ENERGY FRACTIONS OF THREE-JET EVENTS IN $P\overline{P}$ COLLISIONS AT $\sqrt{S} = 1.8$ TEV

David Fernández Connor

A DISSERTATION

in

PHYSICS

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

1991

Hugh H. Williams

Supervisor of Dissertation

H T Fortune

Graduate Group Chairperson

MATH-PHYSICS

COPYRIGHT

David Fernández Connor

1991

DEDICATION

For Francisco Fernández-Dorado (1896-1990).

ACKNOWLEDGMENTS

I thank Brig Williams for giving me the opportunity to work on CDF, one of the premier experiments in this field. Steve Hahn, Rick van Berg and John Cooper were helpful at the beginning of grad school.

I am obliged to Rob Carey who went over all the early manuscripts of this thesis and the article which will be published shortly. He also did some calculations for the article for which I am indebted. Marjorie Shapiro, John Huth and Rob Plunkett were very helpful in criticizing the analysis and polishing up the article, as well as being understanding when I was finishing the degree. Brenna Flaugher gave helpful advice when needed. I also wish to acknowledge the QCD/jet group at CDF, especially Dave Brown and Rick St. Denis, for developing an understanding of the detector without which this analysis would have been less convincing. The entire CDF collaboration merits recognition for the work that went into the detector project.

I thank the Fermilab Accelerator Division for their exceptional performance in the operation of the Tevatron and the Antiproton Source.

I thank my professors at Dickinson College for introducing me to physics, especially Priscilla Laws and Neil Wolf. Lastly I thank my family, my friends and my ultimate teammates (∅♀) for their companionship throughout graduate school. Thanks to Lucia *et al.* for a cool summer. Thanks to Matt, Bill and Laura for putting me up at the end so I could finish these last details at Penn. Special thanks go to

Bato and Kieda, Tammy and Tro, Rob, Josie, Kevin, Steph, Patrice, Greg and Rian, Shawn, and Monica.

This work was supported in part by the Department of Energy under contract number DE-ACO2-76ER03071 and other contracts; the National Science Foundation; Istituto Nazionale di Fisica Nucleare, Italy; and the Ministry of Science, Culture and Education of Japan.

ABSTRACT

ENERGY FRACTIONS OF THREE-JET EVENTS IN P$\overline{\text{P}}$ COLLISIONS AT $\sqrt{S} = 1.8$ TEV

David Fernández Connor

Dissertation Supervisor: H. H. Williams

The leading-order calculation of Quantum Chromodynamics (QCD) for the production of three-parton final states was tested by investigating three-jet events at the Fermilab Tevatron at $\sqrt{s} = 1.8$ TeV. The integrated luminosity used for this study was 3.9 pb^{-1}. The three jets were Lorentz boosted to their center-of-mass (COM) frame, and the fraction of the total COM energy each jet carried was calculated. There were 4 826 events after cuts were made to eliminate detector bias. The distributions of the energy fractions were found to be inconsistent with a three-body phase space decay model and consistent with QCD.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the past twenty years a successful theory of the structure of hadrons and how they interact with each other has been developed. The Quantum Chromodynamics (QCD) parton model holds that all hadrons (*e.g.* pions and protons) are composed of partons (quarks and gluons). The strong force governs interactions between partons and binds partons together inside hadrons. When two hadrons interact at low energies (below a few GeV), their interaction is governed by the force between two dipoles of strong force charge (this is similar to the Van der Waals force between two electrically neutral atoms). This interaction can be modeled by the exchange of a quark-antiquark pair such as a meson. In highly energetic inelastic collisions, the interaction is dominated by the "hard" scatter of a parton from one of the hadrons with a parton from the other hadron. The strong coupling is weak at high energies ($\alpha_s \approx 0.2$ for jet $E_t >$ 10 GeV), so cross sections may be calculated with QCD using a Born approximation perturbation expansion, and then compared to experimental results. The outgoing scattered partons cannot exist in free space, so they are converted into several hadrons and are observed as collimated streams of particles called *jets*. The general properties of QCD, such as the existence of the gluon and its spin, have been verified[1, 2]. Tests of QCD are now being refined in precision. This thesis is concerned with one such test, the production of three energetic partons in the final state.

The topology of 2 → 3 processes, where three partons are produced from the collision of two partons, can be parameterized by the fraction of the total energy each final-state parton has, along with the orientation of the three-jet plane. There is a well understood QCD calculation for these processes which can be used to predict the differential cross sections of the energy fractions[3]. To compare the QCD calculation with data, two effects must be taken into account: (1) the transformation of the final-state partons into jets and (2) the detector's resolution and non-uniformity. Both these effects are well modeled for a high energy jet but become less understood as p_t, the jet momentum transverse to the beam, falls below 10 GeV. Therefore all three jets must be required to have at least 10 GeV p_t. However, a set of cuts depending solely on the jet p_t's will introduce biases in the energy fraction distributions. These biases would be amplified since the differential cross section falls precipitously as an inverse power of p_t for any single jet. Many reconstructed parton energies would be too high due to statistical fluctuations from the two aforementioned effects. (It might also be too low but the slope of the p_t spectrum would make upward fluctuations in energy more common.)

To alleviate this problem, cuts were made on the parameters of the three-jet event topology (the energy fractions and the orientation of the three-jet plane) and the invariant three-jet mass. These three-jet cuts were made tight enough so that for practically all events that pass these cuts, all three jets had a transverse momentum above 10 GeV. Thus biases from the single jet-p_t cuts were avoided. Although there are also resolution effects from the three-jet cuts, they are not as severe as the effects arising from the single jet-p_t cuts since: (1) the distributions of the three-jet topology variables are not as steep as the p_t spectra and (2) to leading order the shape of the differential cross sections of the energy fractions do not depend on the invariant three-jet mass cut.

This method of three-jet event analysis was applied to data taken by the Collider Detector at Fermilab (CDF) from the fall of 1988 to the spring of 1989. The total integrated luminosity of this data run was 4.6 pb^{-1}. The principles of this method had been developed by the UA1 collaboration at the Sp$\bar{p}$S collider at CERN[4]. The method resembles three-jet analyses done earlier at e^+e^- machines[5]. A similar analysis had also been done at the CERN Intersecting Storage Rings (ISR)[6]. The energy fraction distributions at these different $p\bar{p}$ collision energies are reproduced qualitatively by QCD. This thesis will present a quantitative comparison between experimental results from CDF and theory.

Chapter 2

Theory

2.1 The QCD Parton Model

The QCD parton model is a subset of the Standard Model[7], the theory which so far best describes natural phenomena at high energies. The Standard Model has twelve fermions listed in table 2.1. Both leptons and quarks interact via the electroweak force. Quarks also interact via the strong force. The electroweak force is mediated by four bosons, γ, W^+, W^- and Z^0, and governs electromagnetic and weak interactions.

The strong force is mediated by the octet of gluons which couples to the six quarks, each quark being a triplet. Each component of the triplet has a unit of strong force

Leptons	ν_e	ν_μ	ν_τ
	e^-	μ^-	τ^-
Quarks	u	c	t
	d	s	b

Table 2.1: Fermions in the Standard Model

charge. The three components are called *color* and are labeled r, g, and b (for red, green, and blue).

Quarks are constituents of visible hadrons. A meson (a hadron which is a boson) contains a quark-antiquark pair. The color of the quark is matched to the anti-color of the antiquark (*i.e.* if the quark has a red color charge than the antiquark has an anti-red color charge). Baryons (hadrons which are fermions) contain three quarks, each quark with a color different from the other two. The wave functions of the physical hadrons are arranged such that, with respect to color, mesons are symmetric and baryons are antisymmetric. Both mesons and baryons are symmetric with respect to the combination of flavor, spin and orbital angular momentum. The flavors of the quarks or antiquarks and their orbital states determine the species of hadron. This model adequately reproduces the spins and parities of hadrons and is consistent with the spin-statistics theorem. Some models which incorporate the masses of the quarks and a model of the forces inside a hadron at rest can also reproduce the masses of most hadrons. Only the proton will be considered in this thesis. It contains two up quarks and one down quark in the lowest energy orbital states with the spins of the quarks aligned such that the total spin of the proton is $\frac{1}{2}\hbar$.

The gluon is a color octet, each component having a color and an anticolor. The structure of the color octet in terms of the original color triplet is

$$r\bar{b}, r\bar{g}, b\bar{g}, b\bar{r}, g\bar{r}, g\bar{b}, \frac{r\bar{r} - b\bar{b}}{\sqrt{2}}, \frac{r\bar{r} + b\bar{b} - 2g\bar{g}}{\sqrt{6}}. \tag{2.1}$$

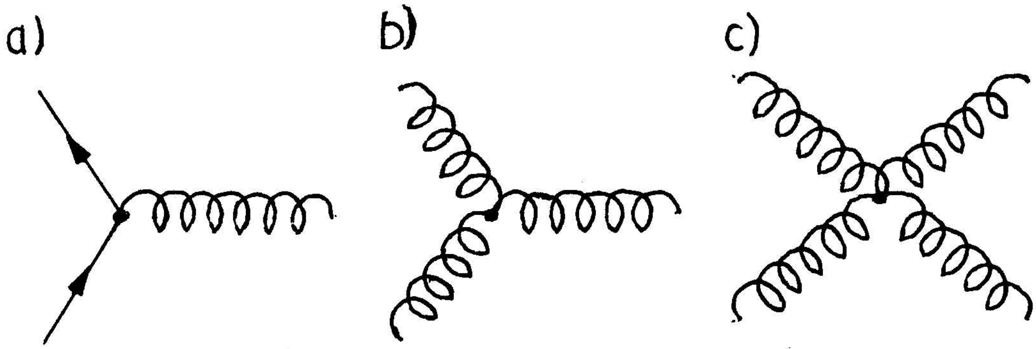

Figure 2.1: QCD couplings for gluons and quarks. a) shows the quark-gluon coupling, b) shows the gluon triple vertex, and c) shows the gluon quartic vertex.

The strong force is governed by the QCD Lagrangian:

$$\mathcal{L}_{QCD} = -\frac{1}{4} F_a^{\mu\nu} F_{a\mu\nu} + \bar{\psi}_j (i\gamma_\mu D_{jk}^\mu - M_{jk})\psi_k \tag{2.2}$$

where

$$D_{jk}^\mu = \delta_{jk}\partial^\mu + ig_s\,(T_a)_{jk} G_a^\mu \tag{2.3}$$

and

$$F_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f_{abc} G_b^\mu G_c^\nu. \tag{2.4}$$

G represents the gluon field, ψ represents the quark fields. The indices $j, k = 1, 2, 3$ refer to the color triplet r, g, and b. The indices $a, b, c = 1, ..., 8$ refer to the color octet in equation 2.1. M_{jk} is the quark mass matrix which is independent of color. g_s is the strong coupling constant and $(T_a)_{jk}$ are the SU(3) generators. This Lagrangian has a quark-gluon coupling shown in figure 2.1a and a gluon self coupling shown in figures 2.1b and 2.1c.

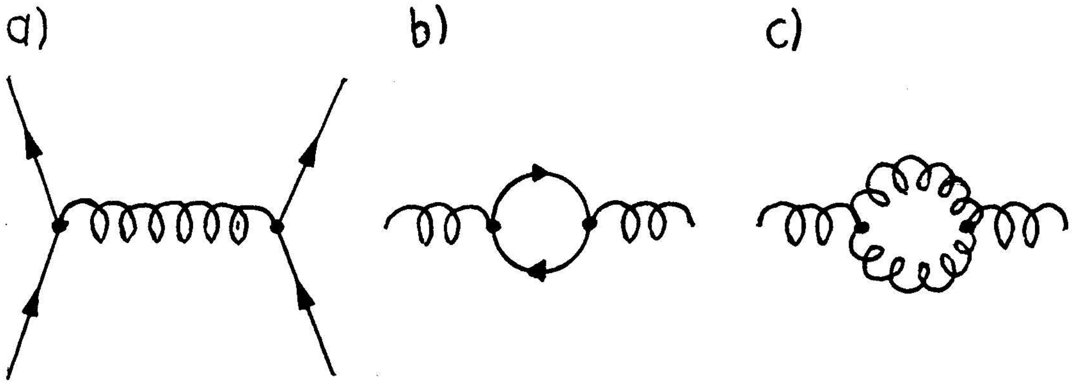

Figure 2.2: Quark-quark interaction. a) is the leading-order Feynman diagram for the QCD interaction between two quarks. Higher order diagrams have modifications to the gluon propagator shown in b) and c).

The gluon self-coupling, which does not have an analog in QED, has far-reaching implications. This can be demonstrated by considering the leading-order quark-quark interaction shown in figure 2.2a. Higher order diagrams include modifications to the gluon propagator term, shown in figures 2.2b and 2.2c. Effects from figure 2.2b will screen the color charge of one quark felt by the other. Effects from figure 2.2c, which arise from the gluon self-coupling, will anti-screen this charge, *i.e.* amplify the effective strong force interaction. The gluon self-coupling terms dominate, thereby making the force of the interaction increase as the distance between the quarks increases, or equivalently, as the momentum of the gluon decreases. A simple perturbation series in $\alpha_s = g_s^2/4\pi$ will not converge because of the contribution from the gluon self-coupling terms. This will be true even for high energy interactions.

The solution is to replace the coupling constant α_s by a *running coupling* $\alpha_s(Q^2)$ which is a function of Q^2, usually assigned to a dynamical variable on the same scale as the momentum of the gluon propagator. A calculation of the running coupling includes gluon and quark self-energy diagrams and vertex diagrams. As $Q^2 \to \infty$ the result is

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2f)\ln(Q^2/\Lambda^2)} \tag{2.5}$$

where f is the number of quark flavors with masses below Q (typically five) and Λ is a scale parameter which is typically 0.2 GeV. This equation is valid when Q is much higher than Λ. So in high energy collisions between quarks (or gluons), a perturbation series in $\alpha_s(Q^2)$ will not diverge. This is called *asymptotic freedom.*

At low energies, equation 2.5 is not valid, and the interaction cannot be calculated with a perturbation series. The interactions between quarks inside hadrons are an example of this. The attractive force would become very strong if a quark tried to leave the hadron. So quarks are effectively confined inside the hadron unless knocked out by a high energy collision. The impossibility of the quark becoming free is called *confinement.*

2.2 p$\bar{\text{p}}$ Collisions

A $p\bar{p}$ collision which produces particles with high momentum transverse to the beam will be dominated by a hard scatter between a parton in the proton and a parton in the antiproton (see figure 2.3). To the first and second leading order the other partons in the proton or antiproton do not participate in the hard process. These partons are

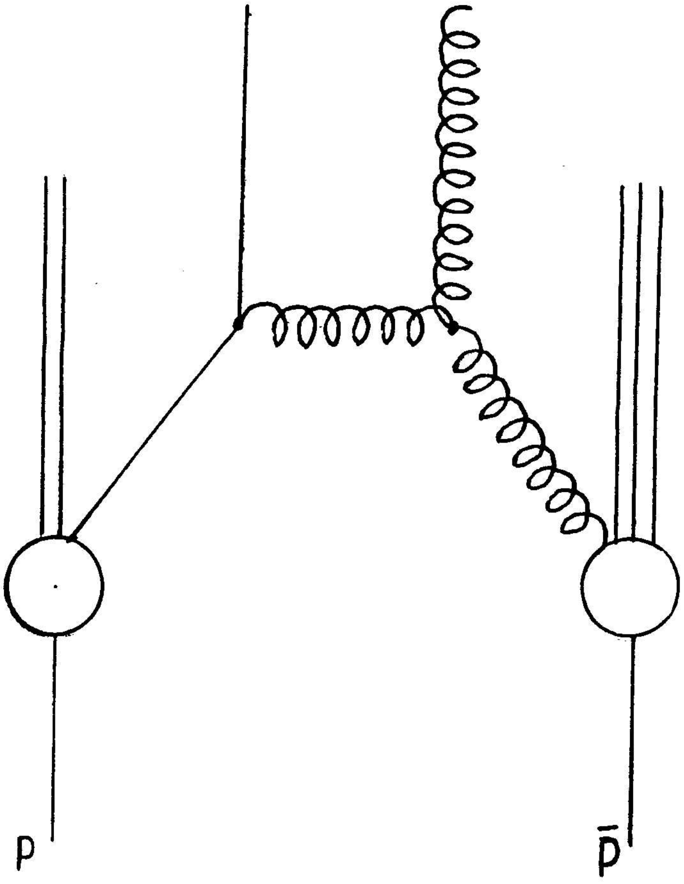

Figure 2.3: Proton-antiproton collisions. The collision is dominated by the hard scatter involving two partons, one from each hadron.

referred to as *spectators.* A parton involved in the hard scatter can be either one of the three quarks which determine the hadron species (*valence quarks*), or a gluon or a member of a quark-antiquark pair. The last two contributions are considered to come from virtual partons which bind the three valence quarks together in the proton; in a hard collision these virtual partons can interact with other partons.

If the energy of the hard scatter is much greater than the energy which binds the parton inside the proton, one can use the *impulse approximation* to calculate the cross section of the $p\bar{p}$ interaction. The partons are assumed to be free and to have negligible transverse momentum. The momentum of the partons are given by momentum distributions which depend only on x, the fraction of the momentum of the proton that the parton carries. These momentum distributions are deduced empirically from deep inelastic scattering of hadrons with each other and with leptons.

Before the collision the partons were tightly bound inside the proton because of confinement. After the scattering takes place the partons are subject to confinement once more. As the distance between the proton remnants increases the color force becomes stronger. This force, while not strong enough to contain the partons inside the original proton, can create quark-antiquark or gluon pairs. An outgoing parton can combine with these and create a color singlet state. For the scattered partons this state manifests itself as a jet, a stream of hadrons with nearly collinear momenta.

The spectator partons will appear in the detector as low p_t particles. Most of the energy from these particles, referred to collectively as the *underlying event*, disappears down the beam pipe. However some energy may appear in the detector, increasing the apparent energy of the jets coming from the hard scattering. It is practically impossible to remove this energy on a jet by jet basis but its average effect can be measured and subtracted from each jet, the penalty being an increase in the systematic uncertainty of the jet's energy.

The hard scatter of the interaction $ab \rightarrow f$ where a and b are the initial partons and f is the final state can be calculated using QCD field theory. However the identities of the initial partons cannot be selected or inferred event by event. Therefore one must calculate the cross section for the interaction $p\bar{p} \rightarrow fX$ using the formula

$$d\sigma(p\bar{p} \rightarrow fX) = \sum_{ab} \iint dx_p \, dx_{\bar{p}} \, f_{a/p}(x_p) f_{b/\bar{p}}(x_{\bar{p}}) d\hat{\sigma}(ab \rightarrow f) \qquad (2.6)$$

where X represents the underlying event and $f_{a/p}(x_p)$ and $f_{b/\bar{p}}(x_{\bar{p}})$ are the parton momentum distributions giving the probability that a parton of type a (b) has a fraction x_p ($x_{\bar{p}}$) of the proton (antiproton) momentum. a and b are summed over all possible combinations of quarks and gluons which can produce the final state f. The term $d\hat{\sigma}(ab \rightarrow f)$ is the differential cross section of the subprocess which produces the final state f from the initial-state partons a and b; it is derived from QCD.

With equation 2.6, one can calculate the $2 \rightarrow 3$ QCD matrix element and the resulting three-jet cross section. The final state f is taken to be three outgoing partons. The differential cross sections from all $2 \rightarrow 3$ subprocesses are summed

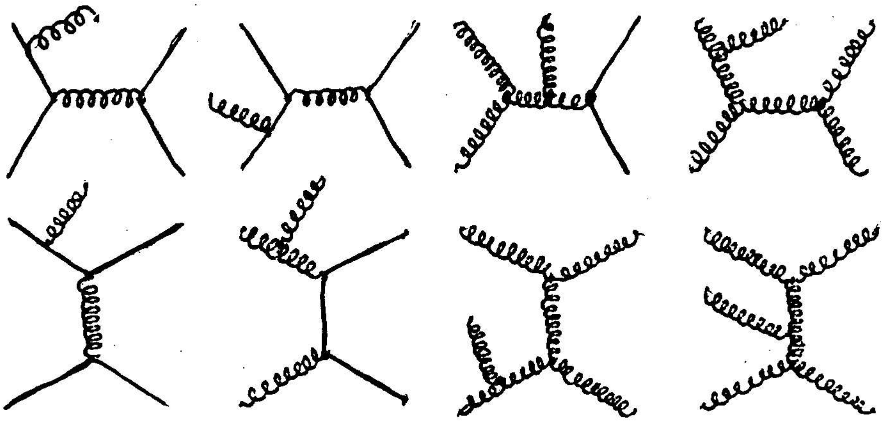

Figure 2.4: Examples of 2 → 3 Feynman diagrams.

and integrated according to equation 2.6. The individual subprocess differential cross sections have been calculated to leading order in α_s by Berends *et al.*[3]. The Feynman diagrams for some of these subprocesses are shown in figure 2.4. The theoretical three-jet cross section can be tested by observing three-jet events at hadron colliders.

2.3 Three-Jet Kinematics and Dynamics

A three-jet process has five partons: two in the initial state (before the hard scatter has occurred) and three in the final state (after the hard scatter). Each final-state parton is identified with one of the three jets. The four-momentum of a final-state parton is not identical to the four-momentum of its corresponding jet, but the two can be related. Following the convention of references [3] and [4], the initial-state partons are labelled 1 and 2, 1 being the faster parton in the lab frame, and the final-state partons are labelled 3 through 5 in order of decreasing energy as measured in the center-of-mass (COM) frame.

Twelve parameters are required to describe a three-jet system in the lab frame since there are three four-momenta. Three of these parameters specify the Lorentz boost from the lab frame to the COM frame. The COM frame of the three jets is not the same as the lab frame due to the variation in initial-state parton momenta and low energy final-state QCD radiation (see section 5.3.1). The former is responsible for the component of the Lorentz boost along the beam axis and the latter for the component transverse to the beam axis. If the Lorentz boost is mostly along the beam axis, it can be specified by a single parameter y^{boost}, which is given by the expression

$$y^{boost} = \frac{1}{2} \ln \frac{x_1}{x_2}. \tag{2.7}$$

This approximation is valid if $y^{boost} \gtrsim 1$. The advantage of making this approximation is that the rapidity of a jet in the COM frame can be calculated by simply adding y^{boost} to the rapidity in the lab frame.

The other nine parameters of the three-jet system specify the properties of the three jets in the COM frame. Of these nine, three are the invariant masses of each jet, three give the angular orientation of the three jets, and three specify the total COM energy and how it is shared among the jets.

The three angles, which are closely related to the Euler angles used to specify the orientation of a rigid body, are θ^*, ψ^*, and ϕ^*. θ^* is the angle between parton-3 and parton-1. ψ^* is the angle between the plane of the three final-state partons and the

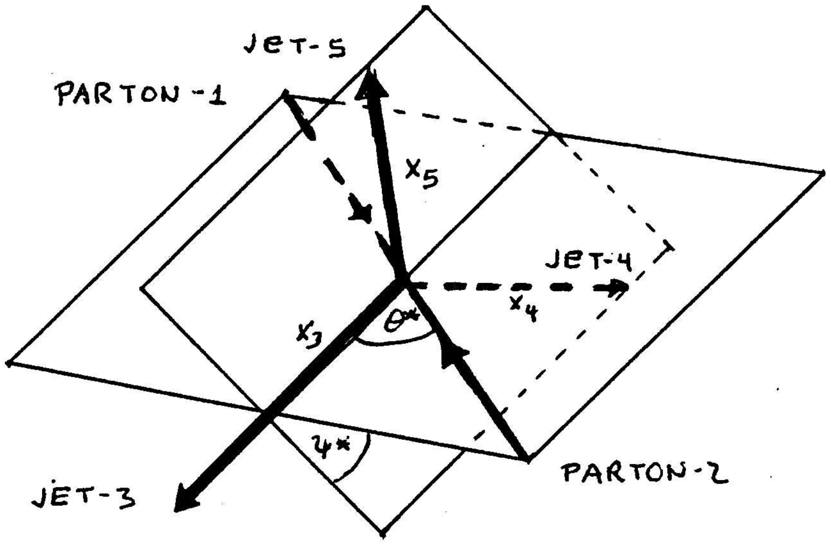

Figure 2.5: Diagram of the three-jet variables in the three-jet COM frame (taken from reference [4]).

plane described by parton-1 and parton-3. ϕ^* is the azimuthal angle of parton-3; it has no dynamic importance.

M_{3j} is the invariant mass or total COM energy of the three partons, which equals the subprocess energy $\sqrt{\hat{s}}$ if there are no more than three jets. The final-state parton energy fractions are x_3, x_4, and x_5 where

$$x_i = \frac{2E_i}{M_{3j}}. \tag{2.8}$$

x_3 varies between $\frac{2}{3}$ and 1, x_4 between $\frac{1}{2}$ and 1, and x_5 between 0 and $\frac{2}{3}$. Naturally, specifying x_3 and x_4 fixes all three energy fractions since $x_3 + x_4 + x_5 = 2$. The angles θ^* and ψ^* and the energy fractions x_3 and x_4 are shown schematically in figure 2.5.

The final leading order 2 → 3 cross section can be represented as

$$\frac{d\hat{\sigma}(2 \rightarrow 3)}{dx_3\, dx_4\, d\psi^*\, d\cos\theta^*} = \frac{1}{4\pi}\frac{[\alpha_s(Q^2)]^3}{\hat{s}}|M'|^2 \tag{2.9}$$

where $|M'|^2$ depends only on x_3, x_4, ψ^* and $\cos\theta^*$ and not on $\alpha_s(Q^2)$ nor $\hat{s}$[4]. The shapes of the single differential cross-sections depend on the Q^2-value selection and the parton structure functions only through the denominator $\hat{s}$ (coupled with M' for different subprocesses) and $\alpha_s(Q^2)$. This dependence is assumed to be negligible.

This analysis will be concerned with the variables x_3 and x_4. A scatter plot (Dalitz plot) of x_4 versus x_3 should be uniformly occupied, if three-jet events are produced by a process which fills phase space equally, such as an isotropic three-body decay. However, because the three-jet matrix element grows very large as $x_5 \rightarrow 0$, the so-called infrared singularity, one expects to see an enhancement of events with large x_3 and x_4.

Chapter 3

Apparatus

The Collider Detector at Fermilab (CDF)[8] detects collisions of protons with antiprotons produced by the Fermilab Tevatron I project. The total energy of the collisions in the lab frame is 1800 GeV. The lab frame is the same as the center-of-mass frame of the proton-antiproton system, *i.e.* the energies of the proton and and antiproton are equal. The two hadron bunches collide with a frequency of 300 000 Hz.

CDF is a general purpose hybrid detector with particle tracking chambers near the interaction region surrounded by calorimetry which is in turn surrounded by a muon detection system (see figure 3.1). Of these three major detector components only the calorimetry approximates total coverage over the entire solid angle. In addition there are time-of-flight counters with limited coverage near the beam pipe 6.5 m away from the interaction region on both sides. All these components are read out by both a data acquisition system (DAQ) and a trigger system. The DAQ digitizes the detector signals and assembles the data to be written on magnetic computer tape. The trigger

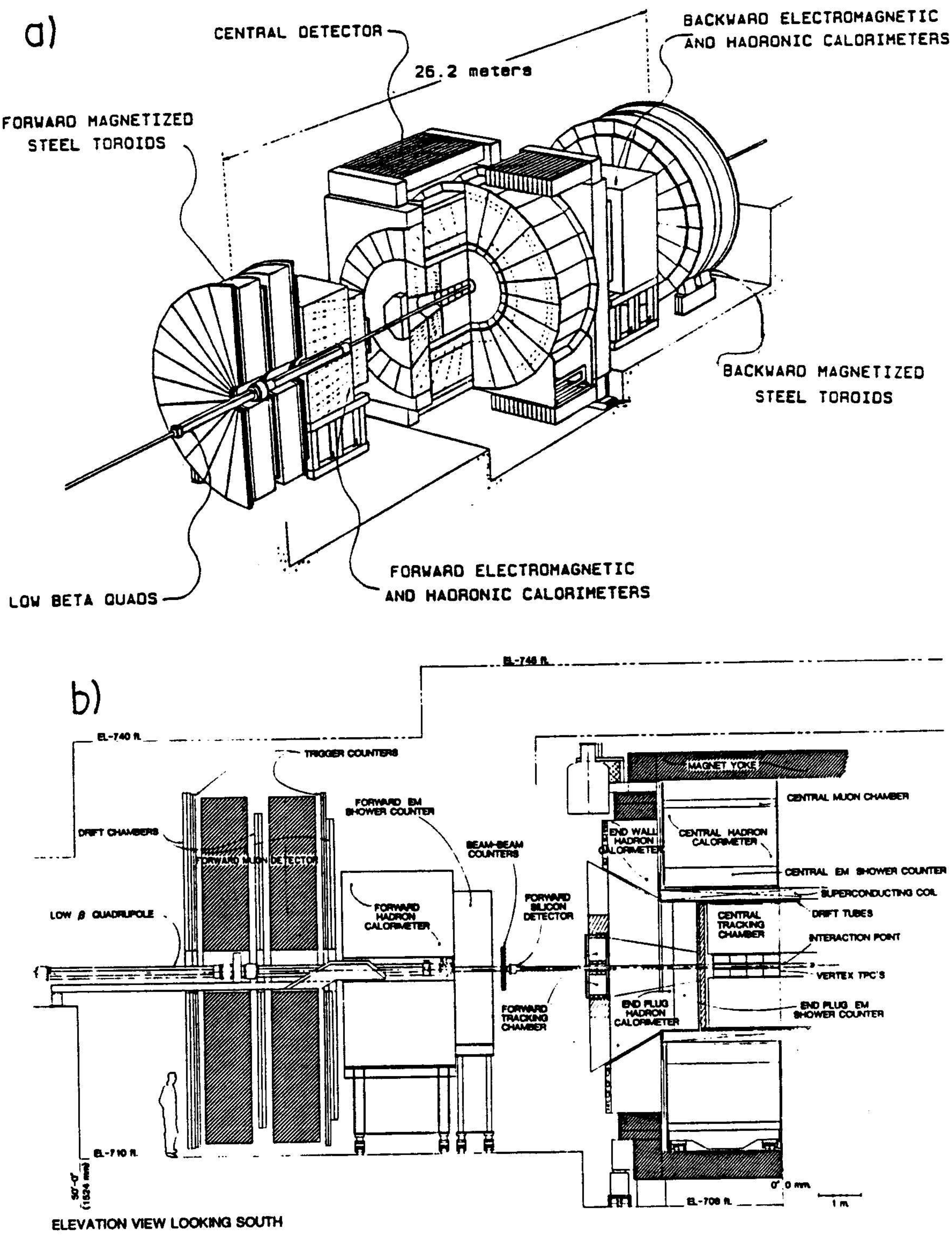

Figure 3.1: Two views of the CDF detector: a) a perspective cutaway view, b) a cross sectional view through one half ($\eta_D > 0$).

analyzes the data online, with both analog and digital processing, and decides which events should be written to tape.

CDF has been used for many studies of jet physics, such as the measurement of the inclusive jet E_t spectra. It has also been used to measure the properties of the intermediate vector bosons through electronic and muonic decays, the direct photon cross section, and bottom physics. It was recently used to search for the top quark. Only the components of the detector which affected the three-jet analysis will be described. These are the trigger, the central tracking systems and the calorimetry. A more complete description of the detector can be found in reference [8].

The coordinate system used for CDF has the z-axis tangent to the proton beam, the x-axis pointing towards the center of the synchroton, and the y-axis pointing up. The proton beam travels in the positive z-direction when in the interaction region, the antiproton beam travels in the opposite direction. Looking from above, the protons travel in a clockwise direction.

3.1 Beam-Beam Counters (BBC)

One must identify a sample of events of which the only requirement is that an inelastic collision has occurred. This can be determined by looking for the existence of particle in very forward angles of the detector. Not all inelastic collisions will satisfy this criterion, but most will. There is a bias introduced by this cut but it is very small.

Hence this sample of events is referred to as the *minimum bias* sample.

The Beam-Beam Counters (BBC) are used to define which events pass the minimum bias cut. The BBC consists of two groups of scintillator paddles located 6.5 m from the nominal interaction point along the beam axis, one group on either side. The paddles are arranged in a square $\sim$ 1 m on a side around the beam pipe (see figure 3.2). Along the vertical or horizontal axis, they cover the pseudo-rapidity range 3.24 to 5.90.

The BBC measures the time of the interaction with an uncertainty of 200 ps. The minimum bias trigger requires a BBC signal within 15 ns of the beam crossing time in both groups of counters. This trigger, the Level 0 trigger, starts the online analysis by the trigger system described in section 3.6. Besides the physics triggers, a sample of events is written to tape with the sole requirement of passing the minimum bias trigger. This minimum bias sample is used to look for physics at low particle energies, as well as giving a sample of events which have isolated low-energy particles used for calorimeter response studies.

3.2 Vertex Time Projection Chamber (VTPC)

The interaction point was measured for each event by the vertex time projection chamber system (VTPC)[9]. A cut was made on the z-position of the interaction point to keep all jets inside the fiducial region of the calorimeter. The VTPC is centered

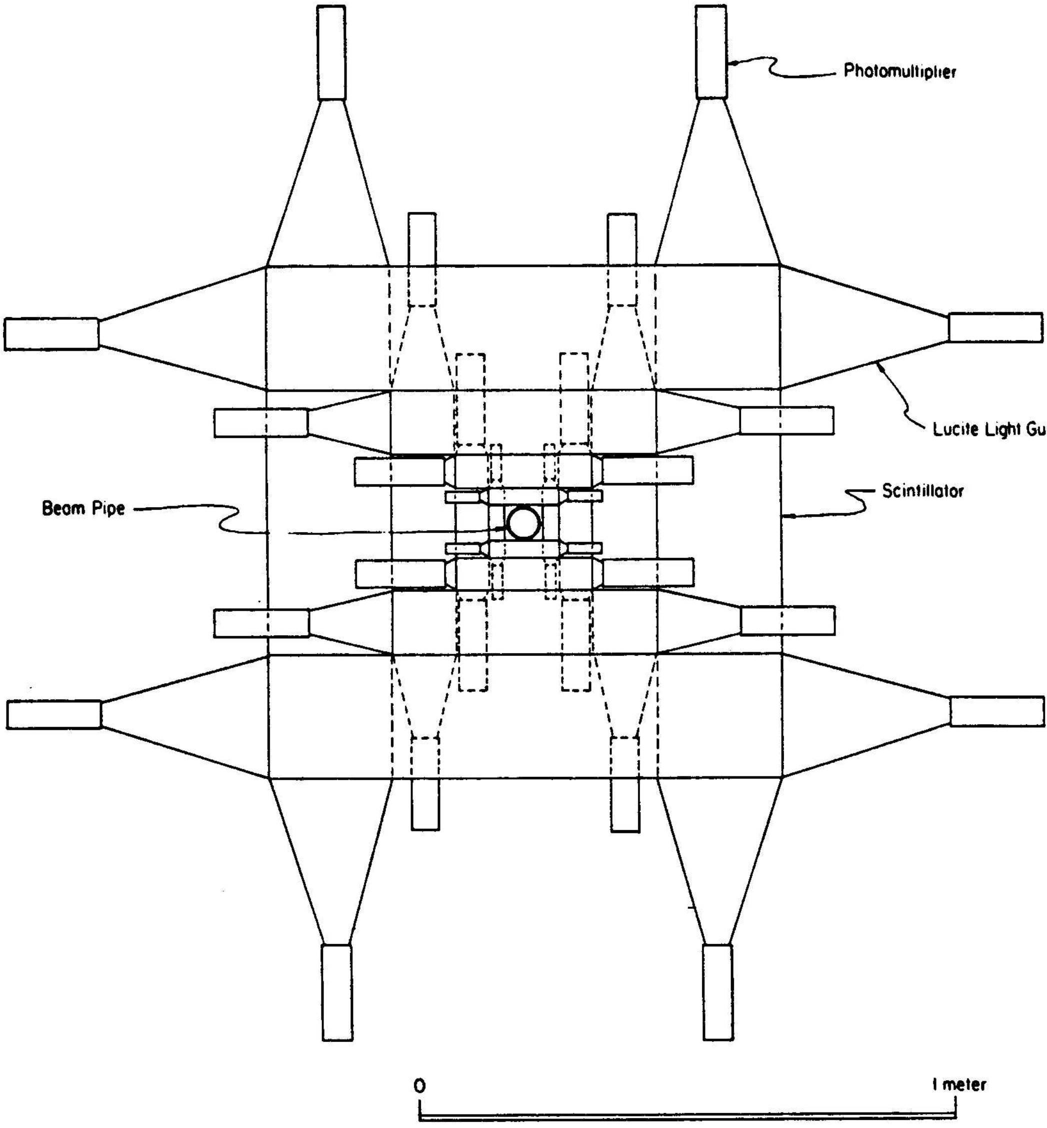

Figure 3.2: A beam's eye view of one of the two beam-beam counter (BBC) planes.

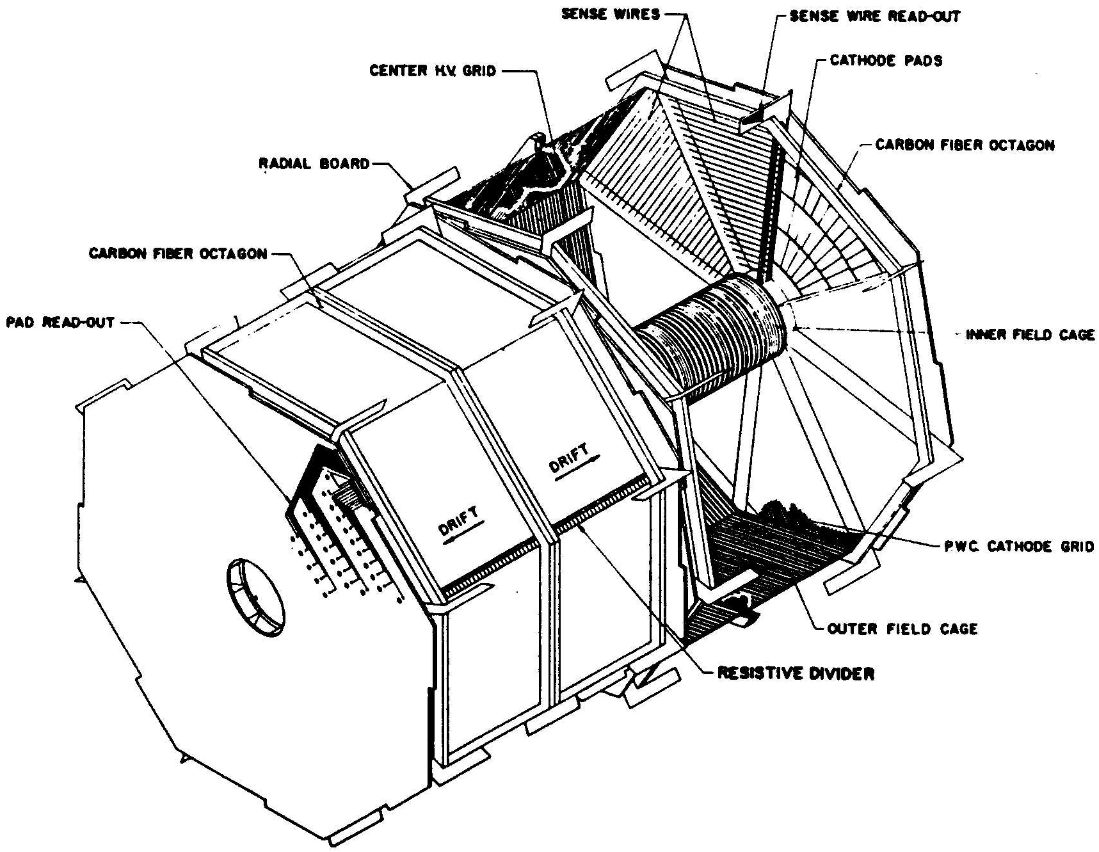

Figure 3.3: Two of the eight vertex time projection chamber (VTPC) modules.

at the nominal interaction point and surrounds the beam pipe for $|z| < 1.4$ m. The distribution of the interaction points along the beam axis is Gaussian with a standard deviation of 35 cm. The interaction region, defined to be the area where an interaction vertex can be, is well covered by the VTPC.

The VTPC is divided up into eight octagonal modules distributed along the beam pipe; each module is rotated 22.5° along ϕ with respect to its neighbors (see figure 3.3). Each module is divided up into two drift regions with the endcaps divided into octants. Each octant has 24 anode wires which measure r-z and 24 cathode pads which measure

r-ϕ. All anode wires were connected to the DAQ but only two modules had the cathode pads connected. The spatial uncertainty per hit in the r-z plane is at most 500 μm. The VTPC measures the z-position of the interaction point with an accuracy of 1 mm.

3.3 Central Tracking Chamber(CTC)

The central tracking chamber (CTC)[10] surrounds the VTPC and is itself surrounded by a solenoid[11] providing an axial magnetic field of 1.412 T. It is used to measure the transverse momentum of charged particles. In the angular region $30° < \theta < 150°$ ($|\eta| < 1.3$) the momentum resolution of the CTC is $dp_t/p_t^2 \leq 0.002$ GeV^{-1}.

The CTC is cylindrical with an axis along the beam line. It has an outer radius of 1.3 m, an inner radius of 0.27 m, and a length of 3.2 m. The chamber has 84 layers of sense wires which are grouped into 9 "superlayers". Five axial superlayers each consist of 12 sense wires parallel to the beam axis; four stereo superlayers each consist of six sense wires tilted by 3° relative to the beam axis. The type of superlayer alternates: the inner most superlayer has 12 wires, the next inner most has six, *etc.* One of the two endplates of the CTC is shown in figure 3.4. The spatial precision per hit is 200 μm in r-ϕ.

The CTC was used in two ways for this analysis: to measure the low-energy response of the calorimeter and to check the detector simulation. The low-energy

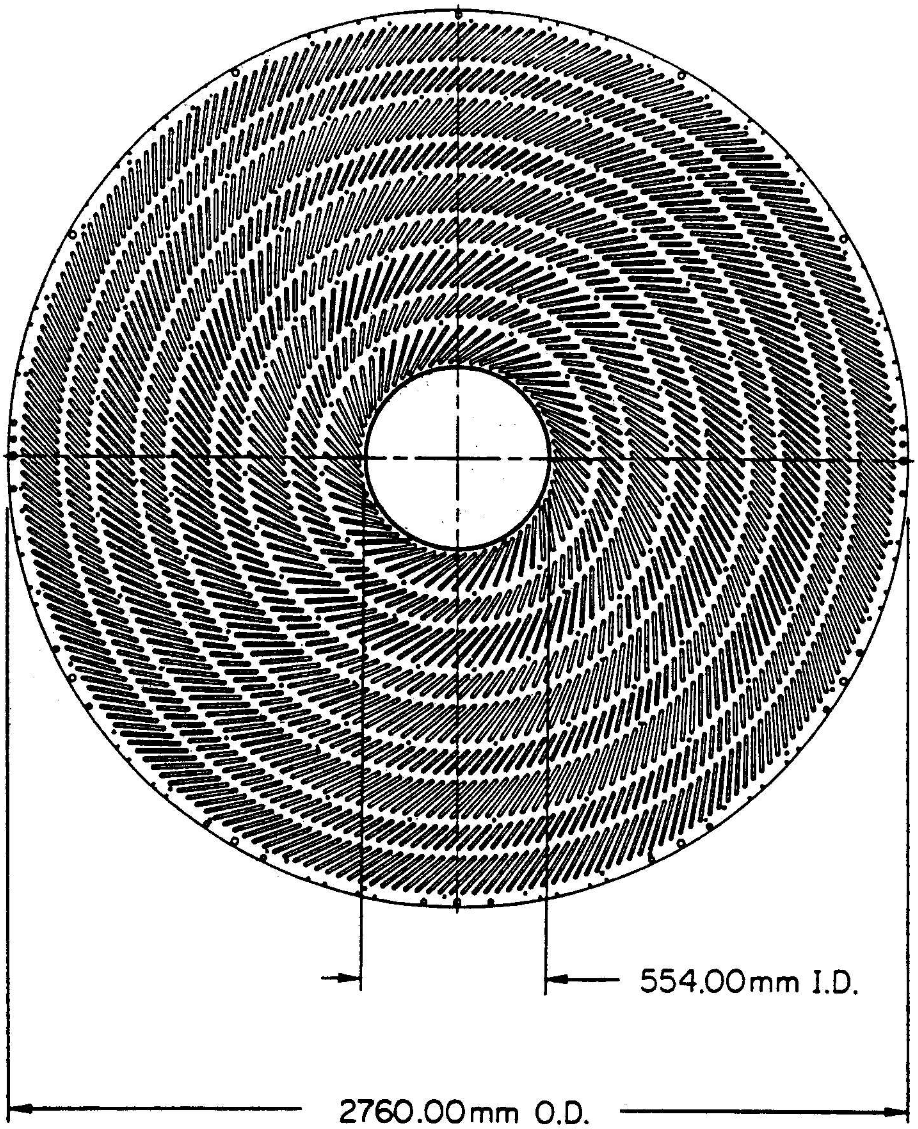

Figure 3.4: An endplate of the central tracking chamber (CTC).

response of the calorimetry was not measured outside the collision hall since the testbeam could not produce particles with low enough energy. Therefore a sample of particles with known momenta was taken from $p\bar{p}$ collisions. The momenta of the particles was measured by the CTC. A full description of this analysis is in section 3.4.4.

The three-jet event sample was simulated in some detail by a physics generation program coupled with a detector simulation program. This calculation is the subject of chapter 7. As part of the checks on this calculation, the four-momenta of particles inside the jets were studied. The CTC was used to measure individual particle momenta inside a jet.

3.4 Calorimetry

The CDF calorimeters were the most important detector subsystems used in this analysis. Jets were reconstructed and their four-momenta determined solely from their energy depositions in the calorimeters. There are four calorimeters in all: the central, endwall, endplug and forward. Each is composed of projective towers segmented in η and ϕ. The towers are approximately 0.1 wide in units of pseudo-rapidity in all regions. They are 15° wide in azimuth in the central region and 5° wide in the plug and forward regions (see figure 3.5). The central and endwall calorimeters cover the range $0 < |\eta_D| < 1.1$, the plug calorimeters cover the range $1.0 < |\eta_D| < 2.4$, and

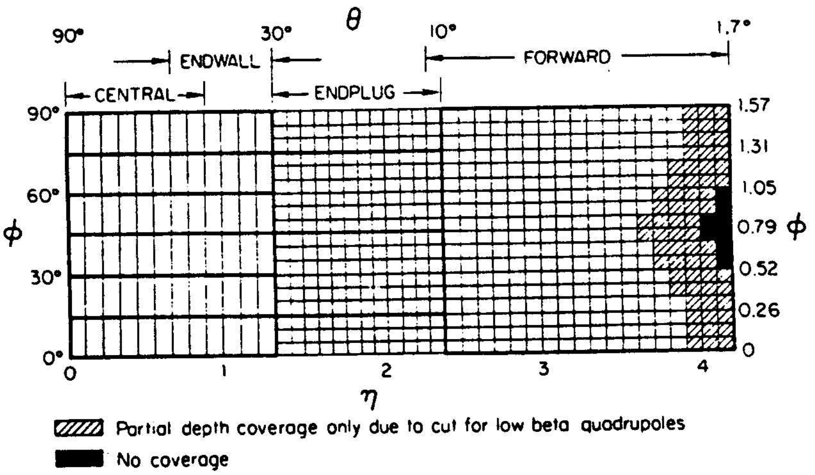

Figure 3.5: Hadron calorimeter towers in one of eight identical quadrants ($\Delta\phi = 90°$, $\eta_D > 0$). The heavy lines indicate module or chamber boundaries. The EM calorimeter towers are the same except that the central EM extends to $\eta_D \leq 1.1$, the plug EM extends to $\eta_D \geq 1.0$ and the forward EM has full coverage out to $\eta_D = 4.2$.

the forward calorimeters cover the range $2.4 < |\eta_D| < 4.2$ ($\eta_D \equiv \ln \cot \theta/2$). Each of these components will be described in detail below.

3.4.1 Central Calorimetry (CEM,CHA,WHA)

The electromagnetic part of the central calorimetry (CEM), which covers the range $0 < |\eta_D| < 1.1$ consists of alternating layers of lead and plastic scintillator. The lead layers are 0.125 in. thick and are clad in 0.015 in. thick aluminum. The scintillator is SCSN-38 polystyrene and is 5 mm thick. At $\eta_D = 0$, there are 31 layers of scintillator interspersed with 30 layers of lead. A normal line to the calorimeter face intersects the beamline at a 90° angle. Therefore any particles detected at $|\eta_D| > 0$ will hit the calorimeter at an oblique angle, increasing the amount of material the particle passes through while interacting with the scintillator. To keep the number of radiation lengths independent of angle, layers of scintillator and lead were substituted with acrylic. The number of radiation lengths the particle has to traverse in the CEM is ~ 18 while the number of absorption lengths is ~ 1. There is a proportional chamber with cathode strip readout (called a *strip chamber*) embedded inside the CEM at a depth of five radiation lengths. It was used to measure the position of EM showers caused by isolated electrons or photons but was not used in this analysis.

Behind the CEM are two iron-scintillator hadron calorimeters, the central hadron (CHA) and endwall (WHA), which cover the ranges $0 < |\eta_D| < 0.9$ and $0.6 < |\eta_D| < 1.1$ respectively. A CHA section wholly contains six towers and shares three towers with a WHA section which has an additional three towers. A figure of the CHA-WHA layout is shown in figure 3.6. The CHA has 32 alternating layers of steel and scintillator. The steel is 2.5 cm thick and the scintillator is 1 cm. The WHA has 15 such layers except the steel is 5 cm thick. The scintillator type is PMMA doped with 8% naphtalene, 1% butyl-PBD and 0.01% POPOP.

Each tower in the CEM, CHA, or WHA is observed by two photomultiplier tubes (PMT's), one on each azimuthal side of the tower. Light coming out of the scintillators is channeled to the PMT's by a series of waveshifters and light pipes.

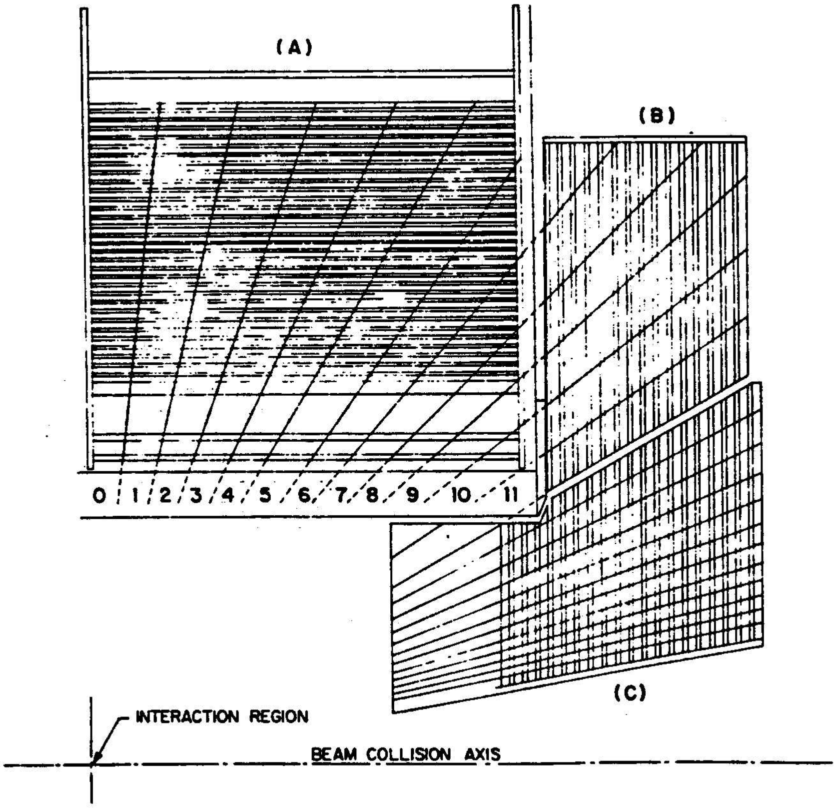

Figure 3.6: Quadrant of the calorimeter where A, B, C show central, endwall, and plug respectively. Towers are numbered from 0 (at 90° in polar direction) to 11 (last tower of the endwall modules). Hadronic towers 6, 7, and 8 are shared between the central and endwall calorimeter.

3.4.2 CEM Calibration

The CEM and CHA together form a mechanical unit which is divided into 48 modules for the detector plus two spares. Due to their shape, modules are referred to as wedges. Each wedge has ten towers of the CEM and eight towers of the CHA, with each tower covering 0.1 units of pseudo-rapidity and 15° of azimuth. The wedges were put into a test beam before the collider run and were exposed to an electron beam of 50 GeV in the center of each tower. The response of the CEM to these electrons was measured along with the response to a ^{137}Cs source.

The source is attached to a loop of nylon coated stainless steel wire which, through a series of pulleys and tubes, passes through both the CEM and the CHA. The source is moved into the CEM and CHA by turning the pulleys with an electric motor. The speed and direction of the motor, as well as its on-off state, are controlled by computer logic circuits located next to the amplifiers and digitizers. These are in turn controlled by the DAQ described below. The source, while it moves through the CEM, is located near *shower maximum*, where most of the energy from a single electron's shower is deposited. The source moves across the face of every tower, passing near the η-ϕ center. While the source is moving the response is measured three times a second. The maximum response with the source in a tower is taken to be the source measurement for that tower. While it is not being used the source is stored inside a lead brick located between the CHA and the WHA support.

Given the response of the CEM to the electron beam and the source, along with the gains of the amplifiers, one can find out the response of the calorimeter inside the collision hall. This is accomplished by exposing the calorimeter once again to the sources and calibrating the amplifiers. Therefore the response of the calorimeter is given by the equation

$$G_E(t) = \frac{G_Q(t)}{G_Q(t_0)} \frac{I(t_0)}{I(t)} G_E(t_0). \tag{3.1}$$

Here $G_E(t)$ represents the *energy gain*, the CEM response in the collision hall in units of GeV per ADC count, and $G_E(t_0)$ is the response to the 50 GeV electron beam in the same units. The measured currents $I(t)$ and $I(t_0)$ are the responses of the calorimeter to the ^{137}Cs sources in the collision hall and the testbeam respectively; they are in units of nA. Likewise $G_Q(t)$ and $G_Q(t_0)$ are the electronic gains of the amplifiers in the collision hall and the testbeam. A different set of amplifiers are used for measuring the CEM response to a high energy electron ($E \gg 500$ MeV) and the electron or photon coming from the ^{137}Cs source ($E \approx 0.5$ MeV). Therefore, it is necessary to account for the gain of the electronics used to measure electron response, G_Q, and the gain of the amplifiers used to measure source response, which are taken into account in the variable I.

The CEM calibration was tested several ways. First, several wedges were calibrated twice in the testbeam and the results of the first calibration were compared with the successive calibration on the same wedge[12]. Figure 3.7 shows a histogram

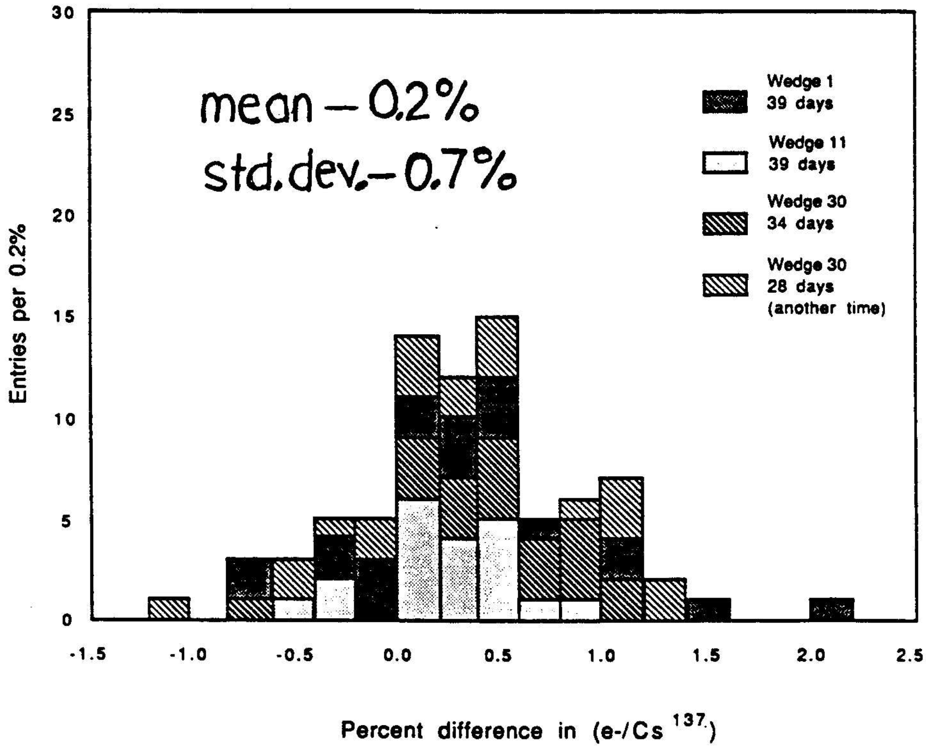

Figure 3.7: Calibration reproducibility. The difference in beam to ^{137}Cs source ratio for each tube is plotted for successive calibration procedures, about 5 weeks apart, for three modules. The deviation of the centroid from zero corresponds accurately to the source decay.

of the comparison. The standard deviation σ comes from the reproducibility of the calibration. This σ should be divided by $\sqrt{2}$ to obtain the uncertainty in the calibration, since each entry really represents two measurements. This shows the calibration is good to within 0.4% over a period of about one month.

Second, a sample of electrons from $p\bar{p}$ data was taken and the response of one PMT was divided by the other for each electron observed. This ratio is called left-over-right (L/R). The tower-by-tower mean of the ratio of responses, $\langle \ln(L/R) \rangle$, is plotted in figure 3.8 with the attenuation length of the scintillator taken into account. The error in the mean for most towers was about 1% (stat.), but any tower with this error above 1.5% (*i.e.* with low statistics) was not plotted, leaving 370 towers. The standard deviation, 2.56%, is the sum in quadrature of the error of the means, known to be 1%, and any effect coming from the reproducibility, which therefore must be 2.3%. As with the first method, each entry is really two measurements. Therefore the calibration in each tower is good to 1.6%, not accounting for effects which affect both PMT's equally.

Third, a sample of electrons which are decay products of $W^{\pm}$'s were used to compare the response of the calorimeter with the momentum of the electron as measured by the CTC. Thus the CEM is tested with another component of the detector independent of the calorimeter. The comparison is parameterized by the quantity E/p, shown in figure 3.9. Because of bremsstrahlung, the electron will emit photons before

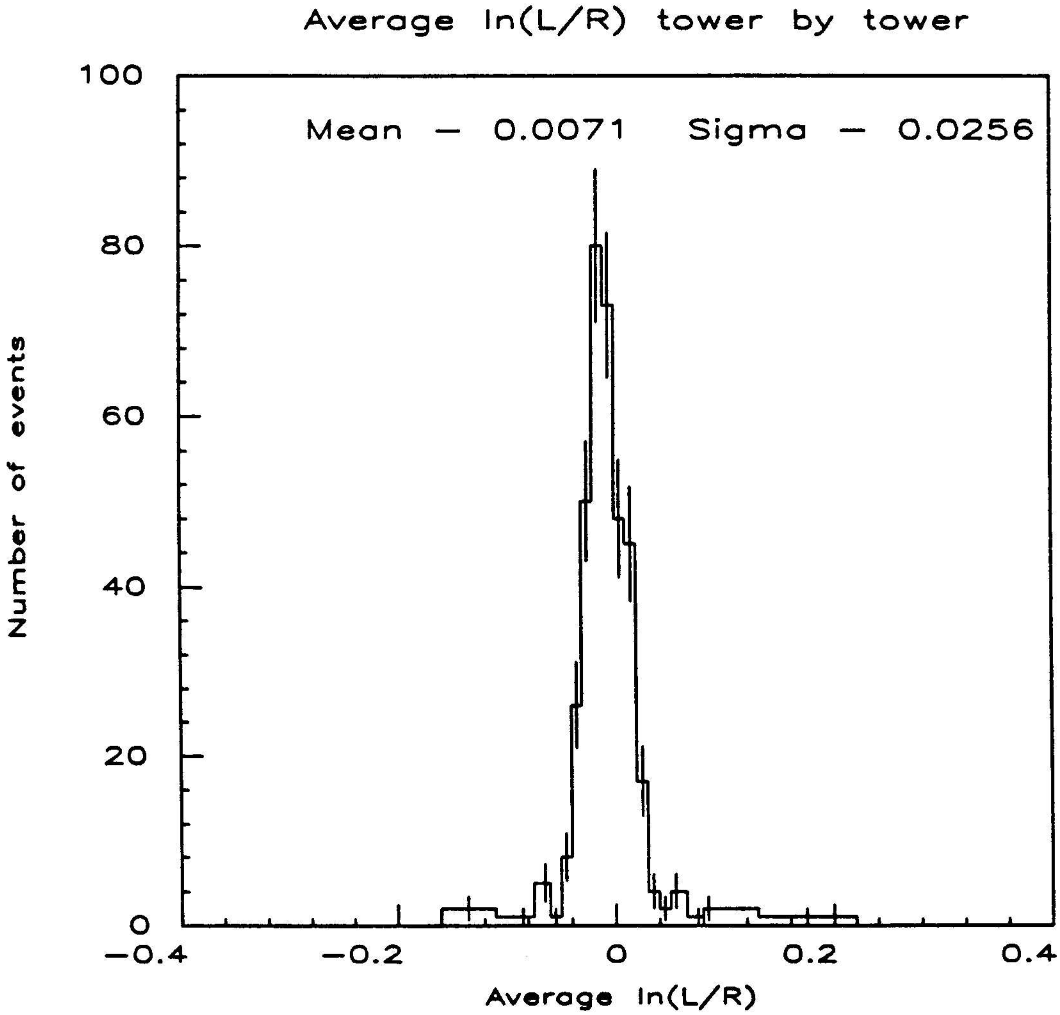

Figure 3.8: Average $\ln(L/R)$ for central EM calorimeter towers. The attenuation length of the scintillator is taken into account.

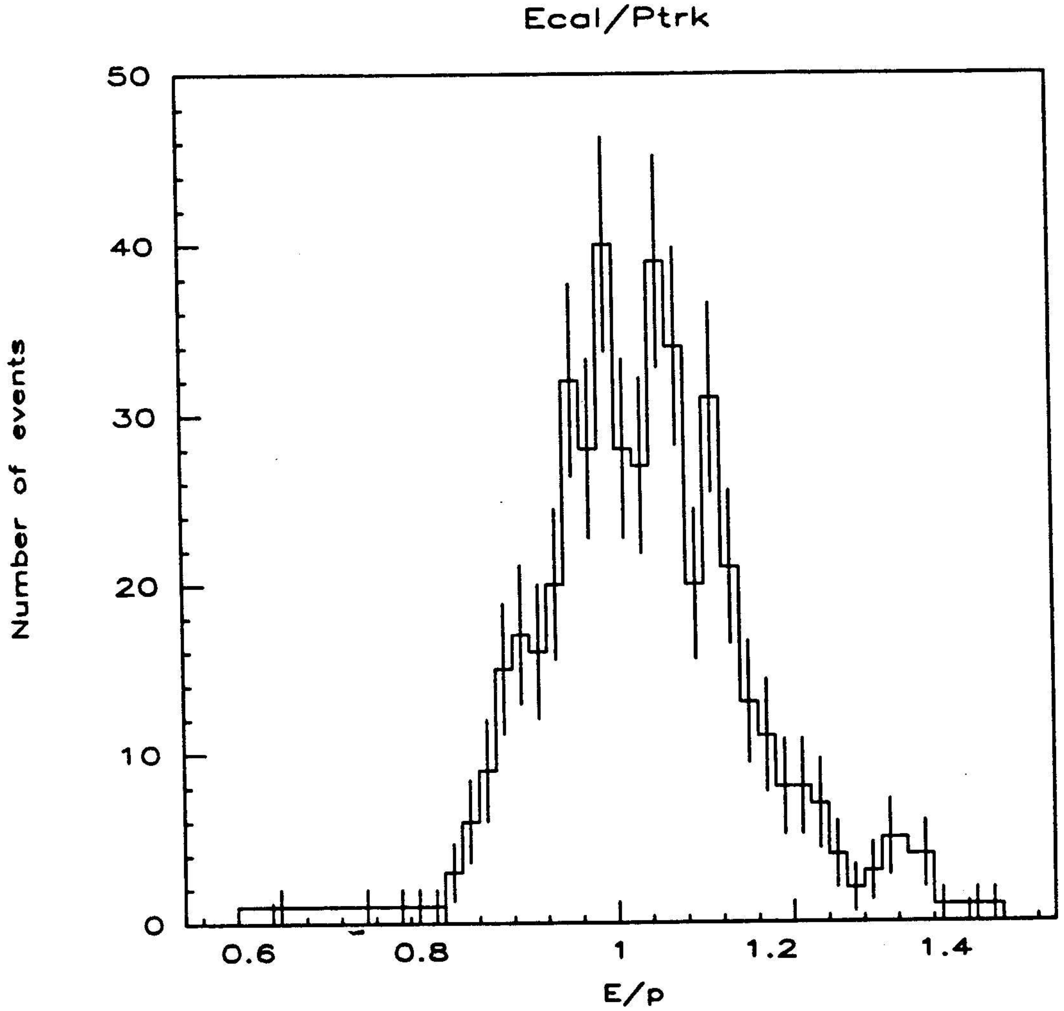

Figure 3.9: E/p—the energy as measured by the calorimeter divided by the momentum as measured by the CTC—for electrons coming from $W^{\pm}$ bosons. The average E/p for events with $E/p \leq 1.4$ is 1.0304 ± 0.0053(stat.), which compares favorably with the theoretical value of 1.0248 ± 0.0012(stat. and syst.).

passing through the CTC. The energy of these photons will usually be detected by the CEM (provided it is not a wide angle bremsstrahlung) but will not be included in the momentum measurement by the CTC. This will cause a shift of about 2% in the plot, along with a high end tail. For all events with $E/p \leq 1.4$, the average E/p is 1.0304 ± 0.0053 (the uncertainty is purely statistical). This is comparable to the theoretical value of 1.0248 ± 0.0012 (systematic and statistical uncertainty), thereby confirming that the CTC and the CEM agree in energy scale to within $\approx 0.5\%$. The theoretical value was derived using a Monte Carlo calculation which takes into account the CTC track momentum resolution, the calorimeter resolution and bremsstrahlung of the electron[13].

3.4.3 CHA and WHA Calibration

The CHA was calibrated in a way similar to the CEM calibration. All CHA towers in all wedges were exposed to a 50 GeV charged pion beam aimed at the tower centers. Only those pions which deposited less than 2 GeV in the CEM are used for the CHA calibration (these are referred to as *minimum ionizing pions*). The response of the CHA to the minimum ionizing pions was measured along with the response to the same ^{137}Cs sources used with the CEM. The CHA calibration was maintained in the collision hall with the ^{137}Cs source using the same method as the CEM calibration. The reproducibility of the CHA calibration was determined by

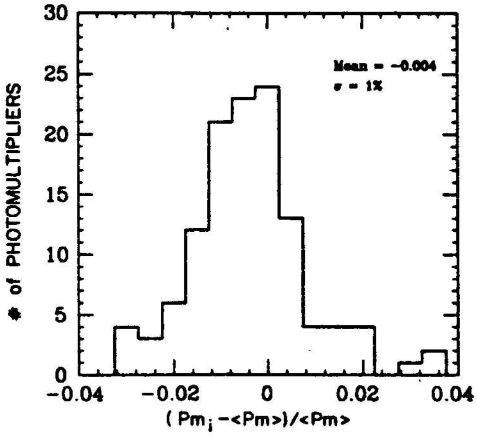

Figure 3.10: Reproducibility of the central hadronic calorimeter calibration plotted as the percentage difference between corresponding pion peaks of 7 recalibrated modules. The peak values are normalized to the reference ^{137}Cs source responses.

calibrating seven wedges in the test beam twice[14], each pair of calibrations was done approximately one month apart. The difference in the successive calibrations averages to zero, and the width of the distribution shows that the calibration in each tower is good to about 0.7% (see figure 3.10). In addition, one wedge (a spare not used in the collision hall) was recalibrated at a different test beam three years after its initial calibration. Only the 12 PMT's which belong to towers not shared with the WHA were considered. The ^{137}Cs source calibration was found to be 1.9% lower than the pion response at the new test beam, with a spread of 1.7%. The systematic uncertainty in the average pion beam momentum was 2% at the new test beam, which may account for the higher pion response. The spread indicates the ^{137}Cs calibration is good to 1.2%, barring effects in the collision hall.

Two WHA modules were calibrated in the testbeam with 50 GeV minimum ionizing pions. This calibration was transferred to the other 48 modules with the use of one ^{137}Cs *line source* (as opposed to a point source). The radioactive material is uniformly distributed inside a 0.5 mm-diameter stainless steel tube. The line source can be inserted into two tubes on either azimuthal side of the tower, each tube centered in polar angle. When the line source is inserted, it simultaneously illuminates all scintillator layers of that tower. The tower's response to the line source is used to maintain the calibration just before data taking. As with some modules of the CHA and CEM, the two WHA modules in the test beam were exposed to beam twice[14]. The line source reproduced the response of each tower to within 2%.

The line source must be inserted by hand, and calibrating all 48 WHA modules with the line source is time-consuming ($\sim$ 1 day). Therefore it cannot be used while taking data since access to the collision hall is restricted. A mounted ^{137}Cs source controlled by a drive mechanism similar to that of the CHA-CEM was installed in each WHA module. During the data taking run, the moving source is used to maintain the calibration.

3.4.4 Non-linear Energy Response of the Central Calorimeters

The calibration above was pegged to 50 GeV electrons and pions. However the calorimeters may not have a linear response to the pion energy. This can affect the jet correction functions described in chapter 5. The non-linearity must be accounted for in the detector simulation program used to determine the jet corrections. Therefore, the non-linearity was measured by exposing the calorimeter to pion beams of different energy and measuring the response. Unfortunately the testbeam line could not produce pions below an energy of 10 GeV. This poses a problem since jets have on the order of ten particles and can have as many as twenty or thirty (see figure 7.2), so even a 50 GeV jet will have many particles below 10 GeV. Therefore, the energy response below 10 GeV must be determined from another source instead of the testbeam.

To address this issue a sample of charged tracks was taken from the minimum bias sample as well as a sample of events in which the trigger required a stiff track in the CTC[15]. All of these charged tracks are assumed to be pions. The range of pion momenta from this sample is $0.75 \leq E \leq 20$ GeV. Each track was required to be well reconstructed in three dimensions, have the impact parameter relative to the beam position be less than 0.5 cm, have the closest point of approach to the z-axis be within 5 cm of the event vertex, have hits on more than 50% of the layers, and have two or more good axial segments (8 or more wires used out of 12) and one or

3.4.4 Non-linear Energy Response of the Central Calorimeters

The calibration above was pegged to 50 GeV electrons and pions. However the calorimeters may not have a linear response to the pion energy. This can affect the jet correction functions described in chapter 5. The non-linearity must be accounted for in the detector simulation program used to determine the jet corrections. Therefore, the non-linearity was measured by exposing the calorimeter to pion beams of different energy and measuring the response. Unfortunately the testbeam line could not produce pions below an energy of 10 GeV. This poses a problem since jets have on the order of ten particles and can have as many as twenty or thirty (see figure 7.2), so even a 50 GeV jet will have many particles below 10 GeV. Therefore, the energy response below 10 GeV must be determined from another source instead of the testbeam.

To address this issue a sample of charged tracks was taken from the minimum bias sample as well as a sample of events in which the trigger required a stiff track in the CTC[15]. All of these charged tracks are assumed to be pions. The range of pion momenta from this sample is $0.75 \leq E \leq 20$ GeV. Each track was required to be well reconstructed in three dimensions, have the impact parameter relative to the beam position be less than 0.5 cm, have the closest point of approach to the z-axis be within 5 cm of the event vertex, have hits on more than 50% of the layers, and have two or more good axial segments (8 or more wires used out of 12) and one or more good stereo segments (4 or more wires out of 6). Only tracks which hit towers with $\eta_D < 0.8$ are used. The track must extrapolate into the central rectangle of the tower which takes up 36% of the area (0.6×0.6) both at the inner and outer radius of the CEM. This was done to minimize the pion shower leaking into the adjacent CEM towers. The calorimeter energy of the track was defined to be the total energy measured by the target CEM tower, plus the corresponding CHA tower, plus the eight CHA towers which share a side or a corner with the target CEM tower. The momentum of the track was measured by the CTC.

Neutral pions may accompany a charged pion. They will be detected only in the CEM since they will immediately decay into two photons. This would add to the calorimetry energy of the charged pion. To reduce this background, the surrounding eight CEM towers are not included in the energy sum. The average of the ratio of the calorimeter energy to the track momentum, E/p, as a function of p is shown in figure 3.11, with the neutral pion background energy subtracted bin-by-bin. The neutral pion background was estimated by looking at the eight CEM towers surrounding the target CEM tower. The average fractional energy per surrounding CEM tower, $\langle E_{background}\rangle/p$, was plotted versus the track momentum p and fitted to a function used for the background subtraction. The proportional energy response at low energies is 40% lower than at high energies. Also shown is the response as measured by the test beam, and the response of the detector simulation described in section 5.3.2. The

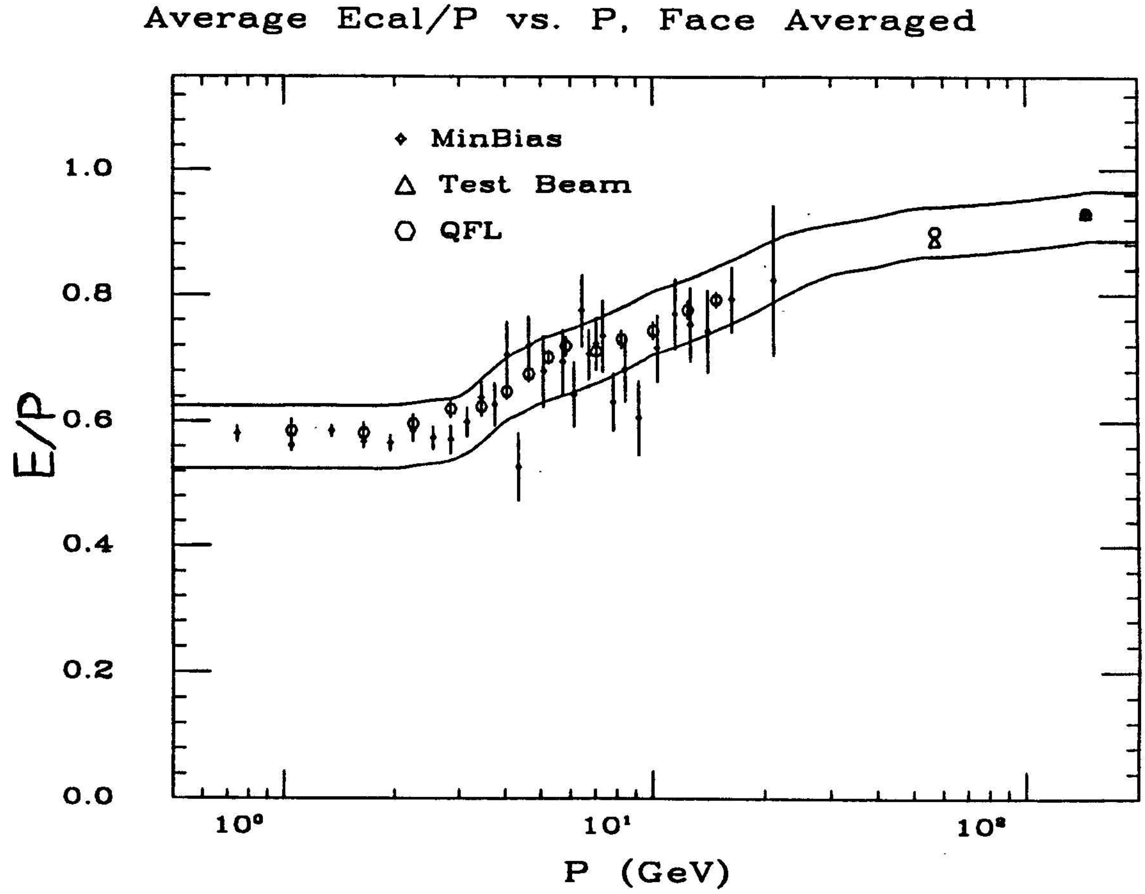

Figure 3.11: E/p for charged pions in minimum bias events, along with the response as measured by the testbeam, and the response of the detector simulation described in section 5.3.2. Curves above and below the points represent the systematic error associated with the minimum bias and test beam measurements.

simulation reproduces the pion response over the range of p that was investigated, as it was designed to do. This simulation was used both to determine the jet corrections described in chapter 5 and the three-jet QCD calculation described in chapter 7.

Another effect which must be understood is the fraction of charged pions which hit the central region of a tower but deposit little or no energy in the calorimetry. This is a function of the pseudo-rapidity of the pion because of the varying amount of dead region (uninstrumented interaction lengths) before the CEM. Figure 3.12 shows the fraction of tracks which leave less than 15% of their momentum in the calorimetry for tracks in the momentum range $1 \leq p \leq 2$ GeV. This fraction is divided by the amount of uninstrumented interaction lengths before the CEM. The curve is consistent with being flat, therefore this effect was assumed to come from the dead region. The undetected energy effect was incorporated into the detector simulation described in section 5.3.2.

3.4.5 Plug and Forward Calorimetry (PEM,PHA,FEM,FHA)

The electromagnetic (hadronic) calorimeters in the plug ($1.1 < |\eta| < 2.4$), and forward ($2.4 < |\eta| < 3.5$) regions consist of alternating layers of lead (steel) and gas proportional tubes. The plug EM calorimeter (PEM) has 34 layers of lead which are each 0.27 cm thick. These are interspersed with proportional tubes which are 0.7×0.7 cm^2. The plug hadronic calorimeter (PHA) has 20 layers of steel which are

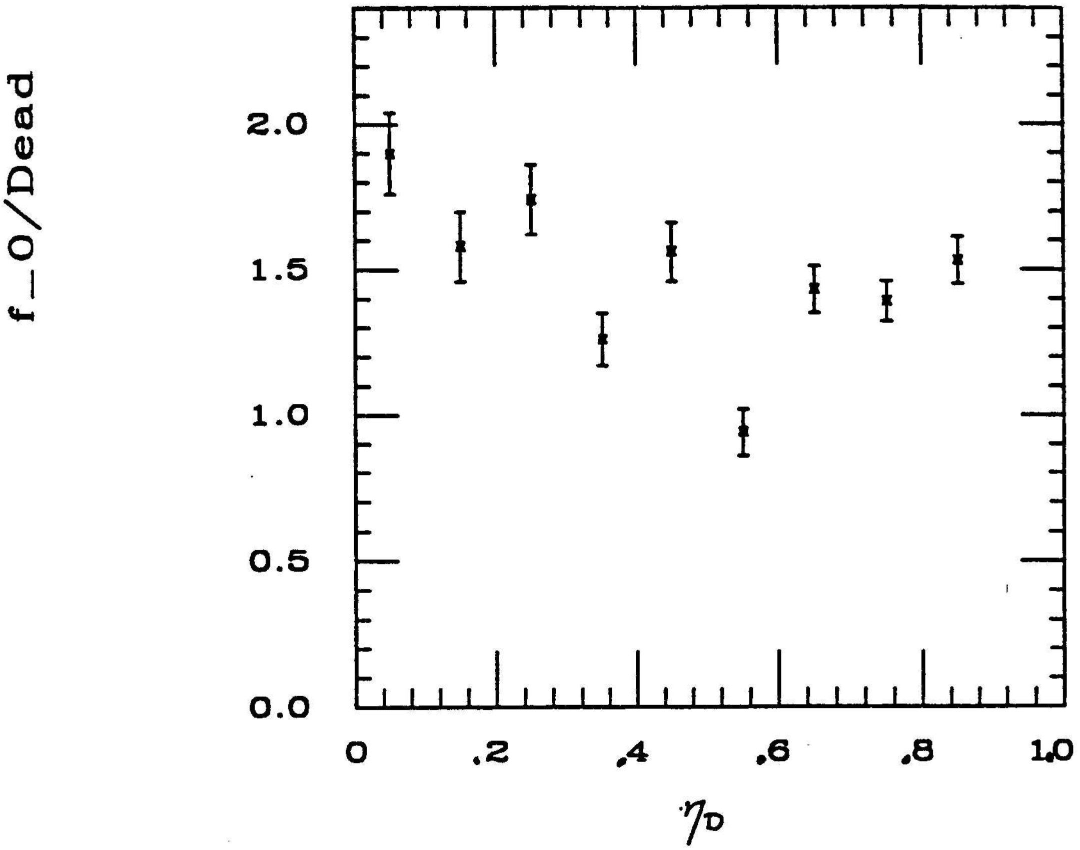

Figure 3.12: Fraction of pions tracks which do not deposit energy in the calorimetry. This is divided by the number of uninstrumented interaction lengths (the dead region) the pion must traverse before entering the calorimeter. The function is flat with respect to η_D, indicating that the probability of a pion not being observed in the calorimeter is strongly correlated with the size of the dead region.

5.1 cm thick. These are interspersed with tubes 1.4 cm thick and 0.8 cm wide. The forward EM calorimeter (FEM) has 30 layers of lead-antimony (94%-6%) 0.48 cm thick interspersed with tubes 1.0 cm thick. The forward hadronic calorimeter (FHA) has 27 layers of steel interspersed with tubes 1.5 cm thick. All gas calorimeters contain a mixture of 50% argon and 50% ethane gas with a small percentage of alcohol added to prevent glow discharge. Each calorimeter was calibrated in the test beam with 100 GeV electrons or pions. To keep track of the gain variations which are produced by changes in pressure and temperature of the gas, the response of *gas gain* tubes to ^{55}Fe was observed both during testbeam running and data taking. Gas gain tubes are proportional tubes located outside the calorimeter supplied by the same gas feed system as the calorimeter.

3.5 Data Acquisition System (DAQ)

Calorimeter responses were amplified and digitized with electronics mounted on the detector. A similar system was used for the rest of the detector. These digital values were collected and assembled by specialized computers called MX's located away from the detector and accessible by experimenters during data taking. These MX's are interfaced with a complex FASTBUS system which assembles the data in a form suitable for processing by a general purpose computer (a VAX 8600). The FASTBUS system also contains the electronics for the trigger system described in the next section. If the trigger passes the event the data is passed to the VAX where it is then written to magnetic computer tape.

3.6 Trigger

The CDF trigger system[16] has four levels of online analysis, numbered 0 through 3 in the order which they are executed. An event is processed by a level provided the preceding level has passed the event. Level 0 is simply the minimum bias trigger described in section 3.1; it passes events at a rate of $\sim$ 50 kHz. If level 0 passes an event, level 1 then looks for an appreciable signal in one of the detector components. For instance, if the scalar sum of the transverse momenta of all particles that deposit energy in the calorimetry exceeds 18 GeV, level 1 passes the event since it may have high-energy jets. If level 1 is passed, level 2 does a more sophisticated analysis of the signal. To keep the level 2 deadtime below 5%, the event rate coming out of level 1 must be less than 5 kHz (it takes about 40 μs for level 2 to process an event). At the same time level 1 cannot eliminate events which would pass level 2. This is why the above threshold was set relatively low. Level 2 then reduces the event rate down to about 1 Hz. This is slow enough for events to be written to tape.

Level 3 can reduce the rate further, but it was not used in a way that directly affects this analysis. Any event passing the required level 2 trigger for this analysis was passed by level 3. In the future, when both the luminosity and the beam crossing frequency increase substantially, level 3 will be indispensable. It will be used to make yet tighter cuts than level 2 to keep the rate down.

To compute calorimeter parameters (*e.g.* ΣE_t) the trigger used solid angle cells which were 0.2 units wide in pseudo-rapidity and 15° wide in azimuth. Each cell's energy was the analog sum of the energies of all the calorimeter towers inside its solid angle. In the central region, the responses of four phototubes were added together for two calorimeter towers, side by side in pseudo-rapidity. For the gas calorimetry, six cathode pad responses were added together since the cell was two towers wide in pseudo-rapidity by three towers wide in azimuth (the towers in the gas calorimetry cover three times less azimuth than those in the scintillating calorimetry). The responses of the phototubes and the cathode pads are transmitted to the trigger electronics by special analog lines which bypass the digitizers. Henceforth, these cells will be referred to as *trigger towers.*

Some of the cathode pads for the plug and forward EM calorimetry were disconnected from the analog lines feeding into the trigger electronics. These corresponded to all the trigger towers with $1.0 \leq |\eta_D| \leq 1.2$ and less than half of the trigger towers with $2.6 \leq \eta_D \leq 2.8$. Due to electrical problems caused by the high voltage distribution system, these cathode pads would occasionally give a high response with no particles present. This noise can be eliminated in the data by looking at the longitudinal shower profile and comparing the pad response to the anode wire response. The trigger does not receive layer information or anode wire signals, therefore the trigger rate might be artificially increased if these pads were not removed from the trigger. The removal of these pad signals may be the cause of the depletion of jets around this region of pseudo-rapidity (see figure 6.2, page 74). However the detector simulation does not show as strong an effect as seen in the data; whereas the data show a depletion of 50 to 100 events ($\sim 0.2\%$ of the event sample), the simulation indicates only 4 events would be removed because of the missing cathode pads. Since the depletion is minor, it was not considered further. No other effect from this modification to the trigger has been observed in this analysis. Note that no other part of the calorimeter has been removed from these trigger towers.

The trigger system computes the ΣE_t of the calorimetry by summing the transverse energy, E_t, of all the trigger towers. The E_t of a trigger tower is defined to be $E \sin\theta$ where θ is the angle between the beam axis and the vector pointing from the nominal beam position to the trigger tower. The value of ΣE_t is available to both levels 1 and 2. The results of a fast hardware track processor, which finds stiff tracks in the CTC, can also be used by both level 1 and level 2. At level 1, only the existence of stiff tracks can be tested. Level 2 can use the momenta of the tracks.

The cluster finder, which is part of level 2, takes all trigger towers with $E_t > 3$ GeV as *seeds* for the clusters, and all trigger towers with $E_t > 1$ GeV as *shoulder towers.* A seed can also be a shoulder tower. The seed tower with the lowest pseudo-rapidity and the lowest azimuth (the x-axis pointing towards the center of the synchrotron) is used to construct the first cluster. All shoulder towers which are adjacent to the seed tower are included in the cluster, not including diagonally adjacent towers, *i.e.* those that only share a corner with the seed tower. Then all shoulder towers adjacent to these towers are included in the cluster. This process is iterated until there are no more trigger towers added. Then the next seed is used to create another cluster in the same fashion. No trigger tower included in a cluster can be included in a subsequent cluster. The E_t of the cluster is defined to be the sum of all the E_t's of the towers in the cluster.

Level 1 passes an event to level 2 provided one of seven conditions is met. These conditions, called triggers, are designed with specific classes of analyses in mind. One trigger, the JET_1_18, requires the analog sum of all trigger tower E_t's greater than 1 GeV to exceed 18 GeV. This was designed with the level 2 jet triggers in mind. CENTRAL_ELECTRON_6_6 requires a central EM trigger tower to exceed 6 GeV. There is also a trigger which passes minimum bias events without cuts at a rate no greater than 0.05 Hz.

Level 2 has about 18 triggers which impose tighter cuts on the detector signals. The trigger used for this analysis was called TOTAL_ET_120. It required the sum of all tower E_t's exceeding 1 GeV to be greater than 120 GeV. The TOTAL_ET_120 trigger was used for most of the data run and passed 466 285 events for a total integrated luminosity of 4.2 pb^{-1}. Due to a cut on the position of the event vertex, described on page 69 in chapter 6, the actual integrated luminosity for the three-jet data sample was 3.9 pb^{-1}. The TOTAL_ET_120 trigger was designed as a general purpose trigger, not specifically for the three-jet analysis.

In addition to this trigger there were three other triggers used in jet analyses, called JET_20, JET_40, and JET_60. The JET_60 passed every event which had a cluster with $E_t > 60$ GeV. The JET_40 passed every 30th event with a cluster of $E_t > 40$ GeV, the JET_20 passed every 300th event with a cluster of $E_t > 20$ GeV. This method of only taking one of a set number of events which passes a trigger is called *pre-scaling.* The rates of the JET_20 and the JET_40 would be too fast for the events to be written to tape. But they are needed for the inclusive jet-E_t spectra analysis, so they were pre-scaled. These two triggers could have been used with the three-jet analysis but they did not produce as much statistics as the TOTAL_ET_120. Other triggers in level 2 were used to pass events for electron analyses (*e.g.* W/Z, J/ψ), muon analyses, *etc.*

Level 3 uses the digitized event data produced by the DAQ and performs an analysis similar to the offline code. The level 3 processor consists of a software processor farm using computers which run compiled FORTRAN code.

Chapter 4

Jet Reconstruction Algorithm

The jet reconstruction algorithm uses a cone of a fixed radius to define a jet. It is similar to previous algorithms used on $Sp\bar{p}S$ experiments and corresponds closely to the parton definitions used in calculating QCD cross sections. There are three stages in the algorithm: preclustering, clustering, and merging or separating the final clusters which overlap. When the algorithm is finished each cluster is assigned a four-momentum.

4.1 Preclustering

Before a precise list of towers is determined for each of the jet clusters in an event, a rough estimate of their positions must be made. This is done by first looking for energetic towers, called *seeds*, and forming preclusters, each precluster consisting of a seed and nearby energetic towers. The first estimate for the position of the jet cluster is the centroid of the precluster.

Before a search for seeds is undertaken, the gas calorimeter towers are summed together in groups of three along ϕ to correspond to the central segmentation; each group is considered one tower during the preclustering. All towers with E_t above 1 GeV are used as seeds. The seed tower with the highest E_t is included in the first precluster. Towers are added into the precluster in an iterative procedure. If a tower with $E_t > 1$ GeV is adjacent to a tower which has higher E_t and is in a precluster, it is included in that precluster. A tower is defined to be adjacent if it shares a side or a corner. The iteration stops when no more towers can be included. Subsequent preclusters are grown from seed towers which have not yet been included in a precluster. After this stage is finished the resulting preclusters are groups of contiguous towers with the towers in each group decreasing in E_t from a central tower to the towers on the edges.

4.2 Clustering

The preclusters are grown into clusters using the true tower segmentation. The E_t-weighted centroid of the precluster is found. A cone in η-ϕ space of radius 0.7 is formed around the centroid. All candidate towers inside this cone are merged in to form a cluster, candidate towers being those with E_t greater than 0.1 GeV. A new centroid is calculated from this new set of towers, based on an E_t weighting. Again, all candidate towers inside the cone around the new centroid are merged in; any towers outside the cone are excluded. The process of recomputing a centroid and finding new towers or deleting old ones is iterated until the tower list remains unchanged.

4.3 Overlapping Clusters

There will be occasions when two or more clusters will be so close together that they have common towers in their respective tower lists after the clustering phase. In this case the clusters should be merged or the common towers should be parceled out among the clusters. Otherwise there is a chance the energy of a cluster may not correspond to the energy of the parton that originated it. The merging criteria must produce as sharp a cutoff as possible in the probability that two clusters will merge as a function of the separation between them in η-ϕ space. A sharp cutoff is consistent with the cuts used to calculate multijet cross sections and with how outgoing partons are merged in next-to-leading order calculations. In addition, the method by which towers are parceled out must change the energy as little as possible from what the energy of the cluster might be if it were isolated. The procedure to handle overlap conditions was developed with these two goals in mind.

There are four possible overlap conditions. The first two cases are trivially handled. If two clusters are distinct, they are left alone. If one cluster is completely contained in another, the smaller of the two is merged into the larger. If the towers have some finite overlap, then an *overlap fraction* is computed as the sum of the E_t of the common towers divided by the E_t of the smaller cluster. If the fraction is above a cutoff (75%) then the two clusters are combined. Otherwise the clusters are kept intact, and each common tower is uniquely assigned to the cluster with the closer centroid. After the common towers have been divided up, the centroids are recomputed. The original common towers are then re-divided according to the new centroids. As with the original cluster finding, the process of centroid computation and tower re-shuffling is iterative, and ends when the tower lists remain fixed.

4.4 Determination of Jet Four-Momenta

Every tower consists of an electromagnetic (EM) and an hadronic section. Each section is assigned a four-momentum. The energy and the magnitude of the momentum are both equal to the total energy deposited in the section; the mass is equal to zero. The direction of the momentum is collinear with the vector pointing from the measured event vertex to the electron shower maximum in the η-ϕ center of the tower if it is an EM section, or the pion shower maximum if it is an hadronic section. The electron shower maximum is defined to be the average depth where most of an electron's energy is deposited, it is located in the EM section. The pion shower maximum is analogously defined for pions; it is located in the hadronic section.

The four-momentum of a cluster is the sum of the four-momenta of the member towers' sections. Other quantities are determined from the four-momentum. For example, p_t, the jet momentum transverse to the beam axis, is given by $p_t = \sqrt{p_x^2 + p_y^2}$.

The transverse energy is defined to be $E_t = E\sqrt{p_x^2 + p_y^2} \Big/ \sqrt{p_x^2 + p_y^2 + p_z^2}$. The jet's mass is defined as it would be for an elementary particle, $m = \sqrt{E^2 - \mathbf{p}^2}$. It is related to the invariant mass of the particles in the jet.

4.5 Comparison of the Reconstruction Algorithm with Other Algorithms

The jet reconstruction algorithm just described is referred to as the *fixed cone* algorithm. It was selected from among four algorithms. The *E_t-dependent cone* algorithm is the same to the fixed cone except the cone radius is set to

$$\Delta R = \min(\frac{12.0}{E_t}, 0.6). \tag{4.1}$$

In the *contiguous tower* algorithm, all towers with E_t above 1.0 GeV are used as seeds. The highest seed tower is included in the first cluster. Candidate towers are added into the cluster in an iterative procedure. If a candidate tower with $E_{t,cand} > 0.1$ GeV is adjacent to a tower which has $E_t > E_{t,cand}/2$ and is in a cluster, it is included in that cluster. The iteration stops when no more towers can be included. Subsequent clusters are grown from seed towers which have yet to be included in a cluster. The contiguous tower algorithm is similar to the precluster stage of the fixed cone algorithm. In fact the same code was used for both algorithms, except with different parameters. The *pairwise merging* algorithm starts with the clusters found in the contiguous tower algorithm and merges pairs of clusters which are within 0.7 units of η-ϕ space.

The reconstruction algorithms were checked with a sample of *mixed* events[17]. For each event in the mixed sample, the information for an isolated cluster in an event from the data was merged into the information from another event. Two effects were studied in the mixed sample. First, the originally isolated cluster might not be detected separately in the mixed event; it might have merged with a cluster in the other event. This effect can be quantified by a merging probability. Second, even if the former cluster was separately detected in the mixed event, it might still lose or gain energy due to its proximity to the latter cluster. Both these effects come into play when two partons become collinear.

The merging probability event is plotted in figure 4.1 as a function of the η-ϕ separation. The cutoff is quite sharp for all algorithms. The change in energy of unmerged jets is shown in figure 4.2. The fixed cone algorithm has the least change in E_t.

All four algorithms have about the same resolution. The contiguous tower and pairwise merging algorithms are subject to anomalies; some clusters are found which span an azimuth greater than 180°, and some clusters are wholly embedded within other clusters. These properties are not consistent with the intuitive picture of a jet as an isolated, localized clump of energy. Mainly on the basis of figure 4.2 and the lack of anomalies, the fixed cone algorithm was selected.

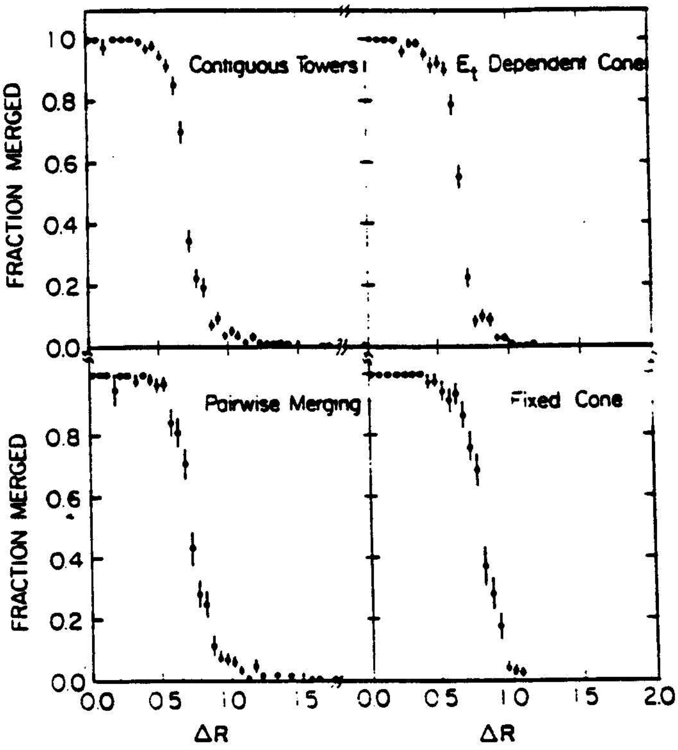

Figure 4.1: Cluster merging probability—the probability that two jets from different events in the unmixed sample might be detected as one jet in a superimposed event. This is plotted as a function of the distance between the two clusters in η-ϕ space.

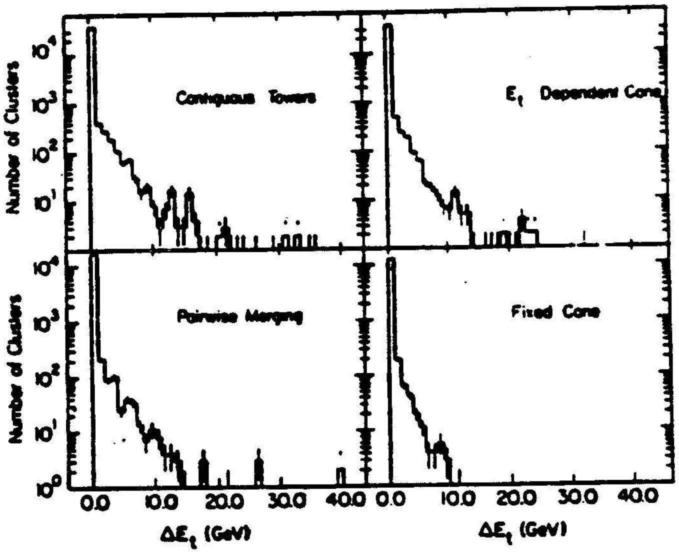

Figure 4.2: The change in a cluster's E_t after event superposition. The fixed cone algorithm shows the least change.

Chapter 5

Jet Correction

5.1 Why Jet Corrections Are Needed

The three-jet analysis uses variables which depend on the partons' four-momenta, not on the clusters' four-momenta as described in the previous chapter. These two sets of four-momenta will have different values for several reasons.

- Energy will be lost at the boundaries between the calorimetry components, particularly those at $|\eta_D| = 0.0, 1.2, 2.4$.
- Any charged particle with a transverse momentum less than ~ 0.2 GeV will not reach the calorimetry due to the solenoidal magnetic field.
- Low energy charged pions will be detected with a lower energy than they really have. This is not the case for neutral pions, whose daughter photons are detected in the EM calorimetry, or high energy particles.

- Energy from the underlying event will be collected inside the cone of the cluster by the jet reconstruction algorithm. This will add energy to the cluster.
- Some particles in the jet will fall outside of the jet clustering cone.

To account for these effects, a three step jet correction was used[18]. First, the variation in the jet response with respect to η_D is eliminated with a scale factor dependent on p_t^{cluster} and η_D. η_D is the cluster's detector pseudo-rapidity, *i.e.* the pseudo-rapidity if the event vertex were in its nominal position. Second, the vector sum of the transverse momentum of all particles inside the cluster cone, p_t^{cone}, is derived from the η_D-corrected transverse momentum. This transformation was determined using Monte Carlo simulation. Third, the average underlying event contribution is subtracted from p_t^{cone} and energy from the parton which falls outside the clustering cone is added in, giving the "true" jet energy p_t^{jet}. The cluster four-momentum (E, p_x, p_y, p_z) is corrected back to the original scattered parton with the scale factor $p_t^{\mathrm{jet}}/p_t^{\mathrm{cluster}}$.

5.2 Pseudo-rapidity Dependence

Dijet events were selected from the JET_20, JET_40 and JET_60 trigger samples by requiring at least two jets with uncorrected $p_t > 15$ GeV. To avoid bias from the online trigger, the sum of the transverse momenta of the two leading jets was required to exceed twice the value of the single jet trigger threshold by 10 GeV.

At least one of the leading two jets was required to lie within the central detector ($0.15 \leq |\eta_D| \leq 0.9$). Additionally, events with more than 1 primary vertex, and events with a primary vertex outside 60 cm were rejected. From an initial sample of about 1 pb^{-1} these cuts resulted in a final sample of approximately 15 000 events, spread almost evenly between p_t of 20 and 150 GeV.

The average missing E_T projection for these events was computed as a function of η_D^{probe}, the detector pseudo-rapidity of the jet which was not required to be in the central detector. The missing E_T projection is defined as the vector sum of all the calorimeter towers' transverse momenta projected on to the dijet axis; it measures the ratio of the effective jet energy scale of the central detector to the probed region. From the missing E_T projection, a p_t-correction factor can be calculated. This factor was fit to a linear function of jet p_t for 18 values of $|\eta_D|$ ranging from 0.05 to 3.35. For *every* jet, correction factors are calculated for the 18 values of $|\eta_D|$ given the jet p_t (those values of $|\eta_D|$ where p would be greater than 900 GeV are not considered). The correction factor for the jet η_D is derived from the set of 18 correction factors using a cubic spline interpolation, thus making the jet response across the continuous range of η_D conform to the central calorimeter response. The correction is shown in figure 5.1. The peaks near $|\eta| = 0$, 1, and 2.3 come from loss of response due to boundaries between calorimeters.

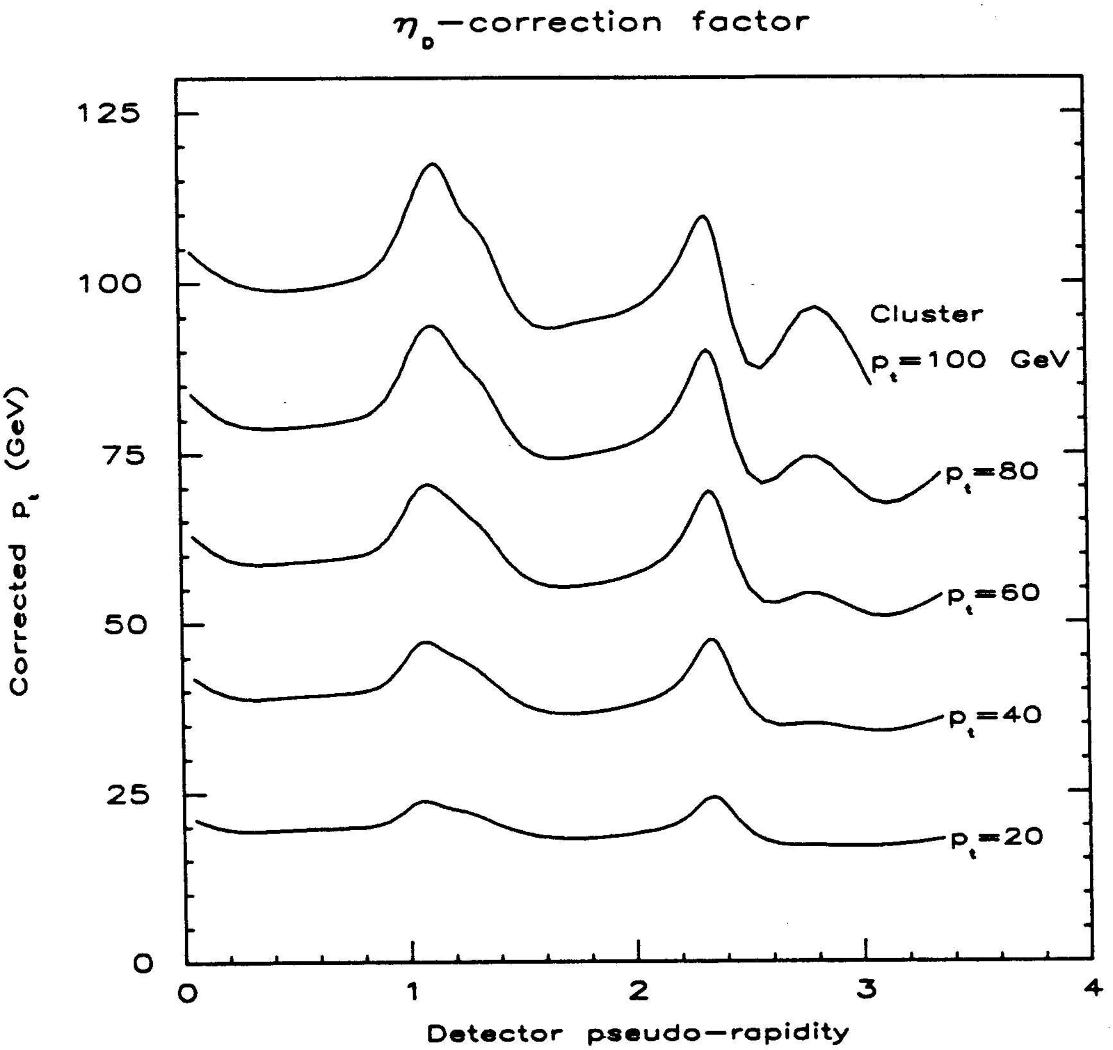

Figure 5.1: Jet correction factor vs. η_D. Note that the correction factor from the central calorimeter response study has *not* been included.

5.3 Central Calorimeter Response Correction

The central jet correction was determined from a Monte Carlo simulation. Events with two outgoing partons of equal and opposite p_t were generated with a uniform distribution in p_t and η with the ranges $10 \leq p_t \leq 700$ GeV and $0.15 \leq |\eta_D| \leq 0.9$. The partons were converted into jets of observable particles and the detector's response to the jets was simulated. The jet reconstruction program was run and the association between $p_t^{cluster}$—the transverse momentum of a given cluster—and p_t^{cone} was used to derive the central calorimeter jet response. This function accounts for the second and third effects listed in section 5.1.

5.3.1 Jet Fragmentation (SETPRT)

Before the detector simulation, the partons must be converted into observable particles. For both the dijet study described here and the three-jet QCD calculation described in chapter 7, this was done with the program SETPRT[19], which also produces particles from the underlying event. The code for both these effects was taken from the two-jet Monte Carlo program ISAJET[20]. After ISAJET generates two outgoing partons from its own 2 → 2 calculation, it converts them into jets in a two stage process. The first stage, called *evolution*, produces diquarks and gluons from the initial- and final-state partons of the 2 → 2 subprocess; this is supposed to approximate hard bremsstrahlung. The second stage, called *fragmentation*, separately converts each of the outgoing partons into observable particles using the Field-Feynman fragmentation algorithm[21]. This division into two stages is artificial; it is motivated by the impossibility of using perturbation theory at fragmentation energy scales. ISAJET then uses the energy left over in the collision ($\sqrt{s} - \sqrt{\hat{s}}$) to generate the underlying event based on an algorithm developed in reference [22].

The evolution was not included in SETPRT for two reasons. First, for this dijet study, it is easier to associate a cluster with a single parton if there are only two generated outgoing partons. Second, for the three-jet QCD calculation, the three outgoing partons were generated using the exact QCD matrix element; so the main effect of the evolution (*i.e.* the third jet) should be taken into account by the third generated parton. Since the fragmentation and underlying event parameters were tuned with evolution present in ISAJET, they were retuned in SETPRT with 2 → 2 events, so that the fragmentation distributions matched dijet data. Only those distributions which depended on the CTC were used in the tuning, so as not to use calorimeter data to correct itself. In addition, a two-jet system receives a boost from low-energy QCD radiation in ISAJET, a phenomenon which is observed in data. This phenomenon, called the k_t-kick, must be put into SETPRT explicitly since QCD radiation (*i.e.* the evolution) is absent. Even though SETPRT was tuned using dijet data, it will be shown in section 7.2 that the phenomena just described are well modeled for three-jet data too.

5.3.2 Full Simulation of the Detector (QFL)

The program QFL is a standard simulation of the CDF detector[23]. Stable particles which are generated using ISAJET or another physics generator are transported through the CTC and charged particles are bent in a magnetic field. Any particles which go through a complete turn of 2π radians in the CTC are considered lost. After leaving the CTC, particles go through the solenoid coil or the CTC endplate and then enter the calorimetry. The solenoid and the endplate are treated as dead areas; particle energy is deposited in this region but this energy is not detected.

The calorimeters are modeled with uniform density, finite sampling thickness being ignored. The calorimeters' non-linearity at low particle energies is modeled to reproduce the results of section 3.4.4. The non-linearity in the plug and forward calorimeters is assumed to be the same as the central calorimeter. This does not affect the central response analysis since jets are confined in the central region. Nor does it affect the three-jet QCD calculation, since various checks of the Monte Carlo did not reveal such effects (see section 7.2). This is confirmed by the agreement between the three-jet QCD calculation which used QFL and another calculation which used the fast simulation described in section 5.5. The latter has the response of the plug and forward calorimetry to low energy particles accounted for. (Since the fast simulation is based on the jet corrections, it cannot be used here.)

The data produced by the CTC, *i.e.* the track parameters, are simulated after the calorimeter data are simulated and the jet clustering algorithm has been run[24]. Track detection efficiency is simulated based on whether the track is located inside a jet and if so, the p_t of that jet. The jet p_t is measured by the jet reconstruction algorithm described in the last chapter, which is why the tracks are simulated *after* the clustering. The track simulation was not used in the jet correction studies, but it was used to check the fragmentation in the three-jet QCD calculation. Henceforth, QFL will be referred to as the *full simulation* in this thesis.

5.3.3 Response Function

After the events were simulated with the full simulation and clustered using the jet clustering algorithm, $p_t^{cluster}$ was related to p_t^{cone}, the vector sum of the transverse momenta of all particles inside the cluster cone. To avoid introducing systematic uncertainties associated with the fragmentation algorithm, p_t^{cluster} is compared to a p_t calculated from the observed particles instead of the p_t of the originating parton. Even if all particles coming from the parton jet were inside the cluster cone and the underlying event were not simulated, the p_t of the parton and p_t^{cone} would not be identically equal. This is because ISAJET rescales the energies of all the particles in order to conserve total momentum and $\hat{s}$. This must be done since partons are massless (or at least have a mass less than 1 GeV) and jets have an invariant mass of $\sim$ 10 GeV. The correction factor $p_t^{\text{cone}}/p_t^{\text{cluster}}$ is fitted to a quadratic equation in

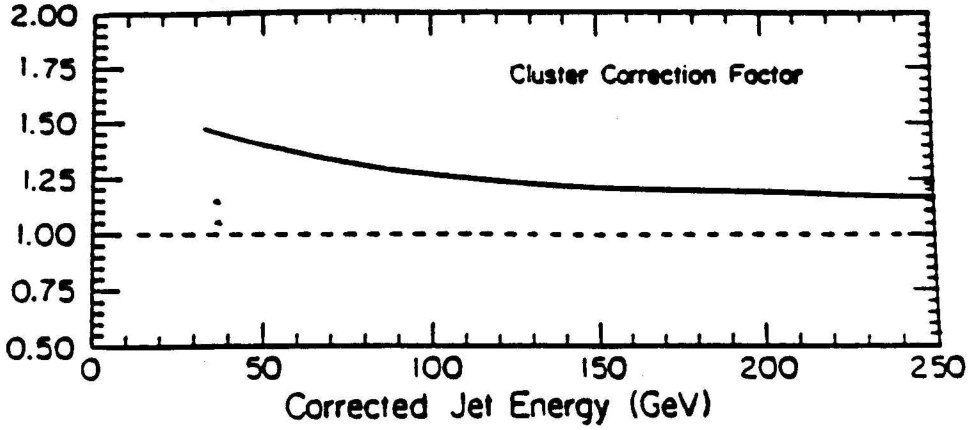

Figure 5.2: Jet correction factor vs. jet energy

$p_t^{\rm cluster}$, this is plotted in figure 5.2 as a function of $p_t^{\rm cone}$.

5.4 Underlying Event and Clustering Corrections

In a leading order calculation, there is no real definition of a jet shape; hence one should correct for energy loss outside of the jet cone assuming a standard shape for jets. For a next-to-leading order calculation, where the jet definition plays a role, it would be a more direct comparison to not include the energy lost outside of the clustering cone, since much of it would be accounted for by the calculation. The out-of-cone correction was quantified by a single parameter which is added to each jet's p_t. The out-of-cone correction was computed by the dijet simulation described in the last section and found to be independent of the p_t of the jet. It is 2.41 GeV after accounting for the underlying event subtraction effect described below. Since this analysis uses a tree level calculation, the out-of-cone correction was applied.

The average underlying event contribution to the jet p_t was determined using a dijet sample described in section 5.2. For each event, a sum was computed of the energies of all candidate towers inside a cone of 0.7 which was 90° away in ϕ from the central jet. The size and shape of the area used for the sum was varied; no statistically significant change in the E_t density was discovered. The underlying event contribution was also found to be independent of the energy of the jets in the event. The average contribution to p_t^{cone} was determined to be 2.76 GeV, corrected for the central calorimeter response as derived in the last section. This contribution was subtracted from all jet p_t's. Some of the jet's energy will appear in the area used for the sum increasing the calculated underlying event contribution. To offset this effect, this jet energy (as determined by the dijet simulation) is included again in the out-of-cone correction quoted above. The sum of the underlying event and the out-of-cone correction is 350 MeV.

5.5 Fast Simulation of the Detector (QDJTMC)

The jet corrections were used to construct the detector simulation program QDJTMC. It is used with physics generation programs which only provide final-state partons, not observable particles. This simulation converts the partons' four-momenta into jet four-momenta using the inverse of the jet correction functions, and then smears the energy of the jets using the jet resolution described below. This thesis will refer to

QDJTMC as the *fast simulation.* It does not simulate the underlying event and cannot be used to determine the properties of particles within jets, but it is much faster than the full simulation.

5.6 Jet Resolution

The energy resolution of the jets does not have much effect on the variables x_3 and x_4 since, to first order, they do not depend on the energy scale, and the three-jet cuts described in chapter 6 keep all events well above the energy thresholds. However it will be briefly described here. The dijet data described in section 5.2 was used to determine the energy resolution. The average difference in the transverse momenta for two jets should depend on how much the responses of the calorimeters differ. The deviation from this average comes from two sources. First, low-energy QCD radiation will provide a transverse boost to the dijet system; this boost is the k_t-kick described in section 5.3.1. Second, the two individual jet energies can fluctuate from their true value independently due to their non-zero energy resolution. The k_t-kick can boost the dijet system in any direction, whereas the resolution effect will only be felt along the dijet axis. Therefore, the resolution can be measured by subtracting the spread in the p_t of the dijet system *transverse* to the dijet axis from the spread in the p_t of the dijet system *along* the dijet axis[25]. Using this method, the resolution for various values of jet p_t and η_D were derived (see figure 5.3).

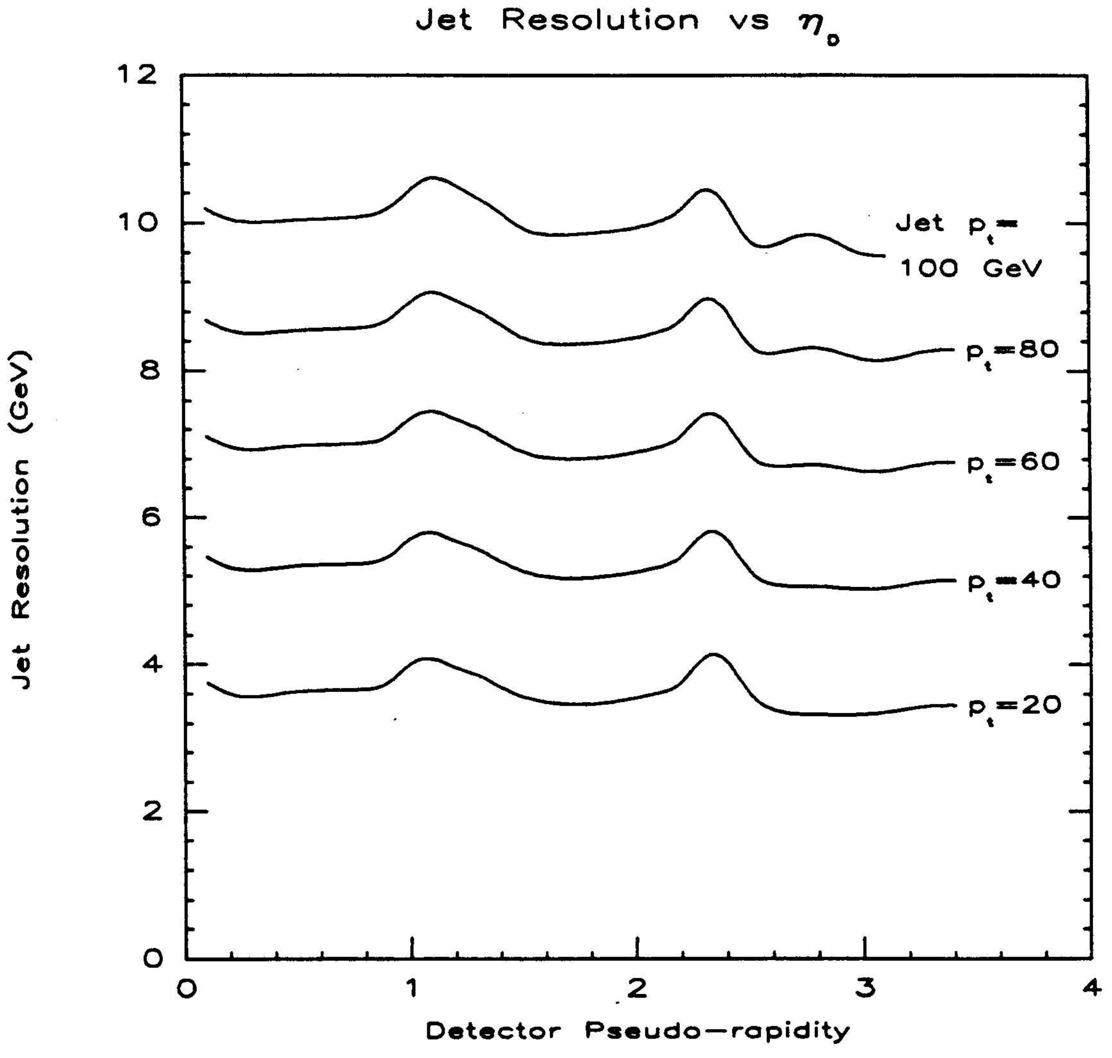

Figure 5.3: Jet resolution for various values of p_t and η_D.

Chapter 6

Event Selection

All events which passed the TOTAL_ET_120 trigger and contained three or more reconstructed jets were analyzed. If there are more than three jets in the event, only the three highest E_t jets are considered. The rest are ignored. There is no cut on the fourth jet or any remaining unclustered energy.

In theory, measuring the distribution of one of the interesting phase space variables should be very easy. One simply chooses a volume of the three-jet phase space, records the events collected within that volume, and integrates over all but that variable to get its distribution. However, because the CDF detector is not ideal, and the jets themselves are not infinitely narrow, one is limited in the choice of phase space volume. Moreover, because the natural variables of the experiment do not coincide with the phase space variables, one must understand how cuts in the former will effect distributions of the latter. Each of the former cuts will be considered below.

- trigger ΣE_t

  The CDF detector triggers on ΣE_t rather than M_{3j}. Given M_{3j}, the ΣE_t of a three-jet event depends strongly on $\cos\theta^*$. As $\cos\theta^*$ grows, the E_t of the leading jet falls, as does the total E_t of the other two jets. One should therefore choose a range of $\cos\theta^*$ and then set the invariant mass cut high enough that any three jet event will have sufficient ΣE_t to satisfy the trigger at $\cos\theta^*_{max}$.

- z_{vtx}

  All events are required to have a primary vertex within 60 cm of the nominal position along the beam axis (*i.e.* $z_{vtx} < 60$ cm). This is to minimize the number of events which have jets leaking through the opening between the plug and forward calorimetry.

- ΔR

  Final-state partons which are closer to each other than ~ 0.8 units of η-ϕ space will be detected as one jet. In other words, jets which are closer than ~ 0.8 will not be separated by the clustering algorithm. This is referred to as *finite angular resolution.* It is equivalent to requiring any pair of the three jets to have a value of ΔR exceeding some minimum, where

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{6.1}$$

The minimum value of ΔR is not well defined, and depending on the fragmentation of the jets, can vary between 0.5 and 1.0. To make the comparison between the data and the QCD calculation more consistent, ΔR is required to be greater than 0.85. This cut can bias the x_3-x_4 Dalitz plot. For events with high x_3 and low x_4, partons 4 and 5 become close together and the event fails the ΔR-cut. This bias is eliminated with a suitable x_3-cut.

- Jet E_t

  Jets must have at least 10 GeV (uncorrected) E_t. This cut tends to eliminate parts of phase space where x_5 is small, that is, where x_3 and x_4 are both large. However, because of the x_3-cut and the fact that $x_4 < x_3$, there is already a lower bound on x_5. More importantly, one of the secondary jet E_t's will fall as ψ^* goes to 0 or 180°. This fall grows steeper as $\cos\theta^*$ gets larger. Once the ranges of M_{3j}, x_3 and $\cos\theta^*$ are selected, one must choose the range of ψ^* so that the two secondary jets always have E_t's in excess of 10 GeV.

- η_D

  To stay well away from the forward boundary of the calorimetry, all three jets must have the detector pseudo-rapidity $|\eta_D| < 3.5$. The combined effect of all the three-jet cuts keeps the three jets well away from the $|\eta_D|$-cut. The most important are the cuts on $\cos\theta^*$ and ψ^*. The $\cos\theta^*$ cut is set to 0.6. Disregarding the boost of the three-jet system and the primary vertex position, the equivalent pseudo-rapidity for the leading COM jet would be $\eta = 0.7$.

As for the other two jets, the ψ^* cut restricts their pseudo-rapidity to be less than ~ 1.5. One needs a boost of $y^{boost} > 2$ to have a non-leading jet with $\eta > 3.5$. However, the M_{3j} cut places a lower limit on the product $x_1 x_2$. This, coupled with equation 2.7 (page 13) and the requirement that $x_{1,2} < 1$, make it impossible to generate such a high boost. The *detector* pseudo-rapidity of a jet, η_D, may increase because the primary vertex is far way from the nominal position. However at $\eta_D \approx 3.5$ the maximum shift in η is 0.1, which is negligible.

The final cuts are:

$$E_{t\ \mathrm{jets}\ 1,2,3} > 10\ \mathrm{GeV}, \tag{6.2}$$

$$z_{vtx} < 60\ \mathrm{cm}, \tag{6.3}$$

$$|\eta_D| < 3.5, \tag{6.4}$$

$$\Delta R > 0.85, \tag{6.5}$$

$$M_{3j} > 250\ \mathrm{GeV}, \tag{6.6}$$

$$x_3 < 0.9, \tag{6.7}$$

$$|\cos\theta^*| < 0.6, \tag{6.8}$$

$$30^\circ < |\psi^*| < 150^\circ. \tag{6.9}$$

Of the 4 866 events remaining after these cuts were made, there were 40 which had a total event energy exceeding 1800 GeV. These 40 were eliminated from the final sample.

In summary, the three-jet cuts, listed in the inequalities 6.6 through 6.9, must be tight enough to so that an event sample passing them would naturally pass the cuts required to keep the detector efficient. These *detector cuts* are the first four inequalities listed above (6.2 through 6.5) as well as the trigger requirement that $\Sigma E_t > 120$ GeV. If there is any bias from the detector cuts, histograms of the cut parameters, E_t, η_D, ΔR, and trigger ΣE_t, should show sharp cutoffs near these cuts.

The histograms of E_t for the three jets all show a gradual turn on resulting from the three jet cuts (figure 6.1). The η_D plots show similar behavior (figure 6.2). There is slight depletion at $|\eta_D| \approx 1.1$. This may be due to the uninstrumented part of the ΣE_t trigger (see page 44). There are no depletions or enhancements in any other region of η_D.

The effect of the finite angular resolution of the jet clustering algorithm may be studied by plotting a histogram of ΔR for the two jets which are closest to each other in η-ϕ space (*i.e.* have the lowest ΔR). This plot is shown in figure 6.3. As with E_t and η_D, there is no sharp cutoff at the ΔR cut. In fact most events have a minimum separation well above the effects of finite angular resolution (see figure 4.1d).

The trigger ΣE_t histogram is shown in figure 6.4. There is no sharp cutoff at the trigger threshold, indicating that there is negligible trigger bias. Figure 6.5 shows the E_t of the leading trigger cluster for the TOTAL_ET_120 sample (*i.e.* the final data sample). Practically all events are above 40 GeV. Therefore, if one uses the

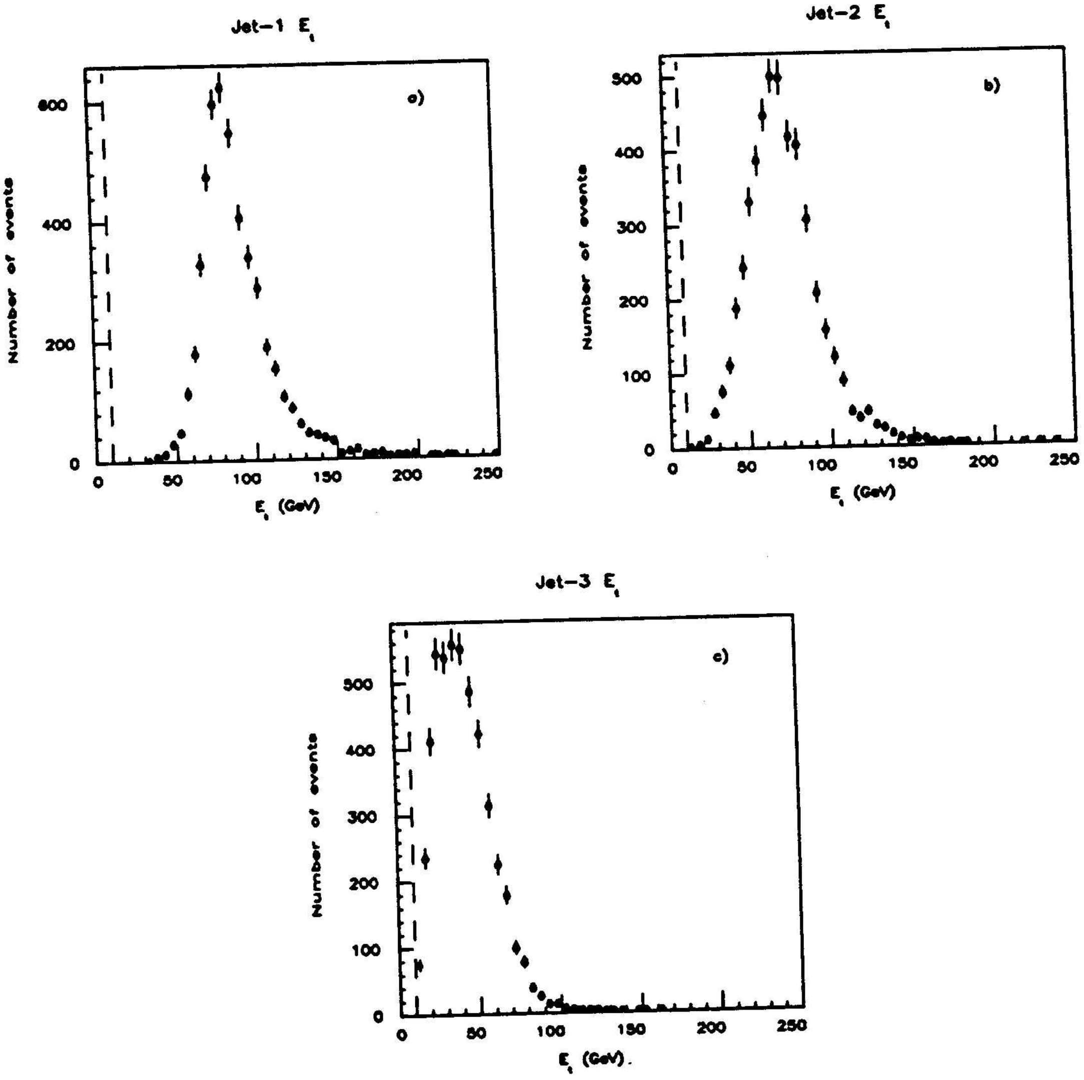

Figure 6.1: Jet transverse energy distributions for the final three-jet sample. a) highest E_t jet in lab frame, b,c) second and third highest E_t jet in lab frame. There is no sharp cutoff near the data cut—shown by the dashed line in the figure—given in equation 6.2.

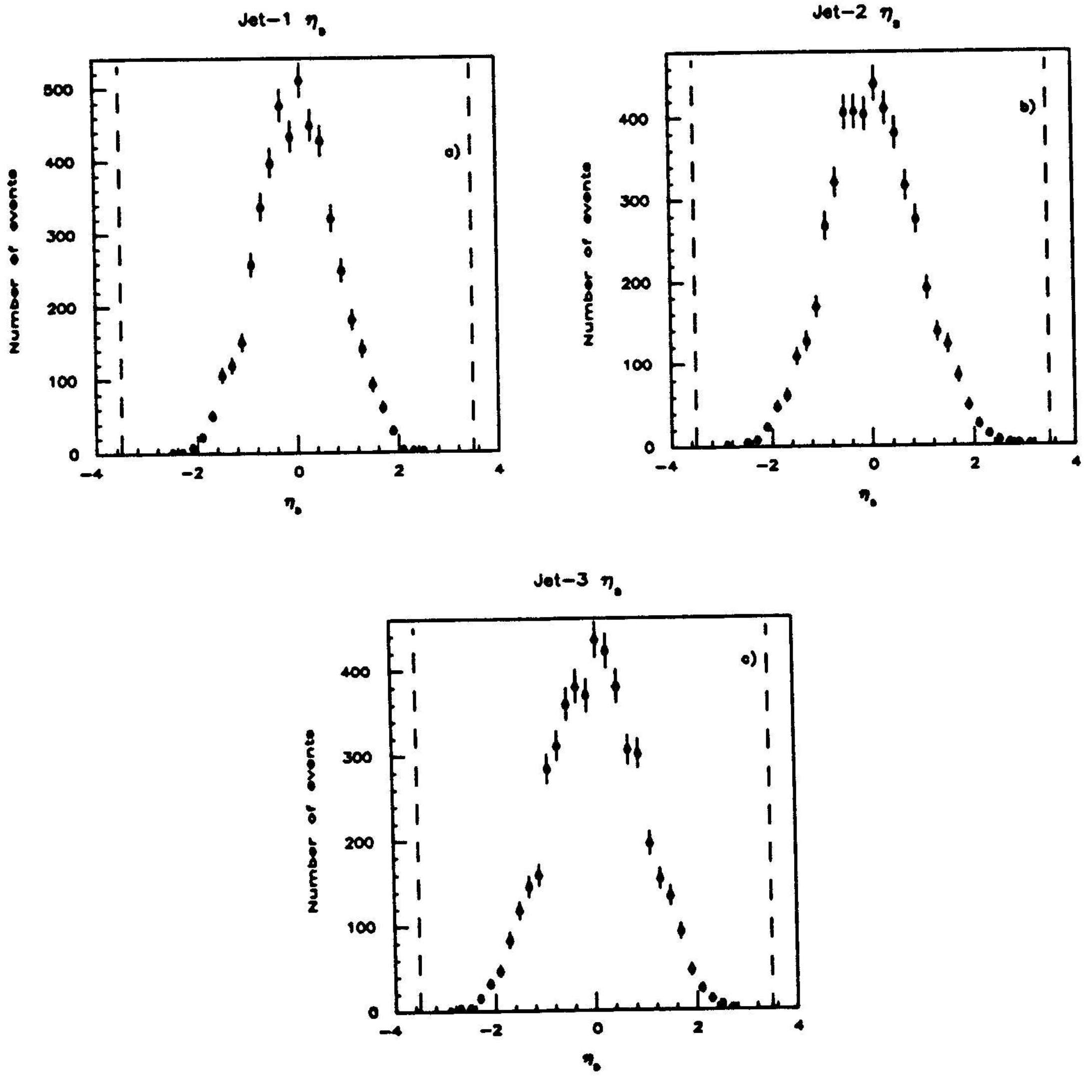

Figure 6.2: Jet pseudo-rapidity distributions for final three-jet sample. a) highest E_t jet in lab frame, b,c) second and third highest E_t jet in lab frame. There is no sharp cutoff near the data cut—shown by the dashed line in the figure—given in equation 6.4. The depletion around $|\eta_D| \approx 1.1$ may be due to an uninstrumented part of the ΣE_t trigger (see page 44).

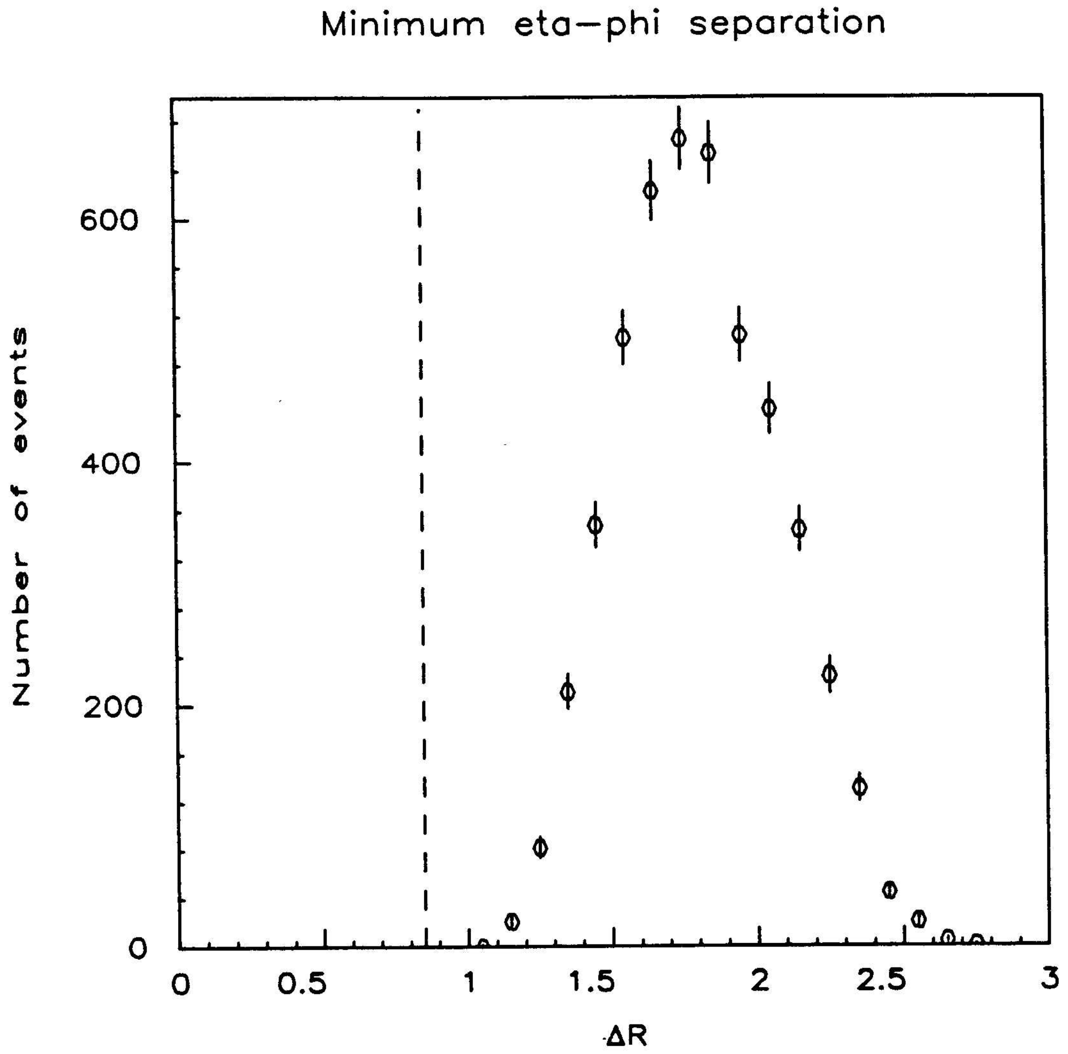

Figure 6.3: The η-ϕ separation of the closest two jets for the final three-jet sample. There is no sharp cutoff near the data cut—shown by the dashed line in the figure—given in equation 6.5.

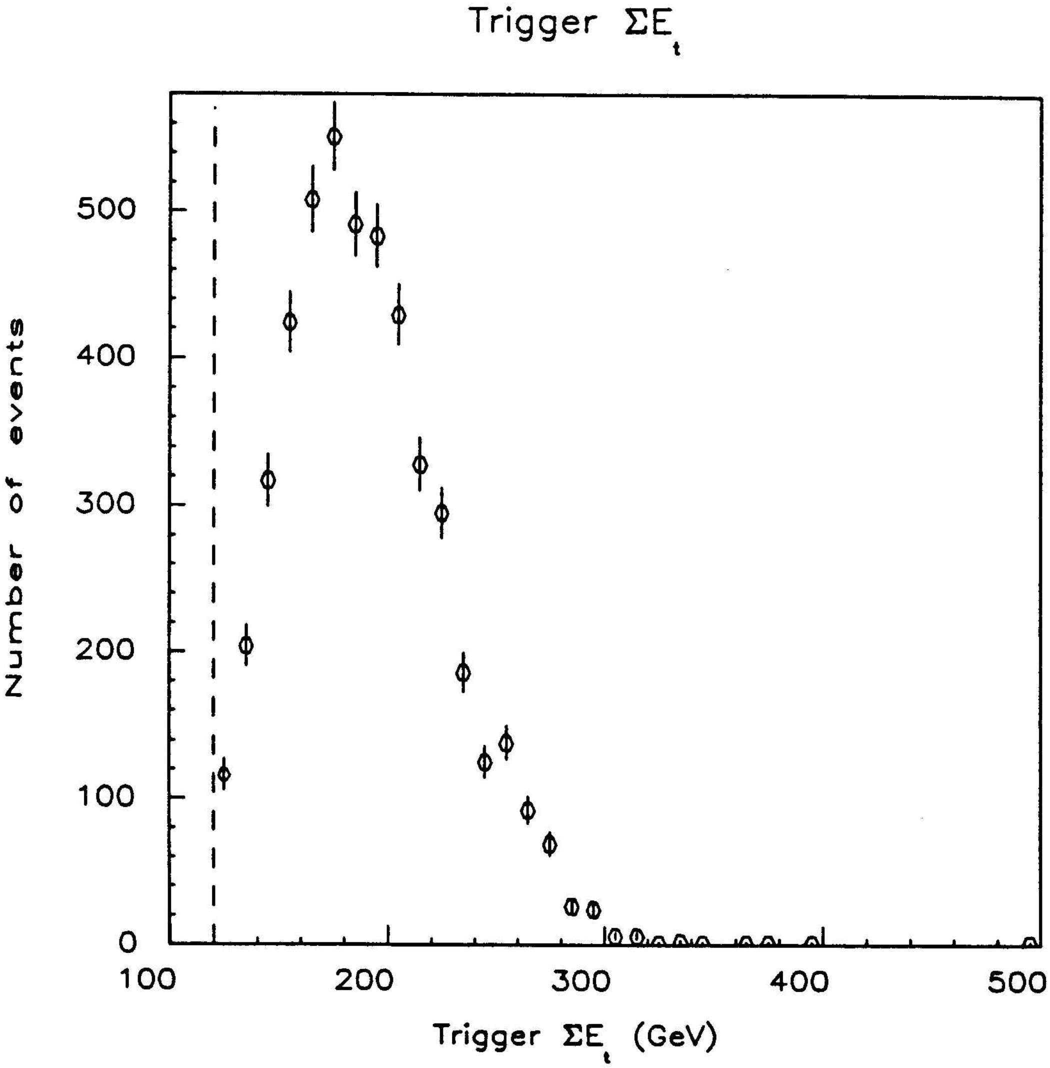

Figure 6.4: The trigger ΣE_t for the final three-jet sample. There is no sharp cutoff near the trigger threshold shown by the dashed line in the figure.

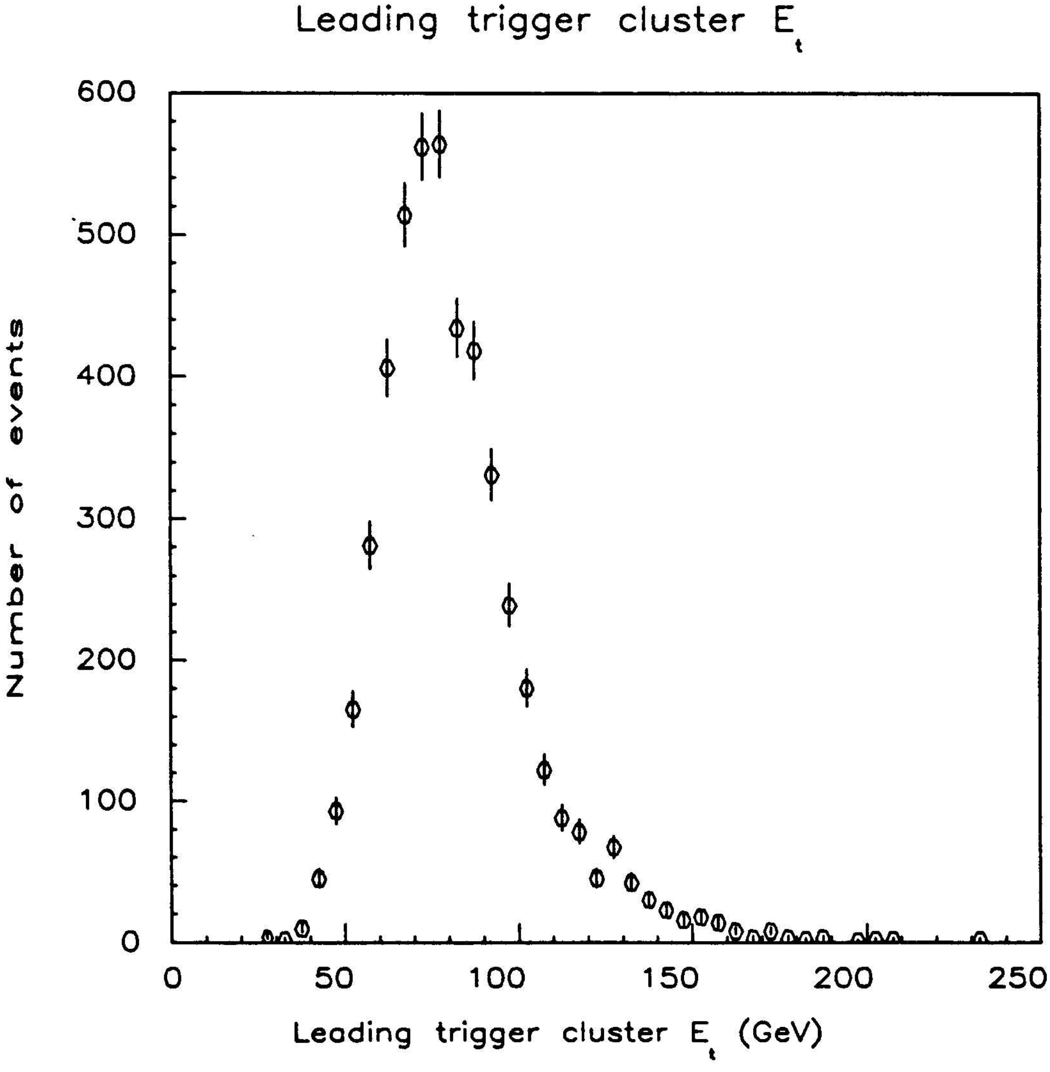

Figure 6.5: Transverse energy of the leading trigger cluster for the final three-jet sample. Most events have an E_t above 40 GeV, indicating there would not be much trigger bias in the JET_40 sample as well as the TOTAL_ET_120 sample.

same cuts for the JET_40 sample as for the TOTAL_ET_120 sample, one should get a sample of events unbiased by the 40 GeV threshold. The trigger ΣE_t can be plotted for the JET_40 sample to see how many events fail the $\Sigma E_t < 120$ GeV cut. This plot is shown in figure 6.6. About 3% fall below 120 GeV. In addition the shape of the distribution is similar to that of the TOTAL_ET_120 sample, indicating that the acceptance between the two samples is similar. As a check on the efficiency of the JET_40 sample, the E_t of the leading trigger cluster is plotted in figure 6.7. As expected, there is a gradual turn on above 40 GeV, indicating negligible bias from the trigger threshold.

As there are only 209 events in the JET_40 sample, it was not included in the final three-jet sample, which already has 4 826 events. The other triggers used in experiment which could have been analyzed with this method were the JET_20 and TOTAL_ET_150. (The latter was used in the beginning of the data run, and required $\Sigma E_t > 150$ GeV.) But these also would produce low statistics compared to the final three-jet sample.

One might argue for eliminating the 692 events which have two or more primary vertices. Two interactions occurring in the same crossing can produce enough ΣE_t to pass the trigger when neither of them has the ΣE_t by itself. The rate is proportional to the product of the cross sections of two ΣE_t triggers with thresholds whose sum equals 120 GeV. This can be compared with the TOTAL_ET_120 cross section multiplied

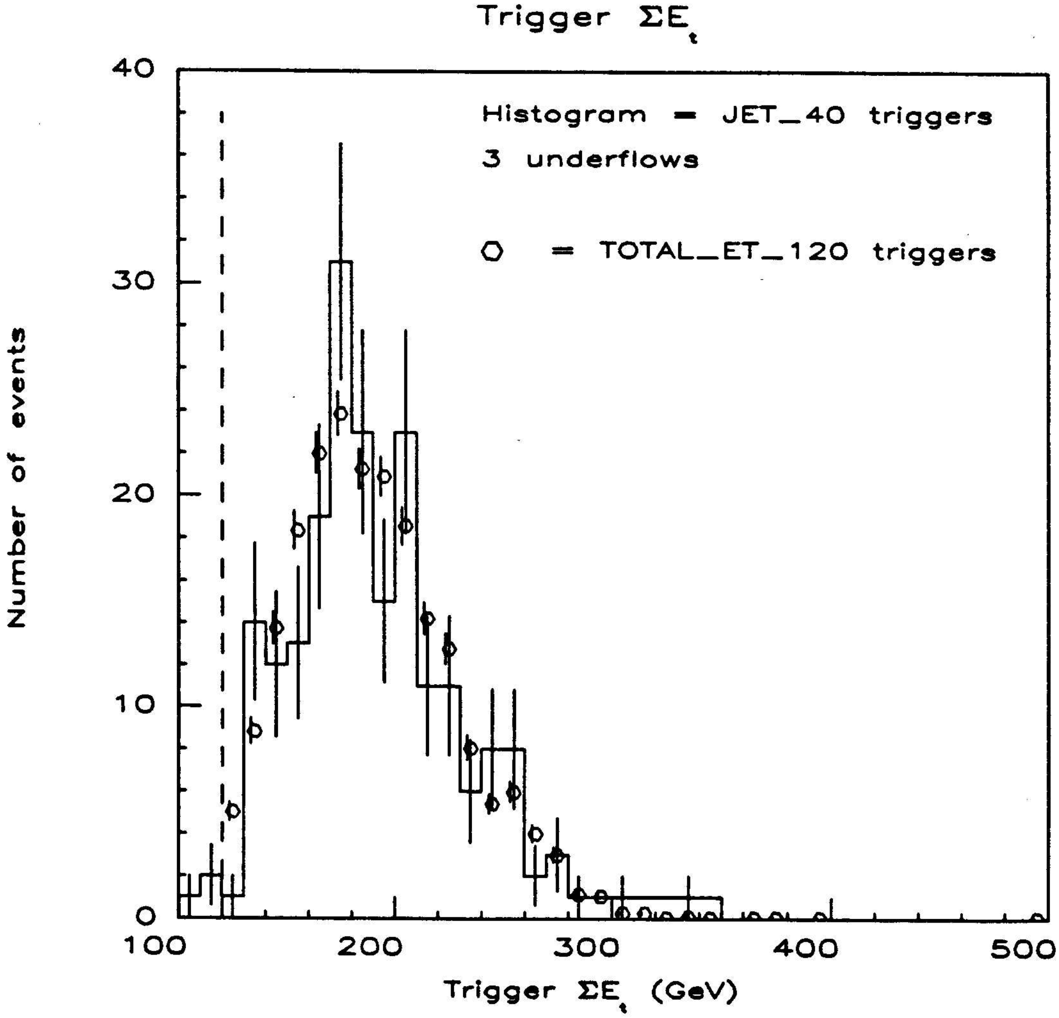

Figure 6.6: The trigger ΣE_t for the JET_40 three-jet sample (histogram), as well as the TOTAL_ET_120 sample (hexagons). Most events in the JET_40 sample have a trigger ΣE_t above the 120 GeV trigger threshold (shown by the dashed line). Only 3% fall below the threshold.

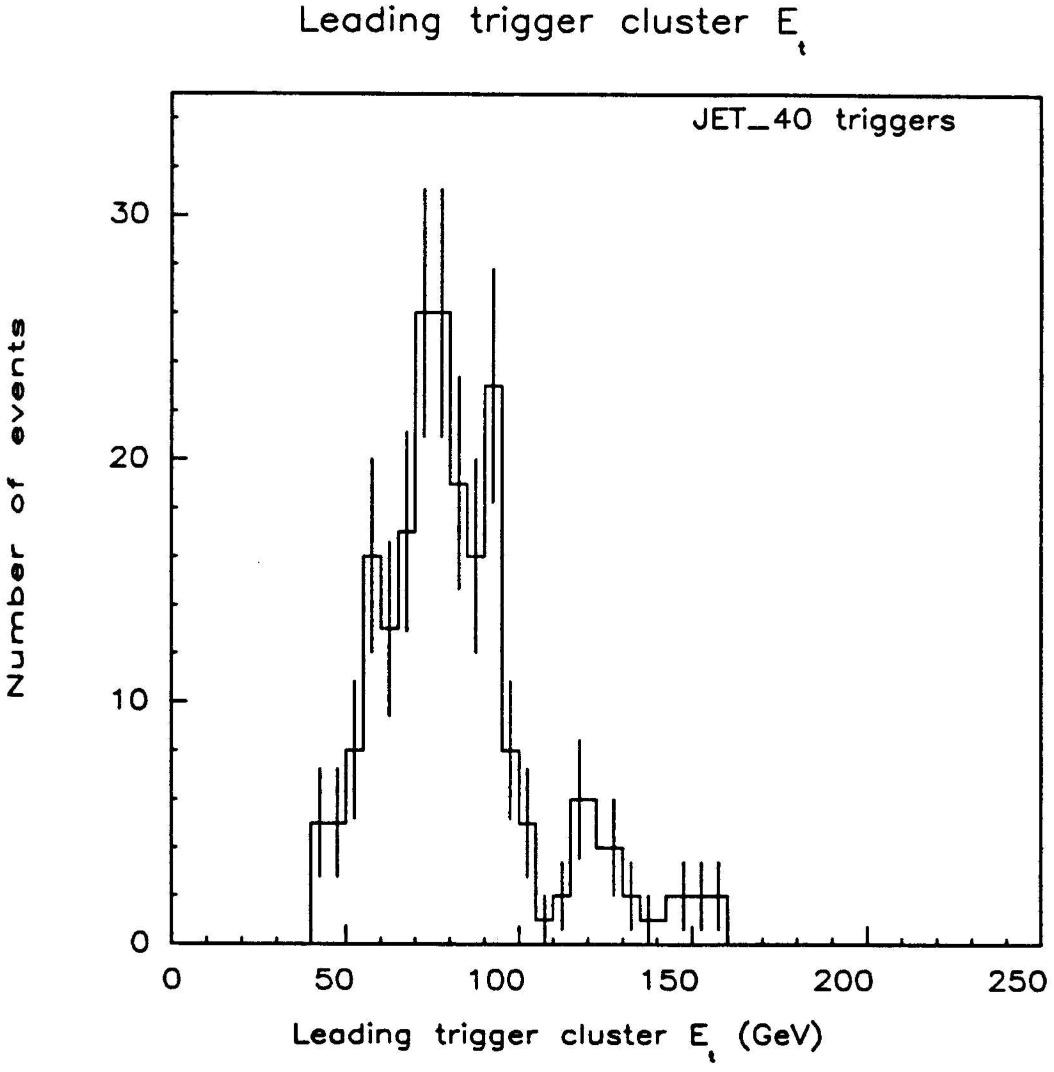

Figure 6.7: Transverse energy of leading trigger cluster for the JET_40 three-jet sample. There is little trigger bias in the sample since there is no sharp cutoff at the 40 GeV trigger threshold.

by the minimum bias cross section. Based on this calculation, about 2% of all events with two primary vertices are expected to come from two hard interactions, the rest should only have one hard interaction. This translates to a background of $\sim 0.3\%$ in the final three-jet sample. Therefore no cuts were made on the number of vertices.

Chapter 7

Monte Carlo Calculations

The energy fraction distributions of the final three-jet sample were compared to a QCD calculation performed using the Monte Carlo technique in several steps. The events were generated at the parton level, the partons were converted into observable particles and then the detector's response to these particles was simulated. Finally the three-jet analysis was applied on the fully simulated events. This chapter will describe the event generation as well as various checks on the simulation. The simulations of the parton fragmentation and the detector were described in sections 5.3.1 and 5.3.2.

7.1 Event Generation

The Monte Carlo program used to generate the three-parton events was PAPAGENO[26]. The momentum distributions of the initial-state partons corresponded to the EHLQ1 structure functions[27]. The momentum scale of the interaction, Q, was taken to be the average parton p_t divided by two. The tree level $2 \rightarrow 3$ matrix elements were calculated by the compact expressions derived by Berends *et al.*[3]. These matrix elements were used to calculate the cross sections of the three-parton configurations which were generated. PAPAGENO is a general purpose Monte Carlo generator, so the three parts of the event generator which were discussed above—the structure functions, the Q^2-value and the process generated—can all be changed. This feature will be used to check the simulation.

PAPAGENO generates the four-momentum of the five partons in the 2 → 3 process by a generalization of the method it uses to produce 2 → 2 events. The variables x_1 and x_2 are generated first with a distribution approximating the structure functions. Then $\cos\theta$ and ϕ for the first two final state partons are generated with distributions tuned to conform to the total matrix element. The tuning is done in a special run before the main calculation. The four-momentum of the third final-state parton is calculated from the first two by using momentum-energy conservation. The order of the partons at this stage is arbitrary. Note that this is *not* a phase space generator. After all the four-momenta of the partons are specified, the generation cuts described below are applied. If the event fails the generation cuts, it is regenerated.

The cross section for this configuration of momenta is calculated using an approximation to equation 2.6 (page 11),

$$\sigma = \sum_{abcde} f_{a/p}(x_1) f_{b/\bar{p}}(x_2) \hat{\sigma}(ab \to cde). \tag{7.1}$$

The sum is over all possible combinations of flavors a, b, c, d and e which do not violate flavor conservation. The subprocess cross section, $\hat{\sigma}(ab \to cde)$, is given by the appropriate 2 → 3 matrix element. For example, for a subprocess with five gluons, the total matrix element for all diagrams is

$$\hat{\sigma}(gg \to ggg) = \frac{27g^6}{8} \frac{\sum_{j=2}^{5}\sum_{i=1}^{j}(ij)^4}{\prod_{j=2}^{5}\prod_{i=1}^{j}(ij)} \sum_{ijklm}{}'(ijklm), \tag{7.2}$$

$$(ij) \equiv p_i \cdot p_j,$$

$$(ijklm) \equiv (ij)(jk)(kl)(lm)(mi).$$

The gluons are numbered 1 through 5: 1 and 2 being the initial-state gluons from the proton and antiproton respectively, and 3 through 5 being the final-state gluons in arbitrary order. This does not necessarily correspond to the ordering scheme introduced in chapter 2, where the partons are numbered in descending energy. The sum $\sum'_{ijklm}$ is over all non-cyclic permutations of 12345.

PAPAGENO is not fully efficient, *i.e.* events generated by PAPAGENO do not populate phase space as QCD would. Therefore PAPAGENO assigns a *weight* to each event. When filling histograms or calculating total cross sections, each event's entry in the histogram or sum is not unity but its weight. The weight of the event is calculated by dividing the cross section of the event's configuration by the probability that this configuration would be generated (for a generator which populates phase space uniformly, this probability would be unity). Because the weights have a large variation event to event, a larger Monte Carlo sample is needed for a statistical accuracy comparable with a similar unweighted calculation. So that the full simulation, which takes a lot of CPU time, can be run on a smaller sample of events, each event was rejected if its weight did not exceed a random number between zero and the maximum weight of all events. All events in the resulting sample have equal weight, *i.e.* these events are unweighted. For other calculations, which were used to check the full simulation, the fast simulation was used, so the events were not unweighted.

Because of the denominator in equation 7.2, the five-gluon cross section has singularities. If any gluon becomes collinear with another parton, the inner product of their four-momenta goes to zero, thereby forcing the denominator to zero. This is called a *collinear divergence.* If any gluon has a zero momentum, the denominator will also be zero, thus causing an *infrared divergence.* Matrix elements for subprocesses which contain two or four quarks have the same gluon divergences, although the gluon divergences for these matrix elements may be unphysical since a zero quark mass approximation was used in the calculation.

These divergences have repercussions for equation 7.1. The cross section could become very large or even infinite for certain parton configurations. Weights could fluctuate wildly from event to event. Therefore the weights should not be calculated for events generated near regions where the QCD matrix element is singular. At the same time, events generated outside the region of phase space enclosed by the data cuts (inequalities 6.2 through 6.9, page 71) must be passed on to the detector simulation. Because of detector and fragmentation effects these events might drift into the region of acceptance.

The infrared divergence is avoided by making a cut on the p_t of a final-state parton. To avoid the collinear divergence two cuts must be used: a separation cut on ΔR to keep two final state-partons from becoming collinear, and a jet-η cut to keep a final-state parton from becoming collinear with an initial-state parton. Finally, in order to avoid simulating a lot of events that would not pass the final cuts, a preliminary cut on the subprocess energy $\sqrt{\hat{s}}$ was added.

The final set of generation cuts for the partons was

$$p_t > 15.5 \text{ GeV}, \tag{7.3}$$

$$\Delta R > 0.7, \tag{7.4}$$

$$|\eta| < 4.0, \tag{7.5}$$

$$\sqrt{\hat{s}} > 200 \text{ GeV}. \tag{7.6}$$

On the average, partons with $p_t = 15.5$ GeV will be detected with a jet-E_t of 9.4 GeV, while all other variables which make up the primary cuts do not change regardless of the partons' energy. Although the cuts listed above are looser than those on the final data sample, they exclude the singular regions.

7.2 Checking the Simulation

Before the x_3-x_4 distributions of the QCD calculation can be compared with the data, other event parameters must be checked so one can have confidence in the simulation.

Not all properties of three-jet events will be matched precisely by the full simulation. But variables which directly measure the four-momenta of the three leading jets must be well reproduced since the x_3-x_4 distributions depend on them.

7.2.1 Fragmentation and Jet Parameters

The characteristics of the individual jets and some of the three-jet parameters will be presented here. Discrepancies in the plots will be discussed at the end of the subsection.

The fragmentation model can be verified by looking at the properties of the individual particles inside the jets. The only way to look at these particles is by using the central tracking chamber (CTC) which detects charged particles. Five parameters of the fragmentation are examined here:

- p_t flow – the scalar sum of the transverse momenta of all particles in a slice of azimuth ϕ as a function of the distance in ϕ from the leading jet,
- z – the fraction of the jet momentum carried by a particle inside the jet,
- p_{perp} – the particle's momentum transverse to the jet axis,
- p_t – the particle's momentum transverse to the beam axis,
- the number of particles inside the jet.

The CTC is most efficient ($\gtrsim$ 80%[24]) for particles in the central region ($|\eta_D| \lesssim 0.8$). Therefore only jets in this region were used. So as not to affect the distributions of the fragmentation parameters, the jets were confined to the regions of the central detector where the response was uniform ($0.1 < |\eta_D| < 0.7$). Only the jets used for the x_3-x_4 calculation (*i.e.* the three highest-E_t jets) were considered. If none of the three leading jets passed the η_D-cut, the event was excluded from the fragmentation histograms.

All events which had a second vertex were eliminated from this sample to avoid any effects from a superimposed minimum bias event. Cuts were made on the track parameters of the particles to insure the CTC had good resolution and efficiency for the entire particle sample. All tracks were required to pass within 5 mm of the primary vertex, go through the sixth superlayer of the CTC, be fully reconstructed by the CTC, and have a p_t of at least 0.4 GeV. A particle is defined to be inside a jet if the particle's momentum at the primary vertex is within 0.7 of the jet axis in η-ϕ space.

The p_t flow plot is shown in figure 7.1a. Every particle's entry in this histogram is weighted by its corresponding p_t. The abscissa corresponds to the difference in azimuth between the particle and the leading jet. The large peak at 0 corresponds to this jet. The peak at 2.7 rad comes from the other two jets. The simulation tends to be more peaked at 0, which can be caused by the simulation jets being narrower

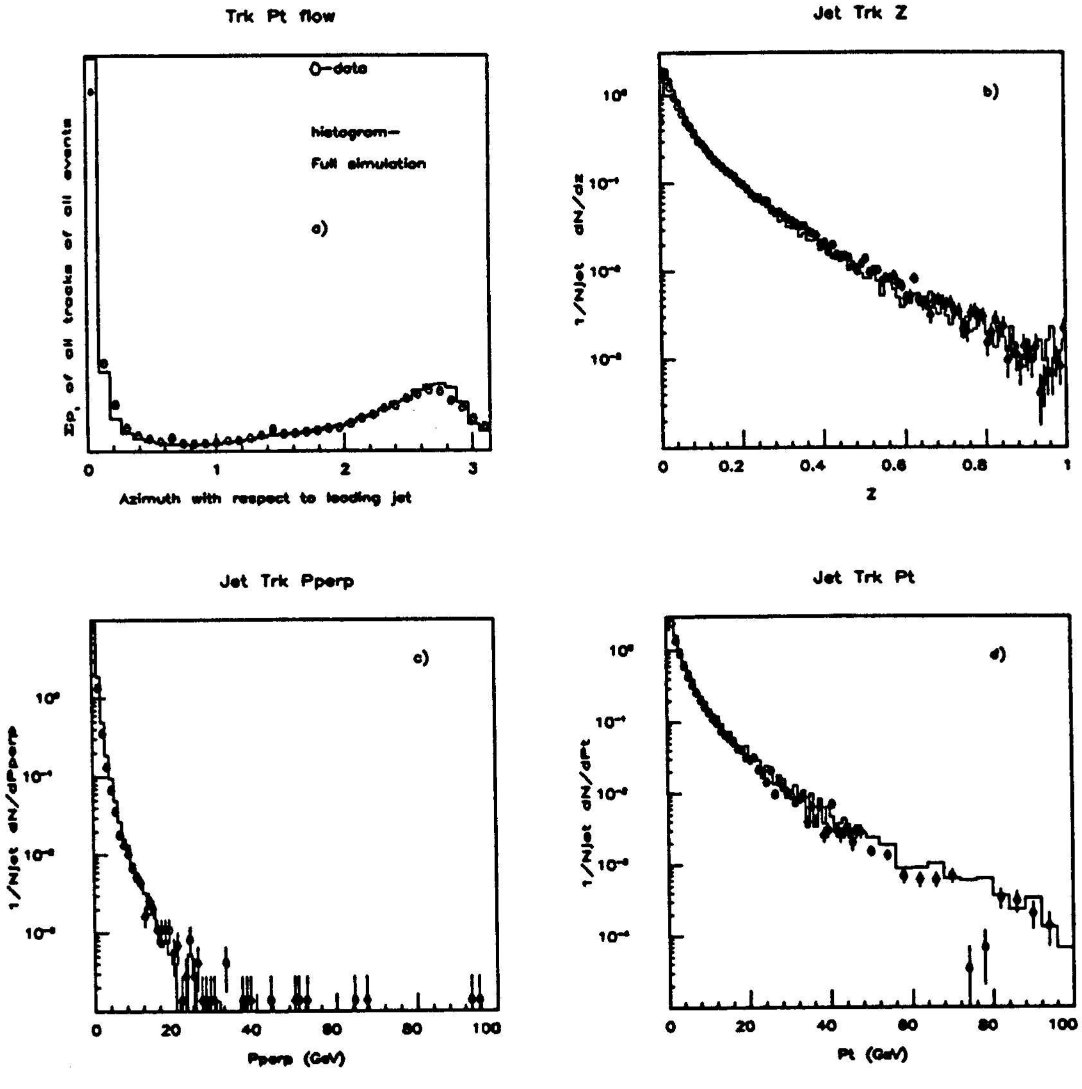

Figure 7.1: Jet fragmentation distributions for data and simulation. a) The p_t-flow plot—the abscissa is the azimuth of a particle with respect to the leading central jet, the value of the bin contents is equal to the sum of the p_t's of the particles that fall within that bin. The simulation is area normalized to the data for this plot (but not for b), c) or d)). b) z – The fraction of the jet momentum carried by a particle inside the jet. c) p_{perp} – Particle momentum transverse to the jet axis for particles inside jets. d) p_t – Particle momentum transverse to the beam axis for particles inside jets.

than they should be.

For each particle z is defined by

$$z = \frac{\mathbf{p}_{trk} \cdot \mathbf{p}_{jet}}{|\mathbf{p}_{jet}|} \Bigg/ \left|\sum \mathbf{p}_{trk}\right| \tag{7.7}$$

where the summation is over all tracks inside the jet. The histogram of z is shown in figure 7.1b. Histograms of p_{perp} and p_t are shown in figures 7.1c and d. The simulation and the data agree quite well. The distribution of number of charged tracks inside a jet is shown in figure 7.2. The data distribution is peaked at around 9 particles, whereas the simulation is peaked around 12. Otherwise the two distributions have the same shape.

The definition of a jet's mass was provided in section 4.4; a brief review will be given here. Energy E_{jet} of a jet cluster is defined as the sum of the energies of the calorimeter towers inside the jet. A jet cluster's momentum $\mathbf{p}_{jet}$ is defined as the vector sum of the calorimeter tower momenta. A tower momentum has a modulus pointing from the primary vertex to the tower and a magnitude equal to the energy deposited in the tower. Thus the jet's mass is defined to be

$$m_{jet} = \sqrt{E_{jet}^2 - |\mathbf{p}_{jet}|^2}. \tag{7.8}$$

The mass is a measure of the width of the jet in η-ϕ space. Jets with a low mass compared to their energy tend to be narrow, jets with a high mass tend to be wide. A plot of the masses of the three jets is shown in figure 7.3. There is agreement in

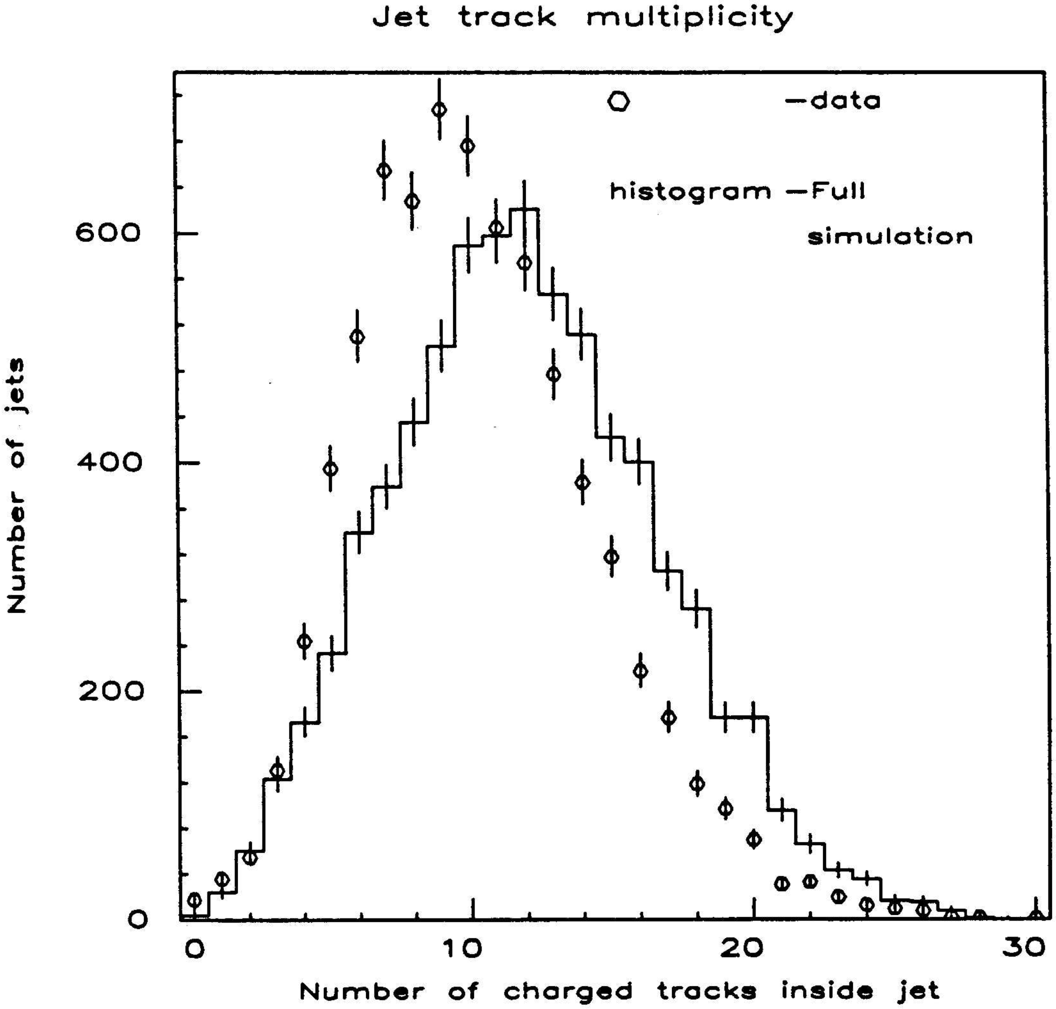

Figure 7.2: Jet multiplicity—number of charged tracks inside a jet—for data and simulation. The simulation is area normalized to the data.

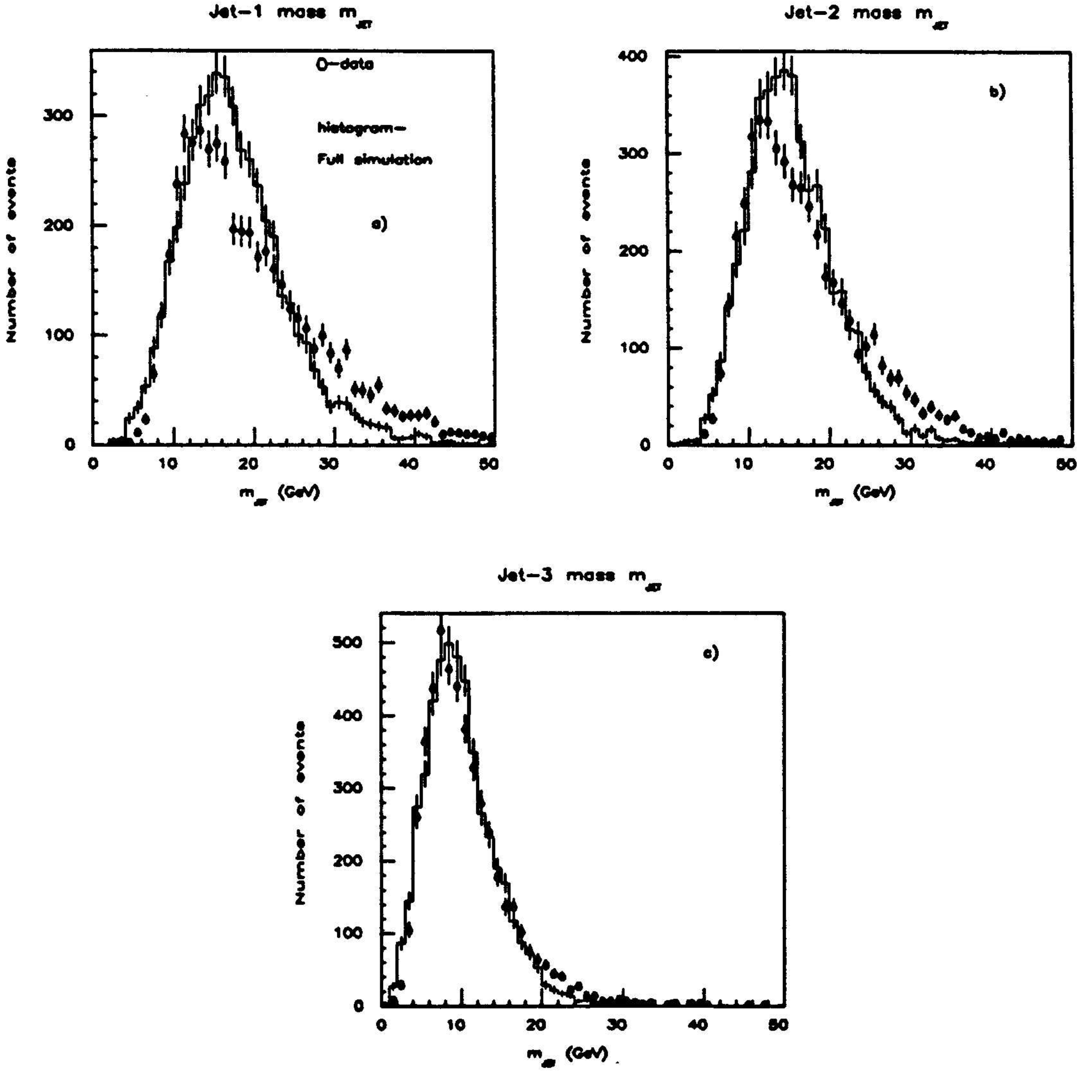

Figure 7.3: Jet masses for data and simulation. The simulation is area normalized to the data.

the mass of the third jet, but for the leading and second-leading jet, the data have higher masses than the simulation. This means that jets in the data sample tend to be wider, a conclusion supported by the p_t-flow plot (figure 7.1a).

Figure 7.4 shows the fraction of the total jet energy detected in the electromagnetic calorimetry (the EM fraction). The data tends to have a higher EM fraction than the simulation, with the difference most pronounced in the leading jet.

Figure 7.5 shows ΔR, the η-ϕ separation of the closest two jets. The data and the simulation agree. There is also agreement in figures 7.6 and 7.7 which plot the jet transverse energy E_t and the pseudo-rapidity η_D respectively. There is a slight discrepancy at $|\eta_D| \approx 1.1$ which may be due to the uninstrumented part of the ΣE_t trigger (see page 44). Agreement between data and simulation for E_t and η_D is important. Of the variables directly measured by the detector, these have the biggest effect on the energy fractions.

Figure 7.8 shows the invariant three-jet mass. As with E_t and η, the two samples agree.

The plots which disagree all have to do with the details of the fragmentation (p_t flow, jet track multiplicity, jet masses, and the jet EM fraction). Therefore the complete physics and detector simulation—PAPAGENO plus SETPRT plus QFL—may simulate the distribution of the jets with respect to each other (*i.e.* bremsstrahlung) but not the fine characteristics of the individual jets, such as the distribution of

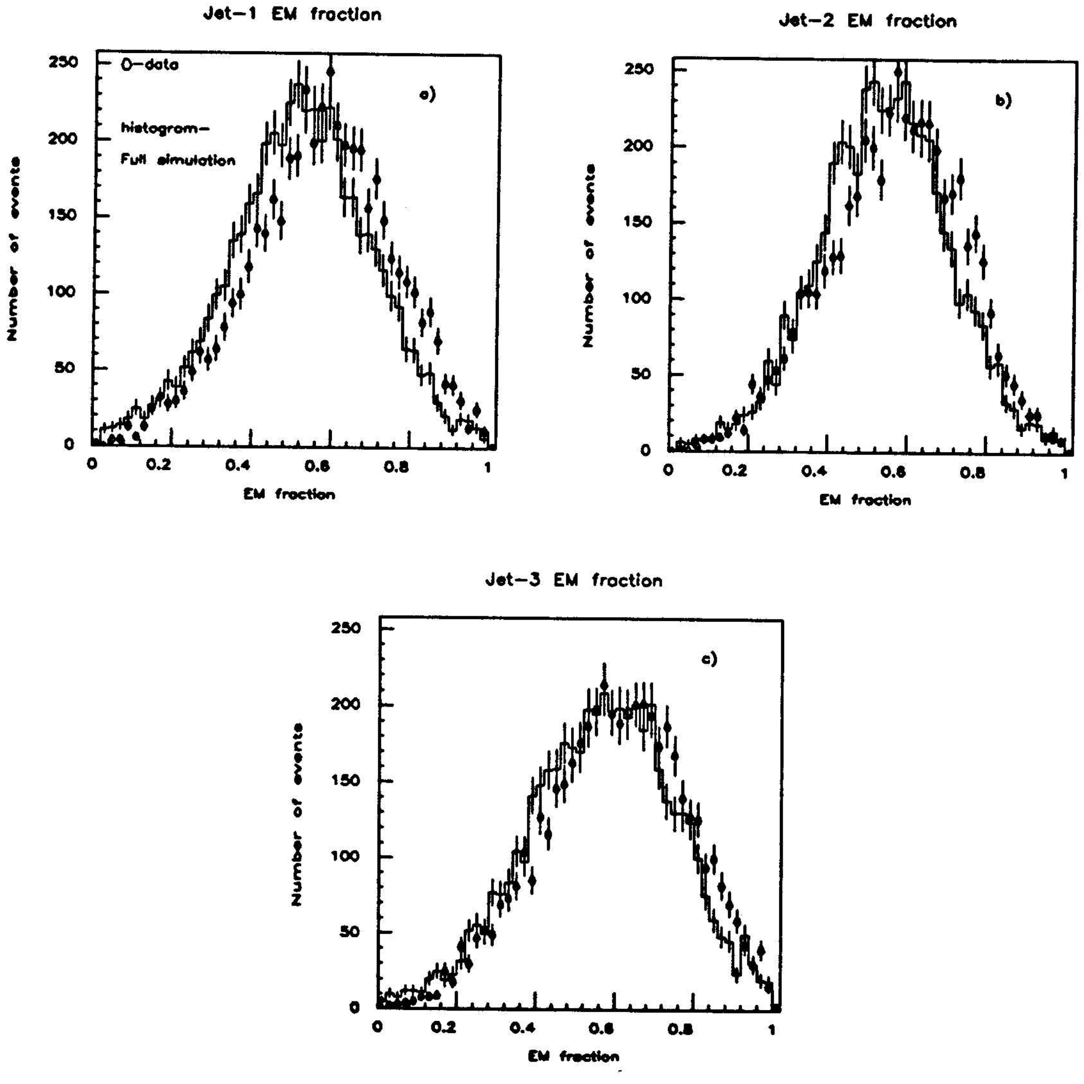

Figure 7.4: Jet EM fraction—the fraction of the total jet energy detected in the electromagnetic calorimeter—for data and simulation. The simulation is area normalized to the data.

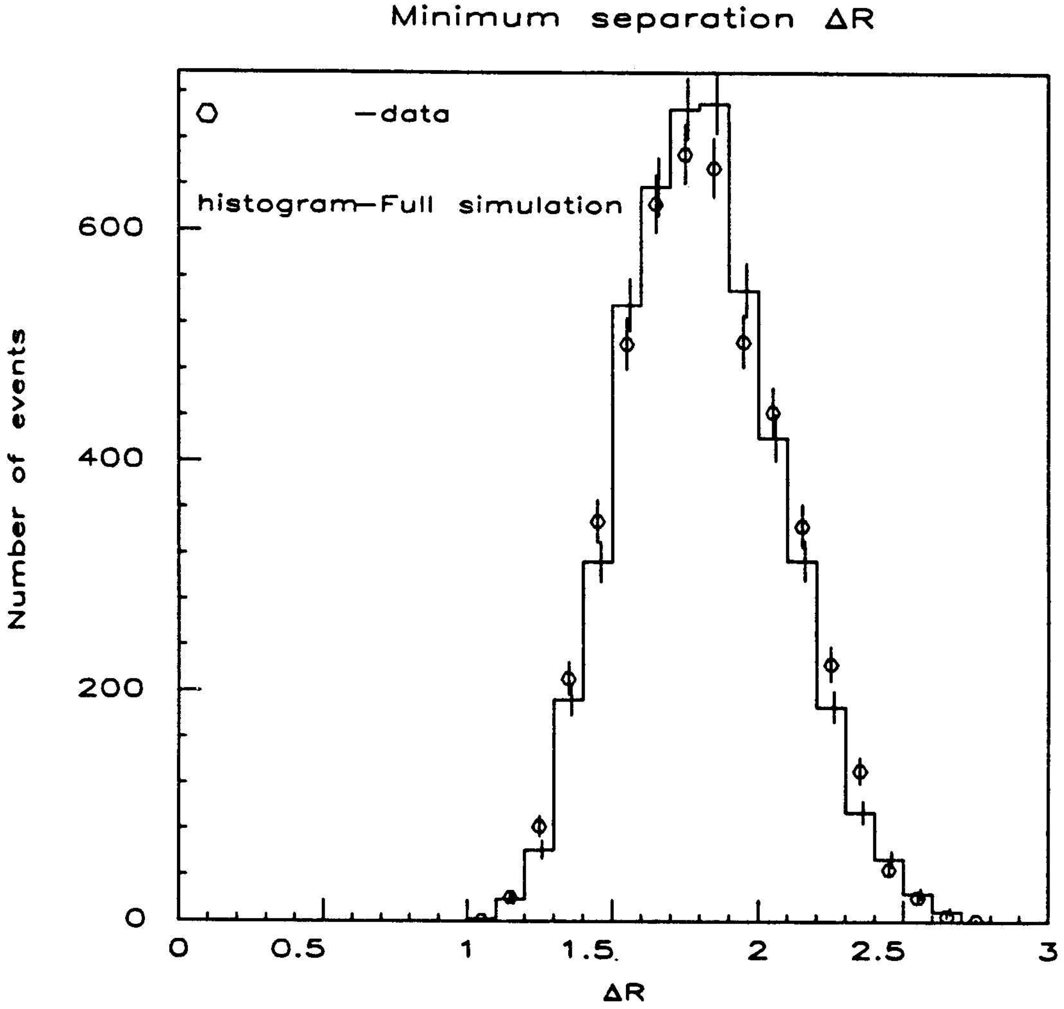

Figure 7.5: The η-ϕ separation of the closest two jets for data and simulation. The simulation is area normalized to the data.

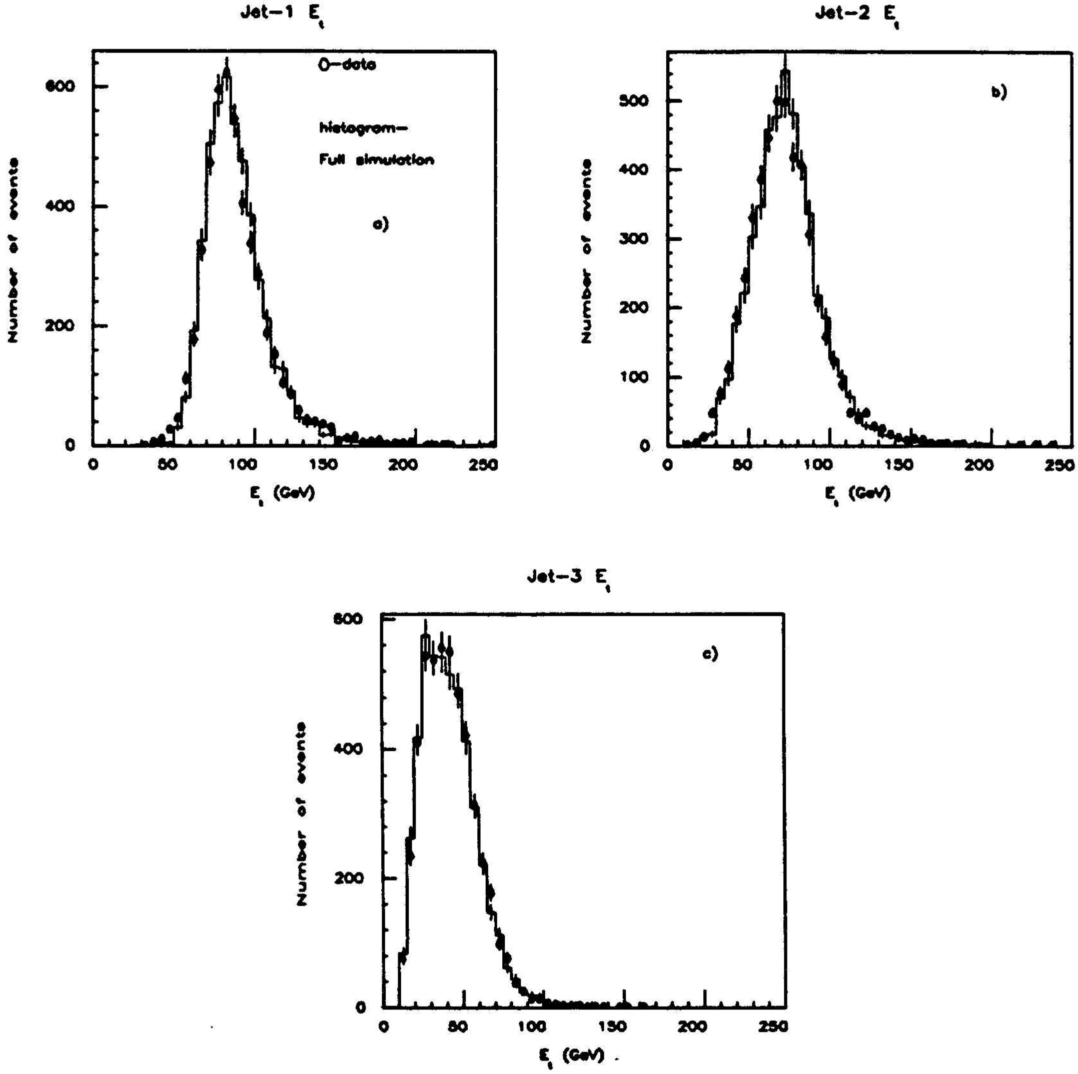

Figure 7.6: Jet transverse energy distributions for data and simulation. The simulation is area normalized to the data.

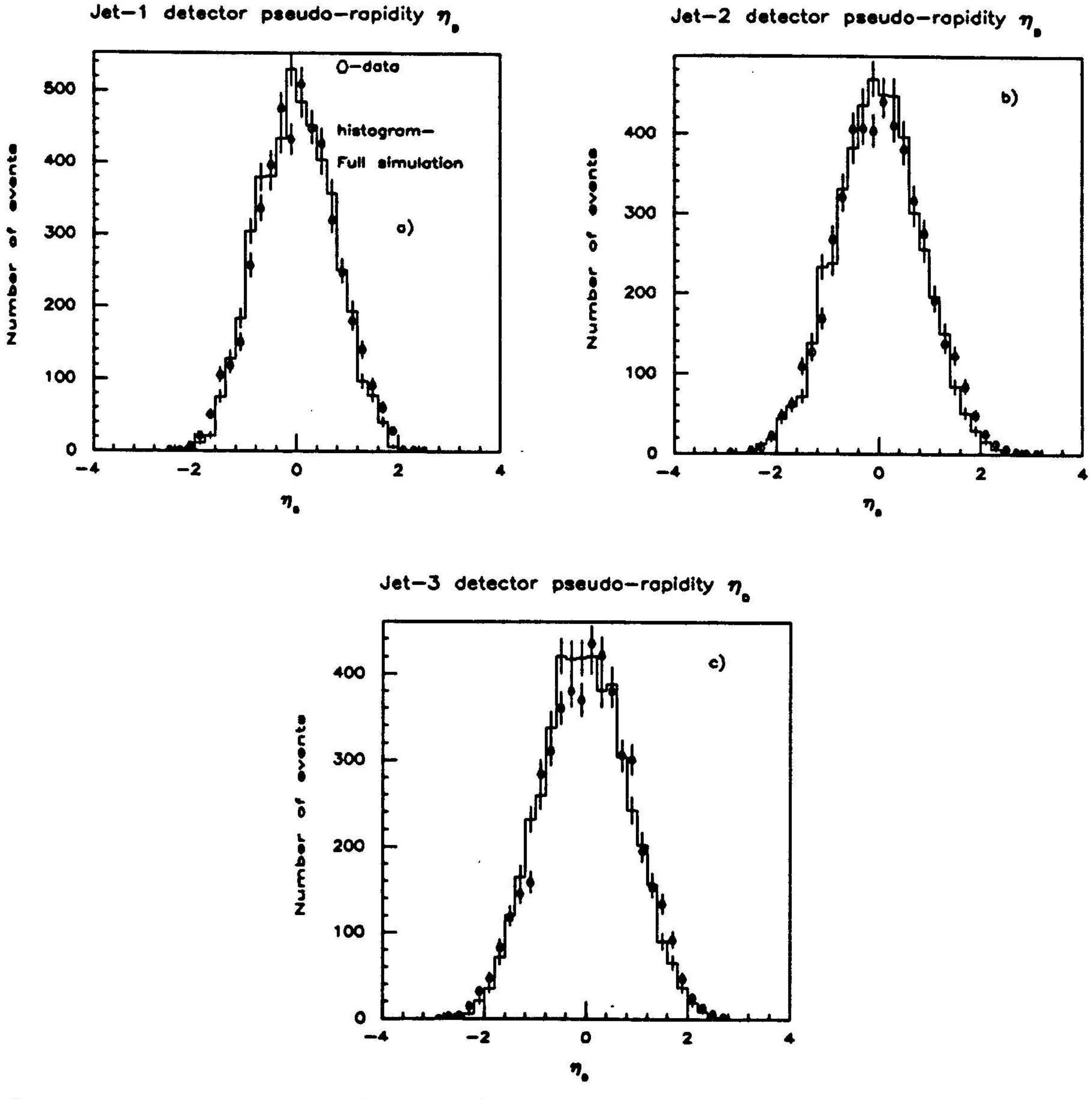

Figure 7.7: Jet pseudo-rapidity distributions for data and simulation. The simulation is area normalized to the data.

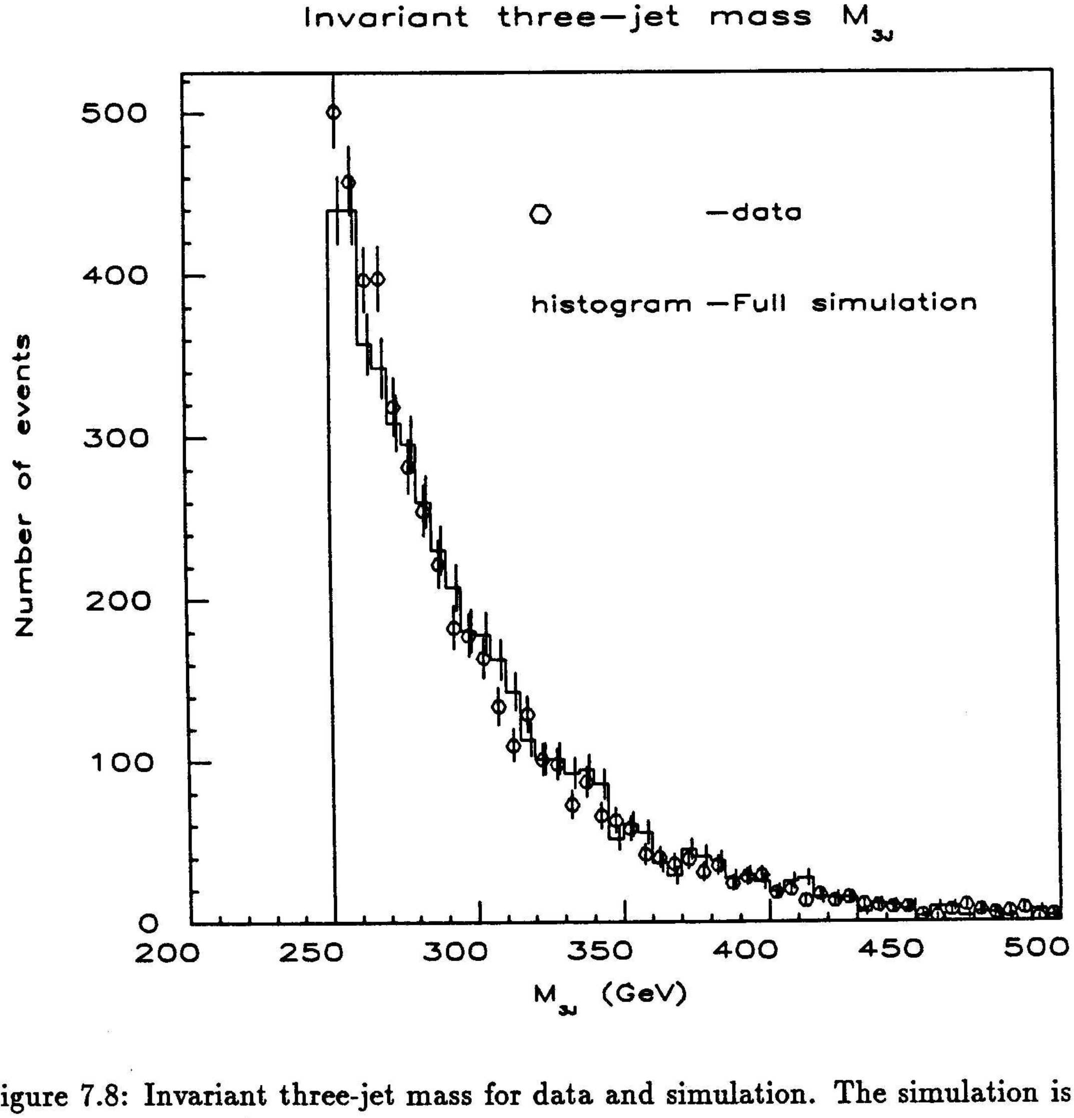

Figure 7.8: Invariant three-jet mass for data and simulation. The simulation is area normalized to the data.

the four-momenta of the particles inside the jets. In effect, this analysis has tried to separate these two parts of the physics and to test only the bremsstrahlung rigorously. The jets are kept well apart, and the variables which are being tested, x_3 and x_4, do not depend directly on the energy distribution inside the jet, only the total energy and location of the jet. In reality, fragmentation and bremsstrahlung are not neatly separable. Instead, the nature of a QCD process is determined by the QCD energy scale Q in equation 2.5 (page 8). Bremsstrahlung corresponds to a high value of Q ($\gtrsim$ 20 GeV) and fragmentation corresponds to a low value of Q ($\lesssim$ 1 GeV). The value of Q which divides processes into those that are bremsstrahlung and those that are fragmentation cannot be precisely defined in calculations, so it is best that these processes have quite different energy scales if they are to be separately simulated.

The QCD energy scale Q is not well defined, but one usually sets it equal to the total energy of the interaction, the momentum transfer, or a related quantity. For this discussion, it can be considered proportional to p_{perp}, the momentum of a particle or jet transverse to the jet from which it was emitted. In fragmentation, p_{perp} does not exceed 20 GeV and rarely exceeds 5 GeV (see figure 7.1c). For a hard bremsstrahlung which would produce the jet in the event with the third highest E_t, the lowest value p_{perp} could possibly be is 5 GeV (assuming the third jet had an energy of 10 GeV and was 1.0 units in η-ϕ space away from the second or first jet); it would most likely exceed 10 GeV. Therefore, these two parts of the jet physics seem to have been separated well enough by the data cuts so that any discrepancies in the fragmentation simulation (*i.e.* SETPRT) should have little effect in the bremsstrahlung simulation (*i.e.* the QCD calculation in PAPAGENO).

7.2.2 Global Event Parameters

A comparison was made between the missing E_T of the two data samples. The missing E_T, $\not{E}_{\mathrm{T}}$, is the vector sum of all the calorimeter towers' transverse momenta. One might expect $\not{E}_{\mathrm{T}}$ to be zero since there is no intrinsic transverse momentum in the $p\bar{p}$ collision. However two effects combine to produce a non-zero $\not{E}_{\mathrm{T}}$. First, a three-jet event receives a transverse boost from low-transverse-energy QCD radiation, an effect explicitly put into the fragmentation simulation (SETPRT). Some of this radiation would appear as very low-energy jets which contain more low-p_t particles ($p_t \lesssim 1$ GeV) than higher energy jets. The proportional energy response for these particles is less than for those of higher energy. Second, the three individual jet energies can fluctuate from their true value independently due to their non-zero energy resolution. One jet may fluctuate high while a jet in the opposite hemisphere may fluctuate low.

A plot of $\not{E}_{\mathrm{T}}$ is shown in figure 7.9. Albeit indirectly, this plot tests the Monte Carlo's simulation of the two effects described above. There is a general agreement between the two samples although the data may have a lower peak. If a slight discrepancy does exist, it should not have much effect on the x_3 and x_4 distributions.

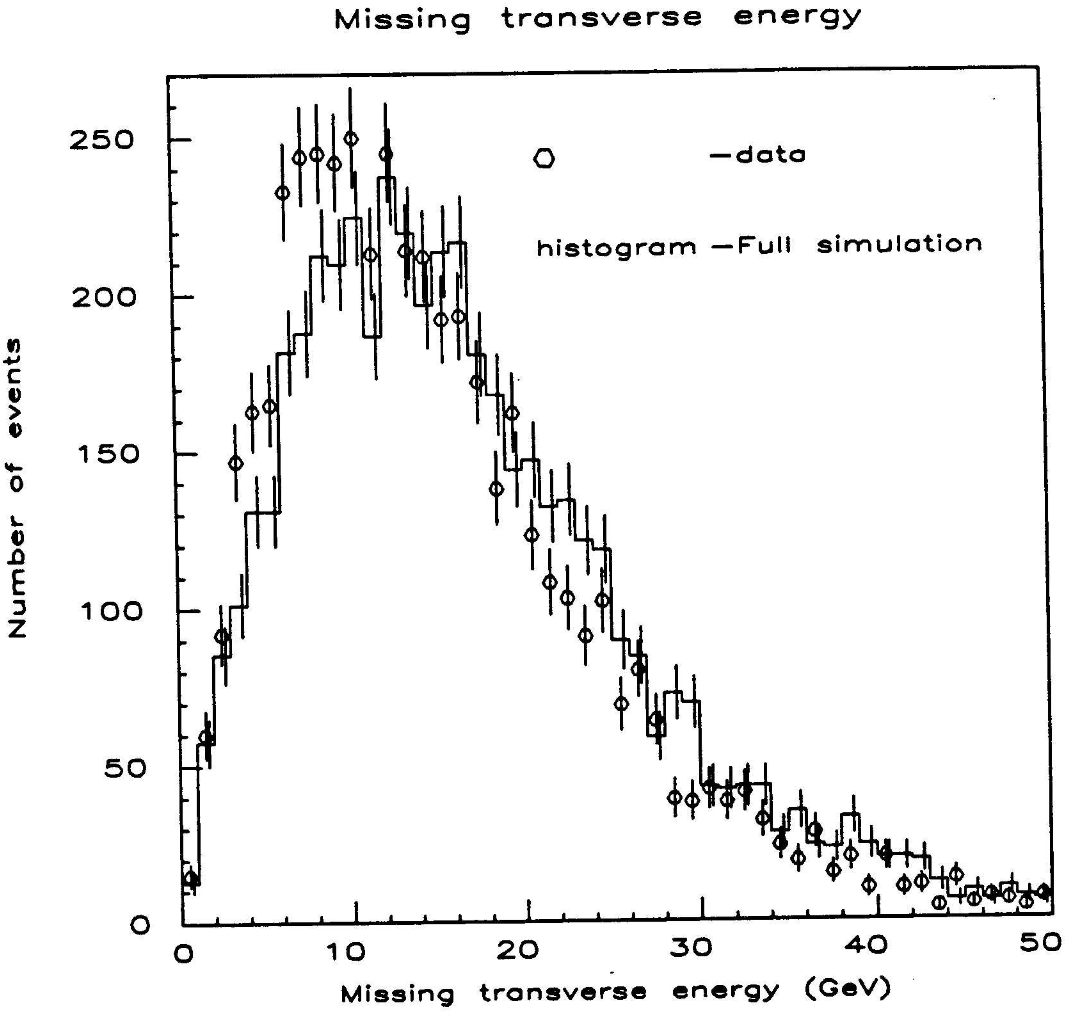

Figure 7.9: Missing transverse energy, $\not{E}_{\mathrm{T}}$, for data and simulation. The simulation is area normalized to the data.

These effects would manifest themselves as a slightly higher contribution to the jet energies by the underlying event, or a small change in the Lorentz boost from the lab frame to the three-jet COM.

Figure 7.10 shows the trigger ΣE_t. The simulation seems to have a narrower peak and a longer tail, but otherwise the two samples agree. Much of the difference in the distributions may be caused by how the ΣE_t is calculated in the trigger electronics. The trigger digitally adds the analog ΣE_t of the five regions of the calorimeter (central and both sides of the plug and forward). Each of the five analog sums can be no greater than 127.5 GeV due to the range of the ADC. Events with a "true" ΣE_t greater than 256 GeV will in general have the trigger underestimate the ΣE_t. In the data, this will deplete the trigger ΣE_t distribution at high values and enhance the distribution at low values. The range of the ADC is not accounted for in the simulation, so this effect is absent. Any discrepancy in the trigger ΣE_t between these two samples should not affect the final result, since the ADC-range effect comes into play above the trigger threshold (120 GeV) and most events have a trigger ΣE_t well above this threshold.

7.3 Checking the Generation Cuts

As noted in the first section of this chapter, the generation cuts must be chosen to keep events which may not pass the equivalent data cuts on the parton level. Because of detector and fragmentation effects, they may pass the data cuts after simulation.

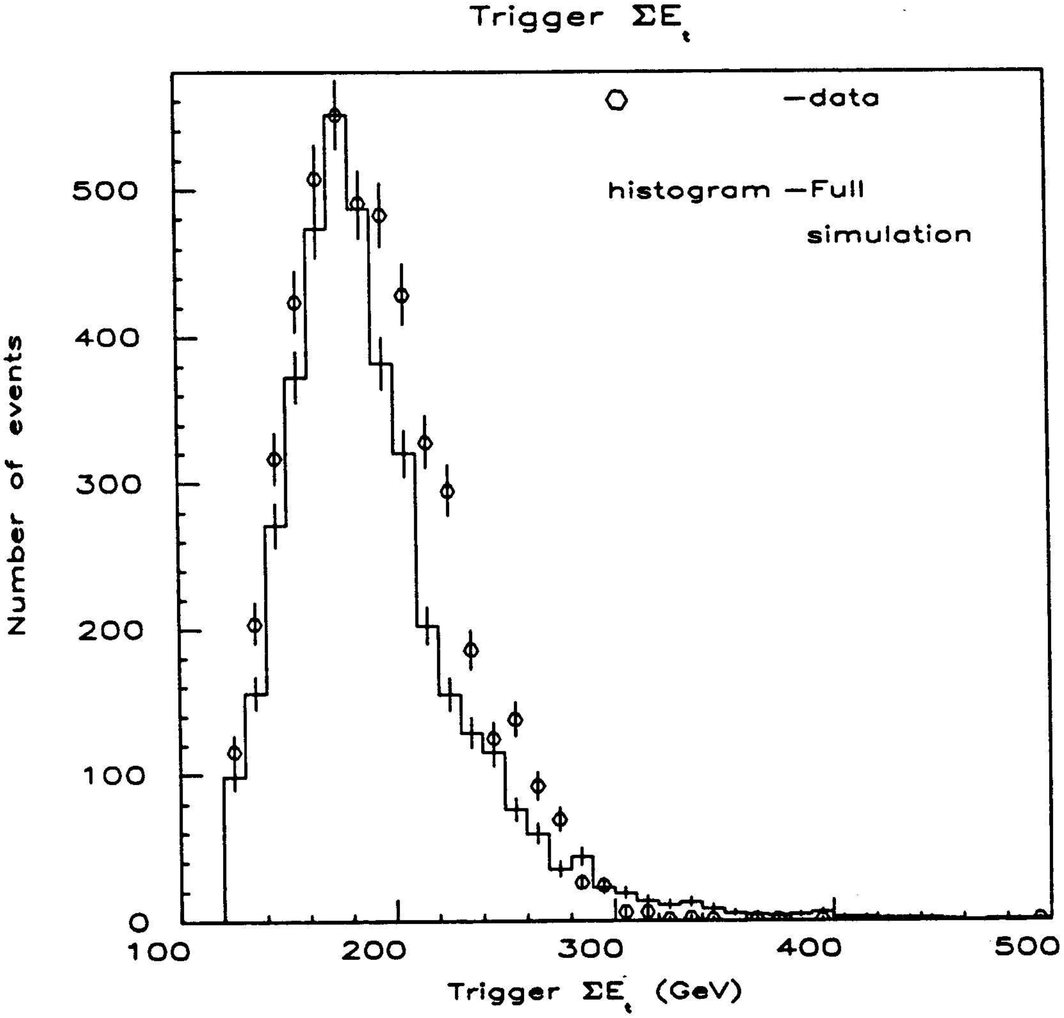

Figure 7.10: Trigger ΣE_t for data and simulation. The simulation is area normalized to the data.

To verify that the generation cuts are loose enough, the dynamic variables of the generated partons were plotted.

Figure 7.11 shows the pseudo-rapidity for the three generated partons, along with the minimum separation ΔR. The three parton-η's are well inside the generation cut of 4.0. Similarly, most events have ΔR well above the 0.7 separation cut. Since there are no sharp cutoffs, the shape of each plot is governed by the detector level data cuts, not the generation cuts.

As an additional check, a faster Monte Carlo with tighter generation cuts was done. The partons were generated with PAPAGENO but the detector was simulated with the inverse jet corrections (see section 5.5). The generation cuts for the fast simulation are

$$p_t > 15.0 \text{ GeV}, \tag{7.9}$$

$$\Delta R > 0.9, \tag{7.10}$$

$$|\eta| < 3.0, \tag{7.11}$$

$$\sqrt{\hat{s}} > 200 \text{ GeV}. \tag{7.12}$$

The distributions from the fast simulation are also shown in figure 7.11. Even with the tighter cuts on η and ΔR, the two Monte Carlo samples agree.

To account for generated events where all the jet energies would fluctuate upward, the cut on $\sqrt{\hat{s}}$ was chosen to be 50 GeV lower than the corresponding data cut on M_{3j}. Figure 7.12 shows $\sqrt{\hat{s}}$ for the full simulation and the fast simulation. In addition

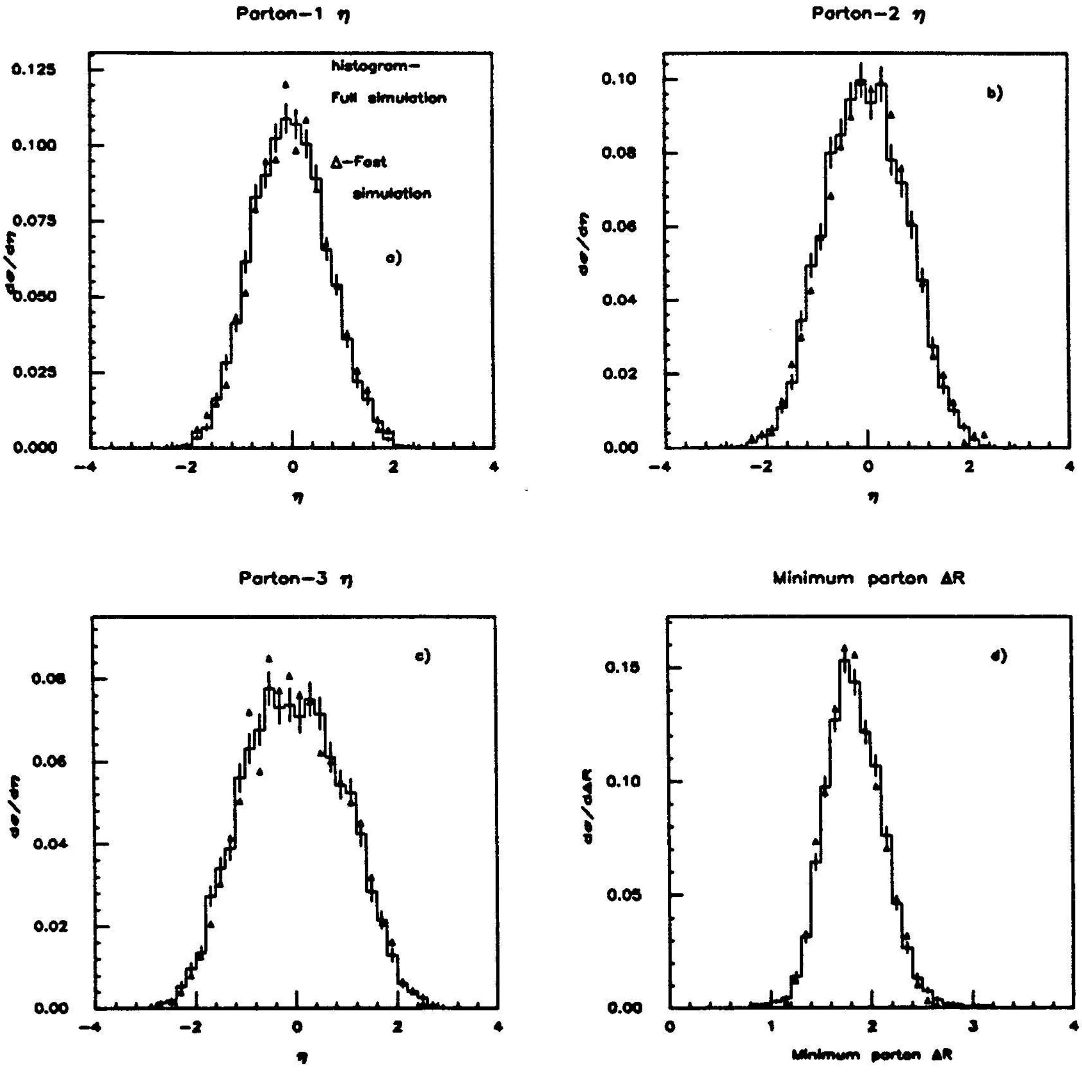

Figure 7.11: Pseudo-rapidity (a, b and c) and minimum separation (d) of generated partons for two simulation programs. The fast simulation is area normalized to the full simulation. Cross sections calculated with the full simulation have a systematic uncertainty of about 50% due to the ambiguity in the Q^2 scale.

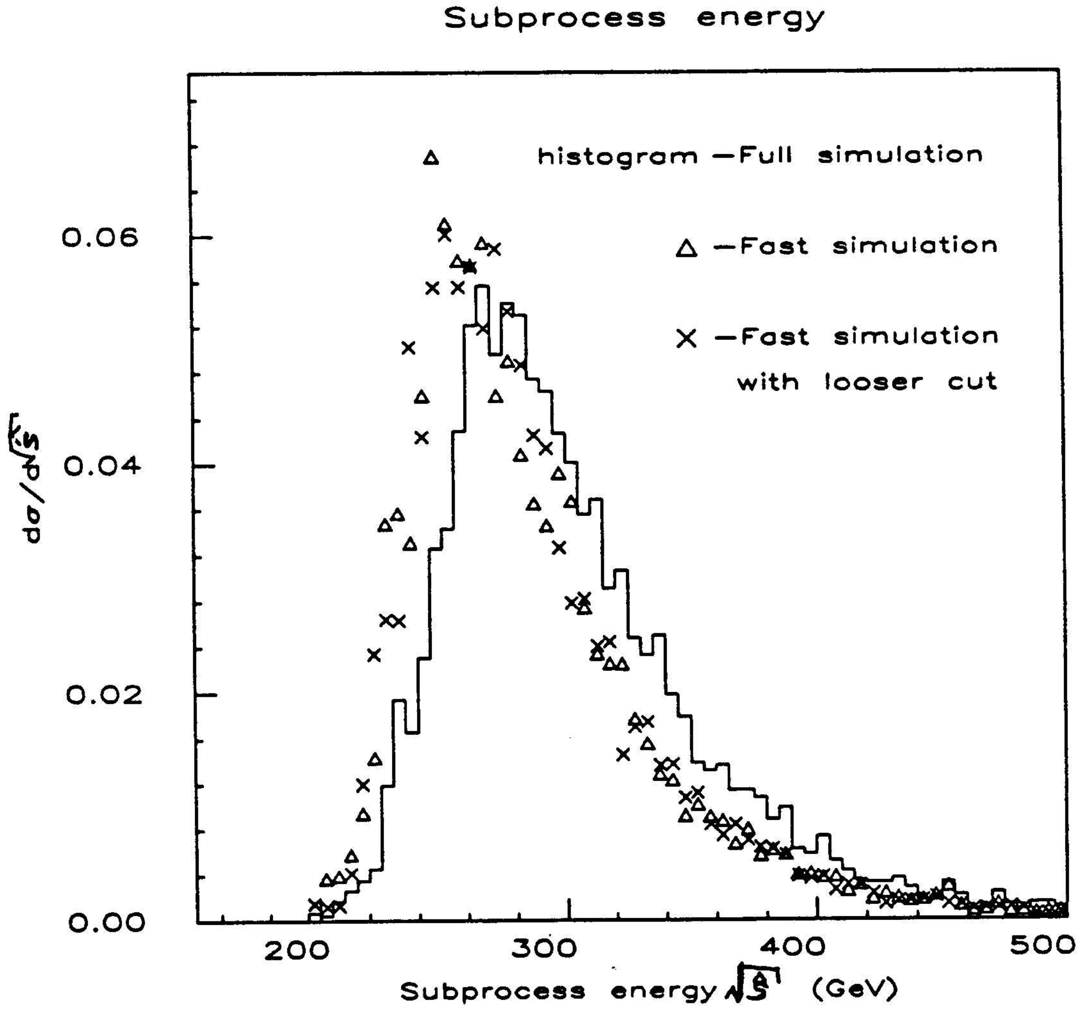

Figure 7.12: Energy of subprocess for different simulations. The fast simulations are area normalized to the full simulation. Cross sections calculated with the full simulation have a systematic uncertainty of about 50% due to the ambiguity in the Q^2 scale.

the distribution from a fast simulation with a $\sqrt{\hat{s}}$ cut of 150 GeV is shown. All three Monte Carlo samples have $\sqrt{\hat{s}}$ above 200 GeV, proving that this cut is sufficiently loose.

The p_t cut in the full simulation corresponds to an uncorrected jet E_t of 9.4 GeV. Considering the resolution is $\sim$ 3 GeV at this E_t, this is rather close to the data cut of 10 GeV. But the three-jet cuts are much tighter than this cut for most topologies. Only 1% of the data have a jet below 15 GeV so the p_t cut is two standard deviations in resolution below the lowest jet E_t for most events. To verify that the parton-p_t cut is really low enough, it was tested in the same manner as the $\sqrt{\hat{s}}$ cut. In addition to the full and fast simulation, a fast simulation was done with the p_t cut at 10 GeV. The parton-p_t's are shown in figure 7.13 for the three simulations. As with the other generation cut parameters, the shape of the distributions is governed more by the data cuts than the generation cuts. In the fast simulation with the looser cut, only $\sim$ 1% of the cross section has a parton with a p_t below the tighter cut.

The full simulation seems to require more parton energy to pass the data cuts. This may due to the non-Gaussian nature of the jet resolution. This should be well modeled by the full simulation, whereas the fast simulation implements a Gaussian resolution for the detected jets. For example, the average uncorrected jet E_t for partons with $p_t = 50$ GeV is 37 GeV. Since the fast simulation applies a Gaussian resolution of 6 GeV, about 0.5% of all generated 50 GeV partons will be converted

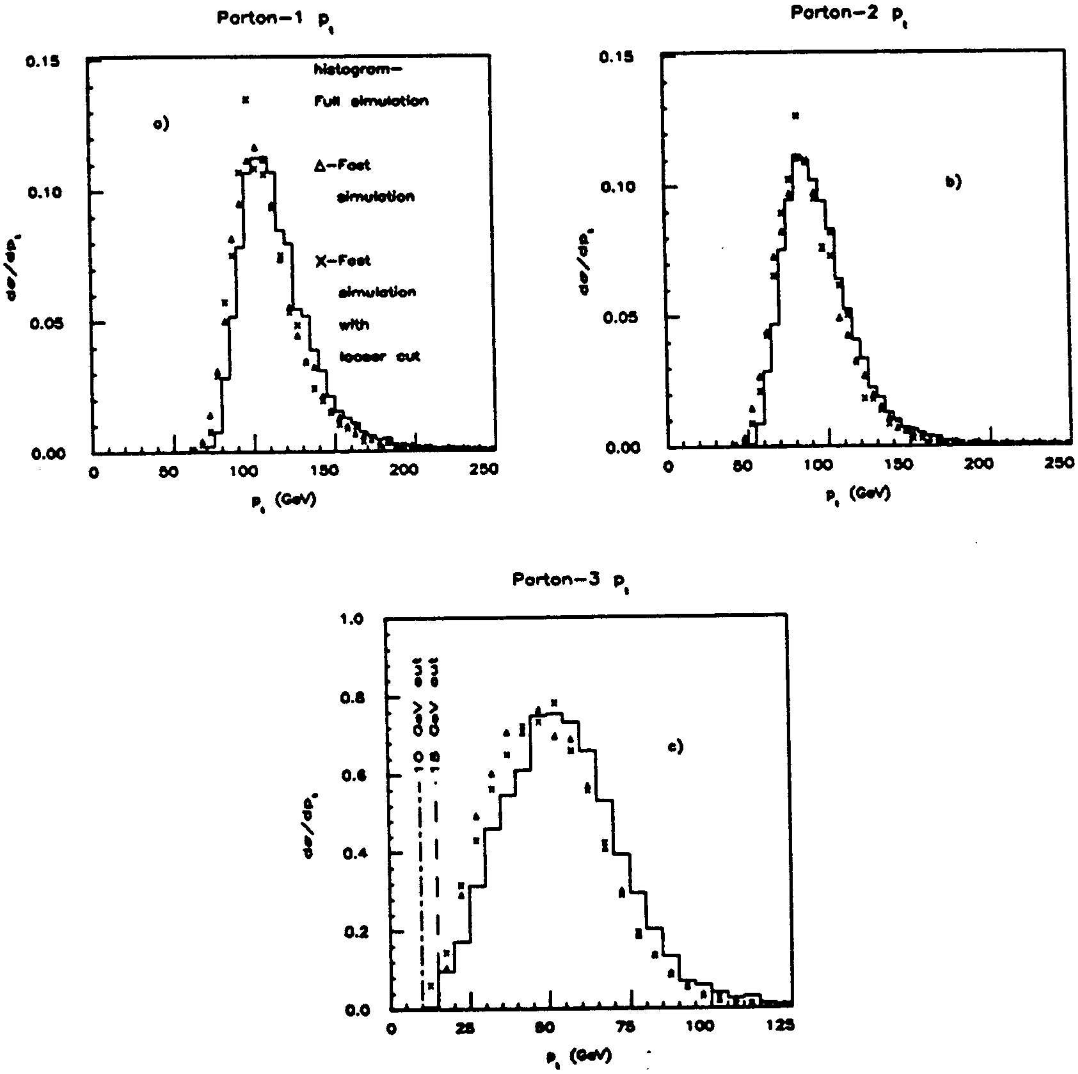

Figure 7.13: Transverse momenta of generated partons for three simulations. All simulations have most of the cross section above the generation cut. The fast simulations are area normalized to the full simulation. Cross sections calculated with the full simulation have a systematic uncertainty of about 50% due to the ambiguity in the Q^2 scale.

into jets with $p_t \lesssim 20$ GeV. However this percentage might be higher in the full simulation if there is a low end tail to the jet E_t distribution. This would correspond to a high end tail in the jet resolution, which has been observed.

Chapter 8

Effects from the Fourth Jet

Of the 4 866 events in the final three jet sample, 4 391 events have a reconstructed fourth jet (the threshold for reconstruction is $E_t \approx 2$ GeV). Of these, 2 235 have the transverse energy of the fourth jet E_t^4 greater than 10 GeV. The fourth jet comes from radiative corrections to $2 \rightarrow 3$ subprocesses (*i.e.* $2 \rightarrow 4$ subprocesses, see figure 8.1).

Since the data is being compared with a calculation to the non-vanishing *leading-order* of α_s, no cut was made on the fourth jet four-momentum. If there were such a cut, it would be hard to compare the data to QCD consistently. As an example, consider an exclusive cut: either requiring E_t^4 not to exceed some threshold, or eliminating all events which have a reconstructed fourth jet. The latter cut is equivalent to the former with a threshold of ~ 2 GeV. For consistency the same cut must be made in the QCD calculation as well as the data. However there are only three final-state partons generated by the calculation. So a fourth jet, if it exists, must be generated by the fragmentation Monte Carlo. The fragmentation does not adequately model

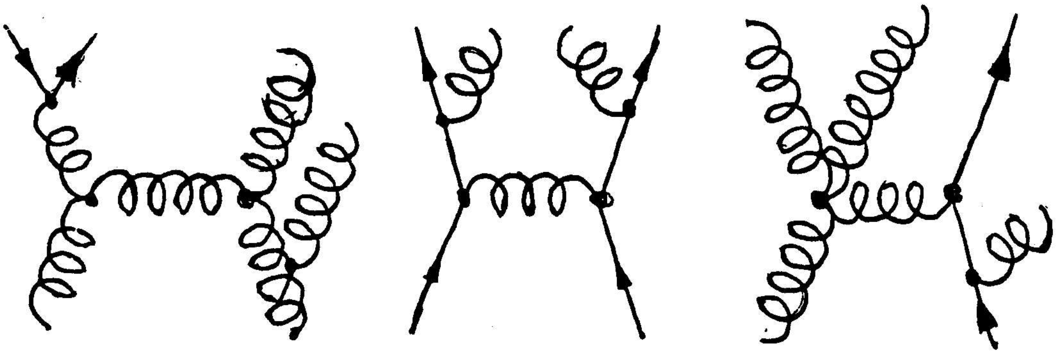

Figure 8.1: Feynman diagrams for radiative corrections to 2 → 3 subprocesses.

the four-momentum distribution of the fourth jet, so the results from the calculation would be incorrect. One could model the fourth jet by extending the calculation to the next highest order of α_s to include 2 → 4 tree diagrams. But then the interference terms between the one-loop 2 → 3 diagrams and their corresponding tree diagrams must be included as well. These one-loop diagrams have yet to be calculated, so a next-to-leading order calculation is impossible at the present time.

Even though the fourth jet is not explicitly treated by the analysis, its effect on the final distributions should be negligible. The 2 → 3 process should be well modeled by the perturbation expansion, so higher order terms should be small. In other words, E_t^4 should be low enough so as not to effect the three leading-E_t jets. To check this the x_3 and x_4 distributions were calculated with a 2 → 4 matrix element using the MMP approximation[28]. The total tree-level 2 → 4 matrix element is approximated by the sum of the following subprocesses: $gg \rightarrow gggg$, $qg \rightarrow qggg$, $q\bar{q} \rightarrow gggg$, $\bar{q}g \rightarrow \bar{q}ggg$,

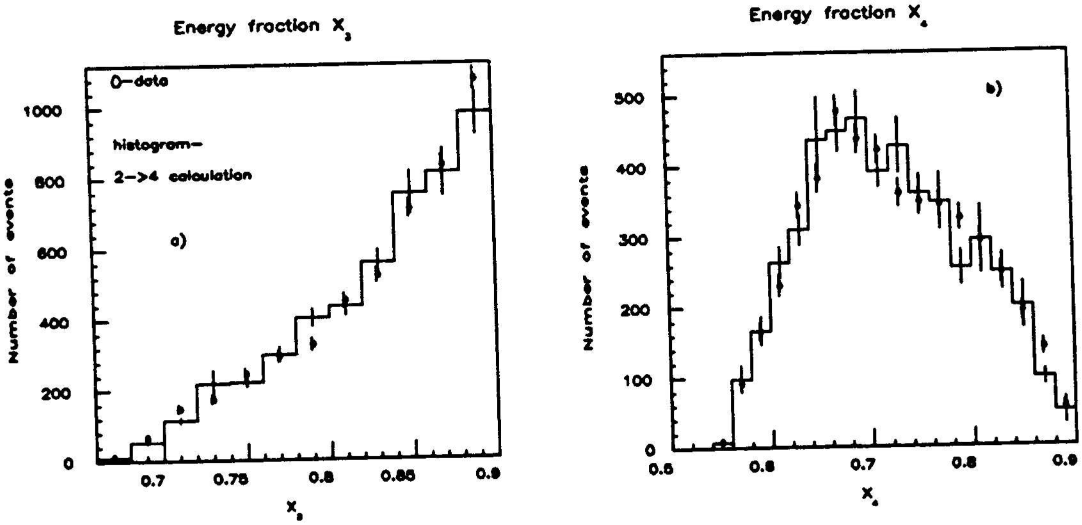

Figure 8.2: Energy fractions x_3 and x_4 for data and 2 → 4 QCD calculation. The QCD calculation is area normalized to the data.

and $gg \rightarrow q\bar{q}gg$. No other 2 → 4 subprocesses are considered. Weighted events were generated using PAPAGENO with primary cuts similar to those used in the 2 → 3 calculation. Three of the partons were required to have a p_t of at least 15.0 GeV, the fourth parton had at least 5 GeV p_t. All four outgoing partons were required to have a pseudo-rapidity $|\eta| < 4.0$ and be separated from each other by 0.9 units of η-ϕ space. The subprocess energy $\sqrt{\hat{s}}$ was required to be above 210 GeV. The cuts on the parton separation and $\sqrt{\hat{s}}$ were tighter than the three-jet calculation to make the Monte Carlo more efficient. The fragmentation and the detector were modeled with the fast simulation.

The x_3 and x_4 distributions from the 2 → 4 calculation are shown in figure 8.2 compared with the data. There is agreement between the two curves. In addition, histograms of E_t^4/E_t^3 and E_t^4/M_{3j} were made for both the data and the $2 \rightarrow 4$ calculation (figures 8.3a and 8.3b). The plots agree except at the low end where E_t^4 is usually less than 10 GeV. The disagreement is due to the poor resolution of the clustering algorithm when $E_t^{\rm jet} < 10$ GeV. If the plots are redone with a cut of $E_t^4 > 10$ GeV the two sets agree (figures 8.3c and 8.3d). This suggests that the behavior of the fourth jet can be explained by a radiative process in the $2 \rightarrow 3$ calculation. Furthermore if there are no cuts on the fourth jet the final distributions should not be affected.

This conclusion is wrong if there is an appreciable contribution from multiparton or rescattering processes. Multiparton interactions have two separate $2 \rightarrow 2$ interactions occurring during a single $p\bar{p}$ collision (see figure 8.4a). Rescattering takes place when one of the outgoing partons in a $2 \rightarrow 2$ process scatters *again* with an as yet unscattered parton in the proton or antiproton (see figure 8.4b). The cross section for these two processes are of the same order of magnitude with rescattering having the lower cross section[29].

To see if one should expect a multiparton signal in the data sample, a calculation was performed using a model of multiparton scattering[29]. For a given two jet event, the chances of another $2 \rightarrow 2$ interaction taking place in a $p\bar{p}$ collision is the cross section for the $2 \rightarrow 2$ interaction divided by the parton flux. Here the parton flux is taken to be the total cross section of the proton. Hence, the multiparton cross section

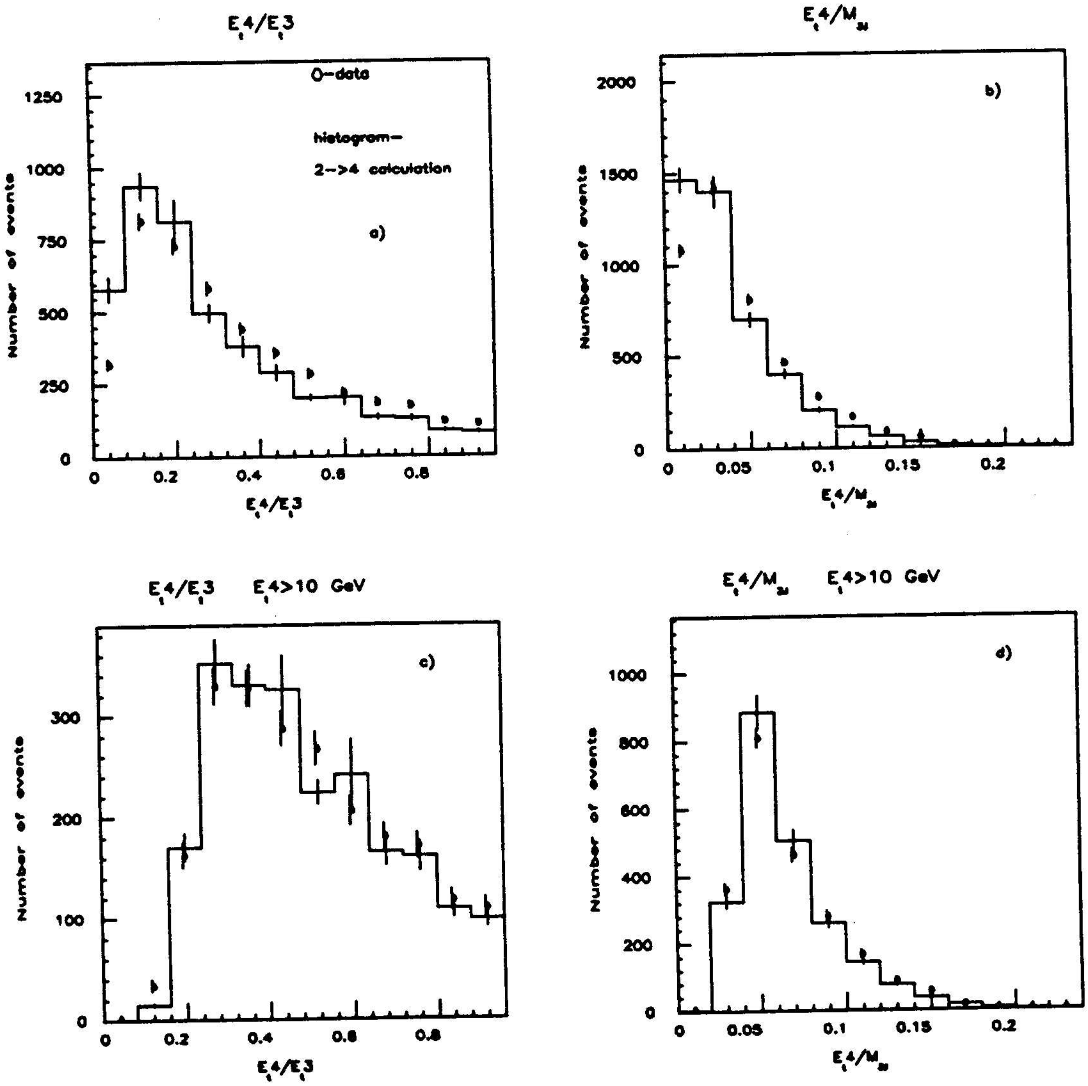

Figure 8.3: Fourth jet transverse energy relative to third jet transverse energy (a and c) and invariant three-jet mass (b and d). a) and b) are without a cut on the fourth jet E_t. c) and d) have a cut of $E_t^4 > 10$ GeV. The QCD calculation is area normalized to the data.

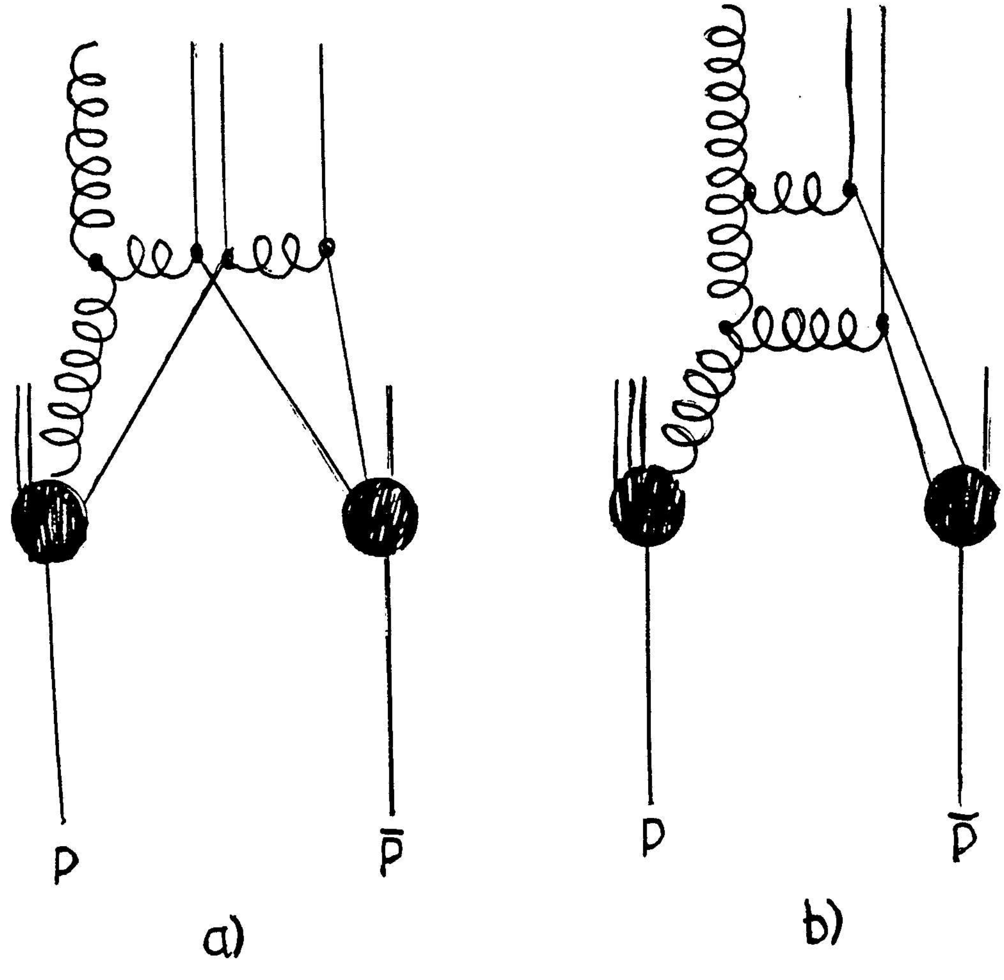

Figure 8.4: a) Multiparton and b) rescattering interactions.

is product of the two two-jet cross sections divided by the total $p\bar{p}$ cross section.

Two assumptions must be pointed out in the theory just described. First, the partons are assumed to be uncorrelated, *i.e.* the momentum distributions of the partons in the second collision are the same as those in the first collision. This assumption is valid provided that the sum of the momentum fractions of two partons coming from one proton is much less then one, and the parton wave functions do not overlap. The first condition is true since the maximum momentum fractions of the partons in typical three-events is about 0.3, and in multiparton interactions the momentum fractions should be less. The second condition was investigated in another thesis[30] and was found to be true for multiparton events at this energy.

Second, if most of the partons are concentrated in the center of the proton, the parton flux may be much less than the total proton inelastic cross section ($\sigma_{\text{tot}} \approx$ 40 mb) thereby increasing the cross section of multiparton events. However the only positive signal for multiparton events measured a parton flux of 5 mb[31], and other experiments, including CDF, set lower limits of the parton flux which were higher[30, 32]. Therefore, for this analysis, the parton flux is assumed to be the same as the total proton inelastic cross section.

PAPAGENO was used to generate two jet events with each outgoing parton required to have p_t greater than 12 GeV and $|\eta|$ less than 4.0. Every two events were merged, thereby creating a Monte Carlo sample with four outgoing partons in each event.

The weights of the merged events were equal to the product of the weights of the two original events. The sum of the weights of all events are in units of nb^2. The total energy of all four partons was required to be less than the total energy of the collision or the event was discarded. In addition the partons were required to be separated by at least 0.7 units in η-ϕ space. Of the 577 578 merged events, 1% failed the total energy cut and 25% failed the separation cut.

The fragmentation and the detector were modeled with the fast simulation. The simulated events were put through the three jet analysis. The total "cross section" of all events passing the three jet cuts was $3.58 \times 10^{-2}\mathrm{nb}^2$. This number was divided by the total cross section of the proton, $\sigma_{\mathrm{tot}} \approx 40$ mb, to get the true multiparton total cross section $\sigma_{MP} \approx 10^{-9}$nb. The total cross section of the three-jet sample is ~ 1 nb, so even if there are large errors in this calculation, the multiparton signal should be practically non-existent.

If there were an appreciable multiparton signal, the x_3 and x_4 distributions would change appreciably (see figure 8.5). The curves for the $2 \rightarrow 3$ QCD calculation (figure 9.2) and the multiparton calculation are quite different, which in principle enables an upper limit on the multiparton cross section to be made (it is usually quoted as a lower limit on the parton flux). From the final three-jet sample, one can calculate the contribution of the multiparton signal to the total three-jet cross section. The lower limit on the parton flux from this calculation is about 4×10^{-7}mb,

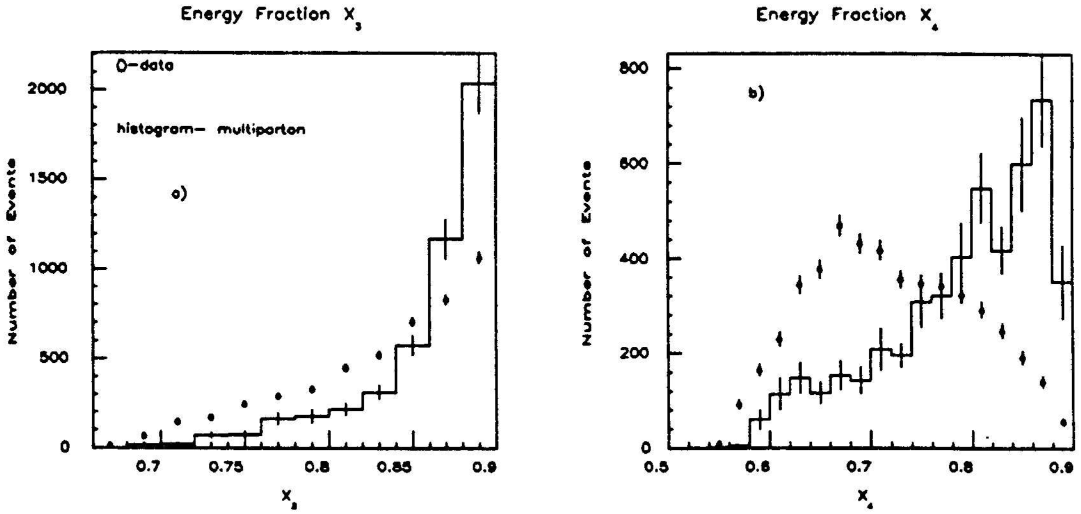

Figure 8.5: Energy fractions x_3 and x_4 for the data and the multiparton calculation. The multiparton calculation is area normalized to the data.

seven orders of magnitude below the observed signal and lower limits discussed above. Therefore, this analysis cannot add anything to the knowledge of the multiparton cross section.

Chapter 9

Results

9.1 Energy Fractions

Figure 9.1 shows the three-jet Dalitz plot along with the projections of x_3 and x_4. The left corner of the Dalitz plot corresponds to events in which all three parton energy fractions are equal. Events from the upper right corner look like dijets with a small third jet. The lower right corner is populated with events in which the second and third highest energy partons each have about 28% of the subprocess energy. There is an enhancement in the density of events in the upper right corner which is consistent with QCD. A phase space matrix element would populate the Dalitz plot uniformly.

The projections along x_3 and x_4 are plotted in figure 9.2 with the prediction of QCD and a phase space matrix element ($|M'| = 1$, equation 2.9, page 15) area normalized to the data. Both QCD and phase space distributions include the effects of the detector simulated with the full simulation. QCD reproduces the x_3 and x_4 distributions quite well while phase space is in sharp disagreement. The χ^2's for the

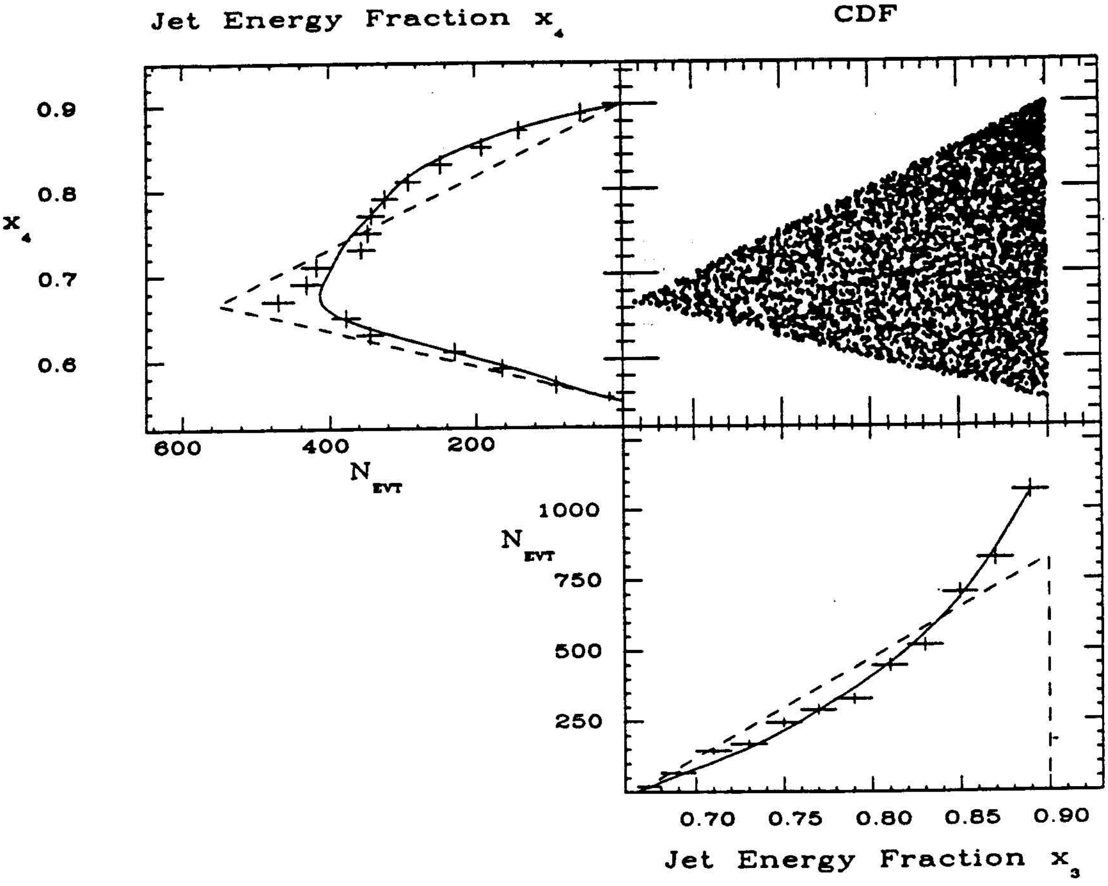

Figure 9.1: Three-jet Dalitz plot. The projections are plotted along with a smoothed QCD curve (solid line) and an area projection (dashed line); both are area normalized to the data.

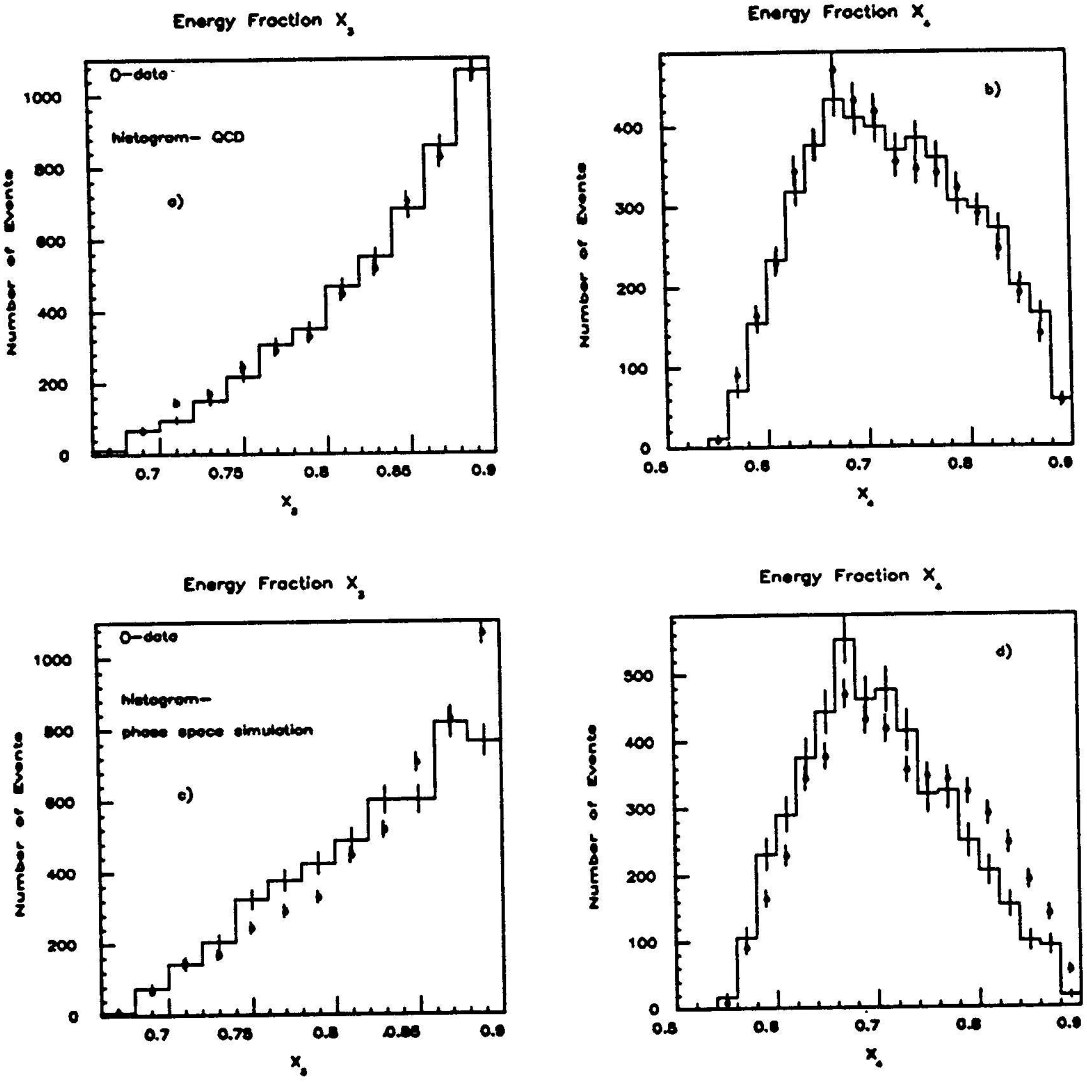

Figure 9.2: Energy fractions x_3 and x_4. a) and b) show the data and QCD; c) and d) show the data and a phase space simulation. In all plots, the calculation (whether QCD or the phase space simulation) are area normalized to the data.

	chi^2	degrees of freedom	confidence level
x_3	15.9	11	85.3%
x_4	13.8	17	32.2%

Table 9.1: χ^2 for comparison between data and the QCD calculation with the full simulation.

comparison between the data and QCD curves are shown in table 9.1. As figure 9.1 confirms, the data is consistent QCD. The phase space matrix element is disproved with a confidence level better than 99.99%; the χ^2's are 65 and 98 for x_3 and x_4.

The systematic error of the QCD calculation with the full simulation was assumed to be negligible compared to the statistical error. If this assumption is valid, the data and QCD would still agree if the detector simulation or the generation cuts were changed. Therefore, as a check, a χ^2 calculation was done with the fast simulation using the cuts on page 104 (expressions 7.9 through 7.12). Additional χ^2-comparisons were done with fast simulation runs having different generation parameters; two of the cuts were loosened ($p_t > 10$ GeV and $\sqrt{\hat{s}} > 150$ GeV), the EHLQ1 structure functions were replaced with MRS1 or MRS2[33], and the Q^2-value was changed from $Q = \sum p_t/6$ to $Q = \sum p_t$. Each of these changes was done in separate runs with the other parameters held constant. Overall there were six runs done with the fast simulation. All of these runs are consistent with the data.

Figures 9.3 and 9.4 show the results from two previous experiments, UA1[4, 34]

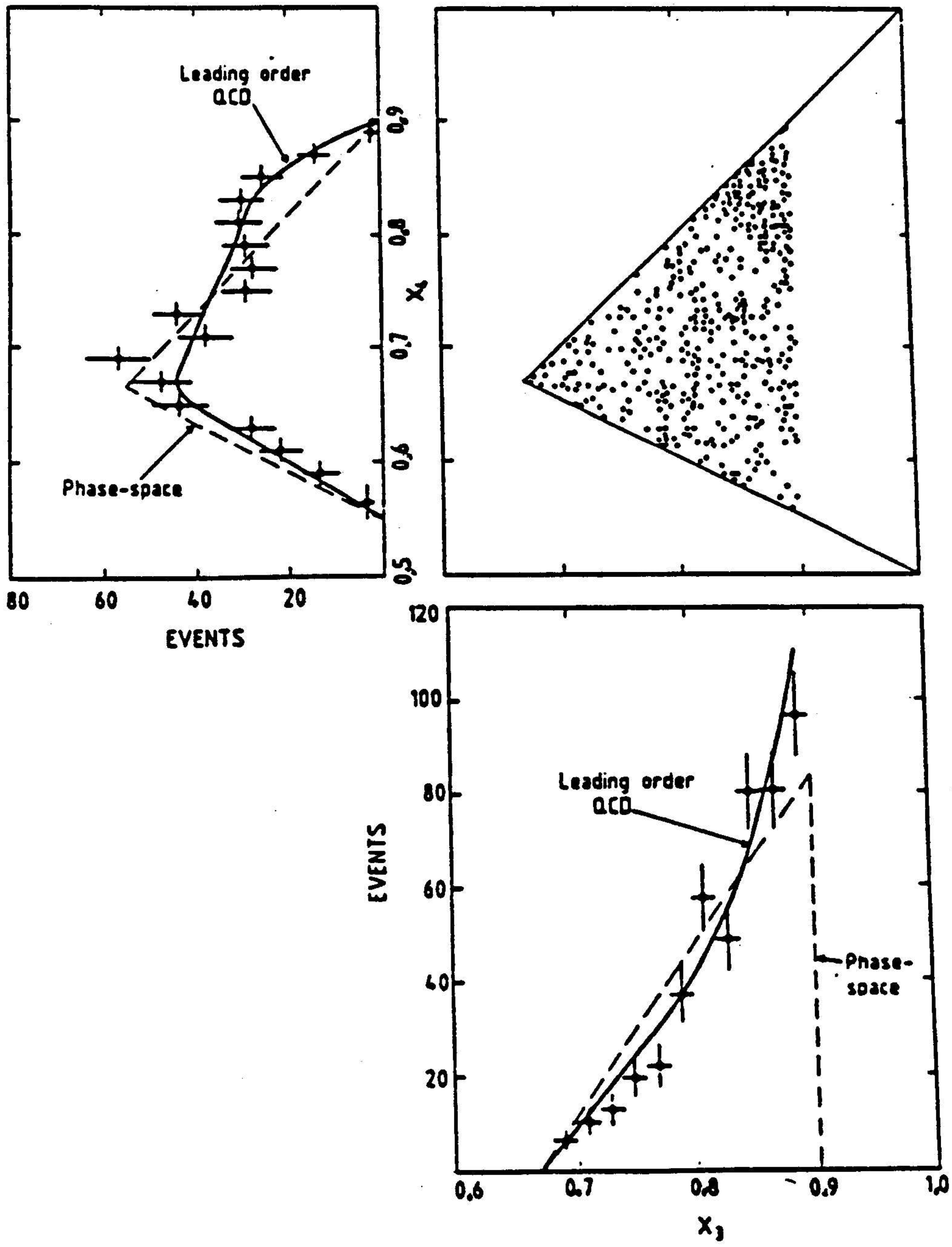

Figure 9.3: Three-jet Dalitz plot from UA1. The data was taken from reference [34].

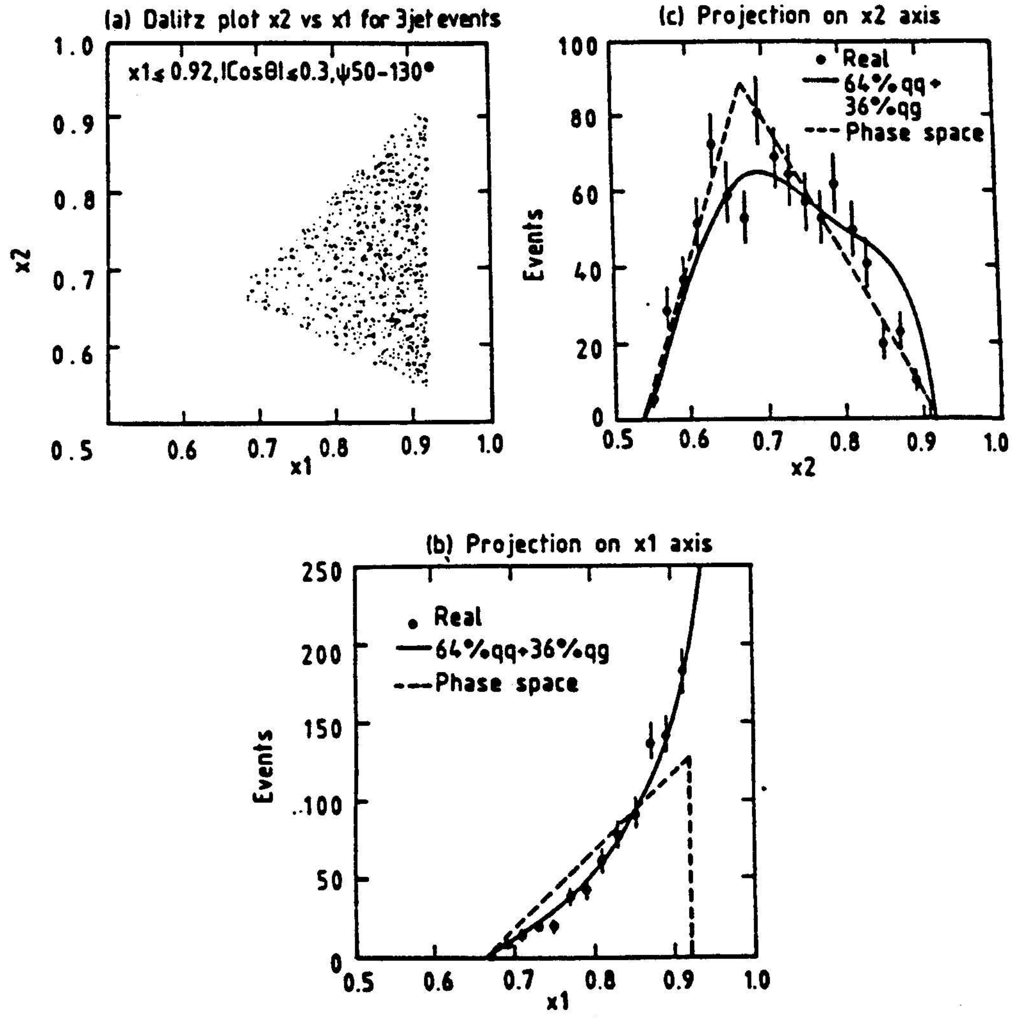

Figure 9.4: Three-jet Dalitz plot from CMOR[6]. The variables x_1 and x_2 in the convention used by CMOR correspond to x_3 and x_4 in this analysis.

and CMOR[6]. The UA1 experiment took $p\bar{p}$ data with energies of $\sqrt{s} = 546$ GeV and 630 GeV at the CERN Sp$\bar{\text{p}}$S. The CMOR experiment took pp data at the CERN Intersecting Storage Rings (ISR) with an energy of $\sqrt{s} = 62.3$ GeV. Both experiments used a jet reconstruction algorithm and a three-jet analysis method similar to that of CDF. These experiments show a qualitative agreement between data and QCD. This analysis, with ten times more statistics, makes the agreement definitive.

The x_3 and x_4 acceptance were calculated with the full simulation. The parton level distributions were computed with cuts on the partons exactly the same as the data cuts on the jets with two exceptions. The E_t cut was raised to 16.19 GeV to correspond to an uncorrected jet E_t of 10 GeV and the η_D cut was applied to the parton's physical pseudo-rapidity, not the η_D of the jet it produced. The distribution of jets reconstructed by the analysis program after full simulation of the detector was divided by the parton level distributions bin by bin. Figure 9.5 shows the acceptance plots. The acceptance is uniform; therefore it was not considered further.

Figure 9.6 shows the x_3 and x_4 distributions from QCD, decomposed into three different initial-state configurations, compared with the data. The separate curves for the different subprocesses were calculated with the fast simulation. The curves for the subprocesses with gg, qg and $\bar{q}g$ in the initial state agree with the data. The curve which represents subprocesses with $q\bar{q}$ in the initial state disagrees with the data and the other curves. This is expected since interactions involving initial-

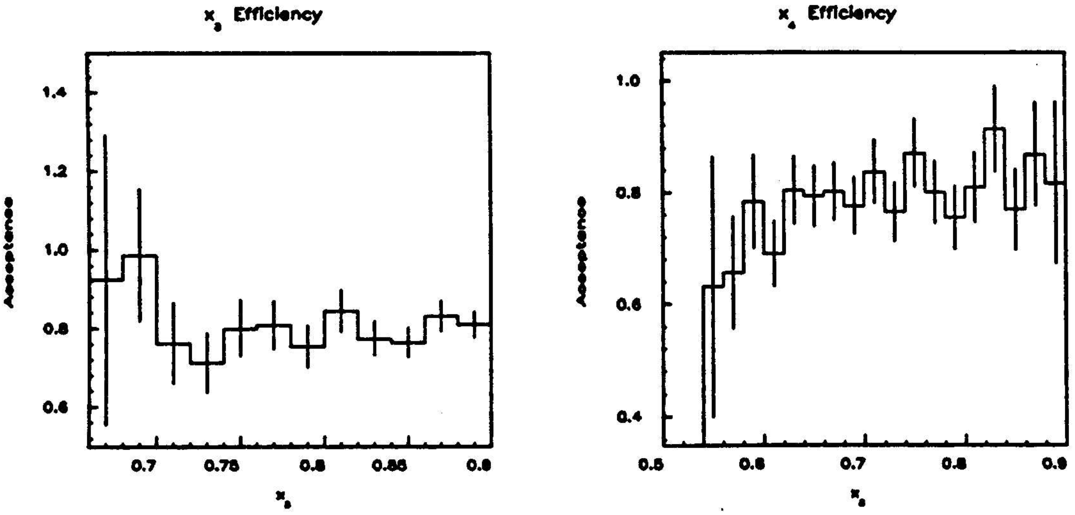

Figure 9.5: Efficiency for x_3 and x_4 as determined by the simulation.

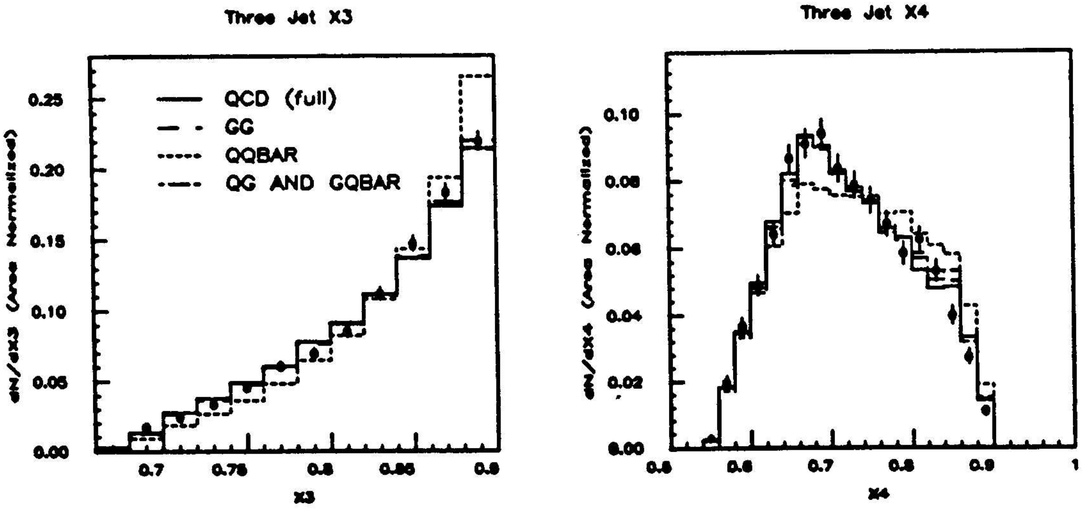

Figure 9.6: Energy fractions x_3 and x_4 for different initial states, compared with the data. All plots are area normalized to the data. The separate curves for the different subprocesses were calculated with the fast simulation.

Subprocess	Contribution to total cross section
$gg \rightarrow ggg$	36.1%
$qg \rightarrow qgg$	21.7%
$g\bar{q} \rightarrow gg\bar{q}$	21.5%
$q\bar{q}' \rightarrow gq\bar{q}'$	4.1%
$q\bar{q} \rightarrow gq\bar{q}$	4.4%
$gg \rightarrow gq\bar{q}$	4.6%

Table 9.2: Contribution to the total cross section from different subprocesses. Only those contributions exceeding 4% are listed; the other subprocesses contribute a total of 7.6%.

gluons are preferred due to the structure functions of the proton. Table 9.2 shows the contributions to the total cross section for different subprocesses. The subprocesses with a gluon in the initial state contribute to 80% of the total cross section.

9.2 Three-Jet Angles

Figure 9.7 show the three-jet angles for both the data and the full simulation. The angles $\cos\theta^*$ and ψ^* were described in chapter 2. They represent the leading jet polar angle and the orientation of the three-jet plane. The variable $\cos\xi_{EK}$ is the Ellis-Karliner angle[35]. It is the angle between parton-4 and parton-3 in the COM frame

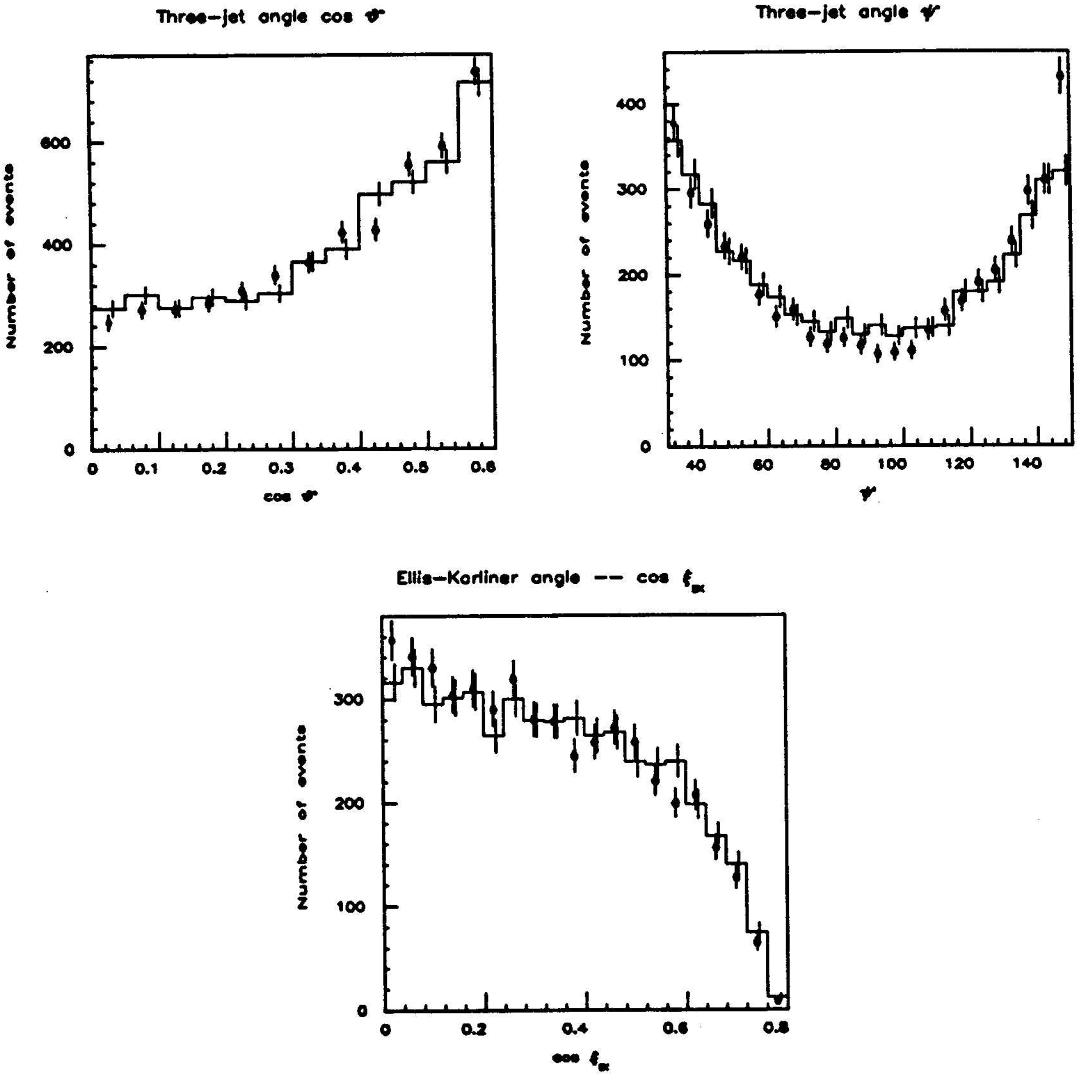

Figure 9.7: Three-jet angles for data and QCD. The QCD prediction is area normalized to the data.

of partons-4 and -5 and is related to the energy fractions by the simple relation

$$\cos \xi_{EK} = \frac{x_3 + 2x_4 - 2}{x_3}. \tag{9.1}$$

The variable $\cos \xi_{EK}$ increases as one approaches the upper right corner of the Dalitz plot. As with the energy fractions, all three plots show agreement between the data and the QCD model.

9.3 Conclusion

The production of three-jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV was studied with the CDF detector. Cuts were made on the kinematic variables of the three-jet system to make the detector acceptance uniform in the area of phase space which was studied. In addition these cuts keep the angles between the jets large enough so that only events with a hard gluon bremsstrahlung were passed. The distributions of the fractional energy carried by the leading two jets in the three-jet COM system, x_3 and x_4, were compared to predictions of leading order QCD calculations.

The QCD parton model predicts the rate of hard bremsstrahlung using a convolution of structure functions and matrix elements coming from the Born approximation. QCD perturbation theory only specifies the matrix elements for the different subprocesses ($gg \to ggg$, $qg \to qgg$, *etc.*). It is the purpose of this analysis to test these matrix elements only, independently of the other parts of the calculation—the structure functions, the momentum scale of the interaction (Q), the parameters used in generating the $2 \rightarrow 3$ events (*i.e.* the generation cuts), and the details of the fragmentation and detector simulation. All of these other parts, except for the fragmentation model, were varied. The final distributions of x_3 and x_4 did not change.

Only one fragmentation model was used[21]. However, the physics of the fragmentation should be independent of the hard bremsstrahlung process, and therefore should not affect the test of the matrix elements. In other words, the internal details of the jets—such as the location and momentum fraction of each particle inside the jet—should affect the distributions of x_3 and x_4 only slightly. To first order the energy fractions depend only on the total energy and location of the jet. To insure that the energy fraction distributions are not affected by the choice of the fragmentation model, it is sufficient to verify that the model used gives a reasonable, though not necessarily exact, reproduction of the data. This has been done; only a few discrepancies were observed and it was concluded that these were negligible.

The calculation of the energy fraction distributions depends mostly on the matrix elements as derived by QCD perturbation theory; it is largely independent of other factors. Furthermore, no acceptance correction is needed for the distributions since the efficiency is uniform. Therefore, this analysis tests QCD perturbation theory as directly as possible. To the precision of the final sample (~ 5000 events for both data and QCD samples), QCD reproduces the data.

Bibliography

[1] MARK-J Collaboration, D. P. Barber *et al.*, Phys. Rev. Lett. **43**, 830 (1979).

[2] TASSO Collaboration, R. Brandelik *et al.*, Phys. Lett. **97**B, 453 (1980).

[3] F. A. Berends *et al.*, Phys. Lett. **103**B, 124 (1985).

[4] UA1 Collaboration, G. Arnison *et al.*, Phys. Lett. **158**B, 494 (1985).

[5] JADE Collaboration, W. Bartel *et al.*, Phys. Lett. **119**B, 239 (1982).

[6] CMOR Collaboration, A. L. S. Angelis *et al.*, Nucl. Phys. B**303**, 569 (1988).

[7] This chapter is partly based on material in Vernon D. Barger and Roger J. N. Phillips, *Collider Physics*, (Addison-Wesley, Redwood City, CA 1987) and Paul Langacker, *Grand Unified Theories and Proton Decay*, Phys. Rep. **72**, 185 (1981).

[8] F. Abe *et al.*, Nucl. Instr. Meth. A**271**, 387 (1988).

[9] F. Snider *et al.*, Nucl. Instr. Meth. A**268**, 75 (1988).

[10] F. Bedeschi *et al.*, Nucl. Instr. Meth. A**268**, 50 (1988).

[11] H. Minemura *et al.*, Nucl. Instr. Meth. **A238**, 18 (1985).

[12] S. R. Hahn *et al.*, Nucl. Instr. Meth. **A267**, 351 (1988).

[13] S. Errede and W. Trischuk, CDF Note #1081, *E/P Calibration of the CEM Calorimeter.* This note is an internal memorandum of CDF (CDF notes can be obtained from the CDF Dept., M.S. 223, Fermilab).

[14] S. Bertolucci *et al.*, Nucl. Instr. Meth. **A267**, 301 (1988).

[15] S. Behrends, M. D. Shapiro and E. Pare, CDF Note #1066, *Single Pion Response in the Central Calorimeter*; S. Behrends, S. Kuhlmann, D. Amidei and F. Rimondi, CDF Note #583, *Response of the Central Calorimeter to Low-Energy Charged Particles.* These notes describe the technical details of the non-linearity analysis and the many checks performed on the conclusions. The results used in the detector simulation are described in section 3.4.4.

[16] D. Amidei *et al.*, Nucl. Instr. Meth. **A269**, 51 (1988).

[17] D. Brown *et al.*, CDF Note #605, *Clustering algorithms and their performance.*

[18] The corrections described in this chapter are later versions (using data taken in 1988) of the work described in the following CDF notes: D. Brown, CDF Note #835, *A Study of Dijet P_t Balancing using 1987 Data*; D. Brown and R. Carey,

CDF Note #755, *Jet Corrections over the full CDF η Range*; D. Brown and R. Carey, CDF Note #754, *A Study of the Underlying Event in Jet Events.*

[19] D. Brown and S. Kannapan, CDF Note #873, *SETPRT, a Single Parton Fragmentation Module.*

[20] F. Paige and S. Protopopesceau, BNL Report No. BNL 38034, 1986 (unpublished).

[21] R. D. Field and R. P. Feynman, Nucl. Phys. B**136**, 1 (1978).

[22] V. A. Abramovskiĭ, V. N. Gribov and O. V. Kancheli, Sov. J. Nucl. Phys., **18**, 308 (1974) (translated from Yad. Fiz. **18**, 595 (1973)).

[23] D. Brown, S. Kannapan and M. Shapiro, CDF Note #753, *QFL Version 2: Improvements and Checks.*

[24] T. Hessing *et al.*, CDF Note #1131, *Monte Carlo Fragmentation Tuning for Jet Energy and Resolution Analysis.*

[25] UA2 Collaboration, P. Bagnaia *et al.*, Phys. Lett. **144**B, 283 (1984).

[26] I. Hinchliffe, in preparation.

[27] E. Eichten *et al.*, Rev. Mod. Phys. **56**, 579 (1984).

[28] C. J. Maxwell, Phys. Lett. B**192**, 190 (1987); C. J. Maxwell, Nucl. Phys. B**316**, 321 (1989); M. Mangano and S. Parke, Phys. Rev. D**39**, 758 (1989).

[29] N. Paver and D. Treleani, Z. Phys. C**28**, 187 (1985).

[30] David Nathan Brown, Ph.D. thesis, Harvard University, 1989.

[31] AFS Collaboration, Z. Phys. C**34**, 163 (1987).

[32] UA2 Collaboration, *A Study of Four-Jet Events at the CERN $P\bar{P}$ Collider*, Proc. XXI Rencontre de Moriond (1986).

[33] A. D. Martin, R. G. Roberts and W. J. Stirling, Phys. Rev. D**37**, 1161 (1988).

[34] Elizabeth Jane Buckley, Ph.D. thesis, University of London, 1986.

Reference [4] contains data taken during a run in 1983. This reference has data from the run in 1983 combined with data from a run in 1984.

[35] John Ellis and Inga Karliner, Nucl. Phys. B**148**, 141 (1979).