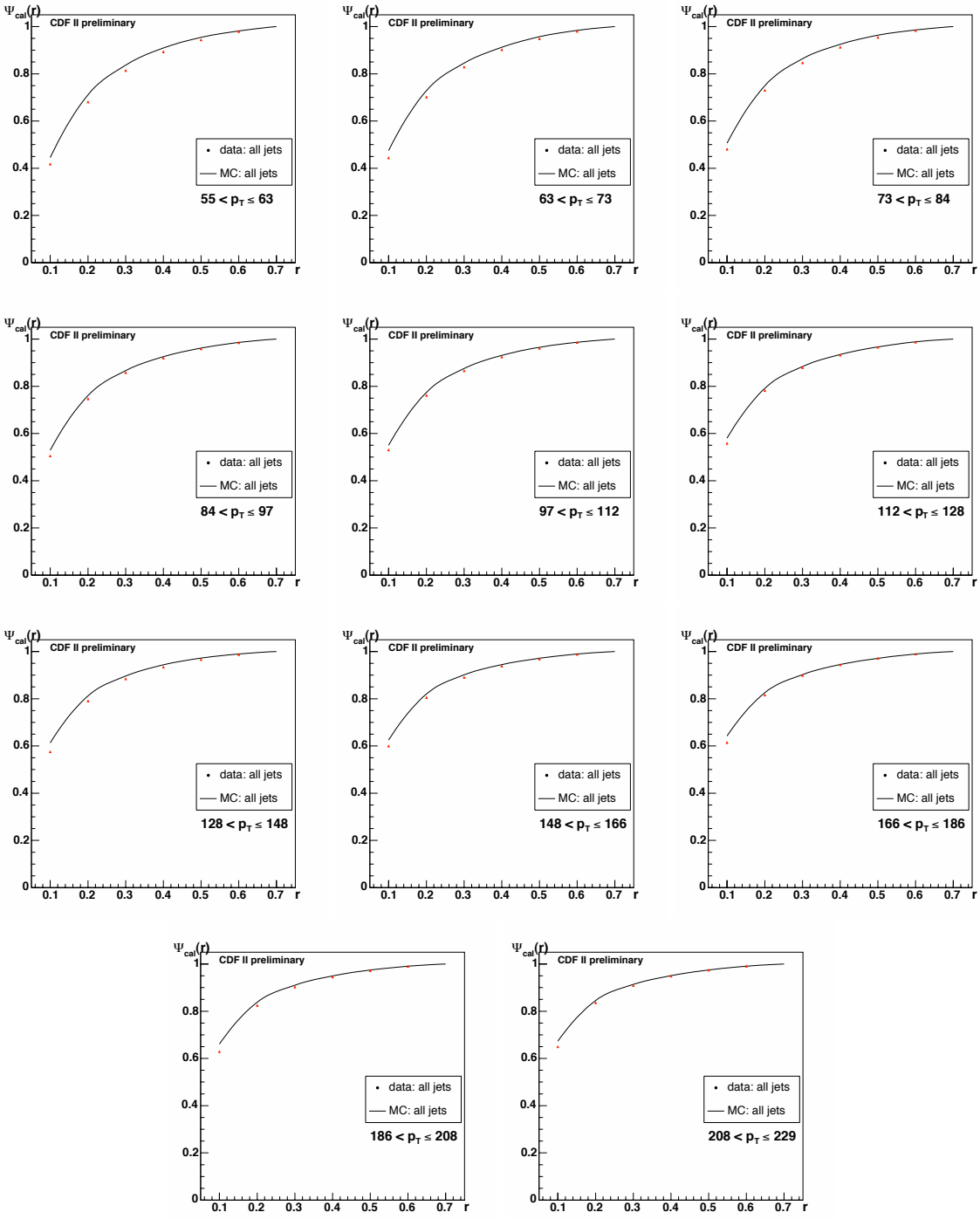

FIGURE C.6: Calorimeter level inclusive jet shapes with a p_T cut on the towers at 0.5 GeV. Comparing my data (red points) to the MC (black line). Only the statistical errors are shown.

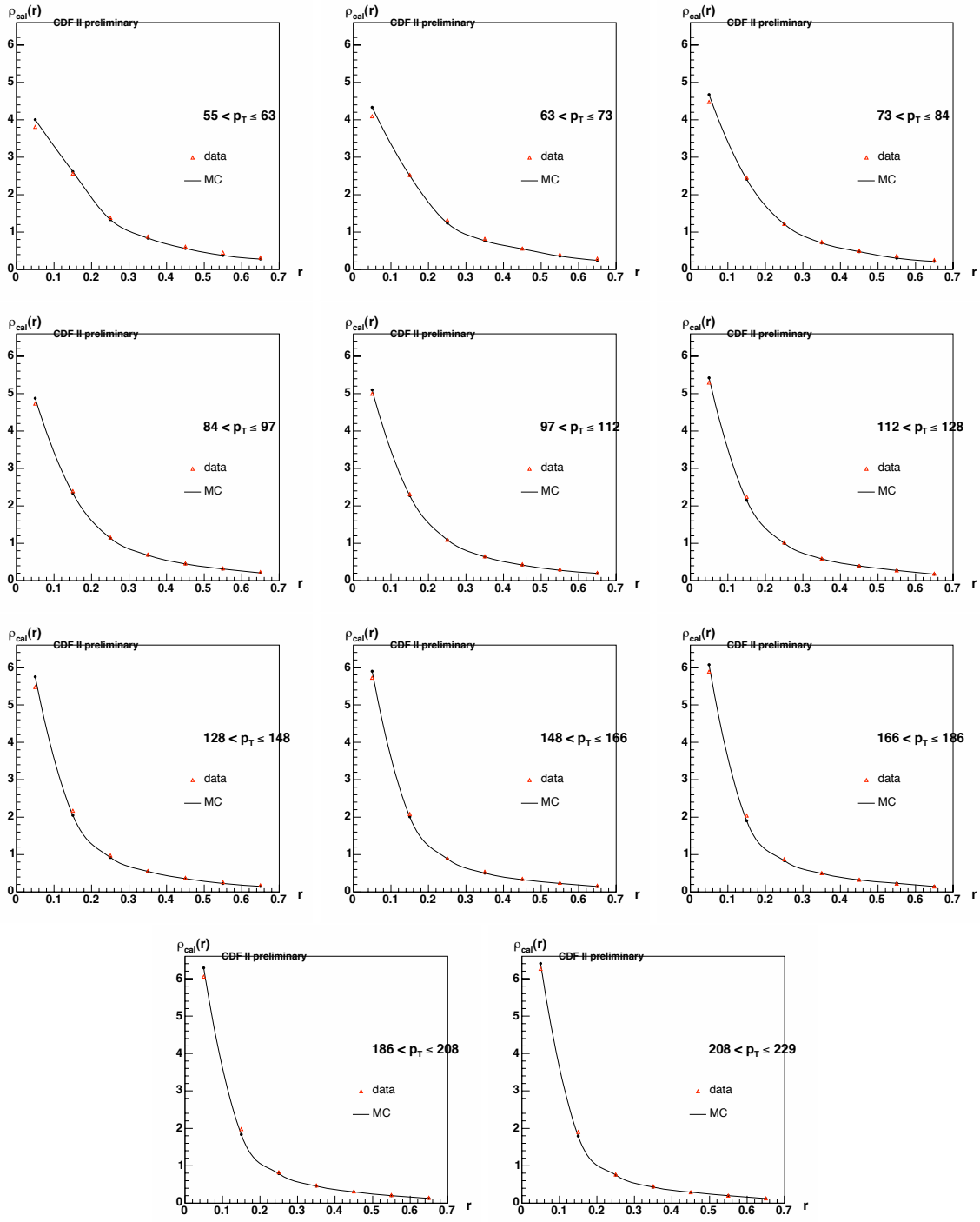

FIGURE C.7: Calorimeter level inclusive differential jet shapes. Comparing my data (red points) to the level MC (black line). Only the statistical errors are shown.

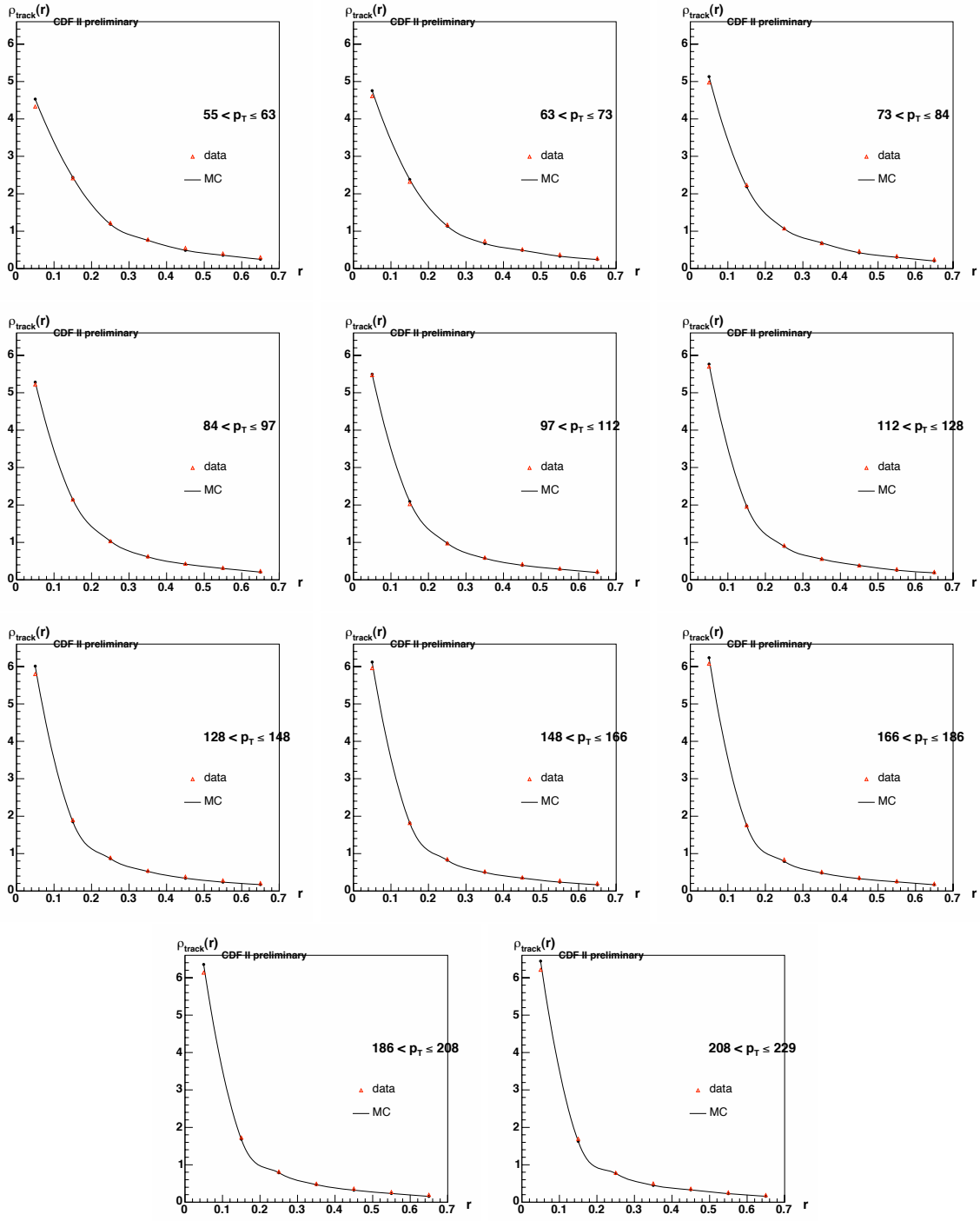

FIGURE C.8: Track level inclusive differential jet shapes. Comparing my data (red points) to the level MC (black line). Only the statistical errors are shown.

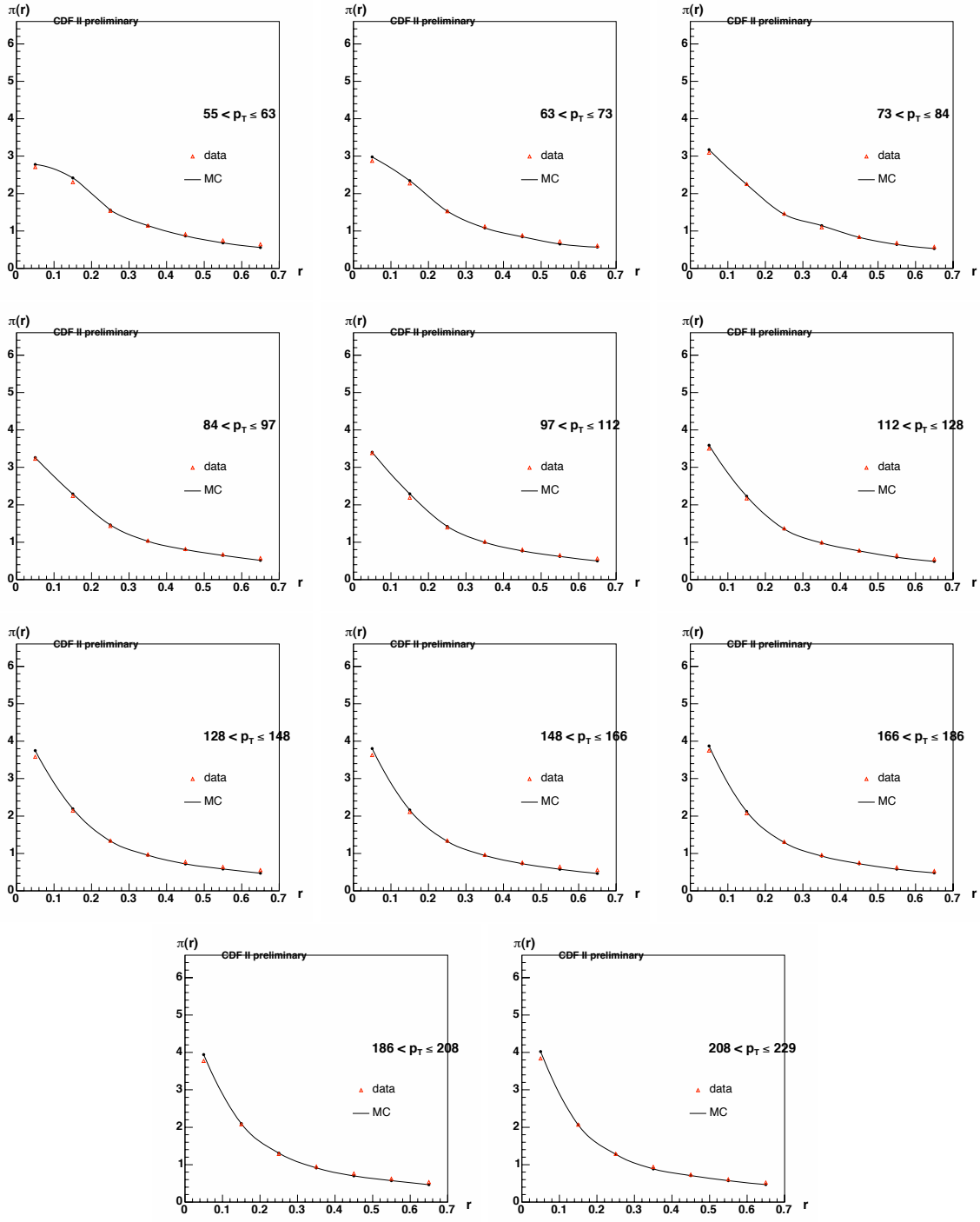

FIGURE C.9: Track multiplicity fraction. Comparing my data (red points) to the level MC (black line). Only the statistical errors are shown.

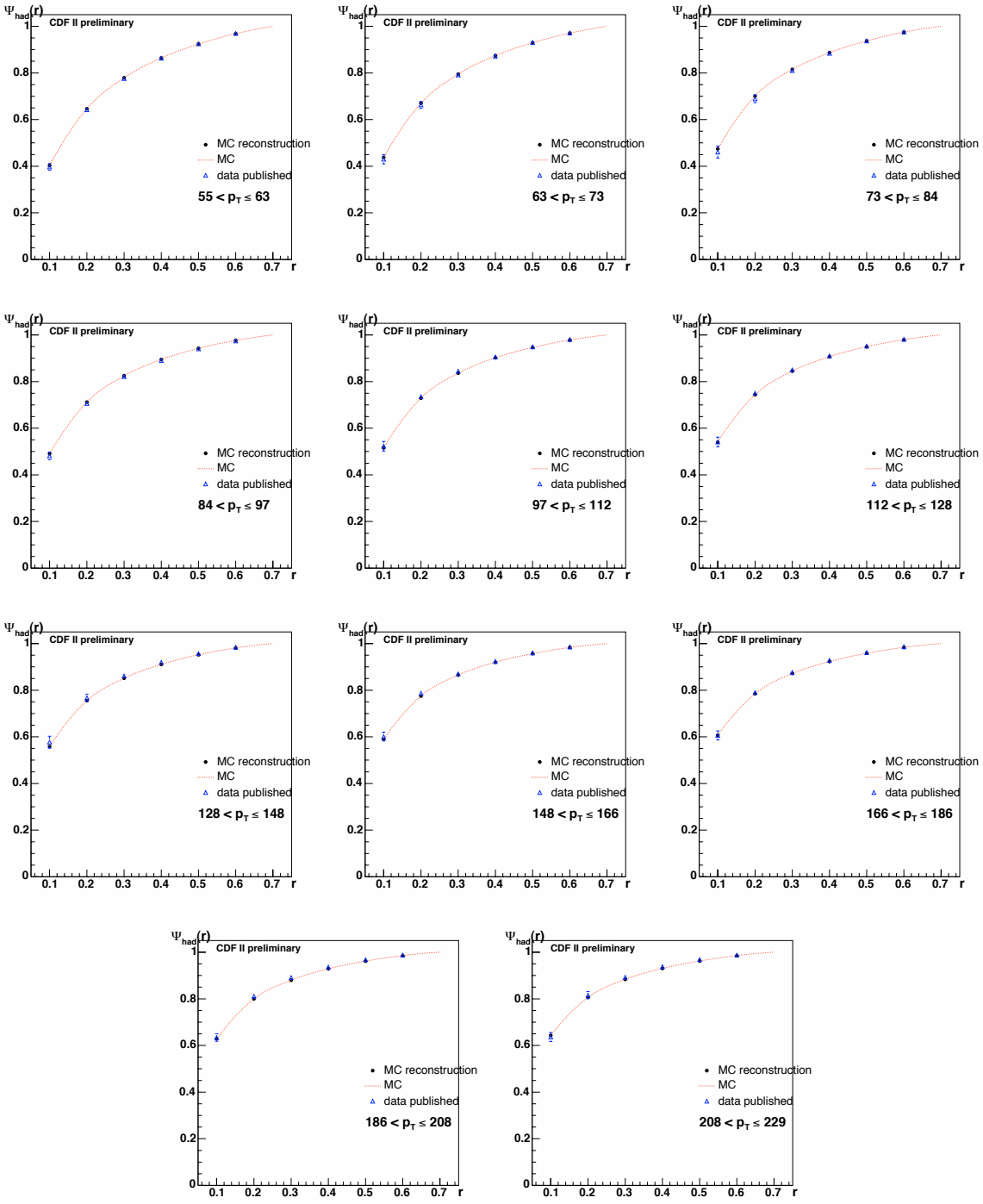

FIGURE C.10: Hadron level inclusive jet shapes. Comparing my reconstructed MC starting from raw shapes using calorimetric towers (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

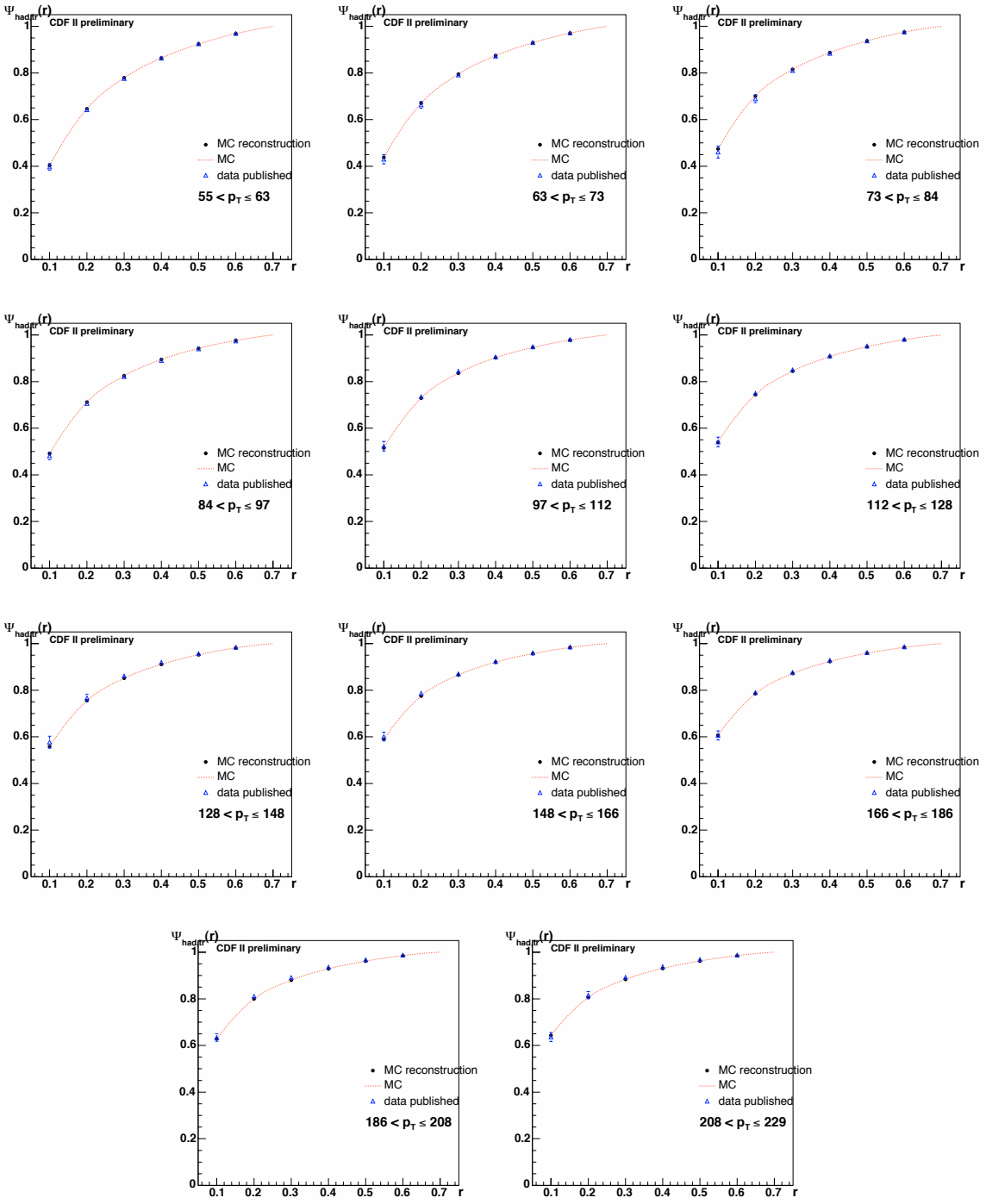

FIGURE C.11: Hadron level inclusive jet shapes. Comparing my reconstructed MC starting from raw shapes using tracks (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

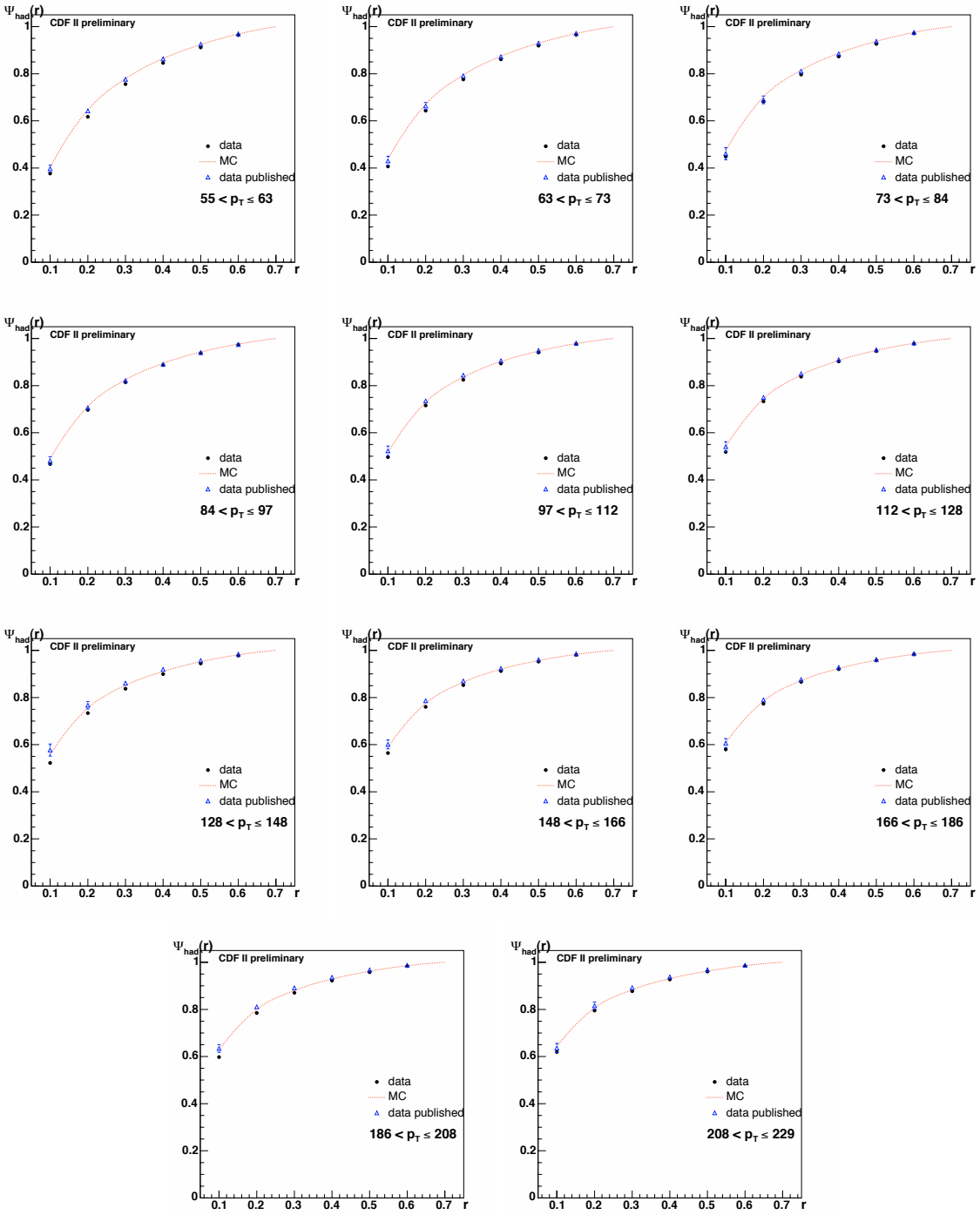

FIGURE C.12: Hadron level inclusive jet shapes. Comparing my results reconstructed from raw calorimeter towers (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

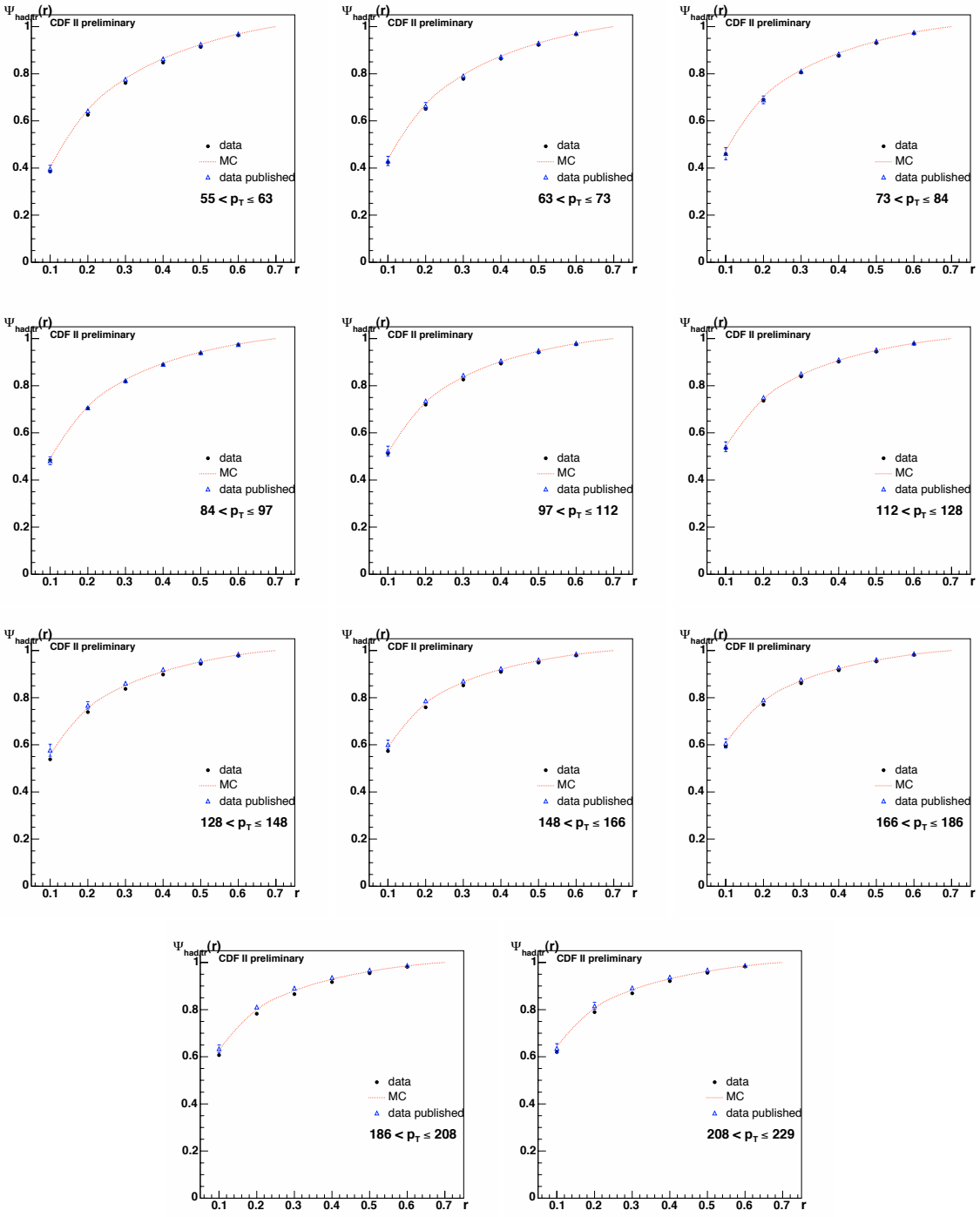

FIGURE C.13: Hadron level inclusive jet shapes. Comparing my results reconstructed from raw tracks (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

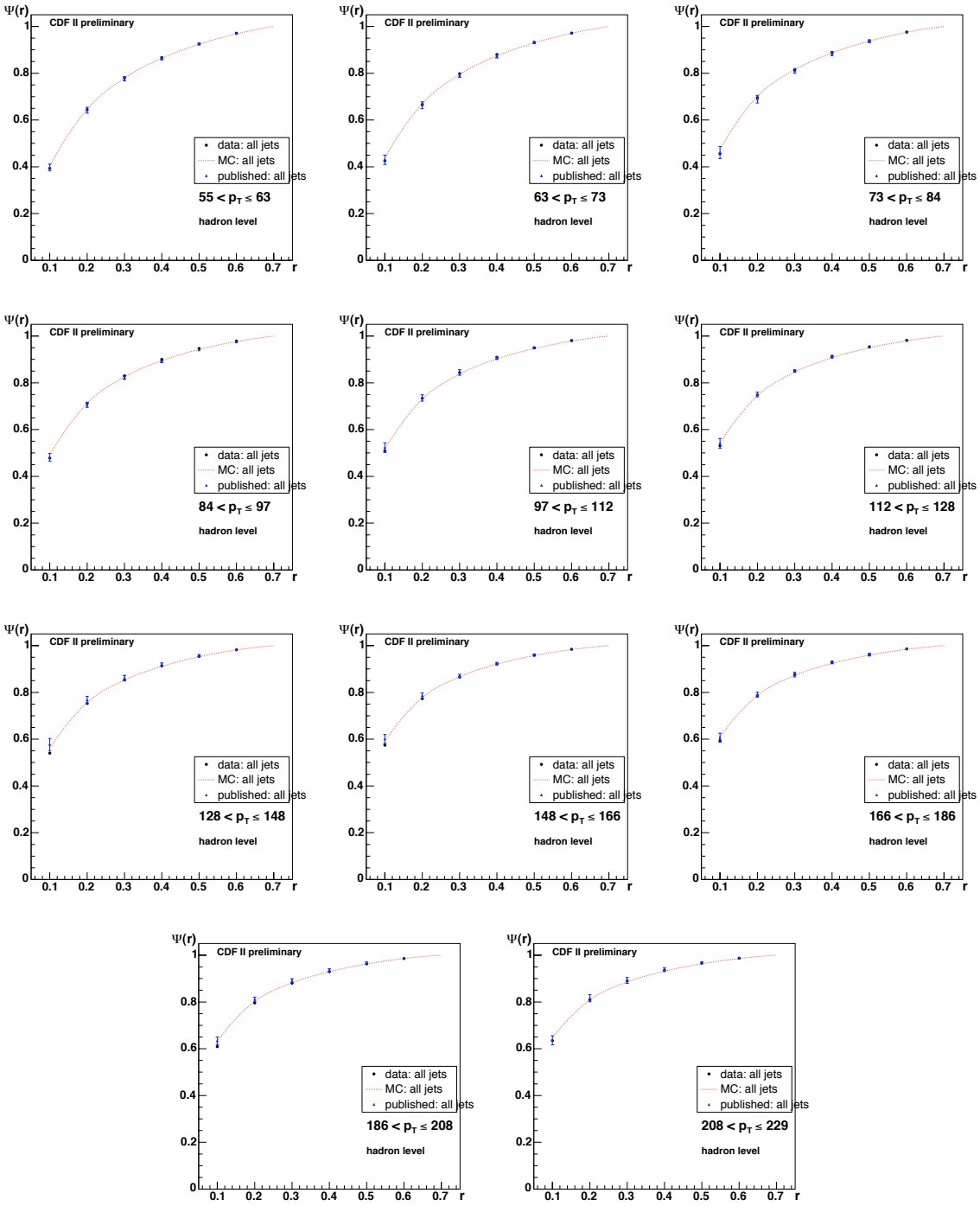

FIGURE C.14: Hadron level inclusive jet shapes. Comparing my results reconstructed from raw calorimeter towers with $p_T > 0.5\,\text{GeV}$ in data but still using the hadron level corrections for the shapes without a p_T cut (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

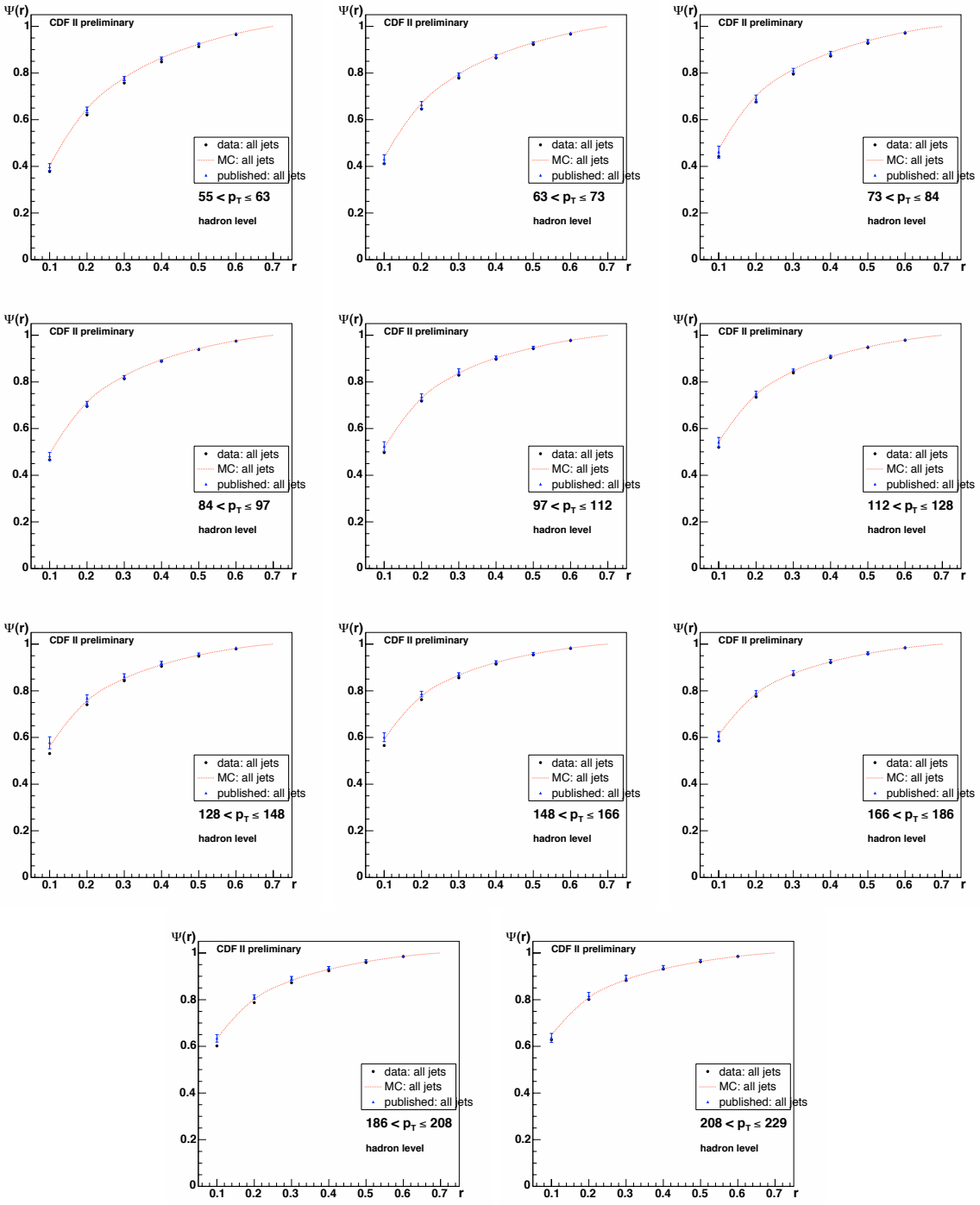

FIGURE C.15: Hadron level inclusive jet shapes. Comparing my results reconstructed from raw calorimeter towers with $p_T > 0.5\,\text{GeV}$ in data (black points) with the published results (blue points) and the Pythia Tune A MC predictions (red curve). Only the statistical errors are shown for my points.

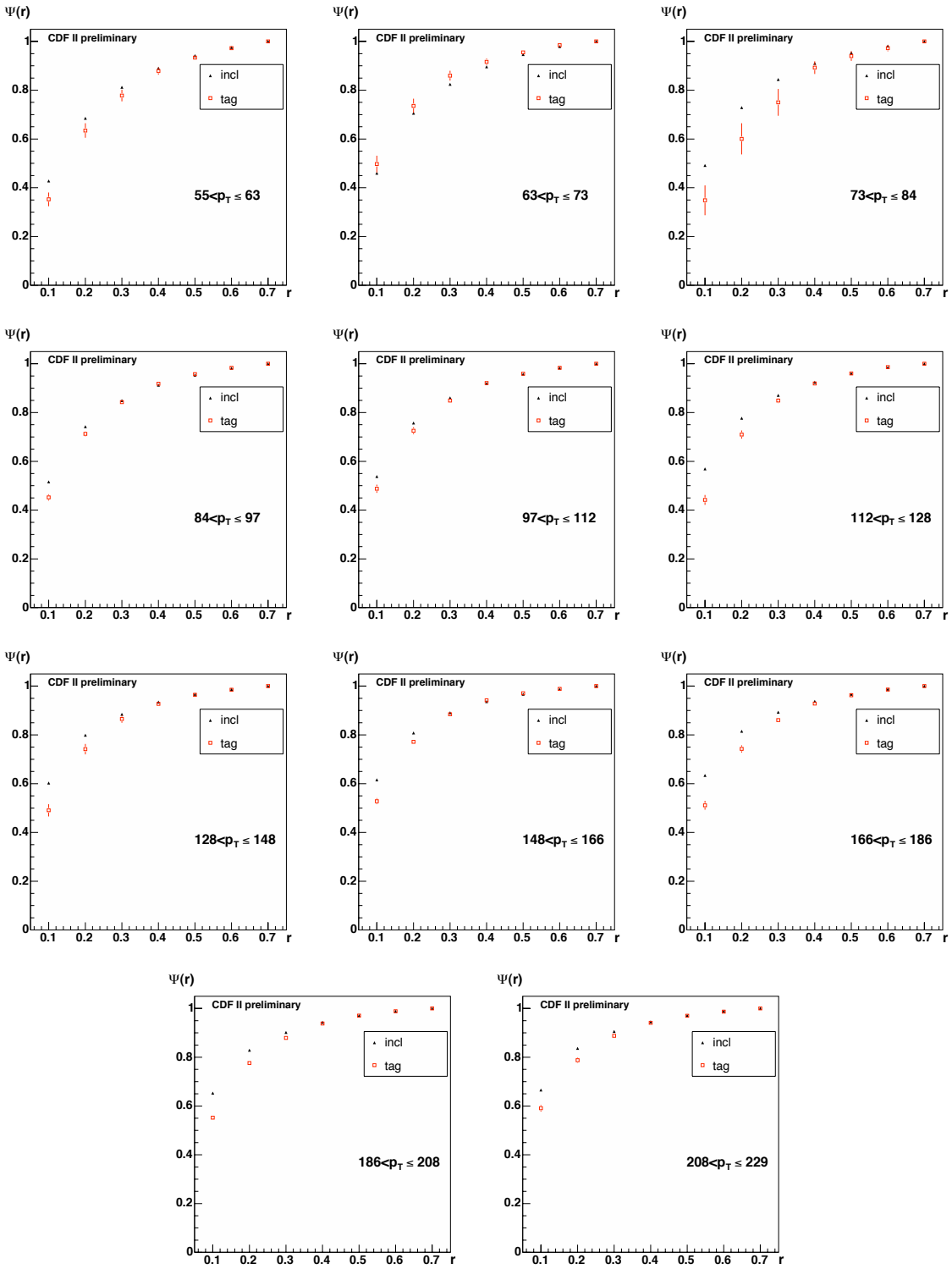

FIGURE C.16: MC Inclusive (and tagged) jet shapes at detector level level using cal towers.

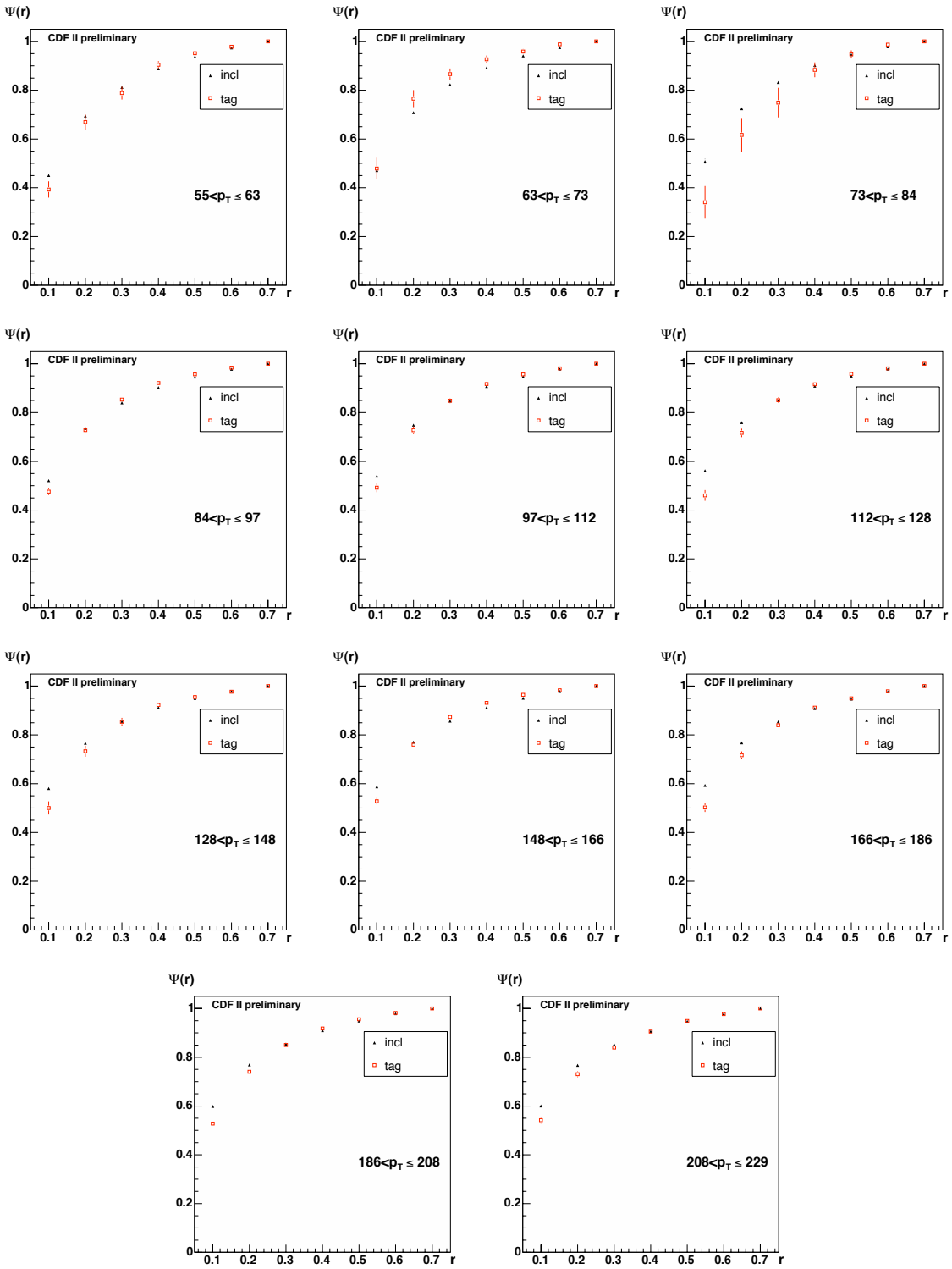

FIGURE C.17: MC Inclusive (and tagged) jet shapes at detector level level using tracks.

Appendix D

Systematic Studies

In this chapter, the systematic studies carried out for this analysis are presented. First, the method used to get an estimate on the error on the c-content of the nonb-jets is presented. A number of the parameters used in the unfolding method are sensitive to the relative fraction of b-jets which have only one b-quark inside the jet cone. This parameter is called $f_{1\mathrm{b}}$ and is defined as

$$f_{1\mathrm{b}} = \frac{N^{1b}_{\mathrm{jets}}}{N^{1b}_{\mathrm{jets}} + N^{2b}_{\mathrm{jets}}}. \tag{D.1}$$

Section D.2 presents a previous study comparing the Pythia Tune A results to NLO calculations.

There are a certain number of standard systematic studies which are carried out for all QCD analysis. These studies, presented and explained in section D.3, are repeated for each of the parameters of the unfolding.

Sections D.4 through D.7 present in turn the systematic effects on each of the parameters in the unfolding. The difference between the final shapes calculated using tracks instead of calorimeter towers is presented in D.8. The final section of this chapter presents a study, which is not included as a source of systematic error, on the effect of changing the rapidity cut from $|Y| < 0.7$ to $0.1 < |Y| < 0.7$. The purpose of that study is to check if the inclusive results, computed using the latter rapidity cuts, can be directly compared to this analysis. It turns out that removing the central rapidity region does not affect the b-jet shapes of the inclusive shapes.

D.1 Estimate of the c-fraction in data

An important source of systematic uncertainty is this fraction of nonb-jets which are c-jets. This will have an effect on the secondary vertex mass templates for nonb-jets which are used for the measurement of the purity of the sample. It will also have an effect on the bias due to tagging on the nonb-jets. Previous studies quote a conservative error of $20-30\%$ [10]. It is important to see if this error can be reduced. This was done by doing a secondary vertex mass fit using three different templates instead of the standard two presented in section 3.3. Different templates were obtained for b-jets, c-jets and light+gluon jets.

The limitation of this method is that the statistics for the c- and light+gluon-jet templates is too low for the fits to converge properly. It is reasonable to expect that the templates should not be very sensitive to the p_T threshold effects. The comparison between the templates obtained using the default MC samples and the higher p_T ones is shown in figure D.1. This shows that the templates are compatible. This could justify the use of higher p_T samples (Pt40, Pt60, Pt90 and Pt120 for the four bins). This assumption is checked by using the higher p_T samples for

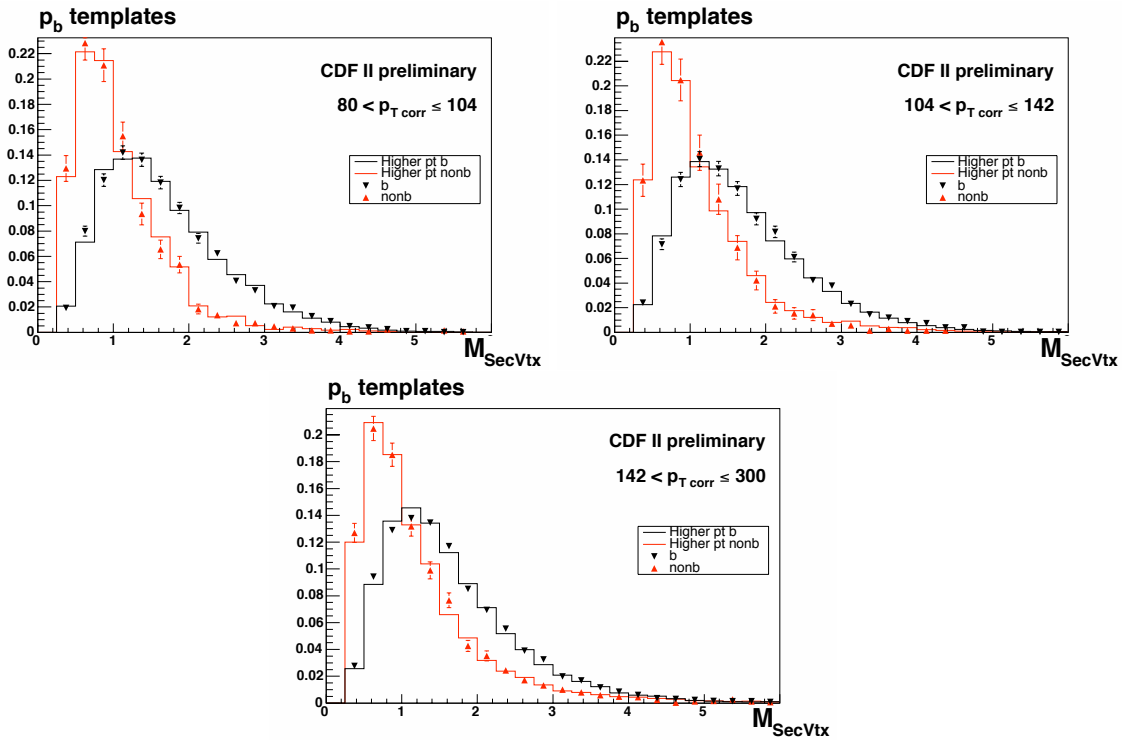

FIGURE D.1: Secondary vertex mass distributions for b- (black) and nonb-jets (red) for all the p_T bins. The standard MC samples, shown as points with the MC statistical errors, are compatible with the higher p_T samples (lines).

the templates in the standard purity extraction method described in section 3.3. This method is not applicable to the lowest p_T bin because the p_T distribution of the Pt40 sample is still rising up to about 60 GeV so it cannot be compared to the p_T distribution of the Pt18 sample. The results obtained using the higher p_T samples are within 2.5% of the results obtained using the standard MC samples, well within the total error on the fit. Table D.1 shows the fit results using the standard p_T templates and the higher p_T ones as well as the difference in the final results. Another check is that the results for the b-jet purity should be the same whether the fit is done using the standard method or using three different templates. The results are found to be within 2% of the standard results, as seen in table D.2. These tests show that the use of the higher p_T templates should not bias in any way the results of the fits.

p_T range	p_b	p_b higher p_T templates	Δp_b
80-104	0.323	0.299	0.024
104-142	0.300	0.275	0.025
142-300	0.242	0.242	0.000

TABLE D.1: Results for p_b using the standard MC samples and the higher p_T ones. Also shown is the difference between these two values for each bin.

The templates for b-, c- and light+gluon-jets using the higher p_T MC samples are shown as lines in figure D.2. The distribution in data, along with the fit to the three templates is shown in figure D.3. This shows that the fit is also very well. Finally, figure D.4 shows the results for

p_T range	p_b	p_c	f_c	f_c in MC	Δf_c	Δp_b
80-104	0.327	0.307	0.456	0.418	0.038	0.004
104-142	0.284	0.245	0.342	0.309	0.033	0.016
142-300	0.240	0.190	0.250	0.254	0.004	0.002

TABLE D.2: Results for p_b, p_c and f_c in data from the fits to data using the higher p_T samples. The f_c fraction in MC is also shown along with the difference between the fit and the MC values. The difference between the p_b values obtained with the standard fit and the 3-way fit are shown in the last column.

the c-content of the nonb templates, f_c as a function of p_T. This is computed by

$$f_c = \frac{p_c}{1 - p_b} \tag{D.2}$$

where p_b and p_c are the fitted b- and c-fractions. This shows that that the c-content of the nonb-jet fraction is well described in MC. The fit gives results, shown in table D.2, that are within 4% of the MC values. In order to account for this 4% difference in the fitted fractions, along with the uncertainty on the fit, about 3%, a systematic error of 5% is considered for the fraction of nonb-jets which are c-jets.

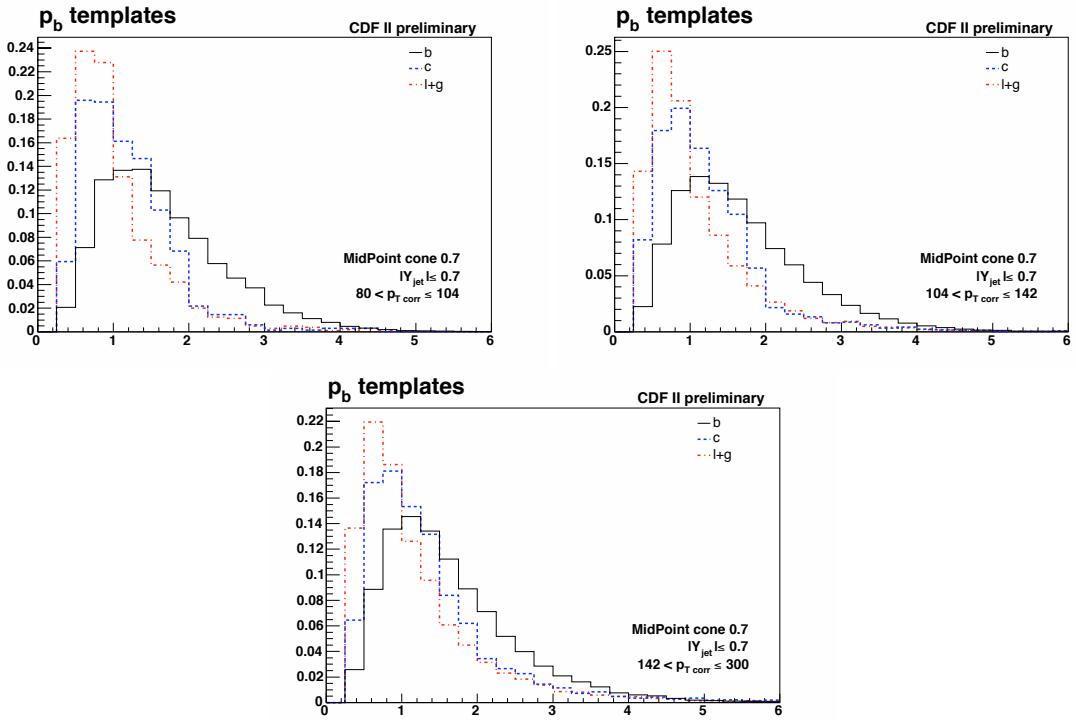

FIGURE D.2: Secondary vertex mass distributions for b- (black solid line), c- (blue dashed line) and light+gluon-jets (red dotted line) for all the p_T bins.

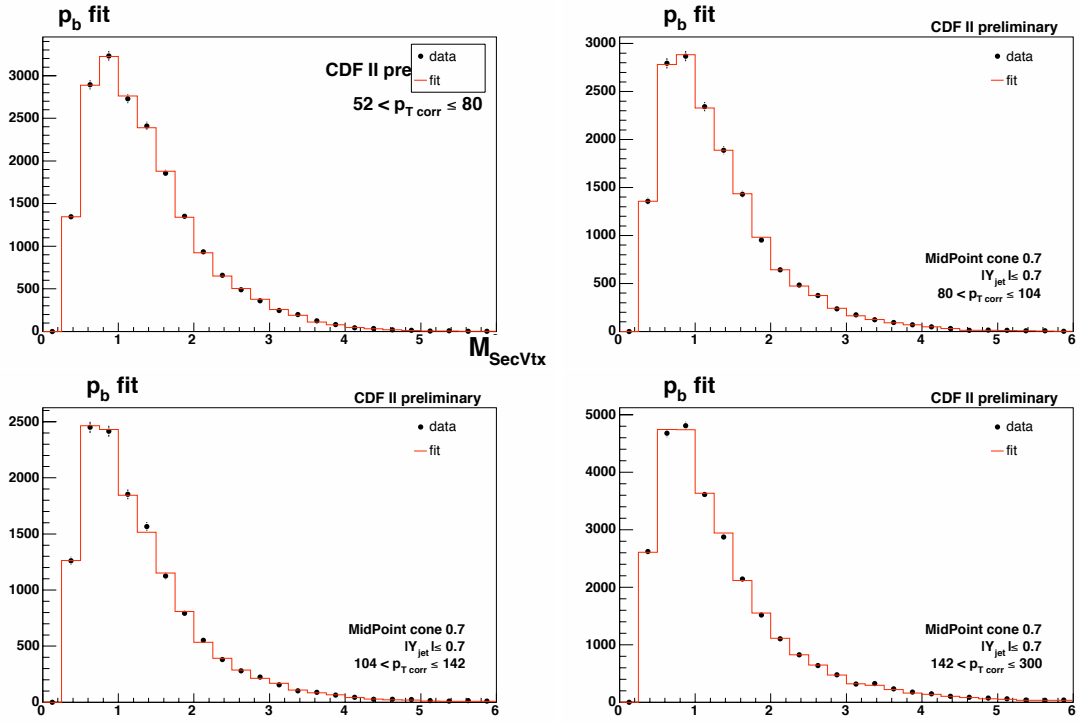

FIGURE D.3: Secondary vertex mass distribution in data (black points) compared to the fitted distribution using the three templates (red histogram) for all the p_T bins.

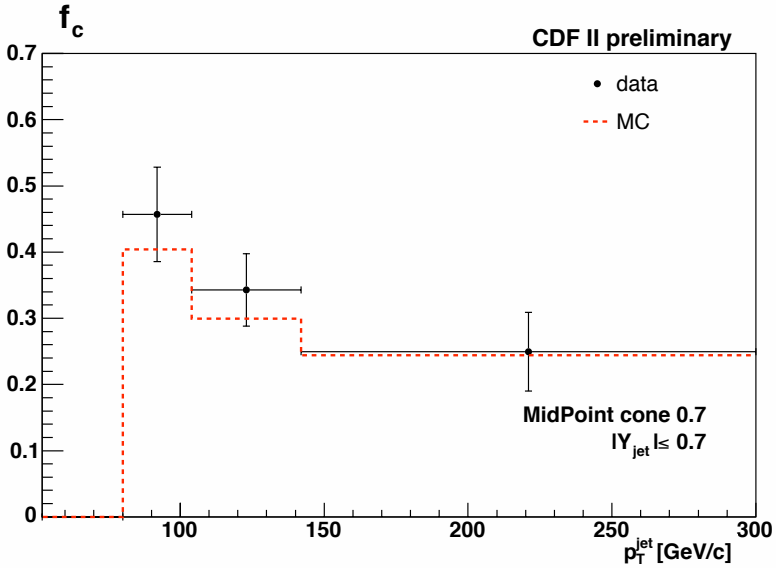

FIGURE D.4: Fraction of nonb-jets which are c-jets as a function of jet p_T (black line) shown along with the values obtained from MC (red dashed line) which are used in this analysis.

D.2 Estimate of error on f_{1b} and f_{1c}

Jets that contain two b-quarks inside the same jet cone are mainly jets which originate from a gluon splitting into a $b\bar{b}$ pair. Such jets are expected to be significantly wider than jets from flavour creation, containing only a single b-quark inside the cone. The fraction of gluon splitting events is not determined empirically but is one of the parameters which must be tuned in Monte Carlo to agree with data. The total fraction of inclusive jets which are gluon jets has been shown, by the inclusive shape analysis, to agree well with the data [2]. One of the aims of my analysis is to determine how well the fraction of gluon splitting events in b-jet production is modelled by the MC.

Many of the distributions used for the unfolding, such as the secondary vertex mass templates, the bias corrections and the hadron level corrections, are different for events that contain one or two b-quarks so the values used in the unfolding are somewhat sensitive to the fraction of events which have a single b-quarks, f_{1b}. In a similar way, the fraction of events with one c-quark inside the jets, f_{1c}, also has an effect on the analysis procedure. It is therefore important to obtain an estimate on the potential error of these fractions and include the variation as a systematic error.

Reference [11] shows a comparison of the fraction of events which contain two b-quarks between Pythia Tune A where the jet algorithm is run at hadron level and true Next to Leading Order (NLO) calculations. This comparison plot is reproduced in figure D.5. The scale choice of the NLO calculation affects the $b\bar{b}$ fraction. The maximum difference with respect to the Pythia Tune A values is obtained with a scale of $\mu = \mu_0/2$. The maximum difference between the Pythia Tune A and the NLO calculation is of the order of 15% (absolute). This 15% is the total absolute systematic error used on both the 1b- and the 1c-fractions. It turns out that when fitting the final b-jet shapes results to an unknown f_{1b} fraction, the results of the fit are within this 15% uncertainty for all the bins, see section 5.4.

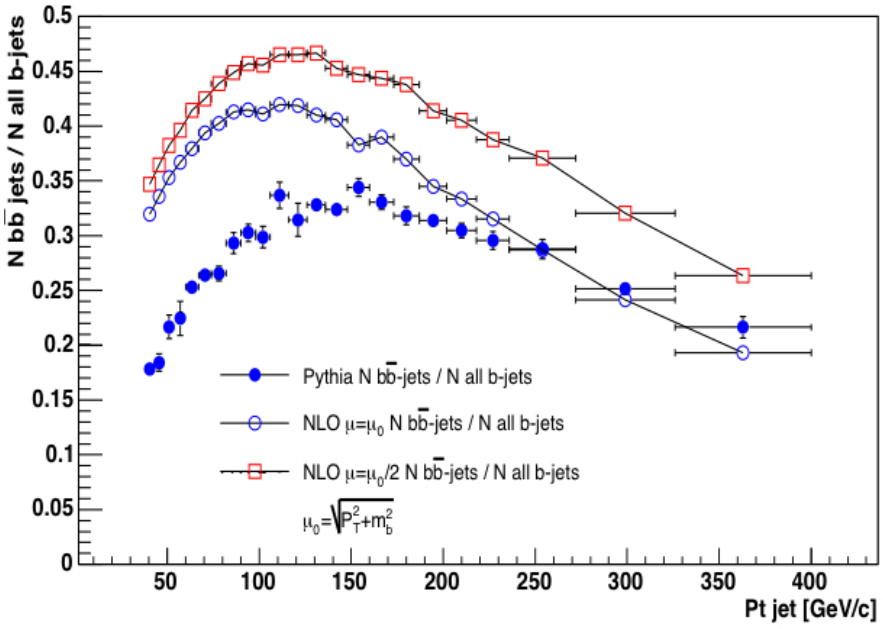

FIGURE D.5: Fraction of b-jets containing two b-quarks inside the jet cone. Pythia Tune A is compared to NLO calculations using both a scale $\mu = \mu_0$ and $\mu = \mu_0/2$

D.3 Standard source of systematic errors

A number of systematic studies are carried out for all, or most, of the parameters of the unfolding as well as on the measured shapes in data. The variations presented here are standard variations used in other QCD analyses [2] [10]. These are systematic effects due to the jet energy scale, the missing E_T significance cut and the $|Z_{\text{vtx}}|$ cut on the primary vertex.

The first of these, the jet energy scale, is the systematic error associated to the uncertainty on the jet corrections as well as the uncertainty on the detector level jets. It is a measure of how well the jet energy corrections, presented in section 2.4 manage to correct back the measured jet p_T back to the hadron level one. The systematic uncertainty on the jet energy scale quoted by the official jet energy corrections is 3% [13]. Another study calculated specifically the jet corrections for b-jets and found results that were in good agreement with the inclusive jet corrections [10]. The comparison between the official jet corrections and the b-jet specific corrections is shown in figure D.6 and D.7 for the inclusive and tagged jets respectively. Also shown in this plot is the p_T distribution of the corrected p_T using the same correction factors as the inclusive jet shape analysis [2]. This shows that the official jet corrections, other than having slightly different threshold effects, are close to the b-jet specific corrections. The systematic error on the jet energy scale should combine both the error quoted by the official jet corrections as well as the uncertainty on the b-jet fragmentation which is 0.6%. Adding these two effects in quadrature leads to a total systematic error on the jet energy scale of 3%. For this analysis, this effect is computed by shifting the p_T of the bins by $\pm 3\%$ and re-computing the different parameters.

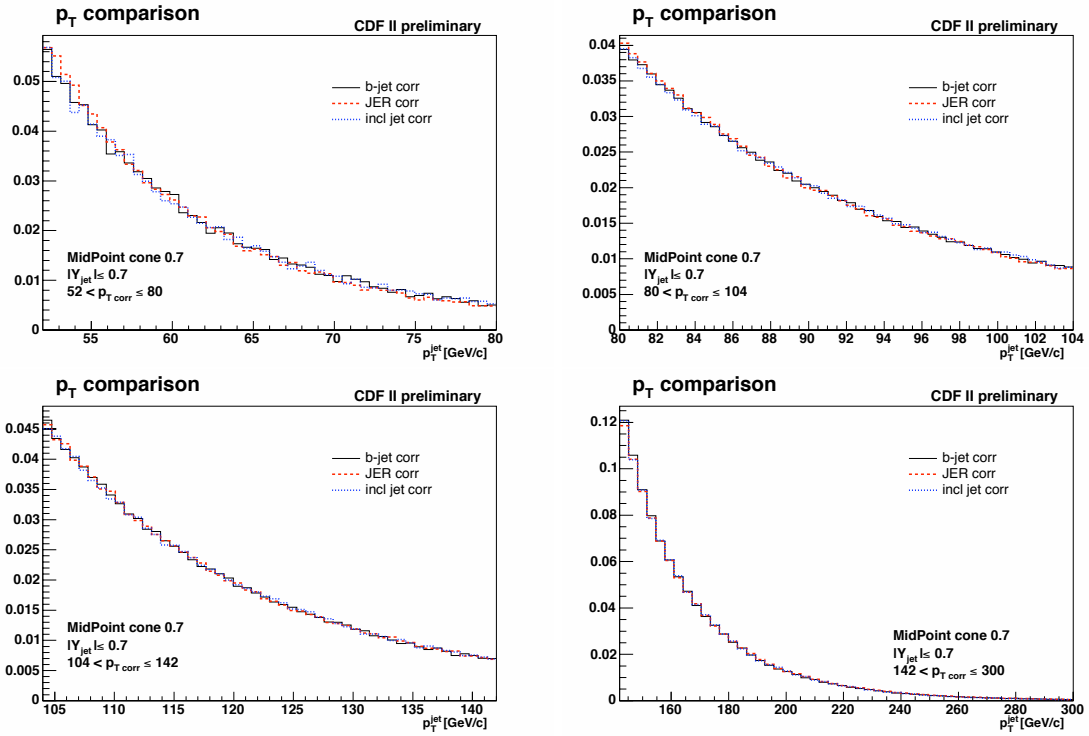

FIGURE D.6: Comparison in data of the corrected p_T distributions of all jets for each dataset obtained with the official jet corrections (red line), the b-jet specific corrections (black line) and the corrections used in the inclusive jet shape analysis (blue line).

The systematic effect on the missing E_T significance cut is taken to be the standard variation of $\pm 15\%$. The cut on the missing E_T significance removes a large fraction of the cosmic events

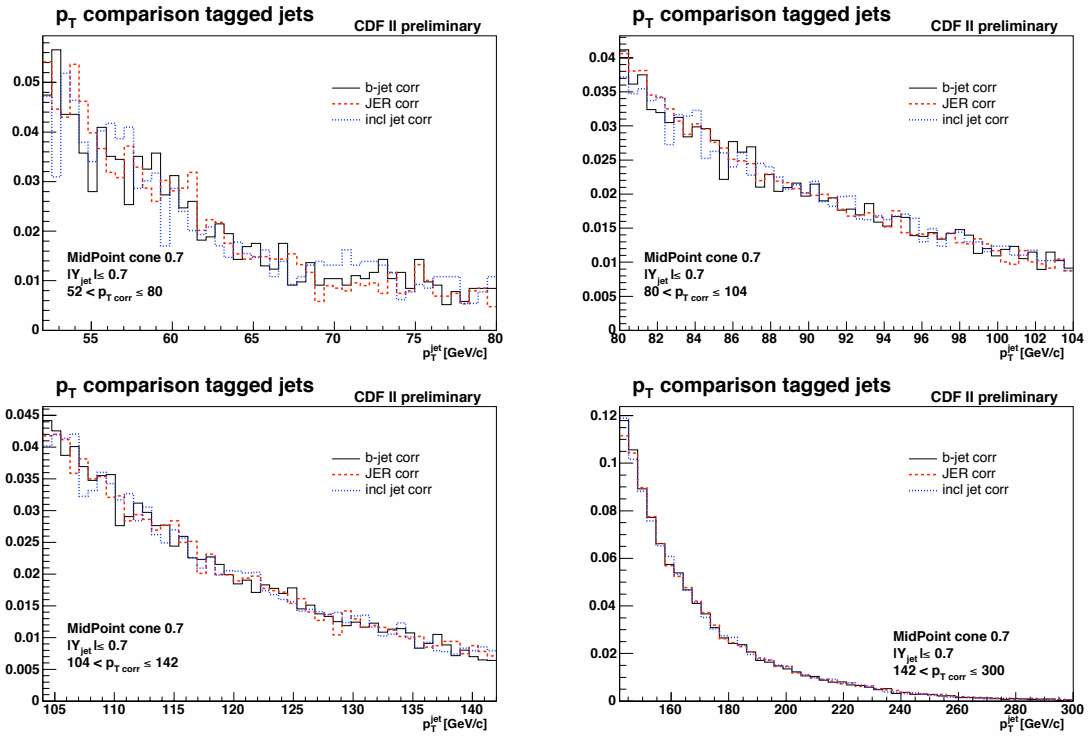

FIGURE D.7: Comparison in data of the corrected p_T distributions of tagged jets for each dataset obtained with the official jet corrections (black line), the b-jet specific corrections (red line) and the corrections used in the inclusive jet shape analysis (blue line).

but the shapes are not expected to be sensitive to the value of the cut because the number of events which are removed by this cut is very small. This systematic effect is thus expected to be small for all the parameters.

The cut on the location of the primary vertex is varied by $\pm 5\,\text{cm}$ around the base cut at $50\,\text{cm}$. Again this effect is expected to be small because the corrections to the jets and secondary vertices due to the location of the primary vertex are well understood in this region and the number of events which are removed by this cut is not large.

D.4 Raw shapes

The raw shapes, measured from data, must be checked against the typical systematic variations. The effects considered for the measured shapes are the following

- p_T bins $\pm 3\%$
- missing E_T significance $\pm 15\%$
- $|Z_{\text{vtx}}|$ cut $\pm 5\,\text{cm}$

D.4.1 Tagged Shapes

The statistical error, total systematic error as well as the individual systematic errors are shown in table D.4. The dominant source of systematic error on the tagged shape is the effect of the jet energy scale. the other effects are negligible in comparison. The statistical errors are of the same order of magnitude as the systematic ones.

D.4.2 Nonb-jet Shapes

In addition to the variations listed above, an additional systematic effect is considered which is the effect of using the inclusive shape instead of the nonb-jet shape. This is basically investigating how good the approximation

$$\Psi_{\text{nonb}}(r) \approx \Psi_{\text{incl}}(r) \tag{D.3}$$

is. The statistical error, total systematic error as well as the individual systematic errors are shown in table D.4. The systematic error due to the use of the inclusive instead of the nonb-jet shape is much smaller than the systematic error due to the jet energy scale. The error on the jet energy scale is the dominant source of systematic error on the inclusive shape. The statistical errors are about a factor of 10 smaller than the systematic errors.

p_T range 52-80 GeV					
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$
0.1	0.0037	0.0051	0.0051	0.0002	0.0002
0.2	0.0034	0.0046	0.0046	0.0001	0.0001
0.3	0.0025	0.0027	0.0027	0.0001	0.0000
0.4	0.0016	0.0018	0.0018	0.0001	0.0000
0.5	0.0010	0.0009	0.0009	0.0000	0.0000
0.6	0.0005	0.0006	0.0006	0.0000	0.0000
0.7	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 80-104 GeV					
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$
0.1	0.0023	0.0041	0.0041	0.0001	0.0000
0.2	0.0020	0.0033	0.0033	0.0001	0.0001
0.3	0.0014	0.0022	0.0022	0.0000	0.0000
0.4	0.0009	0.0012	0.0012	0.0000	0.0000
0.5	0.0005	0.0008	0.0008	0.0000	0.0000
0.6	0.0003	0.0004	0.0004	0.0000	0.0000
0.7	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 104-142 GeV					
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$
0.1	0.0025	0.0039	0.0039	0.0001	0.0000
0.2	0.0021	0.0019	0.0019	0.0001	0.0000
0.3	0.0014	0.0005	0.0005	0.0001	0.0000
0.4	0.0009	0.0002	0.0002	0.0000	0.0000
0.5	0.0005	0.0003	0.0003	0.0000	0.0000
0.6	0.0003	0.0002	0.0002	0.0000	0.0000
0.7	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 142-300 GeV					
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$
0.1	0.0019	0.0036	0.0036	0.0001	0.0001
0.2	0.0015	0.0036	0.0036	0.0000	0.0001
0.3	0.0010	0.0023	0.0023	0.0001	0.0000
0.4	0.0006	0.0013	0.0013	0.0000	0.0000
0.5	0.0003	0.0007	0.0007	0.0000	0.0000
0.6	0.0002	0.0003	0.0003	0.0000	0.0000
0.7	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE D.3: Statistical and systematic errors on the raw tagged shapes for each p_T bin and each r bin. Also shown are the individual sources of systematic errors.

p_T range 52-80 GeV						
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$	$\Psi_{\rm nonb}(r) - \Psi_{\rm incl}(r)$
0.1	0.0007	0.0075	0.0073	0.0000	0.0001	0.0015
0.2	0.0006	0.0064	0.0063	0.0001	0.0001	0.0012
0.3	0.0005	0.0044	0.0044	0.0001	0.0001	0.0008
0.4	0.0004	0.0028	0.0028	0.0000	0.0000	0.0006
0.5	0.0002	0.0017	0.0016	0.0000	0.0000	0.0004
0.6	0.0001	0.0007	0.0007	0.0000	0.0000	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 80-104 GeV						
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\%$	met. sig $\pm 15\%$	$\Psi_{\rm nonb}(r) - \Psi_{\rm incl}(r)$
0.1	0.0005	0.0059	0.0056	0.0001	0.0001	0.0018
0.2	0.0004	0.0039	0.0036	0.0000	0.0001	0.0015
0.3	0.0003	0.0024	0.0022	0.0000	0.0000	0.0011
0.4	0.0002	0.0014	0.0013	0.0000	0.0000	0.0007
0.5	0.0001	0.0007	0.0006	0.0000	0.0000	0.0003
0.6	0.0000	0.0002	0.0002	0.0000	0.0000	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 104-142 GeV						
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$	$\Psi_{\rm nonb}(r) - \Psi_{\rm incl}(r)$
0.1	0.0005	0.0047	0.0044	0.0001	0.0001	0.0017
0.2	0.0004	0.0029	0.0026	0.0001	0.0001	0.0012
0.3	0.0003	0.0017	0.0014	0.0001	0.0001	0.0010
0.4	0.0002	0.0011	0.0009	0.0001	0.0000	0.0007
0.5	0.0001	0.0005	0.0004	0.0001	0.0000	0.0004
0.6	0.0000	0.0001	0.0001	0.0001	0.0000	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 142-300 GeV						
r bin	stat	syst.	$p_T \pm 3\%$	$\lvert Z_{\rm vxt}\rvert \pm 5\,{\rm cm}$	met. sig $\pm 15\%$	$\Psi_{\rm nonb}(r) - \Psi_{\rm incl}(r)$
0.1	0.0004	0.0048	0.0042	0.0001	0.0001	0.0023
0.2	0.0003	0.0031	0.0026	0.0001	0.0001	0.0017
0.3	0.0002	0.0017	0.0014	0.0000	0.0000	0.0010
0.4	0.0002	0.0009	0.0007	0.0000	0.0000	0.0006
0.5	0.0001	0.0004	0.0003	0.0000	0.0000	0.0003
0.6	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE D.4: Statistical and systematic errors on the raw inclusive shapes for each p_T bin and each r bin. Also shown are the individual sources of systematic errors.

D.5 Purity

Before going into detailed systematic error calculation on the purity, it is important to be able to extract the contribution of the statistical error in the total error on the purity quoted by the fitting algorithm. The fitting algorithm returns a total error which takes into account both the statistical errors of the templates and the data curves. In order to extract the error due only to the statistics of the data, the statistics of the templates are artificially increased until the contribution to the total error of the fit is negligible. The error contribution due to the statistics of the templates goes roughly as $1/\sqrt{N}$ so by increasing artificially the statistics of the Monte Carlo templates, one is effectively reducing the error of the templates to zero. This method is similar to that shown in [10]. In practise this is done by scaling the template histograms for b- and nonb-jets by a factor of 10^3. This factor is motivated by looking at figure D.8 taken from [10] which shows the total error of the fit as a function of the factor by which the templates are scaled for one particular bin in p_T. The blue line shows the limit to which the error tends. This is taken as the statistical error. This shows that a scaling of 10^3 is definitely sufficient to assume the error quoted is the statistical error.

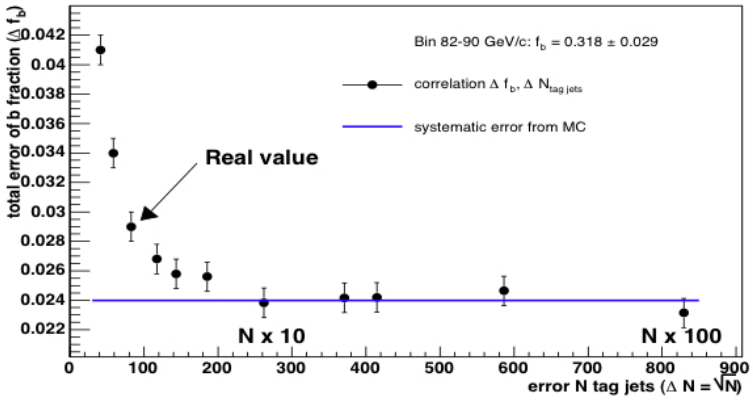

FIGURE D.8: Error on the purity of the sample returned by the fitting algorithm as a function of the artificial scaling applied to the MC templates.

The error on the fit when using the scaled histograms is taken as the statistical error, $\Delta p_{\mathrm{b}}^{\mathrm{stat}}$. The error due to the MC statistics is thus given by

$$\Delta p_{\mathrm{b}}^{\mathrm{MC\,stat}} = \sqrt{(\Delta p_{\mathrm{b}}^{\mathrm{fit}})^2 - (\Delta p_{\mathrm{b}}^{\mathrm{stat}})^2} \tag{D.4}$$

where $p_{\mathrm{b}}^{\mathrm{fit}}$ is the total error quoted by the fitting algorithm.

Having extracted the statistical error and the systematic error due to the MC statistics, other sources of systematic error must be investigated. These are

- p_T bins $\pm 3\%$
- missing E_T significance $\pm 15\%$
- $|Z_{\mathrm{vtx}}|$ cut $\pm 5\,\mathrm{cm}$
- $f_{1\mathrm{b}} \pm 15\%$
- $f_{1\mathrm{c}} \pm 15\%$
- $f_{\mathrm{c}} \pm 5\%$

The difference in the purity and templates between the different MC samples also needs to be investigated. Because the statistics of the Herwig MC samples are not very large, it was necessary to carry out the same procedure as detailed in section D.1. The higher p_T Herwig nonb-templates were used. As for the c-fraction extraction, the lowest p_T bin cannot be considered because of the different p_T distributions. The Herwig templates are within statistical errors of the Pythia Tune A ones. Figure D.9 shows the comparison between the templates using Pythia Tune A (solid lines) and Herwig (points). The fit results using the higher p_T Herwig samples are within the quoted fit errors of the Pythia Tune A results as can be seen in table D.5.

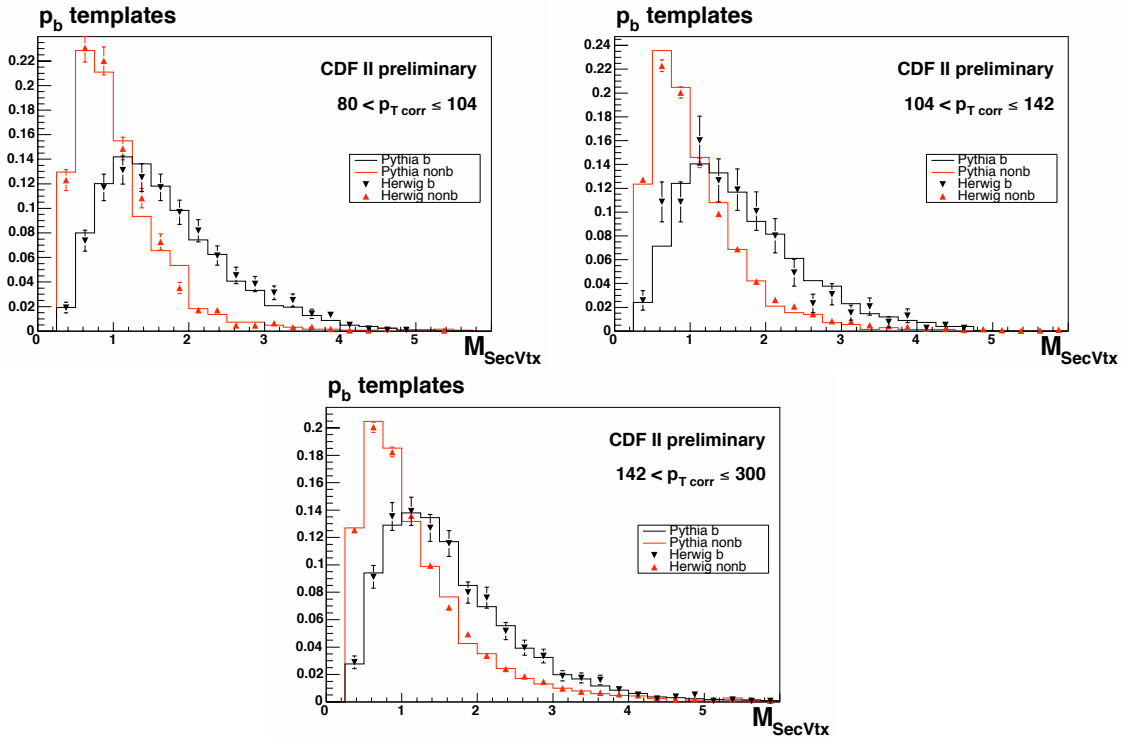

FIGURE D.9: Secondary vertex mass distributions for b- (black) and nonb-jets (red) for all the p_T bins. The full lines represent the Pythia Tune A MC templates whereas the points represent the Herwig higher p_T samples.

	Herwig		
p_T range	p_b	fit error	Δp_b
80-104	0.342	0.024	0.019
104-142	0.327	0.032	0.027
142-300	0.253	0.020	0.011

TABLE D.5: Purity fits results using the Herwig templates. The total fit error quoted by the fitting algorithm is shown as well as the difference between the fit results and those obtained with Pythia Tune A. The Herwig results are within errors of the Pythia Tune A ones.

A similar procedure can be applied to investigate the effect of the fragmentation function and PDFs on the secondary vertex mass templates. Because only b-filtered samples are available for the Pythia Tune A MC samples with the CTEQ6L PDFs as well as those with the Peterson fragmentation function, the comparison is only done between the b-jet templates. The b-jet template for the default Pythia Tune A templates are shown in figures D.10 and D.11 along with the b-jet template obtained using the Pythia Tune A with CTEQ6L PDFs for the former

and the Pythia Tune A with the Peterson fragmentation functions for the latter. The statistics for the comparison to Pythia Tune A with the Peterson fragmentation for the lowest bin are not very high. This shows that at least for b-jets, where we expect the difference to be maximal, the templates are within errors of eachother.

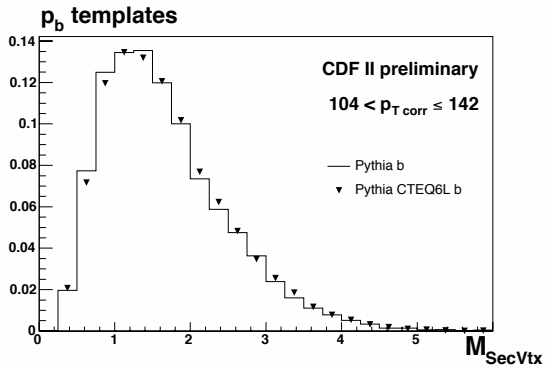

FIGURE D.10: Secondary vertex mass distributions for b-jets. The line represents the default Pythia Tune A templates whereas the points represent the Pythia Tune A using the CTEQ6L PDFs templates.

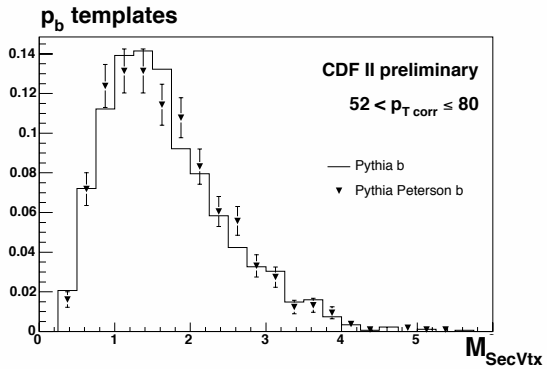

FIGURE D.11: Secondary vertex mass distributions for b-jets. The line represents the default Pythia Tune A templates whereas the points represent the Pythia Tune A using the Peterson fragmentation function templates.

The total effect on the purity due to the use of templates using Pythia Tune A is thus taken to be negligible. The statistical and systematic errors on the purity are shown in table D.6. Also shown are the different contributions to the systematic effect due to the contributions mentioned above. The dominant source of systematic error on the purity is the error due to the MC statistics. This is dominated by the low statistics available for the nonb-templates. The error on the purity is not a dominant source of systematic on the final b-jet shapes.

p_T range	stat.	syst.	MC stat	$p_T \pm 3$	$\lvert Z_{\mathrm{vtx}} \rvert \pm 5$	met sig ± 15	$f_{1\mathrm{b}} \pm 15\%$	$f_{1\mathrm{c}} \pm 15\%$	$f_c \pm 5\%$
52-80	0.0150	0.0424	0.0322	0.0246	0.0016	0.0004	0.0074	0.0091	0.0041
80-104	0.0097	0.0348	0.0192	0.0213	0.0015	0.0004	0.0119	0.0130	0.0086
104-142	0.0101	0.0320	0.0231	0.0076	0.0010	0.0003	0.0142	0.0098	0.0115
142-300	0.0089	0.0358	0.0239	0.0122	0.0011	0.0012	0.0068	0.0167	0.0152

TABLE D.6: Statistical and systematic errors on the purity for each Pythia Tune A bin. Also shown are contributions to the total systematic error from the different effects.

D.6 Biases due to SecVtx tagging

All the standard sources of systematic errors are investigated both for the bias due to tagging on b-jets and on nonb-jets. These are

- MC statistics
- Jet energy scale $\pm 3\%$
- Missing E_T significance $\pm 15\%$
- $|Z_{\text{vtx}}|$ cut $\pm 5\,\text{cm}$

D.6.1 Biases due to tagging on b-jets

On top of the above mentioned effects, the $f_{1\text{b}}$ fraction affects the overall bias because the biases due to tagging on 1b and 2b jets are not the same, as was shown in figure 3.12. It is therefore necessary to include the effect of the variation of $f_{1\text{b}}$ as a systematic error.

The total systematic error on the bias due to the tagging on b-jets is shown in table D.7 along with the different contributions due to the different effects. The dominant source of systematic error on the bias on b-jets due to the tagging requirement is due to the MC statistics.

D.6.2 Biases due to tagging on nonb-jets

On top of the effects mentioned in the beginning of this section, some additional systematic effects need to be considered. These are

- Effect of using the inclusive jet shapes instead of the nonb-jet shapes in the unfolding
- $f_{1\text{c}} \pm 15\%$
- $f_c \pm 5\%$

The total systematic error is shown in table D.8 along with the different contributions from the different sources of systematic error. The dominant sources of systematic error on the bias on nonb-jets due to the tagging requirement are the MC statistics and the $f_{1\text{c}}$ fraction.

p_T range 52-80 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig $\pm 15\%$	$f_{\text{1b}} \pm 15\%$
0.1	0.0228	0.0213	0.0065	0.0004	0.0002	0.0049
0.2	0.0117	0.0112	0.0001	0.0009	0.0001	0.0032
0.3	0.0082	0.0072	0.0002	0.0008	0.0000	0.0038
0.4	0.0055	0.0045	0.0000	0.0006	0.0000	0.0032
0.5	0.0030	0.0026	0.0001	0.0003	0.0000	0.0016
0.6	0.0013	0.0011	0.0000	0.0001	0.0000	0.0006
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 80-104 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig $\pm 15\%$	$f_{\text{1b}} \pm 15\%$
0.1	0.0109	0.0109	0.0006	0.0004	0.0001	0.0000
0.2	0.0076	0.0065	0.0006	0.0002	0.0001	0.0038
0.3	0.0055	0.0042	0.0007	0.0002	0.0000	0.0035
0.4	0.0036	0.0026	0.0003	0.0001	0.0000	0.0024
0.5	0.0018	0.0014	0.0002	0.0001	0.0000	0.0011
0.6	0.0007	0.0006	0.0003	0.0000	0.0000	0.0002
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 104-142 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig $\pm 15\%$	$f_{\text{1b}} \pm 15\%$
0.1	0.0127	0.0117	0.0034	0.0008	0.0002	0.0034
0.2	0.0074	0.0072	0.0010	0.0003	0.0002	0.0011
0.3	0.0051	0.0045	0.0009	0.0003	0.0001	0.0022
0.4	0.0033	0.0028	0.0005	0.0003	0.0001	0.0016
0.5	0.0019	0.0017	0.0003	0.0002	0.0001	0.0008
0.6	0.0008	0.0007	0.0000	0.0001	0.0000	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T range 142-300 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig $\pm 15\%$	$f_{\text{1b}} \pm 15\%$
0.1	0.0066	0.0064	0.0008	0.0004	0.0002	0.0014
0.2	0.0058	0.0040	0.0004	0.0002	0.0002	0.0042
0.3	0.0052	0.0025	0.0002	0.0001	0.0001	0.0046
0.4	0.0038	0.0015	0.0001	0.0001	0.0000	0.0034
0.5	0.0020	0.0009	0.0002	0.0001	0.0000	0.0017
0.6	0.0006	0.0004	0.0001	0.0000	0.0000	0.0004
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE D.7: Total systematic error on the bias due to tagging on b-jets for each p_T bin for each bin in r. Also shown are contributions to the total systematic error from the different effects.

p_T 52 − 80 GeV							
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{c}} \pm 15\%$	$f_{\text{c}} \pm 5\%$
0.1	0.0665	0.0440	0.0113	0.0015	0.0017	0.0477	0.0092
0.2	0.0355	0.0260	0.0009	0.0013	0.0006	0.0236	0.0052
0.3	0.0195	0.0156	0.0019	0.0010	0.0004	0.0111	0.0033
0.4	0.0101	0.0086	0.0021	0.0005	0.0002	0.0041	0.0023
0.5	0.0052	0.0048	0.0002	0.0002	0.0001	0.0017	0.0010
0.6	0.0019	0.0018	0.0004	0.0001	0.0000	0.0003	0.0004
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 80 − 104 GeV							
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{c}} \pm 15\%$	$f_{\text{c}} \pm 5\%$
0.1	0.0357	0.0167	0.0084	0.0007	0.0003	0.0297	0.0069
0.2	0.0172	0.0093	0.0074	0.0012	0.0001	0.0120	0.0032
0.3	0.0089	0.0054	0.0036	0.0009	0.0001	0.0058	0.0018
0.4	0.0041	0.0029	0.0004	0.0003	0.0000	0.0027	0.0008
0.5	0.0023	0.0016	0.0005	0.0002	0.0000	0.0014	0.0004
0.6	0.0010	0.0008	0.0003	0.0000	0.0000	0.0005	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 104 − 142 GeV							
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{c}} \pm 15\%$	$f_{\text{c}} \pm 5\%$
0.1	0.0297	0.0209	0.0081	0.0021	0.0000	0.0188	0.0046
0.2	0.0149	0.0125	0.0014	0.0005	0.0000	0.0077	0.0019
0.3	0.0078	0.0068	0.0001	0.0002	0.0000	0.0036	0.0011
0.4	0.0042	0.0038	0.0006	0.0004	0.0000	0.0017	0.0005
0.5	0.0023	0.0020	0.0010	0.0002	0.0000	0.0006	0.0002
0.6	0.0009	0.0009	0.0001	0.0001	0.0000	0.0002	0.0001
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 142 − 300 GeV							
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{c}} \pm 15\%$	$f_{\text{c}} \pm 5\%$
0.1	0.0147	0.0107	0.0031	0.0005	0.0003	0.0095	0.0004
0.2	0.0076	0.0064	0.0018	0.0006	0.0001	0.0036	0.0002
0.3	0.0043	0.0036	0.0014	0.0002	0.0001	0.0019	0.0003
0.4	0.0021	0.0020	0.0003	0.0001	0.0000	0.0008	0.0001
0.5	0.0013	0.0011	0.0005	0.0000	0.0000	0.0005	0.0001
0.6	0.0006	0.0005	0.0002	0.0000	0.0000	0.0002	0.0000
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE D.8: Total systematic error on the bias due to tagging on nonb-jets for each p_T bin for each bin in r. Also shown are contributions to the total systematic error from the different effects.

D.7 Hadron Level Corrections to the jet shapes

The systematic effects considered for the hadron level corrections to the b-jet shapes are the following

- MC statistics
- Jet energy scale p_T $\pm 3\%$
- Missing E_T significance $\pm 15\%$
- $|Z_{\mathrm{vtx}}|$ cut $\pm 5\,\mathrm{cm}$
- $f_{1\mathrm{b}} \pm 15\%$

The total systematic error for each p_T bin and r bin is shown in table D.9 along with the contributions due to the different effects. The dominant source of systematic error on $C_{had}(r)$ factors is the $f_{1\mathrm{b}}$ fraction which dominates by a factor of 4-5 the other systematic errors.

p_T 52 − 80 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{b}} \pm 15\%$
0.1	0.0433	0.0109	0.0035	0.0011	0.0027	0.0417
0.2	0.0263	0.0067	0.0007	0.0006	0.0013	0.0254
0.3	0.0152	0.0045	0.0000	0.0005	0.0007	0.0145
0.4	0.0083	0.0029	0.0004	0.0004	0.0004	0.0078
0.5	0.0040	0.0017	0.0005	0.0002	0.0003	0.0036
0.6	0.0015	0.0008	0.0001	0.0001	0.0001	0.0012
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 80 − 104 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{b}} \pm 15\%$
0.1	0.0364	0.0062	0.0097	0.0007	0.0012	0.0346
0.2	0.0187	0.0038	0.0026	0.0003	0.0006	0.0181
0.3	0.0100	0.0025	0.0008	0.0001	0.0003	0.0096
0.4	0.0052	0.0016	0.0003	0.0001	0.0002	0.0050
0.5	0.0025	0.0009	0.0000	0.0001	0.0001	0.0023
0.6	0.0009	0.0004	0.0003	0.0001	0.0000	0.0008
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 104 − 142 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{b}} \pm 15\%$
0.1	0.0270	0.0066	0.0033	0.0011	0.0016	0.0259
0.2	0.0139	0.0040	0.0030	0.0005	0.0008	0.0130
0.3	0.0075	0.0027	0.0010	0.0004	0.0005	0.0069
0.4	0.0043	0.0017	0.0015	0.0004	0.0003	0.0036
0.5	0.0021	0.0010	0.0006	0.0002	0.0002	0.0017
0.6	0.0008	0.0004	0.0000	0.0001	0.0001	0.0006
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p_T 142 − 300 GeV						
r bin	total syst.	MC stat	$p_T \pm 3\%$	$\lvert Z_{\text{vtx}}\rvert \pm 5\,\text{cm}$	met. sig ±15%	$f_{1\text{b}} \pm 15\%$
0.1	0.0354	0.0034	0.0014	0.0004	0.0016	0.0352
0.2	0.0175	0.0021	0.0001	0.0002	0.0008	0.0174
0.3	0.0093	0.0014	0.0005	0.0001	0.0004	0.0092
0.4	0.0050	0.0009	0.0003	0.0001	0.0002	0.0049
0.5	0.0023	0.0006	0.0003	0.0000	0.0001	0.0022
0.6	0.0007	0.0002	0.0000	0.0000	0.0000	0.0007
0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE D.9: Total systematic error on the hadron level corrections to the b-jet shapes for each p_T bin for each bin in r. Also shown are contributions to the total systematic error from the different effects.

D.8 Comparing final shapes obtained using calorimeter level towers and tracks

The hadron level b-jet shapes obtained starting from the raw calorimetric shapes are compared to those obtained starting with the tracks. To be included inside a jet, the tracks must pass certain cuts

- $0.5 < p_T^{\text{track}} < 100\,\text{GeV}$
- track $|Z_0| < 2\,\text{cm}$ with respect to the primary vertex

These cuts are the same ones as used for the inclusive analysis [2].

Figure D.12 shows a comparison between data and Pythia Tune A MC of the raw inclusive shapes measured using tracks. Figure D.13 shows the same thing but for tagged jets. As for the case of calorimetric jets, the agreement between data and MC is not perfect for the inclusive jets. The data and MC for the tagged jets are also different, as would be expected from the fact that there are different fractions of b-jets in data and MC.

The bias corrections to correct for the tagging requirement and the hadron level corrections are computed specifically for the track level shapes. The hadron level b-quark jet shapes are computed separately for the tracks and the calorimeter tower shapes and the differences between these shapes are taken as a systematic uncertainty.

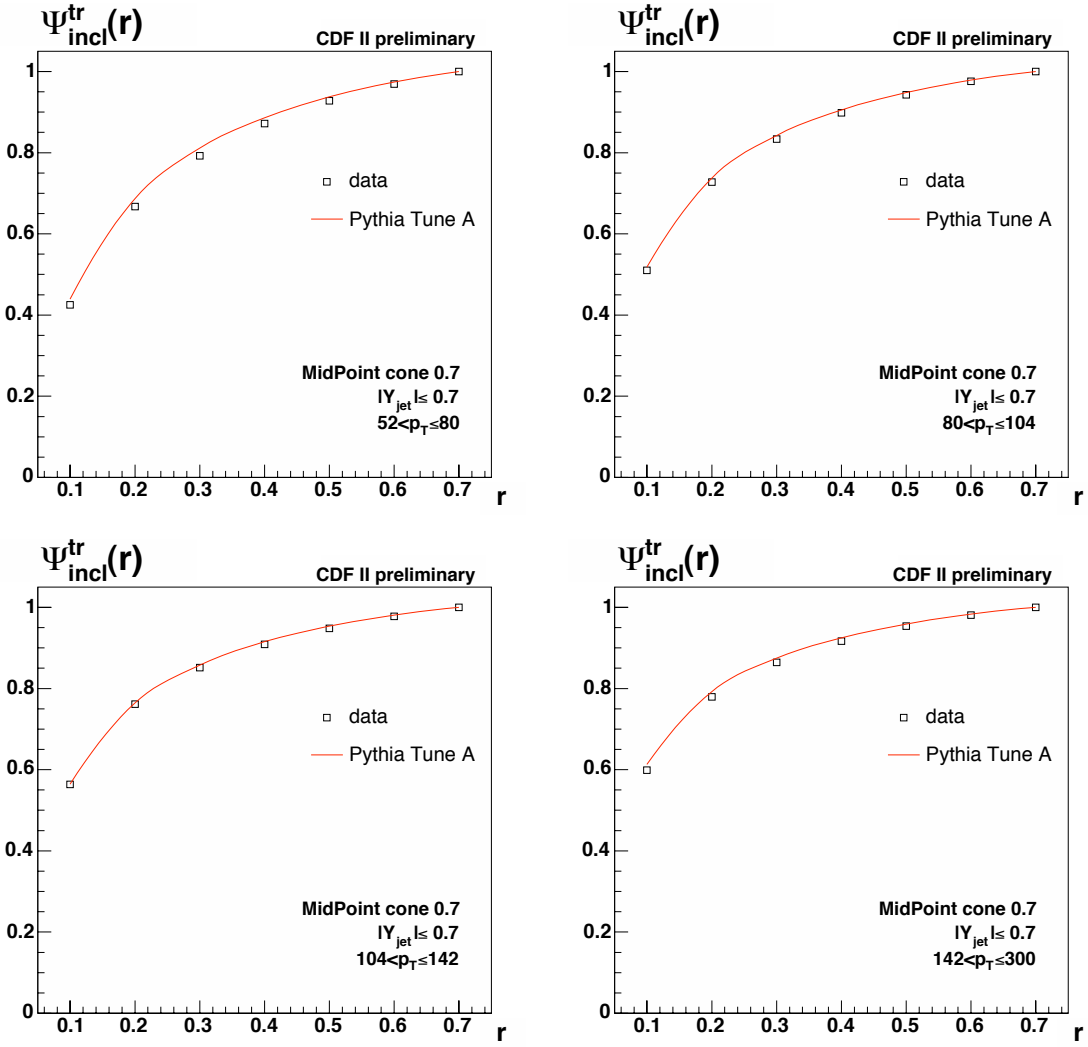

FIGURE D.12: Raw integrated shapes for inclusive jets measured using tracks. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

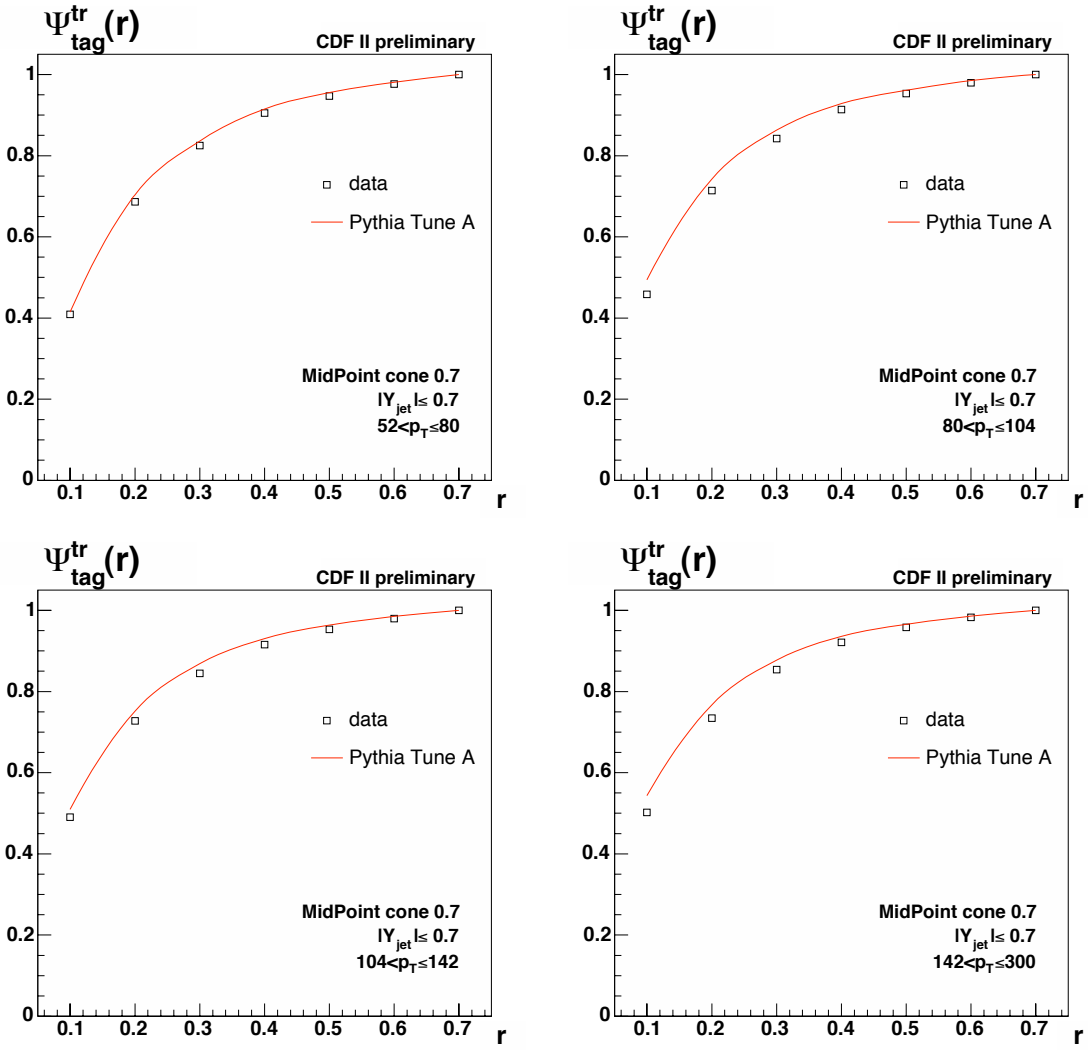

FIGURE D.13: Raw integrated shapes for tagged jets measured using tracks. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

p_T $52 - 80\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had,\,track}}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.353	0.301	0.052
0.2	0.609	0.572	0.037
0.3	0.760	0.733	0.027
0.4	0.844	0.832	0.012
0.5	0.907	0.905	0.002
0.6	0.955	0.950	0.005
0.7	1.000	1.000	0.000
p_T $80 - 104\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had,\,track}}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.356	0.352	0.004
0.2	0.595	0.583	0.011
0.3	0.746	0.730	0.016
0.4	0.845	0.830	0.015
0.5	0.918	0.906	0.011
0.6	0.966	0.961	0.005
0.7	1.000	1.000	0.000
p_T $104 - 142\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had,\,track}}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.451	0.419	0.033
0.2	0.644	0.634	0.010
0.3	0.750	0.752	0.002
0.4	0.854	0.843	0.011
0.5	0.919	0.910	0.009
0.6	0.966	0.963	0.003
0.7	1.000	1.000	0.000
p_T $142 - 300\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had,\,track}}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.413	0.429	0.016
0.2	0.642	0.653	0.010
0.3	0.771	0.769	0.002
0.4	0.867	0.854	0.012
0.5	0.931	0.922	0.009
0.6	0.976	0.967	0.009
0.7	1.000	1.000	0.000

TABLE D.10: Hadron level b-quark jet shape starting from raw shapes using tracks. Also shown is the b-quark jet shape starting from calorimeter jets shapes and the difference in the final results between these two methods.

D.9 Comparing final shapes obtained using calorimeter level towers with $p_T > 0.5\,\mathrm{GeV}$

The hadron level b-jet shapes obtained starting from the raw calorimetric shapes with the default cut on the p_T of the towers of 0.1 GeV are compared to those obtained with a cut at 0.5 GeV.

Figure D.14 shows a comparison between data and Pythia Tune A MC of the raw inclusive shapes measured with the additional p_T cut. Figure D.15 shows the same thing but for tagged jets. As for the default case, the agreement between data and MC is not perfect for the inclusive jets. The data and MC for the tagged jets are also different, as would be expected from the fact that there are different fractions of b-jets in data and MC.

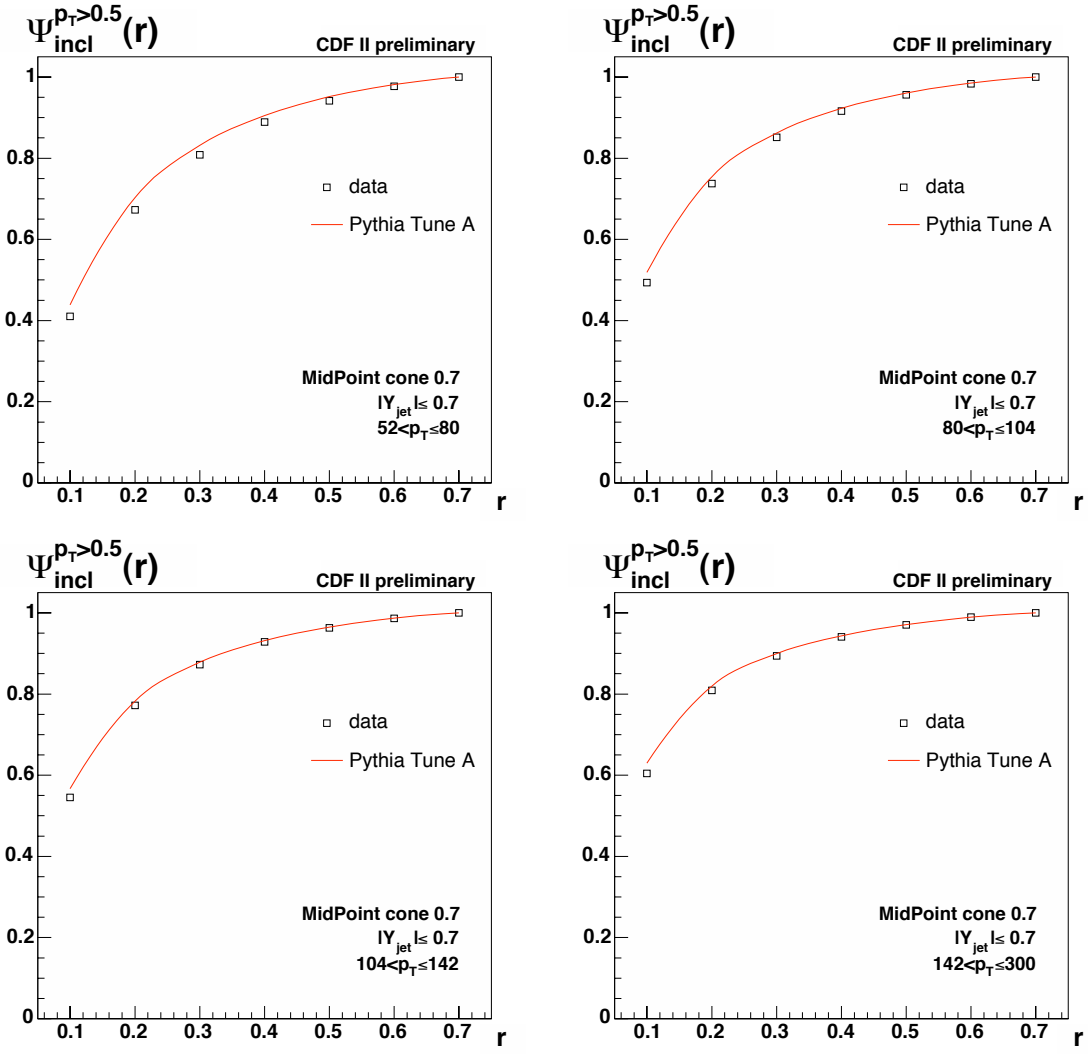

FIGURE D.14: Raw integrated shapes for inclusive jets measured using calorimeter towers with cut on the p_T of the towers at 0.5 GeV. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

The bias corrections to correct for the tagging requirement and the hadron level corrections are computed specifically for the raw shapes measured with this additional p_T cut on the towers. The final values for the hadron level shapes are show in in figure D.11 along with the values for the default cut on the tower threshold and the difference between these two values. The

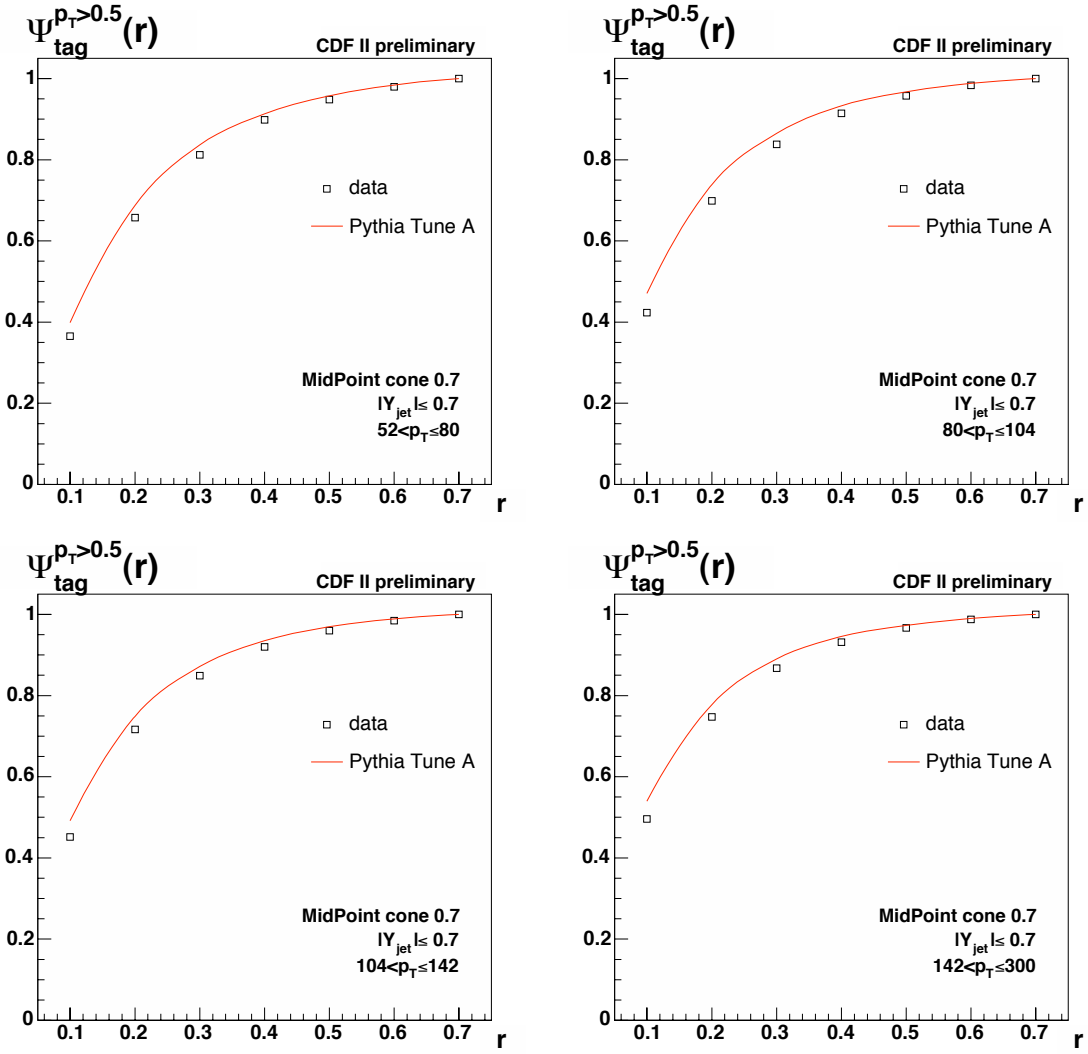

FIGURE D.15: Raw integrated shapes for tagged jets measured using calorimeter towers with cut on the p_T of the towers at 0.5 GeV. The Pythia Tune A MC predictions (red line) are compared to the data (black points). Only the statistical errors are shown.

difference between these two values is of the same size or smaller than the total statistical error. The difference between the

p_T $52 - 80\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had},\,0.5}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.303	0.301	0.002
0.2	0.578	0.572	0.006
0.3	0.740	0.733	0.007
0.4	0.837	0.832	0.005
0.5	0.910	0.905	0.005
0.6	0.952	0.950	0.002
0.7	1.000	1.000	0.000
p_T $80 - 104\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had},\,0.5}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.353	0.352	0.001
0.2	0.585	0.583	0.001
0.3	0.733	0.730	0.003
0.4	0.833	0.830	0.003
0.5	0.909	0.906	0.003
0.6	0.963	0.961	0.002
0.7	1.000	1.000	0.000
p_T $104 - 142\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had},\,0.5}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.419	0.419	0.000
0.2	0.634	0.634	0.000
0.3	0.754	0.752	0.001
0.4	0.844	0.843	0.001
0.5	0.911	0.910	0.001
0.6	0.964	0.963	0.001
0.7	1.000	1.000	0.000
p_T $142 - 300\,\mathrm{GeV}$			
r bin	$\Psi^{\mathrm{b}}_{\mathrm{had},\,0.5}(r)$	$\Psi^{\mathrm{b}}_{\mathrm{hac,\,cal}}(r)$	$\Delta\Psi^{\mathrm{b}}(r)$
0.1	0.429	0.429	0.000
0.2	0.653	0.653	0.000
0.3	0.769	0.769	0.001
0.4	0.855	0.854	0.001
0.5	0.924	0.922	0.002
0.6	0.970	0.967	0.002
0.7	1.000	1.000	0.000

TABLE D.11: Hadron level b-quark jet shape starting from raw shapes using calorimeter towers with cut on the p_T of the towers at 0.5 GeV. Also shown is the b-quark jet shape starting from the default calorimeter jets shapes and the difference in the final results between these two methods.

D.10 Total systematic errors

The hadron level b-jets shapes are computed from equation 3.7 and the total systematic and statistical errors are computed from the statistical and systematic errors of the different parameters of the equation. The total systematic errors are shown in table D.12 for each p_T bin and r bin. Also shown are the various contributions from the individual parameters.

The dominant source of systematic errors vary as a function of the p_T bin. For all p_T bins, the tagging bias on nonb-jets is at least one of the dominant sources. The other sources of systematic errors are: the difference between the final shapes reconstruction from raw track shapes and from raw calorimeter shapes for the lowest and highest p_T bins and the hadron level corrections to the b-quark jets for the highest p_T bin.

$52 - 80\,\mathrm{GeV}$

r bin	tot	stat tot	syst tot	Statistical errors			Systematic errors						
				$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	Δb_{b}	Δb_{nonb}	ΔC_{b}	$\Delta\Psi_{\mathrm{had\,trS}}$	$\Delta\Psi_{\mathrm{had\,0.5}}$
0.1	0.076	0.010	0.075	0.010	0.001	0.001	0.014	0.012	0.000	0.006	0.048	0.014	0.052	0.002
0.2	0.062	0.010	0.061	0.009	0.001	0.000	0.013	0.011	0.000	0.006	0.042	0.016	0.037	0.006
0.3	0.044	0.007	0.044	0.007	0.001	0.001	0.008	0.008	0.000	0.006	0.029	0.012	0.027	0.007
0.4	0.025	0.005	0.024	0.005	0.001	0.001	0.005	0.006	0.000	0.004	0.017	0.007	0.012	0.005
0.5	0.013	0.003	0.013	0.003	0.000	0.000	0.003	0.003	0.000	0.003	0.010	0.004	0.002	0.005
0.6	0.007	0.002	0.007	0.001	0.000	0.000	0.002	0.001	0.000	0.001	0.004	0.001	0.005	0.002
0.7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

$80 - 104\,\mathrm{GeV}$

r bin	tot	stat tot	syst tot	Statistical errors			Systematic errors						
				$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	Δb_{b}	Δb_{nonb}	ΔC_{b}	$\Delta\Psi_{\mathrm{had\,trS}}$	$\Delta\Psi_{\mathrm{had\,0.5}}$
0.1	0.040	0.007	0.039	0.007	0.001	0.001	0.012	0.010	0.000	0.004	0.033	0.013	0.004	0.001
0.2	0.032	0.006	0.031	0.006	0.001	0.002	0.009	0.007	0.000	0.004	0.024	0.011	0.011	0.001
0.3	0.025	0.004	0.025	0.004	0.001	0.001	0.006	0.005	0.000	0.004	0.014	0.008	0.016	0.003
0.4	0.018	0.003	0.018	0.003	0.000	0.001	0.004	0.003	0.000	0.003	0.007	0.004	0.015	0.003
0.5	0.013	0.002	0.013	0.002	0.000	0.001	0.002	0.001	0.000	0.002	0.004	0.002	0.011	0.003
0.6	0.006	0.001	0.006	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.002	0.001	0.005	0.002
0.7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

$104 - 142\,\mathrm{GeV}$

r bin	tot	stat tot	syst tot	Statistical errors			Systematic errors						
				$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	Δb_{b}	Δb_{nonb}	ΔC_{b}	$\Delta\Psi_{\mathrm{had\,trS}}$	$\Delta\Psi_{\mathrm{had\,0.5}}$
0.1	0.052	0.008	0.052	0.008	0.001	0.000	0.012	0.009	0.000	0.005	0.035	0.012	0.033	0.000
0.2	0.030	0.007	0.030	0.007	0.001	0.001	0.006	0.006	0.000	0.005	0.024	0.009	0.010	0.000
0.3	0.017	0.005	0.017	0.004	0.001	0.001	0.001	0.004	0.000	0.004	0.015	0.006	0.002	0.001
0.4	0.015	0.003	0.015	0.003	0.000	0.001	0.001	0.002	0.000	0.003	0.009	0.004	0.011	0.001
0.5	0.011	0.002	0.011	0.002	0.000	0.001	0.001	0.001	0.000	0.002	0.005	0.002	0.009	0.001
0.6	0.004	0.001	0.004	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.002	0.001	0.003	0.001
0.7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

$142 - 300\,\mathrm{GeV}$

r bin	tot	stat tot	syst tot	Statistical errors			Systematic errors						
				$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	$\Delta\Psi_{\mathrm{tag}}$	$\Delta\Psi_{\mathrm{incl}}$	Δp_{b}	Δb_{b}	Δb_{nonb}	ΔC_{b}	$\Delta\Psi_{\mathrm{had\,trS}}$	$\Delta\Psi_{\mathrm{had\,0.5}}$
0.1	0.040	0.008	0.039	0.008	0.001	0.002	0.014	0.012	0.000	0.003	0.026	0.016	0.016	0.000
0.2	0.030	0.006	0.029	0.006	0.001	0.002	0.014	0.008	0.000	0.004	0.018	0.012	0.010	0.000
0.3	0.018	0.004	0.018	0.004	0.001	0.002	0.009	0.005	0.000	0.004	0.011	0.007	0.002	0.001
0.4	0.016	0.003	0.016	0.002	0.000	0.001	0.005	0.003	0.000	0.003	0.006	0.004	0.012	0.001
0.5	0.011	0.002	0.011	0.001	0.000	0.001	0.003	0.001	0.000	0.002	0.004	0.002	0.009	0.002
0.6	0.010	0.001	0.010	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.002	0.001	0.009	0.002
0.7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

TABLE D.12: Total, statistical and systematic errors on the final b-jet shape along with the contributions from each parameter.

D.11 Changing the rapidity cut to $0.1 \leq |Y| \leq 0.7$

In order to compare the final b-jet shapes to the previously published inclusive jet shapes, it is important to see if the different cut requirements on the rapidity of the jets have an effect on the measured shapes. Figure D.16 shows a comparison, in data, between the measured, raw, shapes using a rapidity cut of $|Y| < 0.7$ (colour?) and those measured removing the central rapidity region, i.e. requiring $0.1 < |Y| < 0.7$ (colour). Figure D.17 shows the same comparison but for tagged jets instead of inclusive jets. Both these figures show that the measured shapes are the same for both cases. This allows us to plot the published inclusive shape results on the same plots as the b-jet shape results.

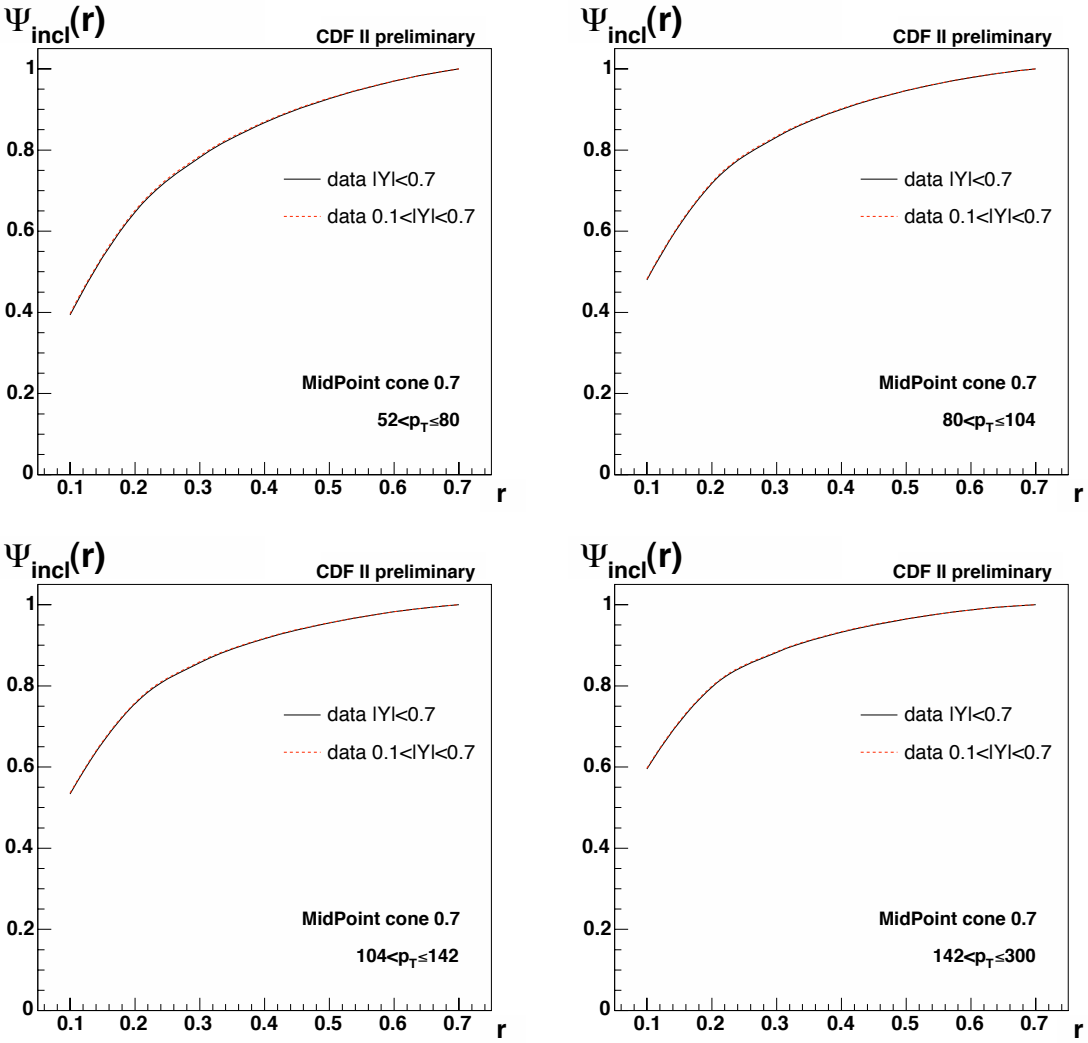

FIGURE D.16: Comparison in data between inclusive shapes with the rapidity cut $|Y| < 0.7$ (colour) and the cuts used for the measurement of the inclusive jet shapes, $0.1 < |Y| < 0.7$.

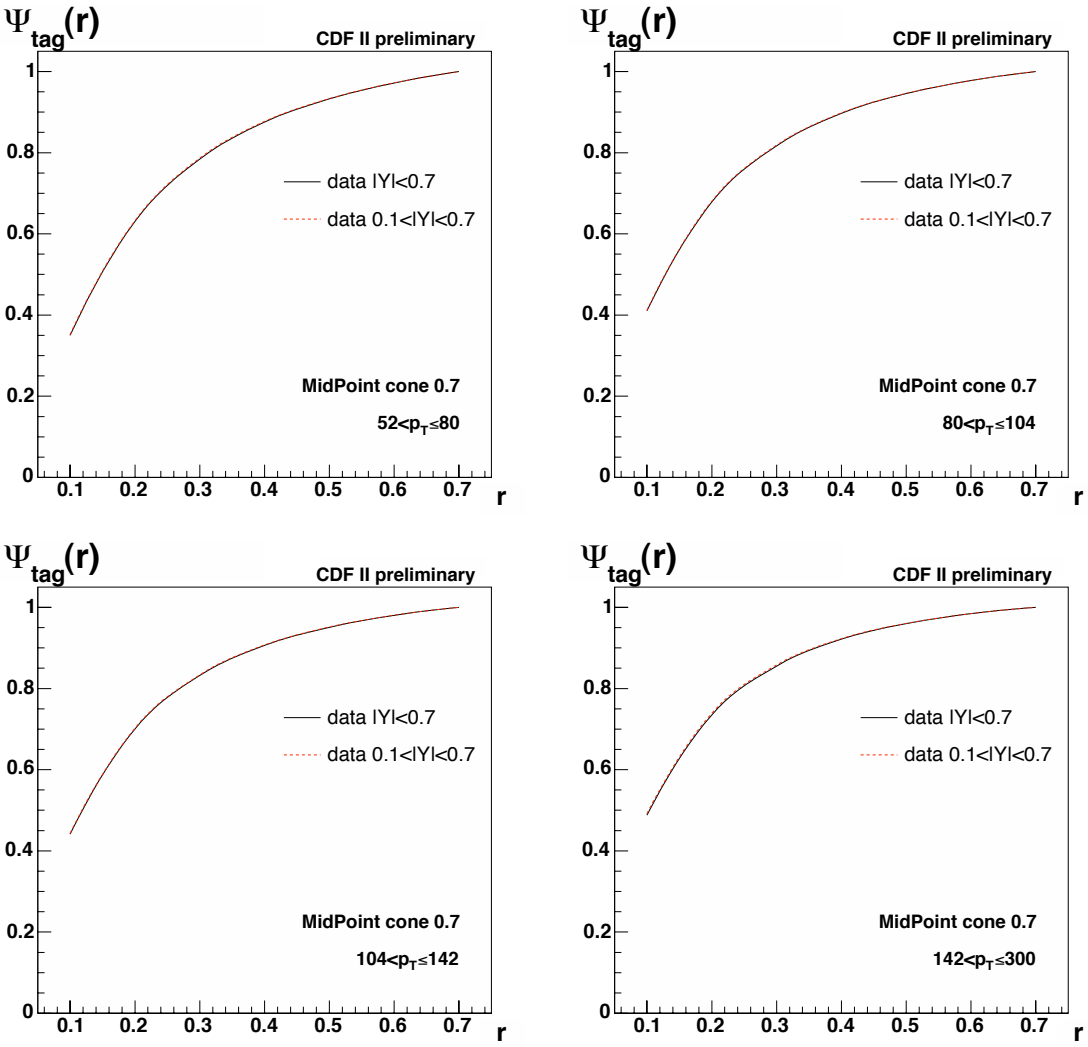

FIGURE D.17: Comparison in data between tagged jet shapes with the rapidity cut $|Y| < 0.7$ and the cuts used for the measurement of the inclusive jet shapes, $0.1 < |Y| < 0.7$.

Bibliography

[1] R. D. Field. The Underlying Event in Hard Scattering Processes. hep-ph/0201192.

[2] D. Acosta *et al* [CDF Collaboration]. Study of Jet Shapes in Inclusive Jet Production in p anti-p Collisions at $\sqrt{s} = 1.96\,\text{Tev}$. *Phys. Rev. D*, 71(112002), 2005.

[3] S. Frixione *et al*. Heavy Quark Production. *CERN-TH/97-16*, 1997. hep-ph/9702287.

[4] H.L. Lai *et al* [CTEQ Collaboration]. Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions. *Eur. Phys. J*, C. 12(375), 2000. hep-ph/990328.

[5] B. Anderson. *The Lund Model*. Cambridge University Press, 1988.

[6] G. Corcella *et al*. Herwig 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (Including Supersymmetric Processes). *JHEP*, 0101(010), 2001. HERWIG 6.5. hep-ph/0011363.

[7] H.L. Lai *et al* [CTEQ Collaboration]. CTEQ6 Parton Distributions With Heavy Quark Mass Effects. *Phys.Rev.D*, 69(114005), 2004. http://user.pa.msu.edu/wkt/cteq/cteq6/cteq6pdf.html.

[8] C. Peterson *et al*. Scaling Violation in Inclusive e^+ e^- Annihilation Spectra. *Phys. Rev. D*, 27(105), 1983.

[9] The CDF Collaboration. J/ψ, $\psi' \to \mu^+\mu^-$ and $B \to J/\psi$, ψ' Cross Sections. *FERMILAB-CONF-94-136-E*, 1994. hep-ex/9412013.

[10] M. D'Onofrio. Preliminary Results on Inclusive b-jet Cross Section. *CDF Note* 6984, 2004.

[11] M. D'Onofrio. Updated Results and Preliminary Comparison with Next to Leading Order Predictions: ADDENDUM to b-jet Cross Section Measurement. *CDF Note* 7806, 2005.

[12] M. D'Onofrio, A. Lister. SecVtx Using the MidPoint Cone Algorithm: Comparison to JetClu. *CDF Note* 7045, 2005.

[13] J. Adelman *et al*. Generic Jet Scale Corrections for Run II. *CDF Note* 7358, 2005.

[14] A. Lister. Measurement of b-quark Jet Shapes at CDF: Inclusive Method. *CDF Note* 8143, 2006.