UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MEASUREMENT OF THE DOUBLE DIFFERENTIAL DIJET MASS CROSS
SECTION IN $p\bar{p}$ COLLISIONS AT $\sqrt{s} = 1.96$ TEV

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

BY

MANDY ROMINSKY

Norman, Oklahoma

2009

MEASUREMENT OF THE DOUBLE DIFFERENTIAL DIJET MASS CROSS
SECTION IN $p\bar{p}$ COLLISIONS AT $\sqrt{s} = 1.96$ TEV

A DISSERTATION APPROVED FOR THE
HOMER L. DODGE DEPARTMENT OF PHYSICS AND ASTRONOMY

BY

_____

Dr. Michael G. Strauss, Chair

_____

Dr. Phillip Gutierrez

_____

Dr. Chung Kao

_____

Dr. Sheena Murphy

_____

Dr. Horst Severini

_____

Dr. William Beasley

Dedication

This thesis is dedicated to my husband for being patient and supportive throughout the process, and to my family and friends, who believed in me the whole way.

Acknowledgments

It is a pleasure to thank the people who made this thesis possible. My advisor, Mike Strauss, for guiding me and answering all my questions, and for letting me make mistakes along the way. The analysis would not have been possible without the hard work of everyone on the DØ collaboration. I would like to particularly thank Don Lincoln for his stimulating conversation on all topics and Markus Wobisch who always made sure I was precise in all aspects of this analysis. I also want to thank the many people I met along the way; Xuan Li, a fellow grad student who helped me understand the basics, Bill Lee and George Ginther, who taught me how to operate the detector, and Mikko Voutilainen, who was a tremendous help with the programming. Finally, I want to thank my friends and family for being so supportive while I worked on this thesis.

Contents

List Of Tables

List Of Figures

Abstract

This thesis presents the analysis of the double differential dijet mass cross section, measured at the DØ detector in Batavia, IL, using $p\bar{p}$ collisions at a center of mass energy of $\sqrt{s} = 1.96$ TeV. The dijet mass was calculated using the two highest p_T jets in the event, with approximately 0.7 fb^{-1} of data collected between 2004 and 2005. The analysis was presented in bins of dijet mass (M_{JJ}) and rapidity (y), and extends the measurement farther in M_{JJ} and y than any previous measurement. Corrections due to detector effects were calculated using a Monte Carlo simulation and applied to data. The errors on the measurement consist of statistical and systematic errors, of which the Jet Energy Scale was the largest. The final result was compared to next-to-leading order theory and good agreement was found. These results may be used in the determination of the proton parton distribution functions and to set limits on new physics.

Chapter 1

Introduction

Hadron colliders probe interactions between partons at small distance scales, which are governed by Quantum Chromodynamics (QCD), the theory of strong interactions in the Standard Model (SM). Since jet production has one of the highest cross sections at hadron colliders, jet properties are ideal final states to study QCD. In this thesis, the inclusive dijet mass is studied as a double differential in rapidity[1] (a quantity related to the polar angle) and dijet mass. The dijet mass cross section at a hadron collider is particularly sensitive to the gluon parton distribution function (PDF), which is not currently well measured. In addition, the dijet mass can be used to search for quark compositeness and decays of exotic particles to a 2-jet final state.

The dijet mass is calculated using the two jets with the highest transverse momentum[2] (p_T) in an event according to the standard invariant mass formula, $M = \sqrt{E^2 - \vec{p}\,^2}$, where the energy and momentum are the sums of the energies and momenta of the two jets. The measurement is performed using approximately 0.7 fb^{-1} of Run II data from the DØ detector located at the Tevatron in Batavia, IL, a proton-antiproton ($p\bar{p}$) collider with a center of mass energy of $\sqrt{s} = 1.96$ TeV. The rapidity has been extended to $|y| < 2.4$, where previous measurements ([1], [2], [3]) only measure out to $|y| < 1.0$. In addition, the largest systematic uncertainty (Jet Energy Scale) has been reduced by a factor of two over the majority of the p_T range as compared to the uncertainty in [1]. The final results are compared to next-to-leading order (NLO) QCD theory and found to be in good agreement.

The thesis is divided into three broad categories: theory, detector and analysis. The theory section will review the Standard Model and the specific parts of Quantum

---

[1] $y = 1/2 \ln[(E + p_z)/(E - p_z)]$

[2] $p_T = \sqrt{p_x^2 + p_y^2}$. The transverse momentum is invariant under longitudinal boosts.

Chromodynamics that pertain to this analysis. The detector section describes both the Tevatron, the accelerator complex where the experiment is performed, and the DØ detector. The final section details the steps involved in calculating the dijet mass and discusses the corrections, uncertainty calculations, and the final results. The first two appendices provide information on common quantities and terms used in particle physics and will be helpful during the course of reading the document. The other appendices contain useful supplementary information.

Chapter 2

Theory

The motivation for this analysis is to verify the Standard Model of particle physics, particularly Quantum Chromodynamics (QCD), and to look for deviations from the SM, indicating new physics. This section will introduce the theory behind the dijet mass measurement. It begins with a discussion of the Standard Model, followed by a more in-depth discussion of QCD. The last section details how the dijet mass spectrum can be used to probe for new physics.

2.1   Standard Model

The Standard Model represents physicists' current best knowledge of the fundamental building blocks of matter based on experimental evidence and theory. It postulates that all matter (and antimatter) is made of 6 leptons, 6 quarks (and their antiparticles) and 4 force mediators. The leptons and quarks are fermions (half integer spins), while the force carriers are bosons (integer spins). The Standard Model is comprised of the theory of Quantum Electrodynamics (QED) combined with weak theory into the electroweak theory by Glashow, Weinberg, and Salam [4], and Quantum Chromodynamics (QCD) by Politzer, Gross, and Wilczek [5].

There are six leptons in the Standard Model; three charged particles (the electron (e), the muon (μ) and the tau (τ)), with three corresponding neutrinos (ν_e, ν_μ, ν_τ). Table 2.1 lists information about each one. The μ and τ have the same basic properties (charge, spin) as the electron, but are more massive. The neutrinos are assumed to be massless in the Standard Model, but in recent years there has been evidence they do have very tiny masses [6]. All leptons obey Fermi-Dirac statistics and interact

via both the electroweak force and gravity. Similarly there are six quarks (see Table 2.2), grouped into three generations, with each successive generation being more massive. Quarks obey Pauli-Dirac statistics and are the only class of particle that can participate in all types of interactions (electroweak, strong, and gravity).

Flavor	Charge (units of proton charge)	Mass (MeV)
electron (e)	- 1	0.511
ν_e	0	< 0.000002
muon (μ)	-1	105
ν_μ	0	< 0.19
tau (τ)	-1	1776
ν_τ	0	< 18

Table 2.1: Charge and mass for leptons [7].

Flavor	Charge (units of proton charge)	Mass
down (d)	$-\frac{1}{3}$	3.5 - 6.0 (MeV)
up (u)	$\frac{2}{3}$	1.5 - 3.3 (MeV)
strange (s)	$-\frac{1}{3}$	105 (MeV)
charm (c)	$\frac{2}{3}$	1.27 (GeV)
bottom (b)	$-\frac{1}{3}$	4.2 (GeV)
top (t)	$\frac{2}{3}$	171 (GeV)

Table 2.2: Charge and mass for quarks [7].

Mediator	Force	Mass (GeV)
photon (γ)	electromagnetic	0
gluon (g)	strong	0
$W^\pm$	weak	80
Z	weak	91

Table 2.3: Symbol, force and mass for the force mediators in the Standard Model [7].

The forces included in the Standard Model are the electromagnetic, weak and strong. The electromagnetic force and the weak forces were combined into the electroweak force by Sheldon Glashow, Abdus Salam, and Steven Weinberg in the 1960's.

4

The mediators for this force are the photon (electromagnetic) and the $W^\pm$ and Z^0 (weak). The gluon mediates the strong force between quarks. Table 2.3 gives relevant information on these force carriers.

While the Standard Model has been successful for the most part in describing experimental results, it does leave some gaps. The largest is its failure to incorporate gravity. In addition, the masses of the particles are not derived from first principles. The Higgs mechanism was introduced to explain particle masses, adding a new particle called the Higgs boson. The Standard Model needs gauge invariance for consistency and the electroweak symmetry must be broken spontaneously by the Higgs field getting a non-zero vacuum expectation value, thus generating the mass for elementary particles. The Higgs boson is part of a weak isospin complex doublet that has four degrees of freedom. After the symmetry is broken, three of those degrees of freedom become the longitudinal components of the $W^\pm$ and Z^0 bosons, generating their masses, and the last degree of freedom becomes the Higgs boson. While this electroweak symmetry breaking is required in order for particles to have mass, the Higgs boson has not been observed, meaning alternative symmetry-breaking theories should be considered. In addition there are theories that expand the Standard Model, such as supersymmetry, which postulates additional particles. In typical supersymmetry models, for each known fermion there is a supersymmetric boson and vice versa. While no evidence of new physics beyond the Standard Model has been observed to date, efforts are underway to put limits on various models. The dijet mass is a measurement that could be used to study some of these models, particularly ones that probe small distance scales.

2.2   Quantum Chromodynamics

The theory probed by this thesis is Quantum Chromodynamics (QCD), the study of the interactions between colored objects[1] via the strong force. In QCD, quarks and gluons have a property called color charge, similar to the electric charge. This color charge produces interactions via gluons analogous to the electric charge producing interactions via the photon. In addition, unlike the electrically neutral photon, the color-charged gluons can interact with each other. Each quark has a color (red, blue,

---

[1]Any object that has a color charge, such as quarks and gluons, is subject to the strong force. In the Standard Model, only quarks and gluons have color charge.

and green) associated with it, antiquarks have anticolors (antired, etc), and gluons have 1 color and 1 anticolor (blue, antired). Thus far in experiments, only colorless combinations of particles have been found, composed either of all 3 colors or a color and anticolor.

Deep inelastic experiments performed at SLAC in the late sixties led physicists to observe that nucleons have internal structure. The basic interaction for these experiments was $eN \rightarrow e'X$, where e is the incoming electron, e' is the outgoing electron, N is the nucleon that is hit by the electron and X is the resulting hadrons. The constituent particles of the nucleon were called partons, and later determined to be quarks. The parton model [8] was developed to explain the results of these experiments. Two important concepts were developed at this time; quark confinement and asymptotic freedom.

Quark confinement explains why we never see direct evidence of quarks and gluons in experiments. There is nothing in the parton model to forbid quarks from being free, so some other mechanism must be causing this. The answer lies with the fact that the QCD potential has an approximately linear behavior at low energies due to the self interaction of gluons that continues to grow as the distance between quarks becomes large. At a certain point (about the radius of the proton), it becomes more efficient to create quark-antiquark or gluon pairs from the vacuum than to continue stretching.
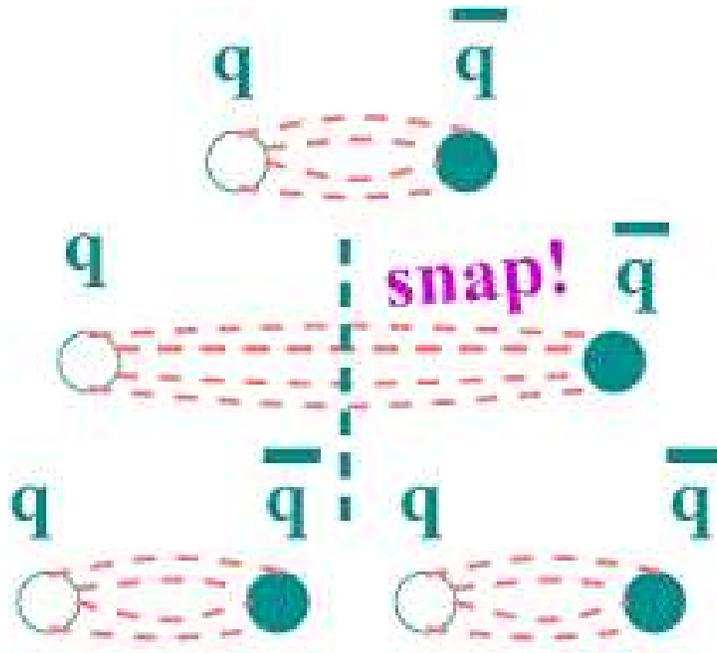
Figure 2.1: Cartoon of a quark and antiquark stretching and producing pairs from the vacuum [9]. Note that this is an example; the same is true for a gluon pair. As the pair is forced farther and farther apart, the potential grows. Eventually, it becomes more efficient to create a new pair from the vacuum and the "string" holding them together breaks.

The second concept to come out of these experiments is asymptotic freedom, developed by Politzer, Gross and Wilczek. Quarks interact via the strong force in general, but at short distances, quarks behave as if they are hardly interacting. How can this be true? The answer is that the strong coupling constant α_s is not actually constant, but is "running". The strong coupling constant depends on the energy transferred (Q), rapidly decreasing as the energy increases. Note that the energy scales as the inverse of distance ($1/r$), so another way to look at the running coupling constant is to say that α_s decreases as the distance between quarks decreases. Equation 2.1 [10] shows α_s

$$\alpha_s(Q) = \frac{4\pi}{(11 - \frac{2}{3}N_f)\ln(\frac{Q^2}{\Lambda^2})} \tag{2.1}$$

where N_f is the number of quark flavors, Q is the energy scale and Λ is the characteristic QCD scale. Figure 2.2 shows the currently measured value of α_s with data combined from various experiments [11].
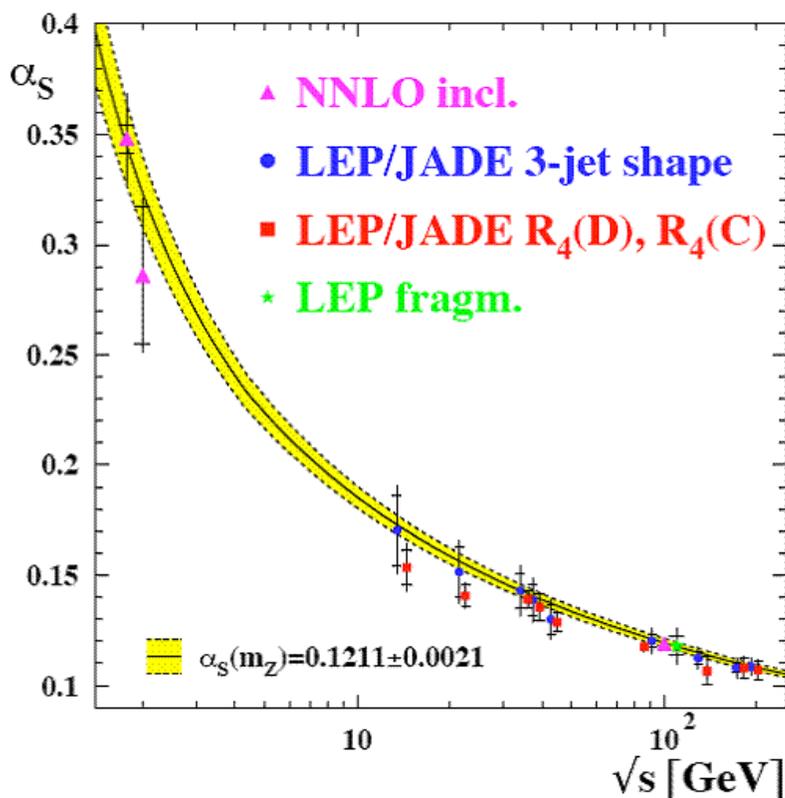
Figure 2.2: The strong coupling constant as a function of energy [11].

2.2.1 Calculations in QCD

QCD can be broken into two types of calculations: nonperturbative and perturbative. Nonperturbative QCD is not calculable, but can be approximated using phenomenological methods and will not be discussed here. However the effects of the nonperturbative physics must be taken into account in data measurements. See Section 2.2.5 for more information on how this is handled. Reference [12] has information on nonperturbative QCD for those interested. Any non-linear formula (such as the one describing quark scattering) can be expressed as a perturbative expansion. As long as the expansion variable is relatively small, the theory is well described by the first few terms. In quantum field theories, the coupling constant can be used as the expansion variable. In QED, the coupling constant is relatively weak ($\alpha_{EM} \approx 1/137$) and perturbation theory works very well. However in QCD, the coupling constant is close to unity at low energies. Instead of each successive term contributing less and less values to a cross section, the terms in the expansion continue to grow and the expansion won't terminate. Perturbative QCD will work at a short distance scale and when the final number of partons are fixed (i.e. $2 \rightarrow 3$ processes), however this is not what is measured at experiments. In order to use perturbative theory more generally in QCD, certain methods are employed and will be discussed in this section. Currently, pQCD has been successful at predicting broad trends in the experimental data, when properly renormalized.

The basic perturbative cross section formula for hard scattering in hadronic collisions is given in Eq. 2.2 [15],

$$\sigma(p\bar{p} \rightarrow X) = \sum_{ij} \int dx_1 dx_2 f_i(x_1, \mu_F) f_j(x_2, \mu_F) \hat{\sigma}(ij \rightarrow X) \qquad (2.2)$$

where f_i, f_j are parton distribution functions[2] (PDFs) of the initial partons and $\hat{\sigma}$ is the partonic cross section, dependent on the parton energies and the renormalization scale, μ_R. The PDFs cannot be calculated; they must be measured and are dependent on the factorization scale, μ_F. The partonic cross section can be calculated using perturbation theory.

There are 2 important concepts hidden in this calculation; the factorization scale (μ_F) and the renormalization scale (μ_R). These two scales are not physical, but

---

[2]Parton distribution functions describe the momentum from the proton is distributed among its constituent partons. These will be discussed in more detail in Section 2.2.2.

must be included in order for us to use perturbation theory to solve QCD problems. The renormalization scale absorbs singularities that come from multiple orders of α_s being included[3] while the factorization scale is used to separate long and short distance calculations (infrared singularities), and distinguishes between confinement and asymptotic freedom. Note that as more and more orders are added, the dependence on these scales should go to zero. They are only used to make the calculations easier. Infrared singularities arise from the emission of low energy gluons which causes the formula to blow up in the limit where the gluon energy approaches zero. To illustrate some of the problems with infrared singularities, consider $e^+e^- \rightarrow q\bar{q}g$, where we have electrons and positrons colliding and a quark, antiquark and gluon in the final state. In this case, we have the differential cross section given in Eq 2.3 [10].

$$\frac{d\sigma}{dx_1 dx_2} = \sigma_0 \frac{\alpha_s}{2\pi} C_F \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)} \tag{2.3}$$

where σ_0 is the leading order cross section, C_F is a group constant[4] and x_i indicates the momentum fraction of the final state quarks. Figure 2.3 shows a diagram of the phase space used in this interaction, assuming the quark masses are zero. The edges represent the case where two of the partons are collinear[5], while the corners show points where two partons are back to back and the third has a vanishingly small energy. It is at these edges of phase space that the calculation breaks down. In Equation 2.3, the allowed values of x_i are between zero and one. If the gluon is soft ($x_3 \approx 0$) then x_1 or x_2 goes to 1, the denominator goes to 0 and we get an infinity, if the gluon is collinear to either the quark or antiquark, then $x_3 \approx 1$ and $x_{1,2}$ also equals one, and $(1-x_i)$ will go to zero, also infinity. Collinear and infrared singularities occur whenever there are many indistinguishable final states, i.e. a $2 \rightarrow 2$ process is mathematically the same as a $2 \rightarrow N$ process. In order to make the math tractable, cutoff scales (μ_F, μ_R) are used.

---

[3]A factor of $\sqrt{\alpha_s}$ is added for each vertex in a Feynman diagram, seen later in this section.

[4]A group constant is part of the non-abelian algebra used in the theory.

[5]Collinear means two of the three partons are moving in the same direction and are indistinguishable from each other.
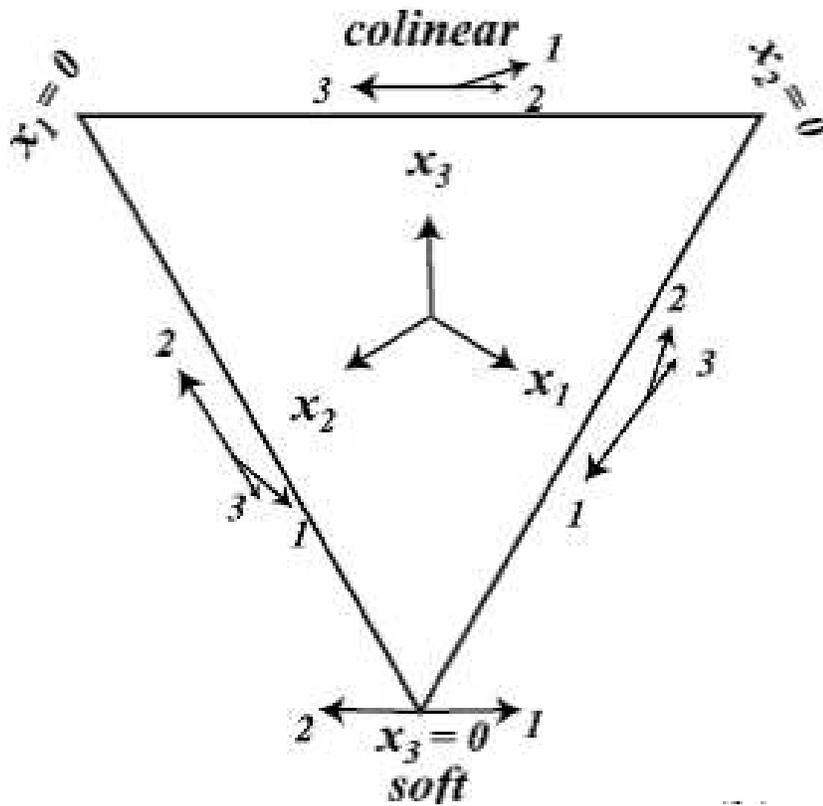
Figure 2.3: Cartoon showing how a $2 \to 3$ process can break down at certain points in phase space to a $2 \to 2$ process [10], meaning that the two states are mathematically indistinguishable.

In order to avoid infrared singularities, the Kinoshita-Lee-Nauenberg (KLN) theorem ([10],[14]) can be invoked. This theorem states that although the scattering matrix may have singularities, the sum of the squares of the scattering amplitude of degenerate states is finite. Degenerate states are states with the same final four vector values and the same quantum numbers are conserved. In other words, degenerate states are mathematically indistinguishable. In QCD, this leads to final states with jets, where there are many particles. A jet with N particles will produce the same cross section as a jet with $N+1$ particles. Variables with this property, called inclusive variables, are useful for studying QCD. The inclusive dijet mass measured in this thesis is such a variable, as are many jet variables and event shape variables [13].

So what about quantities that are not infrared safe? These can be calculated using factorization. Factorization breaks up an interaction into a long distance, non-infrared safe variable and a short distance variable that can be calculated by pQCD. This is best seen with an example. Much like the example of infrared safety, we start with $e^+e^- \to A(p) + X$, where $A(p)$ is a hadron or parton with momentum p. The cross section can be written as [13]

$$\frac{d\sigma}{dz\, d\cos\theta} = \frac{\pi\alpha^2}{2s}[F_A^T(x,Q)(1+\cos^2\theta) + F_A^L(x,Q)\sin^2\theta] \tag{2.4}$$

where F_A^i are structure functions[6] dependent on x, the fractional momentum carried by the partons and Q, the momentum transferred. In addition, the structure functions depend on the strong coupling constant α_s. The angle θ is the center-of-mass scattering angle of the incoming hadron A. We can rewrite F_A^i as [13]

$$F_\alpha(Q,x,m) = \sum_a \int \frac{dz}{z} \hat{F}_a(\frac{x}{z}, \frac{Q}{\mu_F}, \alpha_s) \otimes D_\alpha^a(z, \frac{m}{\mu_F}, \alpha_s) + \mathcal{O}\left(\frac{m^2}{Q^2}\right) \tag{2.5}$$

Written in this form, we have broken the structure functions into a calculable function from hard partonic scattering, $\hat{F}$ and fragmentation functions D_α^a, which can be measured, and $\mathcal{O}\left(\frac{m^2}{Q^2}\right)$, which refers to higher order terms.

Collinear singularities are dealt with by rewriting the parton distribution functions using renormalizable group equations. These will be discussed more in Section 2.2.2.

---

[6]Structure functions are used to characterize the cross section into an easy to understand format. The parton distribution functions are derived from these. For a good accounting of how this works, see [13].

Now that we have the tools for dealing with singularities, we can actually perform calculations. Since this is a perturbative theory, each calculation can be broken up into orders of α_s. It is helpful to use the Feynman diagrams shown in Figure 2.4 when discussing the various levels of calculation.
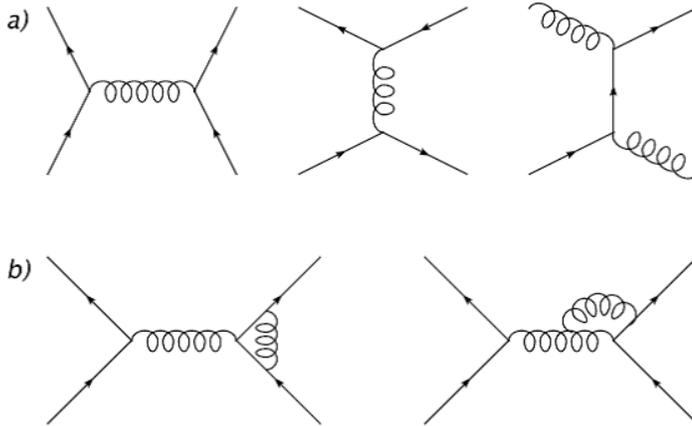
Figure 2.4: Examples of Feynman diagrams for various orders of QCD processes. The top row shows leading order (LO) diagrams and the bottom shows next-to-leading order (NLO) diagrams [16].

The first order is called leading order (LO) or Born level calculations, (a) in Figure 2.4. This is the easiest to do, only taking into account the tree level calculations, using Feynman diagrams. However, this level of calculations does not describe the interaction very well, only giving us an idea of the shape of the distribution. At leading order, all the jets are assumed to have only one parton in them and there is no way to get a handle on the internal structure of the jet. Additionally, there are logarithmically divergent terms that must be renormalized away, introducing a non physical factor. For these reasons, additional orders in calculations are needed. The next level is next-to-leading order (NLO), which is a more precise calculation, (b) in Figure 2.4. NLO takes into account all leading order diagrams and loop diagrams that result from virtual interactions. Each successive order of calculation brings more precision to the measurement, as well as more information about the substructure of the jet. However, there is a trade off in difficulty. As more orders are included, the integrals used to calculate the diagrams become more complicated. Currently, theorists can calculate up to NNLO (next-to-next-to-leading-order), with work underway to calculate higher orders [17].

To summarize, QCD calculations are broken into two types: nonperturbative and perturbative. Perturbative QCD is used to calculate theoretical predictions that have been confirmed by experimental results. In order to use pQCD, infrared singularities are removed by factoring the cross section into calculable and non-calculable (but measurable) parts and collinear singularities are removed by rewriting the PDFs. In addition, the renormalization scale (μ_R) and factorization scale (μ_F) are used to set the energy at which pQCD can be used.

2.2.2 The Proton Structure and Parton Distribution Functions

We first understood the structure of hadrons through deep inelastic scattering experiments. These experiments proved that hadrons such as the proton have internal structure. Partons all carry a certain fraction of the momentum of the proton, characterized by parton distribution functions (PDFs). In the simplistic view, protons are made of two up quarks and a down quark. However, when PDFs were first measured the total momentum of the three quarks only added to about 35% of the momentum of the proton. Another half of the remaining momentum comes from gluons, while approximately 15% of the momentum comes from "sea" quarks. These are pairs of quarks and antiquarks, usually up or down, that can pop in and out of the vacuum briefly because of interactions between gluons. Figure 2.5 shows a cartoon of the proton structure, as it is currently understood.
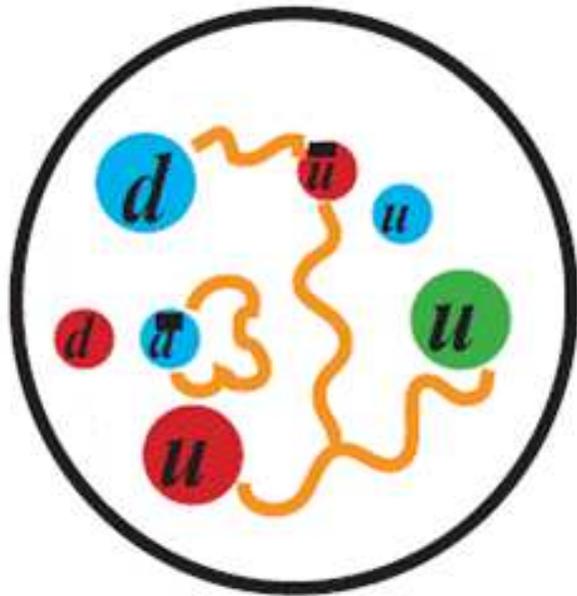
Figure 2.5: Cartoon showing the proton structure. The valence quarks are shown as the large red, green and blue circles. The squiggly lines are the gluons and the sea quarks are the red and blue pairs of small circles [16].

Parton distribution functions provide a way to study partons inside hadrons by describing their momentum distributions. PDFs must be universal (able to be used in any calculation), separate long and short range physics scales, and contain the non perturbative part of the interaction. They are used to absorb the singularities for each order of α_s. However, this loses the advantage of asymptotic freedom, since the singularity ends up being a constant, instead of disappearing. Therefore all these values need to be resummed using a set of renormalizable group equations (RGE), called the DGLAP[7], shown in Eq. 2.6 [18].

$$Q^2 \frac{dG_q(x, Q^2)}{dQ^2} = \frac{\alpha_s(Q^2)}{2\pi} \int \frac{dy}{y} [P_{qq}(y)G_q(x/y, Q^2) + P_{qg}(y)G_g(x/y, Q^2)] \quad (2.6)$$
$$Q^2 \frac{dG_g(x, Q^2)}{dQ^2} = \frac{\alpha_s(Q^2)}{2\pi} \int \frac{dy}{y} [P_{gq}(y)G_q(x/y, Q^2) + P_{gg}(y)G_g(x/y, Q^2)]$$

Using these equations will also remove collinear singularities.

PDFs are determined using global fits to data. Various groups, such as CTEQ [22] and MSTW [23], take experimental data in a parameterized form and evolve them using the DGLAP equations. They then vary the parameters and repeat, until an optimal fit to data is obtained. They use many different types of data, such as hadronic jet production, deep inelastic scattering and vector boson production. Hadronic jets are particularly useful at determining the gluon parton distribution function [18].

Figure 2.6 shows the PDFs as they are currently measured for the proton. The valence quarks carry the most momentum at high x, as expected, while the gluon PDF dominates at low x.

Figure 2.7 shows PDF sets from CTEQ and MSRT compared to the new MSTW PDF set. The dijet mass analysis can be used to reduce the uncertainties shown in the plot as bands around the central lines.

---

[7]These equations were developed by Gribov and Lipatov (1972) [19], Altarelli and Parisi (1977) [20], and Dokshitzer (1977) [21].
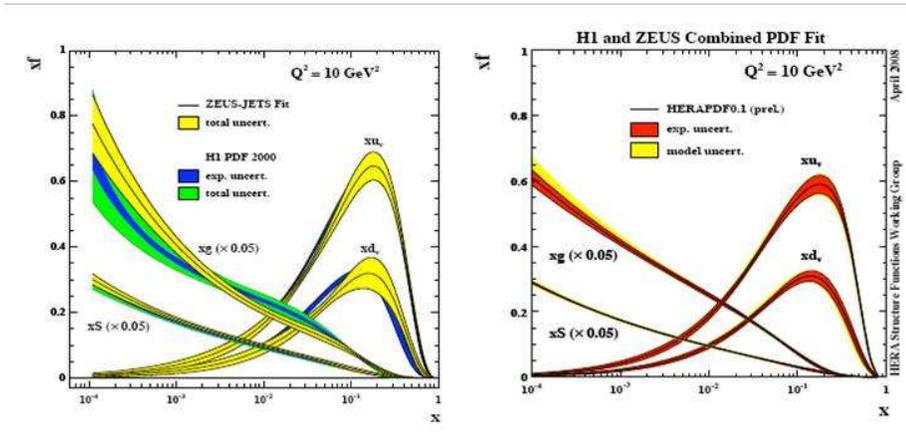
Figure 2.6: The proton PDFs as currently measured. The plot on the left shows a comparison of PDFs sets from HERA and CTEQ and the right plot shows a comparison between HERA and MSTW PDFs [24].
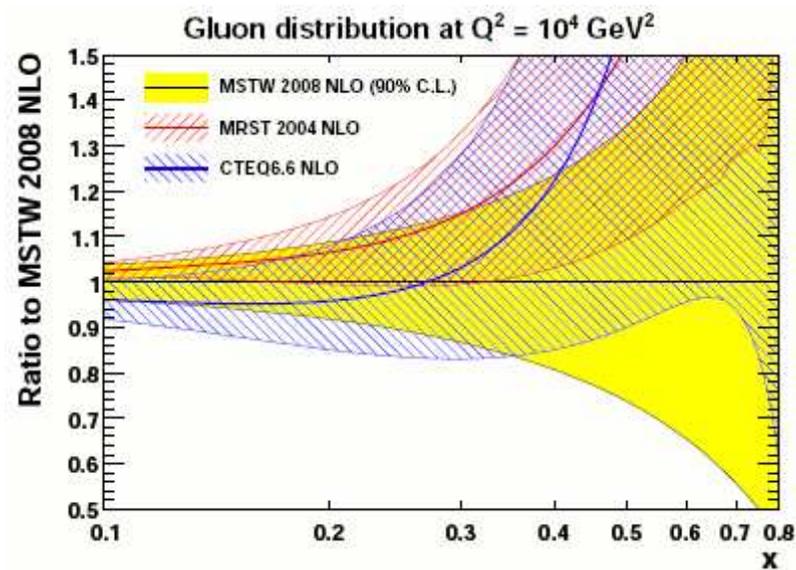
Figure 2.7: Figure from MSTW collaboration comparing various PDFs to the latest measurement from MSTW [23].

2.2.3 Jets

Ultimately, we strive to measure the energies of the particles deposited in the calorimeter after the hard scatter. To do this, the particles are formed in to jets, where a jet is defined as a collimated group of particles using jet algorithms. The dijet mass is then measured from the two highest p_T jets in an event. This section will describe how jets are formed and what type of algorithms are used to define them.

Figure 2.8 shows a simplistic cartoon of what happens during a hadronic collision that produces jets at $p\bar{p}$ colliders, to leading order. The protons and antiprotons approach, characterized by their parton distribution functions, and one parton from each participates in the hard scatter ($\hat{\sigma}$). The remaining partons may participate in soft collisions, which needs to be corrected for and will be discussed in Section 2.2.5. After the collision, the partons fragment according to fragmentation functions ($D(z, \mu_F)$) and produce a shower of particles. These then combine into hadrons at a long distance from the collision in a process called hadronization. This is not a completely calculable feature of the theory, but can be determined using fits to experimental data. The hadrons formed in the hard scatter proceed to interact with material in the detector and produce more particles. The jets used in this measurement are formed from these particles.
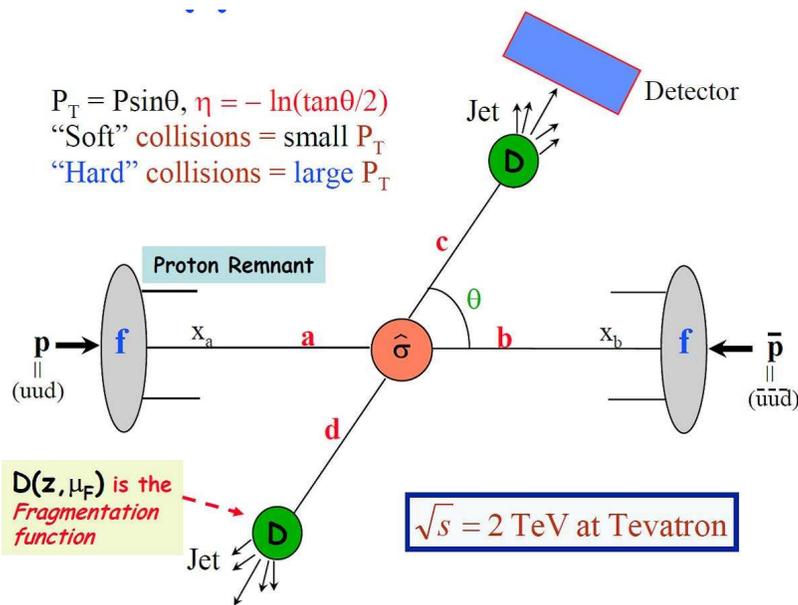
Figure 2.8: Cartoon illustrating the simple nature of leading order approximations.

Once the particles interact with the detector, they are grouped into jets according to jet finding algorithms. There are certain desirable attributes all algorithms should have [26]:

- Algorithms should be fully specified. This means all kinematic variables for the jets and corrections (i.e. underlying event) should be defined. All special features should be fully defined as well, such as preclustering, splitting and merging.

- Algorithms should be theoretically well behaved, i.e. infrared and collinear safe[8].

- Algorithms should be independent of the detector.

- Algorithms should be independent of the order of the jets (parton, particle or detector).

More specifically, there are desirable theory behaviors:

- Algorithms are free of infrared singularities and insensitive to soft radiation.

- Algorithms should be free of collinear singularities and find jets that are insensitive to collinear effects.

- Algorithms should be invariant under longitudinal boosts.

- Algorithms should be stable at the boundary of the phase space for the event.

and desirable experimental features:

- Algorithms shouldn't make resolution or angle bias corrections any larger.

- Algorithms should be stable with increasing luminosity, i.e. the algorithm shouldn't be affected by multiple hard scatters in an event.

- Algorithms should use computing power efficiently, although this is not as important as making sure the algorithm is fully specified.

- Algorithms should maximize the efficiency of reconstruction in order to ensure all jets of interest are found.

---

[8]Infrared safety and collinearity are discussed in Section 2.2.1.

- Algorithms should be easy to calibrate.

- Algorithms should be straightforward to implement.

Jet finding has two steps: finding jet constituents and calculating their kinematic variables. Recombination schemes determine how the particles are grouped together. For example, in the E-scheme, the particle's four momentum is used. Two particles are combined into one by adding together their four momentum, as long as they are within a certain distance of each other. This scheme is used for massive jets. Another recombination scheme, E_0, uses the energy of the particle instead of the full four vector. This produces massless jets. For more information on recombination schemes, see [26].

There are two main types of algorithms used to find jets; cone algorithms and clustering algorithms. Clustering algorithms, such as the k_T algorithm [9], group particles together using energy. There is no defined shape, which can make calibration hard for detectors. In this type of algorithm, jets are formed by adding together pairs of partons/particles/towers by relative transverse momentum (k_T). A parameter D, typically between 0.5 and 1, is used to determine the cut off value for the jets. After this point, a new jet is formed. These types of algorithms are used at lepton colliders and to a lesser extent at hadron colliders. They have the advantage of being infrared safe (all energy in an event is assigned to a jet), no overlapped jets, and they are less sensitive to hadronization effects than cone algorithms. However, they have the drawback of being difficult to calibrate, due to the fact that there is no defined shape and expensive computationally, since all the energy is accounted for. At lepton colliders, this isn't as large a problem, because events tend to be cleaner and easier to group all the energy. At hadron colliders, the computing is very expensive, since events are very messy. Figure 2.9 shows a flow chart explaining the k_T jet finding algorithm.

The cone algorithm sums up energy within a cone of radius $\mathcal{R}$ defined by

$$\mathcal{R} = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.7}$$

Note that while the cone radius is defined here by η, in the Run II cone algorithm, η has been replaced by y. Cone algorithms have the advantage of being easy to calibrate, since the jet shape is constant, and can be easier on computing resources, but cones are not inherently infrared safe, because some of the energy in the event
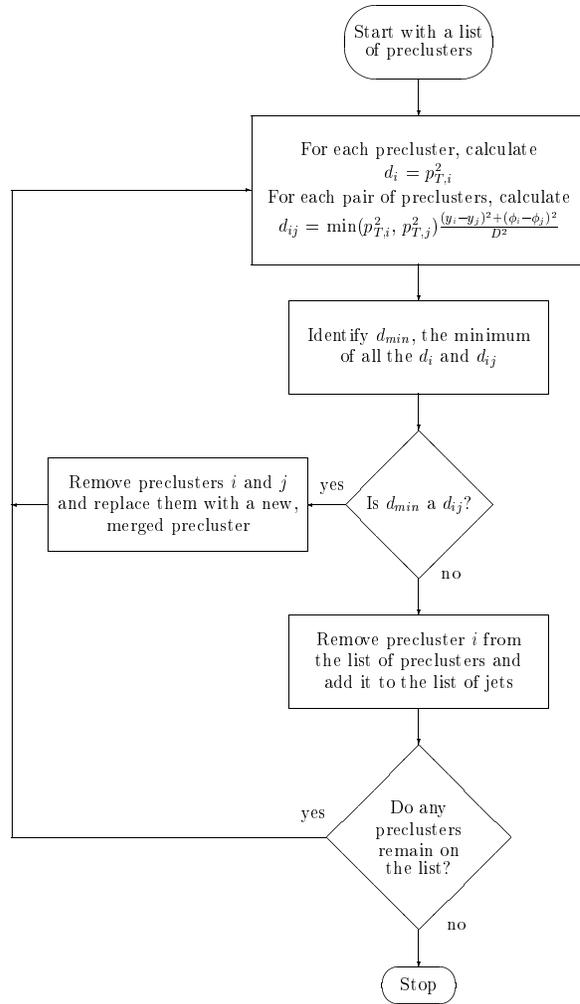
Figure 2.9: A flow chart for the k_T algorithm [26].

The flow chart contains the following elements:

Start with a list of preclusters

For each precluster, calculate
$$d_i = p_{T,i}^2$$
For each pair of preclusters, calculate
$$d_{ij} = \min(p_{T,i}^2, p_{T,j}^2)\frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{D^2}$$

Identify d_{min}, the minimum of all the d_i and d_{ij}

Is d_{min} a d_{ij}?

yes → Remove preclusters i and j and replace them with a new, merged precluster

no → Remove precluster i from the list of preclusters and add it to the list of jets

Do any preclusters remain on the list?

yes

no → Stop

won't be reconstructed. However, there are certain modifications, discussed below, that will make cone algorithms infrared safe. A cone algorithm works by looking at an initial geometric cone in $\eta - \phi$ space and calculating the energy weighted centroid of all the particles within that space. A new point in $\eta - \phi$ is chosen based on energy centroid. This process is iterated until a stable centroid is found, where the energy deposit centroid aligns with the geometric center of the cone. The starting point for these cones can either be seedless, meaning all the energy in an event is looked at, or can use seeds. Seeds are towers/partons/particles above a certain p_T threshold, so the algorithm only looks certain places. Seedless algorithms are infrared safe and not sensitive to collinear singularities, but are very computationally intensive. In Run I, DØ used the Snowmass algorithm [27], which defined jet variables using massless four vectors and a seed based search pattern. This algorithm has the disadvantage of overlapping jets, so split/merge routines are necessary. These routines set a limit on the energy (say 50%). If the shared energy is greater than 50% of the jet energy, it's merged, otherwise, a new jet is formed. For more information on how the Snowmass algorithm and cone algorithms in general work, see [27], [9].

This analysis uses the DØ Run II cone algorithm [26]. The measurements performed in [3] and [1] used massless jets and the Snowmass [27] jet finding algorithm. In order to determine massive jets, each cell is defined as a four vector: $(E_i, p_{x,i}, p_{y,i}, p_{z,i})$, where E_i is the energy in cell i and the direction comes from the vector pointing from the interaction point to the center of the cell. The full jet four momentum vector$(E_J, p_{x,J}, p_{y,J}, p_{z,J})$ can be calculated from the cell four vectors, in a recombination scheme called E-scheme[26]. The E recombination scheme simply adds the cell four vectors to form the jet four vector.

In order to optimize computing, the Run II algorithm uses a seed based cone algorithm with the addition of midpoints, which allows a seed-based algorithm to approximate a seedless algorithm. In this algorithm, towers are only considered if they have a higher energy than some threshold ($E_T^{tower} > E_T^{seed}$). This allows the algorithm to only focus on a few towers instead of all and greatly reduces computing power. Midpoints are starting positions added to the list of seeds that are determined by $p_j + p_i$ (momentum of the seeds). It is sufficient to look at jets where $\Delta R > 2R_{cone}$, since widely scattered proto-jets won't be clustered. Depending on what is found, the energy will be incorporated into its own jet or merged with an existing jet. This way all the energy gets accounted for and the effects from soft/collinear gluons are

minimized. The jet finding cone algorithm used in Run II is summarized in the Fig 2.10.
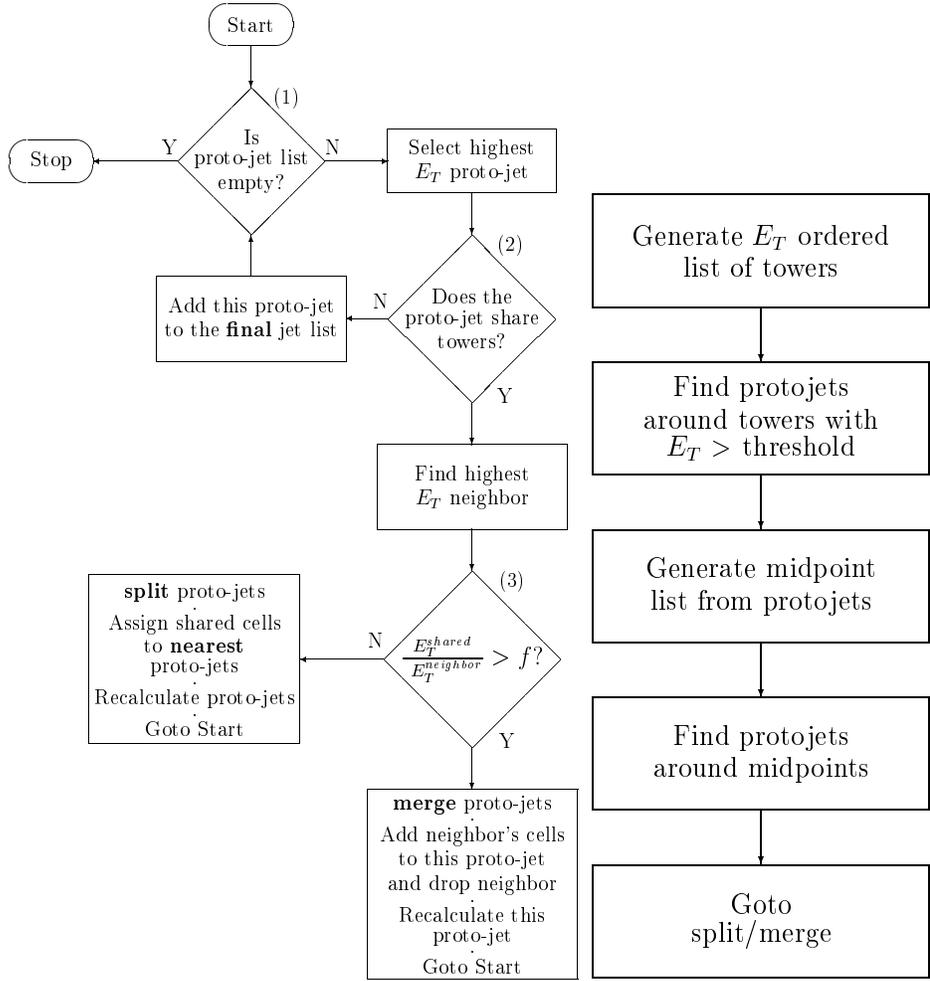
Figure 2.10: The flow chart on the left shows a standard seeded jet finding cone algorithm and the right shows how to include midpoints [26]. Together, they approximate a seedless cone algorithm. This is the algorithm used for Run II jets.

2.2.4   Monte Carlo Simulations

Experimentalists measure objects (jets) in the detector, that contain both perturbative and nonperturbative parts. Theorists calculate cross sections of partons that are perturbative. In order to compare, theory and experiment must be brought to the same level of computation. While experiments should always correct back to particle level, to remove detector effects, theory may be compared at either parton or particle levels. In order to compare, corrections to theory and experiment are determined using Monte Carlo simulations. Figure 2.11 shows a carton of how the levels of calculation are related.
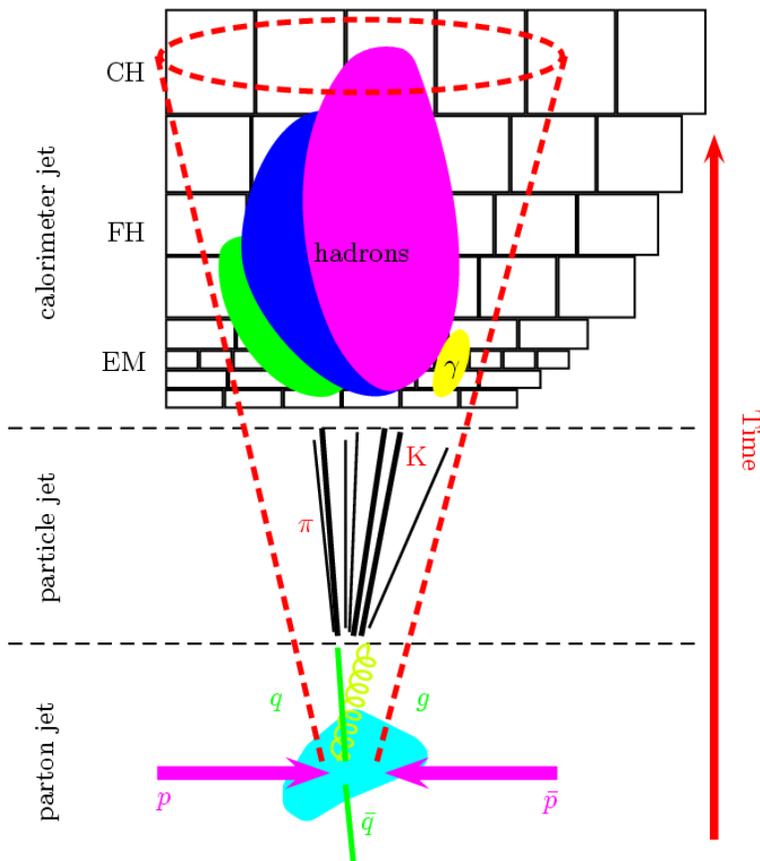
Figure 2.11: Cartoon showing three different levels of calculation; parton level by theorists, calorimeter level by experimentalists and particle level, where the theory and data is compared.

Monte Carlo simulations are computational algorithms that use random sampling to generate a data set and are used in many fields, from physical chemistry to finances.

24

In particle physics, the MC simulations used most often are called event generators and generally include the following phases [28]:

- *Hard Process*: For example, $2 \rightarrow 2$, $2 \rightarrow 3$ interactions. These can be calculated perturbatively.

- *Parton Shower*: When a parton exits the collision, it will lose some of its energy to the surrounding color field and create pairs of partons: gg, $q\bar{q}$, gq, $g\bar{q}$.

- *Hadronization*: Where the initial partons in the interaction form hadrons.

- *Multiple Scatter/Soft Underlying event*: These are nonperturbative effects, discussed more in Section 2.2.5.

Figure 2.12 shows the process the MC event generator goes through. Before the collision, the protons are shown with three lines indicating the valence quarks. As the protons collide, one parton from each participates in the hard process, described by parton distribution functions. These colliding partons interact and the energy is used to form several more pairs of partons in the parton shower. The resultant partons are then grouped into color neutral hadrons (hadronization). In addition, the MC event generators simulate the underlying event, with the energy coming from the remnants of the colliding partons, multiple parton interactions[9], and pile-up[10] included. However, the description of pile-up is usually poor in event generators and is not used in experiments[11].

In the end, the event generator has produced particles as output, which is what we ultimately want to measure. This output is then put through experiment-specific software to be reconstructed into jets and another simulation that applies detector effects[12]. The specific detector simulation for this analysis will be described in detail in Section 4.3.3.

In this analysis, the MC event generator used is PYTHIA 6.419 [29]. PYTHIA comes with many parameters that can be changed to suit a specific analysis; in this case, we use a specific set of parameters called tune QW [30], which is a particular

---

[9]Where the other partons in the colliding protons interact.

[10]Energy from other protons in the beam crossing interacting.

[11]To simulate pile-up in experiments, minbias events (events with one inelastic collision) are overlaid on top of the event generator output.

[12]At that point, we call it "reco" MC or reconstructed MC.
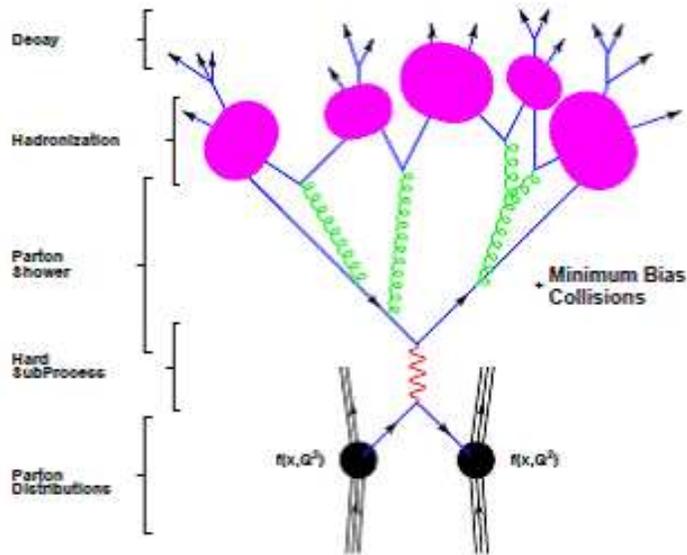
Figure 2.12: The typical MC simulation used by particle physicists' factorize the relevant processes in a collision into distinct phases that are handled separately [28].

set of PYTHIA parameters that have been fitted to Tevatron data. In addition, this analysis uses the leading order parton distribution functions calculated by the MSTW [23] group, which uses Run II jet data from the Tevatron[13] in addition to results from other experiments.

2.2.5 Nonperturbative Corrections

As mentioned previously in Section 2.2.1, nonperturbative QCD is not calculable, but when we measure jets, both the perturbative and nonperturbative parts of QCD are included and can not be separated. Data is always corrected to particle level, to remove detector dependence, while theory is calculated at the parton level. We have chosen to correct the theory to particle level as well, which involves correcting the theory with nonperturbative effects, instead of correcting the data back to the parton level. There are two types of non perturbative effects calculated in this analysis:

---

[13]Note that it is our choice to use the MSTW PDFs as input to our PYTHIA simulation. These were chosen because they are the most recent set available and include recent Tevatron results. Because PYTHIA is a leading order simulator, only LO PDFs are used.

hadronization and underlying event. In both cases, PYTHIA MC is used to determine the size of the correction.

The hadronization correction is necessary because of the hadronization procedure described in Section 2.2.3. The parton participating in the hard scatter has evolved into many tens of hadrons and the energy from these hadrons is deposited in the calorimeter and reconstructed into jets. But not all of the energy from the hard scatter goes into the hadrons that form the jet. Some of the energy in the hadronization process was used by the other partons in the proton and so this is an energy loss. To determine this effect, the energy from parton jets (jets before hadronization) in MC is compared to the particle jet energy (after hadronization) in MC. The size of the effect ranges from 5% to 10%.

When protons and antiprotons collide, usually only one parton from each is involved in the actual hard scatter. The other partons, called spectator partons, can still leave energy in the event by interacting with the outgoing partons. The spectator partons also can have their own soft interactions[14] which will cause additional radiation in the event. This is called the underlying event. This additional energy can be included in the energy of the cone defining the jet. To determine the effect of this on the dijet mass, a ratio of MC with multiple parton interactions to MC without multiple parton interactions is calculated. Events with multiple parton interactions involve both the hard scatter and the underlying event, so this ratio will determine how much energy is contained in the underlying event. The size of this correction ranges from 5% to 20%.

Figure 2.13 shows the size of the hadronization and underlying event corrections in the dijet mass analysis. The total correction is the linear fit to the product of these two corrections [37]. The corrections are applied to the NLO theory used when comparing to data, as described in Section 4.5. The total correction ranges from 5% to 10%. The uncertainties for the underlying event and hadronization correction factors are taken to be 50% of the correction factor. The total uncertainty is found by adding the underlying event and the hadronization correction uncertainty in quadrature.

---

[14]Meaning they participate in interactions where they don't collide head on, but glance off another parton.
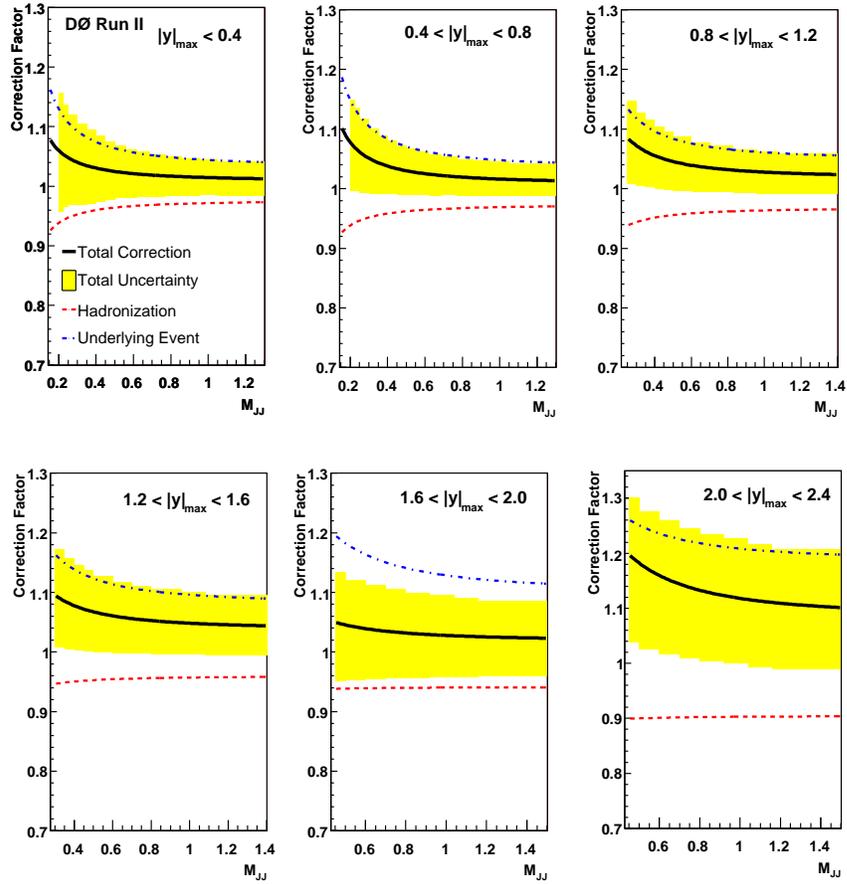
Figure 2.13: Nonperturbative corrections to theory. The total correction is the linear fit to the product of the hadronization and underlying event corrections. The total uncertainty is shown as a yellow band and is calculated by adding the individual uncertainties in quadrature.

2.2.6 New Physics

The dijet mass is sensitive to new physics, such as quark compositeness or heavy exotic particles that decay to two jets. When quarks were first postulated, it was in response to the large number of hadrons that had been discovered. Similarly, the three generations of quarks hints there might be something even more fundamental. Quark compositeness is the theory that quarks are not fundamental particles, but are made of smaller particles sometimes called preons [1] that interact via a new interaction called metacolor. Below some energy scale Λ, quarks would behave like elementary particles, while above this scale, they could be resolved into constituent particles in experiments similar to the deep inelastic scattering experiments used in the discovery of quarks. A signature of quark compositeness would be an excess at high dijet mass in the central region of the detector. Currently, the limit on quark compositeness has been set at 2.0 TeV with a confidence level[15] of 95% [25].

In addition to quark compositeness, dijet resonances could be found when measuring the dijet mass cross section. If quarks are composite particles, excited states of their constituents would be possible and show up as bumps in the dijet mass. Also, Z' and W' which are heavy states of the Z and W bosons from the Standard Model could decay to two jets, implicating extra SU(2) or U(1) gauge symmetries. A recent paper by CDF has placed limits on several exotic particles that decay to dijets [31], with the results shown Figure 2.14.

---

[15]This is a measure of how well we think we understand the limits set on a measurement.
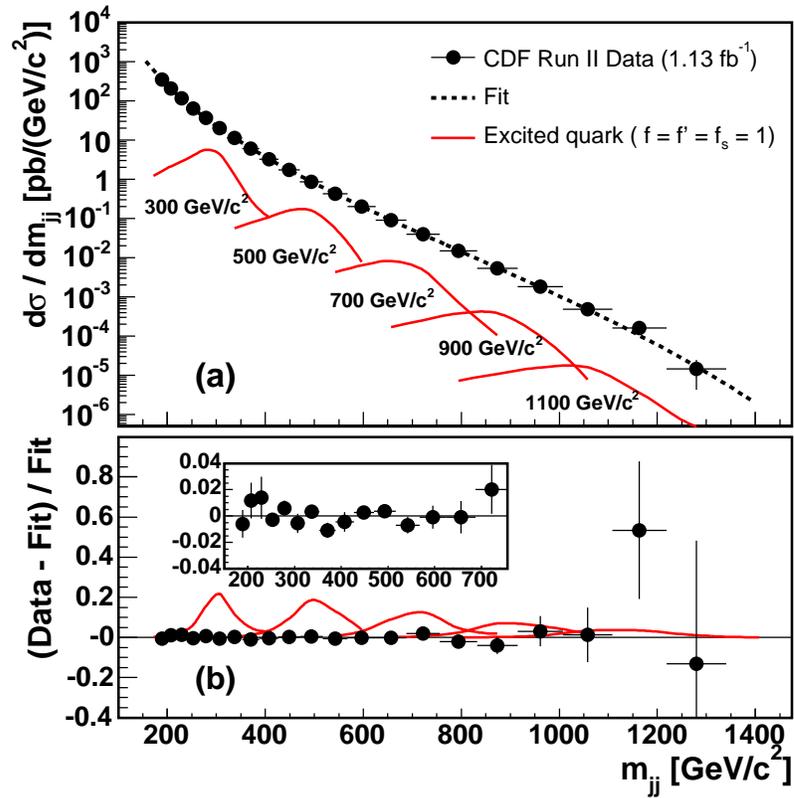
Figure 2.14: Results from CDF's recent paper on dijet resonances [31]. No evidence of new physics was found.

Chapter 3

Experimental Apparatus

This measurement was performed at the DØ detector located at Fermilab in Batavia, IL. In this section the Tevatron and the DØ detector are described.

3.1 Accelerator

The National Accelerator Laboratory was founded in 1967 using funding from the US Atomic Energy Commission with Robert Wilson acting as the first director. Located in Batavia, Illinois, the lab was renamed the Fermi National Accelerator Laboratory (Fermilab) in 1974, in honor of Enrico Fermi. The Tevatron was commissioned in 1983 and marked the start of colliding beam experiments at the lab. Up to this point the lab had conducted fixed target experiments, with the Main Ring providing beams of particles colliding with stationary targets. With the Tevatron installed, Fermilab started proton - antiproton ($p\bar{p}$) collisions in 1985. The first experiment to use counter rotating colliding beams was CDF (Collider Detector at Fermilab). In 1992 DØ (named for its position in the ring) was installed. In 2001, the Main Injector was built and both experiments underwent major upgrades. This is considered the start of Run II. Run IIb started in 2006 after a shutdown to upgrade the detectors and continues to the time of this thesis publication.

Several important discoveries took place at Fermilab. The bottom quark was discovered there in 1977 (non colliding beams). The top quark was found in 1995 and evidence of the last neutrino (the τ neutrino) predicted by the Standard Model was found in 2000. In addition to fixed target experiments and colliding beam experiments, there are several neutrino and astrophysics experiments conducted at Fermilab.

The following sections give a brief description of the accelerator complex, which is shown in Fig 3.1. Anyone interested in a more detailed explanation should consult [32].
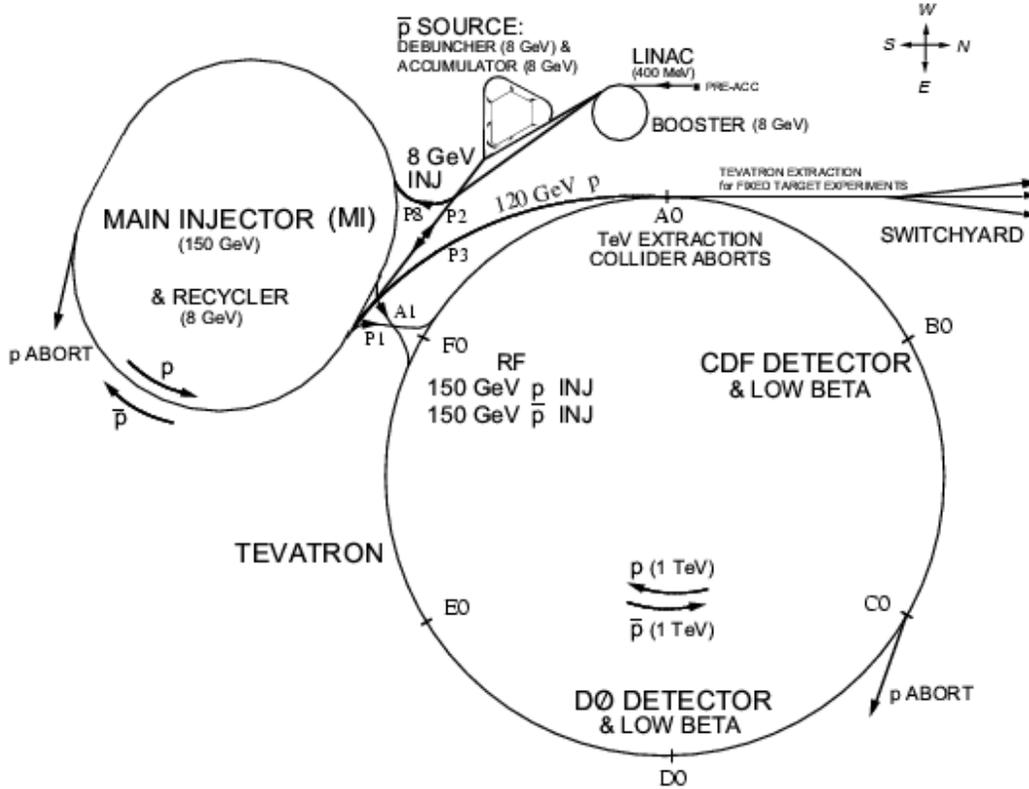
Figure 3.1: The full accelerator chain, showing the energies achieved at each stage of acceleration.

The accelerator complex includes the following components:

- Preaccelerator consisting of the hydrogen source and the Cockcroft-Walton generator. (0-750 KeV[1])

- Linac (750 KeV - 400 MeV)

- Booster (400 MeV - 8 GeV)

- Main Injector (8 GeV - 150(120) GeV)

- Tevatron (150 GeV - 980 GeV)

- Recycler and antiproton source

---

[1]The definition for KeV, MeV, etc may be found in Appendix A

3.1.1 Pre-accelerator

The pre-accelerator consists of two parts: the H$^-$ source and the Cockcroft-Walton generator. Hydrogen atoms are fed into a magnetron source producing H$^-$ ions that are then sent to the Cockcroft-Walton generator. The Cockcroft-Walton is a commercially built generator, which charges capacitors in parallel and discharges them in series, creating the necessary positive 750 KeV potential. The H$^-$ ions are then sent to the Linac.

3.1.2 Linac

The Linac (linear accelerator) takes the beam of H$^-$ ions from the pre-accelerator and accelerates them from 750 KeV to 400 MeV. It consists of two types of accelerators: a drift tube linac and a side-coupled linac. The drift tube linac (DTL) is 79 meters long and has 5 electrically resonant cylindrical copper clad tanks, which resonant at 201.24 MHz. Each tank has its own RF (radio frequency) cavity that produces 5 MW of power and has between 23 and 59 drift tubes. When the particles are in the drift tubes, they are shielded from the RF, while in the gaps, they are impacted by the RF, effectively accelerating the particles. The particles accelerate outside the tubes and decelerate inside. The phase of the particles is defined as the phase of the electric field when it is in the center of the resonant cell. The phase is carefully controlled to optimize the efficiency of transportation to the booster and the acceleration efficiency. The side coupled linac (SCL) makes up the last 7 resonant chambers and resonates at 805 MHz. While the drift tube accelerates the beam during each cycle , the SCL only accelerates it every fourth cycle to compensate for the difference in resonance. Otherwise, the principles of acceleration are the same.

3.1.3 Booster

The Booster is the first of 5 synchrotrons in the accelerator complex. A synchrotron is a particle accelerator that uses radio frequency (RF) cavities to accelerate the particles and magnetic fields to keep them moving in a circular path. The Booster is a proton synchrotron that takes the 400 MeV beam from the Linac and strips the electrons off the H$^-$ ions. The booster then accelerates the beam to 8 GeV and sends

it to the Main Injector. The Booster is 151 meters in diameter, has 96 combined function dipole/quadrapole magnets and 17 dual gap ferrite-tuned cavity resonators.

The first apparatus the beam from the Linac encounters is a RF debuncher that helps to control the momentum spread. It takes approximately 6 turns from the Linac to fill the Booster, meaning 6 cycles where particles are added. During each turn the new beam needs to be completely merged with the existing beam already circulating. This is accomplished by using a "dogleg" - 2 adjacent dipole magnets with opposite polarity. The H^- beam arrives and passes through the first magnet which sends it through a foil that strips the electrons from it. The protons are then merged into the existing beam, while any H^- still left are sent to a beam dump.

Once the Booster is full (about 3×10^{12} protons) the RF stations are tuned to the proper phase to accelerate them to the final 8 GeV. This is accomplished using focusing - bending - defocusing - bending cycles with the magnets.

3.1.4   Main Injector

The Main Injector (MI) is an approximately circular synchrotron built in the mid 1990's, to take over for the older Main Ring, which used the same tunnel as the Tevatron. With the advent of colliding beams at Fermilab, many problems arose by having the Main Ring and the Tevatron in the same tunnel. The Main Ring was decommissioned to alleviate these problems. The Main Injector can accelerate beams of protons up to 120 GeV (for the anti proton source) or 150 GeV and sends beams of protons and antiprotons into the Tevatron. The MI is made of 6 sections with 18 accelerating centers and is about half the circumference of the Tevatron and 7 times the circumference of the Booster. It runs in several different modes, including $\bar{p}$ production, shot setup, and protons for NuMI. In $\bar{p}$ production mode, it takes a beam of protons from the Booster and accelerates it to 120 GeV, in a process called stacking. In slip-stacking, the currently used method, two such beams are combined. The 120 GeV beam is then sent to the $\bar{p}$ source. Shot setup is when beams are accelerated for injection into the Tevatron for colliding beams. In this case the protons are taken from the Booster and accelerated to 150 GeV. Seven bunches are taken at a time, accelerated to 150 GeV and then, using a process called coalescing, combined into 1 "bunch". This is repeated 36 times. The antiprotons are loaded 4 bunches at a time, but NOT combined and accelerated up to 150 GeV. This is done 9 times, ending with

a total of 36 bunches of antiprotons. The end product is a 36×36 store, meaning 36 bunches of protons colliding with 36 bunches of antiprotons, spaced 396 nanoseconds apart. A store will typically last for 24 hours. Another mode of operation is for NuMI, a neutrino experiment that uses 120 GeV protons. The final mode of operation is for studies, such as diagnosing problems after a premature end to collisions. In this case any beam can be used.

3.1.5 Tevatron

The Tevatron was built in 1983 in the same tunnel used by the now decommissioned Main Ring, where its original purpose was to accelerate beams for fixed target experiments. At that time, the Tevatron accelerated beams from 400 GeV up to 800 GeV . It has a 4 mile circumference and is classified as a superconducting magnet synchrotron and is the only synchrotron at the lab to be cryogenically cooled. It takes the protons/antiprotons and accelerates them to 980 GeV, giving a center of mass energy of $\sqrt{s} = 1.96$ TeV for the colliding beams for Run II. The beams collide at two points, D0 and B0 (CDF), called luminous regions. The Tevatron is not a perfect circle; at each "0" position there is a straight section used for acceleration. It has 8 cavities used for acceleration, which occupies about 20 meters of the circle, the rest of the space is used to curve the beams.

3.1.6 Anti-proton Source and Recycler

The final piece of the accelerator complex is the antiproton source and recycler. The antiproton source is made of 3 pieces: the target station, the debuncher and the accumulator. The target station is a nickel target that the 120 GeV protons from the MI strike on their way to the rest of the machine. The protons hitting the target produce many secondary particles. Using magnets to pick out the charge and momentum, 8 GeV antiprotons are collected and sent to the debuncher, a triangular shaped synchrotron with a radius of 90 meters. Once they reach the debuncher, they are kept at 8 GeV and cooled using stochastic cooling. This consists of taking a signal from the opposite side of the debuncher, applying it to the near side and repeating. The antiprotons are then sent to the accumulator, the storage ring for the antiprotons. The accumulator is in the same tunnel as the debuncher with a radius of about 75

meters. The antiprotons are kept in the accumulator until they are extracted to the Main Injector and accelerated to 150 GeV for use in colliding beams.

3.2 DØ Detector

This analysis uses the DØ detector to measure collisions at the Tevatron. The DØ detector was built in 1992 during the Tevatron's Run I and continued running until 1996. It underwent substantial upgrades from 1996 - 2001 for Run II. During the upgrade, a new tracking system and 2 T magnet were installed and the calorimeter and muon systems were upgraded, along with the trigger and data acquisition system (DAQ). At DØ a right handed coordinate system is used with the z-axis oriented along the proton beam, and the y-axis pointing up. We use ϕ for the azimuthal angle and θ for the polar angle, and r is the distance perpendicular to the beam pipe. Two commonly used variables in analyses are η (pseudorapidity), defined as $\eta = -\ln[\tan(\theta/2)]$, and rapidity, defined as $y = 1/2 \ln[(E + p_z c)/(E - p_z c)]$. In this analysis, we use y, η, and η_{det}, which is the same as η, but measured from the geometric center of the detector[2]. The term central is used to describe rapidities of $|y| < 0.8$, ICR (intercryostat region) for $0.8 < |y| < 1.6$ and forward for $1.6 < |y|$; the same is true for η and η_{det}. In addition, the transverse momentum is commonly used in analyses and defined as $p_T = \sqrt{p_x^2 + p_y^2}$. The following subsections describe the various detector components briefly. More information may be found in [33], [34], [35] .

The full Run II DØ detector is shown in Fig. 3.2 and consists of the following components:

- Tracking detector

- Magnet

- Calorimeter

- Muon system

---

[2]Rapidity is a quantity used in physics analyses, while pseudorapidity usually refers to the detector geometry. In the massless limit, these two quantities are interchangeable. However, even with massive particles or jets, they are close enough that experimentalists will often use them interchangeably.
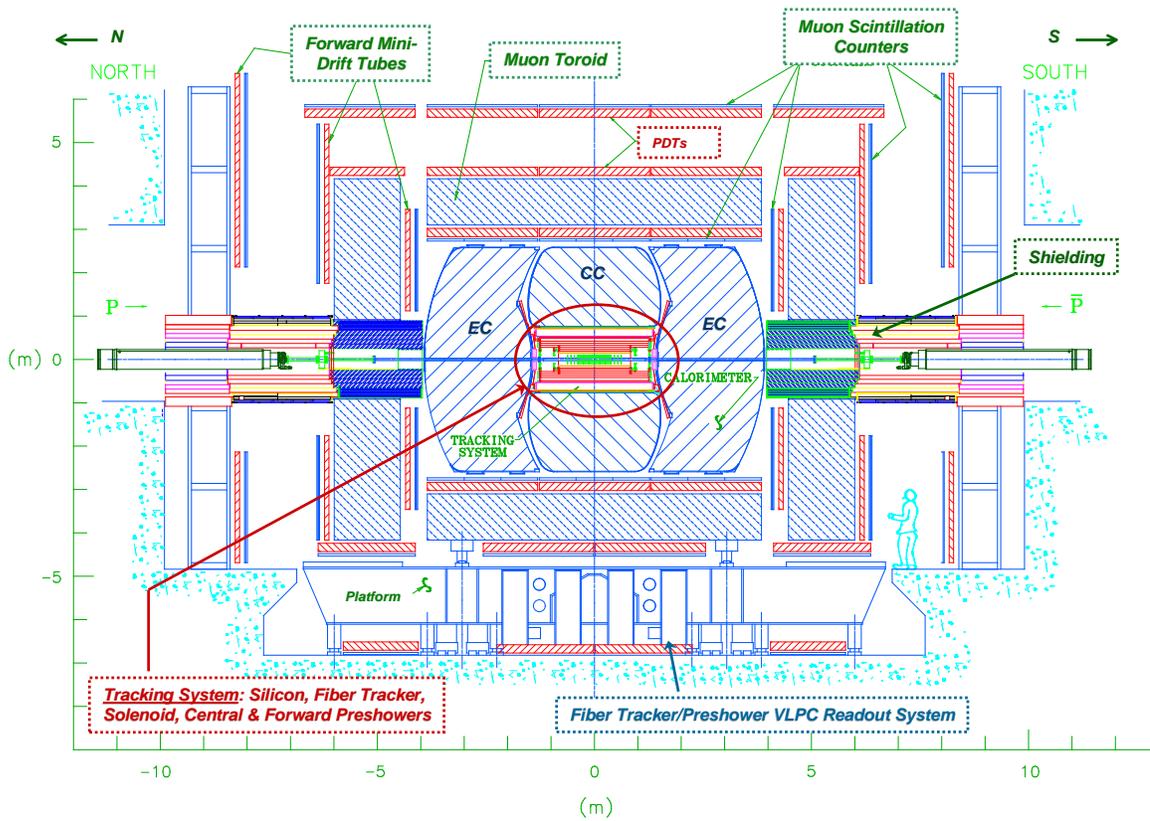
Figure 3.2: The upgraded Run II DØ detector showing the new tracking system.

3.2.1 Tracking

Having a good tracking system is essential for momentum and position measurement of charged particles and for other uses such as electron identification. The tracking system consists of the SMT (silicon microstrip tracker) and the CFT (central fiber tracker). Both components were added during the shutdown between Run I and Run II. One of the purposes of the tracker is to find the location of the primary vertex. It can measure the interaction point (IP) with a resolution of 35 micrometers and is good for measuring jet E_T, lepton p_T (transverse momentum) and missing transverse energy, $\not{E}_T$. The SMT provides information to the Level 2 and Level 3 trigger system while the CFT provides information to all 3 trigger levels (L1, L2, L3).

The Run II DØ tracking system is shown in Fig. 3.3.
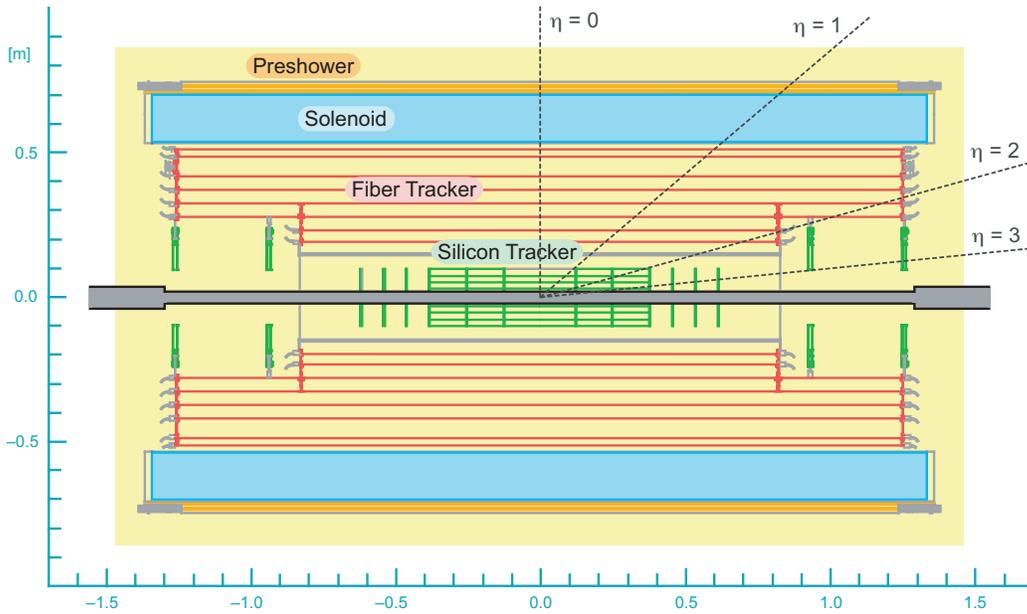
Figure 3.3: The new Run II DØ tracking system. The SMT is the closest detector to the beam pipe, followed by the CFT and the magnet.

3.2.1.1 SMT

The SMT consists of a n-type silicon semiconductor with a p-type silicon semiconductor embedded. The detector is reverse biased until it is fully depleted (no charge carriers). The electric field present then separates the charges that are produced by ionization, producing a current. The current is relayed to the electronics, where it

38

is amplified and digitized. The main purpose of the SMT is to measure the primary vertex and triggers on displaced vertices from b-quark events.

The SMT was designed to do tracking and vertexing and was constrained by what would fit in the existing space, which was not designed to include the solenoid magnet. The SMT is set up along the full interaction region (25 cm) and the detectors interspersed perpendicular and transverse to the region. The final design has 6 barrels in the central region surrounded by disks (F-disks and H-disks). Each barrel had four layers originally, and a 5th layer (layer 0) was added at the beginning of Run IIb. The layers consist of ladders made of silicon modules that do the actual detecting. Layers 1 and 2 have 12 ladders while layers 3 and 4 have 24.

Each barrel is capped by a disk, called an F-disk. These consist of 12 double sided wedge detectors. There are three F-disks on each end to cover the high $|z|$ area. In the most forward positions of each end are two H-disks. The H-disks are made of 24 full wedges, composed of two back to back single sided wedges.

In total, there are 432 ladders, 144 F-disk wedges, and 96 full H-disk wedges. The F-disks have a pitch of 50/62.5 μm (p/n), are 7.93 cm in length with an inner radius of 2.57 cm and outer radius of 9.96 cm. The H-disks have a pitch of 40 μm for the inner sensors, with a length of 7.63 cm, an inner radius of 9.5 cm and an outer radius of 26 cm. The outer sensors for the H-disks have a pitch of 80 μm and a length of 6.33. The barrels have a pitch of 50/153.5 μm (p/n) for layers 1 and 3 and a pitch of 50/62.5 μm for layers 2 and 4. The length of the central barrels is 6.0 cm. The outer barrels have a pitch of 50 μm for layers 1 and 3, a pitch of 50/62.5 μm (p/n) for layers 2 and 4, and their length is also 6.0 cm. There is a total of 912 readout modules and 792,576 channels. The readout electronics consist of 128 channel SVXIIe readout chips.

3.2.1.2   CFT

The next detector away from the beamline is the central fiber tracker (CFT). The CFT is made of scintillating fibers that give off light when a charged particle passes through them. The photons are collected by visible light photon counters (VLPCs), that send out a digitized signal to the electronics. The CFT, like the SMT, gives accurate readings on the momentum of the particle and reconstructs the primary (interaction point) and secondary (vertex from a particle that travels before decaying) vertices.

The momentum resolution at 1 GeV is about 1.5%, and about 11% at high p_T. The impact parameter resolution is 50 microns at 1 GeV and 10 microns at 50 GeV.

The CFT at DØ is made of 8 concentric cylinders with scintillating fibers mounted on them. The two inner cylinders are shorter to accommodate the SMT which is inside the CFT. The length of the inner cylinders is 1.66 m and the outer cylinders have a length of 2.52 m. Each cylinder has two doublets, one oriented along the beam (z) direction called the axial layer and one oriented along a stereo angle of $\phi = -3$ (v) or $\phi = 3$ (u) called stereo layers. From inner to outer the layers go $zu - zvzu - zvzu - zvzu - zvzu - zv$ where v indicates the negative stereo angle and u indicates the positive stereo angle. The fibers are made of clear scintillating fibers that are doped with a fluorescent dye and are 835 μm thick. These attach to clear fiber wave guides that feed out to the front end electronics (the VLPCs). The fibers are only read out on one end; the other end is covered with sputtered aluminum having a reflectivity of $\approx 90\%$. The axial layers are used at Level 1 for a hardware trigger, while levels 2 and 3 use the full CFT readout information to make trigger decisions.

3.2.1.3 Magnet

Surrounding the tracking system is a 2 Tesla superconducting solenoid magnet. The magnet was designed to maximize the momentum resolution and tracking pattern recognition while taking into account the available space (diameter of 1.42 m, length of 2.73 m), safe operation in both polarities, be able to quench[3] safely, and to have a uniform magnet field over the whole volume. The magnet was designed to have a thickness of 1 radiation length. The solenoid is wound with two layers in order to achieve the linear current density needed for the magnetic field. In order to keep the field as uniform as possible, the current density in the windings needs to be larger at the ends. To achieve this, the conductors are narrower at the ends of the coil. The conductors are made of superconducting Rutherford-type cables of Cu:NbTi strands with pure aluminum for stability. The only joints in the coil are at the four points (2 in each layer) where the conductors changed from wide to narrow.

---

[3]A quench occurs when the temperature of the magnets increases almost to the point where they are no longer superconducting. In order to avoid damaging the magnets, the beam is directed into a chunk of iron and stopped safely.

3.2.2 Preshower

The first detector beyond the magnet is the preshower, composed of a central preshower (CPS) and a forward preshower (FPS). The purpose of the preshower is to aid photon and electron identification and can be used as either a calorimeter or a tracking detector. It is useful in matching tracks to EM shower information in the calorimeter. Both preshower detectors are made of triangular scintillating strips. These strips are overlayed so that a particle may traverse more than one strip and will provide full coverage. In the middle of each strip is a WLS (wavelength shifting) fiber which is read out of only one end; the other end is diamond polished and silvered. The readout for the preshower connects to the CFT readout electronics.

The CPS is made of three concentric cylinders located between the solenoid and the central calorimeter. The 3 layers are in an axial-u-v arrangement where $u = 23.774°$ and $v = 24.016°$. There are a total of 2560 readout channels for each layer.

The FPS is located on the face of the end cap calorimeters, between the luminosity monitors and the ICD (intercryostat detector). Each is composed of two layers; a MIP (minimum ionizing particle) layer and a shower layer. Between the MIP layer and the shower layer is a lead-stainless steel absorber of two interaction lengths. The MIP layers are the nearest (in z) to the interaction point. Particles can be identified based on how they shower and leave tracks. For example, electrons will shower in the absorber, leading to an energy deposit in the shower layer. Each detector has four measuring planes arranged in $u - v$ pattern. These measuring planes are made of eight 22.5° wedges of active material and eight wedges of inactive material. The WLS is run through to the edge of the detector and then fed into the VLPC cassettes located in the CFT. The absorber is made of 48 wedge shaped pieces to facilitate easier handling.

3.2.3 Calorimeter

The calorimeter is the main detector used for this analysis. Calorimeters are used to measure the energy deposited by the particles passing through them. They achieve this by either measuring the energy at discrete intervals (sampling) or by measuring the full energy deposited. The calorimeter is also designed to help determine which direction a particle came from. At DØ a sampling calorimeter is used and hasn't changed from its original designed described in [35]. The calorimeter was designed to

b) CPS GEOMETRY

6.884 mm

6.100 mm

2 Layers of Mylar (0.025 x 2 = 0.050 mm)

c) FPS GEOMETRY

5.402 mm

7.820 mm

2 Layers of Mylar (0.025 x 2 = 0.050 mm)

60.00°

60.00°
6.858 mm Ref

WLS Fiber
Insert
Ø 0.991 mm

5.994 mm Ref
5.309 mm

1.981
mm

2.972 mm

5.944 mm

R = 0.635 mm

6.858 mm Ref

a) CPS - FPS SCINTILLATOR GEOMETRY
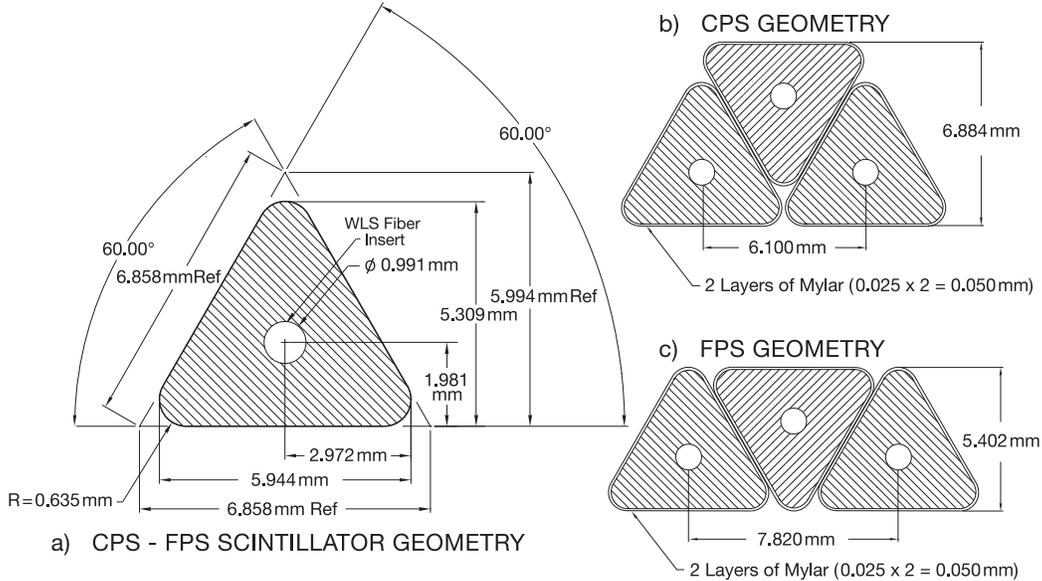
Figure 3.4: The upgraded Run II DØ preshower detector geometry [34]. The scin-tillator strips for both the forward and central preshower are shown on the left (a), while the CPS geometry is shown in (b) and the FPS geometry is shown in (c).

measure the energies of electrons, photons and jets[4], as well as distinguish between the two, since there was no tracking system to help determine particles types in the Run I detector. Calorimeters can be segmented transversely to help determine particle direction and segmented longitudinally, which provides information on the type of shower. This information can be used in particle identification. For more information on how calorimeters work in general, see [36].

The DØ calorimeter is made of uranium/liquid argon and is composed of three separate detectors (one central and two endcaps (EC)), housed in individual cryostats (See Fig. 3.5). In each cryostat, uranium plates are placed in liquid argon filled chambers. Liquid argon was chosen based on its radiation hardness, the ability to segment the detector, and the ease of calibration. The drawbacks to using liquid argon include the fact that each calorimeter must be housed in cryostats, leading to areas where there are gaps in instrumentation. In each detector there is an electromagnetic (EM) section, fine hadronic, and coarse hadronic sections, located radially away from the interaction point. The EM section has the thinnest uranium plates at 3 mm in the central and 4 mm in the EC. The plates are made of nearly pure depleted uranium.

---

[4]See Section 2.2.3 for more information on jet algorithms.

The fine hadronic is made of a 6 mm uranium-niobium alloy plates. The coarse hadronic has plates of 46.5 mm copper in the central region and stainless steel plates in the forward sections. At $\eta = 0$ in the central region, the calorimeter (all layers) is 7.2 nuclear interaction lengths and at the smallest angle in the forward direction, the calorimeter is 10.3 nuclear interaction lengths. The calorimeter is segmented in $\eta - \phi$ space by towers of 0.1 x 0.1 at $\eta < 3.2$, beyond that, the towers increase to 0.2 x 0.2.

The central calorimeter consists of 32 EM segments (CCEM) that have 4 readout sections at 2.0, 2.0, 6.8 and 9.8 radiation lengths[5] for a total of 20.5 radiation lengths (0.76 interaction lengths). Beyond that the calorimeter consists of 16 fine hadronic (CCFH) segments with 3 readout sections (1.3, 1.0, 0.9 interaction lengths) and 16 coarse hadronic (CCCH) segments with one readout section at 3.2 interaction lengths. The total number of signals for the central calorimeter are 10,400 (CCEM), 3500 (CCFH) and 700 (CCCH).

The endcap calorimeter has one EM (ECEM) segment with 4 readout sections at 0.3, 2.6, 7.9 and 9.3 radiation lengths. The hadronic is broken into an inner hadronic (ECIH) portion with a fine hadronic section that has four readout sections at 1.1 interaction lengths each and a coarse hadronic segment with one readout section at 4.1 interaction lengths. There are also middle and outer hadronic sections in the endcap calorimeter. The middle hadronic has four fine hadronic segments with readout sections at 0.9 interaction lengths each and a single coarse hadronic section with a readout at 4.4 interaction lengths. The total signals received from the endcap calorimeter is 7488 (ECEM), 5216 (ECIH), 1856 (ECMH) and 960 (ECOH).

Since the calorimeter is divided into three sections, each in their own cryostat, there needs to be a detector to fill the gaps. The Inter cryostat detector (ICD) was placed between the cryostats of the central and endcap calorimeters, in the massless gap regions. These regions have identical readout boards to the rest of the calorimeter, but no uranium plates or liquid argon. This is traditionally a difficult region of the detector to use in measurements because the instrumentation is not as comprehensive as and has a different resolution from the rest of the calorimeter.

---

[5]A radiation length is the length it takes an electron to lose all but $1/e$ of it's energy.

DØ LIQUID ARGON CALORIMETER

END CALORIMETER

Outer Hadronic
(Coarse)

Middle Hadronic
(Fine & Coarse)

Inner Hadronic
(Fine & Coarse)

Electromagnetic

**CENTRAL
CALORIMETER**

Electromagnetic

Fine Hadronic

Coarse Hadronic

1m
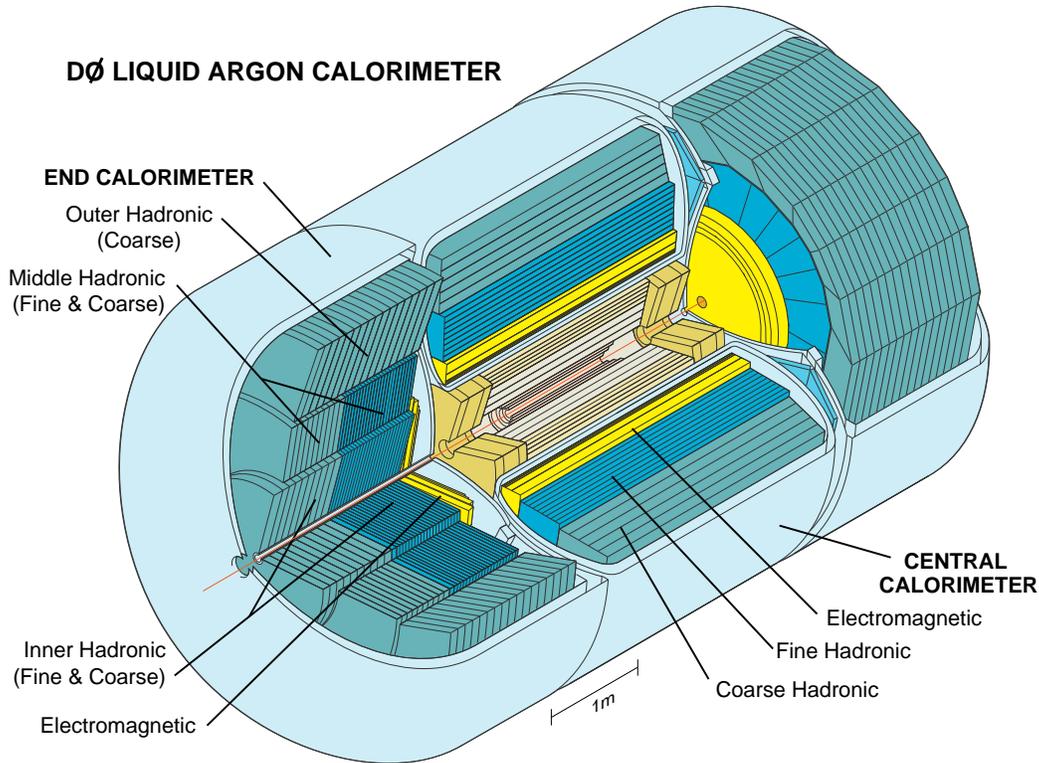
Figure 3.5: The upgraded Run II DØ calorimeter.

3.2.4 Muon system

Just beyond the calorimeter is the muon detector, which is used to detect muons. The muon detector consists of a central muon detector and toroid (both present during Run I) and a new forward muon detector. The central detector covers $\approx$ $|\eta| < 1.0$ and the forward detector extends this out to $|\eta| = 2.0$. Cosmic caps and bottom scintillation counters were added to help reduce background from cosmic rays. $A\phi$ scintillation counters were also added to the central muon detector. Using the scintillation counters helps with the triggering of events while the drift chambers are useful for precision coordinate measurements. These combine to make a robust system.

The toroid is composed of two "C" sections around the central calorimeter and two end magnets so that the total shape is a box around the calorimeter. There are square holes in the end toroids to accommodate the beam pipe. The center toroid is made of twenty coils with ten turns each while the end toroids are made of eight coils with eight turns each. The polarity is regularly reversed and runs at a magnet

field that is about 6% smaller than the Run I value of 2.0 T because the current was reduced to save money and minimize quenching.

The central muon system consists of the central toroids, the PDTs (partial drift tubes), the cosmic cap, bottom scintillation counters, and the $A\phi$ scintillation counters. There are three layers of PDTs; the A-layer is located inside the toroid while the B and C layer are located just beyond. The B and C layers have three decks of drift cells while the A layer has four decks. There is an anode wire at the center of each cell while vernier cathode pads are above and below the wires to get hit position information. The cosmic cap and the bottom scintillation counters surround the PDT layers. These detectors are linked to the Tevatron clock, providing timing information that helps to reduce the background from cosmic rays. The $A\phi$ detectors are located between the toroid and the central calorimeter and are used to identify muon and to reject back scatter from the forward sections. The signals from the $A\phi$ detector are matched with hit information from the CFT for a Level 1 trigger.

The forward muon system consists of the end toroidal magnets, MDTs (mini drift tubes), scintillation counters, and shielding around the beam pipe. The MDTs are used in the forward detector because they have a shorter drift time for electrons, radiation hardness, and coordinate resolution. Similar to the central detector, the wire chambers are arranged in three layers - A (located inside the toroid), B, and C (both located outside the toroid), with the A layer closest to the interaction region. All three layers are divided into 8 octants. The B and C layers consist of three planes of tubes and the A layer consists of four planes of tubes. The planes of tubes are oriented along the toroidal magnetic field lines. The scintillation counters are also made of three layers. The layers are divided into octants that have 96 counters each. The ϕ segmentation matches the CFT trigger sectors at 4.5°.

There are three sources of background that can affect the muon system. One source of background affects the A layers and comes from scattered proton and antiproton fragments that interact with the calorimeter or the beampipe. The two other sources of background affect the B and C layers and come from scattered proton and antiproton fragments that interact with the Tevatron low-beta quadrapole magnets and from halo interactions in the tunnel. Shielding was installed in Run I to help reduce the beam halo background. In Run II more shielding was installed to help reduce the other types of backgrounds.

3.2.5 Luminosity Monitor

The luminosity monitor provides a measurement of the luminosity to the Tevatron for the DØ interaction region and also provides halo rates[6]. The luminosity is measured using inelastic $p\bar{p}$ collisions. The detector consists of two arrays with 24 plastic scintillation counters with PMT (photo multiplier tubes) readout. They are located just inside $|z| = 140$ cm, in front of the end calorimeter and located radially between the forward preshower and the end cap calorimeter.

The luminosity is determined using Eq. 3.1

$$\mathcal{L} = \frac{f * \bar{N}_{LM}}{\sigma_{LM}} \tag{3.1}$$

where f is the beam crossing frequency, $\bar{N}_{LM}$ is the number of inelastic interactions per beam crossing, and σ_{LM} is the effective cross section of the luminosity monitor, which accounts for the acceptance and efficiency of the monitor. It should be noted that $\bar{N}_{LM}$ is normally greater than 1, due to multiple collisions in a single beam crossing. To account for these extra collisions, the fraction is taken between $\bar{N}_{LM}$ and the number of beam crossings with no $p\bar{p}$ collisions and Poisson statistics are used to find the average.

As of June 2009, DØ has recorded 6 fb^{-1}, 2 fb^{-1} of which was recorded in 2008 alone.

3.2.6 Trigger system

Due to high luminosity and high interaction rate, we cannot record every event that occurs in the detector. Therefore, a three level trigger system is used to help determine which events are recorded to tape. The trigger system is managed by a dedicated computer called TCC1 (trigger control computer) at Level1, TCC2 at Level 2, and TCC3 at Level3. These computers then talk to COOR, which is the computer that coordinates between the various subsystems of the detector.

The Level 1 (L1) stage is a hardware trigger. For example, the L1 calorimeter triggers look for certain energy deposits above instrument limits. The L1 CTT and L1 Muon look for tracks to tie to muons above a certain p_T. The L1 trigger is

---

[6]When there are colliding beams, occasionally particles interact with the beam pipe, creating a "halo" around the beam pipe. The size (rates) of these haloes are monitored. If they are too large, it can indicate that the beam is not in proper position.

controlled by the Trigger Framework (TRGFR) computer. The TRGFR handles the communications between COOR, the trigger control computer (TCC) and the front-end electronics. It can take up to 128 specific triggers. These are made up of "OR" terms, which determine if the crossing had a valid event. These are called the physics triggers. There are also beam triggers, which are based on beam conditions, that are associated with the 128 physics triggers. Both the physics and beam triggers must be satisfied for an event to be passed to Level 2 (L2). The L1 trigger system reduces the rate to 2 kHz.

The L2 trigger stage uses a combination of hardware triggers and embedded microprocessors. There are detector specific subsystem crates[7] and a global crate used to issue a final decision. At this stage, physics objects can be correlated between several subsystems to make a more informed decision. If an event passes one of the 128 L1 trigger terms and passes the scripts imposed by L2, the event is passed onto Level 3 (L3). At this point the input rate is reduced to 1 kHz.

The L3 trigger system makes a decision based on physics objects and relations between these objects (for example, the angle between two objects). This is accomplished by using special scripts called filter tools. These tools unpack the data, locate hits, check calibrations, find clusters of energy and object reconstruction. An event must pass these filters in order to be recorded to tape where the input rate is reduced from 1 kHz to about 100 Hz.

In addition, prescale sets are used to determine which type of physics triggers are recorded. These are based on luminosity. For example, groups looking for new physics will need higher luminosities where exotic particles are more likely to be produced, while groups studying b-physics can use lower luminosity beams. The prescale is a number that tells the system to fire on this type of event one out of so many times. For example, the system would fire on a jet with $p_T > 45$ GeV only once in every 100 times it is seen. In some cases, a trigger is unprescaled, meaning every time an event happens it gets recorded.

---

[7]A crate is the collection of front end boards that handle information from a specific source. For example, the L2CAL crate handles all the information coming in from the calorimeter.

3.2.7 Data Acquisition System

The data acquisition system (DAQ) takes the data from the readout crates to the L3 processing nodes and communicates between the L3 farm and COOR. COOR takes care of run control operations by keeping a current configuration of the detector at all times. It controls run transitions and configurations. COOR sends information to the various subsystems (called downloaders) and makes sure that the users sending information (clients) to it do not issue conflicting commands.

Chapter 4

Analysis

This section describes the details of the dijet mass analysis. The data and event selection criteria are discussed, along with the corrections and the final results.

The cross section was calculated using Eq. 4.1,

$$\frac{d\sigma}{dM_{JJ}d|y|_{max}} = \frac{N_{evt}}{L \cdot \sum_i \varepsilon_{vtxeff} \cdot \Delta(M_{JJ})\Delta(|y|_{max})}C \tag{4.1}$$

where N_{evt} is the number of events after mass rescaling, L is the luminosity, $\sum_i \varepsilon_{vtxeff}$ is the vertex efficiency per event i, $\Delta(M_{JJ})$ is the mass bin width, $\Delta(|y|_{max})$ is the rapidity bin width and C is the correction factor determined from DØJetSim. Each one of these terms will be discussed throughout the following sections.

4.1   Data Set and Event Selection

This analysis uses ≈ 0.7 fb^{-1} of data from the Run IIa data set (runs 191000 - 213064) taken from 2004 to 2006. After the data was recorded, it was processed offline to correct for known issues, such as bad cells in the calorimeter. A smaller data set that does not include any cable swap data[1] or early Run II data was chosen because this data set is very well understood. Additionally, it was used in other analyses([37], [38]), allowing for easy combination and comparison of results. The data format for this sample is roottuples, which are created by the program ROOT [39]. This is a common histogramming program used by high energy physicists to quickly and easily process data into histograms.

---

[1]During early Run IIa, it was discovered that several of the calorimeter sections had the wrong cabling. This caused the data quality to be poor as we could not be sure we were mapping to the correct portions of the calorimeter.

The events are chosen to have at least two jets and the two jets with the highest transverse momentum (p_T) are used to calculate the dijet mass. These are often called the leading jets. In addition, there are quality cuts imposed to make sure the events are not bad due to noise or other effects. The data is considered "bad" by looking at the behavior of the detector; for example, the calorimeter may have a tower that fires more often than it should. This can affect the measurement of the jet energy and so the event or run[2] is removed. In some cases, a run may be bad for certain types of analyses while working perfectly fine for others. For example, a b-physics analysis might need the muon detector to be working, while a jet analysis may only need the calorimeter, so the muon system is not necessary. Bad events or runs are determined by several different groups, such as the data quality group or the Jet/MET group. For instance the data quality group looks at the recorded data and tries to remove all bad runs and bad LBNs (Luminosity Block Numbers) from it. The data is required to pass these selection criteria:

• Official Run II data quality cuts (v2006-10-10), including cuts on bad LBNs (Luminosity Block Numbers) and bad runs.

• The calfail flag is set to false. This flag removes events that are marked bad because of a problem in the calorimeter.

• $\not{E}_T/p_T^{leading} < 0.7$. This cuts out signals that come from cosmic rays hitting the detector.

• $n_{vtx} \geq 1$, $|z_{vtx}| < 50$ cm, and $n_{trks} \geq 3$, which ensure the vertex is in a fiducial region of the detector.

In addition to event quality cuts, the two leading jets in the analysis must pass "good" jet criteria. A complete description of the jet ID quality cuts can be found in [46], which is summarized here. The first set of cuts is on the fraction of energy deposited in the EM calorimeter (emf). The first cut of this set is designed to remove any overlap between jets and EM objects (such as electron or photons). It is the most inefficient cut, removing about 1% of the jets. The remaining cuts remove events with anomalously low emf due to effects like a jet being formed out of noise in the hadronic calorimeter.

• emf < 0.95 **and**

• emf> 0.05 **or**

---

[2]A store is broken up into smaller parts called runs. These are used to change prescales, to fix problems and to keep the data streams from getting too large.

- $0.13 > |(|\eta_{det}| - 1.25|) + \max(0.4 \times (\sigma_\eta{}^3 - 0.1))$ (in the no EM gap) **or**
- emf > 0.03 and $|(|\eta_{det}| - 1.25)| < 1.5$ (in the no EM gap) **or**
- emf > 0.04 and $2.5 < |\eta|$ (forward region)

The next set of cuts addresses the removal of noisy jets from the coarse hadronic calorimeter (chf). The resolution is poor here and not much energy is expected, so upper limits on how much energy is found in the coarse hadronic calorimeter are imposed. The values listed below represent the amount of energy allowed to be in the chf.

- chf < 0.4 **or**
- chf < 0.6 and $0.85 < |\eta_{det}| < 1.25$ (in ECMH) and $n_{90}{}^4 < 20$ **or**
- chf < 0.46 and $|\eta_{det}| < 0.8$ (central) **or**
- chf < 0.33 and $1.5 < |\eta_{det}| < 2.5$ (endcap excluding forward region)

The final cut is on the L1 confirmation term[5]. This cut was designed to remove jets formed out of noise from the precision readout.

- $E_{L1} > 55$ GeV **or**
- $f_{L1} > 0.5$ **or**
- $f_{L1} > 0.35$ and $p_T^{uncorr} < 15$ GeV and $1.4 < |\eta_{det}|$ **or**
- $f_{L1} > 0.1$ and $p_T^{uncorr} < 15$ GeV and $3.0 < |\eta_{det}|$ **or**
- $f_{L1} > 0.2$ and $p_T^{uncorr} \geq 15$ GeV and $3.0 < |\eta_{det}|$ where $f_{L1} = \frac{p_{T,L1}}{p_T^{uncorr}(1 - chf - ccmg - ecmg)}$

and $ccmg, ecmg$ represents the fraction of energy deposited by the jet in the central calorimeter massless gap(ccmg)[6] and the EC calorimeter massless gap(ecmg), respectively.

After making good event and good jet selections, the only other cut applied to the data is that the p_T of the second leading jet must be greater than 40 GeV. This cut ensures that the analysis is done in a p_T range where the data corrections are well understood.

---

[3]This refers to the width of the jet.

[4]The number of towers with less than 90% of the jet's energy

[5]The L1 confirmation term is used to distinguish jets formed from noise and actual jets at the L1 trigger reconstruction phase.

[6]The region outside the cryostats are referred to as massless gaps, because there is no material there for the particles to interact with.

4.2 Analysis Bins

This section describes how the analysis bins were chosen for presentation. The cross section is shown as a double differential with six rapidity bins divided into varying mass bins. Each bin is required to only have data from one trigger which helps ensure there are no discrepancies due to the different luminosities between triggers.

4.2.1 Rapidity bins

After making the event selections, the data sample is then divided into six rapidity[7] regions based on $|y|_{max}$, where $|y|_{max}$ is the rapidity of the jet with the highest rapidity value. The rapidity regions are:

- Two central regions: $|y| < 0.4$ (CC1) and $0.4 < |y|_{max} < 0.8$ (CC2)
- Two ICR regions: $0.8 < |y|_{max} < 1.2$ (IC1) and $1.2 < |y|_{max} < 1.6$ (IC2)
- Two forward regions: $1.6 < |y|_{max} < 2.0$ (EC1) and $2.0 < |y|_{max} < 2.4$ (EC2)

In all cases, the jet with the smaller $|y|$ is allowed to be anywhere less than $|y|_{max}$. For example, the dijet mass may be composed of one jet with $|y| = 0.3$ and a second jet with $|y| = 0.7$. In this case, the event would be placed in the second central region. Note that the highest p_T jet is not necessarily the most forward. The choice of these particular regions is a historical one. A previous analysis measuring the inclusive jet p_T used these rapidity bins and we wanted to be able to compare results. Also, these rapidity bins are wide enough that the resolutions do not change significantly in a single bin and narrow enough to give a precise measurement. The central region (up to $|y| < 0.8$) is the most straightforward to measure, as it only encompasses one well understood detector. The forward regions $1.6 < |y| < 2.4$ have a worse resolution, which makes the measurements more difficult than the central, but the finely segmented detector in that region offsets the poor resolution. The middle region (ICR, $0.8 < |y| < 1.6$) is the most difficult to measure because the resolutions are extremely poor and involve several detectors with different instrumentation. Figure 4.1 shows the calorimeter divided into pseudorapidity segments. Rapidity and η_{det} use the same pictorial representations.

---

[7]The rapidity is defined in Section 3.2.

η=0.0   0.2   0.4   0.6   0.8

1.0

1.2

1.4

1.6

1.8

2.0

2.2

2.4
2.6

3.2

4.1
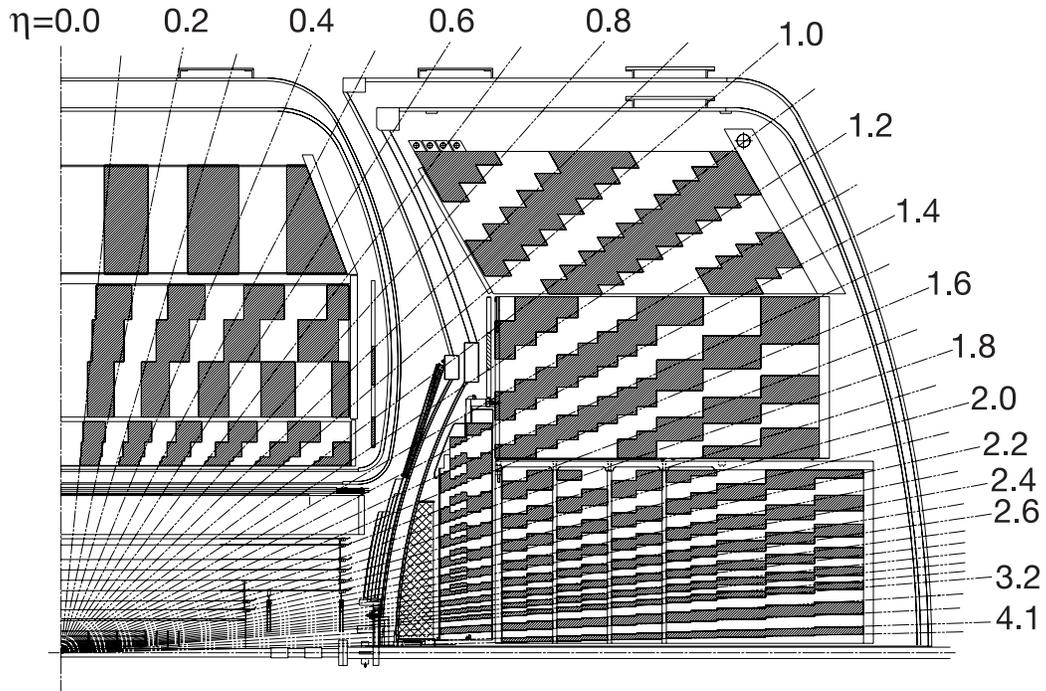
Figure 4.1: DØ calorimeter with the η regions shown. The central region is $|\eta| < 0.8$, intermediate is $0.8 < |\eta| < 1.6$, and the forward region is $1.6 < |\eta| < 2.4$. This analysis further splits the rapidity regions in to bins of 0.4.

4.2.2 Triggers

After defining the rapidity regions, the next step is to determine which triggers are used in each region. This analysis uses data from two specific sets of triggers designed to find jets. The first set records the events based on a minimum p_T value for a single jet at Level 3. These triggers are referred to as single jet triggers and are called JT_25TT_NG, JT_45TT, JT_65TT, JT_95TT and JT_125TT. The second set uses a minimum dijet mass to determine what is recorded at Level 3 and are called dijet mass triggers (JT_L3M250, JT_L3M430) . We want to make sure these triggers are efficient, meaning that the jet has a high probability of firing the trigger, defined in this analysis as events with $> 99\%$ trigger efficiency. By having an efficient trigger, we are cutting out events that do not contain the jets we are interested in, while keeping almost all the jets we are interested in. A previous study [40] had determined the absolute trigger efficiency per event for the single jet triggers. This is done by comparing the single jet triggers to an independent sample; in this case, a set of muon triggers with no calorimeter requirements. At each level (L1, L2, L3), the efficiency is calculated according to Eq. 4.2:

$$\epsilon_{L1} = \frac{\# \text{ jets with event passing reconstructed L1}}{\# \text{ jets}} \tag{4.2}$$

$$\epsilon_{L2} = \frac{\# \text{ jets with event passing reconstructed L2}}{\# \text{ jets with event passing reconstructed L1}}$$

$$\epsilon_{L3} = \frac{\# \text{ jets with event passing reconstructed L3}}{\# \text{ jets with event passing reconstructed L1 and L3 tool running}}$$

By comparing to the unrelated set of triggers, we can find a p_T where the jet trigger is fully efficient. Since the jet triggers are highly correlated, once we determine where the lowest trigger is efficient, we can compare the next highest trigger to get a relative efficiency. This allows us to determine which p_T (or mass) region each trigger can be used in.

These individual efficiencies are then multiplied together to form a total efficiency for the trigger. Ideally, the probability of any jet firing the trigger would be independent of other jets and independent of the three different levels of trigger. The individual probability for each level of trigger in this case is given by Eq 4.3. For an event with multiple jets, the probability would be given by Eq 4.4, where P_i is given

by Eq. 4.3. Note that these are for trigger matched objects which means that the reconstructed jet is matched to a trigger object.

$$P_i(p_T, \eta) = P_{i,1}(p_T, \eta) P_{i,2}(p_T, \eta) P_{i,3}(p_T, \eta) \tag{4.3}$$

$$P_{uncorrelated}^{event} = 1 - \prod_i (1 - P_i(p_{T,i}, \eta_i)) \tag{4.4}$$

The probabilities in Equations 4.3 and 4.4 are ideal cases[8] and would work at L2 and L3 of the trigger system, since the jet object is formed at that point. However, at the first level of trigger, there is only a requirement for some number of towers above a certain p_T threshold. In this case, a single jet object may not have the required number of towers, making a single jet trigger efficiency hard to determine. However, if there is more than 1 jet in the event, the towers from the 2 jets would combine to fire the trigger. In this case, the trigger efficiency is an event efficiency, not a single jet efficiency. In practice the event efficiency is higher than the single jet efficiency. Figure 4.2 shows an example of this concept.
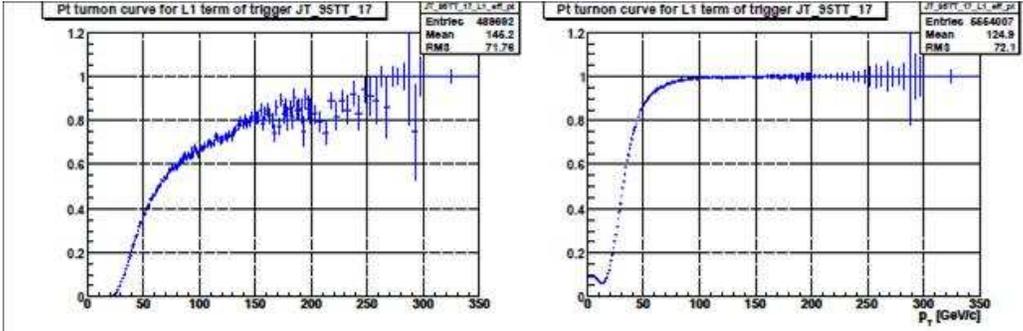
Figure 4.2: Figure (a) shows the efficiency if we only consider a single jet in the event, while(b) shows how much more efficient the trigger is if all jets in the event are used [40].

The preceding discussion outlines the method used for the inclusive jet analysis [38], where the p_T was the variable studied. In this analysis, since the data sample is the same, the same triggers can be used and the absolute trigger efficiency does not need to be recalculated. The single jet triggers are used in the two Central and two ICR regions and the dijet mass triggers are used in the two EC regions. The reason

---

[8]Meaning that the jets are distinguishable from one another at all trigger levels.

for this choice is that in the EC region, the inclusive jet triggers aren't really efficient in mass, as seen in Fig. 4.3. All triggers are taken from trigger list versions v12 - v14[9] and the properties for each trigger can be found in Appendix C.
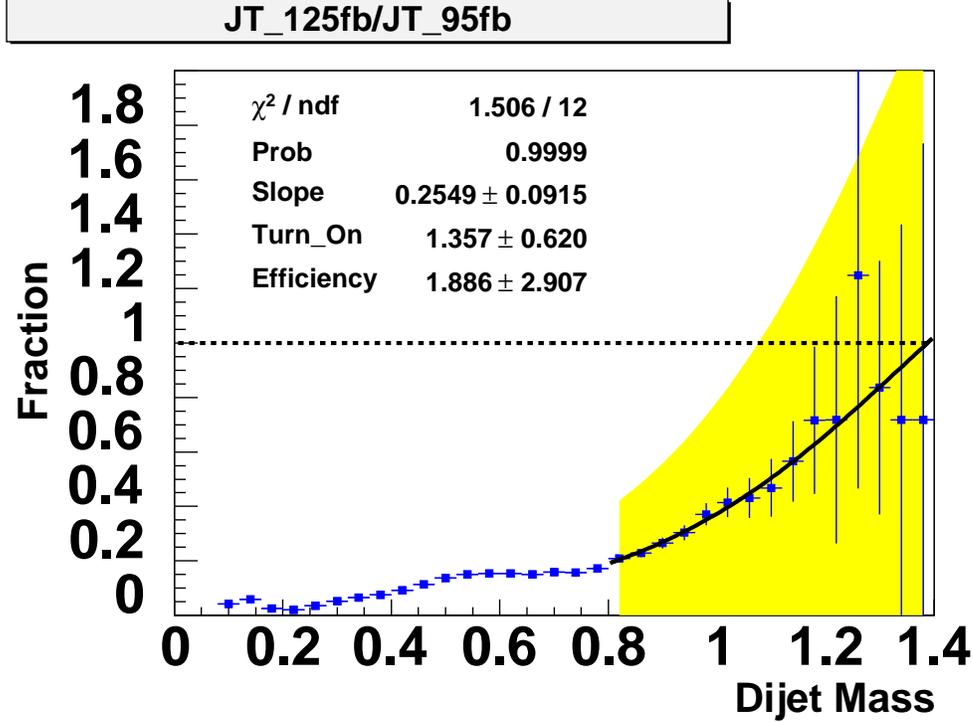
Figure 4.3: Trigger turn-ons for JT_125TT in the most forward bin ($2.0 < |y|_{max} < 2.4$). The ratio of the two triggers turns on very slowly and never plateaus within the kinematic range of the analysis. For this reason, the dijet mass triggers are used in the two most forward rapidity bins.

In order to find efficient triggers for each mass range, the turn-on curves and relative efficiencies for the triggers in mass, instead of p_T are needed. Since the absolute efficiency has already been determined, it is valid to use the relative efficiency to determine when a trigger has an efficiency that is approximately 100%. The trigger efficiencies and turn-ons are determined using consecutive triggers, i.e. JT_25TT is used to determine the efficient region in mass for JT_45TT. In order to minimize trigger inefficiencies, the turn-on is fitted with Eq. 4.5,

$$f(x) = \frac{\varepsilon}{1 + \exp((t - x)/s)} \tag{4.5}$$

---

[9]The trigger list is incremented when something changes. Calibrations can change, new triggers added, old triggers removed, etc.

where ε is the efficiency of the plateau region, t is the turn on (meaning the 50% point) and s is the slope. The plateau is then fitted with a constant value above a certain mass to make sure that the efficiency is around 100%.

All trigger turn ons are shown in Appendix D, which also has the information in tabular format (Table D.1). Fig. 4.4 shows two examples, one fitting the knee region (Eq. 4.3) and one fitting the plateau with a zeroth order polynomial. Table 4.1 shows the mass range each trigger is used in.
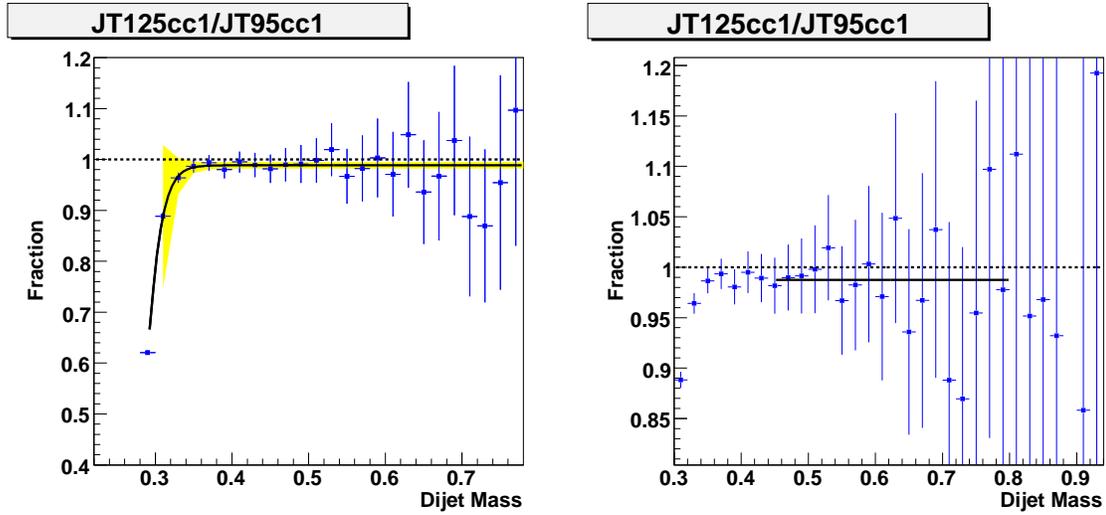
Figure 4.4: Trigger turn-ons for JT_125TT where the left shows the fit to the knee region and the right shows the fit to the plateau to show that trigger is efficient for the mass range it is used in.

The mass spectra from the various triggers overlaid are shown in Fig. 4.2.2 where the triggers have been scaled by their respective luminosity.
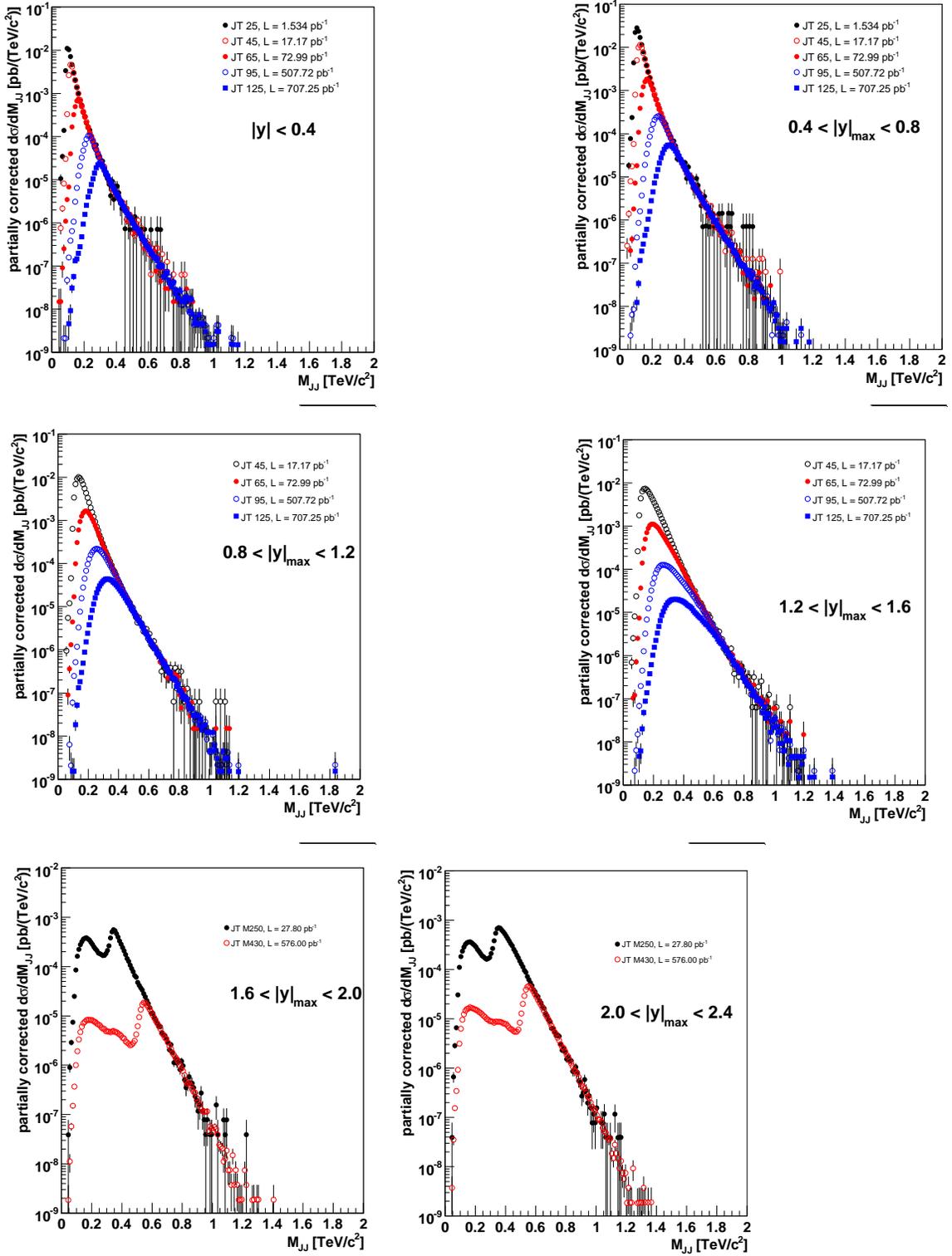
Figure 4.5: Mass spectra from the overlaid triggers used in the 6 rapidity regions, scaled by the appropriate luminosity.

Trigger	Luminosity (pb^{-1})	Rapidity Region	Mass Range (TeV)
JT_25TT	1.5	$|y| < 0.4$	0.150 - 0.200
JT_45	17.2	$|y| < 0.4$	0.200 - 0.250
JT_65	73.0	$|y| < 0.4$	0.250 - 0.350
JT_95	507.7	$|y| < 0.4$	0.350 - 0.450
JT_125	707.3	$|y| < 0.4$	> 0.450
JT_25TT	1.5	$0.4 < |y|_{max} < 0.8$	0.150 - 0.200
JT_45	17.2	$0.4 < |y|_{max} < 0.8$	0.200 - 0.250
JT_65	73.0	$0.4 < |y|_{max} < 0.8$	0.250 - 0.350
JT_95	507.7	$0.4 < |y|_{max} < 0.8$	0.350 - 0.450
JT_125	707.3	$0.4 < |y|_{max} < 0.8$	> 0.450
JT_45	17.2	$0.8 < |y|_{max} < 1.2$	0.250 - 0.400
JT_65	73.0	$0.8 < |y|_{max} < 1.2$	0.400 - 0.500
JT_95	507.7	$0.8 < |y|_{max} < 1.2$	0.500 - 0.700
JT_125	707.3	$0.8 < |y|_{max} < 1.2$	> 0.700
JT_45	17.2	$1.2 < |y|_{max} < 1.6$	0.300- 0.500
JT_65	73.0	$1.2 < |y|_{max} < 1.6$	0.500 - 0.600
JT_95	507.7	$1.2 < |y|_{max} < 1.6$	0.600 - 0.800
JT_125	707.3	$1.2 < |y|_{max} < 1.6$	> 0.800
JT_L3M250	27.8	$1.6 < |y|_{max} < 2.0$	0.450 - 0.700
JT_L3M430	576.0	$1.6 < |y|_{max} < 2.0$	> 0.700
JT_L3M250	27.8	$2.0 < |y|_{max} < 2.4$	0.450 - 0.700
JT_L3M430	576.0	$2.0 < |y|_{max} < 2.4$	> 0.700

Table 4.1: Regions of mass and rapidity where each trigger is used.

4.2.3 Mass bins

The final mass bins are chosen based on 3 requirements:

• Only one trigger is used in each mass bin, which cuts down on inefficiencies due to triggers having different prescales.

• The bin width is chosen to be about twice the RMS of the mass resolution. A simple study on the mass resolution was conducted on the QCD MC sample. A profile plot was made with the difference between the reconstructed mass and the particle mass on one axis and the particle level mass on the other. A plot was then made by taking different mass slices and plotting the RMS. This was then fitted with a line and the bin widths were determined from there. These bins were used as initial starting values for the simulation.

• The bin purity and bin efficiency are about 50% (see Fig. 4.6), which were calculated using DØJetSim (discussed in Section 4.3.3). The bin purity and efficiency are defined as:

  – efficiency = reconstructed and generated events in a particular bin / all events *generated* in that bin

  – purity = reconstructed and generated events in a particular bin / all events *reconstructed* in that bin
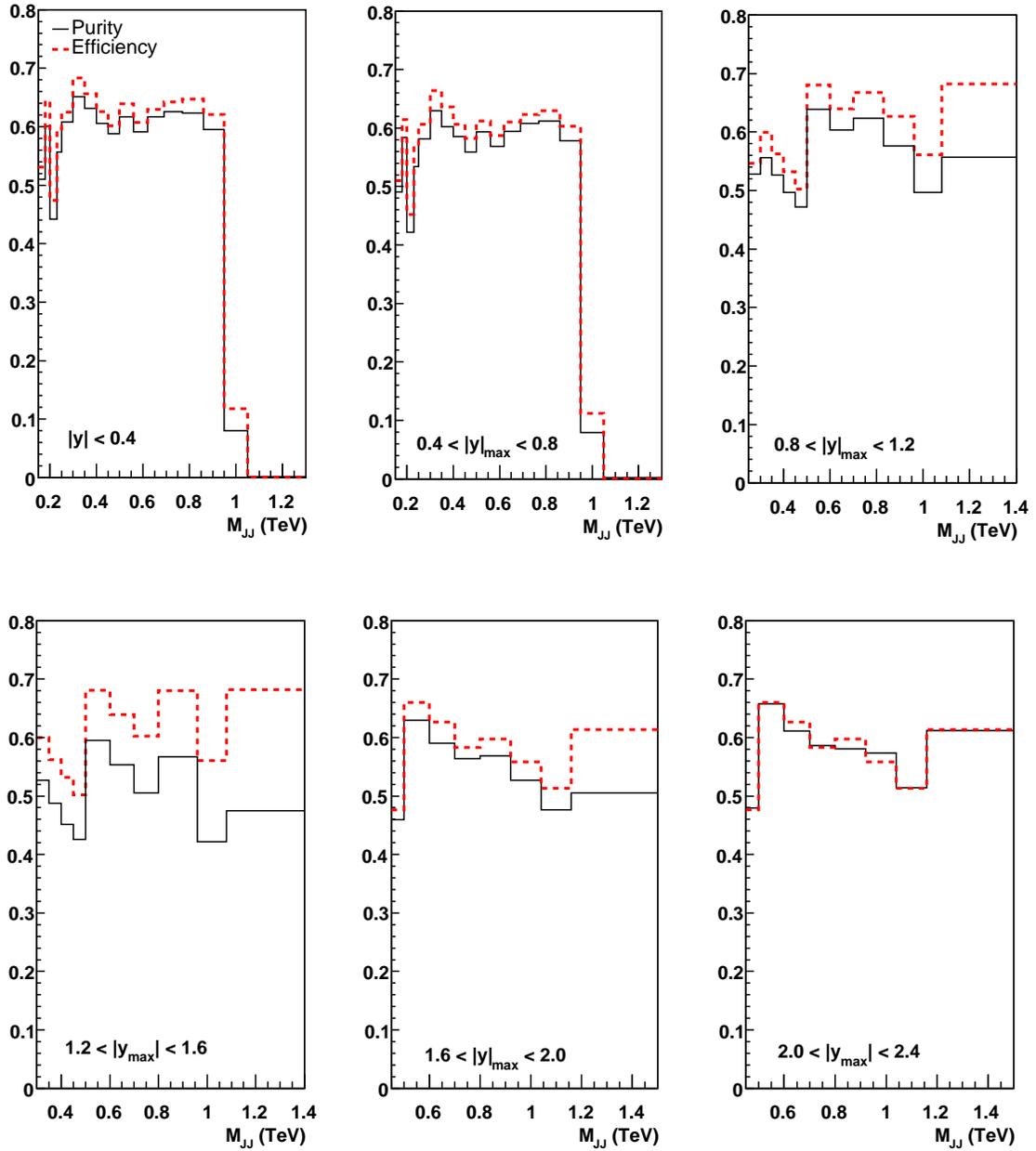
Figure 4.6: Bin purity and efficiency for each of the six rapidity regions, with the mass rescaled (See Section 4.3.2). The y-axis is the purity/efficiency and the x is the dijet mass in units of TeV.

The final mass bins and number of events per bin are shown in Tables 4.2 through 4.7, in Section 4.5.

4.3   Corrections

Ideally the detector would measure quantities as they occur in nature. In the real world, the detector is not perfect and what we measure is distorted compared to the actual quantity. In this analysis, we ultimately want to measure the particle level energies, meaning the energies of the actual particles as they enter the detector. Due to resolutions, what should be a delta function at 100 GeV, for example, is measured as a gaussian centered around 100 GeV. By carefully measuring resolutions in data, we can "unfold" the data, meaning we attempt to remove detector effects from the data. In a steeply falling spectrum, such as the dijet mass, events tend to migrate preferentially from a lower mass bin to a higher mass bin, leading to a distribution that has a steeper slope than the distribution would have if the resolution were perfect. This effect is shown in Fig 4.7.
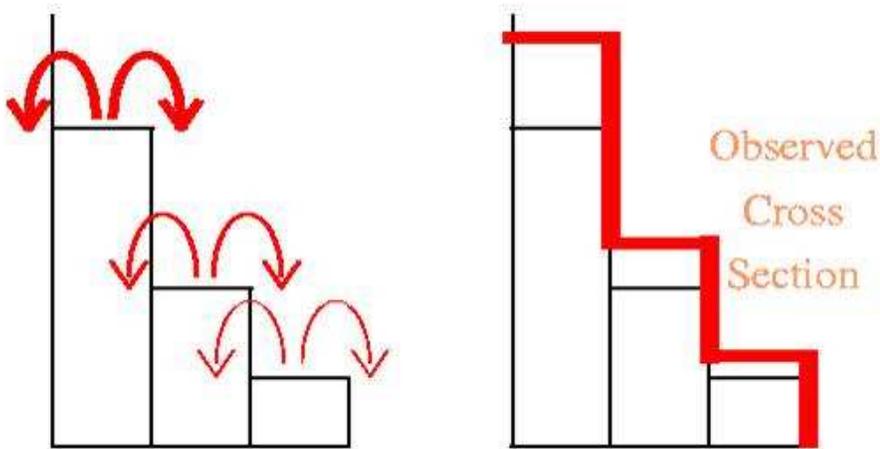
Figure 4.7: Cartoon showing how the events tend to migrate to a higher mass bin due to the steeply falling spectrum.

There are many methods to unfold data. One of the most popular methods is the use of an ansatz function as seen in [2], [3] and [16]. This method uses an ansatz function which has several free parameters. The function is chosen based on a priori knowledge of what the truth spectrum looks like. The parameters are then fitted to

the data distribution over several iterations. In this analysis, a MC reweighting and rescaling scheme is used instead. This has the advantage that no prior knowledge of the truth distribution is needed. The drawback is that MC does not always describe data and in the past has done an extremely poor job. However, MC simulations have improved to the point where it is valid to use this method.

4.3.1 Data corrections

The data is corrected for two effects before comparing with Monte Carlo; Jet Energy Scale (JES) effects and the vertex efficiency. These are standard corrections that many analyses in the collaboration use. In addition to these two corrections, the dijet mass is rescaled in both the data and the reconstructed MC as part of the unfolding scheme.

4.3.1.1 Vertex Corrections

The vertex is chosen based on the three cuts described in the Section 2, shown here as a reminder:

- $n_{vtx} \geq 1$
- $n_{trks} \geq 3$
- $|z_{vtx}| < 50$ cm

The second cut ensures that the vertex is good, while the third cut keeps the vertex in the high efficiency tracking region. While a simple way to calculate the vertex efficiency is to take the ratio of good events whose vertex passes these cuts to all good events, this method has a bias. The resolution is expected to be worse for high $|z_{vtx}|$ events than for low $|z_{vtx}|$ events. This bias could lead to the efficiency estimate being too low. A better way to calculate the vertex efficiency is with a study of the luminous region of the detector, which is based on luminosity and run number. Such a study was done by H. Schellman [41] and the results are applied in this analysis, using an application written by M. Voutilainen. The uncertainty due to this inefficiency is about 0.5%. The vertex acceptances are shown in Fig. 4.8 where the average effect from the vertex acceptance is about 92%.
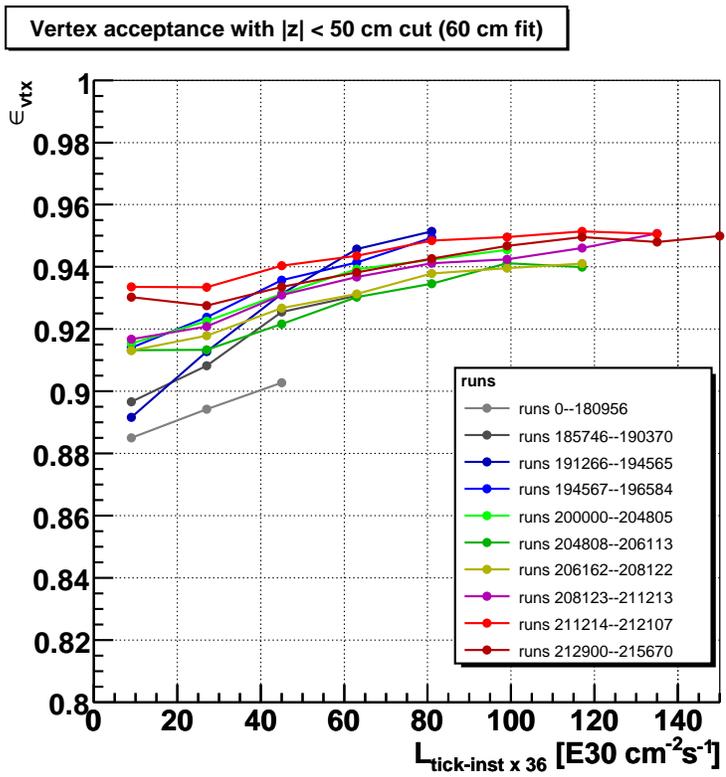
Figure 4.8: Vertex acceptance calculated with minbias events.

4.3.1.2 Jet Energy Scale

The goal of the Jet Energy Scale (JES) is to correct the calorimeter energy of jets back to whatever their energies were before interacting with the detector (particle level). Note that this is not the same as unfolding the data completely. There are other effects that must be considered. But on average, applying the JES returns the calorimeter jet energy to the particle level energy. The formula used to correct the jets is (Eq 4.6):

$$E_{jet}^{particle} = \frac{E_{jet}^{caloriemeter} - O}{F_\eta \cdot R \cdot S} \, k_{bias} \tag{4.6}$$

where O is the offset, R is the absolute response, F_η is the relative offset, S is the showering correction and k_{bias} is a correction for all other biases. These terms are discussed briefly below, more information can be found in [42] and [43]. A good description for the casual reader can also be found in [45].

In a beam crossing, there are several effects that lead to energy deposits in the calorimeter. There is the hard scatter that produces the jets we are interested in, but there is also the possibility that multiple interactions (MI) between the colliding beams will lead to additional energy in the event (noise). Additionally, there is extra energy that comes from electronic or uranium noise. Even when there is no beam in the machine, the calorimeter cells will still measure energy. Most of this is subtracted out using electronic calibrations, but there is usually some left over. Pile-up[10] can occur because the calorimeter's preamplifier works slower than the beam crossing. This means that an interesting signal may come in on top of the previous signal and the wrong baseline will end up being subtracted. In order to measure these effects, two samples are used. To test MI, minbias (MB) events are used. These are events that are triggered by the luminosity monitors, indicating an inelastic collision. Minbias events with 1 vertex are subtracted from MB events with multiple vertices to help determine multiple interactions. To determine the effect from noise, zero-bias (ZB) events are used. These have no specific trigger requirements and events are required to have no vertices. This is an easy way to get at the actual energy measured by the calorimeter when there are no actual interactions, so noise from uranium decays and electronics can be determined. The total offset correction is calculated for a cone

---

[10]Note that is different than the pile-up discussed in 2.2.4.

with a specific η by adding all the calorimeter towers in ϕ together. The final formula for the offset (O) is (Eq. 4.7):

$$O(N_{pv}, \eta, \mathcal{L}_{inst}) = MI(N_{pv}, \eta, \mathcal{L}_{inst}) - MI(N_{pv} = 1, \eta, \mathcal{L}_{inst}) + NP(N_{pv}, \eta, \mathcal{L}_{inst}) \quad (4.7)$$

where N_{pv} is the number of primary vertices and $\mathcal{L}_{inst}$ is the instantaneous luminosity. The amount of offset energy for several MI is shown in Fig 4.9.
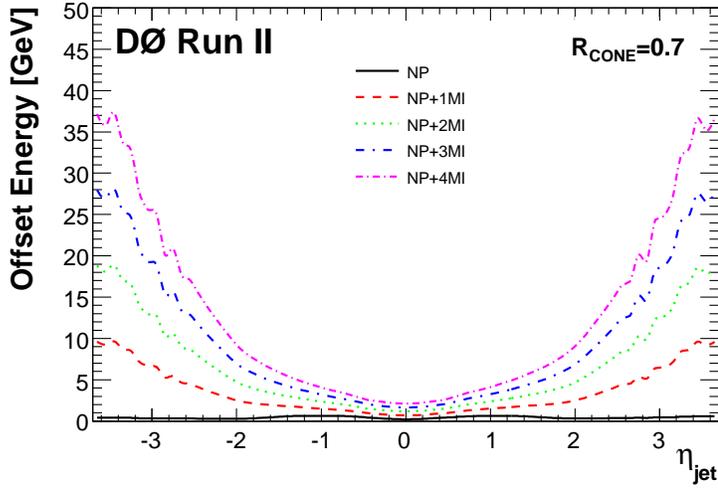
Figure 4.9: Final offset correction for Jet Energy Scale for $R = 0.7$ jets with different vertex multiplicities.

The largest numerical correction to the JES is the absolute response correction (R). The absolute response correction takes into account how the detector affects the data. The particles produced in the hard scatter will hit material in the tracking system before depositing energy in the calorimeter and some of the particles will be bent because of the magnetic field and not included in the jet as they should be. Also, the DØ detector is a non compensating calorimeter, which means that electromagnetic particles will have a different response (linear) to the calorimeter than hadronic particles, which have a more logarithmic response to the calorimeter. Since some of the detector effects are difficult to simulate in Monte Carlo, the response correction was calculated separately for data and MC. In MC, we can, in theory, compare the reconstructed jets (after the detector) with the true particle level energies. In data, a method called missing E_T projection fraction (MPF) is used. This method uses a tag object (a jet, Z, or γ) and compares the response with a probe object (a second jet). In an ideal detector, $\overrightarrow{p}_{T,tag} + \overrightarrow{p}_{T,recoil} = 0$, where the probe jet is part of the hadronic recoil, but doesn't account for all the energy in the recoil[11]. Our detector is not ideal and so we end up with something more like $\overrightarrow{p}_{T,tag}^{measured} + \overrightarrow{p}_{T,recoil}^{measured} = -\not{E}_T$. Using this information and the fact that $p_{T,i}^{measured} = R_i p_{T,i}$, where R_i is the response, the fractional response is given by Eq. 4.8

$$\frac{R_{recoil}}{R_{tag}} = 1 + \frac{\overrightarrow{\not{E}_T}^{measured} \cdot \eta_{T,tag}}{p_{tag}^{measured}} \qquad (4.8)$$

for a certain η. If we also make the assumption that $R_{recoil} \sim R_{jet}$, we can get the response of the jet we are looking for. This approximation is valid, as long as we require the tag and probe objects to be exactly back to back ($\Delta\phi > 3.0$) and we require the photon to have $|\eta| < 1.0$ and the jet to have $|\eta| < 0.4$. Figure 4.10 shows the final absolute response. The relative response correction (F_η) is a way to correct jets in all η regions by comparing a probe jet to the central response correction. The same MPF method is used to determine this and Figure 4.11 shows the final correction due to the relative response.

The next correction applied is the showering correction (S), which corrects for the net energy flow across the boundary of the cone used for the jet. This energy can come from particles that shower into the cone [12], but don't belong to the jet or from particles

---

[11]The remaining energy is from other jets involved in the event.

[12]Note that this is different from parton showering, discussed in 2.2.4, which occurs due to parton interactions. Here, showering occurs due to interactions with the detector material.
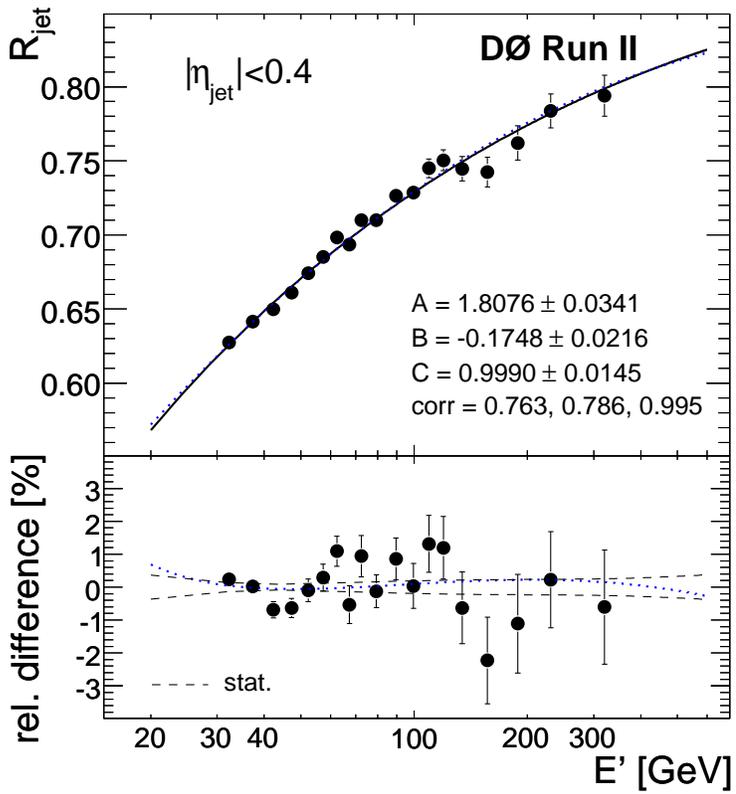
Figure 4.10: Final absolute response correction for Jet Energy Scale for $R = 0.7$ jets.
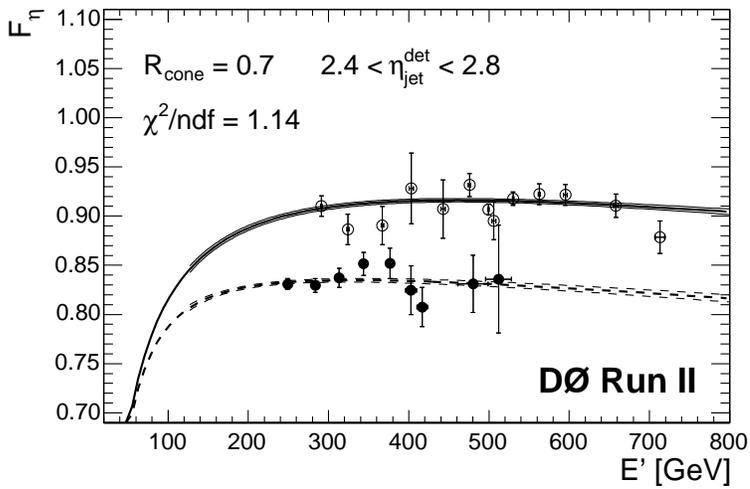
Figure 4.11: Final relative response correction for Jet Energy Scale for $R = 0.7$ jets.

that belong to the cone but not all their energy is contained by it. This showering comes from detector effects and not gluon radiation. The data and MC showering corrections are calculated separately. The data method is described here; for the MC method, please see [45]. The sample used to obtain the showering correction in data is the same as that used for the response correction (MPF). The events are back to back $\gamma + 1$ jet, with one and only one vertex to cut out multiple interaction effects. A tag and probe method is used here, but unlike the response correction, the tag object is allowed to be in any η region. The correction is determined using the energy distribution in rings of $\Delta R(y, \phi)$ with respect to the jet axis; this is called the jet energy profile. The jet energy profile was determined using the same jet reconstruction algorithm discussed in Section 2.2.3. The MC is used to determine the energy from the particle jets reconstructed in the cone and energy from particle jets not in the cone. Several studies were done using various templates to make sure using the MC was valid in the data. The final showering correction is given by the ratio of the visible energy in a cone of $\Delta R < R_{cone}$ compared to the visible energy that comes from the total particle jet energy. This correction is given in Eq. 4.9.

$$\frac{E_{jet}^{measured}}{E_{particlejet}^{measured}} = \frac{E_{particlejet}^{measured, \Delta R < R_{cone}}}{E_{particlejet}^{measured}} + \frac{E_{notparticlejet}^{measured, \Delta R < R_{cone}}}{E_{particlejet}^{measured}} \quad (4.9)$$

Fig. 4.12 shows the motivation for using a showering correction and Fig. 4.13 shows the correction due to showering.

The final piece of the JES correction is the additional biases k_{bias}. Additional biases comes from additional energy in the event that isn't subtracted out or is introduced by the previous corrections discussed. For instance, there can be additional energy that comes from low energy in a jet that helps a cell pass the zero-suppression threshold. To determine the size of this, the offset and response is compared with and without a zero-bias overlay. The effect is small, but large enough to warrant a correction. The MPF method also introduces several biases, all at the $\sim 1\%$ level. These biases come from an incomplete understanding of the event structure. Since jets are only reconstructed above a threshold of $p_T = 6$ GeV, there can be unclustered energy left over in events. Also, since photons are more narrow objects than jets, the response can be off. For more information on the additional biases, see [42].

Since the JES uncertainties are so small, the corrections derived on the $\gamma + jet$ sample are not applicable to all samples. This is mainly due to the difference between quark initiated jets and gluon initiated jets, shown in Fig. 4.14. Since this analysis is
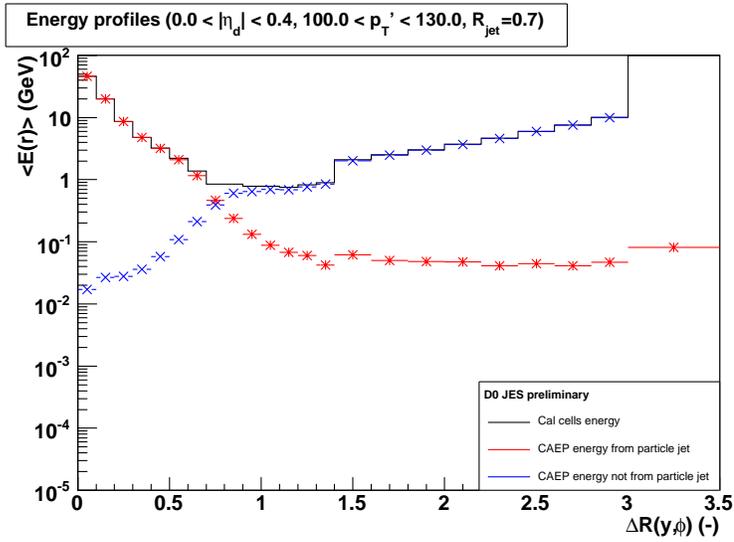
Figure 4.12: Showering motivation for the Jet Energy Scale. The blue stars shows the energy not from particle jets in the cone, while the red crosses shows energy from particle jets in the cone.
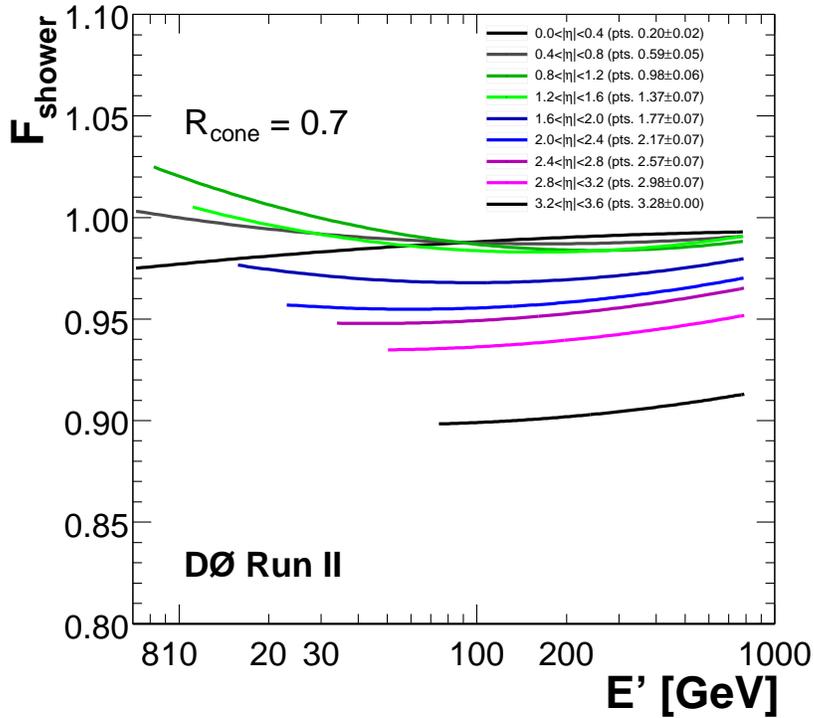
Figure 4.13: Showering correction for the Jet Energy Scale for $R = 0.7$ cone jets.

done using dijet events, which are dominated by gluons jets at low energies, additional corrections are needed [43]; Fig. 4.15 shows a simulation of the difference in gluon initiated jets in the $\gamma + jet$ sample vs the dijet sample. These corrections were determined using the relative response bias method for the forward regions and using MC simulations for the central region. Also, since the jets for Run II are massive, the correction for the transverse momentum is different from the energy and from the angles. These differences were studied in detail with the result being all 4 components of the jet 4 momentum having their own correction factors. This version of the Jet Energy Scale is known as J4S and the uncertainties associated with it are given in [43] (change to appendix). Before the jets are used in this analysis, they are corrected and the dijet mass is calculated from the corrected jet values.
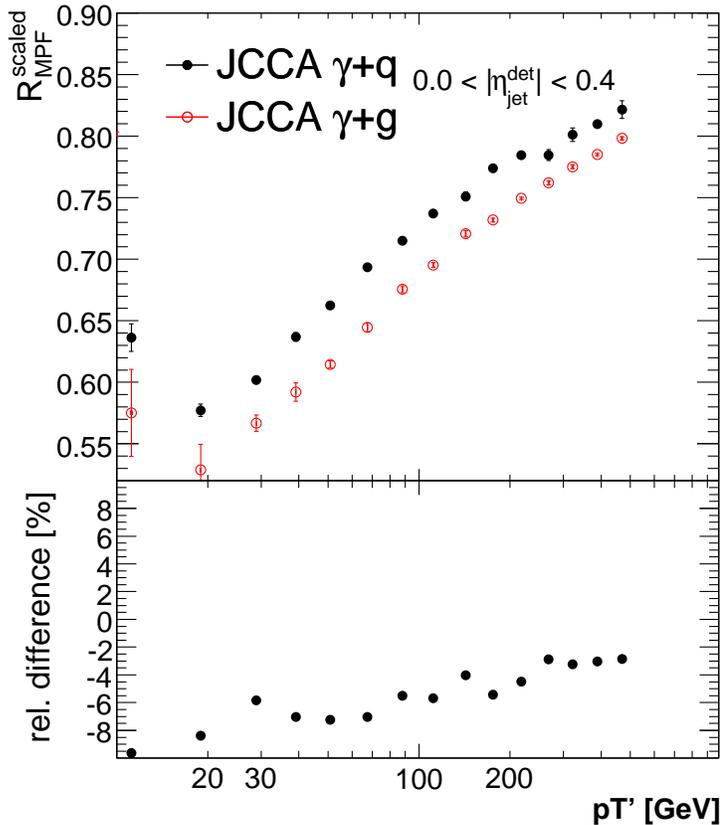
Figure 4.14: Simulated difference in the JES between quark and gluon initiated jets, using MC.
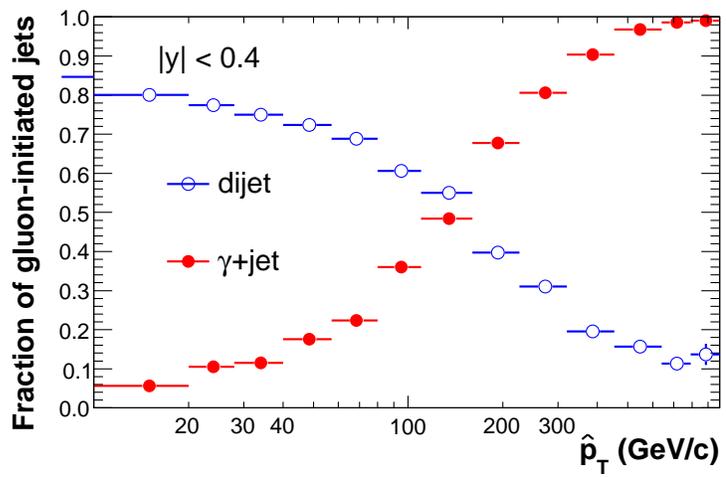
Figure 4.15: Fraction of gluon initiated jets in the $\gamma + jet$ sample (open circles) vs the dijet sample (closed circles).

The final J4S correction factor for jets with $p_T = 100$ GeV is shown in Figure 4.16.
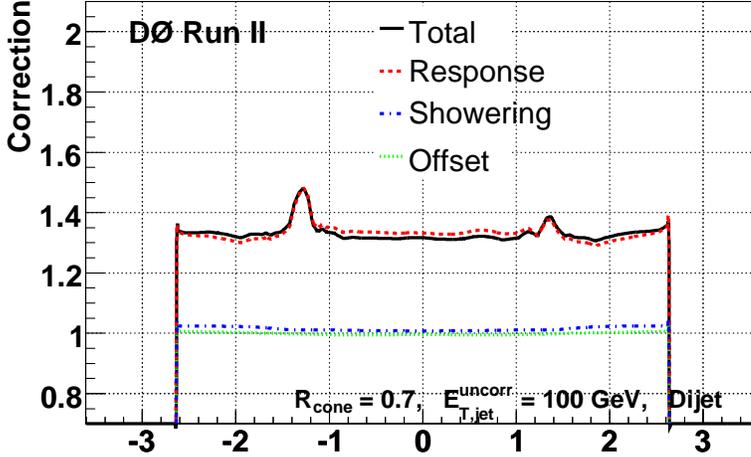
Figure 4.16: J4S correction factor as a function of η for jets with $p_T = 100$ GeV.

4.3.2 Mass Rescaling

In addition to the J4S and vertex efficiency corrections, this analysis uses mass rescaling [37]. While the JES corrects the four momentum to its most likely mean value, there is still a large background from events migrating from low mass bins to high mass bins. This is a consequence of the steepness of the dijet mass spectrum. This rescaling will move events that are close to the bin edge back to a more appropriate bin. Usually, the procedure is to use an ansatz method or reweighting to make the MC match the data. This works as a general approach. But if we want to do a bin by bin correction, as is done in this analysis, the reco MC values need to map to the same values in the truth MC on a bin by bin basis. Using the rescaling will lower the efficiency of finding an event in a bin, but will improve the purity as seen in Fig 4.17.

To rescale the mass, a scale factor dependent on the reconstructed mass and y_{max}

$$M_{JJ}^{rescaled} = M_{JJ}^{reco} - (0.004 + 1.875 * |y|_{max})(M_{JJ}^{reco})^2 \tag{4.10}$$

is used. The rescaling equation is determined using two dimensional plots such as Fig 4.18 for each rapidity region. Ideally, the reco mass and the truth mass should be completely correlated. The equation is chosen such that after it's applied to the

73
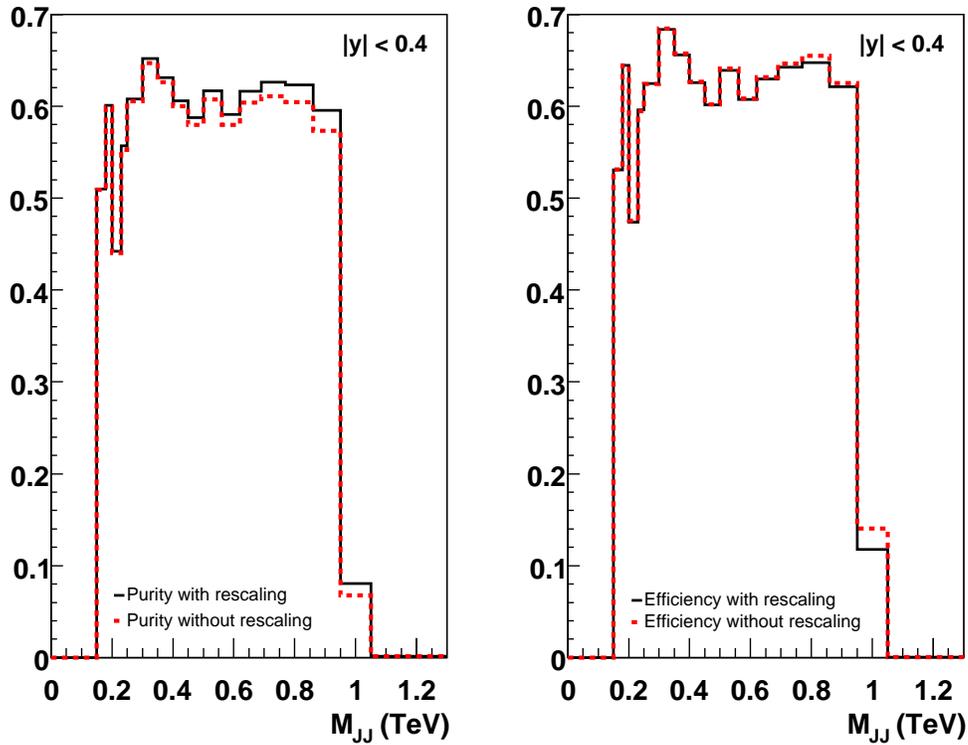
Figure 4.17: Comparison of the efficiency and purity with and without rescaling in the Central region. The effect is similar in the ICR and Forward regions. The dijet mass is shown along the x-axis and the purity/efficiency is along the y-axis. The rescaling improves the purity, as seen on the left, but some efficiency is lost, as seen on the right.

reco mass, the mean value of the two dimensional plot lies along a diagonal line at 45 degrees (see Fig 4.18). In addition, the same function must be applicable in all rapidity regions, so a rapidity dependence is included. A rescaling is chosen to balance a high purity with a high efficiency. Note that the rescaling is applied to both reco MC and data, after an appropriate reweighting has been determined.

The main advantage to rescaling is that the statistical uncertainties will more accurately represent the true number of events produced in a certain bin. Fig 4.18 shows how well the rescaling does in the first central bin (CC1), while Fig 4.19 shows the rescaling in IC1.
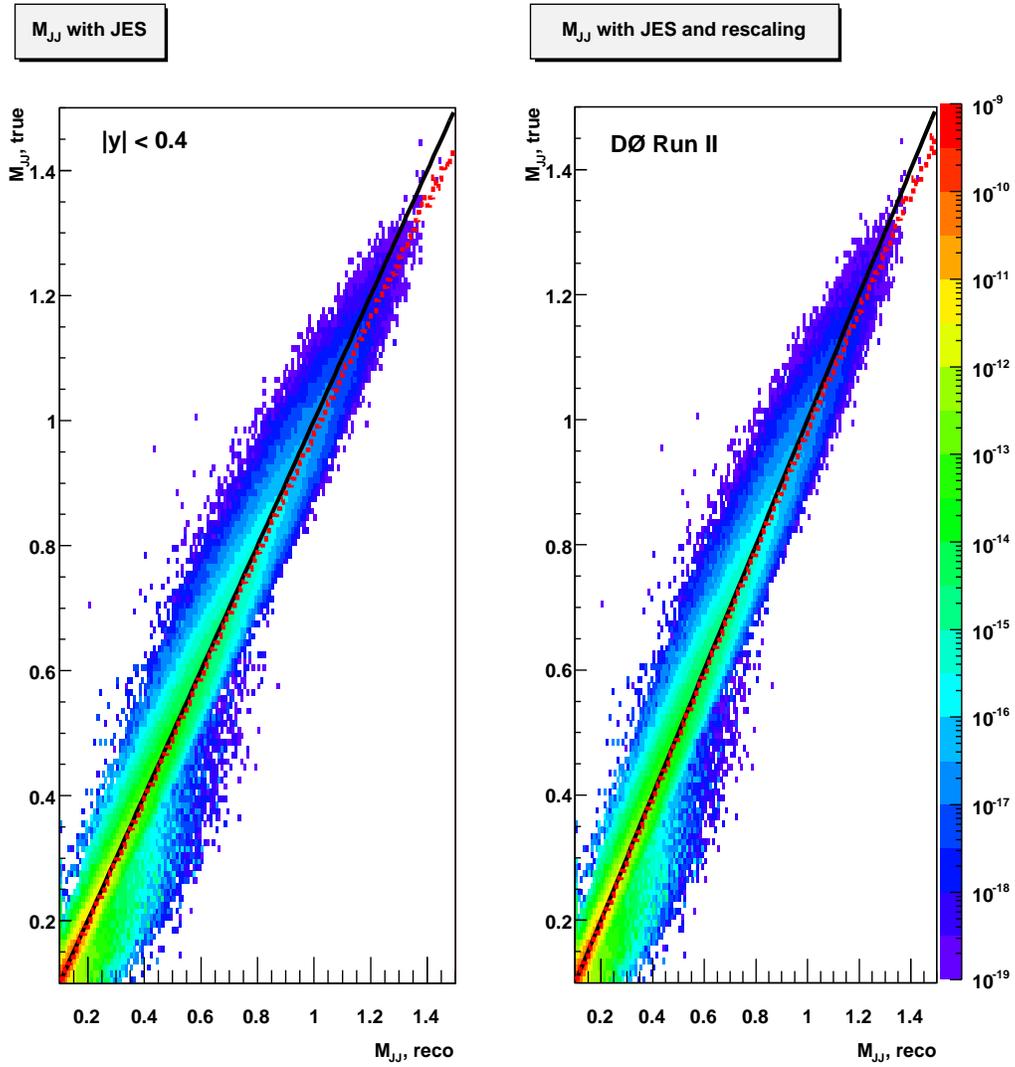
Figure 4.18: The reco mass vs the true mass with and without rescaling in region CC1. The solid black line shows a perfect diagonal while the dashed line is the relationship between reco and true mass. The plot on the left shows that the reco and true mass are off from each other, particularly at high mass, while the plot on right shows the agreement after rescaling.
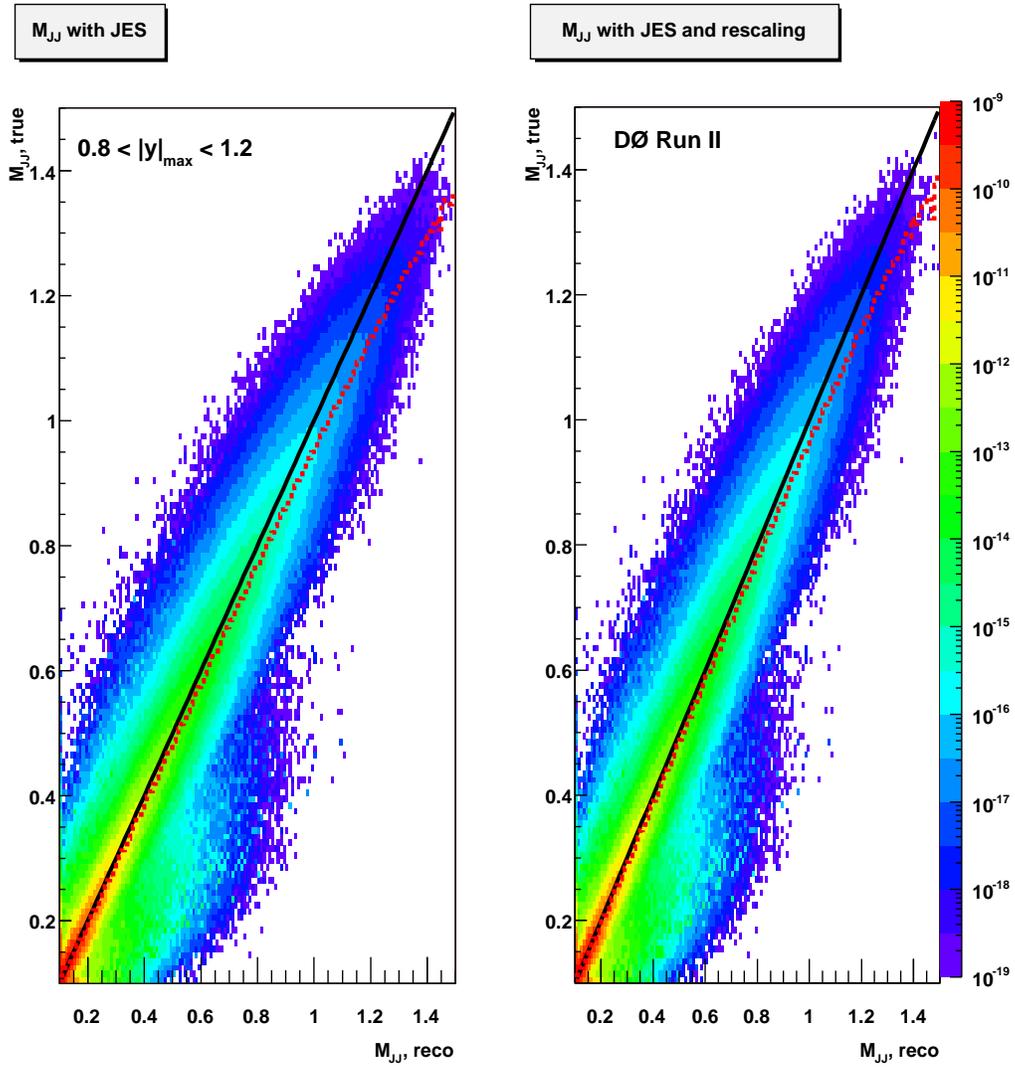
Figure 4.19: The reco mass vs the true mass with and without rescaling in region IC1. The solid black line shows a perfect diagonal while the dashed line is the relationship between reco and true mass. Appendix E shows similar plots for the other analysis rapidity bins.

The partially corrected data with the vertex efficiency, J4S and the mass rescaling applied is shown in Fig 4.20. The MC mass spectra from DØJetSim is shown as well. More information on rescaling can be found in [37] and [44].
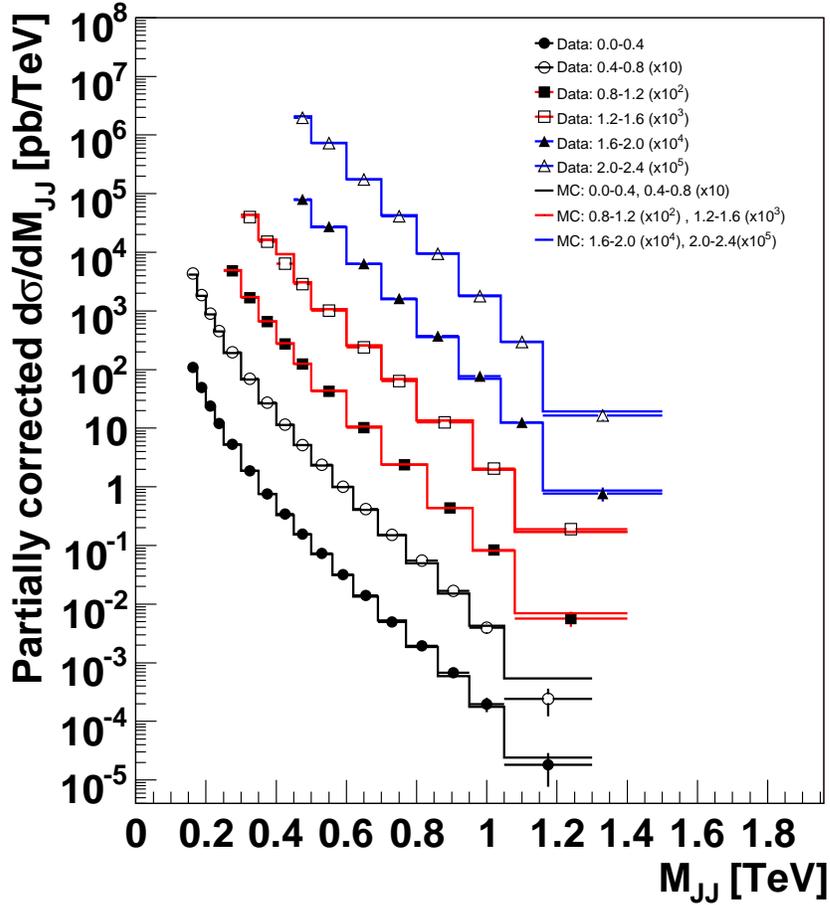
Figure 4.20: Partially corrected data, which includes J4S, vertex efficiency and the mass is rescaled. The mass spectrum from MC is also shown.

4.3.3 Detector Effects, Correction Factors, and Control plots

Once data is corrected for J4S, vertex efficiency and mass rescaling, corrections for remaining detector effects are determined using a MC simulation. The final unfolding corrections can then be calculated. These corrections were determined using DØJetSim [47], which was developed by M. Wobisch and is summarized here.

DØJetSim is a fast MC that simulates the DØ detector using data based measurements used exclusively for DØ jet measurements. JetSim uses PYTHIA 6.419 [29] with tune QW [30] and MSTW2008 LO PDFs [23], where the jets are reconstructed with the DØ cone algorithm, as input. Using various data studies, it models all relevant detector effects, such as p_T resolutions [48] and η and ϕ resolutions [49]. All the variables relevant to the dijet mass (see list below) are made into control plots. The reconstructed MC control plots from JetSim is then compared to data control plots that has been corrected with J4S and vertex efficiency, to test that the simulation accurately describes the dijet mass (observable) in data. Note also that both samples have the mass rescaling applied. Control plots are created in a low mass range ($M_{JJ} < 0.7$ TeV) and a high mass range ($M_{JJ} > 0.7$ TeV) to show this agreement. The control distributions are listed below.

- The z - position of the vertex: z_{vtx}.

- The highest p_T jet in the event: $p_{T,max}$.

- The second highest p_T jet in the event: $p_{T,min}$.

- The p_T of the third jet in the event: $p_{T,3}$.

- The rapidity of the most forward jet of the leading pair and rapidity bin: $|y|_{max}$.

- The rapidity of the second most forward jet of the leading pair and rapidity bin: $|y|_{min}$.

- The ratio of the second leading to leading jet in p_T: $p_{T,2}/p_{T,1}$.

- The ratio of the third leading to the second leading jet in p_T: $p_{T,3}/p_{T,2}$.

- The difference in ϕ between the 2 leading jets in p_T : $|\phi_1 - \phi_2|$

- The η distribution from the center of the detector (0,0,0): η_{det}.

The results of the comparisons can then be used to determine a weighting function for the MC, which is described in Section 4.3.3.1. In this particular analysis, the mass itself was reweighted and the control plots allow us to see that the reweighting affects the variables in an expected way. In general, any variable from the control plots may be used to determine the reweighting. An example control plot before reweighting

is shown in Fig 4.21, while the control distribution after reweighting is shown in Fig. 4.22. The full set of control plots can be found in Appendix F. In all cases, the control plots are shown in a low mass ($M_{JJ} < 0.7$ TeV) and a high mass ($M_{JJ} > 0.7$ TeV) region, for all 6 rapidity regions. JetSim does a good job of describing the first two jets, but does a poor job of describing the third jet angular distributions. The p_T are described fairly well, as seen in the control plots. The poor description of the third jet is believed to be because of a weakness in PYTHIA, but is not expected to have a large impact on this analysis. The second and third jets, if close in p_T can swap, which can result in a change to the dijet mass spectrum. This change can happen when a far forward, low p_T third jet changes the rapidity bin the event is placed in. The jet swapping will mostly effect the low mass bins, since it is rare for three high p_T jets to be in an event. In the central region, the second and third jet are close in p_T less than 1% of the time. In the ICR, they are close 1-2% of the time and in the forward they are close 4-6% of the time. The overall change to the mass spectrum is not expected to have an impact on the final mass spectrum.
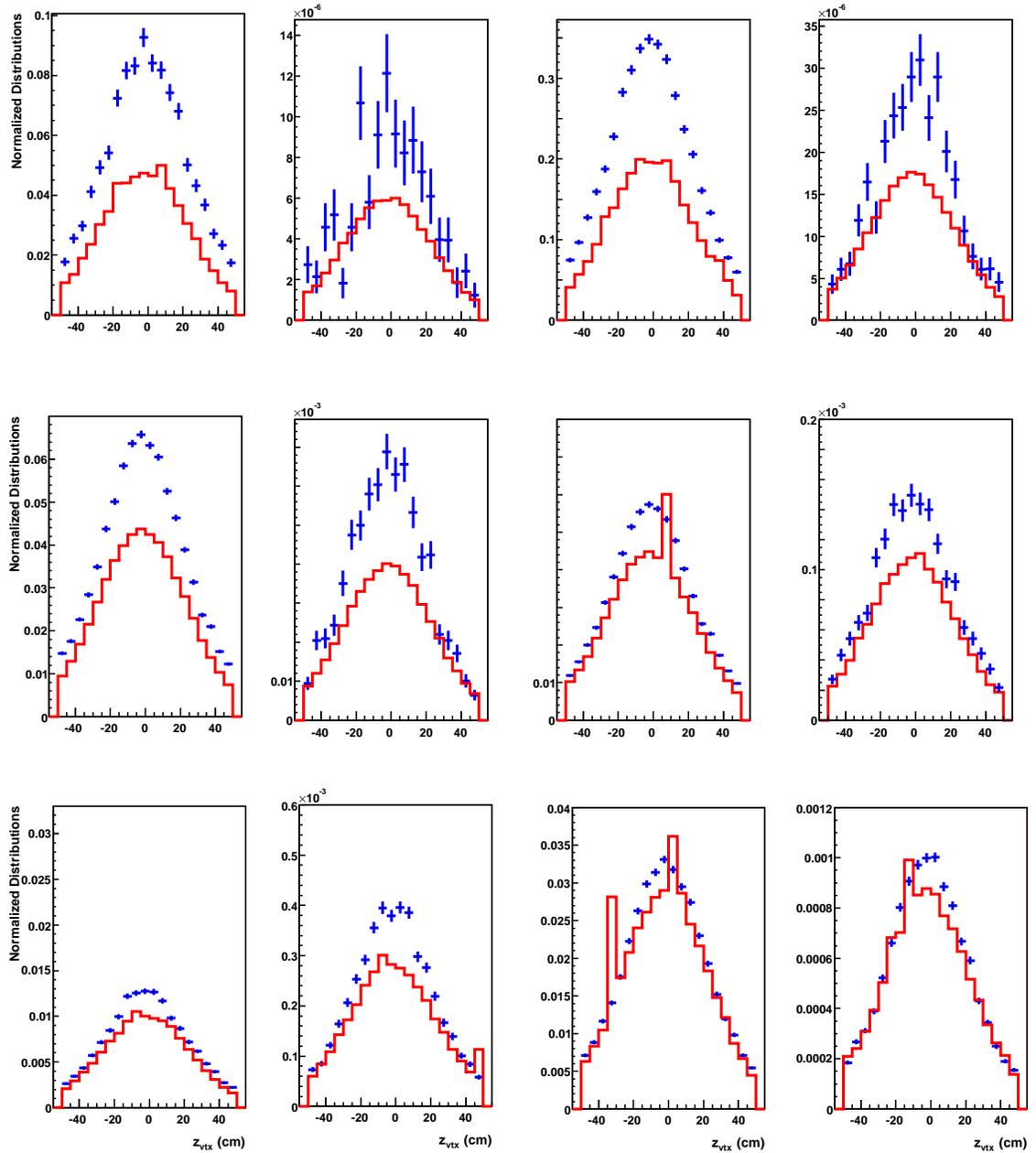
Figure 4.21: Control distributions for $|z_{vtx}|$ before reweighting. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass. This is pattern is followed for all subsequent control plots.
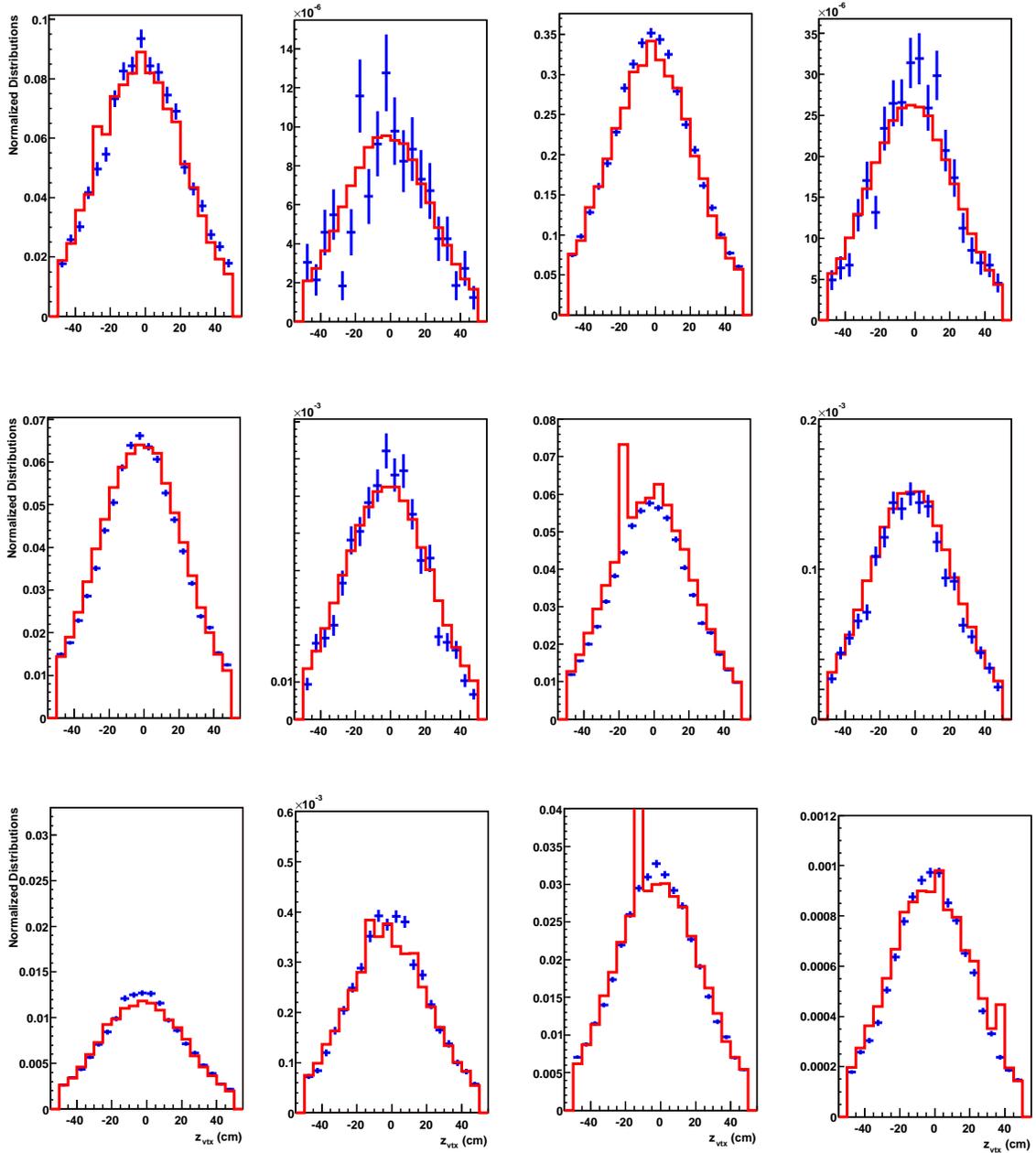
Figure 4.22: Control distributions for $|z_{vtx}|$, after reweighting. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass. This is pattern is followed for all subsequent control plots.

4.3.3.1 Reweighting

The goal of unfolding is to remove detector effects from the measured energies. In order to determine the detector effects, a MC simulation is used and must accurately model the data. If the reco MC does not look like the data, we can't trust that the correction factor used will accurately return the truth distribution. The reco MC is reweighted to data until it matches. After that, the reweighting is applied to the truth level MC, which is then run through the same reconstruction process as the data. After several iterations, the reco MC will look like the data within uncertainties and we can trust our simulation to unfold the data. The procedure used to determine the reweighting is described in this section.

To reweight the reco MC, first data was divided by unweighted reco MC (Fig. 4.23). The mass was fitted with a 2nd order polynomial shown in Eq. 4.11 where A, B, and C are linear functions in $|y|_{max}$ (Eq. 4.11 through Eq. 4.14). In order to get the ICR region to agree, parameter B required a gaussian in addition to the linear term. The gaussian covers the ICR region while going to zero in the central and forward regions.

$$f(M_{JJ}) = A + BM_{JJ} + CM_{JJ}^2 \tag{4.11}$$

$$A = 2.17 - 0.51|y|_{max} \tag{4.12}$$

$$B = -1.99 + 0.92|y|_{max} + 0.316\exp\left(-\frac{(y-1.2)^2}{0.115}\right) \tag{4.13}$$

$$C = 1.51 - 0.64|y|_{max} \tag{4.14}$$

The parameters A, B, and C were tuned until the reweighted MC was flat when compared to data.

The final data to MC mass ratios are shown in Fig 4.24 and the reweighting function is shown in Fig. 4.25.
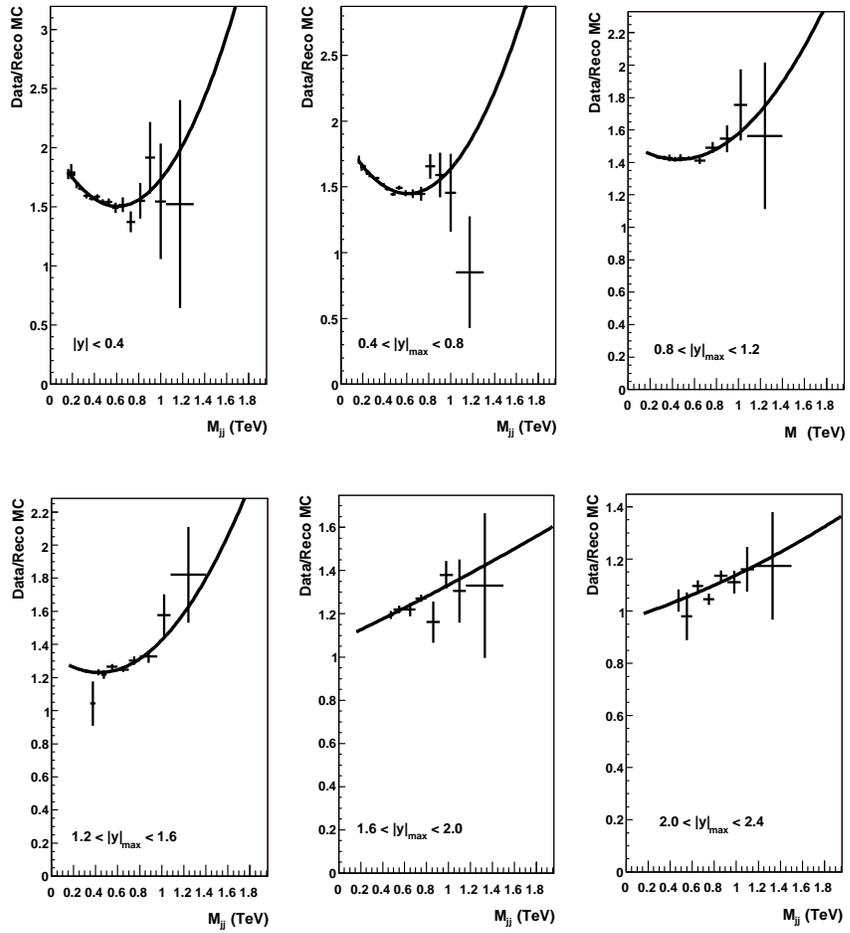
Figure 4.23: Data/MC mass ratios showing the unweighted MC compared to the partially corrected Data, with a quadratic fit.
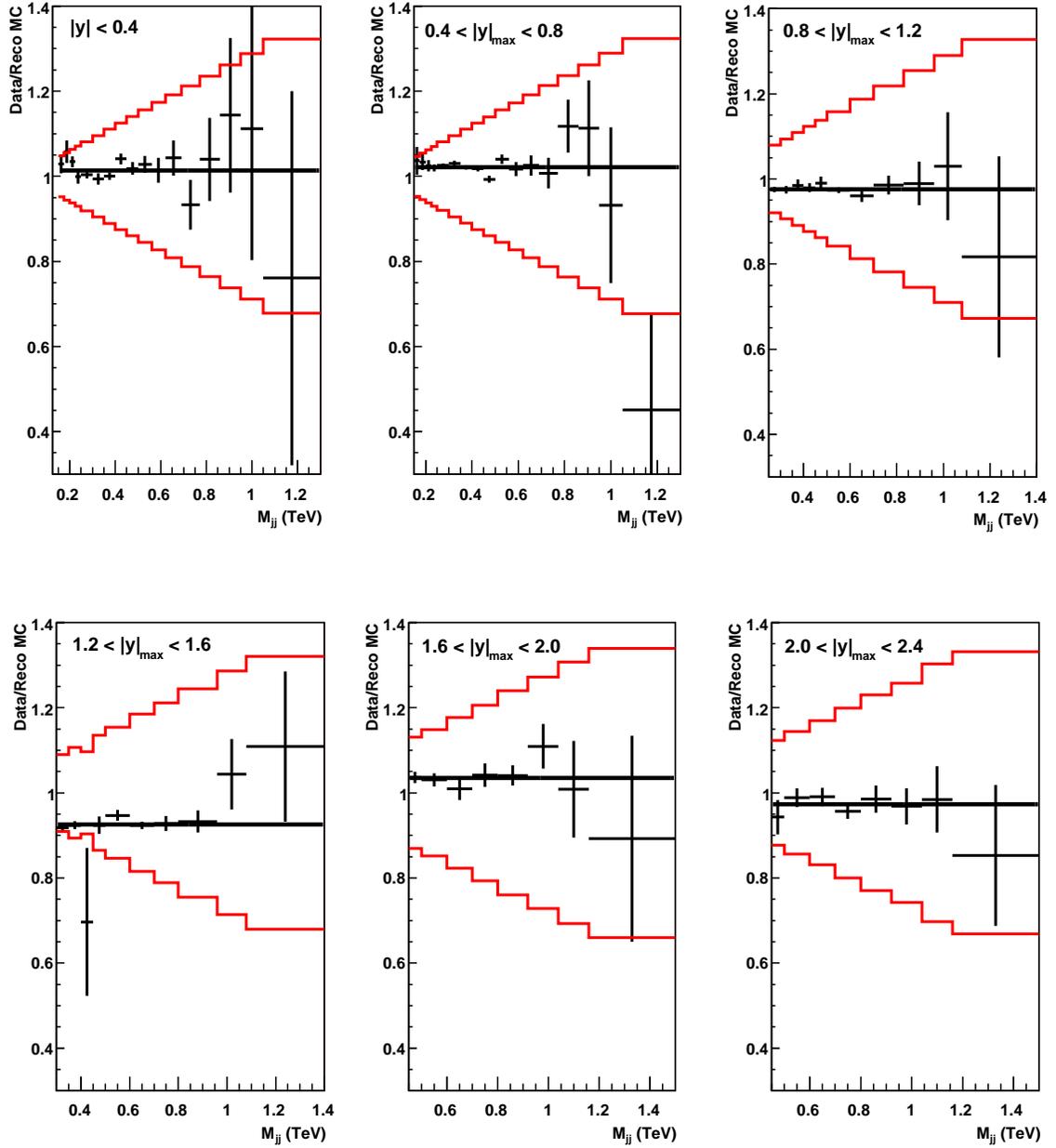
Figure 4.24: Data/MC mass ratios showing the weighted MC compared to the partially corrected Data. The outer red bands show the variations in reweighting used to calculate the reweighting uncertainty.
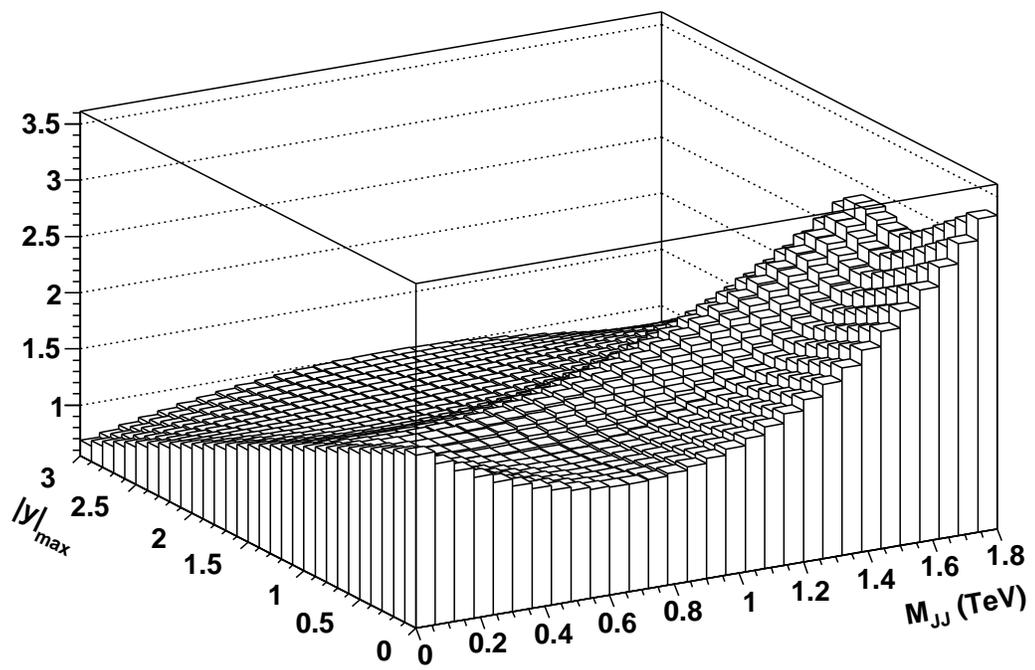
Figure 4.25: Reweighting function (Eq. 4.11) used in JetSim.

4.3.4 Correction Factors

Once the reco MC describes the data, the reco MC can be used with the particle level MC to determine the unfolding factor (also called correction factor). The JetSim package provides the final correction factors for all effects from the detector, which are described below. These correction factors are then applied to the data to correct back to particle level. More detailed descriptions can be found in [37].

• The p_T resolution was first studied by M. Voutilainen as a function of y, defined from the reconstructed vertex. Since JetSim is a detector simulation, all the inputs need to be in detector variables. Therefore, the p_T resolutions were redone using the variable η_{det}, which is defined from the center of the detector, and is discussed in more detail in Section 4.3.5.

• The η and ϕ resolutions were taken from a study done by Z. Hubacek [49]. In this study, $\eta_{ptcl} - \eta_{det}$ versus jet p_T distributions were fitted with a double gaussian. The widths of the two gaussians plus f, the fraction of area underneath them were then fitted with reasonable functions in p_T. The results from the η resolutions are also used for the ϕ resolutions.

• If the position of the primary vertex is shifted from the nominal position, a difference in the η_{jet} and η_{det} distributions will result. This is taken into account in the simulation.

• Misvertexing happens when the wrong event vertex is identified, which will change the jet's reconstructed four momentum. The simulation takes this into account by generating two vertices - one that is correct and one that is wrong. The jet's three momentum components can then be recalculated and the effect calculated.

• The last effect to account for is the jet ID efficiency. For dijet events, the MC is weighted by the product of the efficiencies for both jets in the event. The jet ID efficiency was calculated from [46]. In addition to detector effects, the simulation takes into account the following technical effects:

• MC reweighting. The Monte Carlo must describe all properties in data, such as the p_T resolution correction, which is dependent on the shape of the mass distribution. In order to determine the uncertainty due to reweighting, the mass spectrum must be varied by multiplying the rescaling equation (4.10) with:

$$f(M_{JJ}) = 1 \pm (M_{JJ}(TeV) * 0.3) \tag{4.15}$$

The factor of 30% is chosen arbitrarily and is considered to be conservative.

- Muon/neutrino corrections in jets. The jet calibration software at DØ removes these before correcting back to particle level. These effects are calculated from PYTHIA and put back in.

The total correction factor is shown in Fig. 4.26, for all rapidity regions. This includes all detector effects and the corrections due to the muon/neutrino energies. In the analysis, the fitted line is applied to the data. The individual correction factors are also shown in Appendix H.
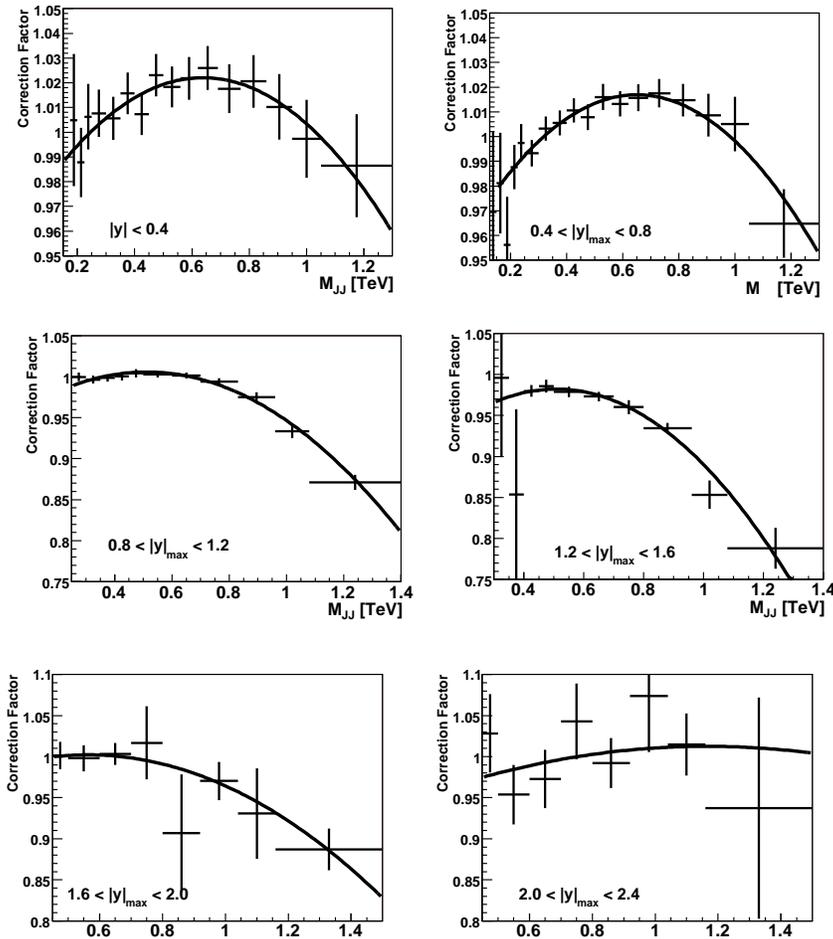
Figure 4.26: Total correction factors for all 6 rapidity regions. The solid black line shows the function that is applied to data, while the points show the bin-by-bin value. The fit smooths out statistical fluctuations.

4.3.5 p_T Resolutions

The p_T resolutions are the largest correction in the unfolding procedure. Normally, the resolutions are derived exclusively from data, but we found that there were small corrections from high p_T tails that affect the data resolutions, but cannot be accurately measured in data. Therefore, these p_T resolutions have additional MC corrections added. This section will give a brief introduction to the methods used; the full documentation can be found in [48].

The core jet p_T resolutions are measured in data from the width of the asymmetry distributions, which is defined in Eq. 4.16,

$$A = \frac{p_{T,2} - p_{T,1}}{p_{T,2} + p_{T,1}} \tag{4.16}$$

using back to back dijet events[13], where $p_{T,1}$, $p_{T,2}$ are the two leading jets in the event and their order is randomized.

The same data sample used in the analysis was used to determine the p_T resolutions, including all the data and event quality cuts. The η_{det} bins were chosen so that the resolution has a smooth behavior over the width of the bin. The jets are binned in 0.4 increments (0.0 - 3.6) in η_{det} except in the ICR region, where finer binning was required due to rapidly changing resolutions. In this region the bins are: $0.8 < |y| < 1.1$, $1.1 < |y| < 1.3$, and $1.3 < |y| < 1.6$. The width of the asymmetry distributions will be denoted by RMS_{asy}.

The Monte Carlo true jet p_T resolutions are described by the RMS of reconstructed p_T (p_T^{reco}) vs particle p_T (p_T^{ptcl}), as defined in Eq. 4.17.

$$RMS_{res} = RMS\left(\frac{p_T^{reco} - p_T^{ptcl}}{p_T^{ptcl}}\right) \tag{4.17}$$

The relationship between the Monte Carlo truth resolution and the asymmetry distribution is defined in Eq. 4.18,

$$\frac{\sigma_{p_T}}{p_T} = RMS_{res} \approx \sqrt{2}[RMS_{asy}] \tag{4.18}$$

where both jets are central. The variable σ_{p_T} is the jet p_T resolution and the variable RMS_{asy} is the asymmetry resolution. Note that Eq. 4.18 becomes Eq. 4.19 in the more general case where one jet is in the rapidity region being probed and the other is in the central reference region $|\eta_{det}| < 0.8$.

---

[13]Back to back jets are defined as jets with $\Delta\phi > 3.0$

$$\frac{\sigma_{tag}}{p_T} = \sqrt{4[RMS_{asy}]^2 - 2[RMS_{asy}]^2_{ref}}, \tag{4.19}$$

The above describes the resolutions assuming there are only two jets in the event. However, in events with more than two jets and additional radiated energy, there are two effects that must be taken into account: particle level imbalance and soft radiation corrections. Particle level imbalance comes from physics effects, such as the particle shower ending outside the fixed cone radius. Soft radiation comes from jets that are in a event, but do not get reconstructed due to the p_T cut in the DØ software ($p_T > 6$ GeV).

The Monte Carlo truth distributions have been shown to have nongaussian tails. Because of this, using a simple gaussian σ is not sufficient to describe the width of the resolutions and the RMS of the distribution is used instead. The tails are not seen in data because Eq. 4.16 is symmetric by constructions, whereas Eq. 4.17 is not. This effect is shown in Fig. 4.27. The top row shows the asymmetry variable in data, and the width of these distributions is RMS_{asy}. The bottom row shows the difference between particle level and reco level in MC truth, where the width of these distributions are RMS_{res}. The tails seen in Fig. 4.27 come from invisible energy at the particle level (muon and neutrino leakage), calorimeter punch-through[14] in MC truth at high p_T in CC and EC, and miscalibration and punch-through in the ICR.

---

[14]Punch-through is caused by the fact that the DØ calorimeter is not deep enough to contain all the energy in a jet.
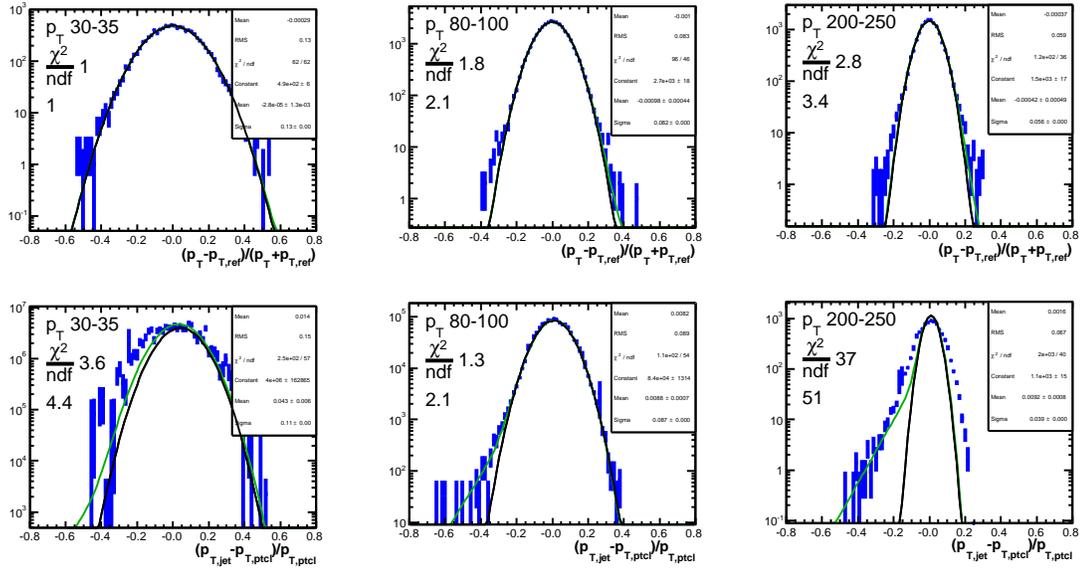
Figure 4.27: Upper row: Gaussian fits to the dijet asymmetry in data for $30 < p_T < 35$ GeV, $80 < p_T < 100$ GeV and $300 < p_T < 400$ GeV, $|y| < 0.4$. Bottom row: Gaussian fits in same p_T bins for true particle level resolution in Monte Carlo (unweighted p_T spectrum). Note that the $\sqrt{2}$ from Eq. 4.18 is included.

Due to these effects, the resolution is best described by two gaussians and an exponential tail as defined in Eq. 4.20, where GPT stands for gaussian-punch-through-tails. Equation 4.23 shows the RMS used to determine the width of the resolutions.

$$g_{GPT}(x, \mu, \sigma, P, \lambda, H, \mu_H, \kappa) \;=\; (1-P)(1-H)g_{det}(x, \mu, \sigma) \tag{4.20}$$

$$+(1-P)Hg_{det}(x, \mu + \mu_H, \kappa\sigma)$$

$$+\frac{P\lambda}{2}\exp\left(\lambda(x-\mu+\frac{\lambda\sigma^2}{2})\right)\cdot\text{erfc}\left(\frac{x-\mu+\lambda\sigma^2}{\sqrt{2}\sigma}\right)$$

$$\langle x \rangle \;=\; \mu - P/\lambda + \Delta\mu, \quad \Delta\mu = H(1-P)\mu_H \tag{4.21}$$

$$\tag{4.22}$$

$$RMS(x) \;=\; \sqrt{(\sigma^2 + H(1-P)[(\kappa^2-1)\sigma^2 + \mu_H^2 + \frac{2P}{\lambda}\mu_H] - \Delta\mu^2 + \frac{P(2-P)}{\lambda^2})}.$$

Figure G.1 shows the parameters from the tail as a function of p_T and for each η region, and are defined as follows:

- μ is the mean of the central gaussian.
- P is the punch through fraction of the jets

$$P = p_T(P_0 + P_1 \cdot p_T) \tag{4.23}$$

- λ is the energy loss constant

$$\lambda = \lambda_0 + \lambda_1 \cdot p_T \tag{4.24}$$

- H is the fraction of jets that end up in the second gaussian
- μ_H is the relative shift with respect to the central gaussian

$$\mu_H = \mu_{H,0} + \log(0.01p_T) \cdot (\mu_{H,1} + \log(0.01) \cdot \mu_{H,2}) \tag{4.25}$$

- κ a scale factor that relates the width of the second gaussian in terms of the central gaussian σ.

$$\kappa = \frac{\sqrt{1.1^2 + 0.794^2 \cdot \kappa_0^2 \cdot p_T + 0.0608^2 \cdot \kappa_1^2 \cdot p_T^2}}{\sqrt{1.1^2 + 0.794^2 \cdot p_T + 0.0608^2 \cdot p_T^2}} \tag{4.26}$$

A table with the values of the tail parameters for all η_{det} is in Appendix G, as well as figures showing how the tail parameters behave as a function of η_{det}. Sigma

is calculated by solving Eq. 4.23 for σ where μ, P, λ, H, μ_H, κ have already been solved and $RMS(x)$ is calculated from fitting the asymmetry (seen in Fig. 4.27) with Eq. 4.27. Note that this σ is the main gaussian's width.

$$RMS_{asy} = \sqrt{\frac{N^2}{p_T{}^2} + \frac{S^2}{p_T} + C^2}, \tag{4.27}$$

where N is the noise term, S is the stochastic term and C is a constant. From this, the final resolutions can be calculate, using Eq. 4.20; Figure 4.28 shows the core gaussian width in the various η_{det} regions.
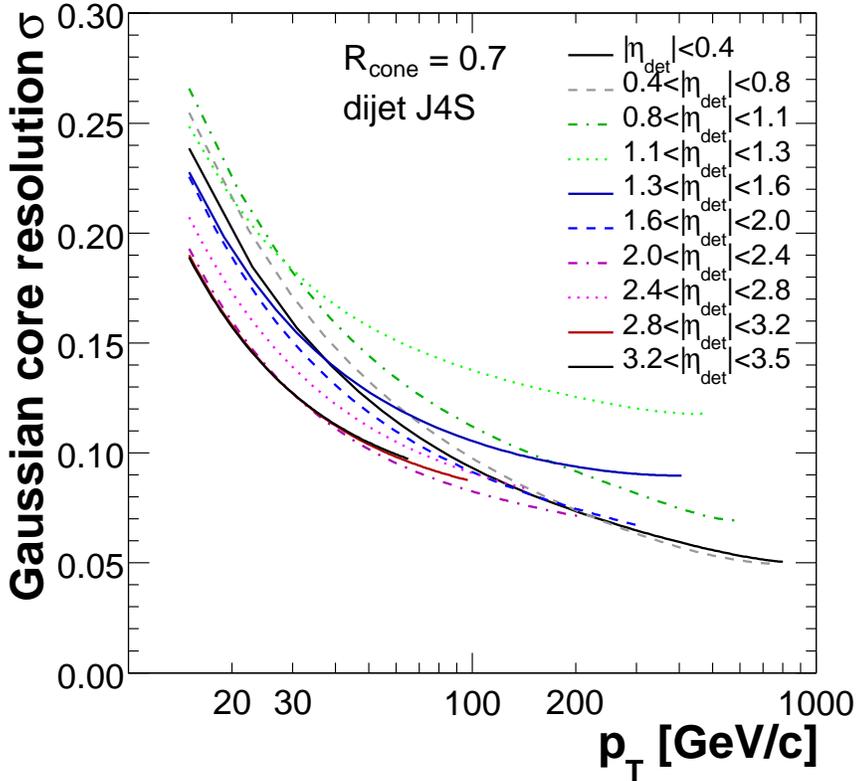
Figure 4.28: σ for the main gaussian using the GPT method

The final resolutions are shown in Fig. 4.29. The resolution is the best in the central and forward regions and slightly worse in the ICR. The added region in the ICR (1.1-1.3) has the worst resolution, as expected.
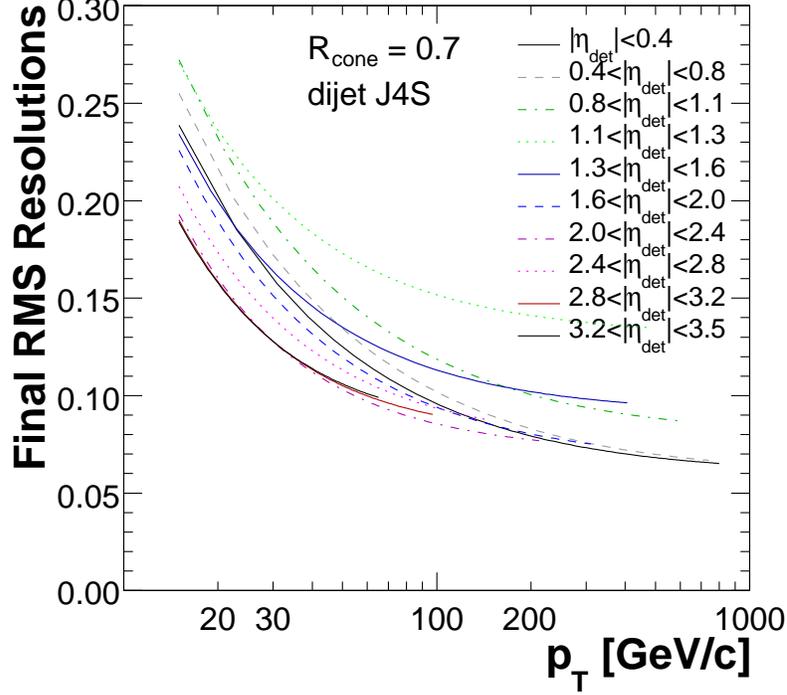
Figure 4.29: Final RMS (GPT) resolution for data as implemented in qcd_jet_caf v01-00-03.

4.4 Uncertainties

There are two types of uncertainties in this analysis; statistical and systematic. Statistical uncertainties refer to the uncertainty due to the finite amount of data we have. In each individual bin, the percent statistical uncertainty is given by $\sqrt{N}/N$, where N is the number of events in that bin. The systematics are more complicated. These come from detector effects, such as how well we know the resolution and the uncertainty due to the luminosity measurement, and characterize any uncertainty in methods used. The systematic uncertainties were obtained using DØJetSim. Note that in some cases, there are fluctuations from highly weighted events that can cause overly large uncertainty factors. Before the uncertainty factors were applied to data, they were fitted with either a quadratic or a quartic function to factor out the behavior of these over weighted events. The uncertainties along with the fits are shown

in Figures 4.30 to 4.39. The J4S uncertainty sources (49 total) and the p_T uncertainty sources (15 total) were added in quadrature (respectively) to give an overall uncertainty due to these effects. They are discussed more below.

● There are 49 correlated uncertainties associated with J4S. The uncertainty due to each source and the total uncertainty added in quadrature is shown in Fig. 4.31. This is the largest uncertainty in the dijet mass analysis. A table with all eigenvectors and a brief description can be found in Appendix I.

● There are 15 uncertainties associated with the p_T resolution. The individual sources and the total uncertainty is shown in Fig. 4.30:

  – Systematic uncertainty for final fit.

  – Systematic uncertainty for soft radiation correction.

  – Systematic uncertainty for particle level imbalance.

  – Systematic uncertainty for closure residual.

  – Systematic uncertainty for noise component in fit for CC1, CC2, ICR1, ICR2, EC1 and EC2.

  – Statistical uncertainty of the fit in all 9 rapidity regions.

The following effects have less well defined sources of uncertainties and it is unclear how they are correlated. To account for this, there is an overall, correlated variation up and down, as well variations on the order of one σ (up and down) separately in each of the six rapidity regions. The final uncertainty for the effect is chosen to be the largest variation up or down. In most cases, it is the overall correlated variation that is the largest. For some of the smaller uncertainties ($\approx 2\%$), a solid dashed line represents the uncertainties applied.

● The η resolutions uncertainties are shown in Fig. 4.32 and are 0.15% to 0.3% in the central regions, 0.2% to 0.5% in the ICR regions and 0.5% in the forward regions.

● The ϕ resolution uncertainties are shown in Fig. 4.33 and are 0.15% to 0.38% in the central regions, 0.03% to 0.07% in the ICR regions and 0.02% to 0.06% in the forward regions.

● The η bias uncertainties are shown in Fig. 4.34 and are 1% in the central regions, 2% in the ICR regions and 4% in the forward regions.

● The z vertex uncertainties are shown in Fig. 4.35 and are 0.06% in the central regions, 0.16% in the ICR regions and 0.03% to 1% in the forward regions.

• The uncertainties due to the misvertexing are shown in Fig. 4.36 and are 0.05% to 0.09% in the central regions, 0.09% in the ICR regions and up to 0.03% in the forward regions.

• Reweighting uncertainties are shown in Fig. 4.37 and is fitted with a quadratic function.

• The jet ID uncertainties are shown in Fig 4.38 and are 0.6% in the central regions, 1.1% in the ICR regions and 0.6% in the forward regions.

In addition to the uncertainties from DØJetSim, there is an overall uncertainty associated with the luminosity [50], which is about 6.1 %, and the uncertainty due to the vertex inefficiency which is about 0.5%. The total uncertainty from all sources except the luminosity is shown in Fig. 4.39.
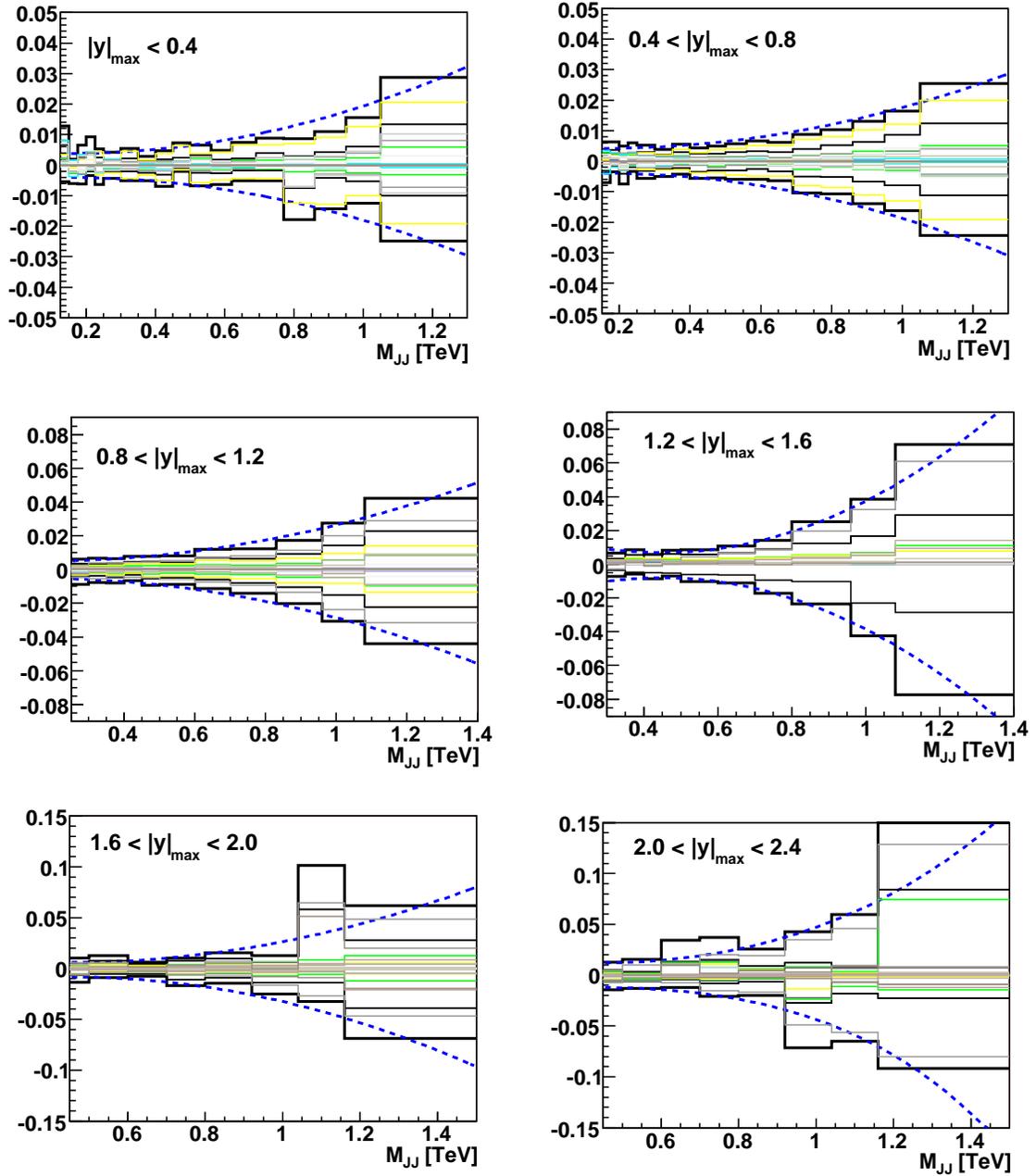
Figure 4.30: p_T resolution uncertainty. The outermost black line shows all 15 sources added in quadrature, while the blue dashed line shows the fit used to determine the final uncertainty.
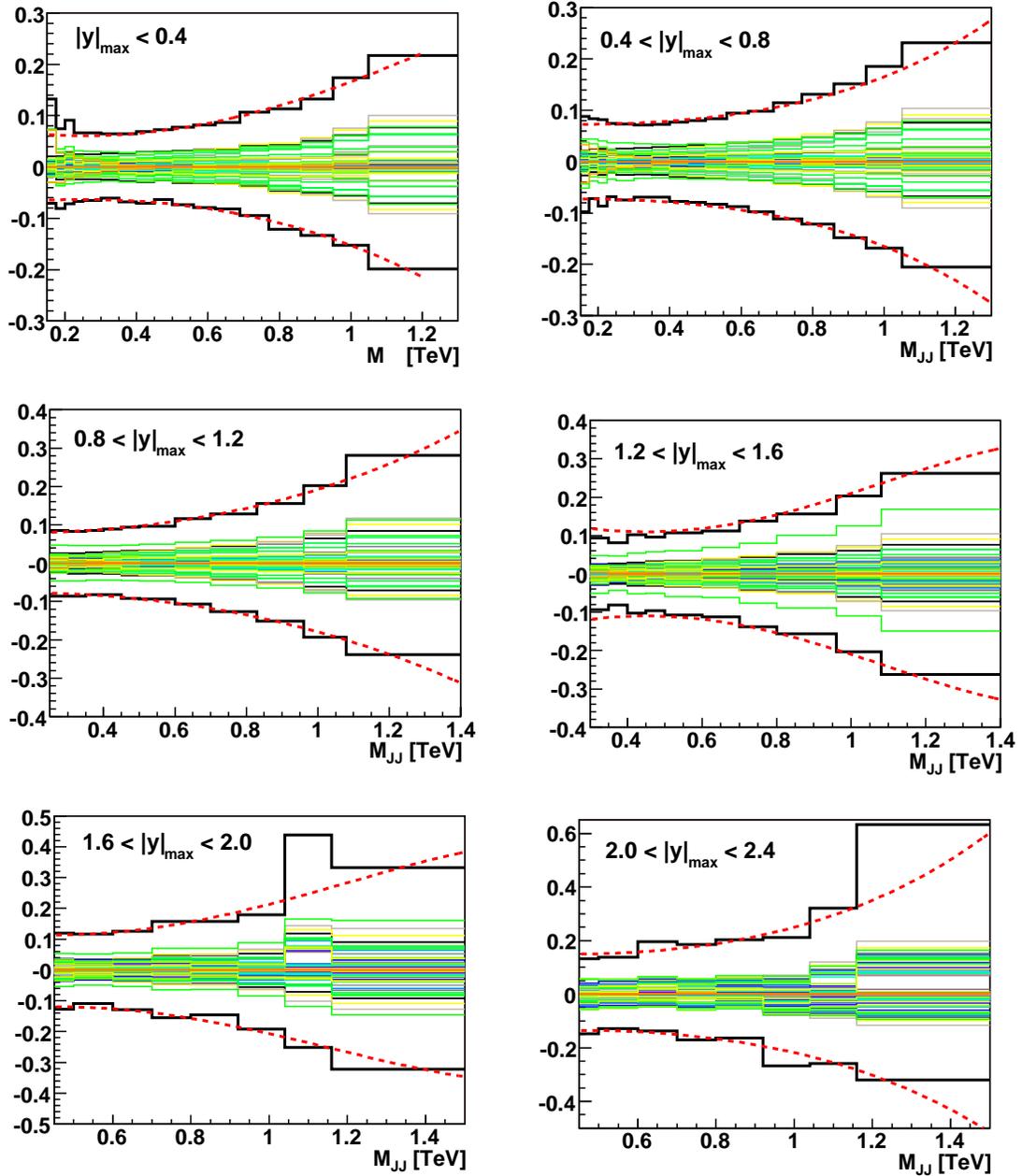
Figure 4.31: JES Uncertainty. The outer black line shows all 49 sources added in quadrature, while the red dashed line shows the fit used to determine the final uncertainty. The y axis is the percentage uncertainty and the x axis is the dijet mass.
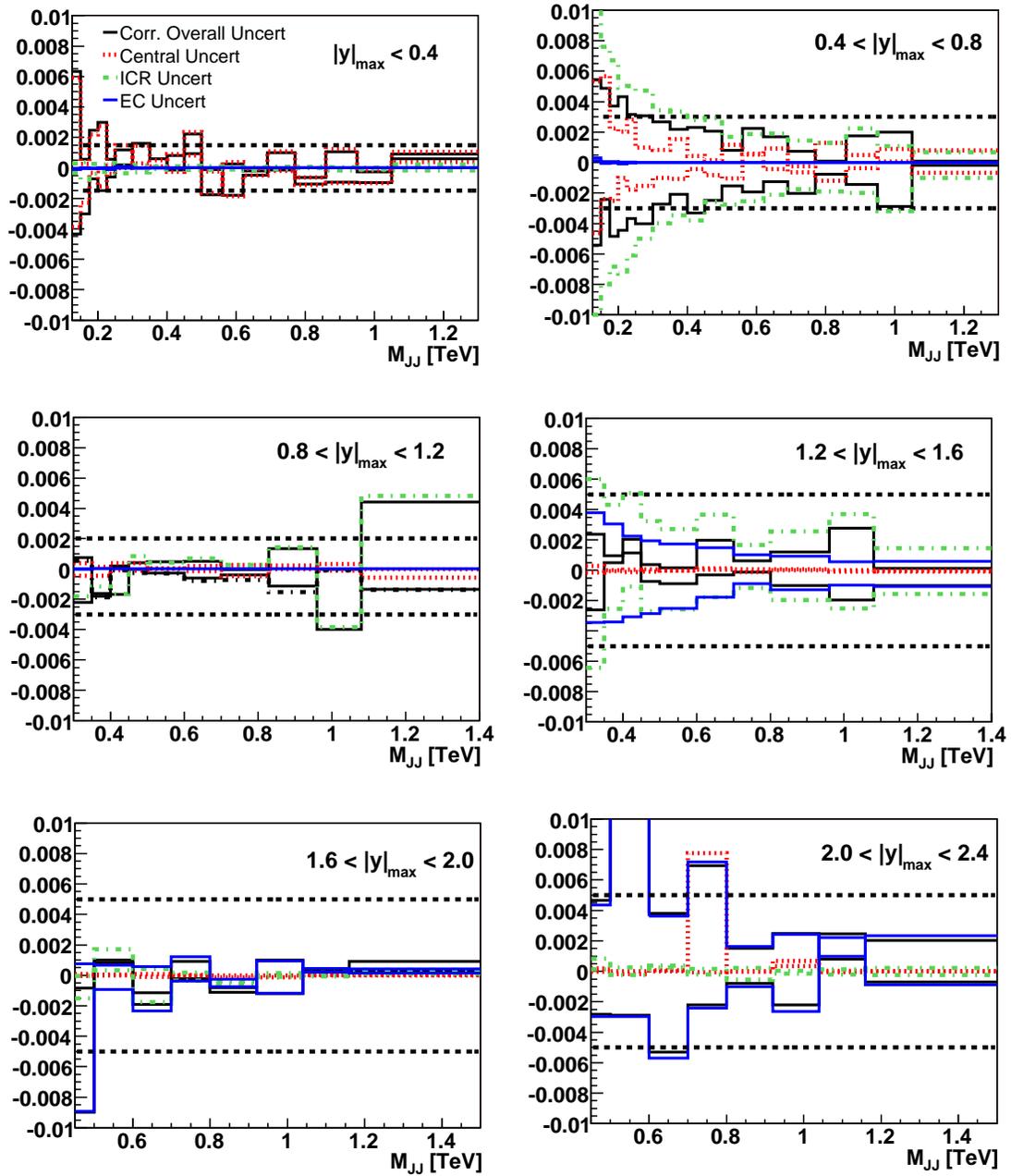
Figure 4.32: Uncertainty due to η resolution corrections. The y axis is the percentage uncertainty and the x axis is the dijet mass.
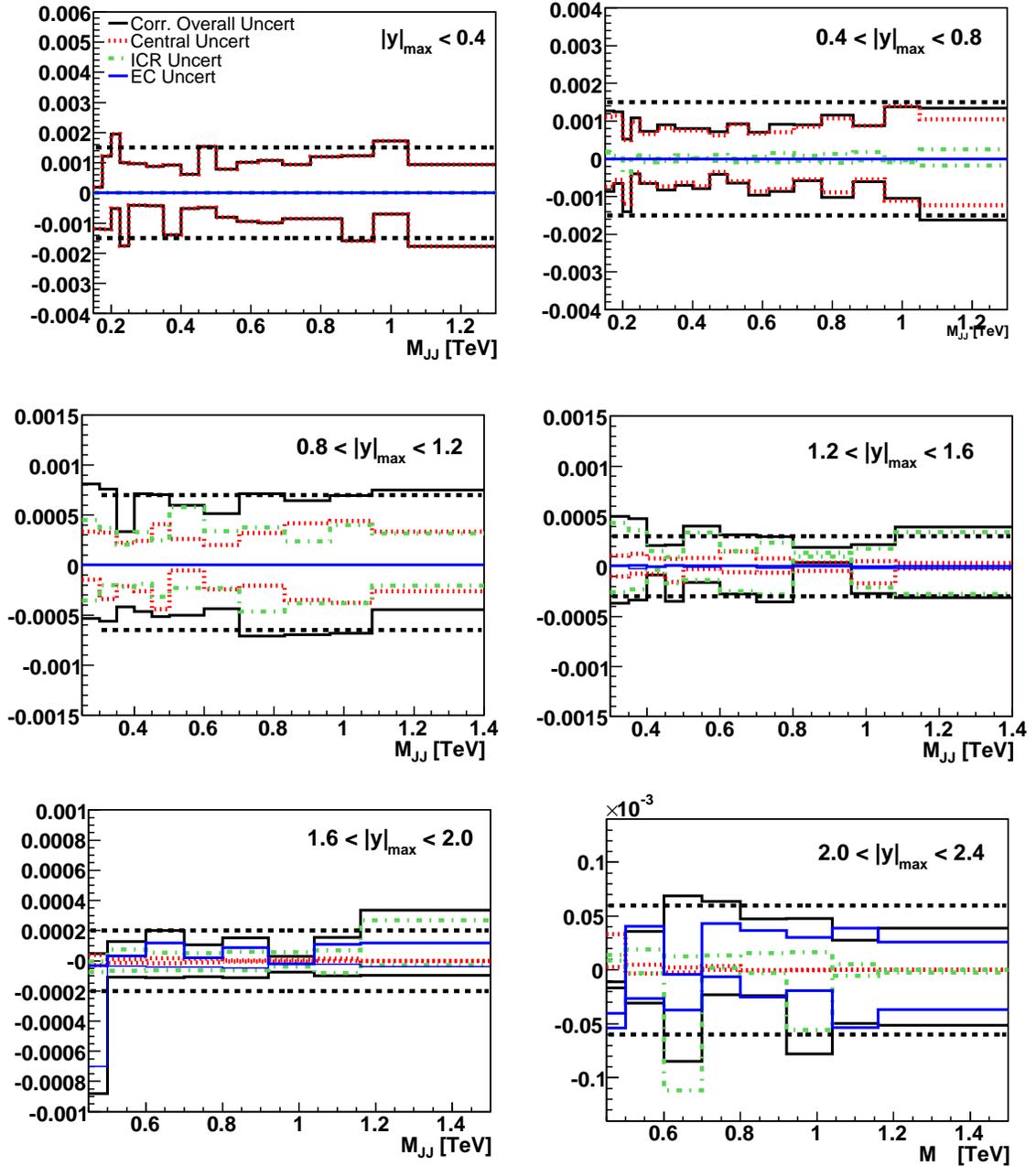
Figure 4.33: Uncertainty due to ϕ resolution correction. The y axis is the percentage uncertainty and the x axis is the dijet mass.
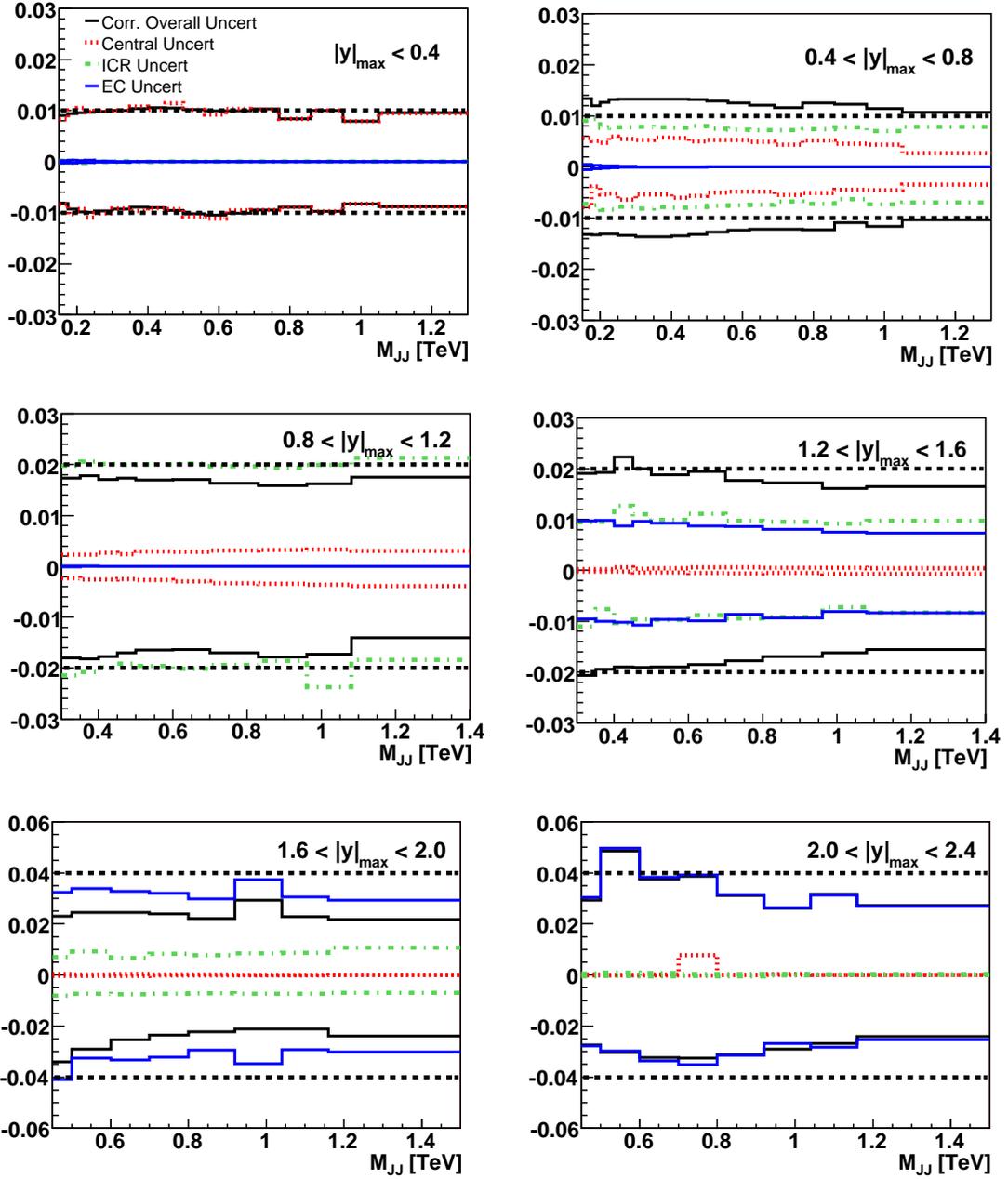
Figure 4.34: Uncertainty due to η bias. The y axis is the percentage uncertainty and the x axis is the dijet mass.
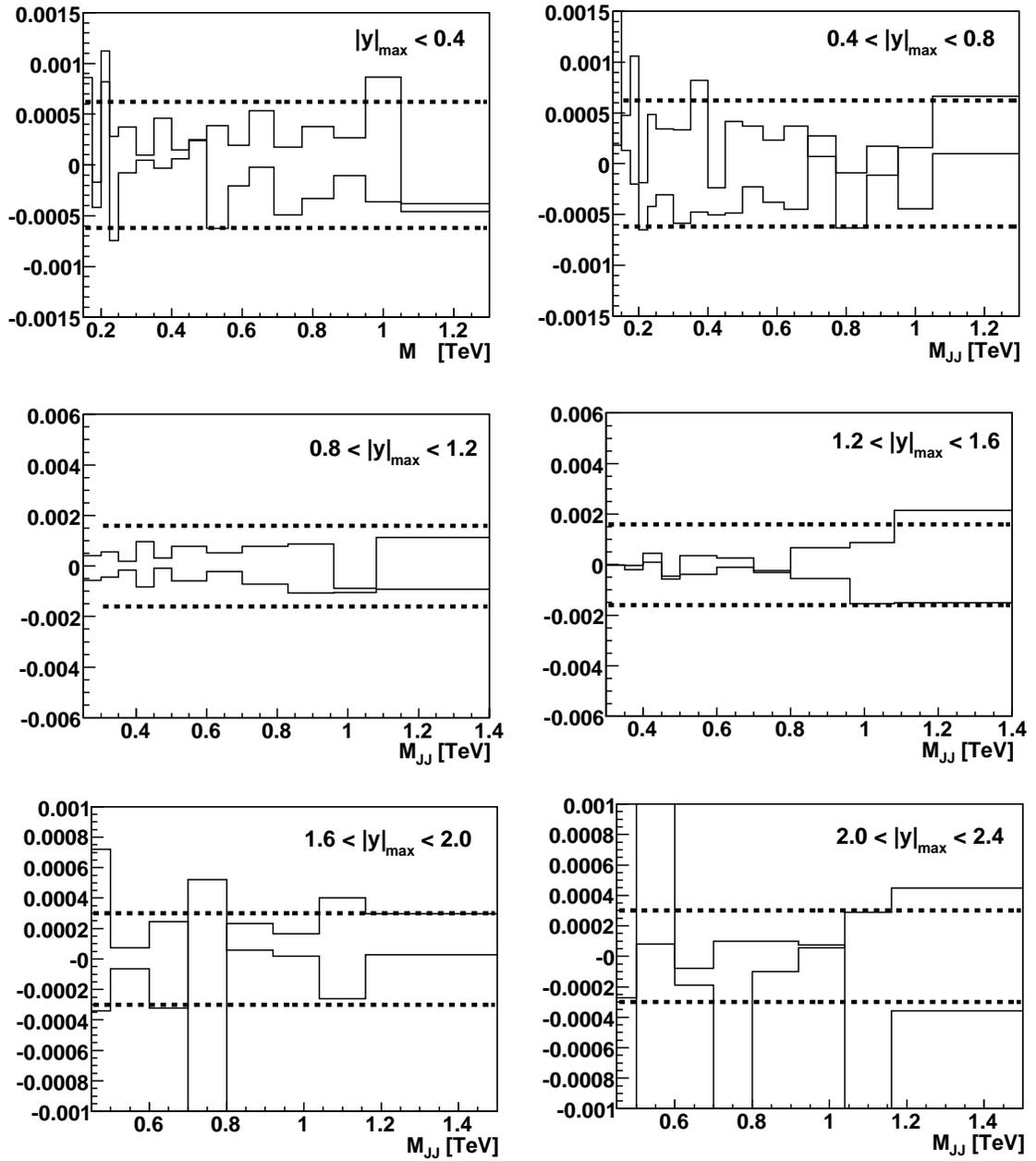
Figure 4.35: Vertex Uncertainty. The y axis is the percentage uncertainty and the x axis is the dijet mass.
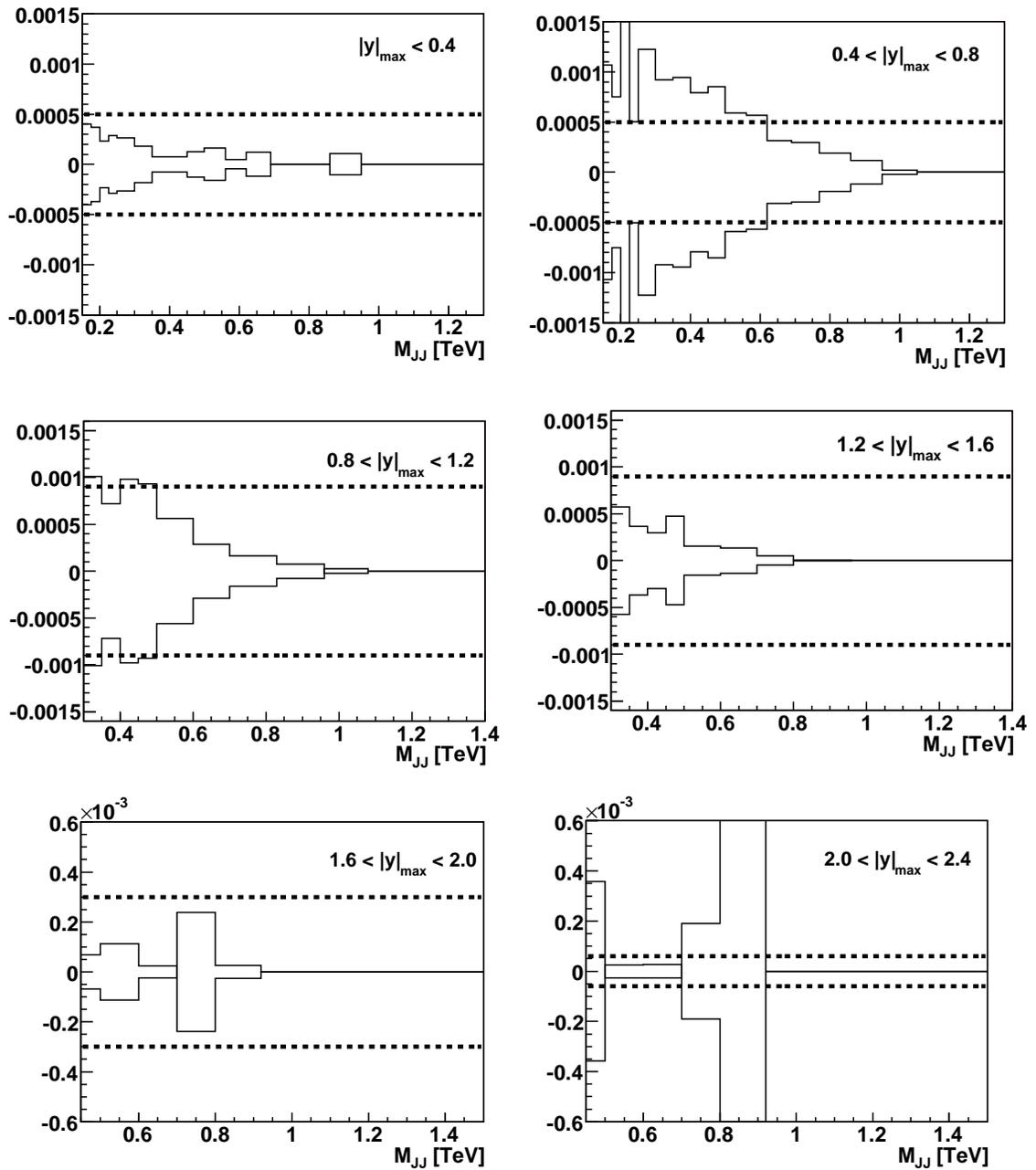
Figure 4.36: Misvertexing uncertainty. The y axis is the percentage uncertainty and the x axis is the dijet mass.
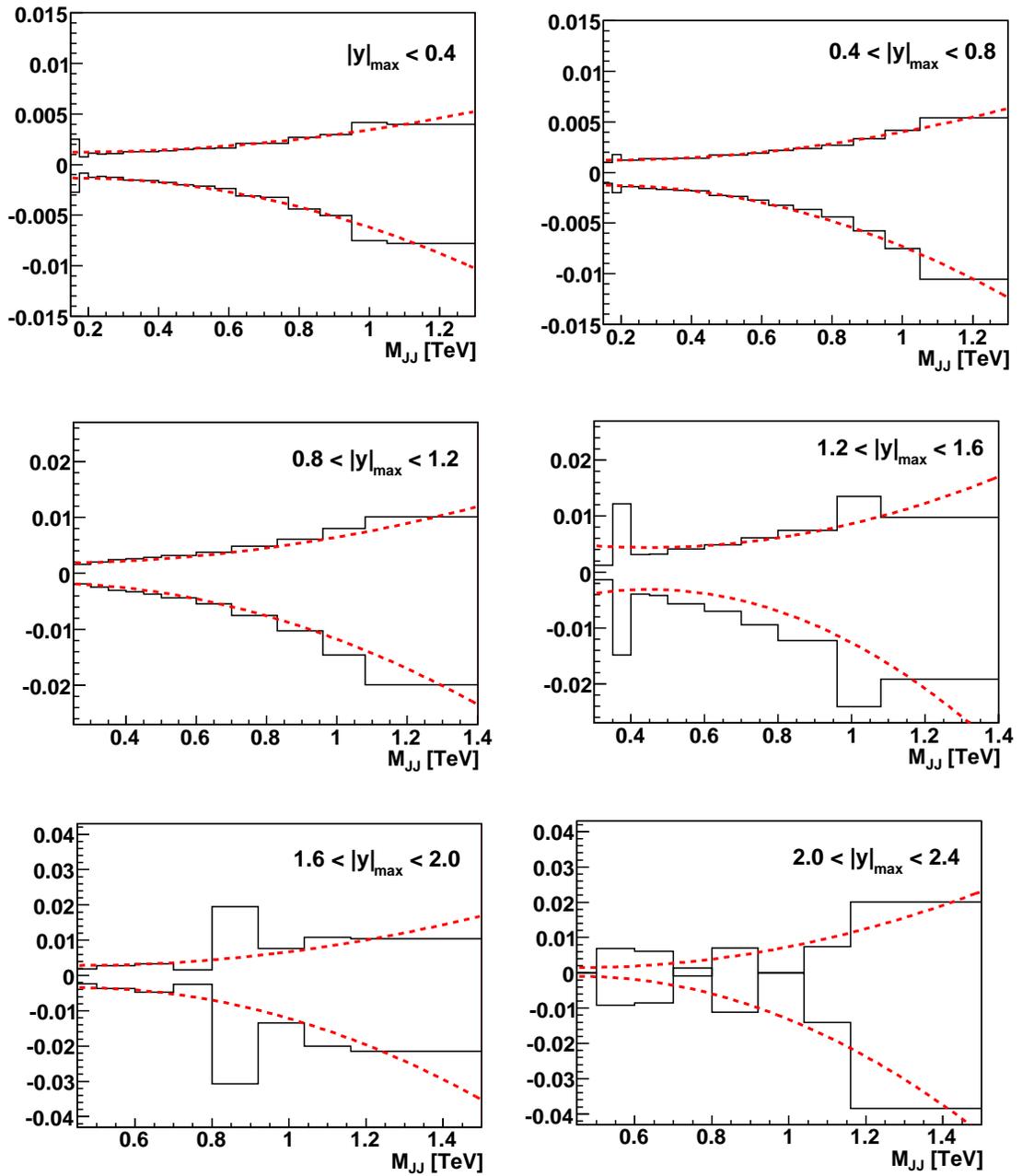
Figure 4.37: Uncertainty due to reweighting the MC, where the red dashed line shows the fit used to determine the final uncertainty. The y axis is the percentage uncertainty and the x axis is the dijet mass.
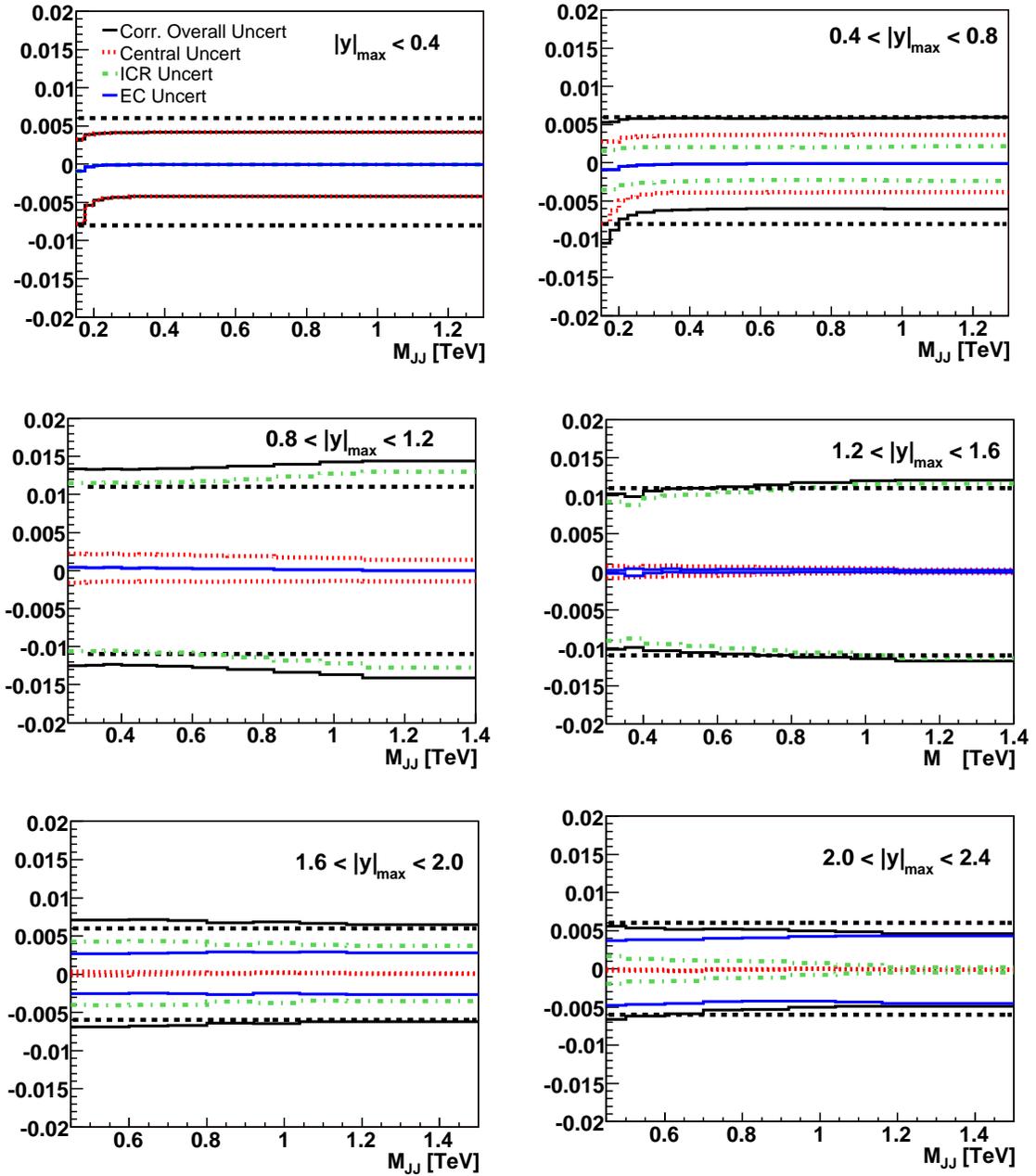
Figure 4.38: Jet ID uncertainty. The y axis is the percentage uncertainty and the x axis is the dijet mass.
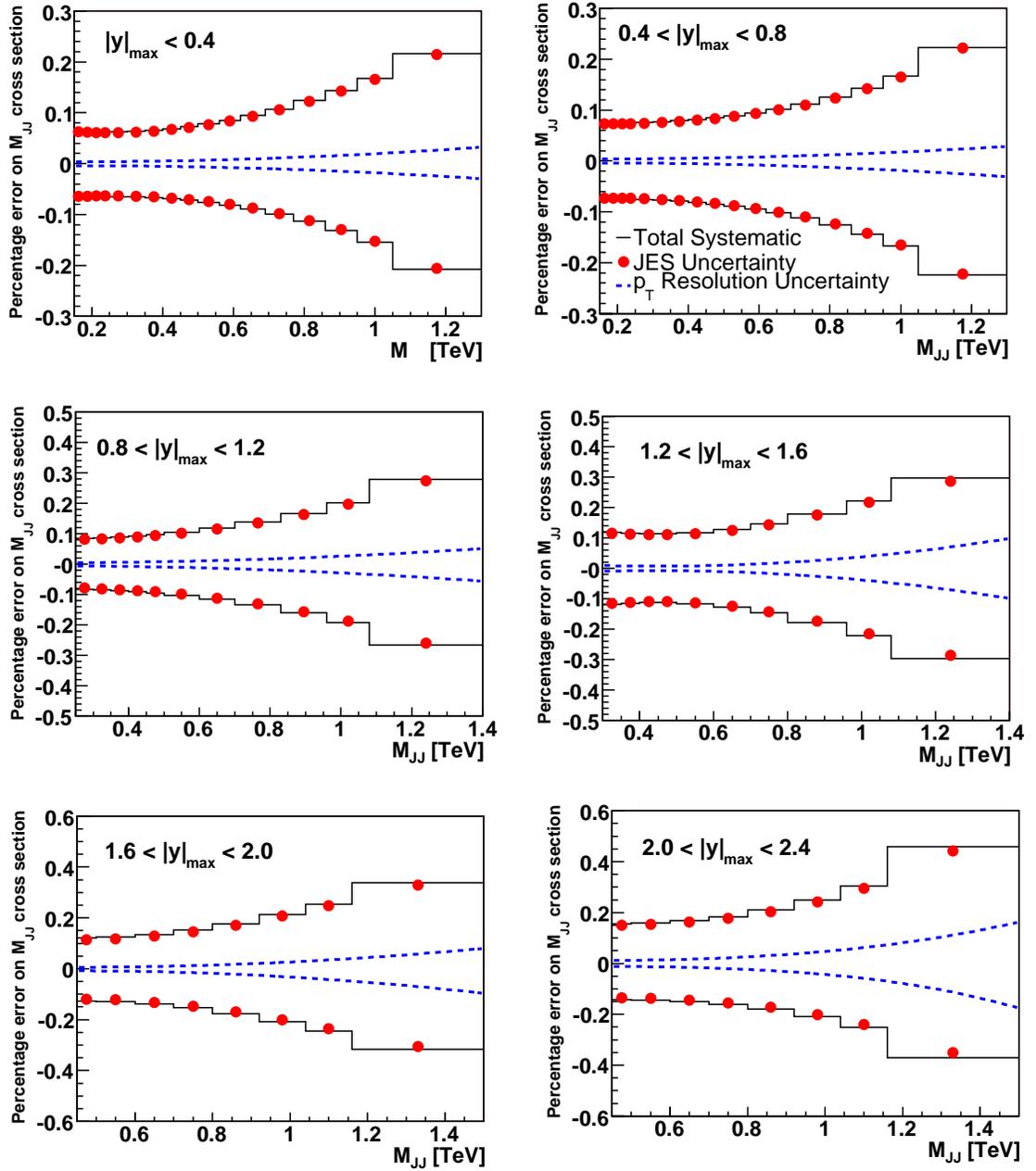
Figure 4.39: Total systematic uncertainty in all 6 rapidity region. All sources were added in quadrature. The y axis is the percentage uncertainty and the x axis is the dijet mass.

4.5 Results and Conclusions

The formula for the double differential dijet mass cross section is given in Eq. 4.28.

$$\frac{d\sigma}{dM_{JJ}d|y|_{max}} = \frac{N_{evt}}{L \cdot \sum_i \varepsilon_{vtxeff} \cdot \Delta(M_{JJ})\Delta(|y|_{max})} \, C \qquad (4.28)$$

where N_{evt} is the number of events after mass rescaling, L is the luminosity, $\sum_i \varepsilon_{vtxeff}$ is the vertex efficiency per event i, $\Delta(M_{JJ})$ is the mass bin width, $\Delta(|y|_{max})$ is the rapidity bin width and C is the correction factor determined from DØJetSim.

The fully corrected data is shown in Fig. 4.40, with MSTW NLO theory predictions overlaid. The data has been corrected for all detector effects and corrections calculated using DØJetSim, as well as dijet JES (J4S) and the vertex inefficiency. The theory has been corrected for nonperturbative effects as described in Section 2.2.5. The renormalization (μ_R) and factorization (μ_F) scales are chosen to be equal and are set at $\mu = p_{T,ave} = (p_{T,1} + p_{T,2})/2$, while $\mu = p_{T,ave}/2$ and $\mu = 2p_{T,ave}$ are used to calculate the variations up and down. The tabulated results are shown in Tables 4.2 through 4.7. The bin centers were determined using the method outlined in [51]. In this method, the truth distributions from JetSim are used to determine the bin centers. Using a grid search method and Brent's method[15], the local minimum of the truth distribution is found and set as the bin center.

The double differential dijet mass cross section was presented in this thesis and is the most precise measurement with the farthest range in rapidity to date. Good agreement between data and perturbative QCD was found in all six rapidity regions, confirming Standard Model predictions. This measurement can then be used to reduce the uncertainties on the parton distribution functions. Reducing these uncertainties will be useful for Monte Carlo simulations at the Large Hadron Collider (LHC) which is the next generation in hadron colliders. Since the systematic uncertainties on data are mostly correlated, the shape of the data divided by theory isn't allowed to change much. QCD is often a large background for other analyses, such as searches for the Higgs boson. Understanding the dijet mass helps other physicists understand their background for a more precise measurement.

No obvious new physics signals are present when the data is compared to next-to-leading order QCD theory. New physics would most likely be observed as either an excess at high mass in the central region or as a bump in the central region. Since

---

[15]Brent's method is a method used to find roots. For more information, see [52].
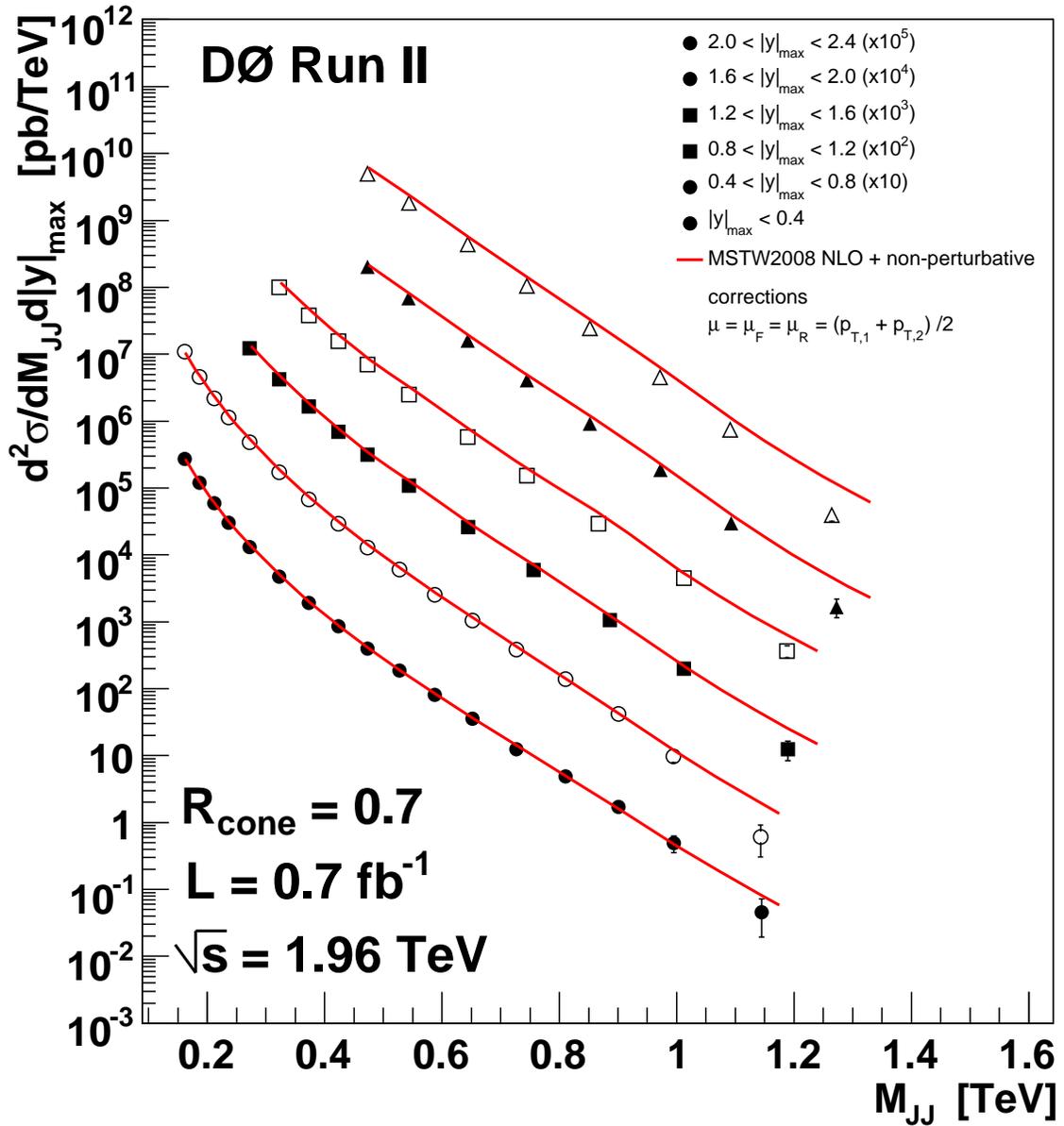
Figure 4.40: Fully corrected data with theory overlaid.

no new signals above the Standard Model were seen, limits can be set on new physics models and this work is in progress.

This analysis was done using the majority of the Run IIa data sample and could be extended by using the Run IIb data set. The DØ detector is expected to almost double the amount of data currently recorded before the end of the Tevatron run. More
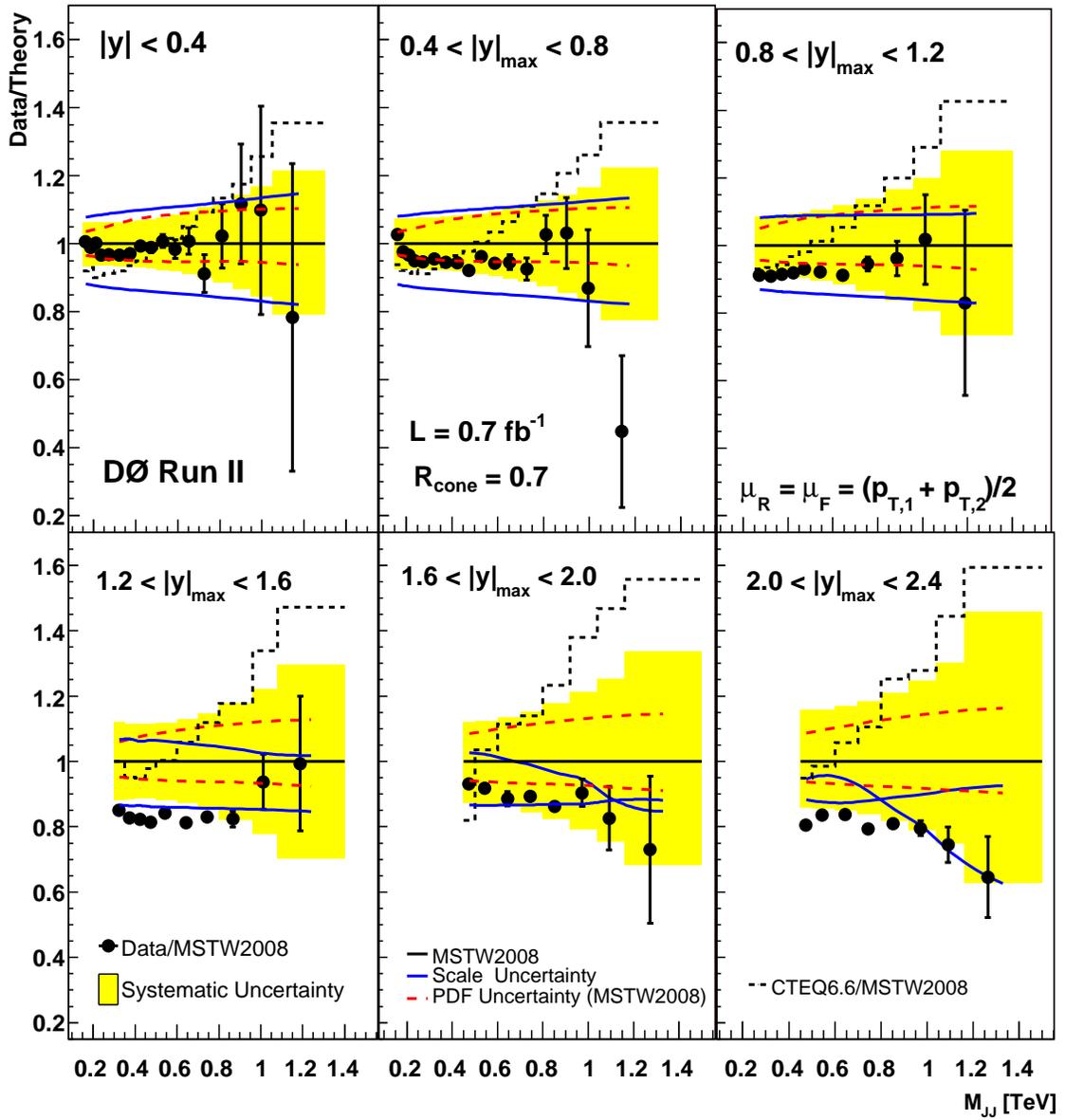
Figure 4.41: Data to theory comparison in all six $|y|_{max}$. The systematic uncertainty from data is shown as a yellow band.

statistics would improve the measurement at high masses. The largest uncertainty for this analysis is the systematic uncertainty, driven by the large uncertainty in the JES.

Mass Ranges	Bin Center	N_{events}	data pb/TeV	sys uncer+ uncer-	stat pb/TeV	Theory	Th Cor had u/e	tot
0.150 - 0.175	0.162	3835	2.737e+05	0.064 -0.066	0.016	2.721e+05	0.928 1.156	1.074
0.175 - 0.200	0.187	1738	1.217e+05	0.063 -0.065	0.024	1.215e+05	0.935 1.137	1.064
0.200- 0.225	0.212	9412	6.002e+04	0.063 -0.065	0.010	5.931e+04	0.941 1.123	1.056
0.225 -0.250	0.237	4711	3.024e+04	0.062 -0.065	0.015	3.106e+04	0.945 1.112	1.050
0.250 - 0.300	0.272	17459	1.322e+04	0.063 -0.065	0.008	1.360e+04	0.950 1.099	1.044
0.3 -0.35	0.323	6156	4.690e+03	0.064 -0.065	0.013	4.861e+03	0.955 1.087	1.037
0.35 - 0.40	0.373	17498	1.896e+03	0.066 -0.067	0.008	1.963e+03	0.959 1.078	1.032
0.40 - 0.45	0.423	7865	8.481e+02	0.068 -0.069	0.011	8.609e+02	0.961 1.071	1.029
0.45 - 0.5	0.473	5050	3.926e+02	0.072 -0.072	0.014	4.013e+02	0.963 1.065	1.026
0.50 - 0.56	0.528	2817	1.838e+02	0.078 -0.076	0.019	1.846e+02	0.965 1.061	1.024
0.56 - 0.62	0.588	1198	7.930e+01	0.085 -0.082	0.029	8.169e+01	0.966 1.057	1.021
0.62 - 0.69	0.652	613	3.504e+01	0.095 -0.089	0.040	3.524e+01	0.968 1.054	1.020
0.69 - 0.77	0.727	245	1.234e+01	0.107 -0.099	0.064	1.370e+01	0.969 1.051	1.018
0.77 - 0.86	0.811	105	4.825e+00	0.124 -0.113	0.098	4.760e+00	0.970 1.048	1.017
0.86 - 0.95	0.901	40	1.694e+00	0.144 -0.131	0.158	1.521e+00	0.971 1.046	1.015
0.95 - 1.05	0.995	10	4.951e-01	0.168 -0.154	0.316	4.485e-01	0.971 1.044	1.014
1.05 - 1.3	1.145	3	4.558e-02	0.217 -0.208	0.577	5.820e-02	0.973 1.042	1.013

Table 4.2: Dijet mass cross section for $|y| < 0.4$. Note that the number of events is for the rescaled mass.

Mass Ranges	Bin Center	N_{events}	data pb/TeV	sys uncer+ uncer-	stat uncer	Theory pb/TeV	Th Cor had u/e	tot
0.15 - 0.175	0.161	15247	1.084e+06	0.074 -0.074	0.008	1.056e+06	0.945 1.128	1.065
0.175- 0.2	0.187	6546	4.670e+05	0.074 -0.074	0.012	4.684e+05	0.949 1.113	1.056
0.2 - 0.225	0.212	35063	2.237e+05	0.074 -0.074	0.005	2.276e+05	0.954 1.098	1.047
0.225 - 0.25	0.237	17856	1.141e+05	0.075 -0.075	0.007	1.185e+05	0.957 1.088	1.040
0.25 - 0.3	0.272	65031	4.915e+04	0.076 -0.075	0.004	5.141e+04	0.959 1.080	1.036
0.3 - 0.35	0.323	22959	1.736e+04	0.077 -0.077	0.007	1.809e+04	0.961 1.073	1.032
0.35 - 0.4	0.373	62843	6.777e+03	0.079 -0.079	0.004	7.160e+03	0.963 1.068	1.028
0.4 - 0.45	0.423	26706	2.895e+03	0.082 -0.082	0.006	3.073e+03	0.964 1.063	1.026
0.45 - 0.5	0.473	16561	1.284e+03	0.085 -0.085	0.008	1.401e+03	0.965 1.059	1.023
0.5 - 0.56	0.528	9194	5.972e+02	0.089 -0.089	0.010	6.253e+02	0.966 1.056	1.021
0.56 - 0.62	0.588	3837	2.499e+02	0.094 -0.094	0.016	2.676e+02	0.968 1.053	1.019
0.62 - 0.69	0.652	1836	1.040e+02	0.102 -0.102	0.023	1.110e+02	0.968 1.048	1.017
0.69 - 0.77	0.726	764	3.781e+01	0.112 -0.112	0.036	4.119e+01	0.969 1.062	1.016
0.77 - 0.86	0.811	315	1.382e+01	0.125 -0.125	0.056	1.353e+01	0.970 1.045	1.014
0.86 - 0.95	0.901	88	4.204e+00	0.143 -0.144	0.107	4.072e+00	0.971 1.041	1.012
0.95 - 1.050	0.995	24	9.897e-01	0.167 -0.167	0.204	1.128e+00	0.971 1.040	1.011
1.050 - 1.3	1.143	4	6.078e-02	0.223 -0.224	0.500	1.359e-01	0.971 1.040	1.011

Table 4.3: Dijet mass cross section for $0.4 < |y|_{max} < 0.8$. Note that the number of events is for the rescaled mass.

Mass Ranges	Bin Center	N_{events}	data pb/TeV	sys uncer+ uncer-	stat uncert	Theory pb/TeV	Th Cor had u/e	tot
0.25 - 0.3	0.272	38097	1.213e+05	0.085 -0.082	0.005	1.332e+05	0.946 1.111	1.067
0.3 -0.35	0.323	13154	4.199e+04	0.087 -0.084	0.009	4.617e+04	0.950 1.101	1.058
0.35 - 0.4	0.373	5125	1.636e+04	0.090 -0.087	0.014	1.795e+04	0.953 1.092	1.052
0.4 - 0.45	0.423	9081	6.877e+03	0.093 -0.090	0.010	7.549e+03	0.955 1.086	1.048
0.45 - 0.5	0.473	4109	3.110e+03	0.098 -0.095	0.016	3.381e+03	0.957 1.079	1.042
0.5 - 0.6	0.544	19803	1.068e+03	0.106 -0.102	0.007	1.173e+03	0.960 1.073	1.037
0.6 - 0.7	0.644	4717	2.565e+02	0.119 -0.115	0.015	2.837e+02	0.961 1.067	1.033
0.7 - 0.83	0.756	1986	5.970e+01	0.139 -0.133	0.022	6.302e+01	0.963 1.063	1.030
0.83 - 0.96	0.886	355	1.087e+01	0.167 -0.160	0.053	1.104e+01	0.964 1.061	1.023
0.96 - 1.080	1.013	64	2.095e+00	0.201 -0.192	0.125	1.949e+00	0.965 1.057	1.025
1.080 - 1.4	1.189	12	1.428e-01	0.279 -0.266	0.289	1.502e-01	0.966 1.054	1.022

Table 4.4: Dijet mass cross section for $0.8 < |y|_{max} < 1.2$. Note that the number of events is for the rescaled mass.

Mass Ranges	Bin Center	N_{events}	data pb/TeV	sys uncer+ uncer-	stat uncert	Theory pb/TeV	Th Cor had u/e	tot
0.3 - 0.35	0.323	31577	1.003e+05	0.119 -0.119	0.006	1.180e+05	0.950 1.144	1.081
0.35 - 0.4	0.373	11939	3.794e+04	0.115 -0.115	0.009	4.585e+04	0.951 1.219	1.075
0.4 - 0.45	0.423	5064	1.609e+04	0.113 -0.113	0.014	1.908e+04	0.952 1.126	1.069
0.45 - 0.5	0.473	2242	7.135e+03	0.113 -0.113	0.021	8.590e+03	0.954 1.110	1.064
0.5 - 0.6	0.544	6773	2.549e+03	0.117 -0.117	0.012	2.969e+03	0.955 1.129	1.058
0.6 - 0.7	0.644	11123	5.957e+02	0.128 -0.128	0.009	7.152e+02	0.956 1.104	1.054
0.7 - 0.8	0.745	2963	1.586e+02	0.146 -0.146	0.018	1.836e+02	0.956 1.099	1.051
0.8 - 0.96	0.867	1315	3.164e+01	0.178 -0.178	0.028	3.566e+01	0.957 1.096	1.048
0.96 - 1.080	1.012	159	5.106e+00	0.222 -0.222	0.079	4.780e+00	0.958 1.092	1.045
1.080 - 1.4	1.188	40	4.767e-01	0.296 -0.297	0.158	3.665e-01	0.958 1.087	1.042

Table 4.5: Dijet mass cross section for $1.2 < |y|_{max} < 1.6$. Note that the number of events is for the rescaled mass.

Mass Ranges	Bin Center	N_{events}	data pb/TeV	sys uncer+ uncer-	stat uncert	Theory pb/TeV	Th Cor had u/e	tot
0.45 - 0.5	0.473	10350	2.012e+04	0.122 -0.126	0.010	2.160e+04	0.939 1.156	1.037
0.5 - 0.6	0.543	7095	6.881e+03	0.125 -0.129	0.012	7.492e+03	0.940 1.145	1.033
0.6 - 0.7	0.644	1645	1.592e+03	0.135 -0.139	0.025	1.797e+03	0.940 1.136	1.030
0.7 - 0.8	0.744	8875	4.104e+02	0.152 -0.154	0.011	4.589e+02	0.940 1.130	1.028
0.8 - 0.92	0.852	2435	9.319e+01	0.177 -0.177	0.020	1.070e+02	0.940 1.124	1.026
0.92 - 1.040	0.972	502	1.929e+01	0.212 -0.208	0.045	2.073e+01	0.941 1.118	1.024
1.040 - 1.160	1.092	83	3.148e+00	0.253 -0.244	0.110	3.599e+00	0.940 1.111	1.022
1.160 - 1.5	1.273	16	1.919e-01	0.338 -0.317	0.250	2.279e-01	0.941 1.109	1.021

Table 4.6: Dijet mass cross section for $1.6 < |y|_{max} < 2.0$. Note that the number of events is for the rescaled mass.

Mass Ranges	Bin Center	N$_{events}$	Data pb/TeV	sys uncer+ uncer-	stat uncert	Theory pb/TeV	Th Cor had u/e	tot
0.45 - 0.5	0.473	25572	4.952e+04	0.156 -0.142	0.006	6.147e+04	0.901 1.230	1.151
0.5 - 0.6	0.544	18812	1.810e+04	0.159 -0.144	0.007	2.166e+04	0.902 1.222	1.138
0.6 - 0.7	0.644	4591	4.366e+03	0.168 -0.150	0.015	5.249e+03	0.902 1.215	1.127
0.7 - 0.8	0.744	22570	1.024e+03	0.184 -0.161	0.007	1.312e+03	0.903 1.209	1.119
0.8 - 0.92	0.852	6342	2.370e+02	0.210 -0.180	0.013	2.988e+02	0.903 1.205	1.113
0.92 - 1.040	0.972	1199	4.438e+01	0.250 -0.209	0.029	5.677e+01	0.903 1.200	1.105
1.040 - 1.160	1.091	203	7.323e+00	0.304 -0.250	0.070	9.878e+00	0.904 1.195	1.098
1.160 - 1.5	1.264	37	4.115e-01	0.458 -0.371	0.164	6.098e-01	0.904 1.193	1.095

Table 4.7: Dijet mass cross section for $2.0 < |y|_{max} < 2.4$. Note that the number of events is for the rescaled mass.

Reference List

[1] B. Abbott, et. al , "Dijet Mass Spectrum and search for Quark Compositeness in p$\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", Phys. Rev. Letters 82 2457 (1999).

[2] A. Kupco, "Measurement of the Dijet Mass Cross Section", DØ Note 4145 (2003).

[3] B. Abbott, et. al, "High-p_T Jets in $p\bar{p}$ Collisions at $\sqrt{s} = 630$ GeV and 1800 GeV", Phys. Rev. D 64, 032003 (2001) [69 pages].

[4] S. L. Glashow, Nuclear Physics 22, 579 (1961).
S. Weinberg, Phys. Rev. Lett. 19, 1264 - 1266 (1967).
A. Salam in *Elementary Particle Theory* ed. by N. Svartholm.

[5] D. J. Gross and F. Wilczek, Asymptotically Free Gauge Theories. I, Phys. Rev. D8 3633 (1973).
H. D. Politzer, Asymptotic Freedom: An Approach to Strong Interactions, Phys. Rep. 14 129 (1974).

[6] Y. Fukuda et. al, "Evidence for oscillation of atmospheric neutrinos", Phys.Rev.Lett.81:1562-1567 (1998).

[7] C. Amsler et al., Physics Letters B667, 1 (2008) , and 2009 partial update for the 2010 edition.

[8] O. W. Greenburg, "The Parton Model", arXiv:0805.2588v3 [hep-ph] (2008).

[9] M. Zielinski, "Jets", talk given at the CTEQ summer school (http://www.phys.psu.edu/ cteq/schools/summer07) (2007).

[10] M. Stratmann, "Lectures on Perturbative QCD", PHENIX Spin Fest at RIKEN Wako, (2005).

[11] S. Kluth,"α_s from LEP", arXiv:0709.0173v2 [hep-ex] (2007).

[12] G. P. Lepage, "Lattice QCD for Novices", Proceedings of HUGS 98, edited by J.L. Goity, World Scientific (2000)

[13] W.K. Tung, "Perturbative QCD and the Parton Structure of the Nucleon", CTEQ summer school (http://www.phys.psu.edu/ cteq/schools/summer06) (2006).

[14] J. Chýla, "Quarks, Partons and Quantum Chromodynamics", Study text. Available at http://www-hep.fzu.cz/Theory/notes/text.pdf (2003).

[15] S. Dawson, "Practical NLO Calculations", talk given at the CTEQ summer school (http://www.phys.psu.edu/ cteq/schools/summer07) (2007).

[16] M. Voutilainen, "Mearsument of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", Ph. D. thesis (2008).

[17] R. K. Ellis, "Perturbative QCD: From the Tevatron to the LHC", talk given at the Fermilab User's meeting, (2009).

[18] J.F. Owens, "Parton Distribution Functions and Global Fitting", talk given at the CTEQ summer school (http://www.phys.psu.edu/ cteq/schools/summer07) (2007).

[19] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438.

[20] G.Altarelli and G. Parisi, Nucl. Phys. B126 (1977) 298.

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

[22] J. Pumplin *et al.*, JHEP **0207** 012 (2002) and D. Stump *et al.*, JHEP **0310** 046 (2003).

[23] A.D. Martin, W.J. Stirling, R.S Thorne, G. Watt, arXiv:0901.0002v2, [hep-ph] (2009), submitted to Eur. Phys. Jour. C.

[24] H. Jung et al., in Proceedings of the workshop: HERA and the LHC workshop series on the implications of HERA for LHC physics, arXiv:0903.3861v2 [hep-ph] (2009).

[25] DØ Collaboration, "Limits on Quark Compositeness from High Energy Jets in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV", Phys. Rev. D 62, 031101 (2000).

[26] G.C. Blazey et al., in Proceedings of the Workshop: QCD and Weak Boson Physics in Run II, edited by U. Baur, R.K. Ellis, and D. Zeppenfeld, Fermilab-Pub-00/297 (2000).

[27] J.E. Huth et al. in Proceedings of Research Directions For The Decade: Snowmass 1990, July, 1990, 30,edited by E.L. Berger (World Scientific, Singapore,1992) p. 134.

[28] M.A. Dobbs, "Les Houches Guidebook to Monet Carlo Event Generators for Hadron Collider Physis", arXiv:hep-ph/0403045v2 (2004).

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

[30] M. Albrow et. al (TeV4LHC QCD Working Group), Tevatron-for-LHC Conference Report of the QCD Working Group, Fermilab-Conf-06-359 (2006).

[31] CDF Collaboration "Search for new particles decaying to dijets in $p\bar{p}$ collisions at $\sqrt{s} - 1.96$ TeV", arXiv:0812.4036v1 [hep-ex] (2008).

[32] Accelerator Concepts Rookie Book, (http://www-bdnew.fnal.gov/operations), (2009).

[33] T. Ferbel, "A Brief Description of the DØ Detector in Run II", (2005) (unpublished).

[34] DØ Collaboration, "The Upgraded DØ Detector", Nuclear Instruments and Methods in Physics Research A 565, pgs 463-537 (2006).

[35] S. Abachi, et al., "The DØ Detector", Nuclear Instruments and Methods in Physics Research A 338 (1994).

[36] T. Ferbel "Experimental Techniques in high-energy nuclear and particle physics"

[37] M. Wobisch, "First Measurement of the Dijet Angular Distributions in the TeV Regime", DØ Note 5616 (2008).

[38] M. Voutilainen, "Measurement of the inclusive jet p_T spectra in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV" , DØ Note 5551 (2008).

[39] ROOT, http://root.cern.ch/drupal/ (2009).

[40] M. Voutilainen, "Jet Trigger Efficiencies for RunIIa", DØ Note 5549 (2006).

[41] H. Schellman, "The longitudinal shape of the luminous region at DØ; DØ Note 5142 (2006).

[42] JES Group "Jet Energy Scale Determination at the DØ Run II (final p17 version)", DØ Note 5382 (2007).

[43] M. Voutilainen, "Jet Four-vector Scale Determination for the Dijets in DØ Run IIa ", DØ Note 5550 (2007) (to be submitted to Nuclear Instruments and Methods in Physics Research A).

[44] G. Zech, "Comparing Statistical Data to Monte Carlo Simulation - Parameter Fitting and Unfolding", DESY-95-113 (1995).

[45] Public website for Jet Energy Scale, (http://www-d0.fnal.gov/phys_id/jes/public_RunIIa) .

[46] A.Harel, "Jet ID Optimization", DØ Note 4919 v1.1 (2005).

[47] M. Wobisch, "DØJetSim - A Parameterized Detector Simulation for Jet Physics", DØ Note 5703 (2008).

[48] M. Rominsky "Jet p_T Resolutions using η_{det}", DØ Note 5908 (2008).

118

[49] Z. Hubacek, "Jet η and ϕ in p17 QCD MC sample", DØ internal document (2009).

[50] T. Andeen, et. al., FERMILAB-TM-2365 (2007).

[51] G. Lafferty, T. Wyatt "Where to stick your data points: The treatment of measurements within wide bins", Nucl. Instr. and Meth. A 355 (1995) 541-547 (1994).

[52] R.P. Brent, "Algorithms for Minimization without Derivatives", Chapter 4. Prentice-Hall, Englewood Cliffs, NJ. (1974).

Appendix A

Common Symbols

The following list is some common symbols used through the thesis. Note that in particle physics the speed of light c is set to 1 in calculations. This means that energy, momentum and mass are all in units of electron volts (eV), which is the kinetic energy of an electron accelerated by a 1 volt of potential.

The DØ detector uses a right-handed coordinate system with positive x pointing into the ring, positive y pointing straight up and positive z pointing along the direction of the proton beam.

polar angle	θ	$\arctan(\sqrt{x^2 + y^2}/z)$
azimuth angle	ϕ	$\arctan(y/x)$
rapidity	y	$1/2 \ln(E + p_z)/(E - p_z)$
transverse momentum	p_T	$\sqrt{p_x^2 + p_y^2}$
dijet mass	M_{JJ}	
strong coupling constant	α_s	
pseudorapitidy	η	$-\ln \tan \theta/2$
detector η	η_{det}	same as η, but measured from the center the detector
transverse energy	E_T	$E/\cosh(\eta)$
missing transverse energy	$\not{E}_T$	
center of mass energy	$\sqrt{s}$	
factorization scale	μ_F	Scale that characterizes infrared singularities
renormalization scale	μ_R	Scale that characterizes uv singularities
fractional momentum of partons	x	
Characteristic scale of interactions	Q	
Radius of cone	ΔR	$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$
barn	b	Unit of cross section; 1 barn $= 10^{-28} m^2$
inverse femtobarn	fb^{-1}	Commonly used in this analysis, the inverse of area

Appendix B

Common Abbreviations and Terms

QCD	Quantum Chromodynamics
PDF	parton distribution functions
SM	Standard Model
MC	Monte Carlo
reco MC	MC that has been reconstructed in the DØ environment so that it mimics our detector.
particle MC	MC that only includes phenomenological effects.
CC1	$|y|_{max} < 0.4$
CC2	$0.4 < |y|_{max} < 0.8$
IC1	$0.8 < |y|_{max} < 1.2$
IC2	$1.2 < |y|_{max} < 1.6$
EC1	$1.6 < |y|_{max} < 2.0$
EC2	$2.0 < |y|_{max} < 2.4$

Appendix C

Properties of Jet Triggers

Table C.1: Trigger definitions used in the analysis. The L2 conditions were not applied all the time during data taking. The L2 size was also changed from 2×2 to 5×5.

Trigger/version	L1 (GeV)	L2 (GeV)	L3 (GeV)
JT_25TT_NG/1–2	2 towers > 3	none	1 jet $E_T > 25$
JT_45TT/1–10	2 towers > 5	none	1 jet $E_T > 45$
JT_65TT/1–13	3 towers > 5	1 jet $E_T > 20$	1 jet $E_T > 65$
JT_95TT/15–16	3 towers > 5	1 jet $E_T > 50$	1 jet $E_T > 95$
JT_95TT/13–14	4 towers > 5	1 jet $E_T > 50$	1 jet $E_T > 95$
JT_95TT/6–12	4 towers > 5	1 jet $E_T > 30$	1 jet $E_T > 95$
JT_95TT/1–5	4 towers > 5	none	1 jet $E_T > 95$
JT_125TT/14–15	3 towers > 5	1 jet $E_T > 60$	1 jet $E_T > 125$
JT_125TT/12–13	4 towers > 5	1 jet $E_T > 60$	1 jet $E_T > 125$
JT_125TT/1–11	4 towers > 7	none	1 jet $E_T > 125$
JT_L3M250/11–16	3 towers > 5	$M_{inv} > 100$	2 jets $E_T > 8$ and $M_{inv} > 250$
JT_L3M430/11–16	3 towers > 5	$M_{inv} > 250$	2 jets $E_T > 8$ and $M_{inv} > 430$

Appendix D

Trigger Turn-on Curves

This section includes both a plot of the trigger turn on curves as well as a the information in table format.

	Trigger	Efficiency	Slope	Turn-on	90%	98%	99%
$|y| < 0.4$	JT_25						
$|y| < 0.4$	JT_45	0.993	0.01367	0.1058	0.136	0.159	0.168
$|y| < 0.4$	JT_65	0.987	.00832	0.159	0.177	0.191	0.197
$|y| < 0.4$	JT_95	0.991	0.0124	0.222	0.249	0.27	0.279
$|y| < 0.4$	JT_125	0.99	0.0155	0.284	0.318	0.344	0.355
$0.4 < |y|_{max} < 0.8$	JT_25						
$0.4 < |y|_{max} < 0.8$	JT_45	0.9972	0.0200	0.118	0.161	0.195	0.20
$0.4 < |y|_{max} < 0.8$	JT_65	0.9942	0.015	0.173	0.205	0.231	0.241
$0.4 < |y|_{max} < 0.8$	JT_95	1.001	0.019	0.239	0.281	0.313	0.326
$0.4 < |y|_{max} < 0.8$	JT_125	1.006	0.025	0.304	0.358	0.400	0.417
$0.8 < |y|_{max} < 1.2$	JT_45	1.015	0.021	0.143	0.189	0.224	0.239
$0.8 < |y|_{max} < 1.2$	JT_65	1.011	0.0318	0.207	0.277	0.331	0.353
$0.8 < |y|_{max} < 1.2$	JT_95	0.9913	0.040	0.284	0.372	0.440	0.468
$0.8 < |y|_{max} < 1.2$	JT_125	1.012	0.0572	0.357	0.483	0.579	0.620
$1.2 < |y|_{max} < 1.6$	JT_45	1.105	0.0731	0.175	0.336	0.459	0.511
$1.2 < |y|_{max} < 1.6$	JT_65	0.9964	0.0381	0.297	0.381	0.445	0.472
$1.2 < |y|_{max} < 1.6$	JT_95	0.974	0.0592	0.403	0.533	0.633	0.675
$1.2 < |y|_{max} < 1.6$	JT_125	1.057	0.109	0.510	0.730	0.900	0.97
$1.6 < |y|_{max} < 2.0$	JT_L3M250	0.9541	0.0104	0.3346	0.356	0.373	0.381
$1.6 < |y|_{max} < 2.0$	JT_L3M430	0.962	0.017	0.535	0.572	0.601	0.613
$2.0 < |y|_{max} < 2.4$	JT_L3M250	0.9902	0.0183	0.335	0.374	0.405	0.417
$2.0 < |y|_{max} < 2.4$	JT_L3M430	0.997	0.084	0.424	0.608	0.750	0.809

Table D.1: Trigger information from turn-on fits for all 6 $|y|_{max}$ regions.
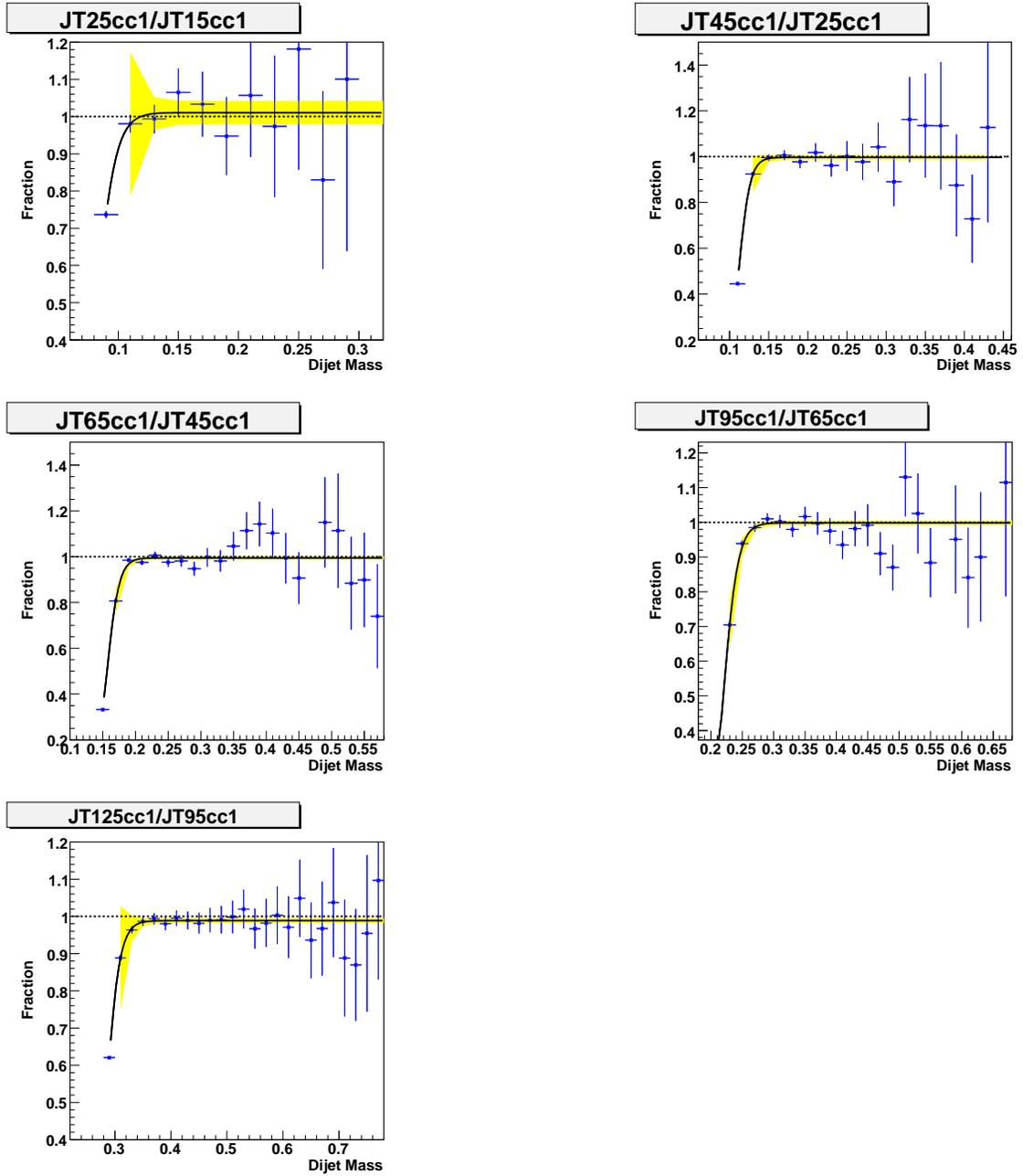
Figure D.1: Trigger turn-on curves for all single jet triggers in the first central region ($|y|_{max} < 0.4$). The "knee" region has been fitted.
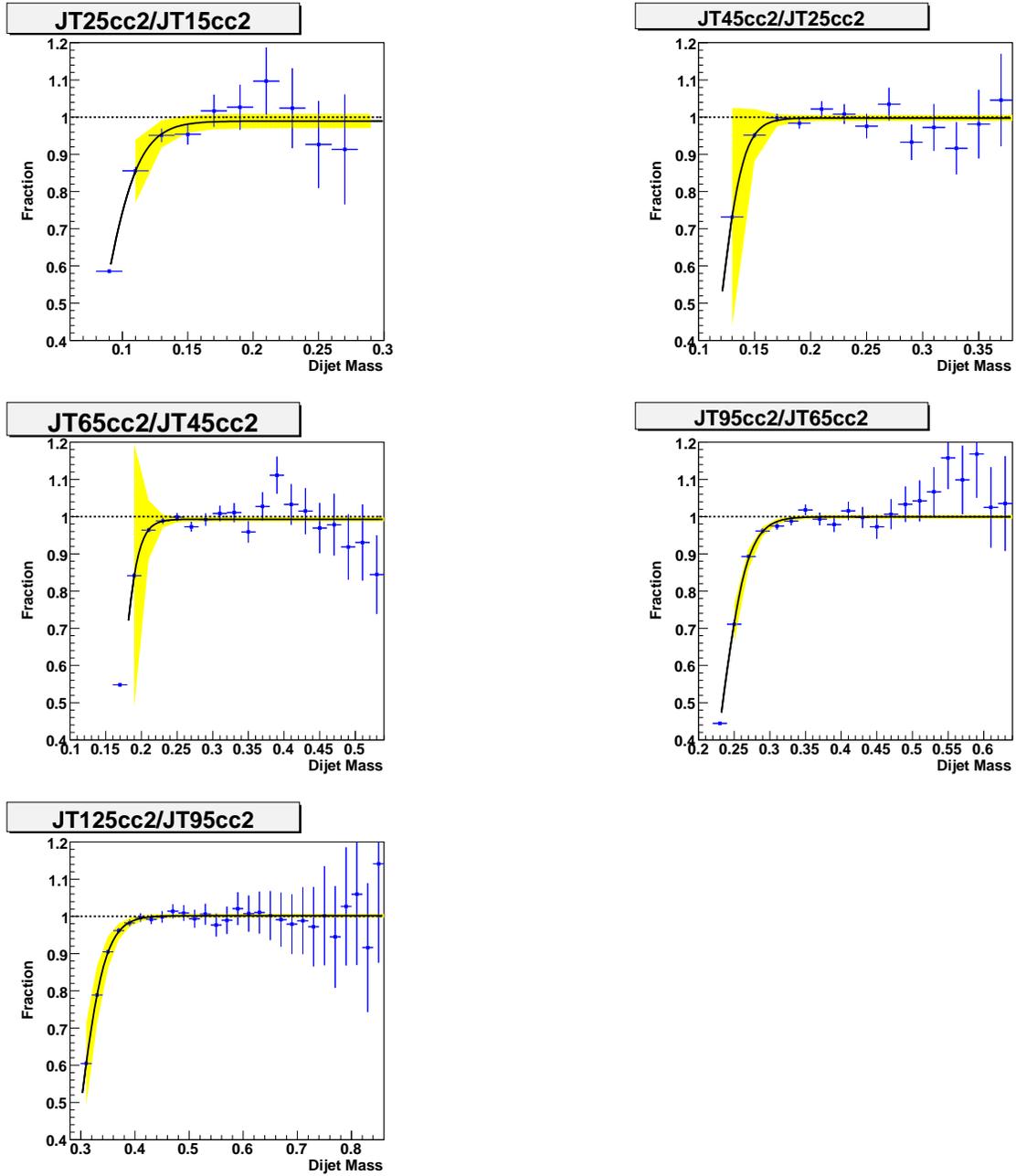
Figure D.2: Trigger turn-on curves for all single jet triggers in the second central region $0.4 < |y|_{max} < 0.8$. The knee region is fitted.
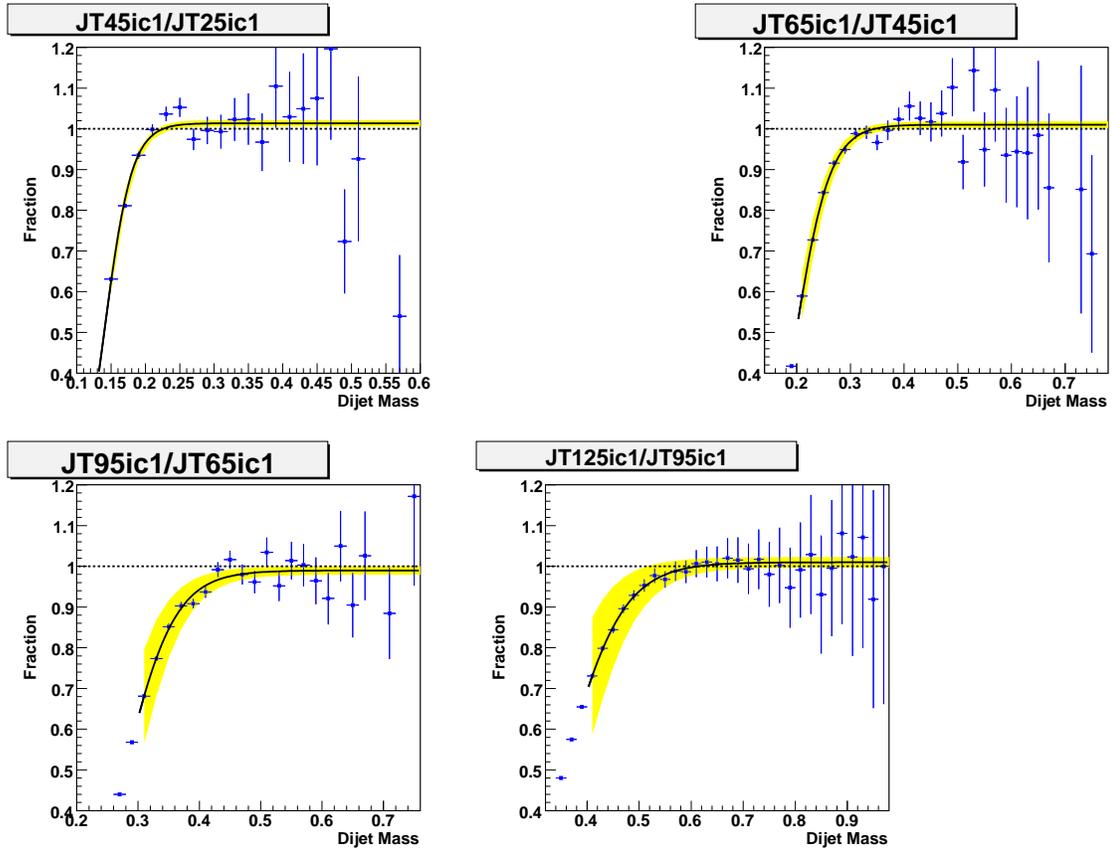
Figure D.3: Trigger turn-on curves for all single jet triggers in the first ICR region $0.8 < |y|_{max} < 1.2$. The knee region is fitted.
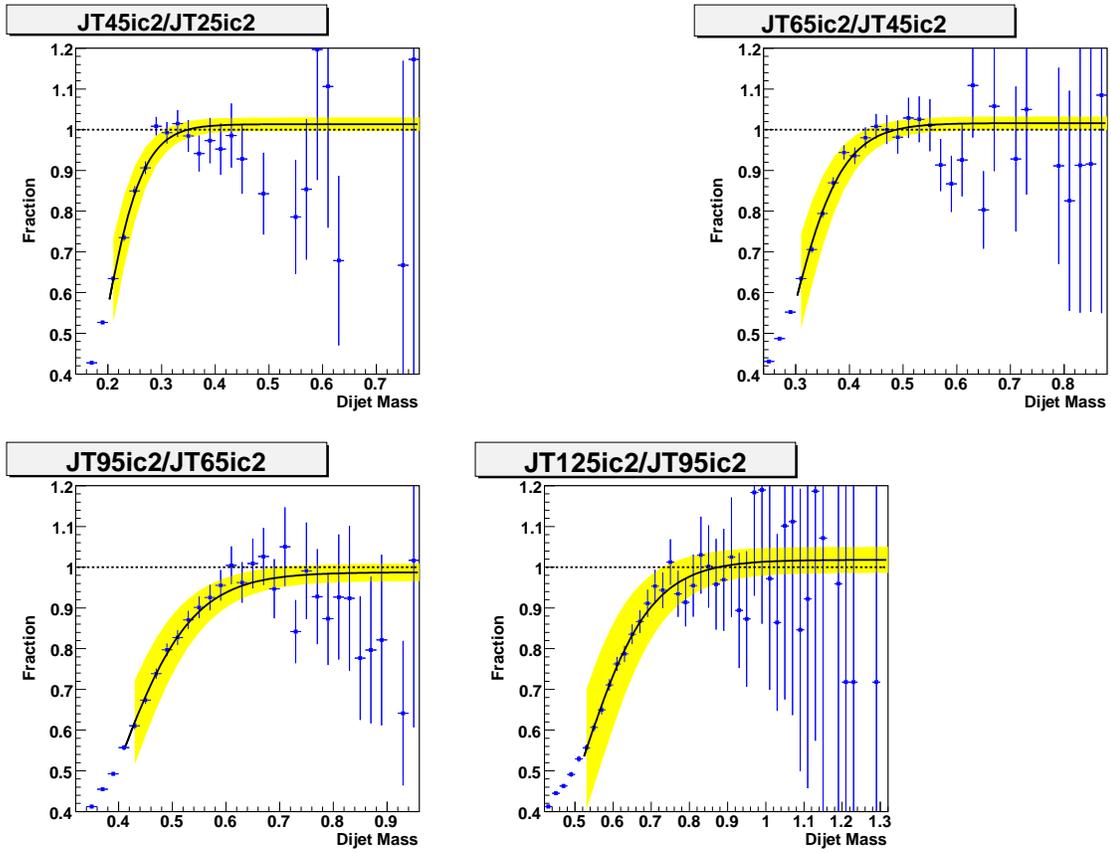
Figure D.4: Trigger turn-on curves for all single jet triggers in the second ICR region $(1.2 < |y|_{max} < 1.6)$. The knee region is fitted.
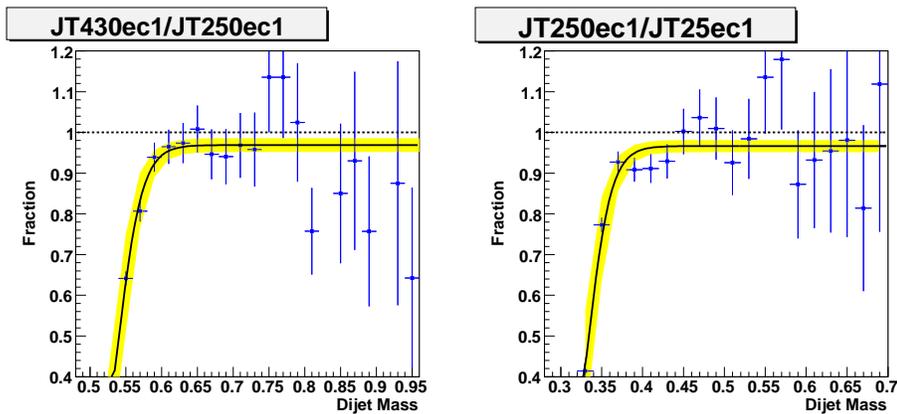
Figure D.5: Trigger turn-on curves for all dijet mass jet triggers in the first EC region $(1.6 < |y|_{max} < 2.0)$. The knee region is fitted.
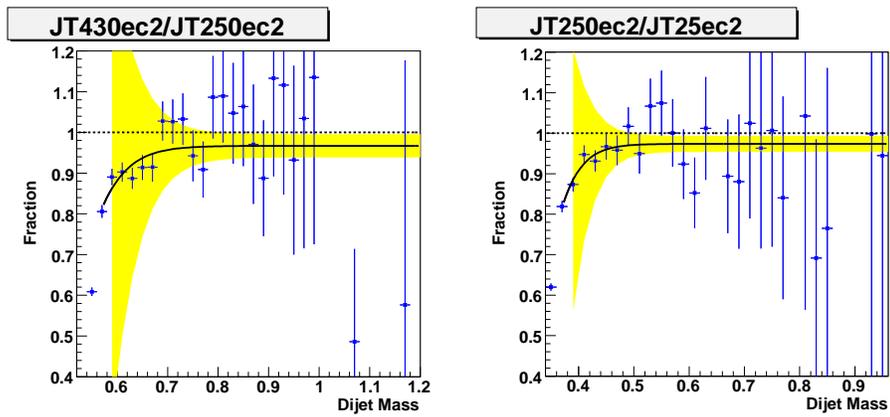
Figure D.6: Trigger turn-on curves for all dijet mass jet triggers in the second EC region $(2.0 < |y|_{max} < 2.4)$. The knee region is fitted.

Appendix E

Rescaled Mass

This appendix shows the rescaling in all the rapidity regions.
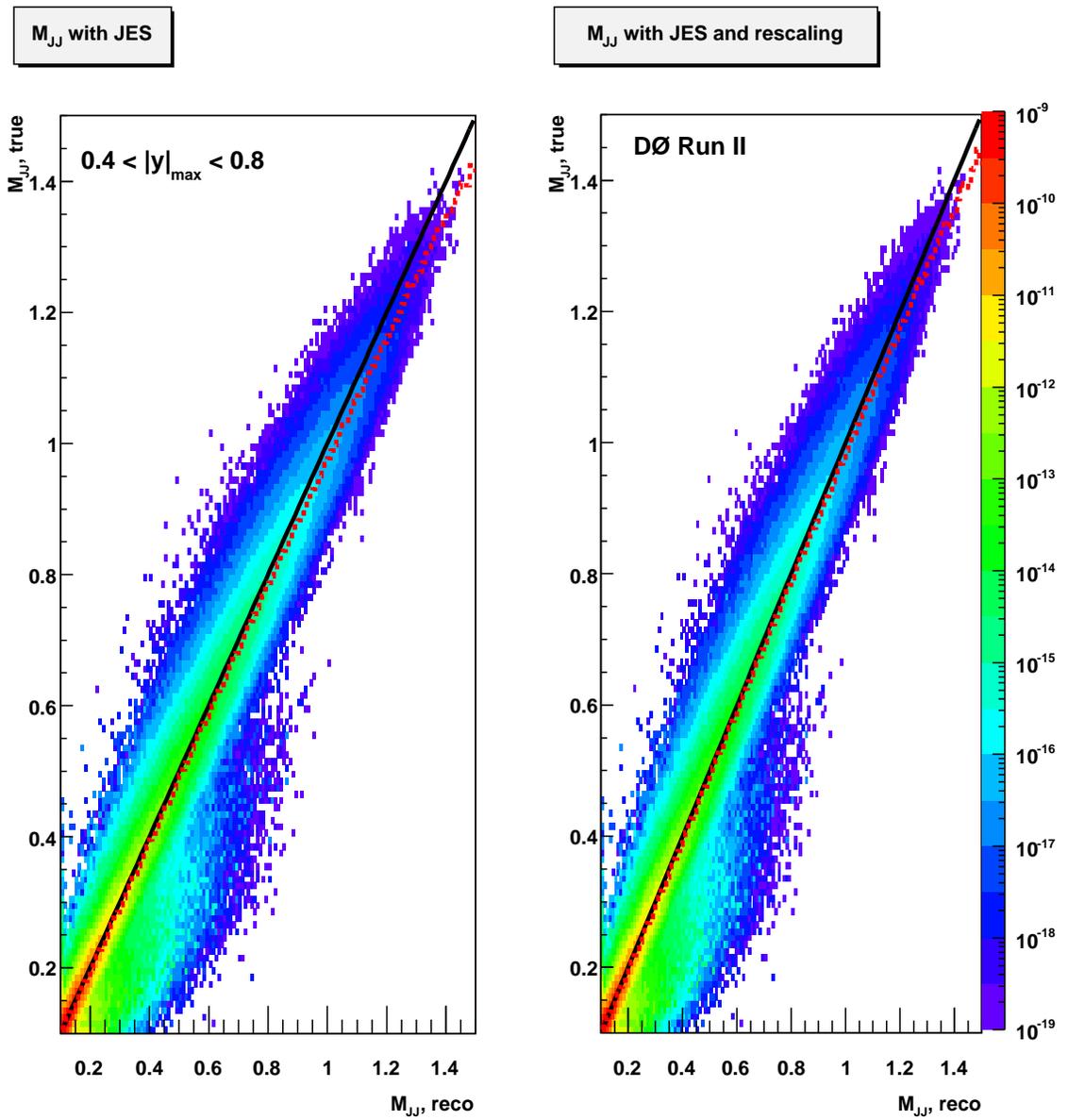
Figure E.1: True vs Reco mass with and without rescaling in the CC2 region.
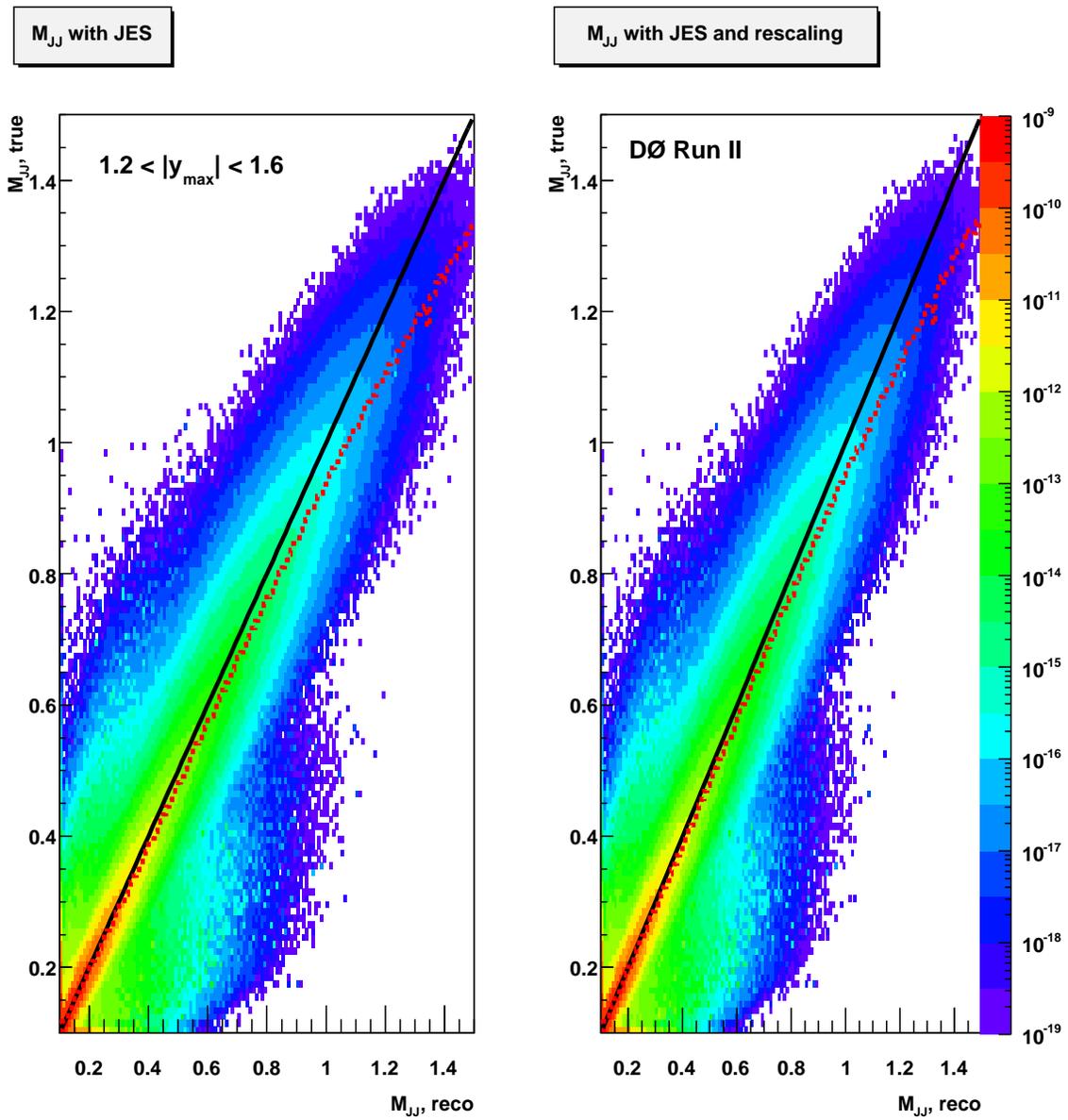
Figure E.2: True vs Reco mass with and without rescaling in the IC2 region.
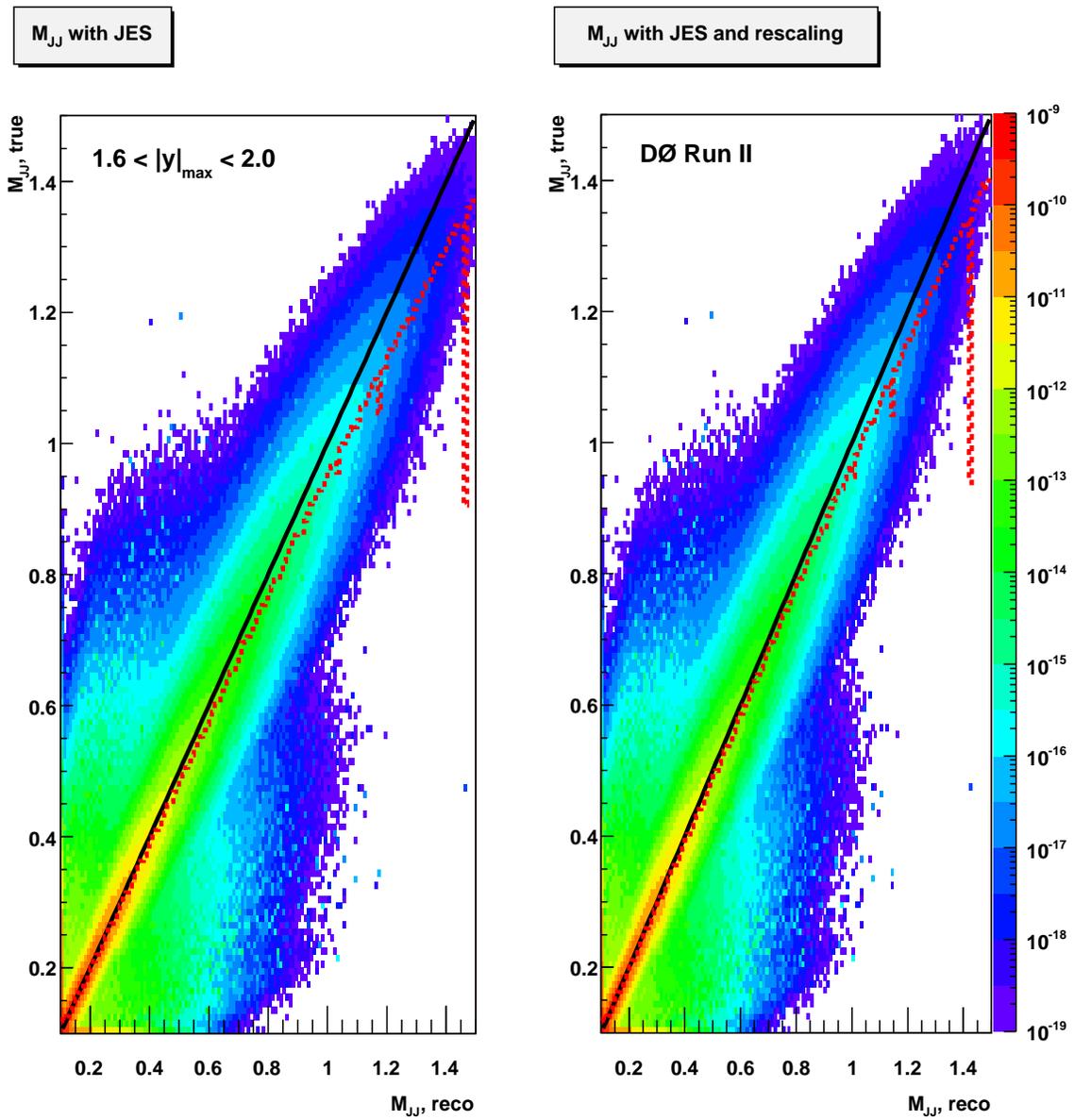
Figure E.3: True vs Reco mass with and without rescaling in the EC1 region.
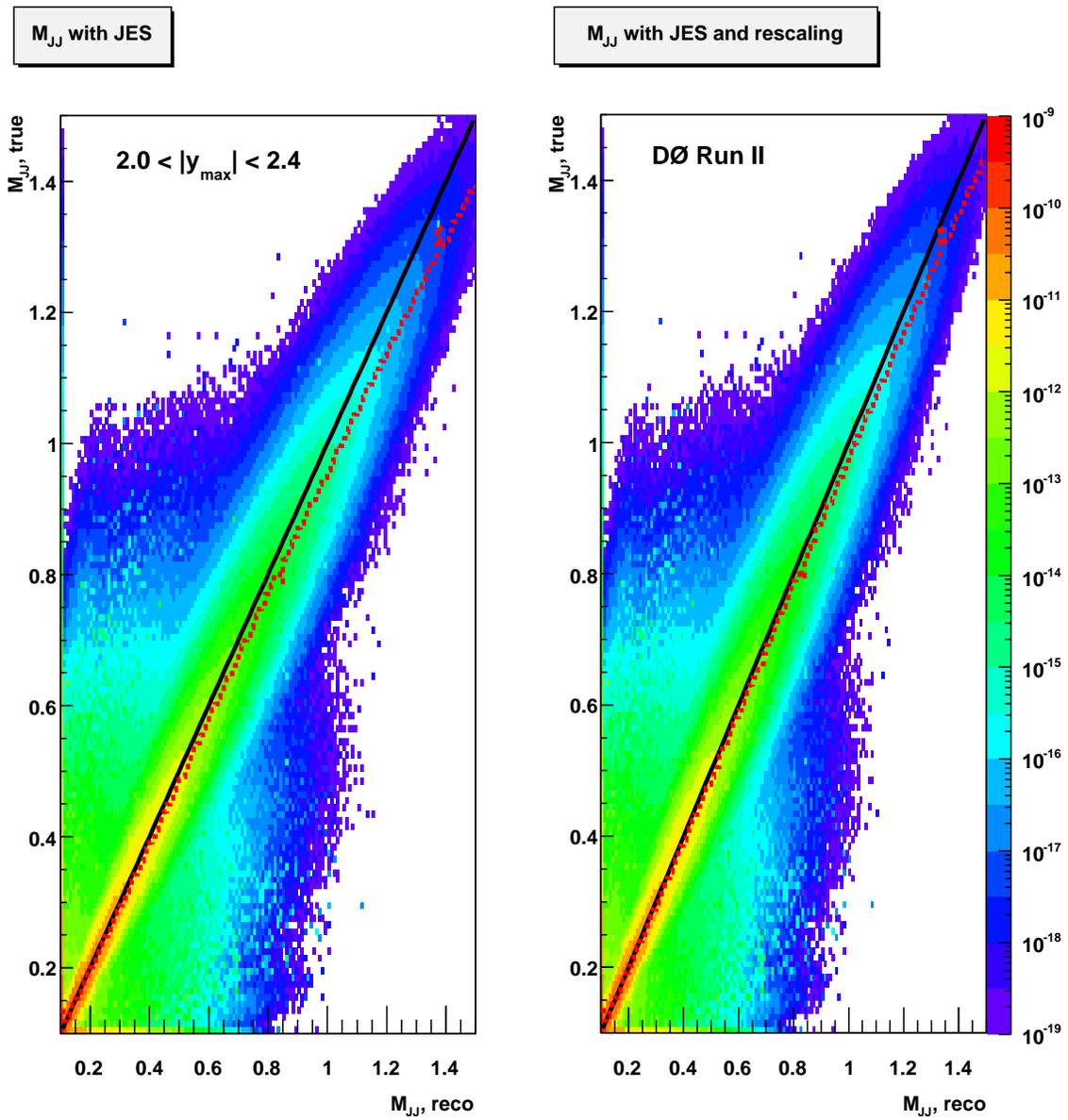
Figure E.4: True vs Reco mass with and without rescaling in the EC2 region.

Appendix F

Control Distributions

The following distributions are the control plots for the dijet mass distribution. The unweighted MC is compared to the data in order to determine which variable will be used in the reweighting. The first set of plots show this. The second set of plots show the reweighted MC compared to data. The raw distributions and the ratios are shown in this ca.
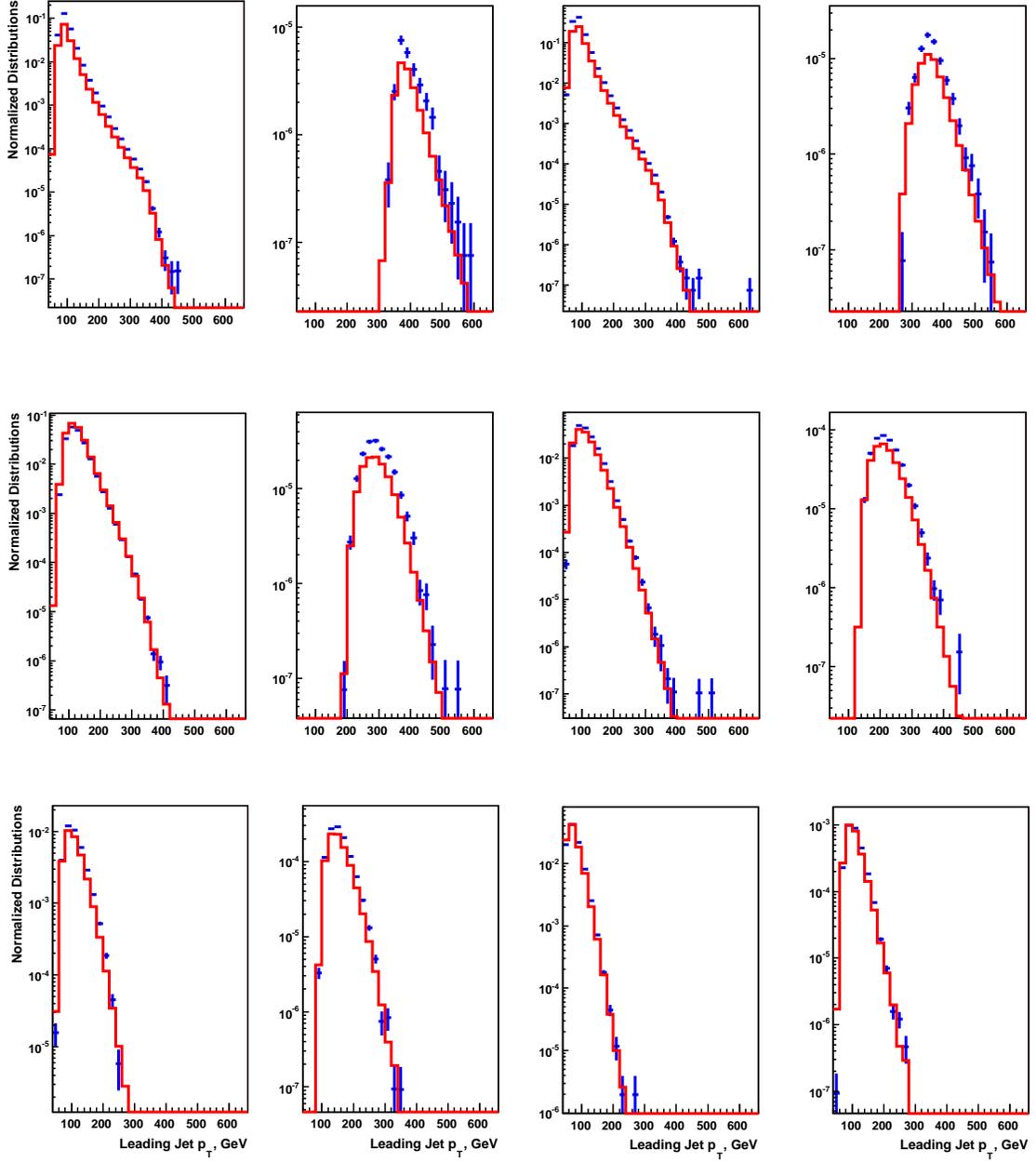
Figure F.1: Control distributions for $p_{T,max}$, no reweighting. The x-axis is $p_{T,max}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
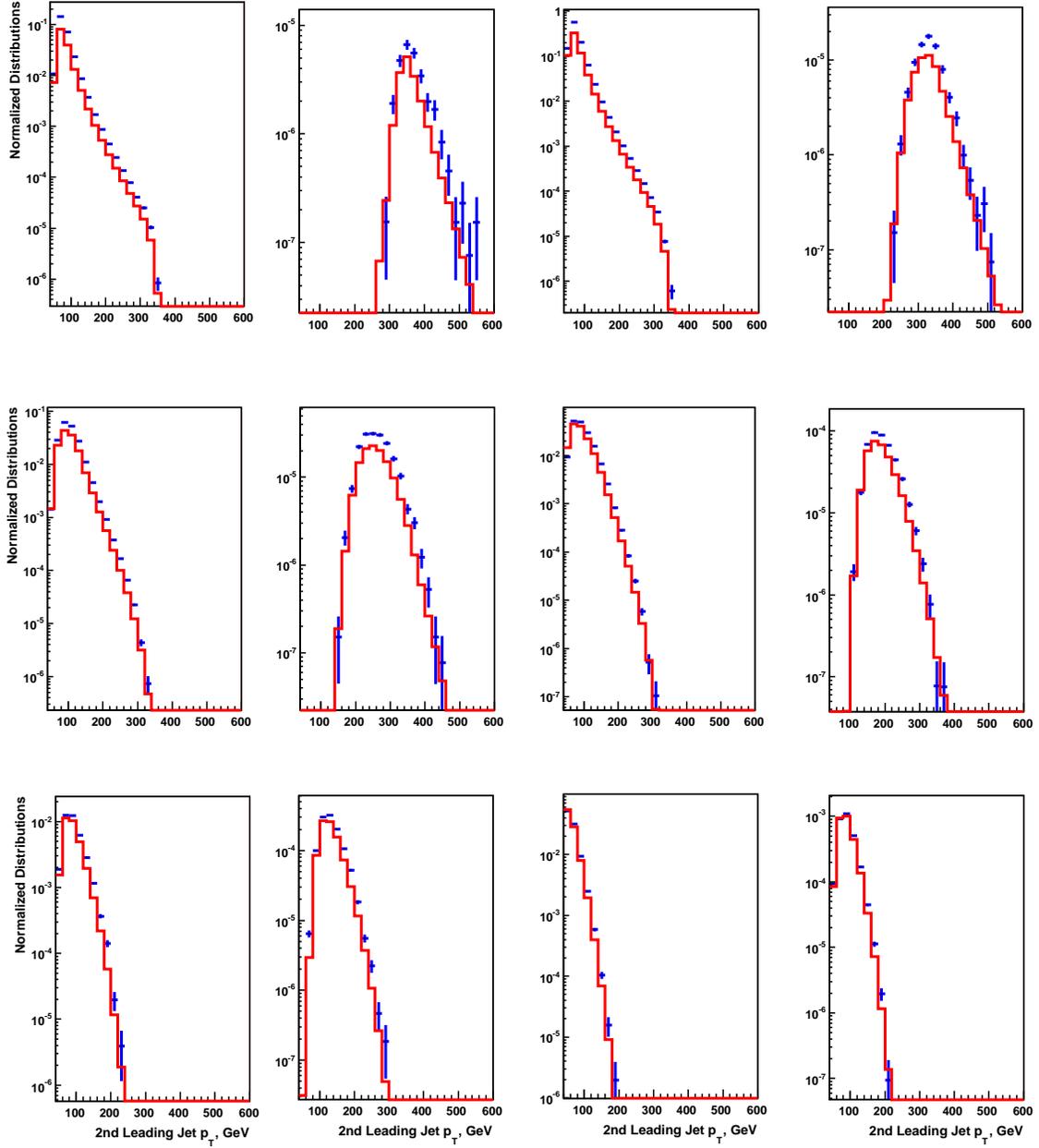
Figure F.2: Control distributions for $p_{T,min}$, no reweighting. The x-axis is $p_{T,min}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.

137
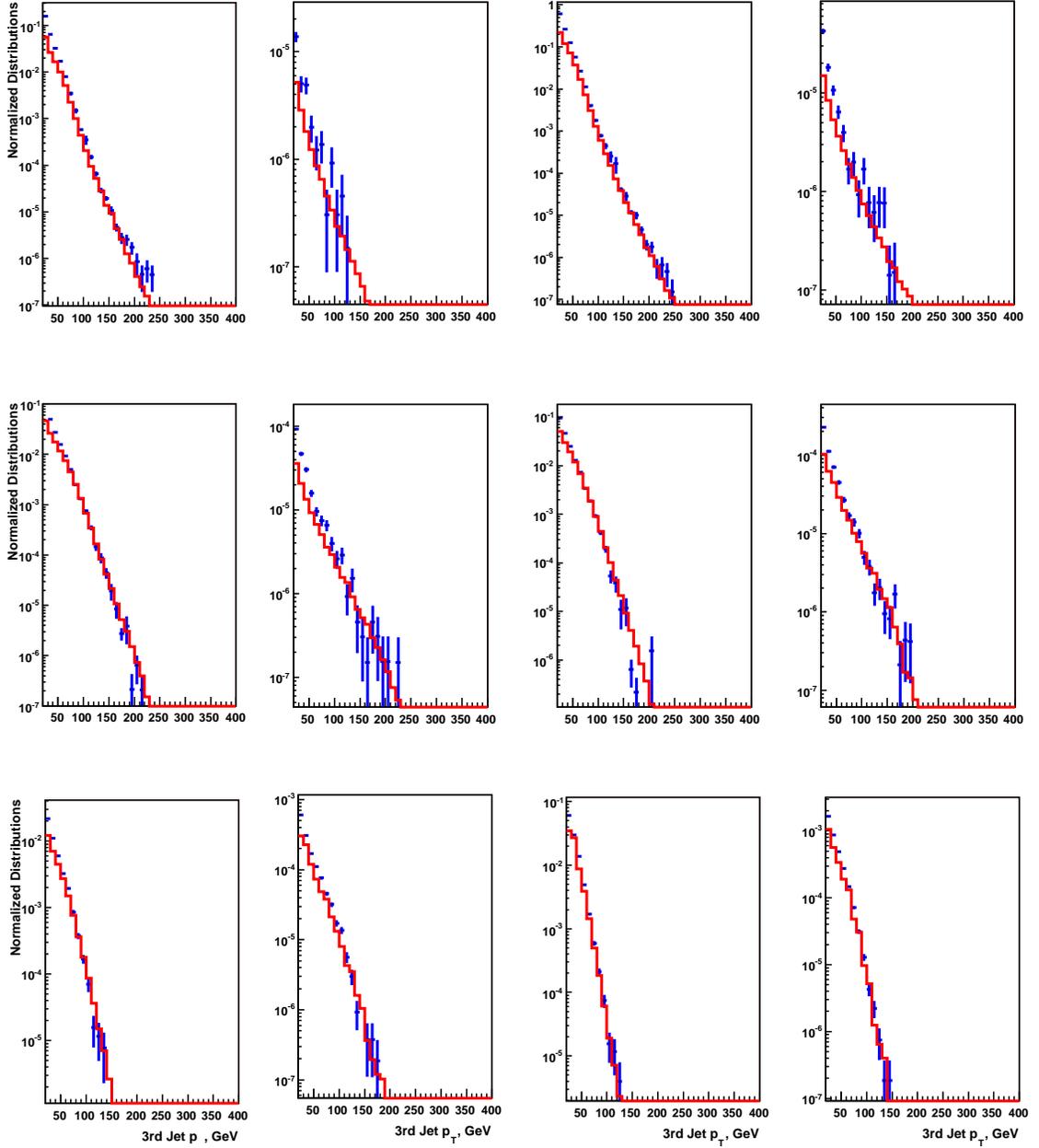
Figure F.3: Control distributions for $p_{T,3}$, no reweighting. The x-axis is $p_{T,min}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
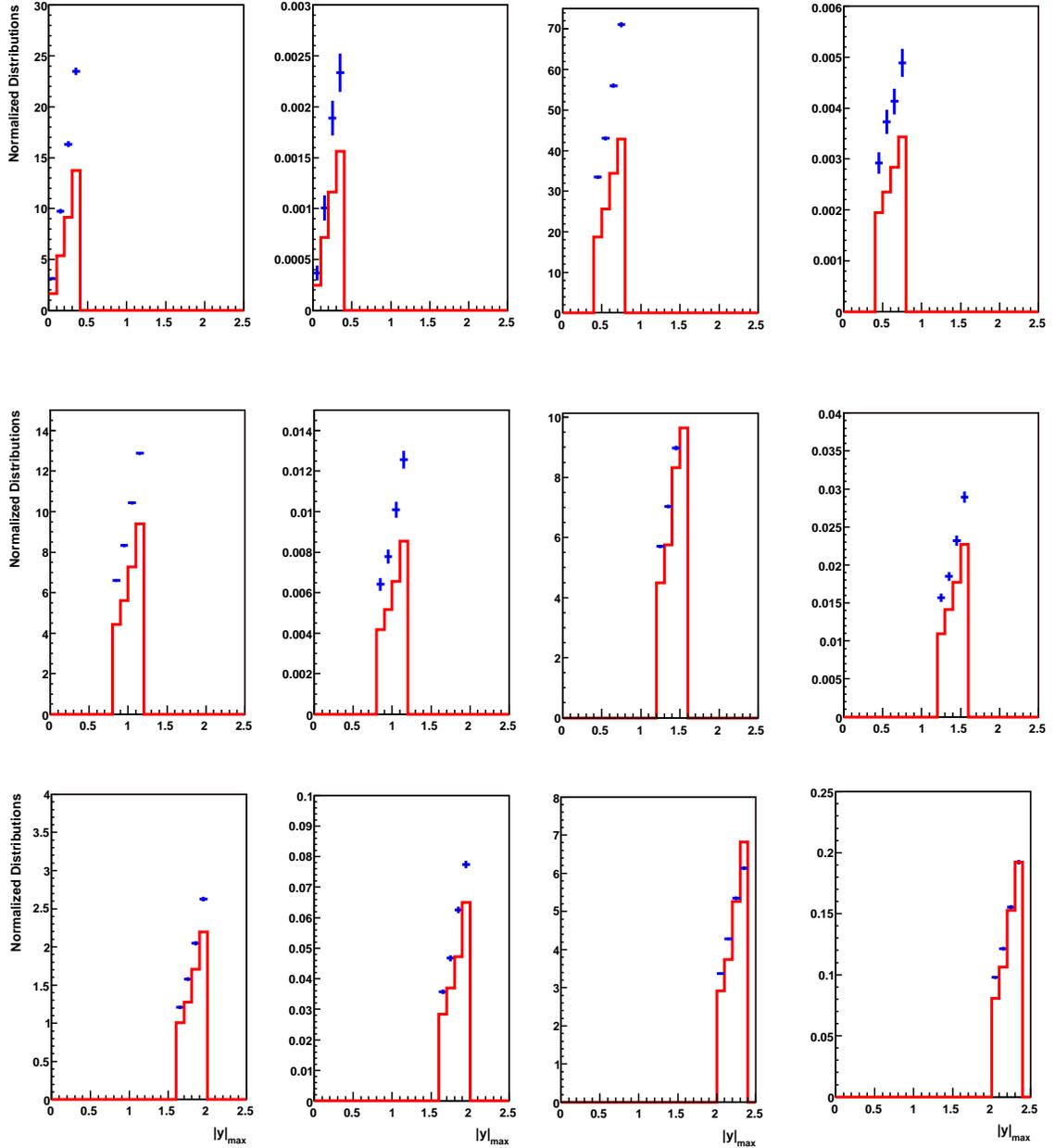
Figure F.4: Control distributions for $|y_{max}|$, no reweighting. The x-axis is $|y_{max}|$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
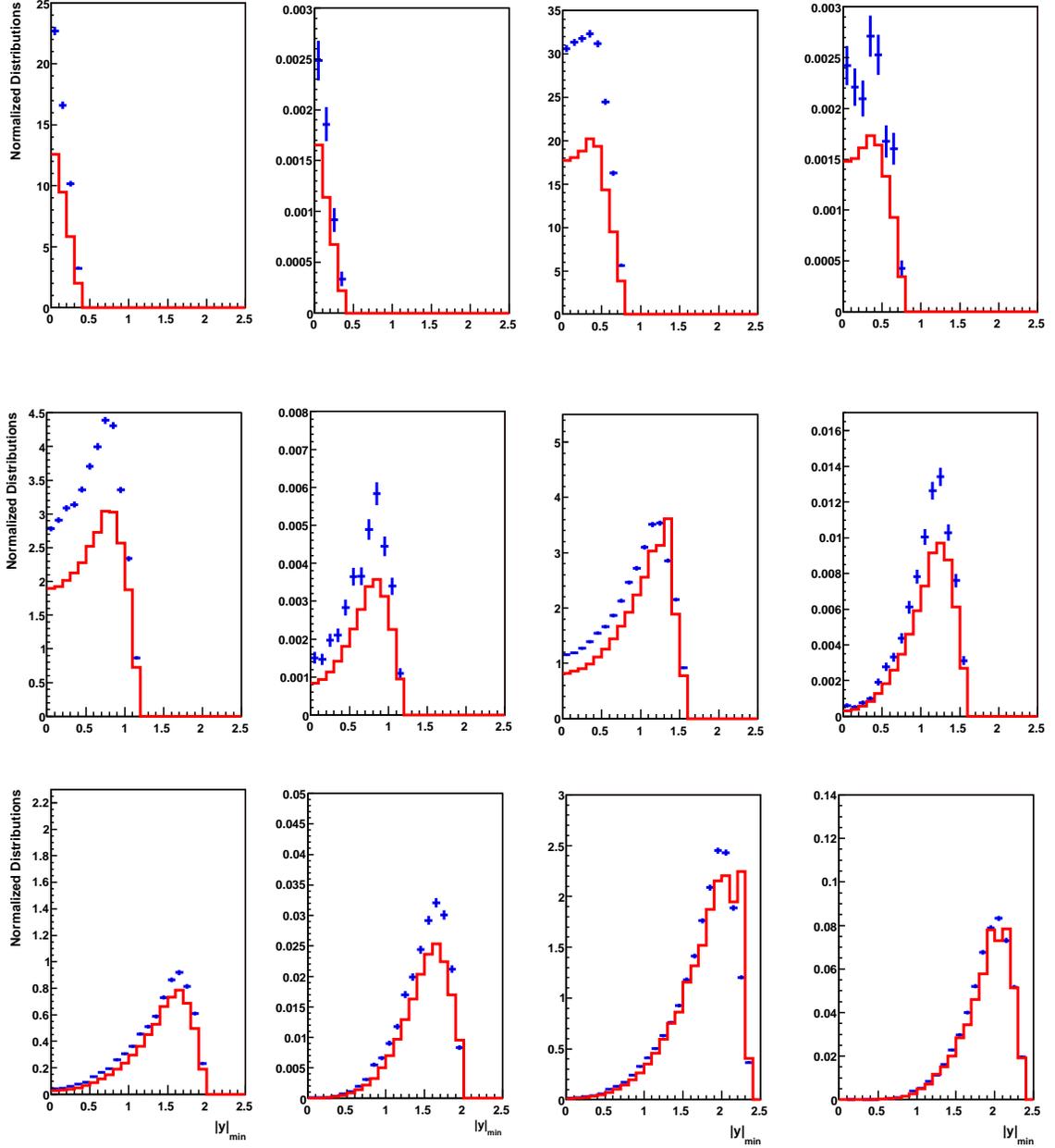
Figure F.5: Control distributions for $|y_{min}|$, no reweighting. The x-axis is $|y_{min}|$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
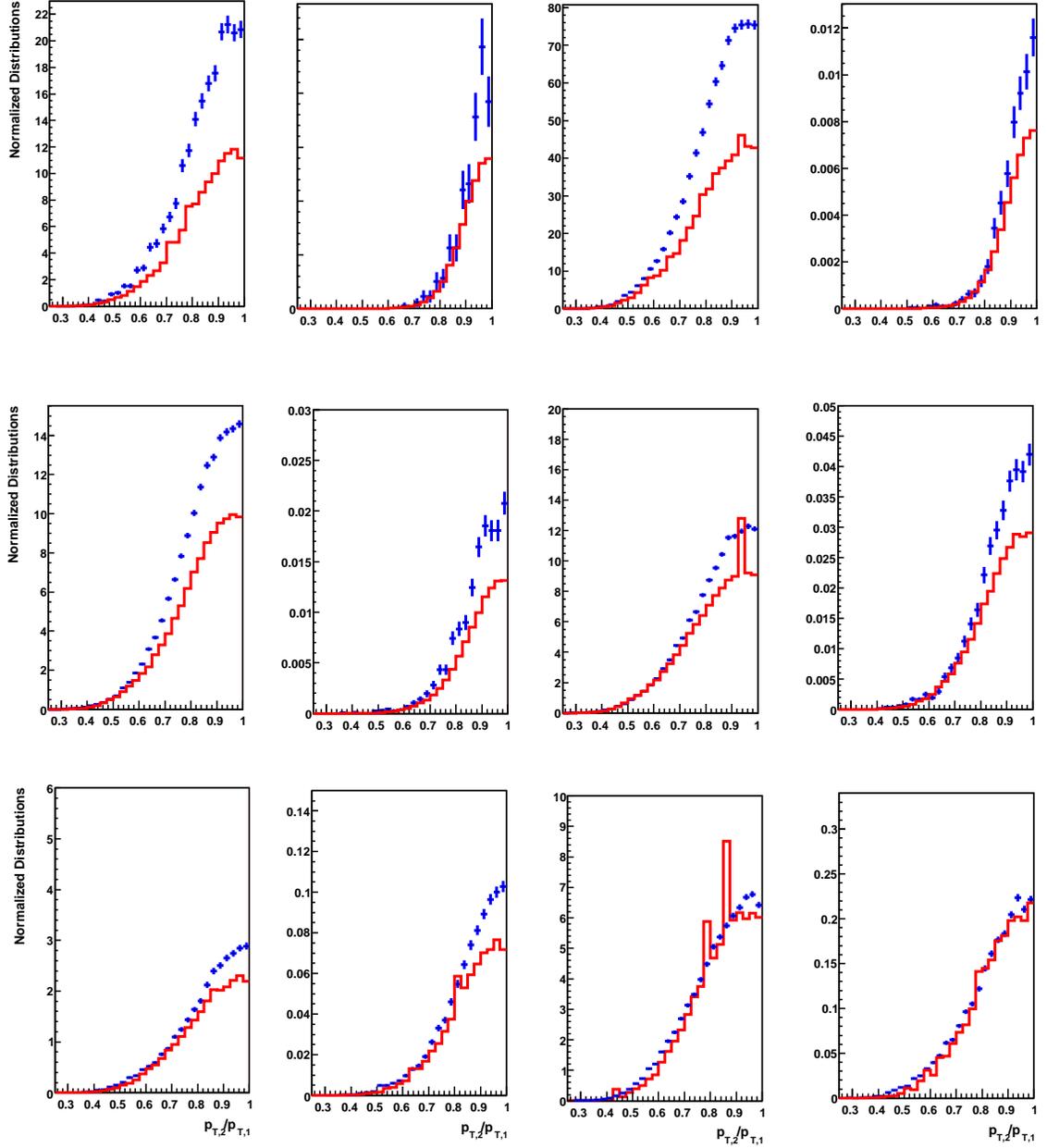
Figure F.6: Control distributions for $p_{T,2}/p_{T,1}$, no reweighting. The x-axis is $p_{T,2}/p_{T,1}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
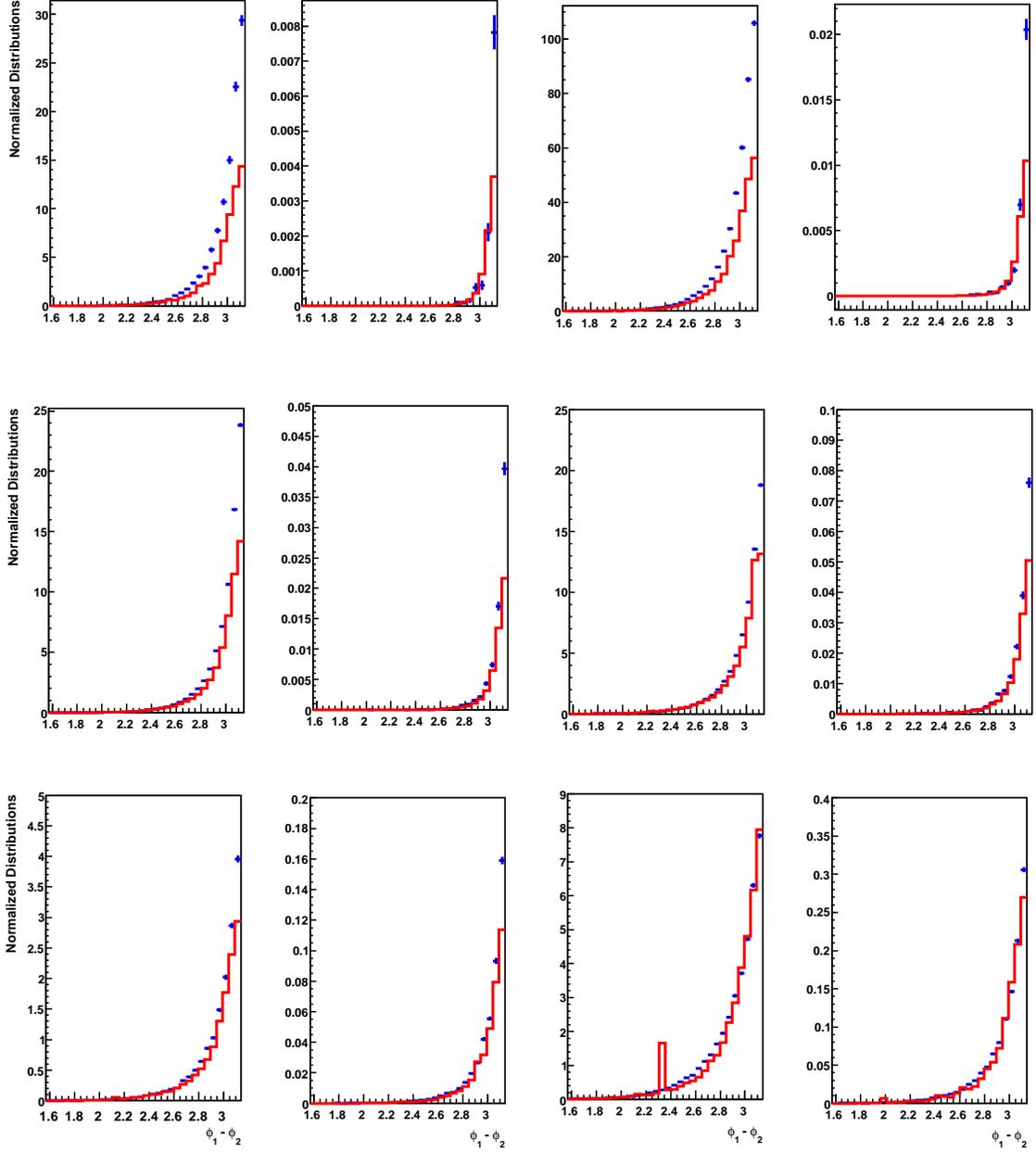
Figure F.7: Control distributions for $\Delta\phi_{1,2}$, no reweighting. The x-axis is $\Delta\phi_{1,2}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
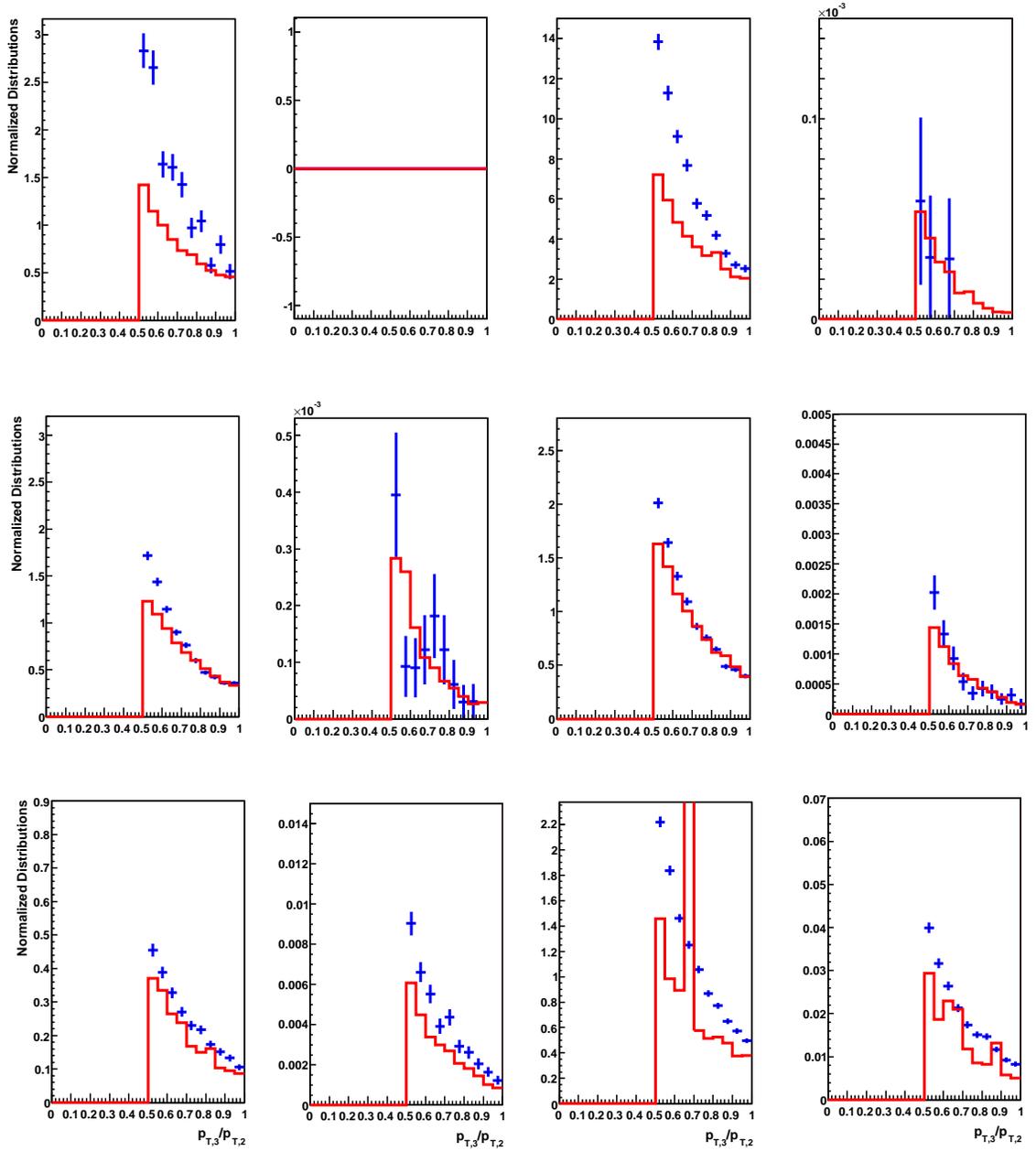
Figure F.8: Control distributions for $p_{T,3}/p_{T,2}$, no reweighting. The x-axis is $p_{T,3}/p_{T,2}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
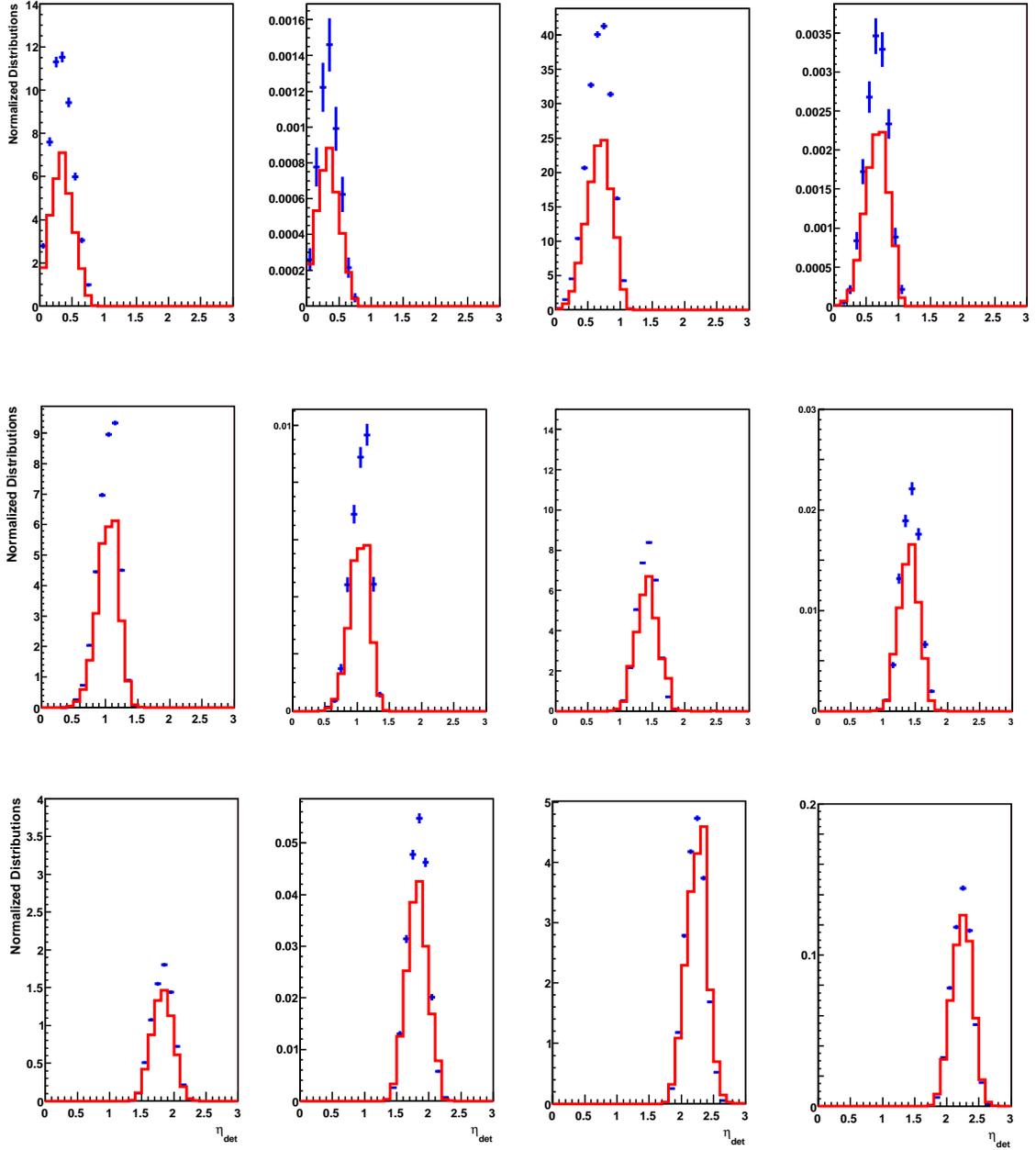
Figure F.9: Control distributions for η_{det}, no reweighting. The x-axis is η_{det}, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
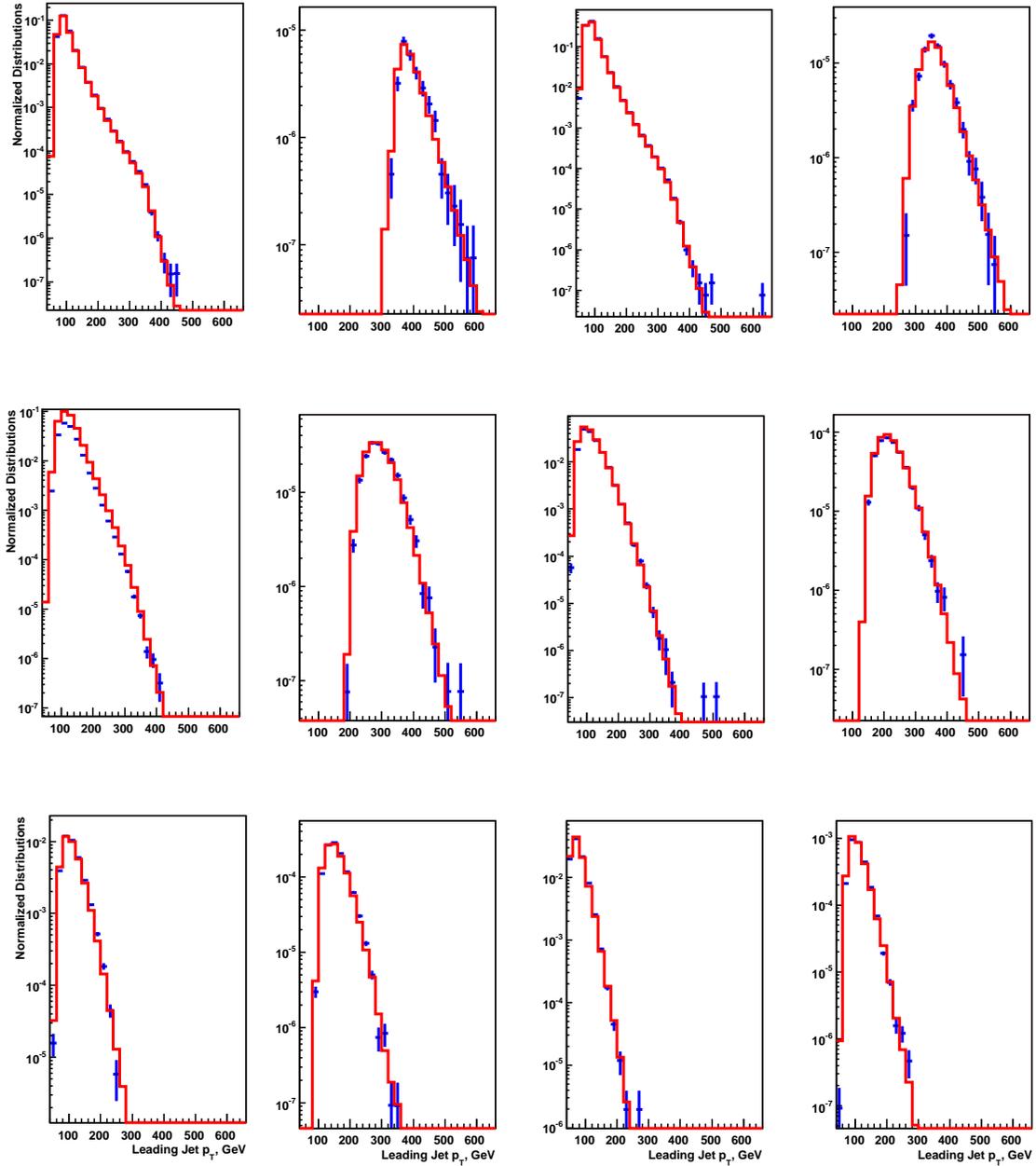
Figure F.10: Control distributions for $p_{T,max}$, after reweighting. The x-axis is $p_{T,max}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
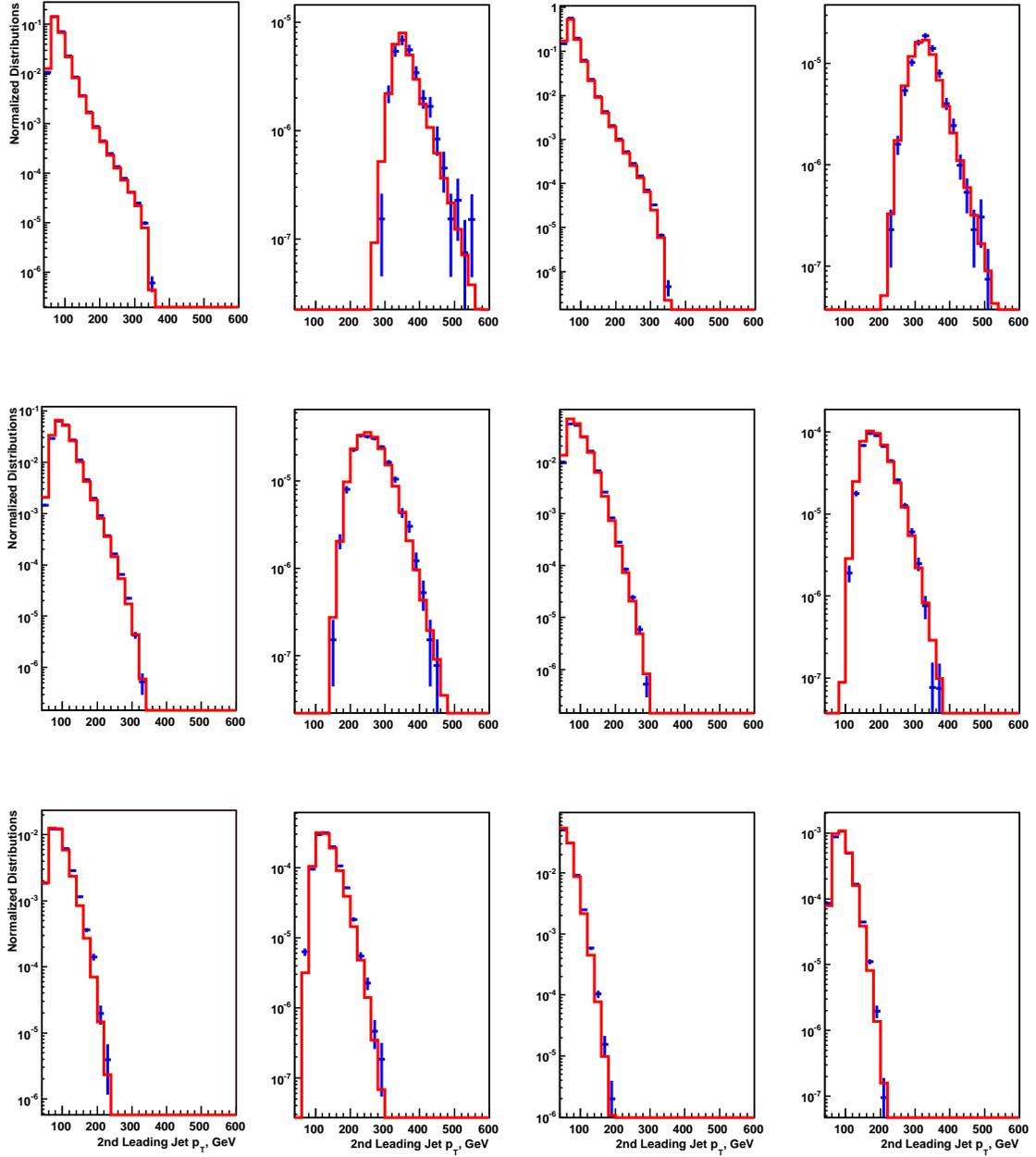
Figure F.11: Control distributions for $p_{T,min}$, after reweighting. The x-axis is $p_{T,min}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.

146
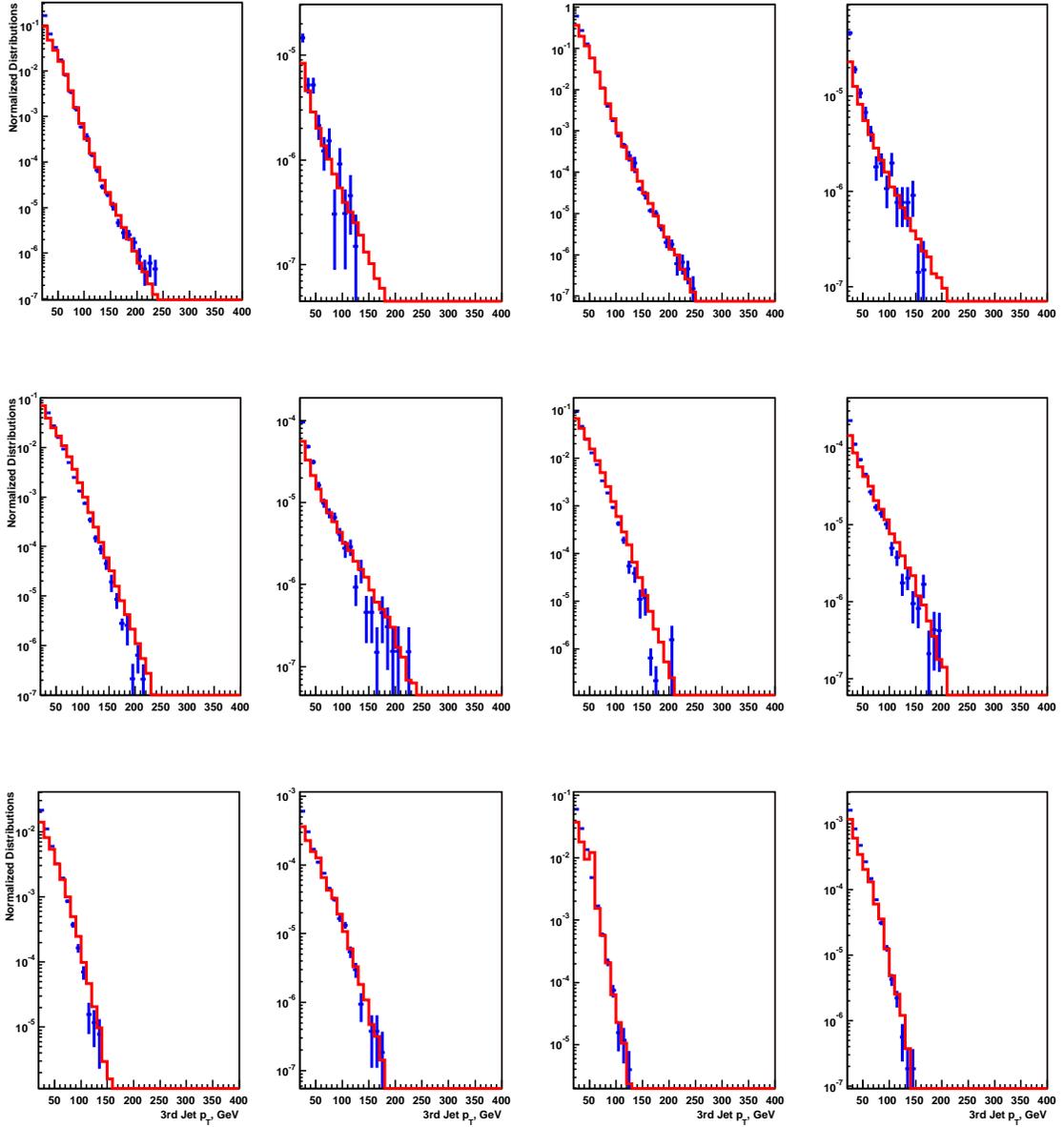
Figure F.12: Control distributions for $p_{T,3}$, after reweighting. The x-axis is $p_{T,3}$ in units of GeV, the y-axis is the normalized number of events per 20 GeV. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
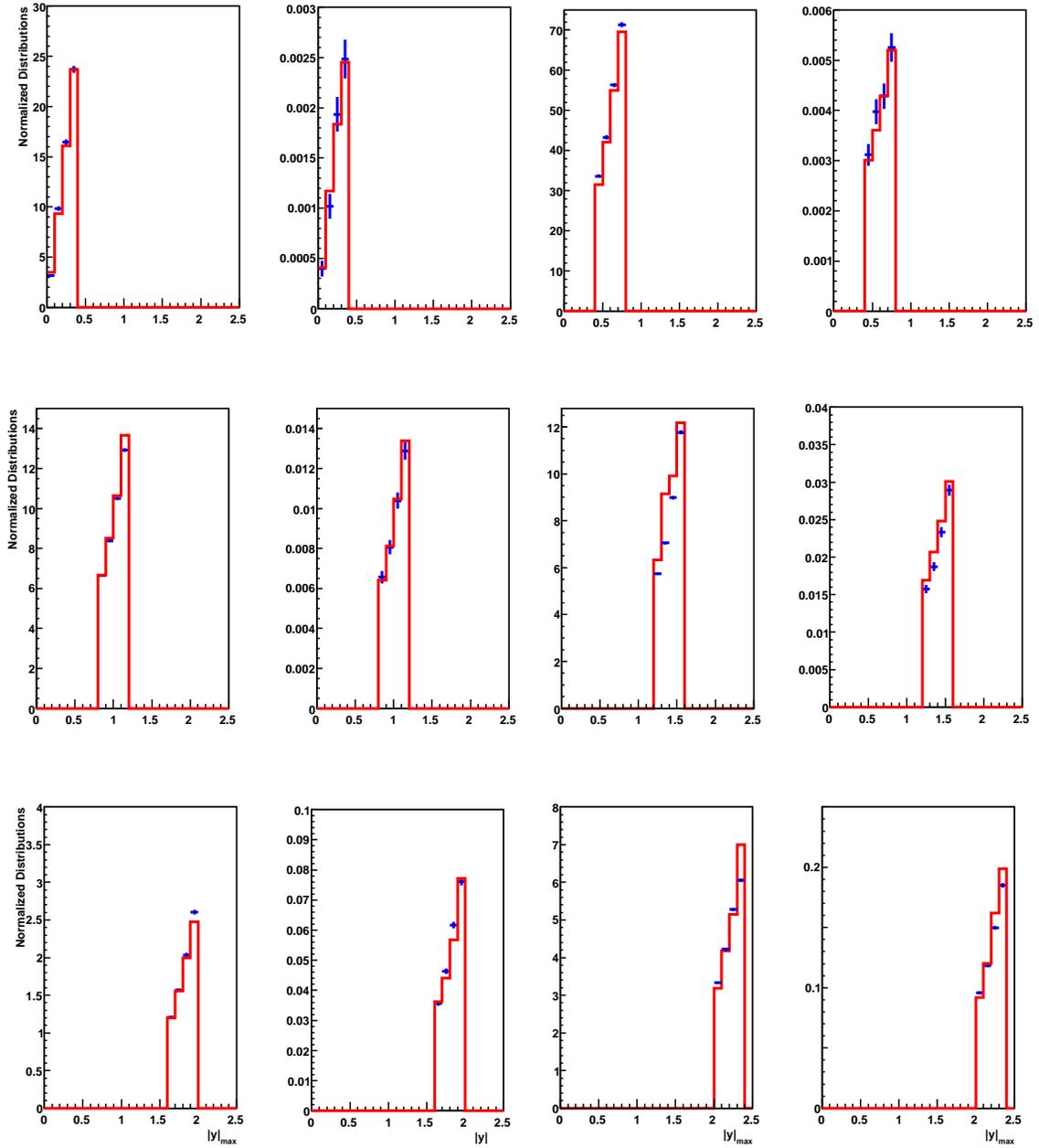
Figure F.13: Control distributions for $|y_{max}|$, after reweighting. The x-axis is $|y_{max}|$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
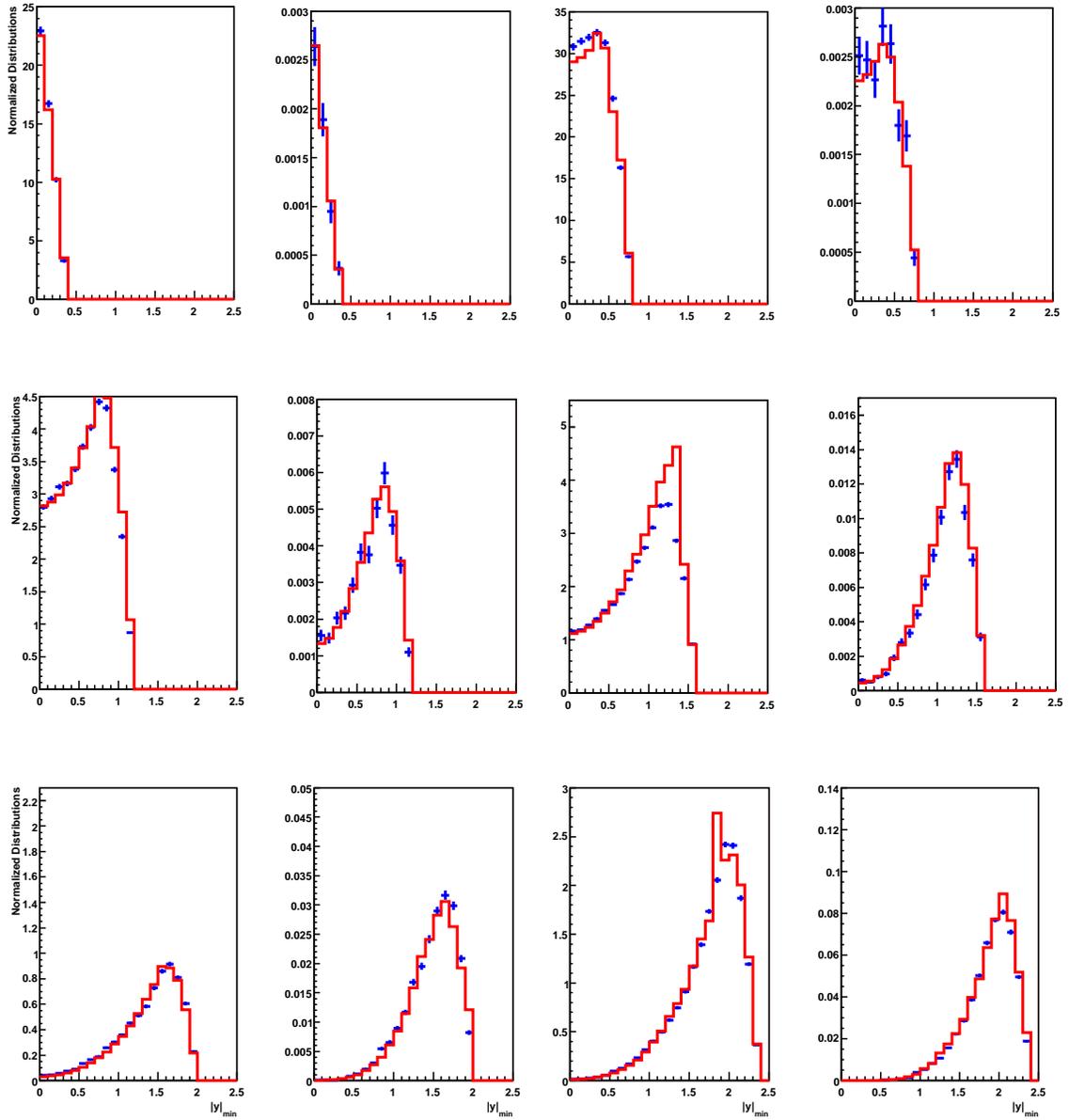
Figure F.14: Control distributions for $|y_{min}|$, after reweighting. The x-axis is $|y_{min}|$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
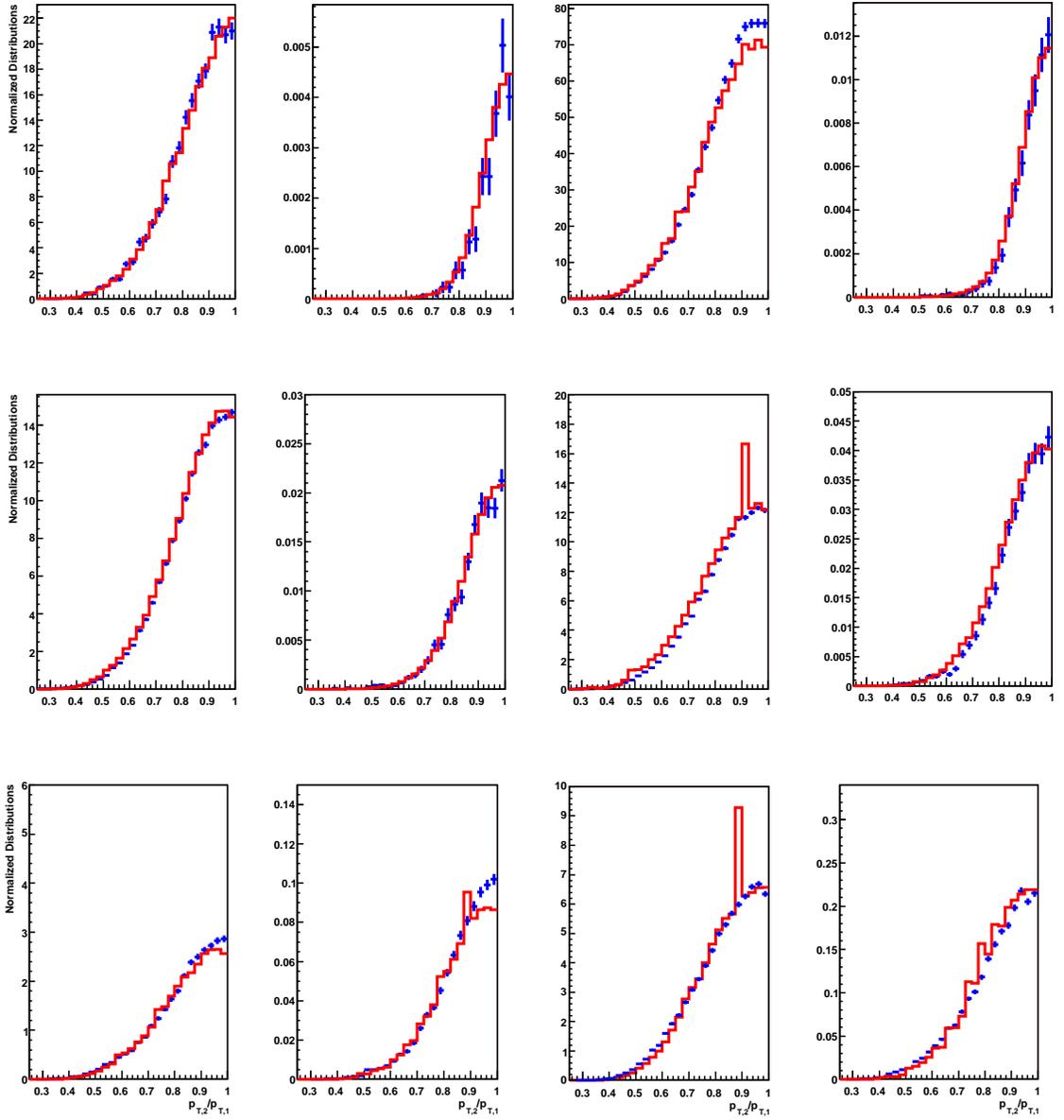
Figure F.15: Control distributions for $p_{T,2}/p_{T,1}$, after reweighting. The x-axis is $p_{T,2}/p_{T,1}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
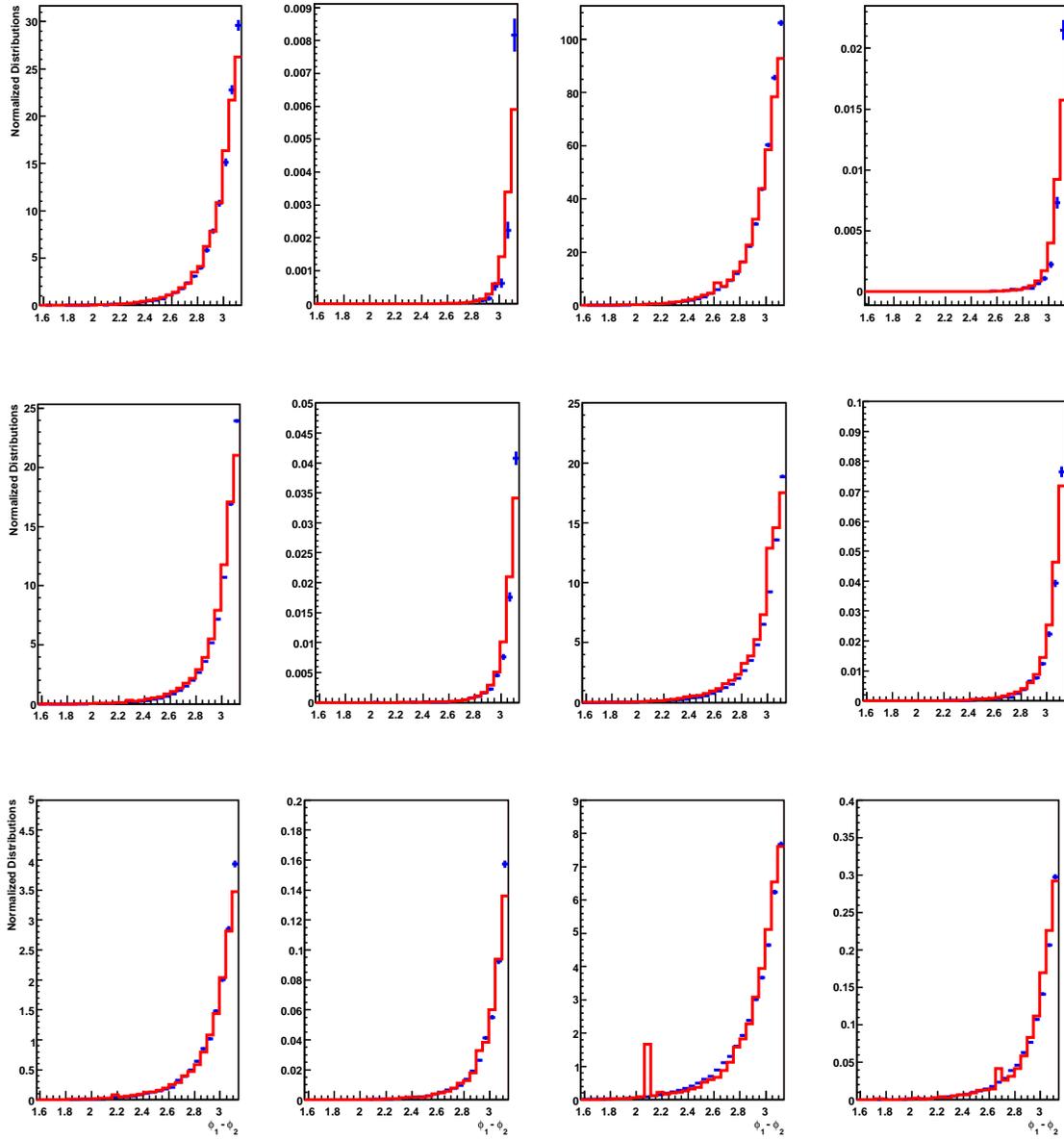
Figure F.16: Control distributions for $\Delta\phi_{1,2}$, after reweighting. The x-axis is $\Delta\phi_{1,2}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
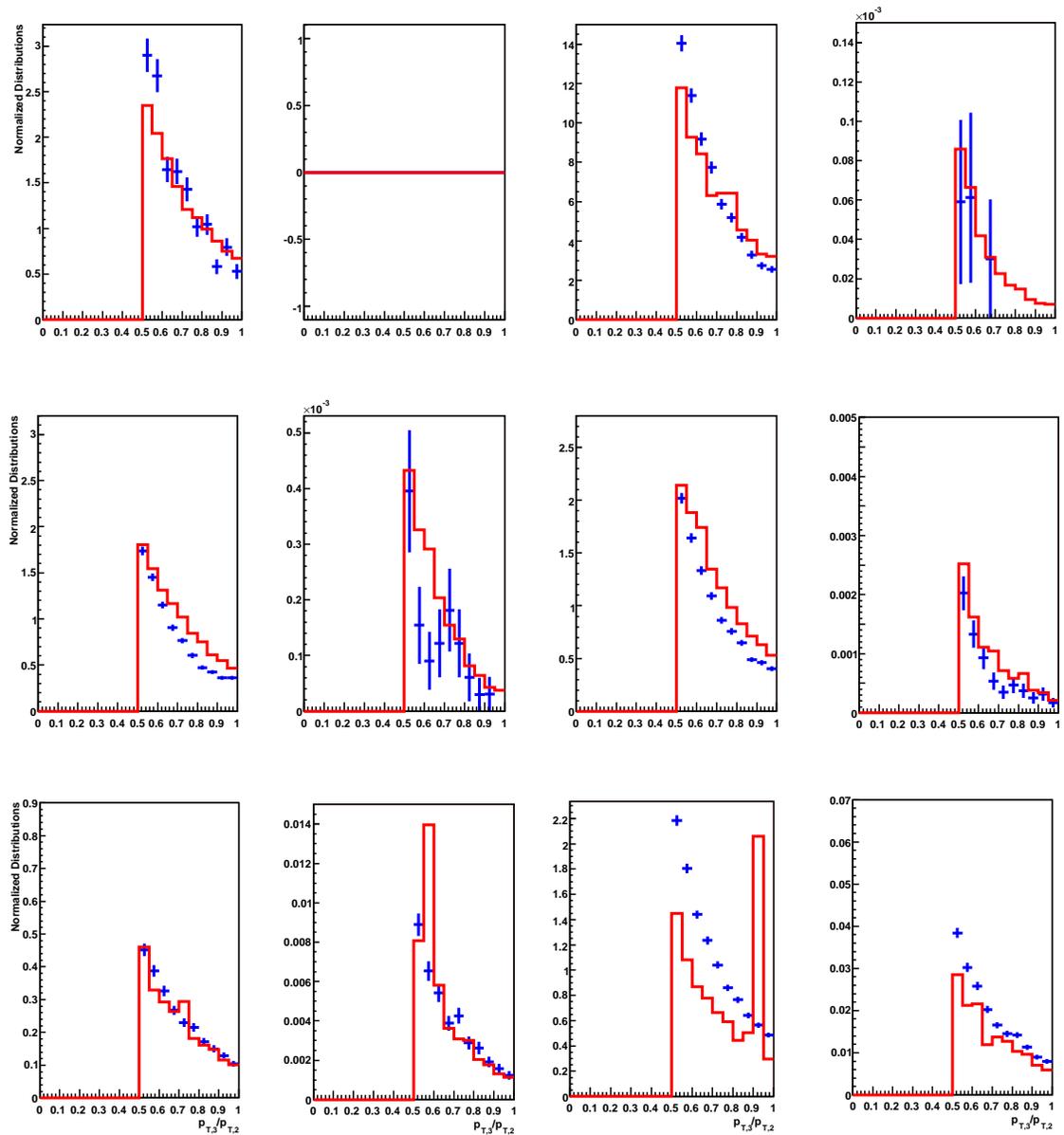
Figure F.17: Control distributions for $p_{T,3}/p_{T,2}$, no reweighting. The x-axis is $p_{T,3}/p_{T,2}$, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.
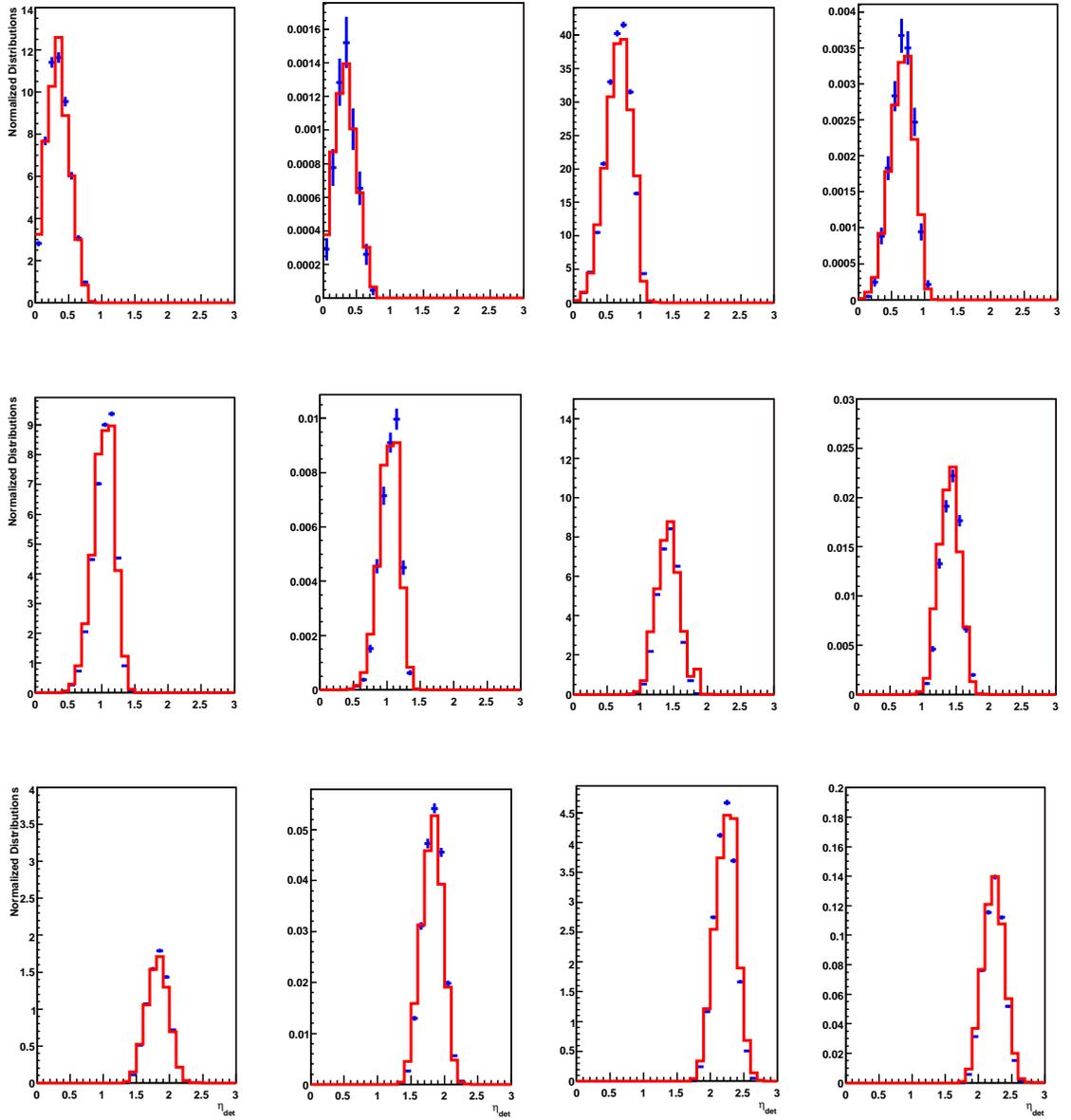
Figure F.18: Control distributions for η_{det}, after reweighting. The x-axis is η_{det}, the y-axis is the normalized number of events. The points represent the data and the histogram is the MC. The top row shows both Central regions (CC1 low mass, CC1 high mass, CC2 low mass, CC2 high mass), the second row shows both ICR regions, and the third row shows the EC regions, also alternating between low mass and high mass.

Appendix G

p_T Resolutions

This appendix contains the supplemental information for the p_T resolutions. Table G.1 lists the value of the parameters used in the smearing equation, while Figure G.1 shows how each tail parameter behaves as a function of η_{det}

	H	P_0	P_1	κ_0	κ_1	λ_0	λ_1	$\mu_{H,0}$	$\mu_{H,1}$	$\mu_{H,2}$
$|y| < 0.4$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0
$0.4 < |y| < 0.8$	0	0.0008	8.76e-08	1	1	11.9	0.0118	0	0	0
$0.8 < |y| < 1.1$	0.024	0.00012	2.18e-06	1.18	2.045	11.9	0.012	0.178	-0.0805	0
$1.1 < |y| < 1.3$	0.11	0.00012	2.18e-06	1.18	2.045	11.9	0.012	0.1	-0.0732	0.014
$1.3 < |y| < 1.6$	0.024	0.00085	-1.21e-06	1.1	1.5	11.9	0.012	0.178	-0.0805	0
$1.6 < |y| < 2.0$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0
$2.0 < |y| < 2.4$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0
$2.4 < |y| < 2.8$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0
$2.8 < |y| < 3.2$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0
$3.2 < |y| < 3.6$	0	0.00041	3.16e-07	1	1	11.9	0.0118	0	0	0

Table G.1: p_T resolution tail parameters for all η regions.
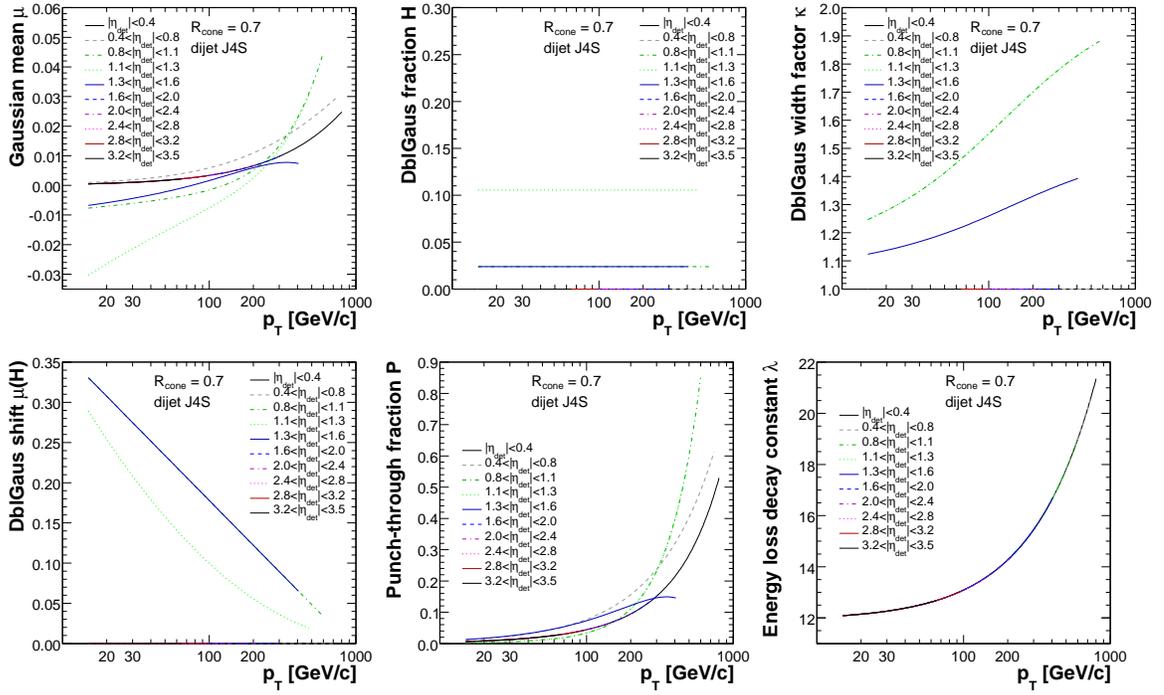
Figure G.1: Tail parameters used in the resolution function (GPT).

Appendix H

Correction Factors

The correction factors due to all the detector effects are shown in this appendix. The total correction factor is determined by add all the sources in quadrature and then fitting to reduce statistical fluctuations.
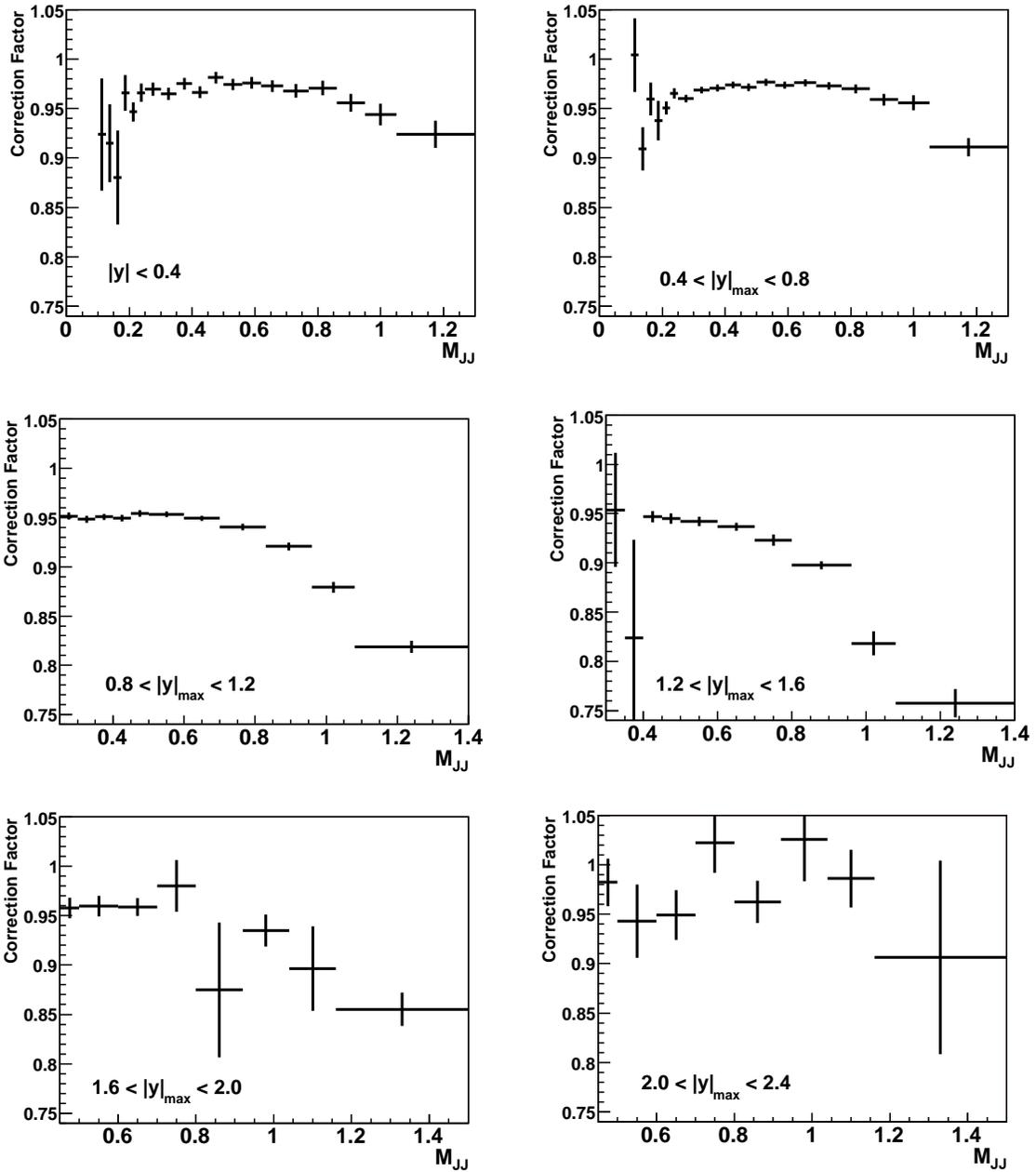
Figure H.1: Correction factor due to p_T resolutions.
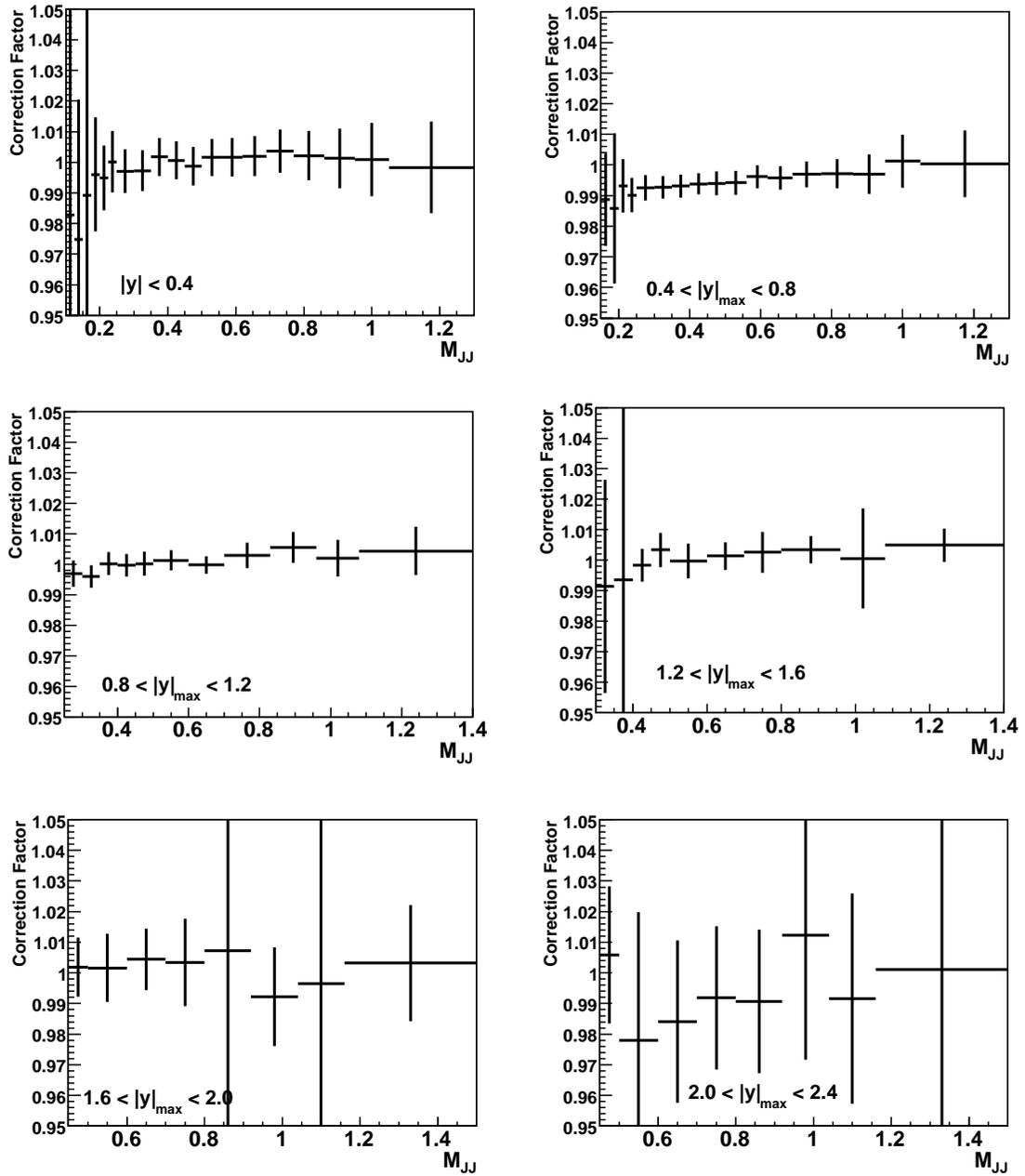
Figure H.2: Correction factor due to η resolutions.
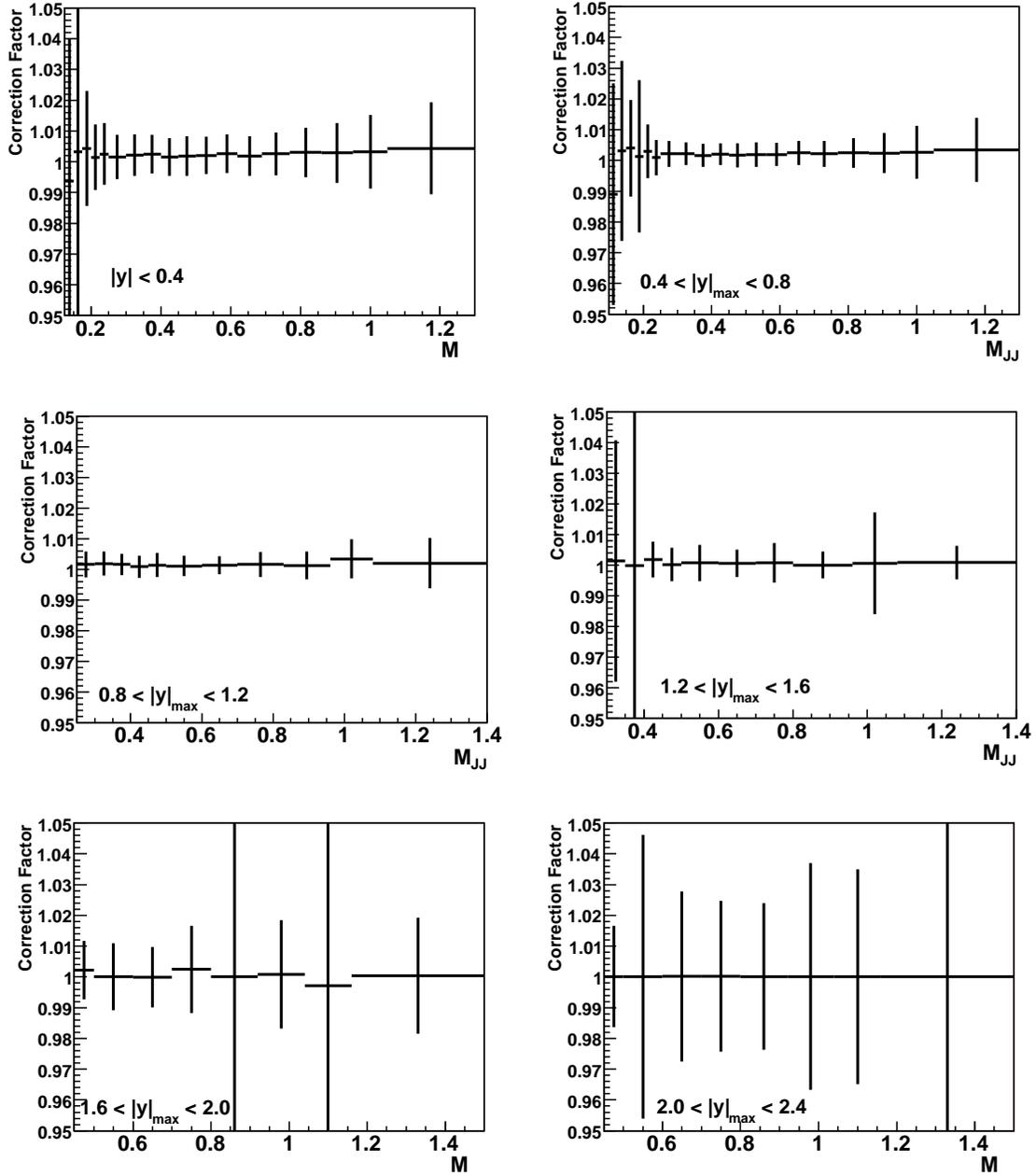
Figure H.3: Correction factor due to ϕ resolutions.
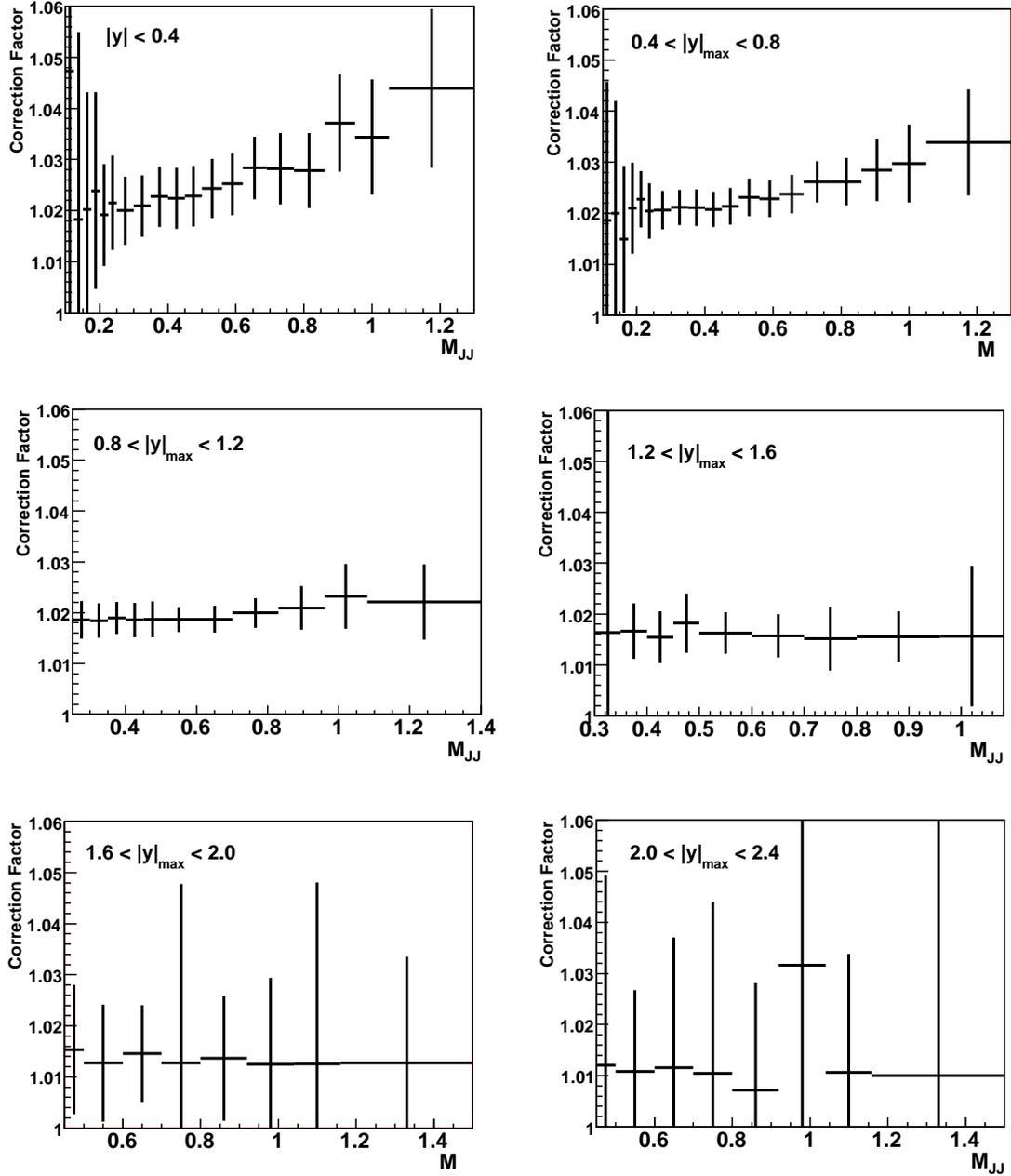
Figure H.4: Correction factor due to muon/neutrino energies.
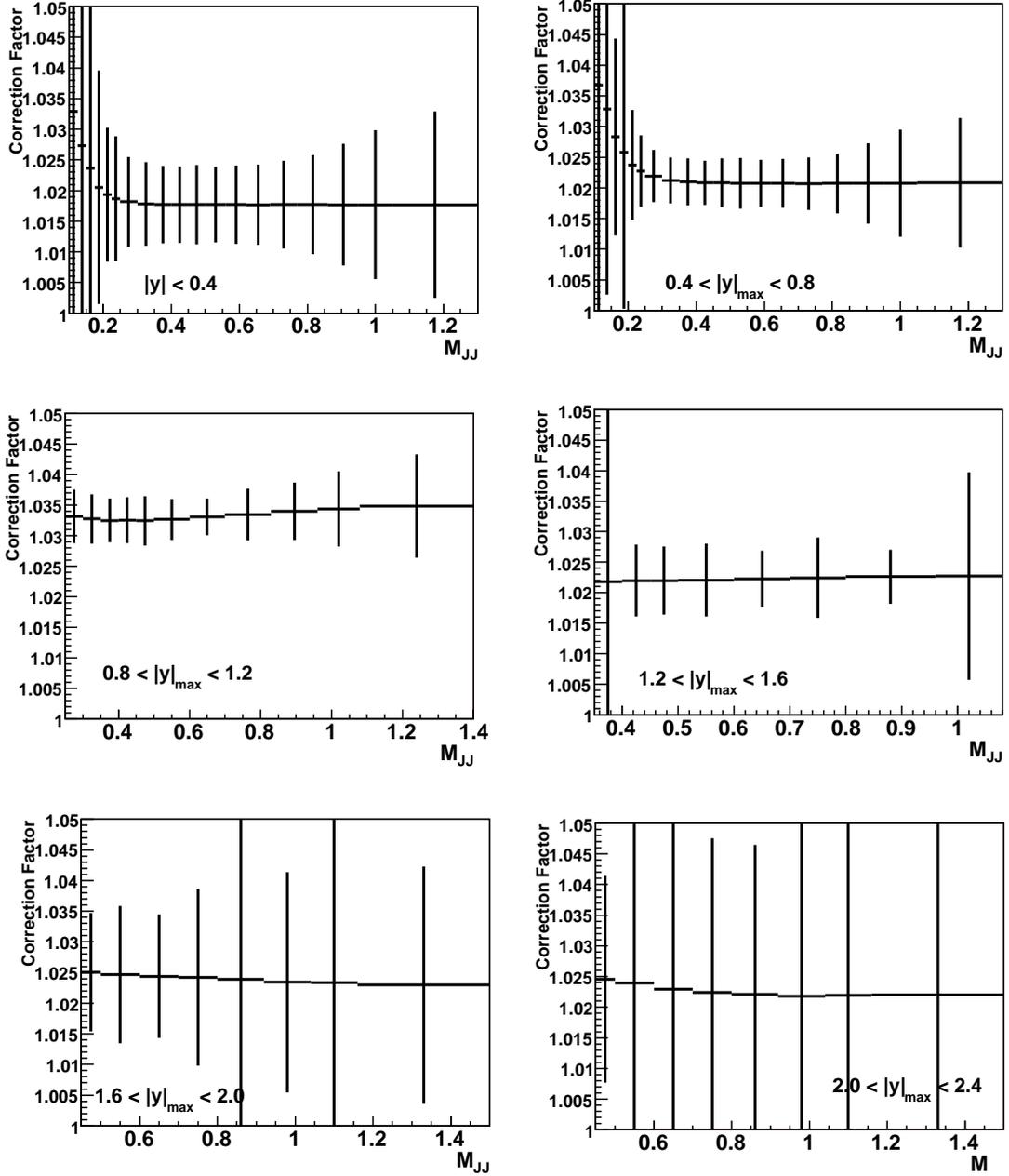
Figure H.5: Correction factor due to jet ID.
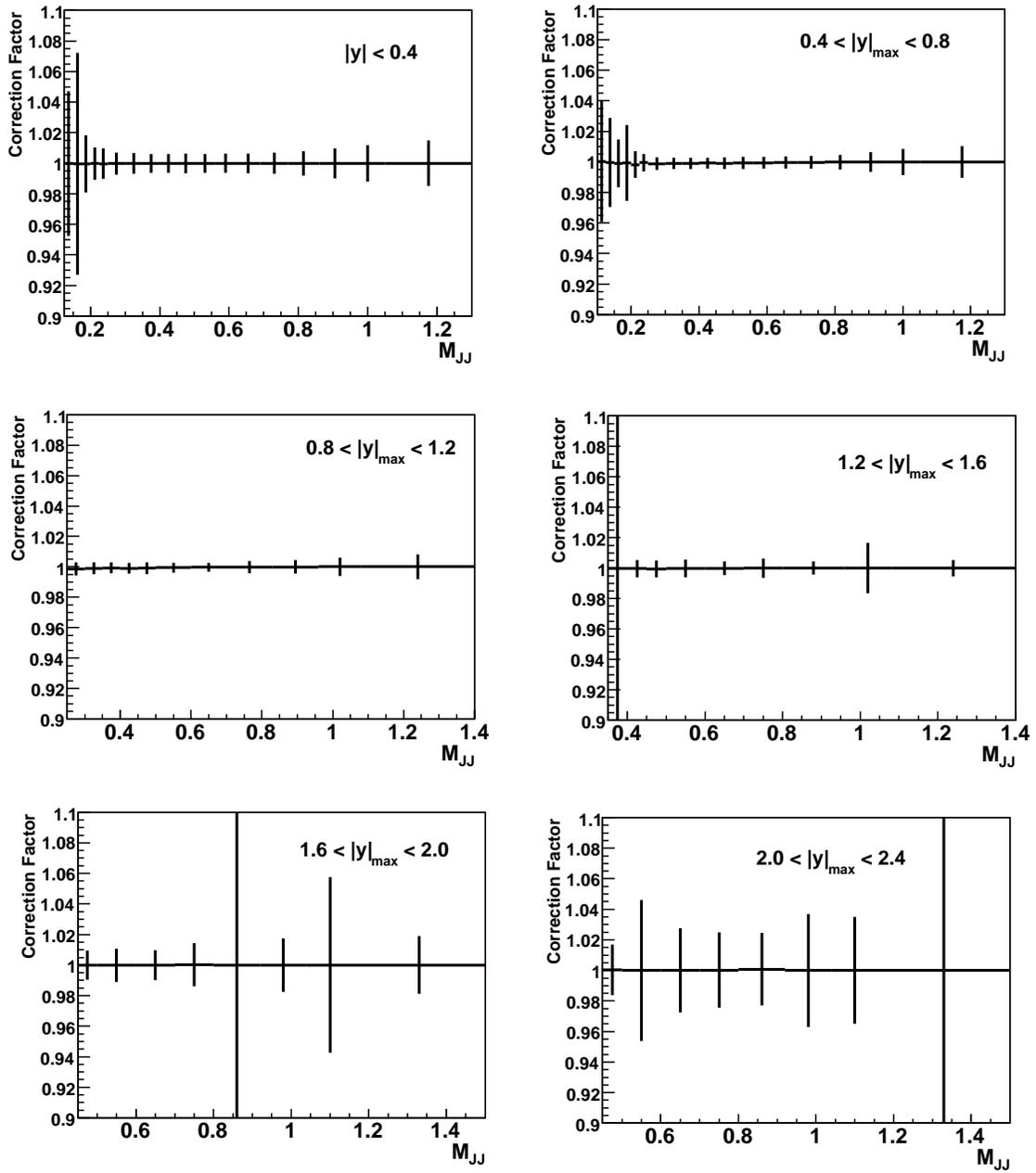
Figure H.6: Correction factor due to misvertexing .

Appendix I

J4S uncertainties

The uncertainties for the JES are composed of 49 eigenvectors. For the dijet JES (J4S), two of those uncertainties (source 007 and source 39) were found to be negligible and not included in the systematics. The table below shows the remaining 47 eigenvectors and a brief description.

Component	Description	Component	Description
jes_000	EM energy scale	jes_026	Zero suppression bias
jes_001	Dead material	jes_027	ZSb number of vertexes
jes_002	Photon energy scale	jes_028	ZSb jet matching
jes_003	Photon sample purity	jes_029	MPF method bias (MPFb)
jes_004	EM-jet background	jes_030	MPFb Pythia vs. Herwig
jes_005	High-p_T extrapolation	jes_031	MPFb scaling
jes_006	PDF uncertainty at high	jes_032	MPF jet matching p_T
jes_008	Fit in CC kRjetCCStat0	jes_033	Detector showering
jes_009	Fit in CC kRjetCCStat1	jes_034	Shw sample purity
jes_010	Fit in CC kRjetCCStat2	jes_035	Shw scaling
jes_011	η-intercalibration in CC	jes_036	Shw jet matching
jes_012	η-intercalibration in IC	jes_037	Shw template fits
jes_013	η-intercalibration in IC	jes_038	Shw Tune A vs. Tune DW
jes_014	η-intercalibration in EC	jes_040	MPFb for dijets
jes_015	η-intercalibration in EC	jes_041	MPFb for dijets
jes_016	η-intercalibration in EC	jes_042	Dijet CC response
jes_017	η-intercalibration in EC	jes_043	Dijet CC response
jes_018	JES resolution bias	jes_044	Dijet CC response
jes_019	η fit in CC	jes_045	Dijet CC response
jes_020	η fit in IC	jes_046	Inclusive jet response
jes_021	η fit in IC	jes_047	Offset
jes_022	η fit in EC	jes_048	Offset systematics
jes_023	η fit in EC	jes_049	empty placeholder
jes_024	η fit in EC		
jes_025	η fit in EC		

Table I.1: J4S systematic uncertainty eigenvectors.