NORTHWESTERN UNIVERSITY

Strongly-Interacting Color-Singlet Exchange in $\overline{p}$-p Collisions at $\sqrt{s} = 1800$ GeV

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Physics

By

Tracy Lea Taylor Thomas

EVANSTON, ILLINOIS

December 1997

ABSTRACT

Strongly-Interacting Color-Singlet Exchange in $\overline{p}$-p Collisions at $\sqrt{s} = 1800$ GeV

Tracy Lea Taylor Thomas

Results are presented from an analysis of the particle multiplicity between high transverse energy jets in $\overline{p}$-p collisions at $\sqrt{s} = 1800$ GeV. The data were collected using the DØ Detector at Fermi National Accelerator Laboratory. We observe an excess of events at low multiplicity which is consistent with strongly-interacting color-singlet exchange. The fraction of events due to color-singlet exchange is measured as a function of the transverse energy and rapidity separation of the jets and is compared to several theoretical models for color-singlet exchange.

In memory of my grandparents, Coy Irene and Louis Kelsey Taylor.

Acknowledgments

This dissertation represents my part in the work of many people. The DØ detector had already successfully completed its inaugural run and was ready for its second run when I joined. I must thank all of the physicists and technicians who put many years of work into designing and building DØ . I would also like to thank my advisor, Heidi Schellman, for providing advice and guidance when needed, and leaving me in the capable hands of Andrew and Brent otherwise. Thanks to Dave Buchholz, who was always helpful and answered many questions on the drives from Northwestern to Fermilab.

I was lucky to inherit an analysis in progress from Brent May and Andrew Brandt. I would like to thank Andrew for the guidance and support that led us to get interesting results soon after I joined the group, and that ultimately led to the successful completion of this work. I would like to thank Brent for being able to use his code to get started quickly with this analysis, for his ideas in our many brainstorming sessions, and for his patience in answering questions whenever they came up. I would also like to thank fellow rap-gapper Bob Hirosky for being a general consultant on many issues. In addition, thanks go to Jill Perkins for allowing me to use her data in my thesis and for letting me blow off occasional steam in her office.

There are many people at Northwestern who helped me get through the hurdles of classes and exams. In particular, I would like to acknowledge Todd Pedlar, Mary Durbin, and Chris Thomas. The four of us made it through classes as a collaborative effort. I must also thank Chris (with whom the collaboration became a lot closer!) for putting up with my stress back then and now, for understanding when I moved to Fermilab for three years, and for providing the emotional support I needed to get through these six years. I would also like to thank Tacy Joffe-Minor for all of her advice (which started before I even came to Northwestern) and for her friendship over the years.

Finally, I would like to thank my family, who contributed to this work more than they think. First, I thank my Grandma and Grandpa Giordano for their support when I was in college (summer storage of myself and my belongings) and now. Thanks to my brother, Jeff, for listening when I needed it and for not making too many physics jokes anymore. My parents deserve thanks for bringing me up to be inquisitive and eager to learn, and for their constant encouragement. My mom taught me to devour books at a young age, and I haven't stopped yet. I would like to thank my dad for never tiring of my incessant "Why, Daddy?" questions, and for answering them all along the years. Both of my parents actively encouraged learning and I owe my academic success to them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Point-like constituents within hadrons were discovered thirty years ago [1]. These constituents became known as partons (part of a hadron), and the parton model formalism for calculating hadron interaction cross sections was developed [2]. In a hadron-hadron collision, the parton model allows us to picture the interaction of hard constituents, or partons, from each hadron instead of viewing the collision of the hadrons as a whole. The proton, for example, is made up of a distribution of partons. These partons may be quarks or gluons and each carries a varying fraction of the total proton momentum; this fraction of momentum is referred to as the parton x. Parametrizations of how much momentum a parton carries are given by parton distribution functions which are determined empirically from many sets of data [3, 4].

Quantum Chromodynamics (QCD) is the theory of strong interactions. QCD describes how the partons interact and are bound together to form hadrons. A degree of freedom called color was introduced to satisfy mathematical requirements of the theory. Quarks and gluons each carry color, but only color-singlet states may exist as free particles in nature. QCD is the dynamics of color; all strong interactions are mediated by the exchange of a colored object (a quark or gluon). The formalism of calculating cross sections using QCD and the parton model has successfully described a large body of collider data [5].

As an example of the variables involved, a center of mass diagram for the interaction $\overline{p} + p \rightarrow jet + jet$ is shown in Fig. 1.1. To understand this process, it is instructive to use the Feynman diagram description (Fig. 1.2), which shows only the partons participating in the hard scatter. The Feynman diagram shows the scattering of two partons via a gluon into two final state partons. Because it carries color, a single parton may not be directly observed. Through fragmentation and hadronization, the strong force turns the final state colored parton into a cluster of colorless particles called a jet. The experimental definition of jets is discussed in Chapter 3.

The interaction of Fig. 1.1 may also be viewed as the final state topology in the detector. Fig 1.3 shows a cartoon of an example for a final state of this interaction. The variables used in the detector cartoon are the pseudorapidity, η, and the azimuthal angle about the beam axis, ϕ. Rapidity is related to the fractional momentum along the beam axis. It is given by $y = \tanh^{-1}(p_z/E)$ and is used because distributions as a function of rapidity are invariant under Lorentz boosts. Pseudorapidity is a good approximation to rapidity for finite angles in the limit of zero mass. Pseudorapidity is related to the polar angle θ by $\eta = -\ln \tan \theta/2$.

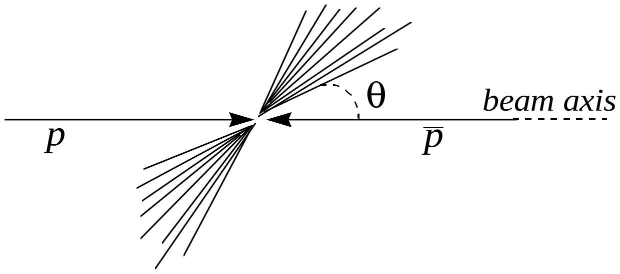

Figure 1.1: $\overline{p} + p \rightarrow jet + jet$ in center-of-mass frame.

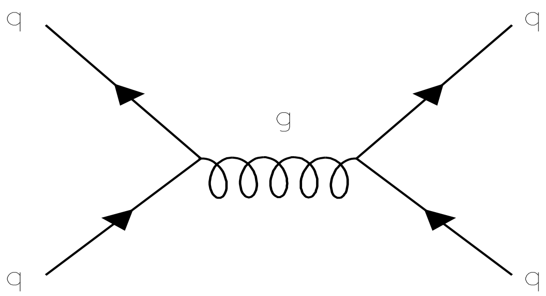

Figure 1.2: Feynman diagram of quark scattering via exchange of a gluon.

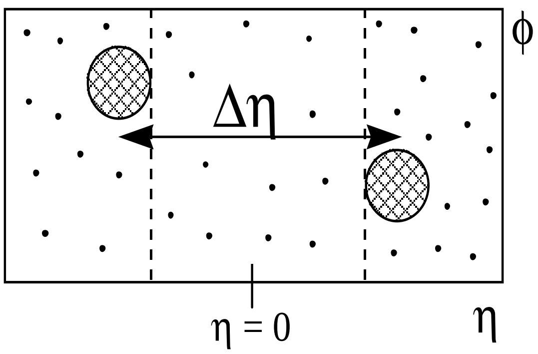

Figure 1.3: Detector cartoon of final state for $\overline{p} + p \rightarrow jet + jet$

The filled circles in the cartoon represent the jets, and the small dots represent particles created through hadronization in the event. In this work, the pseudorapidity separation between the centers of the two highest E_T jets in the event is defined as $\Delta\eta \equiv |\eta_{jet1} - \eta_{jet2}|$.

This work is concerned with a phenomenon not (yet) described by QCD. Recent experimental evidence has arisen for a strongly-interacting color-singlet process, in which the exchanged object carries no net color and has quantum numbers of zero. This type of interaction is outside the scope of current perturbative QCD descriptions. A more complete understanding of QCD or completely new physics may be needed to describe the observed data. The experimental tag for color-singlet exchange is the presence of a rapidity gap. The following sections will outline why rapidity gaps are a signal for color-singlet exchange, what quantities are already measured and what will be measured using them.

1.1 Higgs Production and Rapidity Gaps

Rapidity gaps are regions of (pseudo)rapidity in which there is no particle production. There are several mechanisms which can produce rapidity gaps; those due to strongly-interacting color-singlet exchange are the concern of this analysis. Using rapidity gaps as a signature for Higgs production via W-W fusion was proposed by Bjorken in 1992 [6]. In this scenario, the Higgs decay products would be the only particles in a region between two tagging jets, thus providing a very clear event signature. Figure 1.4 shows the event topology for Higgs production via W-W fusion with a rapidity gap. To use this event topology to tag Higgs production, however,

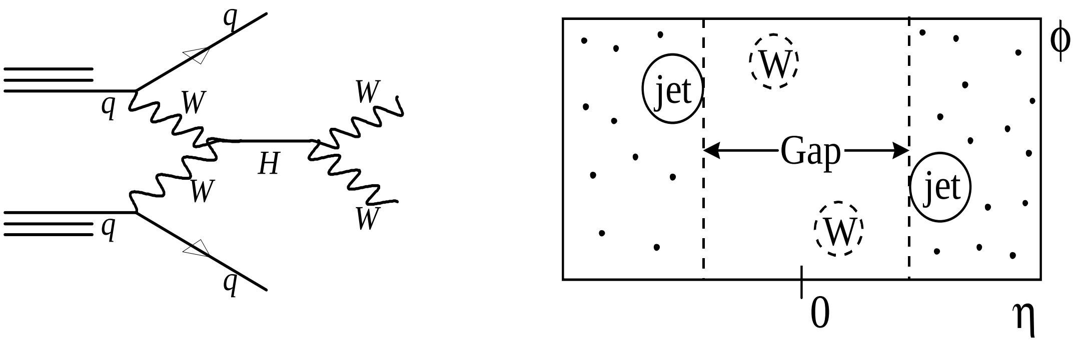

Figure 1.4: Higgs Production with a rapidity gap: The region marked 'Gap' contains only the W-decay products.

other mechanisms for producing dijets with a rapidity gap must be understood. One important contribution is the background from a strongly-interacting color singlet (referred to as the Pomeron).

1.2 Color Flow and Rapidity Gaps

Rapidity gaps arise as a consequence of the color flow in an event. If a colorless object, such as a photon, W, or Z boson, is exchanged in the t-channel (Fig. 1.5), radiation in the phase space between the scattered partons is suppressed. When a colored object is exchanged, the color connections between the parton from the proton and the remnant of the antiproton will lead to hadronization and particles will typically fill the phase space between the scattered partons (Fig. 1.6).

This phenomena has been made clear by Zeppenfeld [7], in a calculation of the rapidity of a radiated gluon in single gluon exchange and in photon exchange. In his calculation, quark jets were generated at $\eta = \pm 3$ and the probability for the

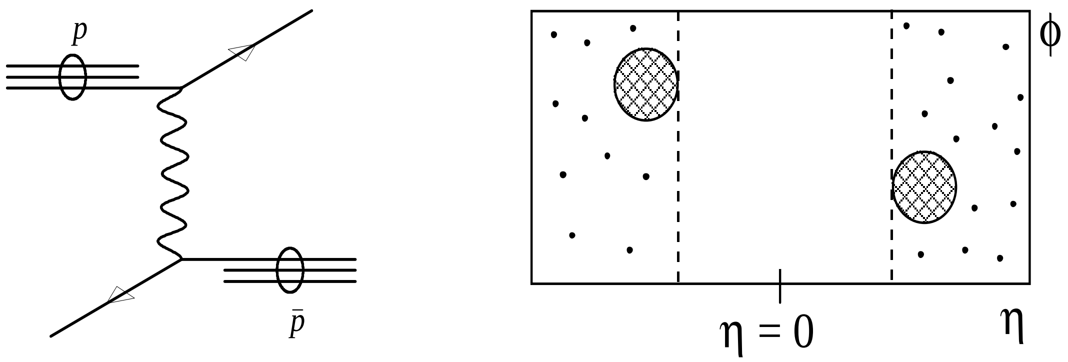

Figure 1.5: Color-singlet exchange event with a rapidity gap.

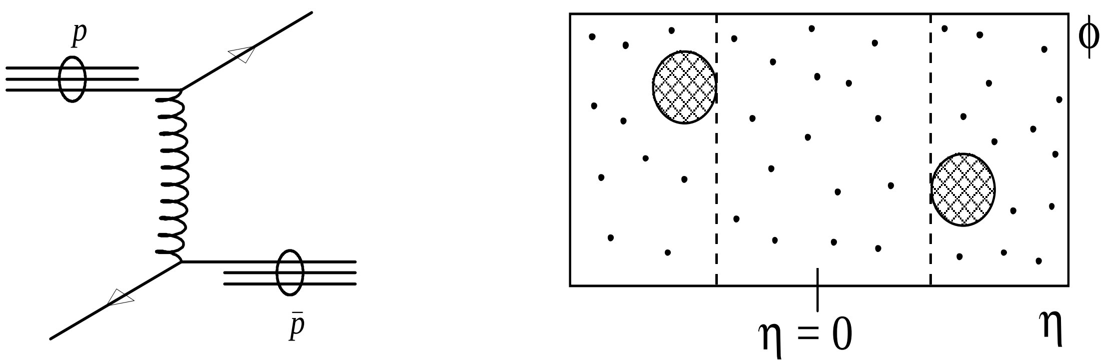

Figure 1.6: Color-exchange event topology.

position of the radiated gluon was determined. Figure 1.7 shows that radiation between the quark jets is highly suppressed in photon (colorless) exchange relative to single gluon exchange. This suppression of radiation in color-singlet exchange events means that such events can be identified by the signature of a rapidity gap between widely separated jets.

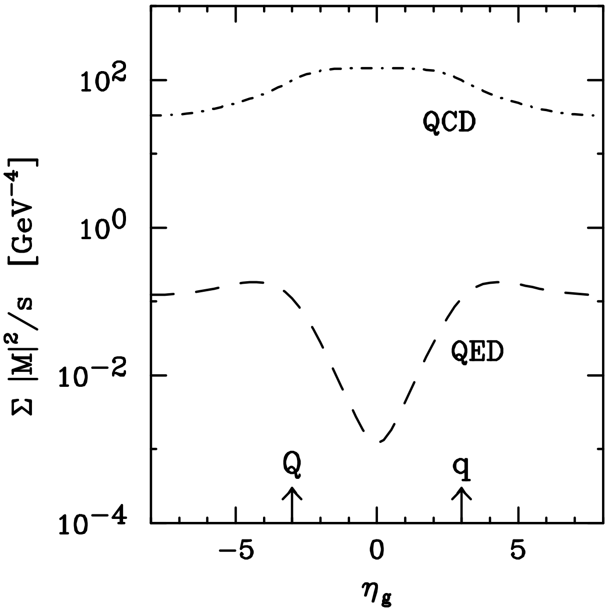

Figure 1.7: Photon exchange (QED) is compared to gluon exchange (QCD). The rapidity distribution of the emitted gluon is shown for quark-quark (q-Q) scattering with quark jets at the rapidities indicated by the arrows. (Figure courtesy of D. Zeppenfeld)

1.2.1 Survival Probability

Not all color-singlet exchange events, however, will result in a rapidity gap. Soft interactions among the spectator partons in the proton and antiproton may fill the phase space with particles (note that this is not the case for the Soft Color model to be discussed in Section 1.5.2). The probability that the soft interactions do not fill the region between the jets with radiation is called the survival probability.

The survival probability is calculated by considering the overlap of the proton and anti-proton in phase space. Given parton densities $\rho(r)$ in the transverse plane, the probability for a hard interaction is given by

$$\sigma_{hard} \backsimeq \sigma_0 \int d^2B d^2r \rho(r)\rho(B-r) \equiv \sigma_0 \int d^2B F(B) \tag{1.1}$$

where B is the impact parameter between the proton and anti-proton and r is the distance from the center of the hadron to the interaction plane. The survival probability is then just the probability that the proton and anti-proton will pass through each other without any other interaction:

$$< |S|^2 >= \frac{\int d^2BF(B)|S(B)|^2}{d^2BF(B)}. \tag{1.2}$$

One common estimate of $< |S|^2 >$ is of the form

$$|S(B)|^2 = e^{-\nu\chi(B)} \tag{1.3}$$

where $\chi(B)$ is a convolution of parton densities and ν is a constant. Several different models [6, 8, 9] have been used to calculate $< |S|^2 >$. These models are constrained by measurements of the total and inelastic cross sections and differ in their assumptions about the effective radius of the parton densities and the parameter ν. The range in resulting survival probabilities is between 8 and 30%. The observable rapidity gap fraction is then

$$f_{obs} = S \cdot \frac{\sigma_{gap}}{\sigma_{dijet}} \tag{1.4}$$

where S is the survival probability, σ_{gap} is the cross section for producing a rapidity gap, and σ_{dijet} is the total color-exchange plus color-singlet cross section.

1.2.2 QCD and Rapidity Gaps

After Bjorken introduced the idea of rapidity gaps as a signal for Higgs production, a large amount of theoretical interest was generated in what he initially considered a background process: rapidity gaps due to strongly-interacting color-singlet exchange. Rapidity gaps in soft diffraction had been observed and studied by many groups including UA1 and UA8. These diffractive processes are characterized by a very small momentum transfer ($\lesssim 1$ GeV^2), so that the proton stays intact after the interaction. The mechanism for producing rapidity gaps in diffractive processes is historically referred to as the Pomeron.

Bjorken proposed rapidity gaps which would be produced by a very high momentum-transfer process (about 1 TeV^2 at the Tevatron). Rapidity gaps of this type were predicted to be due to strongly-interacting color-singlet exchange, but had not yet been observed. These processes are not calculable in the current QCD formalism and were not present in QCD Monte Carlo generators. Therefore, theoretical interest arose in trying to predict the cross section and rate for the production of rapidity gaps, but experimental evidence was needed to confirm the validity and provide input to the predictions.

1.2.3 Other Sources of Rapidity Gaps

There are several other mechanisms which can produce rapidity gaps between jets. The expected fraction of rapidity gaps $f = \sigma_{gap}/\sigma_{dijet}$ for each of these sources is given below.

- *Electroweak Color-Singlet*: The t-channel exchange of a photon, W, or Z boson can produce rapidity gaps with a rate given by [10]:

$$\frac{\sigma_{gap}(EW)}{\sigma_{dijet}} \sim 7 \times 10^{-4}.$$

- *Color-exchange fluctuations*: If the jet separation is small, a color-exchange event may have a rapidity gap because of a fluctuation in the particle multiplicity. The probability of forming a rapidity gap due to fluctuations from color-exchange decreases approximately exponentially with jet separation [6]:

$$\frac{\sigma_{gap}(color)}{\sigma_{dijet}} < 10^{-4} \quad (\Delta\eta > 4).$$

- *Other quark scattering*: Rapidity gaps can be produced in a color-exchange event by quarks which scatter at large angles ($\theta >> 90°$) in the t-channel or at small angles in the u-channel. Both of these processes are highly suppressed relative to t-channel exchange for large jet separation [11]:

$$\frac{\sigma_{gap}}{\sigma_{dijet}} \lesssim 10^{-5} \quad (\Delta\eta \sim 4).$$

1.3 Summary of Published Results

Bjorken's paper and ideas on the two gluon model for the Pomeron led DØ to begin a program of rapidity gap physics. Specially designed forward dijet triggers implemented in the 1992-93 Tevatron run made the measurements possible by providing a large sample of dijets with a large rapidity separation. A measurement of the experimental rapidity gap fraction; defined as the fraction of events with zero calorimeter towers with energy greater than 200 MeV, was published by DØ in 1994 [11, 12]. An upper limit of 1.1% was placed on the fraction of events with no particles between dijets with rapidity separation $\Delta\eta > 4.4$.

Subsequently, both DØ and CDF used fits to the multiplicity distribution of particles between the dijets to subtract the color-exchange background and extract the fraction of color-singlet exchange events. CDF measured a fraction of $(0.85 \pm 0.12(\text{stat})^{+0.24}_{-0.12}(\text{syst}))\%$ [13]. DØ measured a fraction of $(1.07 \pm 0.10(\text{stat})^{+0.25}_{-0.13}(\text{syst}))\%$ [14]. In addition, DØ used the measured color-singlet fraction to exclude pure electroweak exchange plus color-exchange background as the source of the excess. Using a conservative estimate, the probability was determined to be less than 10^{-10} that the observed color-singlet fraction was due to electroweak exchange plus color-exchange background. These measurements were evidence for the presence of a strongly-interacting color singlet.

The measurements in this analysis are a probe of the nature of the strongly-interacting color singlet. Measurements of the color-singlet fraction as a function of jet E_T, of the pseudorapidity separation of the jets ($\Delta\eta$), and of the center of mass energy will give information on the process involved in the exchange. These measurements are compared to several models for color-singlet exchange to determine which models are able to describe the data. This thesis will describe the measurement of the color-singlet fraction as a function of E_T and $\Delta\eta$. The thesis of Jill Perkins [15] will describe the measurement of the fraction at two different center of mass energies as well as provide detail on the lowest E_T data in the measurement of the color-singlet fraction versus E_T.

1.4 Multiplicity Distributions

The particle multiplicity distribution can be used to measure the fraction of color-singlet exchange. The multiplicity produced between the dijets in a color-singlet exchange event is naively expected to be equal to zero. However, there are many effects which can place a small number of particles between the jets. Detector effects, such as noise, can lead to the detection of spurious particles. Backscattering of particles off of other parts of the detector may also place a small number of particles between the jets. Physics effects may also contribute to a non-zero multiplicity between the jets in a color-singlet exchange event. Jet fragmentation may result in particles outside of the defined radius of the jet. Spectator interactions, or interactions among the soft partons in the event, are expected to produce a minimum bias-like multiplicity distribution[1]. Some events with spectator interactions may, therefore, have only a few particles between the jets and be classified as color-singlet exchange events.

The multiplicity distribution for color-exchange events is known to be a smooth distribution with a non-zero mean due to radiation produced by the color connections in the event. This distribution has been shown to have the shape of a Negative Binomial or similar distribution. The mean multiplicity rises as the jet separation

[1] Minimum bias refers to a trigger on inelastic interactions. Such events mainly result from interactions among soft partons which will typically not produce hard physics objects like jets. The particle multiplicity in a minimum bias interaction is peaked at a few particles.

is increased. The mean multiplicity is also a function of detector efficiency; if the detector efficiency is low, then the mean multiplicity will also be low.

The color-singlet exchange component can be observed as a low multiplicity excess on top of the color-exchange background in the particle multiplicity distribution. To achieve maximum separation of the signal and background, the detector efficiency needs to be high. This requirement corresponds to a low energy threshold for particle detection. In addition, the jets need to be widely separated in rapidity to push the mean of the color-exchange multiplicity distribution away from the low multiplicity signal region.

1.4.1 Negative Binomial Distribution Parametrization

Negative Binomial Distributions (NBD's) have been successfully used to describe experimental charged particle multiplicity distributions in experiments ranging from fixed target to p-p data at the ISR to $\bar{p}$-p data at UA5 [16]. These experiments correspond to center-of-mass energies from 11–540 GeV. A NBD fits the charged multiplicity distribution well for inelastic, non-single diffractive events in every case. There is some theoretical motivation for the success of these fits given by Giovannini and Van Hove [17]. They propose that the NBD-like particle multiplicity distribution arises from cascade processes which occur in connection with jet fragmentation.

Charged particle multiplicity distributions from higher center-of-mass energies are not well fit by a single NBD due to the presence of a high multiplicity shoulder [18, 19, 20]. They are, however, well fit by a superposition of two or more NBD's. For e^+-e^- collisions, this is interpreted as being due to a 'soft' and 'hard' component having different underlying multiplicity distributions. For $\overline{p}$-p collisions at $\sqrt{s}=$ 1800 GeV, a high multiplicity shoulder is interpreted as being due to additional minijet activity in the minimum bias spectrum. Particle multiplicity distributions from a wide range of experiments and energies have been successfully fit by either a single NBD or a superposition of two or more NBD's [18].

1.5 Color Singlet Models

Several models exist to explain strongly-interacting color-singlet exchange. Perturbative QCD motivated models are the two gluon model first proposed by Bjorken and an extension which adds dynamics using BFKL formalism [21]. Soft color rearrangement is a non-perturbative QCD model which has been extended from descriptions of charmonium production to consider rapidity gaps [22]. A model containing completely new physics, the presence of a new gauge boson, also proposes to explain the observed rapidity gap data [23].

Each of these models will be discussed in more detail in the sections which follow. All of the current color-singlet models are at leading order ($2 \rightarrow 2$ processes), meaning there can only be two jets in the final state. The data, of course, is equivalent to all orders in perturbation theory and frequently has more than two jets in the final state. The restriction of two jets in the final state in the theory may have a large effect on the predicted gap fraction (see Section 7.1.4). One common factor of all the models is that the resulting gap fraction depends on the initial parton distributions through various factors. These dependences as well as the dynamics present in some of the models govern their predictions for the behavior of the color-singlet fraction.

1.5.1 Two Gluon Model

The two gluon model depicts the Pomeron as two perturbative gluons in a color-singlet state [6]. In the simple two gluon model, the gap fraction is predicted to be $\sigma_{gap}/\sigma_{dijet} \sim 0.1 \cdot S$ [6].

The total dijet cross section can be written as:

$$\frac{d\sigma_{dijet}}{dx_A dx_B} = \prod_{i=A,B} \left[G(x_i, Q^2) + \frac{4}{9} Q(x_i, Q^2) \right] \hat{\sigma}_{dijet} \tag{1.5}$$

where $\hat{\sigma}_{dijet}$ is the hard scattering cross section, and $G(x, Q^2)$ and $Q(x, Q^2)$ are the distribution of gluons and quarks in the proton. The $(G + \frac{4}{9}Q)$ term is referred to as an effective parton distribution function. Combridge and Maxwell showed [24] that, to a good approximation, a single effective parton-parton interaction can describe jet production. This allowed jet production cross sections to be expressed in the simple form given in Equation 1.5. The cross section for two gluon color-singlet exchange is given by

$$\frac{d\sigma_{singlet}}{dx_A dx_B} = \prod_{i=A,B} \left[G(x_i, Q^2) + \left(\frac{4}{9}\right)^2 Q(x_i, Q^2) \right] \hat{\sigma}_{singlet}. \tag{1.6}$$

Therefore, the ratio of the singlet to total dijet cross section is

$$\frac{\sigma_{singlet}}{\sigma_{dijet}} = w(x_A) w(x_B) \frac{\hat{\sigma}_{singlet}}{\hat{\sigma}_{dijet}} \tag{1.7}$$

where $w(x)$ represents the relative weighting of singlet to octet and is given by:

$$w(x) = \frac{G(x,Q^2) + (\frac{4}{9})^2 \sum_f Q_f(x,Q^2)}{G(x,Q^2) + \frac{4}{9}\sum_f Q_f(x,Q^2)}. \tag{1.8}$$

The relative suppression of two gluon color-singlet exchange to color-octet exchange in this model is then w^2. For g-g scattering, $w^2 = 1$, while for q-q scattering, $w^2 = (4/9)^2 \cong 0.2$. Thus the gap fraction due to initial q-q scattering is expected to be around 20% of that from pure g-g scattering. Since the two gluon singlet couples more strongly to gluons, the gap fraction will fall as a parton x increases. The parton x increases as the jet E_T and separation ($\Delta\eta$) are increased; thus the simple two gluon model predicts that the gap fraction falls as both the jet E_T and $\Delta\eta$ rise.

BFKL Pomeron

Del Duca and Tang extended the simple two gluon model by considering the BFKL dynamics of ladders of exchanged gluons [21]. The dynamics in the BFKL calculation add to the effects from the changing parton distribution functions. In this model, the gap fraction as a function of E_T falls more rapidly than that from the simple two gluon model. The fraction rises at large $\Delta\eta$ but the actual value of the $\Delta\eta$ at which it rises is uncertain in the formalism.

1.5.2 Soft Color Model

The soft color model [22] proposes that rapidity gaps are formed when the color in a single gluon exchange is canceled by non-perturbative quarks and gluons at large times and distances from the hard interaction. This model was introduced to explain charmonium production and asserts that the color-singlet charmonium state is not formed at the short distance perturbative scale, but instead depends on large distance interactions of quarks and gluons long after the perturbative interaction. It has been successful in describing certain aspects of charmonium production.

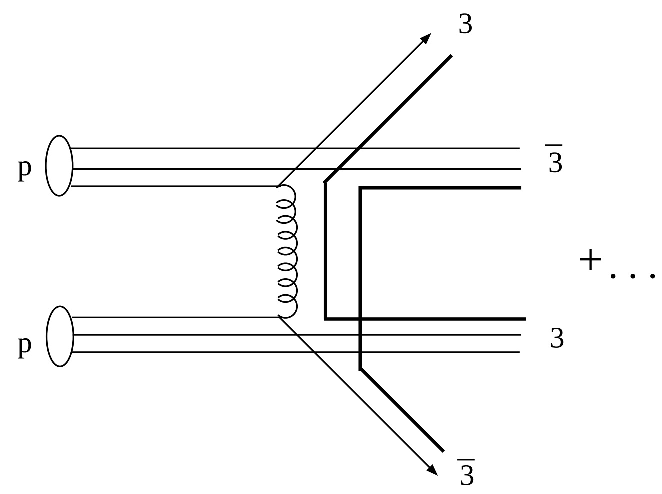

Figure 1.8: Soft color picture for the formation of a rapidity gap for quark-quark scattering in hadron collisions. Rearrangement of color results in an effective color-singlet exchange in the final state after hadronization. (Figure courtesy of J. Perkins)

This concept of color as a non-perturbative phenomenon has been extended to the production of rapidity gaps. Figure 1.8 shows an example of how a rapidity gap can be formed through the soft color mechanism. The probability to form a rapidity gap from q-q scattering is higher than that for g-g scattering because quarks have less possible color combinations, so it is statistically more likely that random quark and gluon interactions can cancel that color. The naive probability to form a gap from q-q scattering is $1/(1+8)^2 \cong 1\%$. The probability to form a gap from g-g scattering is not as clear, but is less than or equal to the rate for q-q scattering. It should be noted that the survival probability has no meaning in this model; in fact, the soft interactions among the spectator partons in the event actually help form the rapidity gap. Thus, the soft color model predicts, in contrast to the two gluon models, that the gap fraction will rise (or stay constant) with increasing parton x.

1.5.3 U(1) Gauge Boson

An alternative explanation of the rapidity gap data is the exchange of a presently unobserved, strongly-interacting U(1) gauge boson [23]. This model proposes to gauge the U(1) symmetry generated by baryon number. When this symmetry is broken, the gauge boson γ_B obtains a mass $m_B < m_Z$. Allowed regions for the γ_B mass (m_B) and coupling (α_B) are mapped out by considering existing data for the hadronic width of the Z boson, the fraction of Z events decaying to jets, the dijet invariant mass peak in $Z \to 4$ jet events, and the Υ(1S) decay. The constraints introduced by the data leave some space open for the γ_B at roughly $m_B \gtrsim 20$ GeV and $\alpha_B \lesssim 0.2$.

The γ_B couples only to baryon number (quarks), so the gap fraction due to γ_B exchange would rise with increasing parton x as more quark initiated processes arise. The fraction of events due to γ_B exchange compared to single gluon exchange is:

$$\frac{\alpha_B^2}{18\alpha_s^2}\left(1+\frac{m_B^2}{E_T^2}\right)^{-2}. \tag{1.9}$$

Using the standard estimates for the survival probability of a perturbative singlet of 0.1–0.3, the rapidity gap fraction due to γ_B exchange would be

$$f(\Delta\eta > 5.4) \sim (0.1-0.3)\frac{4\mathrm{x}10^{-2}}{\left(1+\frac{m_B^2}{E_T^2}\right)^2}\left(\frac{\alpha_B}{0.1}\right)^2. \tag{1.10}$$

Suitable choices of m_B and α_B could give a gap fraction of about one percent. The dynamics of Equation 1.10 lead to a faster rise in the gap fraction as a function of E_T than that from parton distribution functions alone. Figure 1.9 shows the E_T dependence for different values of m_B assuming a survival probability of 10% and a coupling $\alpha_B = 0.2$ which is independent of the E_T scale in the relevant range.

1.6 Measurement of $f(E_T)$ and $f(\Delta\eta)$

Measuring the color-singlet fraction as a function of the E_T and $\Delta\eta$ of the jets can help distinguish between the different color singlet models. The predicted color-singlet fraction depends on the initial parton distribution functions in all models. Events with jets at high E_T and large rapidity separation are expected to have a larger fraction of initial q-q scattering than events with jets of lower E_T or $\Delta\eta$.

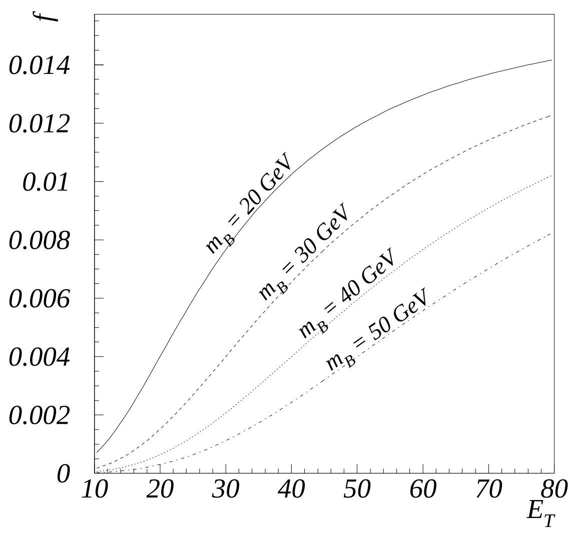

Figure 1.9: Color-Singlet fraction versus E_T in the U(1) Gauge Boson model (independent of parton distribution function effects).

The leading order expression for the parton x is:

$$x = \frac{E_T}{\sqrt{s}} e^{|\eta|} \tag{1.11}$$

where E_T and η are the average E_T and η of the two leading jets. The relative amount of initial quarks and gluons is calculated using CTEQ4M [3] parton distribution functions from the average value of x and E_T for each E_T and $\Delta\eta$ bin in the measurement. The relative amount of initial quarks changes from 60% to 86% from the lowest to the highest E_T bin and from 68% to 86% from the lowest to the highest $\Delta\eta$ bin (see Chapter 7 for more details). Although q-q scattering dominates in all cases, the fraction of initial q-q scattering increases as E_T and $\Delta\eta$ increase. An additional measurement of the fraction at two different center of mass energies is also a probe of different parton x regions. When the jet E_T and η are held constant (at a lower E_T than the previous measurement), the fraction of initial quarks increases from 52% at $\sqrt{s} = 1800$ GeV to 70% at $\sqrt{s} = 630$ GeV. For more details on this measurement, see Refs. [15, 25]. These three measurements together will be compared with the current color-singlet models to determine which models can describe the data.

1.7 Outline of thesis

Chapter 2 gives a brief overview of the DØ Detector and describes features which are relevant to rapidity gap measurements. Chapter 3 discusses the experimental identification of jets. In Chapter 4, the data collection and the cuts used to define the final data samples are described. The observation of color-singlet exchange and a study of the features of color-singlet events are discussed in Chapter 5. Chapter 6 gives the methods used in this analysis including details on the definition of the color-singlet fraction, the fitting method used to define it, and the corrections applied. Chapter 7 gives results for the measurement of the color-singlet fraction and outlines the estimates for each source of systematic error. It also presents results for the measurement of the color-singlet fraction as a function of jet transverse energy and rapidity separation. Finally, Chapter 9 concludes with an interpretation of the data.

Chapter 2

The DØ Detector

The DØ detector is one of two multipurpose collider detectors at the Tevatron at Fermilab. DØ was designed to provide excellent identification of electrons, muons, jets, and missing transverse energy. Although the main focus of the DØ detector is the study of high p_T phenomena, the superior calorimeter makes DØ an ideal place to study jets and QCD.

The DØ detector (Fig. 2.1) consists of three main components: the central tracking system, the calorimeter, and the muon system. The innermost part of the detector is the central tracking system. In order from the beam pipe, the central tracker consists of the vertex drift chamber, the transition radiation detector, and the central drift chamber. Two forward drift chambers sit perpendicular to the beam in the forward direction. The uranium-liquid argon calorimeter is in three sections: the central calorimeter in the region $|\eta| < 1$ and two end calorimeters covering up to $|\eta| = 4$. Massless gaps and intercryostat detectors provide coverage for the region $0.8 \leq |\eta| \leq 1.4$. The muon system measures muon tracks down to

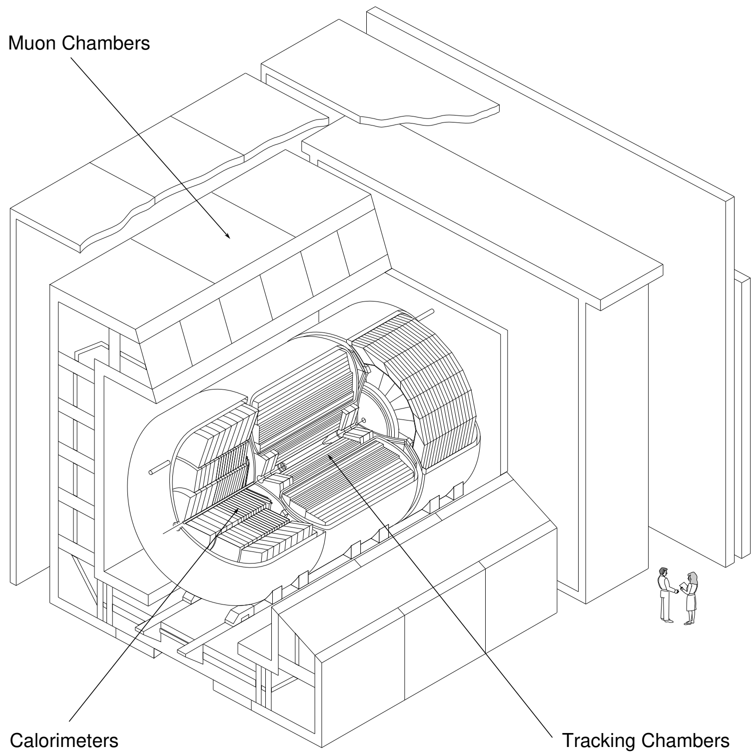

Figure 2.1: An isometric view of the DØ detector.

angles of about three degrees with proportional drift tube chambers interspersed with five solid iron toroidal magnets. Detailed descriptions of the entire detector and its components are given in Ref. [26]. The following sections will focus on detector elements used for this analysis.

2.1 Detector Coordinates

The DØ coordinate system is defined so that the z-axis is along the proton direction, the y-axis points vertically up, and the x-axis points out toward the center of the accelerator. A polar coordinate system can also be used in which r is the perpendicular distance from the beam axis, and θ and ϕ are the polar angles defined with respect to the $+z$ axis. A commonly used coordinate is the pseudorapidity $\eta \equiv -\ln\tan\theta/2$. The pseudorapidity is a good approximation for the true rapidity $y = 1/2\ln{(E+P_z)/(E-p_z)}$ for finite angles in the limit of zero mass. Pseudorapidity is a good approximation for the rapidity of jets and will be referred to as rapidity in this analysis.

2.2 Level Ø

The Level Ø detector is the first level of the trigger system. It is primarily used to indicate the presence of inelastic collisions and to measure the luminosity for the experiment. Level Ø consists of two hodoscopes of scintillation counters mounted in front of each end calorimeter. It has partial coverage for $1.9 \leq |\eta| \leq 4.3$ and full coverage for $2.3 \leq |\eta| \leq 3.9$. The Level Ø hodoscopes provide, at the trigger level, the indication of an inelastic collision, a measurement of the z vertex, and an indication of whether multiple proton-antiproton interactions occurred in a single crossing.

The Level Ø determination of whether a multiple interaction occurred is referred to as the multiple interaction flag. This quantity is determined from the difference in the width of the timing distributions between the North and South Level Ø hodoscopes. A large time difference is indicative of more than one $\overline{p}$-p interaction. Cuts are made on this timing variable to indicate the probability of a single or multiple interaction. The distribution of this timing variable along with

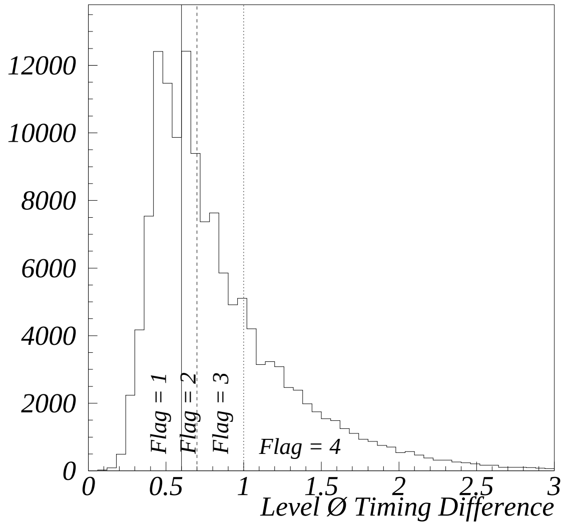

Figure 2.2: Difference in width of timing distribution between the North and South Level Ø . Lines indicate cuts for the multiple interaction flag.

the cuts used to define the multiple interaction flag is shown in Fig. 2.2. The possible values of the multiple interaction flag are:

- A flag value of 0 means there was no inelastic interaction.
- A flag value of one means 'most likely' a single interaction,
- two means 'likely' a single interaction,
- three means 'likely' a multiple interaction and
- four means 'most likely' a multiple interaction.

2.3 Calorimeter

The DØ calorimeter is a sampling calorimeter with uranium, stainless steel, or copper as the absorber and liquid argon as the sampling medium. The calorimeter is contained in two cryostats (Fig. 2.3); the central calorimeter covers the region $|\eta| < 1$ and the end calorimeters cover out to $\eta = 4$. The calorimeter towers are finely segmented for a typical size of 0.1×0.1 in η-ϕ space. The tower centers are projective; that is, the center of each tower lies on a ray projecting from the center ($z = 0$) of the detector. Figure 2.4 shows the η coverage, the segmentation, and the projectivity of the calorimeter.

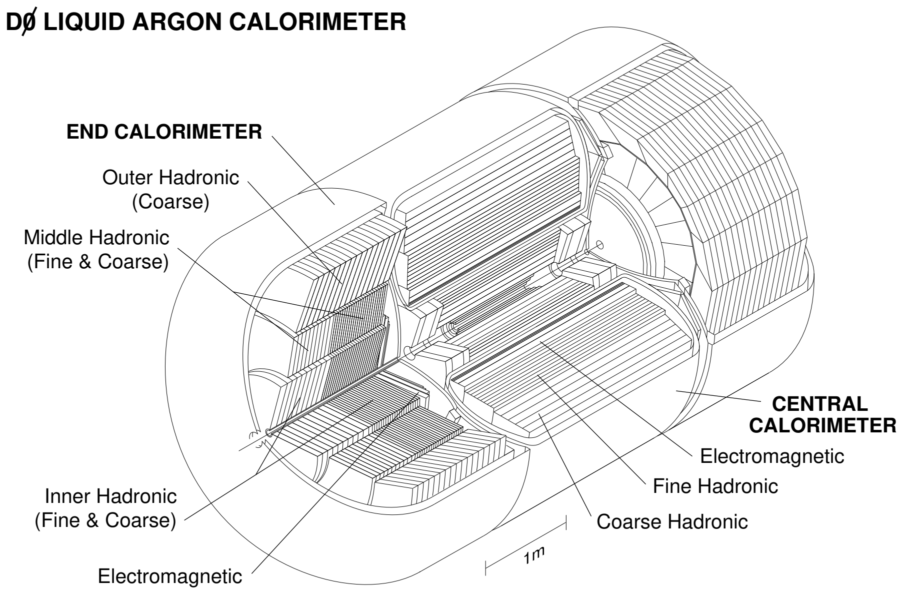

Figure 2.3: An isometric view of the DØ calorimeters and the central tracking.

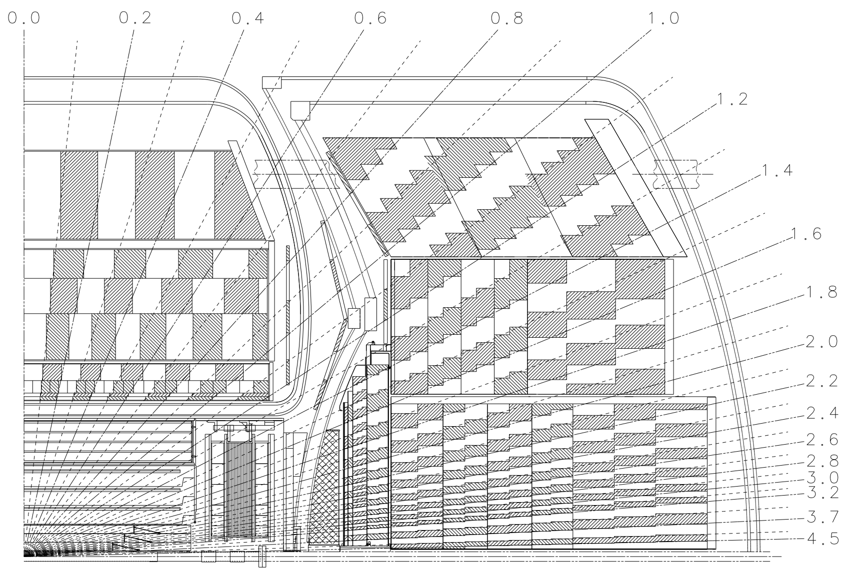

Figure 2.4: View of a portion of the calorimeters showing the transverse and longitudinal segmentation. Each rectangle is a separate cell for readout. The numbered lines are indications of the pseudorapidity η.

2.3.1 Central Calorimeter

The central calorimeter consists of three concentric cylinders: the Electromagnetic (EM), the Fine Hadronic (FH), and the Coarse Hadronic (CH) calorimeters. The signals from the EM calorimeter are segmented into four longitudinal sections for the readout at radiation lengths of 2.0, 2.0, 6.8, and 9.8. The Fine Hadronic calorimeter has three longitudinal segments while the Coarse Hadronic calorimeter consists of one segment.

2.3.2 End Calorimeters

The end calorimeters extend the calorimetry to the forward region. The end calorimeters consist of EM and Hadronic calorimeters. The EM calorimeter readout is again segmented longitudinally at 0.3, 2.6, 7.9, and 9.3 X_0. The Inner Hadronic (IH) calorimeter is positioned directly behind the EM section and consists of four readout layers of fine hadronic and one layer of coarse hadronic modules. The Middle Hadronic (MH) calorimeter forms a concentric ring around the IH calorimeter and also has four readout layers of fine hadronic and one layer of coarse hadronic modules. Finally, the Outer Hadronic (OH) calorimeter surrounds the MH calorimeter and has one layer of coarse hadronic readout.

2.3.3 Intercryostat Detectors

The region between the cryostats for the central and end calorimeters contains a large amount of dead material. To provide coverage in this area and to correct for the energy deposited in uninstrumented regions, scintillation counter arrays called Intercryostat Detectors (ICD) were mounted on the front surface of the end calorimeters. An additional source of instrumentation in the Intercryostat region is supplied by the Massless Gaps. These are signal boards with no absorber surrounded by liquid argon gaps and are mounted inside the cryostats on the surfaces of the central Fine Hadronic, end Middle Hadronic, and end Outer Hadronic modules.

2.3.4 Pileup in the calorimeter

After preamplification, output signals from the calorimeter are input to a baseline subtractor (BLS) shaping and sampling circuit. The signals are sampled just before a beam crossing and 2.2 μs after (there are 3.5 μs between beam crossings). The difference is a dc voltage proportional to the charge collected. Baseline restoration (or the time to discharge the capacitors in the BLS) takes a few μs. In a high luminosity environment, there is an increased probability that the previous crossing had an inelastic interaction. If, in the previous crossing, a readout cell was populated by an inelastic event and the baseline does not restore to zero charge, then the baseline reading taken before the current crossing will be high. If the cell in question does not have any energy deposited in the current crossing, then the baseline keeps falling, so the charge measured is negative. This phenomena is called 'pileup', or 'negative energy'. The amount of negative energy measured is luminosity dependent.

2.3.5 Noise and Low Energy Particle Detection

Noise in the calorimeter comes from two sources: uranium noise and electronics noise. The uranium noise is due to beta decay in the uranium absorber plates and produces a non-Gaussian distribution with a long tail. Noise from the electronics gives a symmetric Gaussian distribution. In addition to single channel noise, there can also be multichannel coherent noise. The number of channels that can be summed before the coherent noise exceeds the random noise is greater than 5000.

The 'pedestal', or noise distribution in each channel is measured approximately once per day by measuring the number of ADC counts when there is no beam in the Tevatron. A Gaussian is fit to the noise distribution from each channel to determine the mean and width of the noise pedestal. These values are used during data taking to suppress a channel if it is within 2σ of the pedestal value. This suppression is called zero suppression and reduces the number of channels read out by about a factor of ten.

Although noise and zero suppression do not have much effect on the measurement of high p_T objects such as electrons and photons, they can have an effect on low energy particle detection. The amount of electronic noise in a channel is proportional to the capacitance of the readout cell, which is proportional to the cell size. The cell size is smallest closest to the beam pipe because the towers are a constant size in η and ϕ. Therefore, the Electromagnetic calorimeter cells are smaller than the Fine Hadronic calorimeter cells, which are smaller than the quite large Coarse Hadronic calorimeter cells. The noise in each of these calorimeter sections is given in Table 2.1 [27]. These values show that the EM calorimeter

Section	Pedestal Width (MeV)
EM	5-20
FH	10-70
CH	25-100

Table 2.1: Noise levels in different calorimeter sections

is best suited for low energy particle detection. The noisier environment of the Hadronic calorimeter could lead to the misidentification of noise fluctuations as low energy particles. In addition, the energy threshold for particle detection in the Hadronic calorimeter is higher than that for the EM calorimeter because a particle must traverse more material before reaching the Hadronic calorimeter.

2.4 Central Tracking

The elements of the tracking system used in this analysis are the Central Drift Chamber (CDC) and the Forward Drift Chambers (FDC). The entire tracking system is shown in Fig. 2.5.

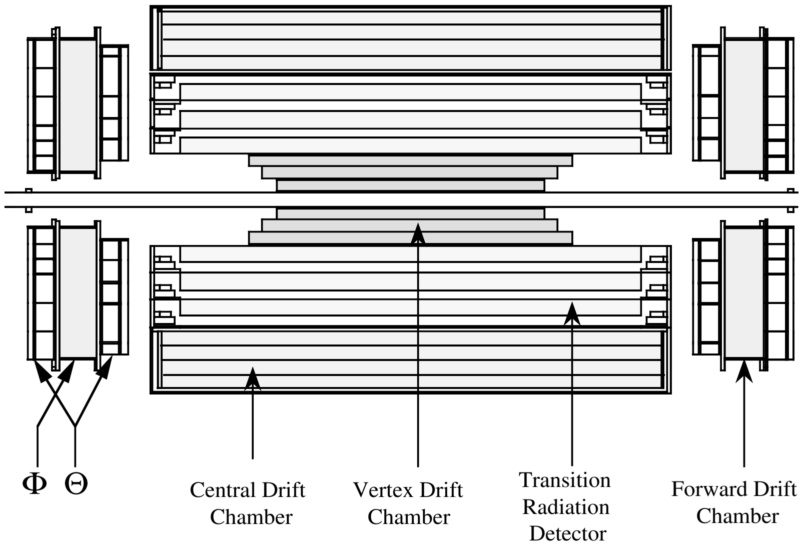

Figure 2.5: The DØ tracking system.

2.4.1 Central Drift Chamber

The CDC consists of four concentric rings of 32 azimuthal cells. Thirty sense wires read out at one end of the chamber provide the r-ϕ position of the hits. Delay lines read out at both ends of the chamber give the z position of the hits. The resolution for the r-ϕ hit position is about 150-200μm and the z resolution is about 2 mm. These resolutions are degraded in a high multiplicity environment.

2.4.2 Forward Drift Chambers

The Forward Drift Chambers provide tracking in the range $1.5 \leq |\eta| \leq 3$. Each FDC consists of three sets of detectors: a Φ module between two Θ modules. The Φ module has radial sense wires and measures the ϕ coordinate. The Θ modules have sense wires and delay lines to measure both the θ and z position of the hits. The r-ϕ resolution for the FDC is about 200μm and the z resolution is 4 mm.

2.4.3 Multiple Scattering

One concern for this analysis is low energy particle detection. In principle, because there is no central magnetic field to bend low energy particles away, the central tracking system can detect particles with arbitrarily low energy. However, particles will lose energy while traveling through the material preceding the CDC and may, through multiple scattering, be lost from the detector. The expression for the scattering angle for multiple scattering is:

$$\phi_{rms} \cong \frac{z}{pv}(21\mathrm{MeV})\sqrt{\frac{\mathrm{X}}{\mathrm{X_0}}} \tag{2.1}$$

where z is the particle's charge, pv is the momentum in MeV, and X/X_0 is the thickness of the scattering medium in radiation lengths. To obtain a rough estimate of the effects of multiple scattering, we use a crude model in which the thickness of the material before the CDC is $X/X_0 = 0.1$ at $\theta = 90$ degrees and $X/X_0 = 1$ elsewhere. Table 2.2 shows the scattering angle for various energy particles. Although the material may not stop very low energy charged particles, they may multiple scatter so far as to not be reconstructed as a track in the tracking

Momentum (MeV)	ϕ_{rms} $X/X_0 = 0.1$	ϕ_{rms} $X/X_0 = 1$
10	38 °	120 °
100	3.8 °	12 °
500	0.8 °	2.4 °

Table 2.2: Multiple Scattering Angles for Low Energy Particles

system. Between the loss from multiple scattering and reconstruction efficiencies, it is estimated that the energy threshold for detecting charged particles in the CDC is about 100–200 MeV [28].

2.5 Triggering and Data Acquisition

Interesting events must be filtered out of the large collision rate of $\overline{p}$-p interactions by a triggering system. The data acquisition system was capable of recording about 2–3 events out of about 10^5 $\overline{p}$-p interactions per second. This was accomplished by using four levels of triggering. There were three hardware levels: Level Ø , Level 1, and Level 1.5 and one software level: Level 2. The Level Ø system was discussed in Section 2.2 as a detector element. This analysis did not use any Level 1.5 triggers.

2.5.1 Level 1

The Level 1 trigger reduced the data rate to about 200 Hz. Digitized information from the calorimeter and muon system were available for the Level 1 trigger decision. Only calorimeter based triggers are used in this analysis.

The calorimeter trigger required a specified amount of energy detected in a group of calorimeter towers. There were two ways to group towers to sum the energy: trigger towers and large tiles. Trigger towers consisted of four calorimeter towers grouped together to form a trigger unit of 0.2×0.2 in η-ϕ space. Large tiles grouped 128 towers together in overlapping tiles of dimension 0.8×1.6 in η-ϕ space. The large tile triggers were implemented to provide a faster turn-on for jet triggers.

2.5.2 Level 2

The Level 2 trigger reduced the data rate to the 2–3 Hz which could be written to tape. Software filters were run on a farm of 48 microvaxes to perform the trigger decision. Jet and track finding algorithms were available at this level. The triggers for this analysis used a Level 2 filter tool which required forward dijets separated by a specified amount in rapidity.

Chapter 3

Jet Reconstruction

A jet is a cluster of energy which originates from a hard partonic interaction when hadronization and fragmentation turn a final state parton into a collimated stream of particles. Jets are typically identified at hadron colliders by using an iterative cone algorithm. At DØ, jets are reconstructed using only the calorimeter. The reconstructed jets are corrected for detector effects back to the 'true' jet energy.

3.1 Event Vertex

The transverse energy, E_T, and the pseudorapidity, η, of a jet depend on the location in z of the event vertex. Information from the central tracking system (CDC and FDC) is used to find the z vertex for the event. In each subsystem, tracks are reconstructed and their position is extrapolated back to the beam axis. The z position of each track is input to a distribution of clusters of z vertex positions. A Gaussian fit is performed on each cluster of z positions to give the mean position: the z vertex. In the CDC, up to three z vertices may be found by fitting separated Gaussian clusters. The vertex which has the most tracks used to define it is identified as the primary vertex. The FDC is only used to find the vertex when the CDC was unable to identify one. Due to the crowded tracking environment in the forward region, the FDC can identify only one vertex.

3.2 Jet Finding Algorithm

Jets are reconstructed at DØ by using an iterative cone algorithm to identify clusters of energy in the calorimeter. The first step in this algorithm is the formation of preclusters. Seed towers are identified by storing the E_T and location (η, ϕ) of every tower with $E_T > 1$ GeV. Preclusters are formed by summing the energy from any tower with $E_T > 1$ GeV within a radius $\mathcal{R} = 0.3$ ($\mathcal{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) from the seed towers. The preclustering continues until all seed towers are in a precluster. For each precluster, an E_T-weighted centroid is defined by:

$$\begin{aligned} \eta_{cluster} &= \frac{\sum_i E_{T_i}\eta_i}{\sum_i E_{T_i}} \\ \phi_{cluster} &= \frac{\sum_i E_{T_i}\phi_i}{\sum_i E_{T_i}}. \end{aligned} \tag{3.1}$$

Jets are clustered by starting with the E_T-ordered list of defined preclusters. A cone of radius $\mathcal{R}$ is drawn around each precluster, the E_T is summed for every tower in that cone, and a new centroid is calculated as in Equation 3.1. DØ uses four cone sizes for jet definition: $\mathcal{R} = 0.3, 0.5, 0.7$, and 1.0 ($\mathcal{R} = 0.7$ is used in this analysis). Cones of radius $\mathcal{R}$ are drawn around each new centroid and the process is repeated until a stable jet center is found.

When the iteration to find the jet center is complete, the jet position is recalculated using:

$$\begin{aligned}
\phi &= \tan^{-1}\frac{\sum_i E_y^i}{\sum_i E_x^i} \\
\theta &= \tan^{-1}\frac{\sqrt{(\sum_i E_x^i)^2 + (\sum_i E_y^i)^2}}{\sum_i E_z^i} \\
\eta &= -\ln\tan\frac{\theta}{2}
\end{aligned} \tag{3.2}$$

Only jets with $E_T > 8$ GeV are retained. In this analysis, the jets are E_T-ordered; the leading jet refers to the highest E_T jet in the event.

If the centers of two jet candidates are less than $2\mathcal{R}$ apart, a decision must be made whether to split the energy into two jets or merge it into one. If the E_T in the overlap region is more than 50% of the smaller jet's E_T, the two jets are merged by summing all towers to find the energy and recalculating the jet center. If the shared E_T is less than 50% of the lower jet E_T, the two jets are split by assigning all towers in the overlap region to the nearest jet and recalculating the E_T and centroid for each.

3.3 Jet Energy Scale

The measured jet energy must be corrected for the non-linearity of the calorimeter for particles with energy less than 10 GeV. The jet energy is also corrected for other detector and physics effects which affect the energy of the jet:

- *noise* - in the calorimeter (both electronic and uranium noise).
- *pileup* - negative energy from the previous crossing.
- *underlying event* - energy from soft interactions unrelated to the hard scatter of interest.
- *showering* - algorithm-dependent correction due to the details of the hadron showering in the calorimeter.

The true jet energy is expressed as:

$$E^{jet}_{particle} = \frac{E^{jet}_{measured} - O}{(1-S)R_{had}}. \tag{3.3}$$

O is the offset factor which corrects for noise, pileup, and the underlying event, S is the calorimeter showering correction, and R_{had} is the hadronic calorimeter response.

The response correction dominates the jet energy correction. This correction was derived from data by using the MPF (Missing E_T Projection Fraction) method. This method relies on conservation of energy in dijet events and photon plus jet events to determine the calorimeter response. The component of missing E_T($\not\!\!E_T$), along the direction of the photon is used to obtain the response:

$$R(E') = 1 + \frac{\vec{\not\!\!E_T} \cdot \hat{n}^{\gamma}}{E_T^{\gamma}} \tag{3.4}$$

where $E' = E_T^{\gamma}\cosh(\eta_{jet})$ and $\hat{n}$ is the direction of the photon. Since both E_T^{γ} and η_{jet} are well measured, any imbalance in the photon-jet E_T is interpreted as a mismeasurement of the jet energy. Dijet events are used instead of photon events to correct high energy jets.

More details about the response correction [1] and all other corrections can be found in References [29] and [30].

3.4 Jet Energy Resolution

The jet energy resolution is measured using transverse momentum balance between dijets [31]. The dijet sample used required two jet events in which the dijets must both have $E_T > 15$ GeV, must be back-to-back in ϕ, and are in the same $|\eta|$ region of the detector. The resolution is parametrized as a function of jet E_T by:

$$\frac{\sigma_{E_T}}{E_T} = \sqrt{C^2 + \frac{S^2}{E_T} + \frac{N^2}{E_T^2}} \tag{3.5}$$

where C is a constant factor due to the calorimeter calibration, S is a factor due to shower fluctuations in the calorimeter, and N is from noise and the underlying event. The resolution depends on details of the detector and is therefore different for the central calorimeter, end calorimeter, and Inter-cryostat region. The jet resolution parameters for the different detector regions are given in Table 3.1.

[1] CAFIX 5.0 was used for this analysis

η region	N	S	C
< 0.4	3.848	0.539	0.027
$0.4 \leq \|\eta\| < 0.8$	4.789	0.423	0.041
$0.8 \leq \|\eta\| < 1.2$	3.067	0.635	0.047
$1.2 \leq \|\eta\| < 1.6$	4.654	0.349	0.055
$1.6 \leq \|\eta\| < 2.0$	3.484	0.556	0.0
$2.0 \leq \|\eta\| < 3.0$	3.729	0.428	0.0

Table 3.1: Jet Energy Resolution Parameters

3.5 Jet Position Resolution

The jet position has a finite resolution due to both detector effects and jet finding algorithm. The position resolution of reconstructed jets was determined by using a Monte Carlo simulation [32]. The resolution in η and ϕ for a 30 GeV jet at $\eta = 2.5$ are 0.07 and 0.10 respectively.

Chapter 4

Data Selection

The data for this analysis was taken during the 1994-1995 Tevatron Collider run. A forward dijet trigger implemented in the 1992-93 run for the rapidity gap analysis was successful in increasing the statistics for opposite-side jet events. This trigger was refined for the 1994-95 run to increase the statistics for the measurement as a function of jet E_T and $\Delta\eta$.

4.1 Level Ø and Level 1 Triggers

The first level of the trigger requires the presence of inelastic proton-antiproton collisions by using the Level Ø scintillation counters. The Level 1, or hardware, trigger requires the presence of energy clusters in the calorimeter to signal jets. The low E_T trigger used trigger towers while the medium and high E_T triggers used large tiles to trigger on jets. The low E_T trigger required two trigger towers with $E_T > 2$ GeV and $|\eta| > 1.0$. The medium and high E_T triggers required two large tiles with $E_T > 12$ GeV and $|\eta| > 1.6$. Additionally, two trigger towers with $E_T > 3$ GeV were also required in the medium and high E_T triggers to reduce the sensitivity of the trigger rate to luminosity. To reduce contamination from multiple interactions, a single interaction cut was applied at the trigger level. This cut used timing information from the Level Ø luminosity counters to select single interactions (see Section 2.2). A control trigger without the single interaction requirement was also taken to study the effects of multiple interactions in the data.

A 'signal' trigger was also implemented which was used to increase the statistics for events with a rapidity gap in order to study the characteristics of color singlet events. This trigger, called jet_gap_veto, was the same as the high E_T trigger at Level 1 with the additional requirement of a veto on central jets. This veto was implemented by rejecting events with any large tiles with $E_T > 5$ GeV in the region $|\eta| < 1.6$.

4.2 Level 2 Filters

The Level 2, or software, trigger requires two jets with E_T and η cuts as given in Table 4.1. Many different filters were used at Level 2 to increase the statistics for the E_T and $\Delta\eta$ measurements. A cut on the scalar sum of energy in the calorimeter of $E < 2000$ GeV was applied to the medium E_T trigger after Run 85277 to further reduce triggering on events with multiple interactions. A same-side jet trigger was implemented at Level 2 which is the same as the high E_T trigger with the $\Delta\eta$ cut removed. The veto trigger had the same Level 2 filters as the medium and high

E_T triggers.

	Trigger name	E_T	$\|\eta\| >$	$\Delta\eta >$
Low E_T	jet_gap_12_qgt	12	1.6	3.2
Medium E_T	jet_gap_lhe	18	1.6	4
High E_T	jet_gap_hhe	25	1.6	4
Same-side	jet_gap_same	25	1.6	—

Table 4.1: Cuts for different Level 2 filters

4.3 Offline Selection Criteria

The following offline cuts are applied to reach the final event sample.

4.3.1 Event Selection

The z vertex of the event is required to be within 50 cm of zero to keep the event centered in the detector. Single interactions are required by using the multiple interaction tool equal to one (see Appendix C for details on the multiple interaction tool). This cut requires a single interaction as indicated by both the timing information from the Level Ø luminosity counters and from the tracking system.

Removal of Problematic Physics Runs

Runs which are known to have problems with either the calorimeter or Central Drift Chamber (CDC) are removed from the data sample. Example of problematic runs are ones in which the high voltage to the CDC was off (or not fully on) and runs in which channels in the calorimeter are known to have been noisy due to malfunctioning electronics.

4.3.2 Suppression of Noisy Calorimeter Towers

Pathologically noisy calorimeter towers will cause the misidentification of particles between the jets, which will affect the measured rapidity gap fraction. Noisy towers in the electromagnetic calorimeter are therefore not counted when measuring the multiplicity of calorimeter towers between the dijets.

These towers are typically noisy over a time period which can be as short as one physics run or as long as several months. To identify noisy towers, the occupancy of EM calorimeter towers with $E_T > 200$ MeV was examined for varying time intervals. If the occupancy of any individual tower was more than 3σ above the mean occupancy of towers at the same detector η, that tower was considered noisy and added to a list to be ignored. Towers are ignored only for the time period in which they are determined to be noisy. Figure 4.1 shows the EM tower occupancy for the entire data sample. A few noisy towers are visible as spikes rising above the nominal occupancy.

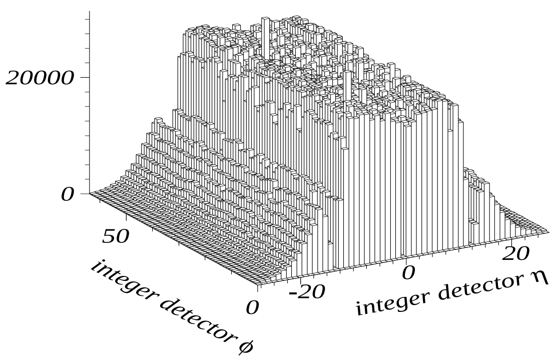

Figure 4.1: Electromagnetic Calorimeter tower occupancy showing a few noisy towers above the background distribution.

4.3.3 Jet Selection

The leading two jets in the event are required to have $|\eta| > 1.9$ for the measurement as a function of E_T and $|\eta| > 1.7$ for the other measurements. The more stringent cut for the E_T measurement is to ensure that variation in jet width as a function of E_T will not affect the measured color-singlet fraction. The leading jets are also required to have a rapidity separation, $\Delta\eta > 4$ and E_T cuts as listed in Table 4.2. The leading two jets are required to pass the following jet quality cuts:

		N_{events}	
Sample	E_T	$\|\eta\| > 1.9$	$\|\eta\| > 1.7$
Low E_T	15	6200	—
Medium E_T	25	32502	—
High E_T	30	71646	101944

Table 4.2: Three samples used for measurement of f_s versus E_T

- *Coarse hadronic fraction*: The fraction of energy deposited in the Coarse Hadronic section of the calorimeter is required to be less than 0.4 to eliminate spurious jets caused by Main Ring energy deposition.

- *Electromagnetic Fraction*: The fraction of energy deposited in the electromagnetic calorimeter must be less than 95% for all jets and greater than 5% for jets with $|\eta| < 1$ or $|\eta| > 1.5$. This cut rejects electrons and photons mimicking jets as well as spurious jets due to noise in the electromagnetic calorimeter .

- *Hot Cell Fraction*: The hot cell fraction is the ratio of the energy in the highest energy cell in the jet to the second highest cell. This fraction is required to be less than 10 to reject jets created by noise in the calorimeter.

4.3.4 Trigger efficiencies

In order to measure a cross section, the triggers should either be used in a fully efficient region, or the efficiency should be well measured. The trigger efficiency is of less concern in this analysis because we measure the fraction of rapidity gap to all events. It is possible that there are some higher order biases which favor gap events over QCD events or vice versa. In order to avoid these biases and retain sufficient statistics for our measurements, we use the triggers where they are at least 50% efficient. The efficiency for dijet triggers was not previously determined, but we can use the values determined for single jet figures as a guide. The low E_T trigger has two jets with $E_T > 12$ GeV at Level 2. The efficiency for a single jet trigger above 12 GeV is above 60% at an E_T of 15 GeV [33]. If both jets are required to have $E_T > 15$ GeV the dijet trigger should have an efficiency of greater than 50%.

The efficiencies of the higher E_T triggers are determined relative to the 12 GeV trigger. Figure 4.2 shows the efficiencies of the medium E_T relative to the low E_T trigger and of the high E_T relative to the medium E_T trigger. The efficiency for the second leading jet is found by dividing the E_T distributions for the two triggers. The efficiency for the leading jet is found the same way, but the leading jet E_T distribution is made for events with the second leading jet already efficient. These plots determine the 25 and 30 GeV cuts on the medium and high E_T triggers.

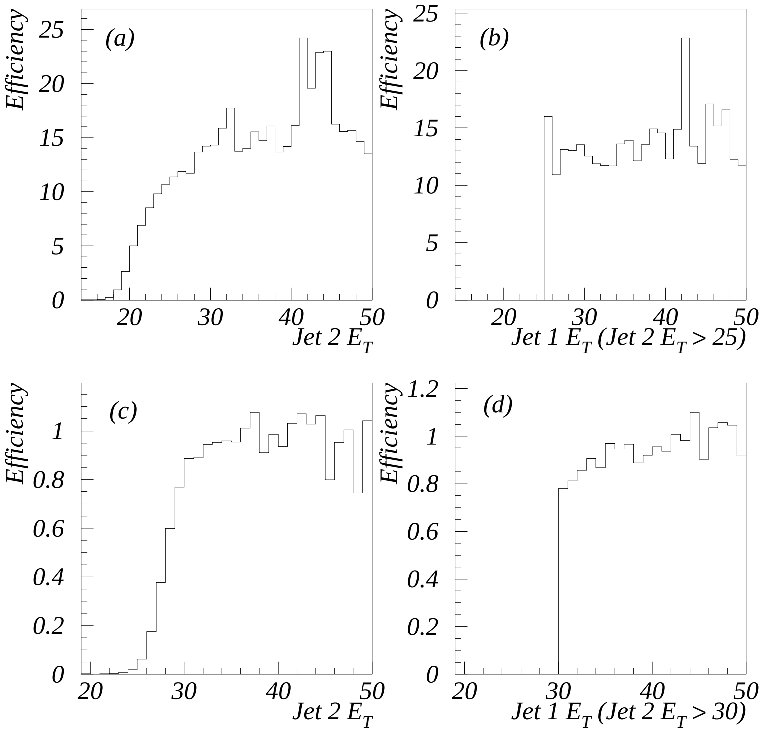

Figure 4.2: (a)Jet_2 E_T(medium trigger)/Jet_2 E_T (low trigger) (b)Jet_1 E_T (medium trigger)/Jet_1 E_T (low trigger) with Jet_2 $E_T > 25$. Note that these 'efficiencies' do not level out at one because of different trigger prescales. (c)Jet_2 E_T (high trigger)Jet_2 E_T (medium trigger) (d)Jet_1 E_T (high trigger)/Jet_1 E_T (medium trigger) with Jet_2 $E_T > 30$.

4.3.5 Final Data Sample

Characteristics of the remaining data sample are shown in Figs. 4.3 and 4.4 for each of the three E_T samples. The dips in the jet ϕ for the two leading jets are due to triggering with the large tiles. For single jet triggers, raising the E_T cut so the trigger is efficient removes the dips in the jet ϕ distribution. This was not the case for the dijet trigger; raising the E_T cut did not affect the shape of the ϕ distribution for the leading jets. Because the calorimeter is uniform in ϕ, it is not expected that the inefficiencies due to the large tiles will affect the measured color-singlet fraction. These figures show that the event sample consists largely of back-to-back dijet events with a small boost and a small difference in jet E_T. Table 4.3 shows, for the high E_T trigger, how each offline cut affects the number of events.

Cut	N_{events}	Percent cut
Trigger	379,875	—
Single Interaction	244,765	36
Vertex Position	227,782	7
$\|\eta\| > 1.7$	187,192	18
$E_T > 30$ GeV	114,833	39
$\Delta\eta > 4$	112,807	2
Jet quality	107,875	4
Bad runs removed	101,944	5

Table 4.3: Number of events for High E_T trigger

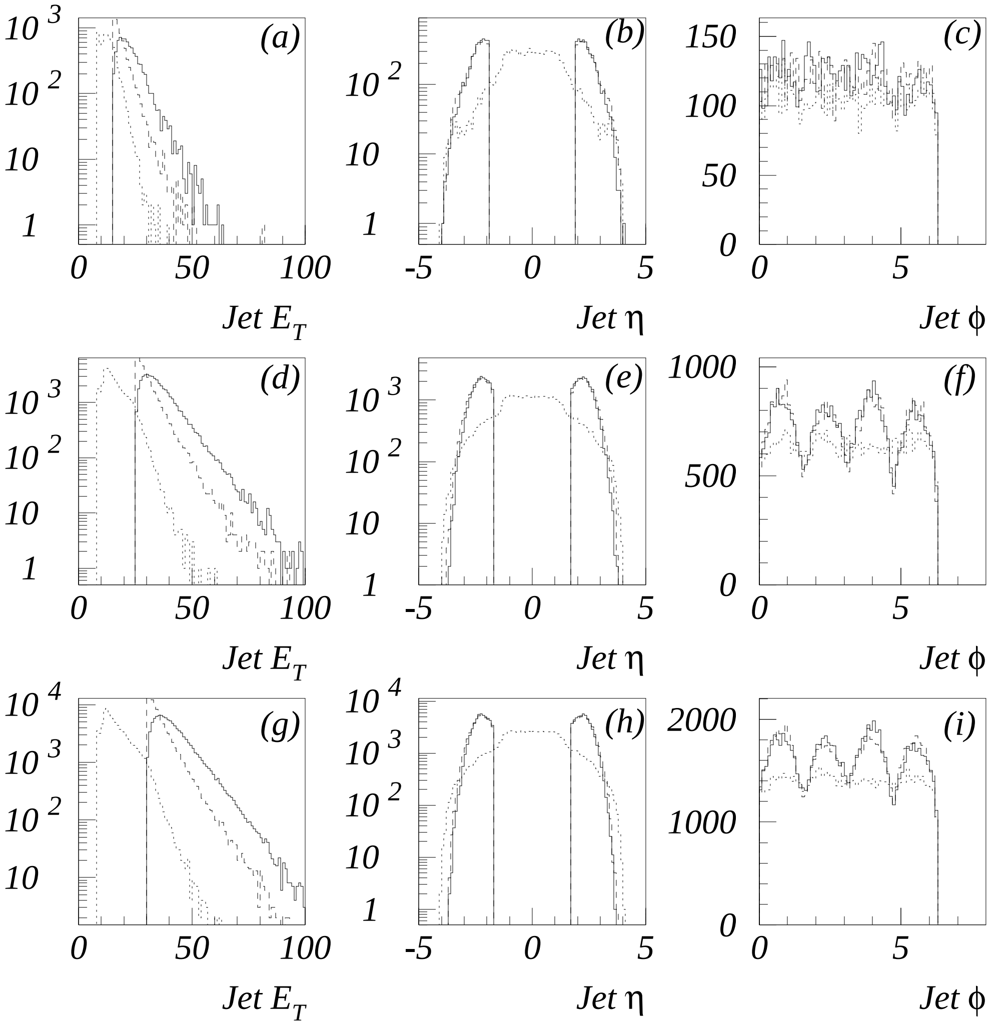

Figure 4.3: Jet Characteristics for Data Samples after Cuts. Solid line is leading jet; Dashed line is second leading jet; Dotted line is Third leading jet. (a)-(c) Low E_T sample; (d)-(f) Medium E_T sample; (g)-(i) High E_T sample.

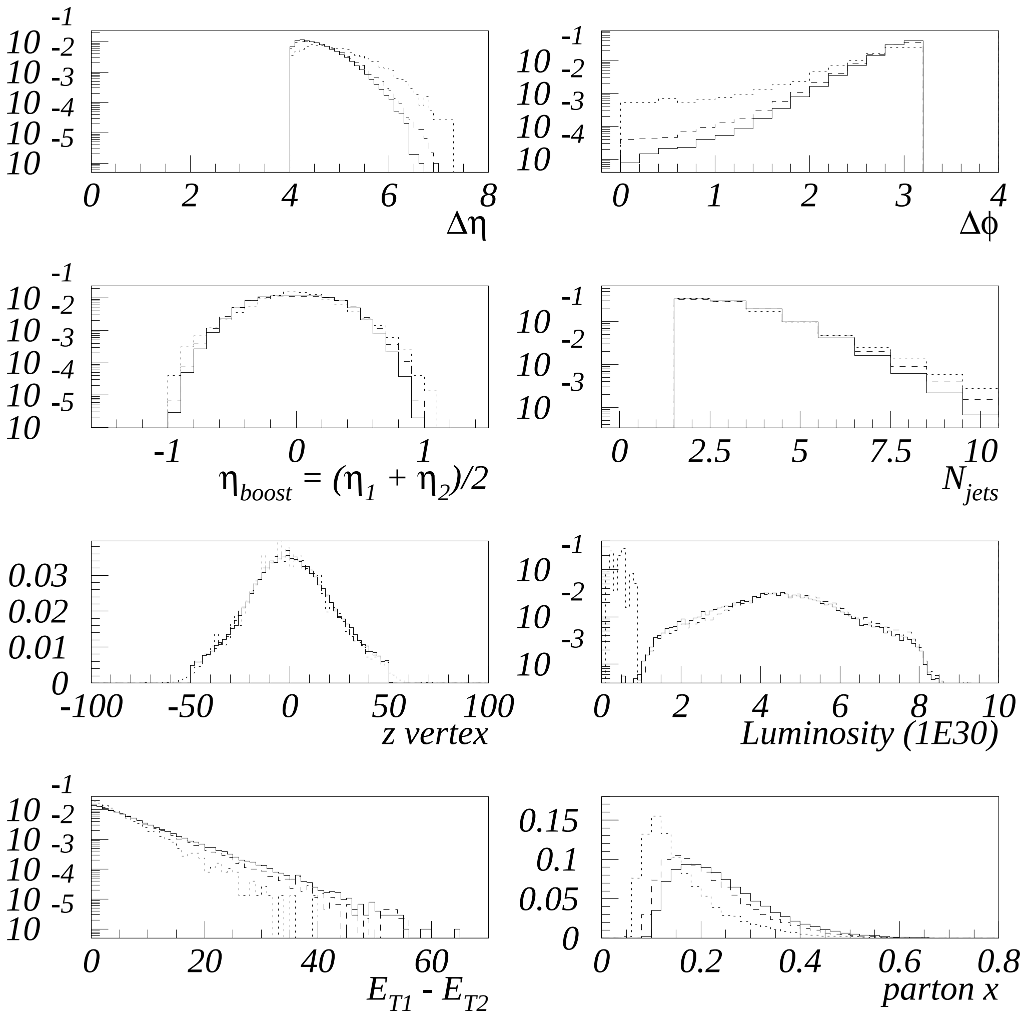

Figure 4.4: Event Characteristics for Final Data Samples: solid line is high E_T sample; dashed line is medium E_T sample; dotted line is low E_T sample.

Chapter 5

Observation of Color-Singlet Exchange

This chapter presents the observation of color-singlet exchange in dijet events. First, the particle tagging methods are discussed. Rapidity gaps in opposite–side dijet events are compared to a sample in which no signal is expected to show that the excess of rapidity gap events is a signal for color-singlet exchange. The features of the color-singlet events are compared to color-exchange events.

5.1 Particle Tagging

The particle multiplicity distribution between the two leading jets is used to find the color-singlet fraction. The features of the particle multiplicity distribution are dependent on the detector used to define it. However, when a fit is used to subtract the color-exchange background, the resulting color-singlet fraction is expected to be independent of the detector used to define the multiplicity distribution (see Section 7.1.2). Three regions of the detector are used to count the particle multiplicity.

- cal - count number of EM towers (with $E_T > 200$ MeV) between the leading jet cone edges.
- cc (central calorimeter) - count number of EM towers in the region $|\eta| < 1$ (jet edges may be further out in η).
- trk - count number of charged tracks in the region $|\eta| < 1.0$

5.2 Observation of Rapidity Gap Signal

The rapidity gap signal is observable as an excess of events at low multiplicity in opposite-side dijet events as compared to events in which no signal is expected. One sample in which no signal is expected contains two forward jets in the same side of the detector (referred to as same-side data). These events are the result of a boosted system in which a parton of small momentum fraction interacted with a parton of large momentum fraction. Central rapidity gaps are not expected in this data because the small η separation between the jets implies that color-exchange is dominant. Also, a rapidity gap due to color-singlet exchange would be between the two forward jets, and not in the central region. Hard single diffraction could produce a central rapidity gap with forward dijets, but this process is suppressed by requiring an inelastic interaction (a beam-beam coincidence in the trigger).

Figure 5.1 shows a two dimensional plot of the number of calorimeter towers (n_{cal}) versus the number of charged tracks (n_{trk}) in the region $|\eta| < 1$ for the opposite-side and same-side dijet data. A striking excess of events at low multiplicity in the opposite-side data as compared to the same-side data indicates the presence of color-singlet exchange. This observation of a new class of events published in Ref. [14] is clearly confirmed by the much larger Run 1b data sample. The color-singlet fraction is measured by parametrizing the one-dimensional n_{cal} multiplicity distribution with a NBD to subtract the color-exchange background. The details of this measurement are given in Chapter 7.

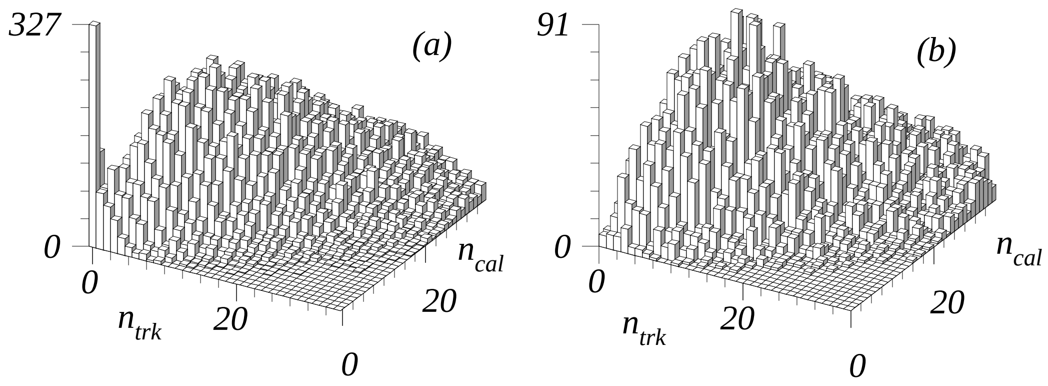

Figure 5.1: n_{cal} versus n_{trk} in the region $|\eta| < 1$ for (a) Opposite-side dijets, and (b) Same-side dijets.

5.3 Studies of Color-Singlet Events

The large statistical sample collected in the 1994-95 run makes it possible to study rapidity gaps events in detail. These studies are interesting because relatively little is known about the nature of strongly-interacting color-singlet exchange. A given rapidity gap event cannot be classified as color-singlet exchange since the multiplicity in a color-exchange event can fluctuate to give a rapidity gap. Therefore, a statistical analysis is done in which the characteristics of a rapidity gap sample are compared to background samples.

The signal sample, which consists of events with $n_{cal} = n_{trk} = 0$, is chosen to have high signal content and low color-exchange background. The color-exchange background is estimated to be about 5% (see Section 6.2 for details on the background estimate). A color-exchange sample consisting of events with $n_{cal} = 3$ and $n_{trk} > 2$ is compared to the signal sample. Events with low multiplicity ('quiet' events) are chosen for this sample to avoid biases in the comparison purely due to differences in the event multiplicity. This sample is made up largely of color-exchange events without a spectator interaction. Using a Single Negative Binomial Distribution fit to the n_{cal} distribution to parametrize the color-exchange contribution gives a color-exchange fraction of greater than 98% for this sample. The signal sample contains 910 events and the quiet color-exchange sample contains 329 events. Finally, an inclusive sample of all high E_T dijet events is also compared to the signal and quiet color-exchange samples.

General characteristics of rapidity gaps events compared to both background samples are shown in Figs. 5.2 and 5.3. The rapidity gap events are largely back-to-back dijet events with little additional radiation. The jet E_T spectrum looks similar for all three samples. The ΔE_T distribution is narrower for the signal and quiet color-exchange samples than for the inclusive sample because the quiet events are mostly two jet events. The $\Delta\eta$ distribution is also quite similar for all three samples. The $\Delta\phi$ distribution shows that the signal sample contains slightly more back-to-back jet events than the quiet color-exchange sample. The small difference is consistent with the difference in the number of jet between the two samples. The jet multiplicity distribution shows that both the signal and quiet color-exchange samples are predominantly two jet events while the inclusive sample contains mainly multijet events. The luminosity profile is quite similar for all three samples, thus showing that the rapidity gap events were not produced during a limited time interval.

All of these characteristics of the color-singlet events are quite similar to those of the quiet color-exchange sample. Both quiet samples have very different features than the inclusive sample. This leads one to conclude that the signal sample event characteristics shown in Figs. 5.2 and 5.3 are features of quiet events and not necessarily unique to color-singlet exchange events.

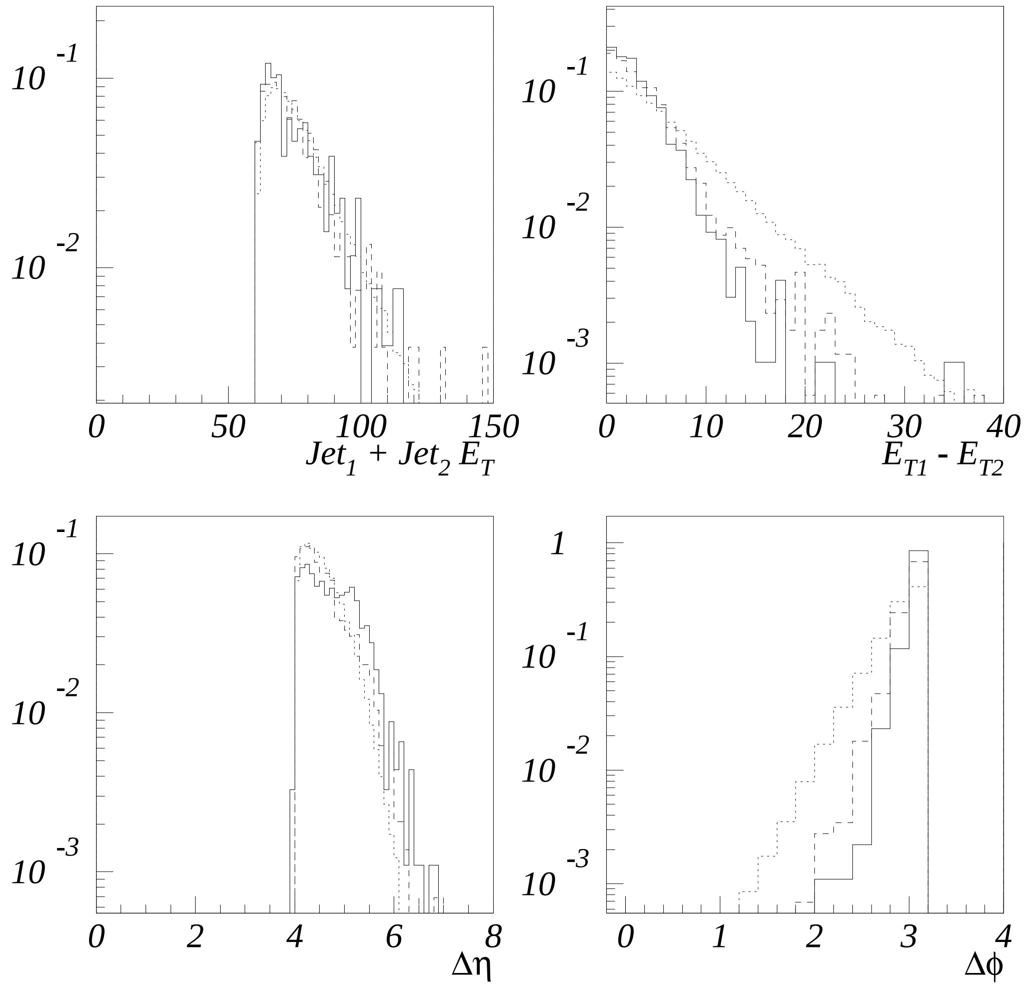

Figure 5.2: Characteristics of three samples of events: solid line is signal sample; dashed line is quiet color-exchange sample; dotted line is inclusive sample.

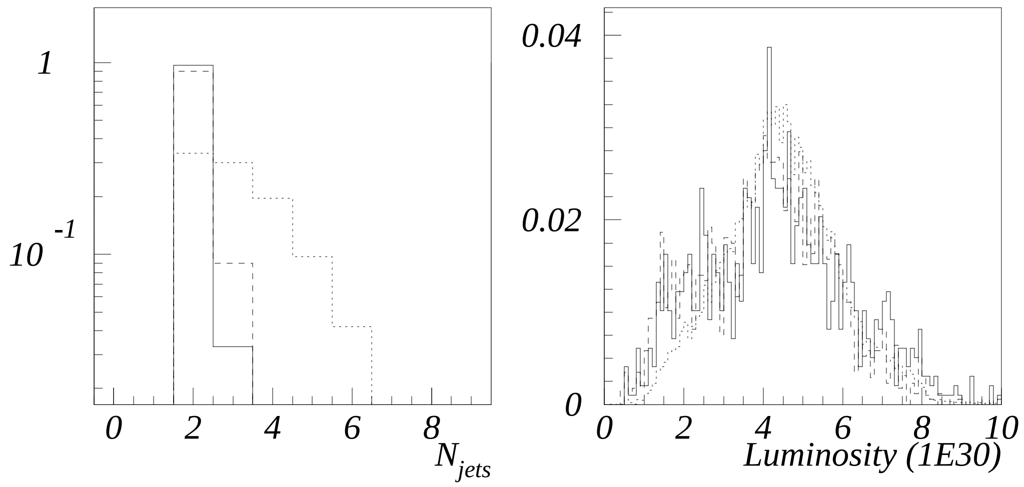

Figure 5.3: Characteristics of three samples of events: solid line is signal sample; dashed line is quiet color-exchange sample; dotted line is inclusive sample.

Another comparison which might have more discriminating power is the particle multiplicity in other regions in the event. The particle multiplicity in the region between the jet and the beam direction ($|\eta| > |\eta_{jet}| + 0.7$), in the jet cone, and in the η band ($\eta_{jet} \pm 0.7$) of the jets, excluding the jet cone are examined (Fig. 5.4 regions C, A, and B, respectively). These regions extend the measurement of the particle multiplicity to the forward region. At high η, a calorimeter tower threshold of $E_T > 200$ MeV becomes quite a high energy threshold ($E > 3.3$ GeV at $\eta = 3.5$). Consequently, the particle multiplicity for these studies is defined as the multiplicity of Electromagnetic calorimeter towers with $E > 200$ MeV .

Figure 5.5 shows the multiplicity in the three different defined regions for the signal sample, the quiet color-exchange sample, and the inclusive sample. A limited inclusive sample of 1450 events was used for these comparisons. The multiplicity

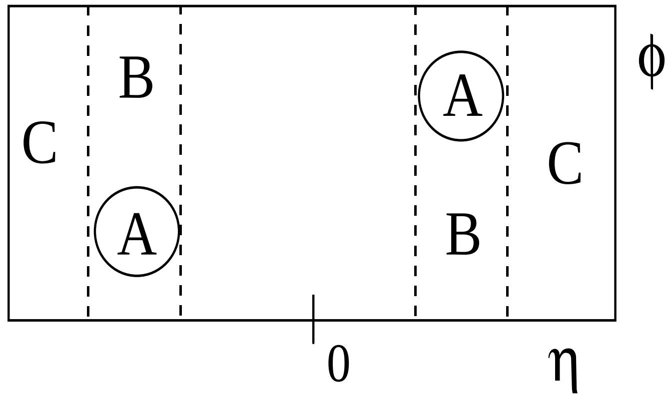

Figure 5.4: Regions used for signal studies: A is in the jet cone; B is the eta band of the jet; C is the beam-jet region.

in the jet cone is similar for all three samples. However, in the band of the jet and in the beam-jet region, the signal sample has a lower multiplicity than the quiet background, which in turn has a lower multiplicity than the inclusive sample. These distributions show that the multiplicity in an event is correlated across η-ϕ: events with a central rapidity gap also have lower multiplicity in other regions than events without a rapidity gap. These correlations are presumably due to differences between color-singlet and color-exchange processes.

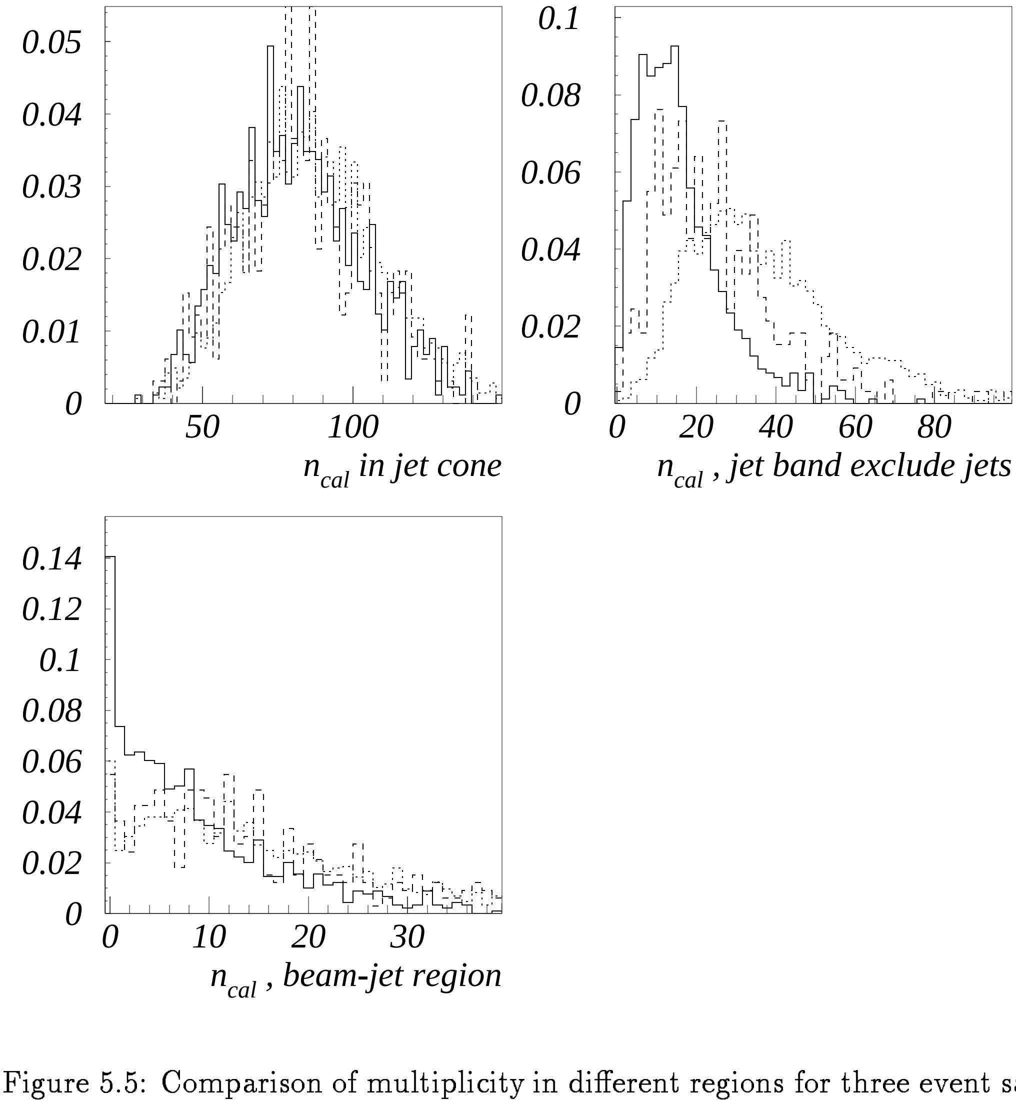

Figure 5.5: Comparison of multiplicity in different regions for three event samples: solid line is signal events; dashed line is quiet background sample; dotted line is inclusive sample.

Chapter 6

Analysis Methods

6.1 Definition of f_s

The color-singlet fraction, f_s, is obtained by fitting the particle multiplicity distribution with a Negative Binomial Distribution (NBD). The NBD is of the form

$$P(n;\overline{n},k) = \alpha\binom{n+k-1}{n}\left(\frac{\overline{n}}{\overline{n}+k}\right)^n\left(\frac{k}{\overline{n}+k}\right)^k \tag{6.1}$$

where $P(n)$ is the probability of observing n particles given the parameters α (a normalization factor), $\overline{n}$ (the mean multiplicity), and k, which is inversely proportional to the width of the distribution. The NBD is a parametrization of the color-exchange background. As discussed in Section 1.4, the color-singlet signal should be visible in the multiplicity distribution as a low multiplicity excess on top of a smoothly falling background. The multiplicity distribution is fit from a starting bin n_0 which gives a good χ^2/df for the fit. The fit is then extrapolated to zero multiplicity to extract f_s. The color-singlet fraction is defined as:

$$f_s = \sum_{n=0}^{n=n_0-1} \frac{data(n) - fit(n)}{N_{total}} \tag{6.2}$$

where N_{total} is the total number of events in the data sample.

A calorimeter multiplicity distribution with a leading edge NBD fit is shown in Fig. 6.1. The solid line is the fit, and the dashed line is the extrapolation of the fit to zero multiplicity. The fractional excess of the data over the fit in the extrapolated region is taken to be the fraction of color-singlet exchange in the data.

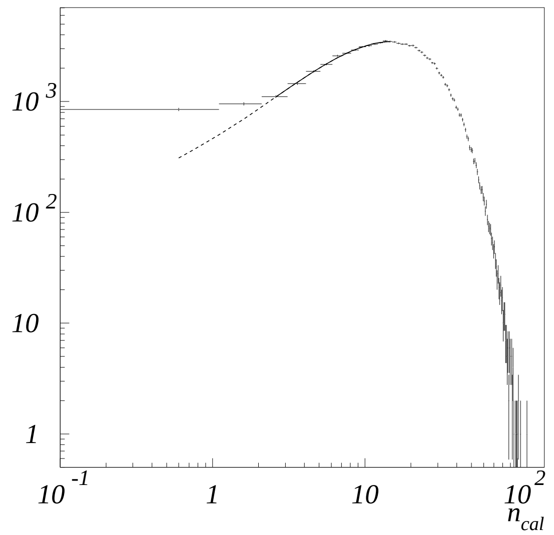

Figure 6.1: Central Calorimeter multiplicity with single Negative Binomial Distribution fit. Solid line is the fit, dashed line is extrapolation of fit to zero multiplicity.

6.2 N_{00} method

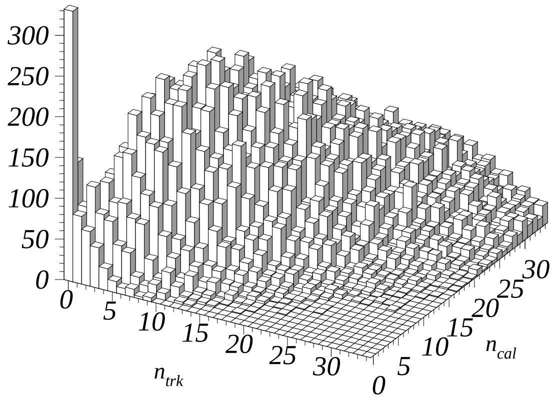

Figure 6.2: n_{cal} vs. n_{trk} for $E_T > 30$ GeV.

As discussed in Chapter 5, the two dimensional plot of the number of calorimeter towers versus the number of charged tracks for $|\eta| < 1$ (Fig. 6.2) shows a striking excess at zero multiplicity. This plot suggests an alternate method for studying the color-singlet fraction using both the calorimeter and tracking information. The fraction of events in the zero track, zero tower (0,0) bin does not require fitting and thus can be accurately measured even for small statistical samples. Using the fitting method, on the other hand, requires larger statistical samples to fit to obtain similar errors. Therefore, the N_{00} method improves the reach and determines the shape of the E_T and $\Delta\eta$ dependence.

The fraction of events with zero towers and zero tracks has a very low color-

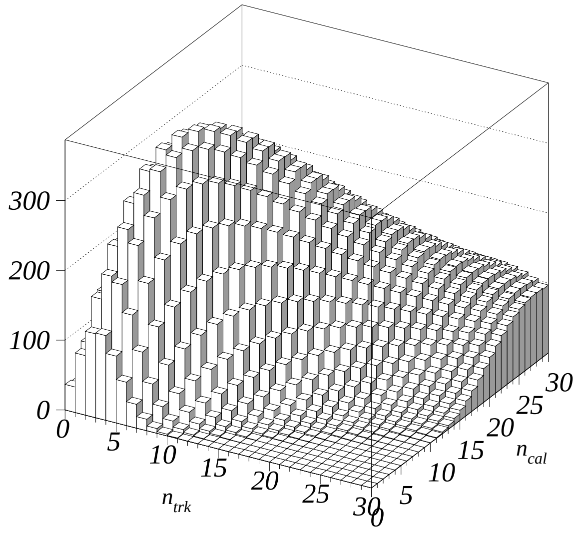

Figure 6.3: Two dimensional Negative Binomial fit to n_{cal} vs. n_{trk} distribution.

exchange background. The background in this bin has been estimated two ways. The first method is to do a two dimensional Negative Binomial fit (Fig. 6.3). The two dimensional distribution is fit from the $(5, 5)$ bin to the $(10, 10)$ bin, and the fit is extrapolated to multiplicity of zero to give a background of about 6%. The second method is to use the same-side data and measure the fraction of events in the $(0, 0)$ bin (Fig. 6.4). This method also gives a small background of about 3%.

The fraction of events with zero towers and zero tracks will be lower than the color-singlet fraction found by fitting because it does not include the color-singlet events which have one tower or one track. Therefore, the color-singlet fraction with this method is defined as: $f_s = f_{00} \cdot C$ where f_{00} is the fraction of events with zero towers and zero tracks and C is a factor which normalizes f_{00} to the color-singlet

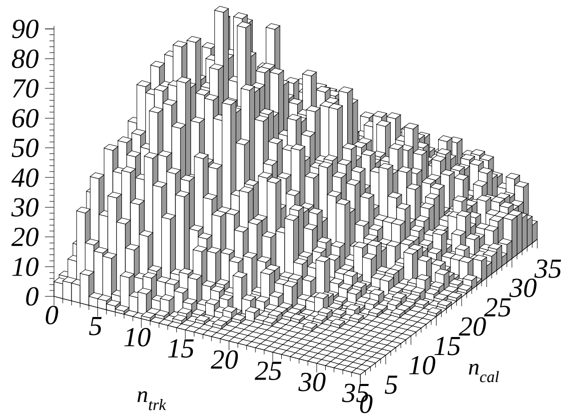

Figure 6.4: n_{cal} vs. n_{trk} for same-side jet events.

fraction found from the fitting method. The normalization factor, C, is defined as $C = \frac{f_s(fit)}{f_{00}}$ This normalization factor will be found for each data sample and applied to every E_T or $\Delta\eta$ bin.

6.3 Confirmation of Background Parametrization

We use two data control samples to demonstrate that the measured excess is not a detector effect and to show that the a leading edge NBD fit is a good parametrization for the multiplicity in color-exchange events. Both of these control samples are expected to have little or no contribution at low multiplicity from color-singlet exchange.

The first control sample is a color-exchange enhanced sample obtained by selecting events with three or more jets. The third jet in the event must be in the η region between the two leading jets. This sample is expected to contain a very small color-singlet exchange component because of the color connections in the event which produce the central jet. The multiplicity of calorimeter towers in the region $|\eta| < 1$, excluding a cone of $\mathcal{R} = 1.0$ around the third jet is shown in Fig. 6.5. Fitting this distribution from $n_0 = 1$ to $n = 14$ gives a color-singlet fraction $f_s = (-0.02 \pm 0.02)\%$, which is consistent with zero.

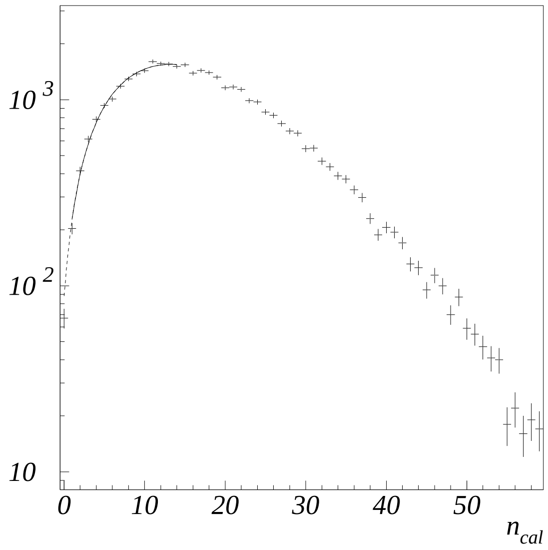

Figure 6.5: Central Calorimeter multiplicity for three jet events where the multiplicity from the third jet is removed.

The second control sample used contains two forward jets in the same side of the detector (referred to as same-side data). The two leading jets are required to have $|\eta| > 1.7$ and $\eta_1 \cdot \eta_2 > 0$. The multiplicity in the central calorimeter ($|\eta| < 1$) is shown in Fig. 6.6. A NBD is again fit from $n_0 = 1$ to $n = 14$ to give a fractional excess of $f_s = (-0.06 \pm 0.02)\%$.

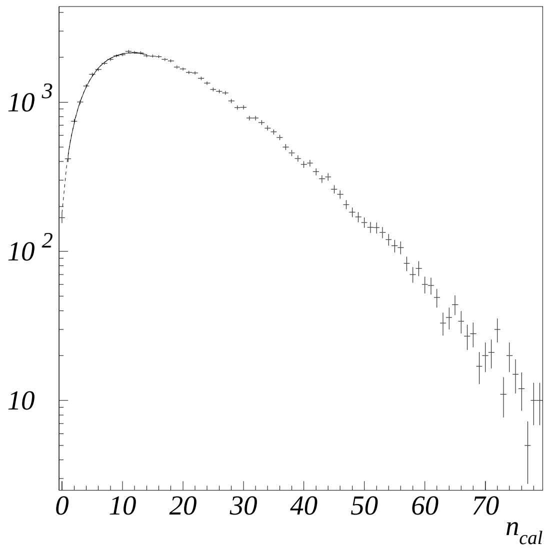

Figure 6.6: Central Calorimeter multiplicity for same-side jet events.

Both of these control samples support the hypothesis that the color-exchange multiplicity distribution is well fit by a leading edge NBD. This data also shows that any excess observed in the opposite-side jet data is not due to a detector effect producing an excess of events with zero multiplicity. Therefore, fitting with a NBD is established as a method to parametrize the color-exchange background. The following section will focus on exactly how these fits should be performed.

6.4 Fitting Method

A Double Negative Binomial Distribution (DNBD) was used to parametrize the color-exchange background for the measurement in Ref. [14]. The relative nor-

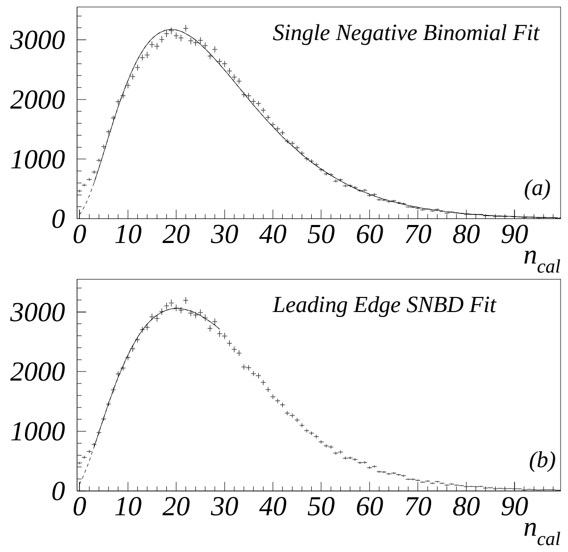

Figure 6.7: Calorimeter multiplicity with Single Binomial Fit of (a) full distribution, and (b) leading edge. Points are data; solid line is the fit; dashed line is extrapolation of fit to low multiplicity.

malization of the two negative binomial distributions was chosen by minimizing the log likelihood of the fit as a function of this normalization ratio. Using this method for the current data has not worked because there is no minimum in this log likelihood distribution. Varying the relative normalization over the full range of possible values shows that no ratio is favored, but the fractional excess is very dependent on the relative normalization. It was shown with the published data that fitting only the leading edge of the multiplicity distribution with a single NBD gave a fraction which agreed within errors with the full DNBD fit.

Figure 6.7 shows both a single negative binomial fit over the range of the distribution and a leading edge fit. The χ^2/df for the full SNBD fit is 306 for 114 degrees of freedom, which gives a confidence level of less than 10^{-6} for the fit. The leading

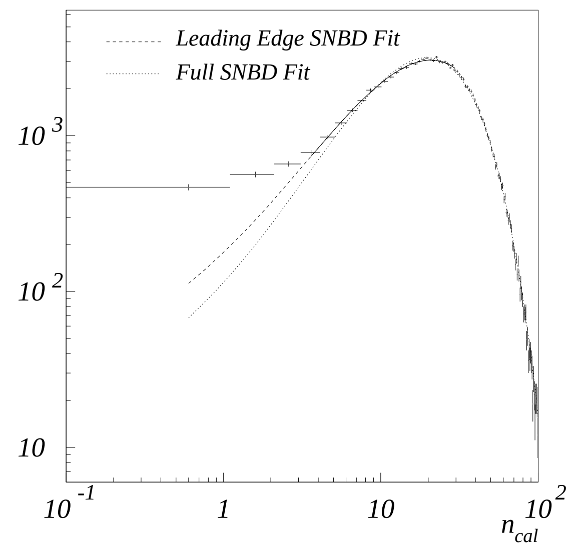

Figure 6.8: Calorimeter multiplicity with Full and Leading edge SNBD fit. The log-log scale emphasizes the low multiplicity region.

edge fit matches the shape of the data much better at low multiplicities. Figure 6.8 emphasizes the low multiplicity region of both fits. The full fit is narrower than the leading edge fit and, therefore, gives a higher signal at low multiplicities.

Another reason to use the leading edge fit is that contamination from multiple interactions affect the shape of the multiplicity distribution at high multiplicities. Figure 6.9 shows the calorimeter multiplicity for low and high luminosity data samples. Both a shift of the mean and the addition of a tail at high multiplicity is evident in the higher luminosity data. Thus, the leading edge fitting method is used for the current measurement because it fits the data better in the region of interest (low multiplicities) and it avoids some effects from multiple interaction contamination.

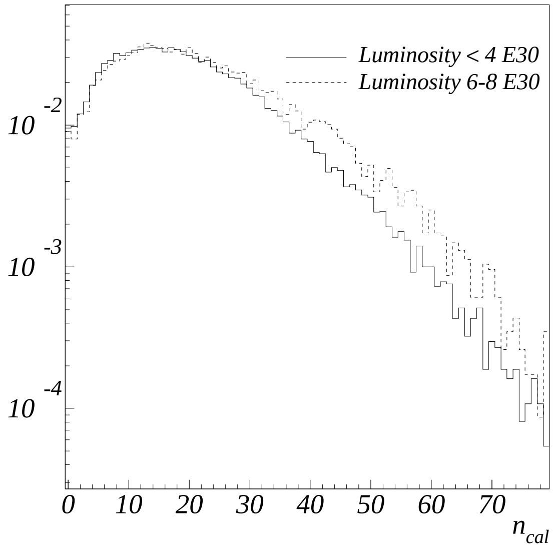

Figure 6.9: Calorimeter Multiplicity for Different Luminosities.

6.4.1 Fitting Studies

To estimate the robustness of the fitting procedure, the leading edge fitting method was studied by generating ensembles of multiplicity distributions and fitting them. A Double Negative Binomial Distribution fits the data well (Fig. 6.10) and is used to study the fitting. The relative weight of each Negative Binomial Distribution is 0.50. An ensemble of distributions based on the data is also generated to study the fit error.

An ensemble of one hundred Double Negative Binomial Distributions is generated. To illustrate the methods of the study, consider the case in which a one percent signal is added to the first two bins: 2/3 in the $n = 0$ bin and 1/3 in the $n = 1$ bin. Each of the one hundred distributions is fit from a range of starting bins, $n_0 = 0$ to $n_0 = 5$. The resulting distribution of the signal fraction found

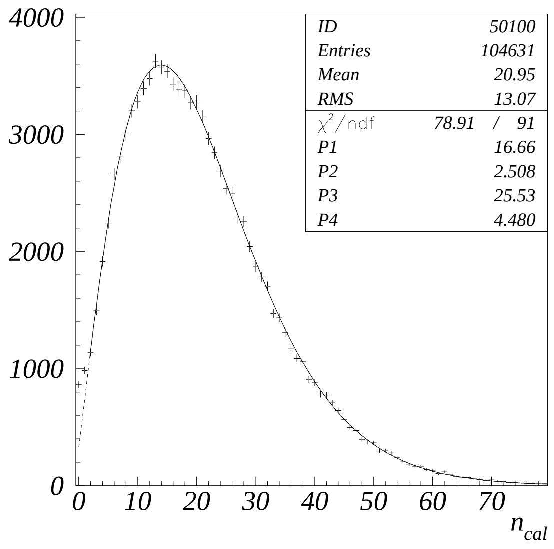

Figure 6.10: Double Negative Binomial Fit to Central Calorimeter Multiplicity Distribution.

from fitting is shown in Fig. 6.11. The rms of the distribution of fit fractions will be used as the fit error. Figure 6.12 shows the confidence limit of the fits for the same range of starting bins. Figures 6.12(a) and (b) show that the fit is bad because the confidence limit is near zero, while Figs. 6.12(c) - (f) indicate a good fit because any value of the confidence limit is equally probable. This example will be discussed fully later in this section. In the tables which follow, the average confidence limit is quoted (an average confidence limit of 50% indicates a good fit).

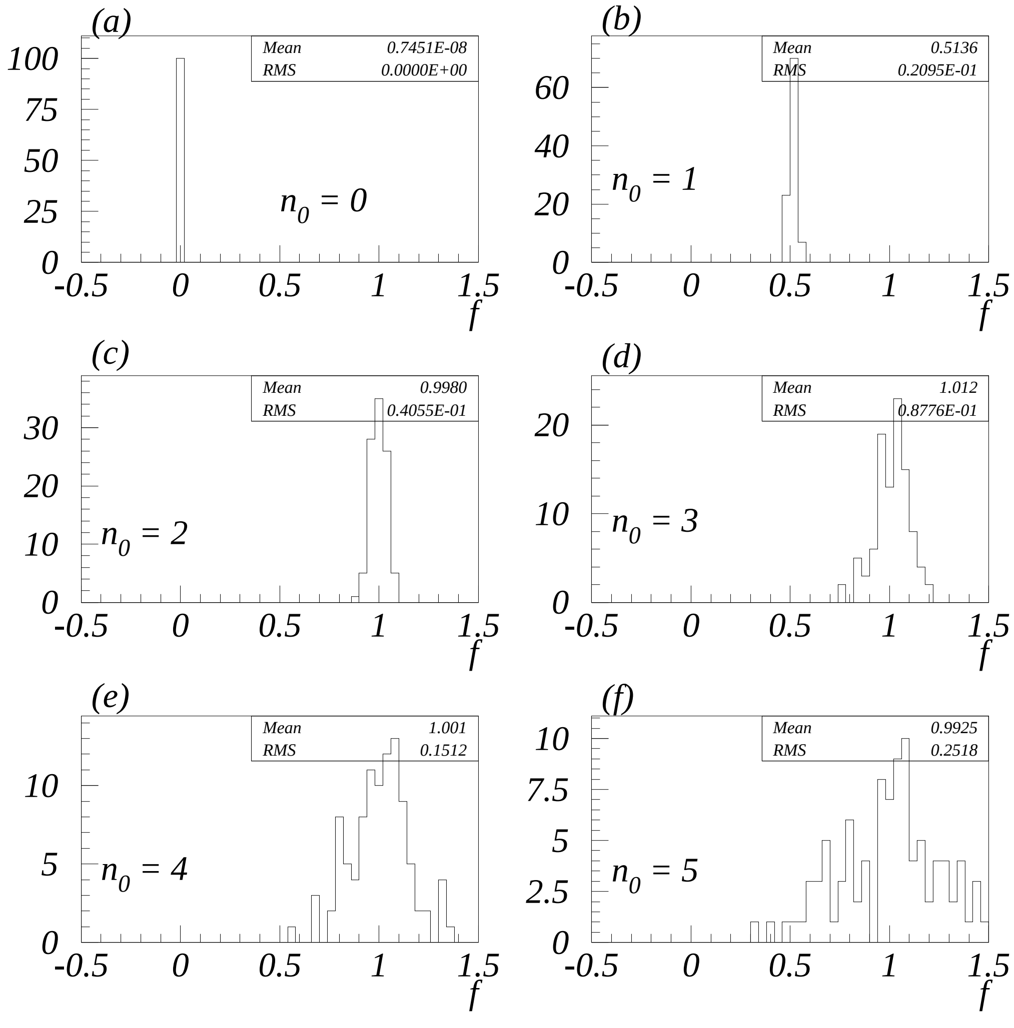

Figure 6.11: Distribution of fit fraction from fitting ensemble of Double Negative Binomial Distributions with one percent signal added((a)-(f) are for starting bin of fit from $n_0 = 0$ to $n_0 = 5$).

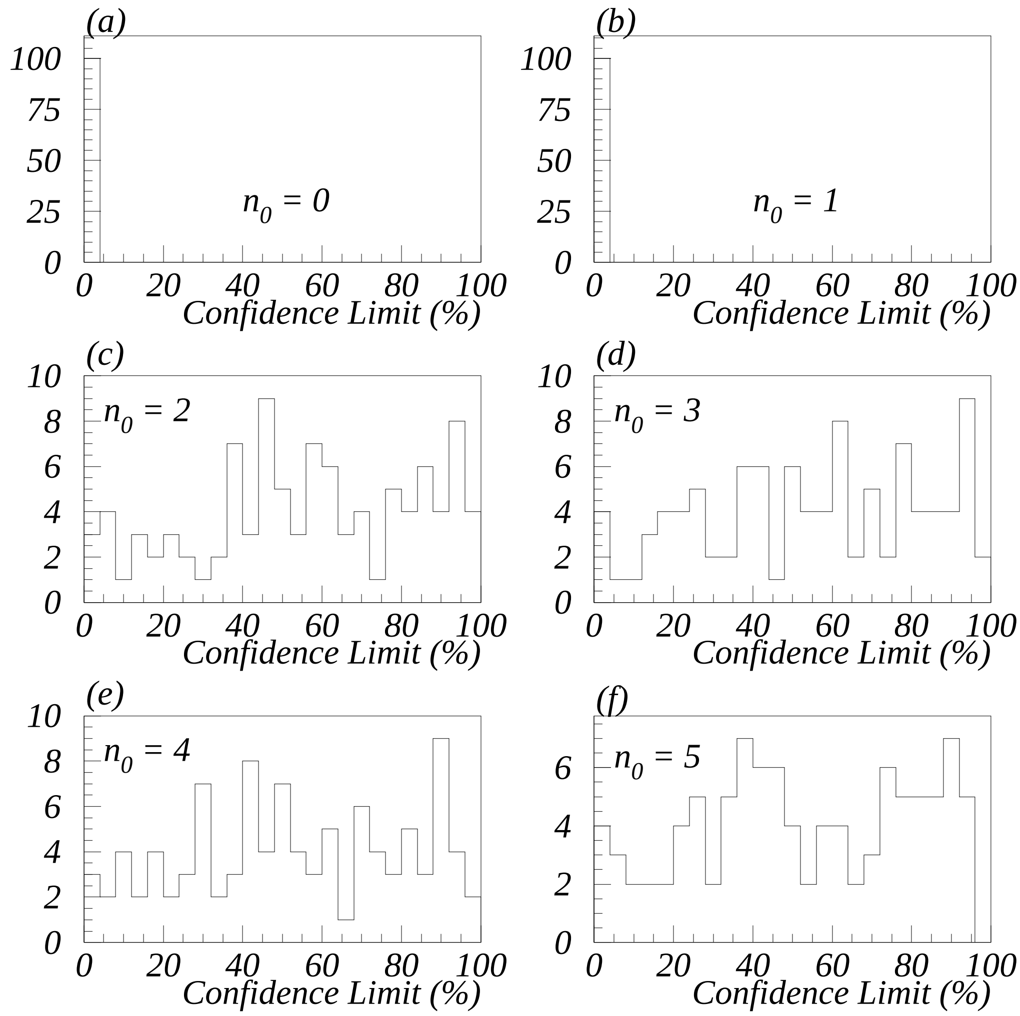

Figure 6.12: Distribution of confidence limits from fitting ensemble of Double Negative Binomial Distributions with one percent signal added((a)-(f) are for starting bin of fit from $n_0 = 0$ to $n_0 = 5$).

First, the fit method is tested by generating one hundred Double Negative Binomial Distributions with no signal. The ensemble of distributions is fit for a range of starting bins from $n_0 = 0$ to $n_0 = 5$ for a fixed ending bin, $n_f = 14$. The fit fraction is consistent with zero and the confidence limit of the fit is good for the range of n_0 (Table 6.1). This shows that the color-singlet fraction should not depend on the starting bin of the fit (as long as n_0 is high enough to include all of the signal). The Double Negative Binomial Distribution ensemble is then fit from $n_0 = 2$ to a range of ending bins: $n_f = 14$ to $n_f = 99$. The fit fraction increases with n_f up to $f_s = .25\%$ for $n_f = 99$ (Table 6.2). Because the color-singlet fraction is biased high as n_f is increased, only the leading edge ($n_f \cong$ peak of distribution) should be fit.

n_0	average $f_s(\%)$	rms	average CL(%)
0	—	—	56.7
1	-0.03	0.02	56.6
2	-0.02	0.04	56.2
3	0.01	0.09	55.1
4	0.001	0.15	53.1
5	0.04	0.25	52.3

Table 6.1: Fraction and confidence limit vs. starting bin of fit for Double Negative Binomial with no signal added.

Next, a signal of one percent is added to the first two bins of the Double Negative Binomial Distribution with 2/3 of the signal in the bin $n = 0$ and 1/3 in the bin $n = 1$. The leading edge is fit with a range of starting bins from $n_0 = 0$ to $n_0 = 5$ (Fig. 6.11) to see if the confidence limit of the fit can be used to choose the correct starting bin (Fig. 6.12).

Table 6.3 shows that, with a large statistical sample, the confidence limit can

n_f	average f_s(%)	rms	average CL(%)
14	0.003	0.04	57.0
19	0.04	0.03	53.8
24	0.04	0.03	38.5
29	0.07	0.03	15.2
49	0.17	0.03	< 2
69	0.22	0.03	< 2
99	0.25	0.03	< 2

Table 6.2: Fraction and confidence limit vs. ending bin of fit for Double Negative Binomial with no signal added.

be used to accurately find the correct n_0 (since fits with $n_0 \geq 2$ have acceptable confidence levels and give a signal near the correct value of 1%). With lower statistics, however, the confidence limit indicates a good fit in a bin which includes signal. One might conclude from the confidence limit alone that $n_0 = 1$ is the correct fit and incorrectly obtain a fraction of 0.5%. Therefore, for samples with low statistics, n_0 should be increased by at least one bin above the nominal value to include all of the signal. Note that including too many bins increases the fit error (rms) and can bias the color-singlet fraction.

	Full statistics (170,000 events)			8,500 events		
n_0	average f_s(%)	rms	average CL(%)	average f_s(%)	rms	average CL(%)
0	—	—	< 2	—	—	< 2
1	0.51	0.02	< 2	0.51	0.09	24.2
2	1.00	0.04	56.2	0.98	0.18	50.2
3	1.01	0.09	55.0	0.95	0.36	48.5
4	1.00	0.15	53.1	0.71	0.50	48.1
5	0.99	0.25	52.6	0.57	0.54	47.2

Table 6.3: Fractional excess and confidence limit (in percent) vs. starting bin of fit for Double Negative Binomial with one percent signal added.

Finally, an ensemble of data distributions is generated by throwing random numbers under the shape of the data multiplicity distribution. Fitting this ensemble illustrates the effect of statistical fluctuations in the data on the resulting fit fraction and error. The ensemble of data distributions is fit from a range of $n_0 = 0$ to $n_0 = 5$ with $n_f = 14$ (Table 6.4). The confidence limit of the fit indicates that the signal is in the first two bins. As n_0 is increased, the signal changes, but its spread also increases. The rms of the distribution of the fraction can be used as the fit error.

n_0	average f_s(%)	rms	average CL(%)
0	—	—	< 2
1	0.48	0.03	< 2
2	0.85	0.05	48.9
3	0.79	0.09	46.5
4	0.65	0.13	46.3
5	0.32	0.33	48.0

Table 6.4: Fraction and confidence limit vs. starting bin of fit for ensemble of data distributions.

These fitting studies show that the ending bin of the fit should be close to the peak of the distribution to fit. The starting bin should be kept low enough to keep the fit error small and high enough to include all of the signal. Although the Double Negative Binomial Distribution is a good fit to our data, it is unclear which combination of Negative Binomial distributions is the correct fit to the data. The ratio of the two Negative Binomial Distributions was changed from 50% to 75% and the results of these studies were unchanged. Because the Double Negative Binomial Distribution fit is ambiguous, these fitting studies are only used as a guide to the fitting method and are not used to pick the absolute starting and ending bins of the fit.

6.5 Determination of Multiple Interaction Contamination

The largest source of systematic error in the published multiplicity analysis was contamination from multiple interactions. This measurement is very sensitive to multiple interactions because even one or two extraneous particles between the jets can spoil a rapidity gap. During the 1994-95 Tevatron run, the average luminosity was higher than for the 1992-93 run. Therefore, multiple interaction contamination is an even greater concern. Using a tight single interaction cut does not remove all the multiple interactions. Some evidence for this contamination is shown in the average multiplicity. Figure 6.13 shows the average multiplicity between the jets for

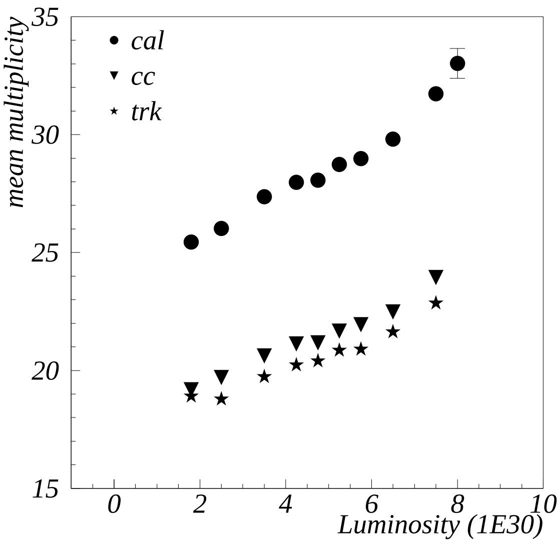

Figure 6.13: Average multiplicity for High E_T sample as a function of luminosity for three tagging methods.

the three different detector regions as a function of luminosity. The contamination is seen clearly by the rise in multiplicity with luminosity. Figure 6.14 shows that the measured color-singlet fraction for the high E_T trigger decreases as a function of luminosity due to this contamination.

A possible method to correct for the residual multiple interaction contamination is to use the total energy measured in the calorimeter. The energy distribution from a single interaction will have a different shape than the distribution from a single interaction with a minimum bias interaction overlayed. A detailed model which accounts for each source of energy in the calorimeter is discussed in Appendix A. Comparing the model for a single interaction to the data gives an estimated multiple interaction contamination of about 12% overall for the high E_T data. Instead of using this estimate for the contamination, however, we use a method to correct the gap fraction for multiple interaction contamination. This correction and a model which demonstrates its validity are detailed in the following section.

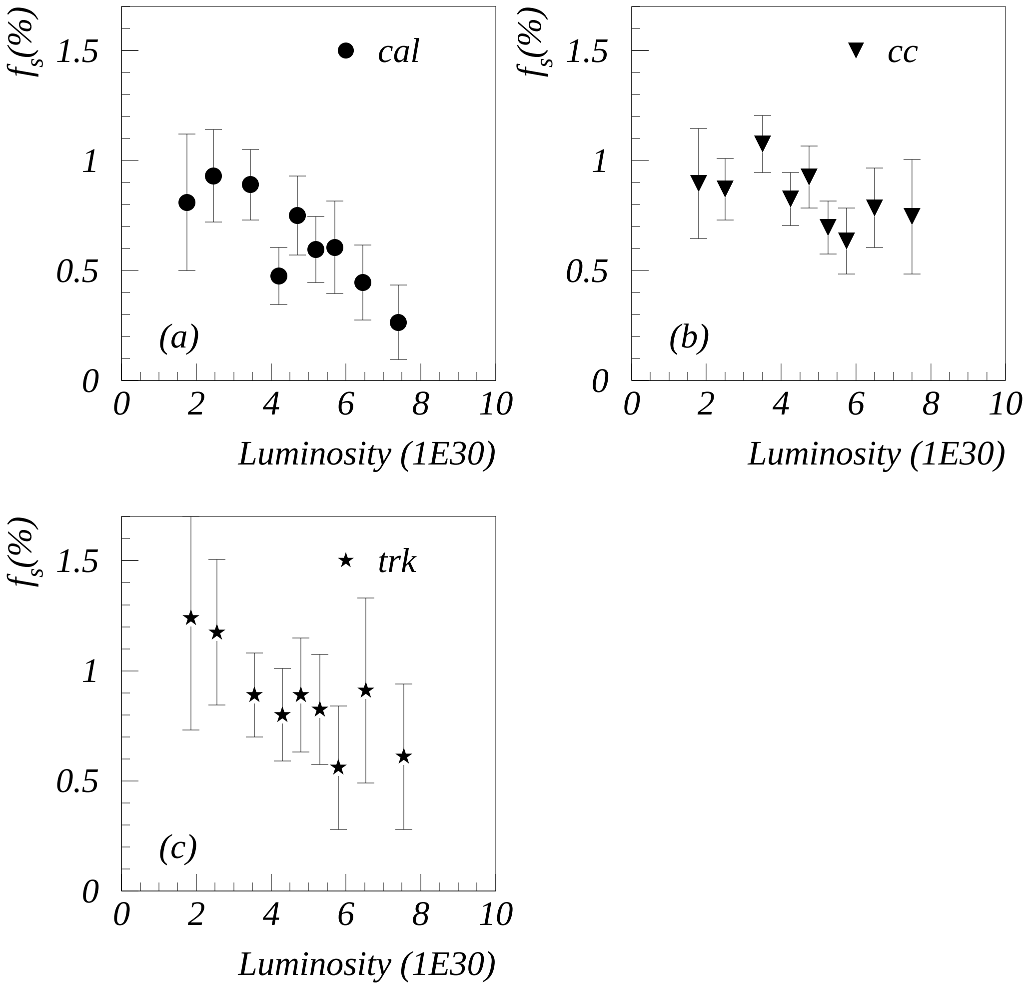

Figure 6.14: Fractional excess as a function of luminosity for High E_T sample when using (a) calorimeter between jet edges, (b) central calorimeter, and (c) tracking to count the multiplicity.

6.6 Luminosity Correction

To use the entire data set, we must find a method in which the color-singlet fraction does not depend on luminosity. The fraction we measure is

$$f_s = \sum_{n=0}^{n=n_0-1} \frac{data(n) - fit(n)}{N_{total}} \tag{6.3}$$

where n_0 is the starting bin of the fit. Because the excess is measured in bins with low multiplicity and multiple interactions will not populate this region, the numerator of the color-singlet fraction has little luminosity dependence. The denominator increases with luminosity as more minimum bias events contaminate the sample.

A new method for finding the color-singlet fraction separates it into two parts: one which depends on luminosity and one which does not. The color-singlet fraction can be expressed as:

$$f_s = \frac{\sum_{n=0}^{n=n_0-1} data(n) - fit(n)}{\sum_{n=0}^{n=n_{sum}} data(n)} \left(\frac{\sum_{n=0}^{n=n_{sum}} data(n)}{N_{total}} \right)_{Lum=0} \tag{6.4}$$

The first term in Equation 6.4 should have very little luminosity dependence if n_{sum} is small. The second term contains the luminosity dependence. This term is found for various luminosities, fitted and extrapolated to luminosity of zero. Therefore, with this method, we are effectively using the denominator (or total number of events) where there is no contamination from minimum bias. Figure 6.9 illustrates the luminosity dependence in the multiplicity distribution. For small values of n_{sum}, there is little luminosity dependence in the data for $n < n_{sum}$, but quite substantial luminosity dependence at higher multiplicity.

The following sections provide support for using Equation 6.4 to correct for the luminosity effects in the data. Section 6.6.1 uses a model for the total multiplicity to show the validity of the correction method. Section 6.6.2 discusses the application of the method to data and shows that the first term in Equation 6.4 is approximately independent of luminosity while the second term is luminosity dependent.

6.6.1 Multiplicity Model

The correction has been tested by using a model of the total multiplicity. The model simulates adding the multiplicity from a minimum bias event to 'uncontaminated' data. The data is at a luminosity of approximately 2E30. The minimum bias multiplicity distribution is parametrized and then convoluted with the data in the following way:

$$mult = P(0)[data] + P(1)[data \oplus MB] + P(2)[data \oplus MB \oplus MB] + \ldots \tag{6.5}$$

$P_{\overline{n}}(k)$ is the Poisson probability for k additional interactions when there are $\overline{n}$ interactions per crossing.

The multiplicity in Equation 6.5 is that which results from including all additional minimum bias interactions that are possible (or one hundred percent acceptance for additional minimum bias interactions). However, the single interaction cut removes much of the minimum bias contamination from the data. The minimum bias acceptance in the model is varied from zero to one hundred percent to simulate the increased acceptance of minimum bias events as a function of luminosity. Other studies indicate that the minimum bias acceptance for our data is about 20%, which results in the 12% estimate for the multiple interaction contamination in Section 6.5.

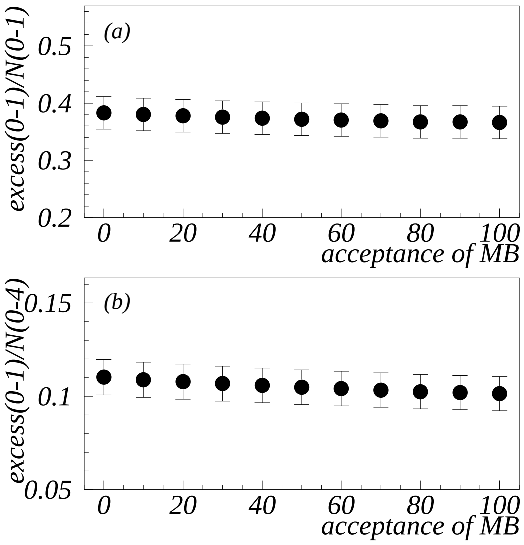

Figure 6.15: (a) $\frac{\sum_{n=0}^{n=1} data(n)-fit(n)}{\sum_{n=0}^{n=1} data(n)}$ and (b) $\frac{\sum_{n=0}^{n=1} data(n)-fit(n)}{\sum_{n=0}^{n=4} data(n)}$ vs. Minimum Bias Acceptance. (First term in Eq. 6.4 for $n_{sum} = 1$ and $n_{sum} = 4$).

The model is used to verify that the first term in the expression for f_s in Equation 6.4 is indeed independent of luminosity by plotting this term as a function of the minimum bias acceptance. Figure 6.15 shows this term for $n_{sum} = 1$ and $n_{sum} = 4$ as a function of the minimum bias acceptance. In both cases, this term is relatively flat with increasing acceptance. In contrast, Fig. 6.16 shows that there is large dependence on minimum bias acceptance when the denominator is the full range ($n_{sum} = 99$).

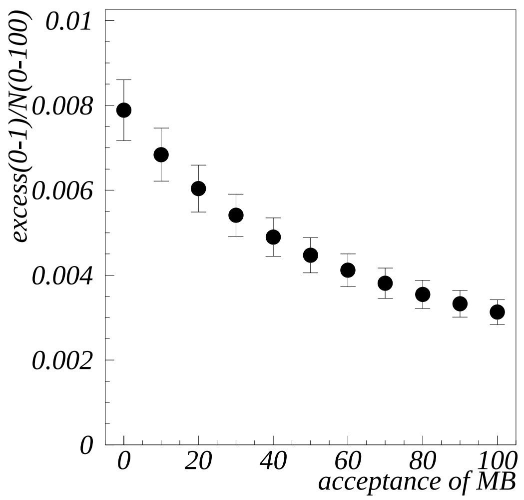

Figure 6.16: $\frac{\sum_{n=0}^{n=1} data(n) - fit(n)}{\sum_{n=0}^{n=99} data(n)}$ vs. Minimum Bias Acceptance. (First term in Eq. 6.4 for $n_{sum} = 99$).

It would be instructive to use the multiplicity model to study the second term of equation 6.4. The denominator, N_{total}, of this term depends on luminosity due to increased contamination from multiple interactions as a function of luminosity. The current model cannot be used to study this term because a fixed number of events is generated to simulate the effect of adding minimum bias events on the *shape* of the multiplicity distribution. We do not know at the present time how to model how many extra minimum bias events contaminate our data sample (with single interaction cuts applied) as a function of luminosity.

6.6.2 Application to Data

The premise of the luminosity correction is that the first term in equation 6.4 is independent of luminosity. It was found this term still has some luminosity dependence when the region between the jet cone edges is used to count the multiplicity. When only the central calorimeter ($|\eta| < 1$) is used to count the multiplicity, there is no luminosity dependence in the first term of equation 6.4. Therefore, all calorimeter measurements of the multiplicity will be done by using the central calorimeter region. This term is shown in Fig. 6.17 for two different values of n_{sum}. In both cases, the result does not depend on luminosity.

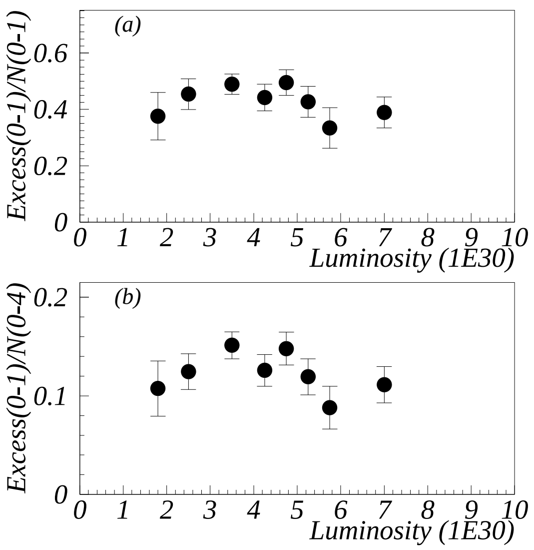

Figure 6.17: (a) $\frac{\sum_{n=0}^{n=1} data(n)-fit(n)}{\sum_{n=0}^{n=1} data(n)}$ and (b) $\frac{\sum_{n=0}^{n=1} data(n)-fit(n)}{\sum_{n=0}^{n=4} data(n)}$ vs. Luminosity. (First term in Eq. 6.4 for $n_{sum} = 1$ and $n_{sum} = 4$).

The second term in equation 6.4 contains the luminosity dependence. This term is plotted as a function of luminosity in Fig. 6.18 for two values of n_{sum}. A line is fit to the data and extrapolated to luminosity of zero. This extrapolation is verified by using data taken at very low luminosity. In Fig. 6.18, the data is

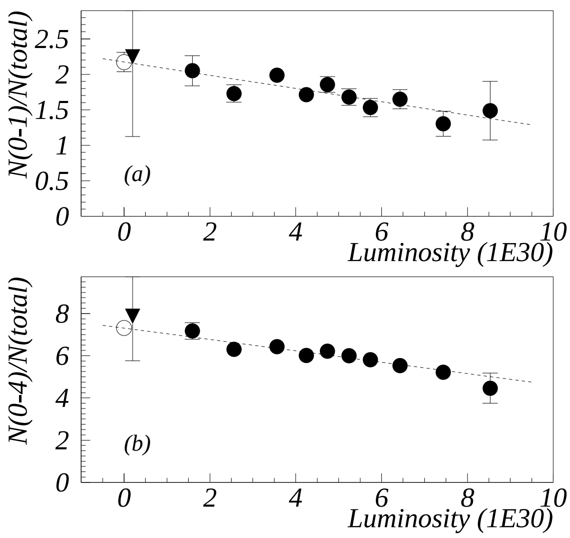

Figure 6.18: (a) $\frac{\sum_{n=0}^{n=1} data(n)}{N_{total}}$ and (b) $\frac{\sum_{n=0}^{n=4} data(n)}{N_{total}}$ vs. Luminosity. (Second term in Eq. 6.4 for $n_{sum} = 1$ and $n_{sum} = 4$).

shown in solid points; the point at luminosity near zero is from the low luminosity data. The open point is the extrapolation of the fit. The extrapolation is $2.17 \pm .14$ using $n_{sum} = 1$ and 7.31 ± 0.25 using $n_{sum} = 4$. The error on the extrapolation is relatively large when only two bins are used to count the multiplicity ($n_{sum} = 1$).

To determine if fitting a different function than a line to the luminosity dependent term gives a significantly different answer, we study the fit to:

$$\frac{\sum_{n=0}^{n=9} data(n)}{N_{total}} \tag{6.6}$$

Integrating to $n_{sum} = 9$ does not introduce significant luminosity dependence in the first term of equation 6.4, but increases the statistics to better determine the shape of the second term. Figure 6.19 (a) shows a linear fit to equation 6.6 and

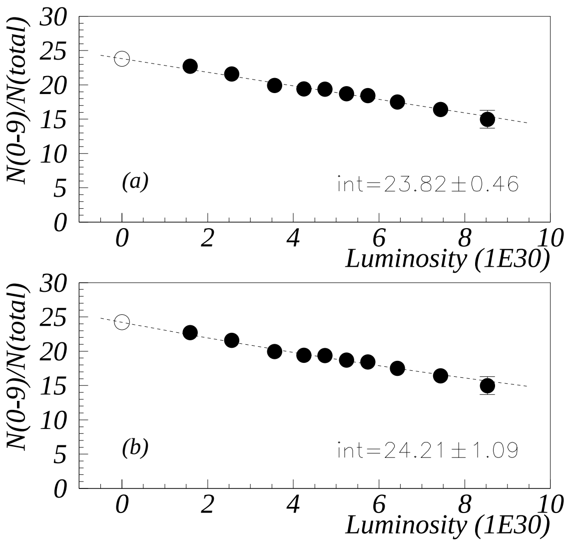

Figure 6.19: $\frac{\sum_{n=0}^{n=9} data(n)}{N_{total}}$ vs. Luminosity with (a) linear fit and (b) second order polynomial fit.

Fig. 6.19 (b) shows a second order polynomial fit. Both fits have a χ^2/df less than 1. The zero luminosity intercepts are 23.82 ± 0.46 for the linear fit and 24.21 ± 1.09 for the polynomial fit. These results do not indicate any significant deviation from a line at low luminosity, so we will use the linear fit to the data.

The color-singlet fraction is determined by combining the two terms. The luminosity corrected color-singlet fraction is given by Eq. 6.4 with $n_{sum} = 4$:

$$f_s = \frac{\sum_{n=0}^{n=n_0-1} data(n) - fit(n)}{\sum_{n=0}^{n=4} data(n)} \left(\frac{\sum_{n=0}^{n=4} data(n)}{N_{total}} \right)_{Lum=0} \tag{6.7}$$

It was shown with both the multiplicity model and the data that the first term is still uncontaminated when integrating to $n_{sum} = 4$. Therefore, we use this slightly larger multiplicity region to integrate over to decrease the error.

6.7 Vertex Correction

The single interaction cut requires a single vertex in the event. In the reconstruction code, if the Central Drift Chamber is used to find the vertex, a maximum of three vertices can be reconstructed. If the Forward Drift Chamber is used, only one vertex can be reconstructed. Rapidity gap events have zero to a few central tracks, so the Forward Drift Chamber will be used to find the vertex in most cases. Even when there are enough central tracks for the Central Drift Chamber to find the vertex, rapidity gap events still only have one vertex because the tracking reconstruction does not find spurious vertices in such a quiet environment.

The background events have enough central tracks that the Central Drift Chamber is almost always used to find the vertex. The vertex reconstruction can be confused by the relatively large track multiplicity in dijet events and reconstruct an additional vertex in a single interaction event. This inefficiency in finding a single vertex must be corrected for since it is only present in the background events. Figure 6.20(a) shows the fraction of events with single vertices for rapidity gap events. As expected, virtually all rapidity gap events have only one vertex. Figure 6.20(b) shows the single vertex event fraction for the full data sample. A second order polynomial is fit to the distribution for all events and extrapolated to zero luminosity. The correction factor is $C_{vtx} = 0.91 \pm .02$, meaning that the reconstruction misidentifies close to 10% of single interactions as having more than one

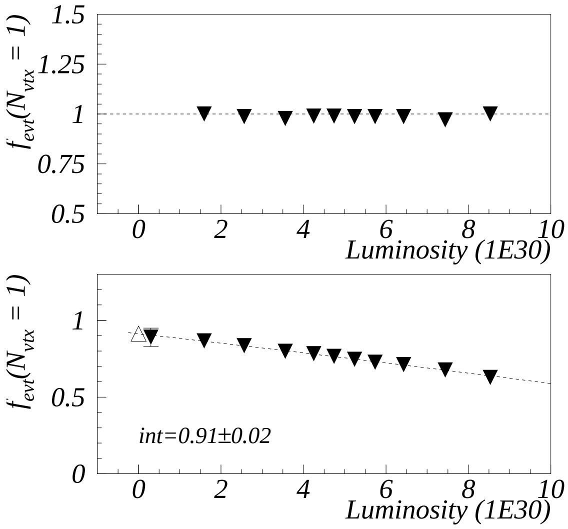

Figure 6.20: Fraction of single vertices versus Luminosity for (a) events with $0 \leq n_{cal} \leq 2$ (b) all events (filled triangles are high E_T data, open triangle is extrapolation of fit to zero luminosity).

vertex. This correction does not change as a function of the E_T or $\Delta\eta$ of the leading jets. With this correction factor, the color-singlet fraction becomes:

$$f_s = C_{vtx} \frac{\sum_{n=0}^{n=n_0-1} data(n) - fit(n)}{\sum_{n=0}^{n=4} data(n)} \left(\frac{\sum_{n=0}^{n=4} data(n)}{N_{total}} \right)_{Lum=0} \tag{6.8}$$

Chapter 7

Results

This chapter gives results on the measurement of color-singlet exchange fraction and its dependence on jet E_T and $\Delta\eta$. The systematic error on the measurement is estimated by considering each source of possible error. Additional systematic errors on the measurement of the color-singlet fraction as a function of jet E_T and $\Delta\eta$ are also discussed.

7.1 Measurement of the Color-Singlet Fraction

The color-singlet fraction is measured using the central calorimeter multiplicity ($|\eta| < 1$) for both jets with $E_T > 30$ GeV. The measured color-singlet fraction is $f_s = (0.85 \pm 0.05(stat) \pm 0.07(syst))\%$. This measurement and all subsequent measurements of the color-singlet fraction have the vertex and luminosity corrections applied as given by Equation 6.8. Figure 7.1 shows the central calorimeter multiplicity with the leading edge fit used to measure f_s. The distribution is fit from $n_0 = 2$ to $n = 14$. The χ^2/df and confidence limit for starting the fit at varying values of n_0 is shown in Table 7.1. Using the fitting studies from Section

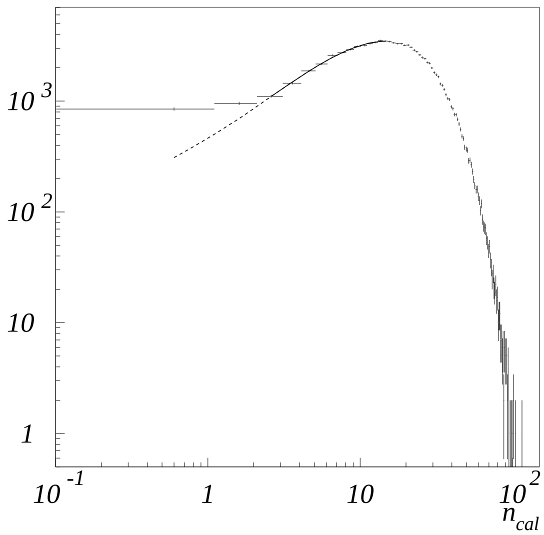

Figure 7.1: Central Calorimeter multiplicity with single Negative Binomial Distribution fit.

n_0	χ^2/df	CL(%)	f_s(%)
0	17.6	$< 10^{-6}$	—
1	4.6	$< 10^{-6}$	0.48 ± 0.04
2	0.93	52.2	0.85 ± 0.05
3	1.02	43.7	0.88 ± 0.06
4	0.88	54.5	0.63 ± 0.07

Table 7.1: χ^2 of Fit for different values of n_0

6.4.1 as guidance leads us to choose $n_0 = 2$ as the starting bin of the fit. The following sections detail the sources of systematic error which contribute to this measurement.

7.1.1 Systematic Error from Cuts

The following systematic studies are done for the high E_T trigger. Each cut is varied in a sensible range and the color-singlet fraction is re-measured. The maximum difference between the new color-singlet fraction and the nominal value is taken as the systematic error on that cut. The color-singlet fraction for each cut variation is shown in Table 7.2. All of the systematic errors are added in quadrature to obtain the final systematic error.

Energy Scale

The jet energy scale (CAFIX 5.0) is varied by $\pm 1\sigma$ and the two leading jets are required to have E_T greater than the high or low corrected value. The estimated systematic error is 2.5%. The central value of the correction in the newer jet energy scale (CAFIX 5.1) is within the errors of the old energy scale. Since varying the energy scale correction within its errors has such a small effect on the measured fraction, we use CAFIX 5.0 for the energy scale corrections.

Jet Quality Cuts

The QCD jet quality cuts were varied by (1) dropping the cut and (2) requiring all jets in the event satisfy the cut. The systematic error estimate is 2.2%.

Jet η cut

The leading jets were required to have $|\eta| > 1.9$ (instead of $|\eta| > 1.7$). The estimated systematic error is 1.9%.

Vertex cut

The cut on $|z_{vertex}| < 50$ cm was relaxed to 100 cm. The estimated contribution to the systematic error is 4.1%.

Nominal	High Escale	Low Escale	No goodjet	All goodjet	$\eta > 1.9$	z_{vtx}
0.849	0.833	0.828	0.853	0.868	0.833	0.884

Table 7.2: Color-singlet fraction measured for variations of systematic cuts

Other systematic errors

In addition to the above sources of systematic error, we assign a systematic error due to the fit and due to the vertex correction. The vertex correction factor was $C_{vtx} = 0.91 \pm 0.02$. This 2% error is included in the systematic error. The fit error, which is the error on the background estimation, has two components. The first component is determined by using the rms of the distribution of fractions found by fitting the ensemble of data distributions (note that the fraction in the fit error table does not have the luminosity or vertex correction applied). This error depends on the statistics of the sample. The error is taken as the rms of the distribution of fit fractions. For the full sample, it is a 4% error. The second component comes from raising the starting bin of the fit by one bin and re-fitting the ensemble of data distributions. This error is the difference between the mean fraction found from starting the fit at n_0+1 as opposed to n_0. The two components of the fit error are combined in quadrature to obtain the final fit error. Table 7.3 shows the fit errors for different size samples which correspond roughly to the statistics in different $\Delta\eta$ and E_T bins. The error which arises from the luminosity correction is included in the statistical error.

Statistics	$< f >$	rms	% error	$< f >_{n_0+1}$	% error	total error (%)
100K	0.791	0.061	3.9	0.835	5	6
74K	0.795	0.076	4.8	0.828	4	6
30K	0.797	0.123	7.7	0.868	9	12
25K	0.795	0.135	8.5	0.862	7	12
20K	0.796	0.152	9.5	0.850	7	12
15K	0.793	0.172	10.8	0.851	7	13
10K	0.782	0.208	13.3	0.802	3	14
6K	0.776	0.254	16.4	0.736	5	17

Table 7.3: Fit errors for different statistical samples

7.1.2 Cross Checks

The following sections detail various systematic checks. These checks are not included in the quoted systematic error; they are used to ensure that the measurement is robust.

Tower Threshold

The calorimeter multiplicity is defined as the number of towers with $E_T > 200$ MeV. This threshold has been varied between 150 and 350 MeV. The resulting color-singlet fraction is shown in Figure 7.2. Although there is some variation of the signal with the tower threshold, there is not a systematic trend.

Particle Tagging

The color-singlet fraction has also been measured by counting the number of charged tracks in the Central Drift Chamber. Figure 7.3 shows the multiplic-

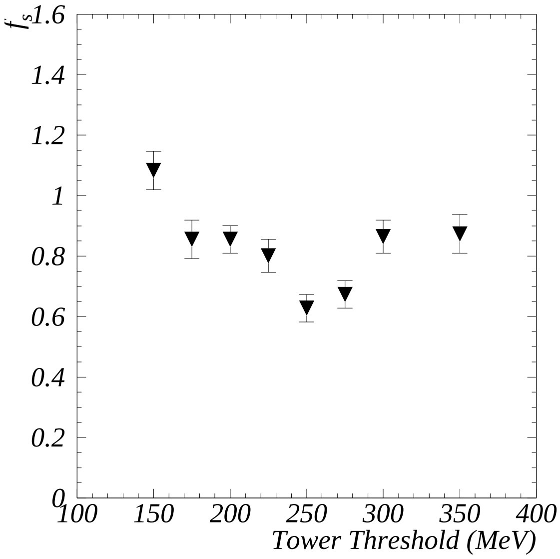

Figure 7.2: Color-singlet fraction versus calorimeter tower threshold.

ity distribution for the number of tracks in the region $|\eta| < 1$. The excess is $(0.85 \pm 0.06)\%$, which is consistent with the calorimeter result.

Another check on the particle tagging method is to make the calorimeter more coarsely divided by clustering adjacent calorimeter towers. This clustering may be closer to the true particle multiplicity because one particle can shower into more than one calorimeter tower. Towers were clustered in a radius $\mathcal{R} = 0.18$ in η-ϕ space. This radius corresponds to clustering towers in a 3×3 array. The fractional excess is $(0.88 \pm 0.05)\%$, which is in good agreement with the nominal value. Because varying the definition of the particle tagging does not significantly affect the measured color-singlet fraction, we do not assign a systematic error due to the tagging method.

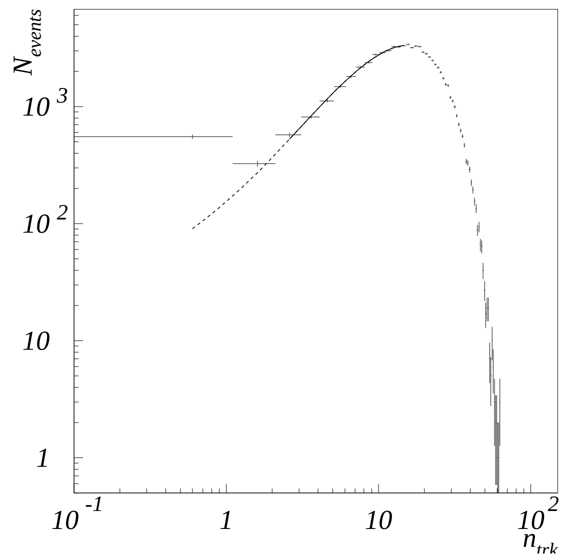

Figure 7.3: Multiplicity of charged tracks in $|\eta| < 1$.

Trigger Biases in ϕ

The data characteristics plots in Section 4.3.5 show that the azimuthal distribution of the leading jets has inefficiencies at the large tile boundaries. These dips could lead to a more back-to-back jet configuration, thus artificially raising the measured color-singlet fraction. This effect has been investigated by both exaggerating the dips and by making the jet ϕ distribution flat. The color-singlet fraction rises to 0.96% when the dips are highly exaggerated, indicating that there is likely some bias. However, $f_s = 0.87\%$ when the ϕ distribution is flattened. Since the nominal color-singlet fraction does not change significantly when the ϕ dips are removed, no systematic error is assigned.

7.1.3 Comparison to Published Result

The current measurement of the color-singlet fraction should be compared to the fraction measured with the central calorimeter in the published multiplicity analysis. The published data was taken during the 1992-93 Tevatron run, referred to as Run 1a. The current measurement is based on 1994-95 data, which will be referred to as Run 1b in the discussion which follows. The published fraction was $f_{s(1a)} = (1.14 \pm 0.11(\text{stat})^{+0.27}_{-0.14}(\text{syst}))\%$. The complete error bar for the 1a and 1b results, $f_{s(1a)} = 1.14^{+.29}_{-.18}$ and $f_{s(1b)} = 0.85 \pm 0.09$ touch at about the one sigma level. There are several factors which could contribute to the discrepancy other than simply a fluctuation.

One major change between the 1a and 1b analyses is the fitting method. A Double Negative Binomial Distribution was fit to the 1a data while a Single Negative Binomial Distribution is fit to the leading edge of the 1b data. Using superior statistics and better modelling, we have shown that the leading edge fit gives a more stable and accurate measure of the fraction. When a leading edge fit is applied to the 1a central calorimeter multiplicity distribution, the fraction drops to 0.9%. Although this one factor seems to explain the discrepancy, when a leading edge fit is applied to the 1a multiplicity between jet cone edges, the fraction rises from 1.07% to 1.2%. The fact that one fit goes up and one goes down could be a fluctuation in the fitting. The 1a statistics were lower than those in 1b, so the fitting error is larger.

Another factor to consider is the difference in the triggers. The 1a trigger used trigger towers to trigger on the jets. The 1b trigger used large tiles. The turn-on of the E_T distribution is slower for trigger towers than for large tiles. This could lead to two possible effects. The first is that both jets were still required to have $E_T > 30$ GeV in the 1a analysis, but the trigger was very inefficient at that E_T. The bias from measuring the fraction where the trigger is very inefficient is not well known, but from the 1b data seems to cause an overestimate of the color-singlet fraction. We have verified that the 1b result does not have a significant trigger bias by using the medium E_T trigger (where it is efficient) to measure the fraction at the same E_T as the high E_T trigger (see Section 6). A second related effect is that the average E_T for the 1a sample is higher than that for the 1b sample. The current measurement of the fraction versus E_T indicates a slight rise with E_T (Section 6).

Any of these several factors or simply a statistical fluctuation can explain the one sigma difference between the current and the published measurement. The large statistical sample and improved fitting methods used in the current measurement put it on very solid ground.

7.1.4 Interpretation of f_s

The measured color-singlet fraction is more than ten times larger than the expected fraction due to Electroweak exchange. This measurement can be used to exclude Electroweak exchange plus color-exchange as the source of the observed color-singlet fraction. The expected fraction of dijet events from Electroweak exchange is determined to be 0.09% from a PYTHIA [34] study which simulates DØ acceptance and efficiency [14]. To exclude Electroweak exchange, we assume 100% survival probability to give the maximum number of observable events. The expected number of color-exchange events is determined from the fit. The number of events with zero calorimeter towers is compared to the number of expected events from Electroweak exchange and the number of expected color-exchange events. The error on the total number of events is a combination of the fit error and the statistical error. The number of standard deviations for the total expected number of events to fluctuate to the number of observed events is given in Table 7.4. The probability that the observed number of events with zero multiplicity is due to Electroweak exchange plus color-exchange background is less than 10^{-10}. Therefore, the measured color-singlet fraction is consistent with strongly-interacting color-singlet exchange.

	$N_{expected}$			
$N_{observed}$	EW	Color-exch.	Total	Number of s.d.
849	71.4	309.1	380.5 ± 26.9	17σ

Table 7.4: Exclusion of Electroweak Exchange

Jet Multiplicity Effects

The amount of radiation in an event has a direct effect on the color-singlet fraction. If the amount of radiation is small, there is less probability to produce particles in the region between the jets, thus raising the gap fraction. Although our data sample tends to contain back-to-back jets, there are more events with three or more jets in our data than two jet events. Figure 7.4 shows the multiplicity distribution for all events, two jet events, and events with three or more jets. Examining Fig. 7.4 at low multiplicity reveals that most of the signal events have only two jets. Therefore, the events with more than two jets contribute largely to only the denominator, N_{total}, of f_s. The color-singlet fraction for two jet events is 2.5%, while the fraction for multijet events is 0.1%.

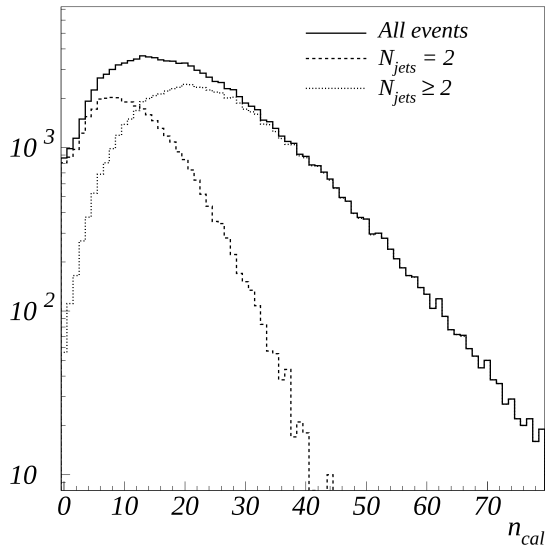

Figure 7.4: Central Calorimeter multiplicity for different number of jets.

The cuts we place on the leading jets (both jets with $E_T > E_T^{min}$) biases the configuration toward more back-to-back two jet events. Since events with only two jets have a higher color-singlet fraction, this cut biases the measured fraction high. If the measurement were done in a more inclusive manner, i.e. one jet with $E_T > 30$ GeV and the second jet with $E_T > 8$ GeV, the color-singlet fraction would be lower. This bias is not a large concern because similar cuts can be placed on the theory. It should be noted, however, that the color-singlet fraction we measure is not the absolute color-singlet exchange fraction, but, rather, the color-singlet fraction for events with a certain topology.

7.2 Color-Singlet Fraction versus jet E_T

The measurement of the fractional excess as a function of the E_T of the leading jets uses all three data samples. It should be noted that the low E_T data used for this measurement is still preliminary. The final data will be discussed and presented in Jill Perkins' thesis [15]. Figure 7.5(a) shows the multiplicity in the central calorimeter ($|\eta| < 1$) for the three data samples. All three distributions look similar in shape, but there is some discrepancy at very low multiplicities. Focusing on the low multiplicity region (Fig. 7.5(b)) shows that the high E_T data has more low multiplicity events than the lower E_T data, which in turn, has more low multiplicity events than the lowest E_T data.

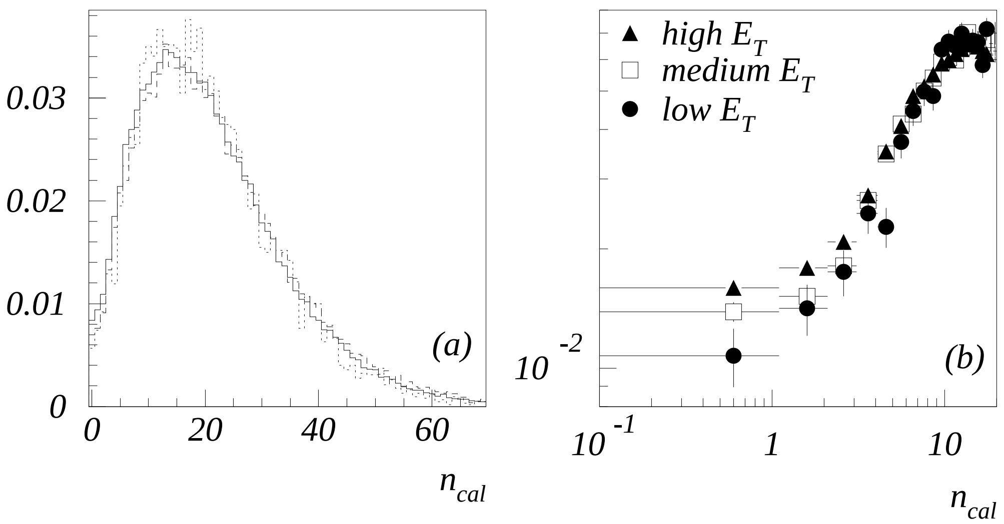

Figure 7.5: Central Calorimeter multiplicity for three E_T samples: (a) Solid line is high E_T;dashed line is medium E_T; dotted line is low E_T (b) log-log scale emphasizes low multiplicity region.

Figure 7.6 shows the measured color-singlet fraction as a function of E_T using the central calorimeter multiplicity and the leading edge fitting method. The inner error bars are statistical and the outer error bars are the systematic plus statistical error added in quadrature. The data is binned as a function of the second leading jet E_T and the points are plotted at the average E_T of the two leading jets. This method of binning the color-singlet fraction eliminates possible acceptance biases which result from triggering on dijets (see Appendix B for details). The color-singlet fraction from the medium and high E_T triggers agrees in the region of overlap. There is not enough data available in the low E_T sample to compare the low E_T and the medium E_T triggers. Figure 7.7 shows the measured color-singlet fraction using the N_{00} method as a function of E_T. The normalization error from normalizing the N_{00} fraction back to the fit fraction is shown in the shaded band for each trigger. Note that the range of the measurement is significantly extended by using this method. There is a slight rise of the fraction as the E_T of the jets is increased.

E_T bin (GeV)	$< E_T >$	$< \|\eta\| >$	f_s fit	f_s N_{00}
$15-25$	21.0	2.41	0.60 ± 0.22	0.60 ± 0.14
$25-30$	29.8	2.39	0.90 ± 0.18	0.87 ± 0.11
$30-35$	35.1	2.35	0.88 ± 0.13	0.76 ± 0.07
$35-40$	40.3	2.32	0.55 ± 0.14	$0.80^{+0.10}_{-0.11}$
> 40	49.9	2.28	1.05 ± 0.22	—
$40-45$	45.5	2.29	—	$1.09^{+0.18}_{-0.30}$
$45-50$	50.7	2.28	—	1.15 ± 0.28
$50-60$	57.4	2.25	—	$1.19^{+0.35}_{-0.37}$
> 60	70.5	2.22	—	$1.21^{+0.61}_{-0.68}$

Table 7.5: f_s versus jet E_T for the fit method and the N_{00} method

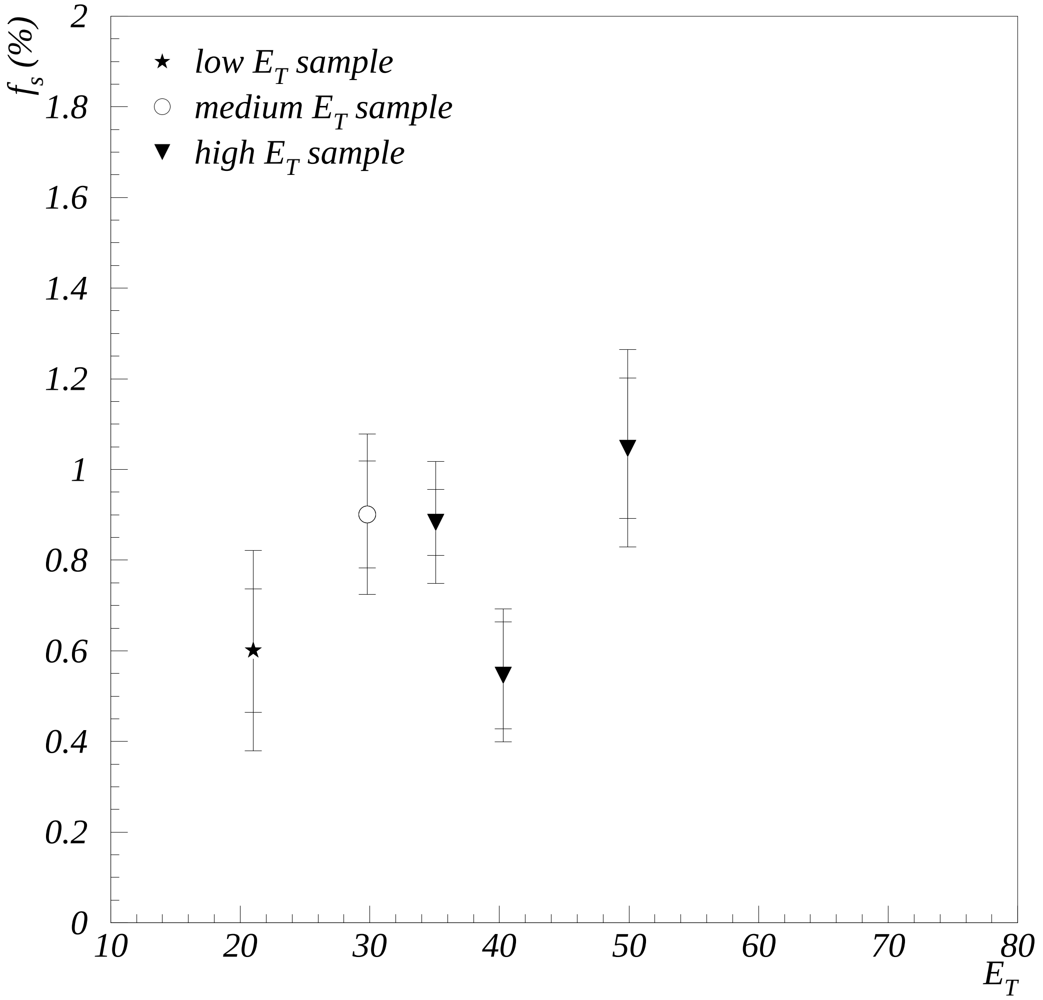

Figure 7.6: f_s as a function of E_T: fit method. Inner error bars are statistical; outer error bars are statistical plus systematic added in quadrature.

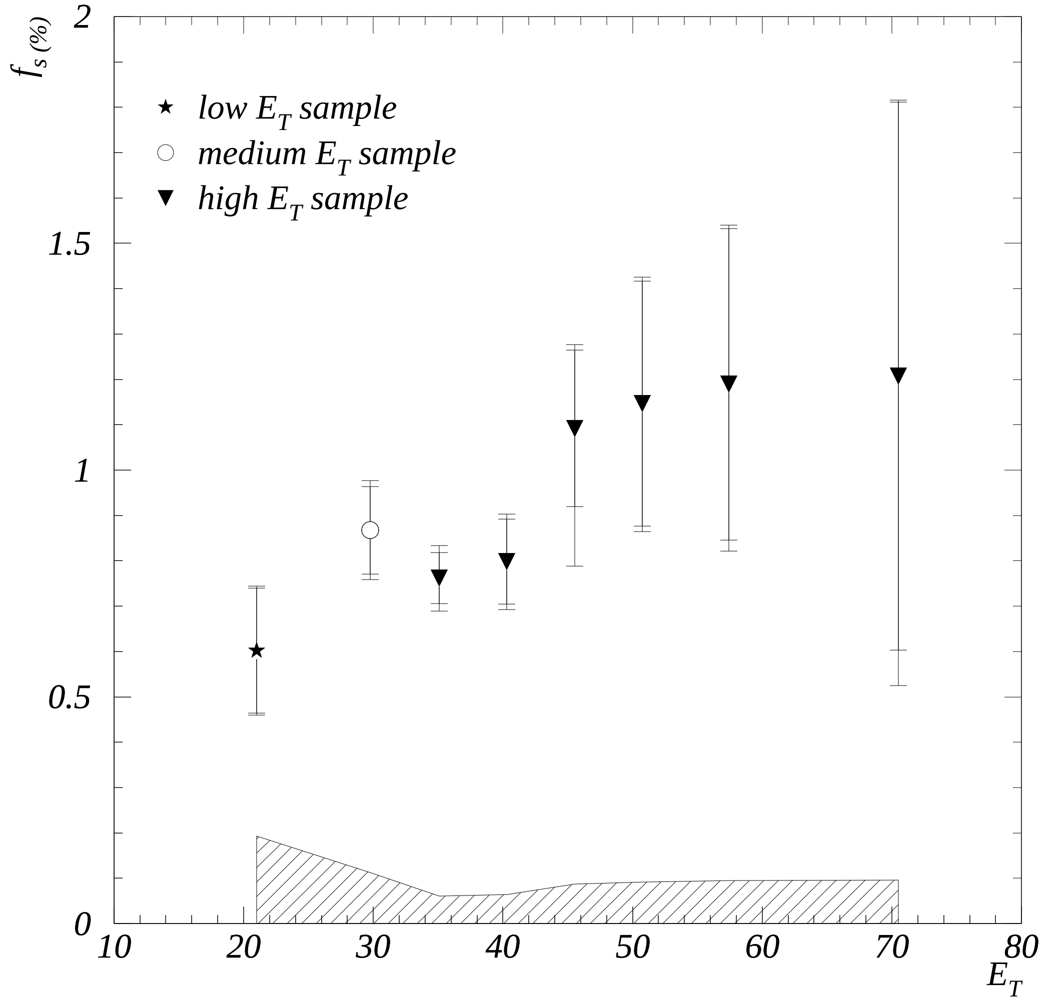

Figure 7.7: f_s as a function of E_T: N_{00} method. Inner error bars are statistical; outer error bars are statistical plus systematic added in quadrature. The shaded band is the normalization error which results from normalizing the fraction of events with zero towers and zero tracks to the color-singlet fraction found from fitting.

7.2.1 Systematic Errors for the N_{00} Method

Using the N_{00} method to extend the range of the measurement of the color-singlet fraction as a function of jet E_T and $\Delta\eta$ introduces additional systematic errors. One additional error is the normalization error. This error arises from normalizing the fraction of events with zero towers and zero tracks back to the color-singlet fraction found with the fit method and is merely the statistical error which results from multiplying the various factors. The normalization error is shown as a hatched band on the plots of f_s versus E_T and $\Delta\eta$.

Another source of systematic error comes from the N_{00} method itself. The fraction of events with zero towers and zero tracks is measured as a function of jet E_T and $\Delta\eta$. Each E_T sample (low, medium, and high E_T) has a separate factor to normalize the N_{00} fraction to the f_s found by fitting. Applying one normalization factor for a large range of E_T and $\Delta\eta$ assumes that the fraction of color-singlet exchange events in the (0,0) bin does not vary as a function of E_T and $\Delta\eta$. It is possible, however, that as E_T and $\Delta\eta$ change, more or fewer color-singlet exchange events may populate bins other than the (0,0) bin. The effect of the color-singlet exchange events moving out of the (0,0) bin would be to introduce a shape in the measured N_{00} fraction as a function of E_T or $\Delta\eta$. If the color-singlet exchange events shift from the (0,0) bin, it is unlikely that they will shift more than one bin in multiplicity. A systematic error is determined by examining the fraction of events with one or fewer calorimeter towers or tracks (the (0,0), (0,1), (1,0), and (1,1) bins added) as a function of E_T and $\Delta\eta$. This measurement is less susceptible to the shape of the color-singlet signal changing, but has a higher color-exchange background of about 15%. The difference in the central value between the fraction of events with one or fewer towers or tracks and events with zero towers and tracks is taken as an additional systematic error on the measurement of N_{00} as a function of E_T and $\Delta\eta$. The maximum percentage change for a single point is about 25% for both the E_T and $\Delta\eta$ measurements. Table 7.6 gives the difference in shape for each bin of the $\Delta\eta$ and E_T measurement. This systematic error is added in quadrature with the systematic error obtained in Section 7.1.1 to obtain the total systematic error on each point. Since this systematic consistently flattens the E_T and $\Delta\eta$ dependence, the systematic error is applied only to the lower side of the total error.

E_T bin	% error	$\Delta\eta$ bin	%error
$15-25$	0.	$4-4.2$	7.5
$25-30$	0.	$4.2-4.3$	6.8
$30-35$	0.6	$4.3-4.4$	5.6
$35-40$	1.7	$4.4-4.5$	2.0
$40-45$	22.0	$4.5-4.65$	12.3
$45-50$	3.7	$4.65-4.8$	5.3
$50-60$	9.2	$4.8-5.0$	15.1
>60	25.7	$5.0-5.2$	16.3
		$5.2-5.5$	16.7
		$5.5-5.9$	26.0
		>5.9	21.1

Table 7.6: Systematic error on shape of N_{00} versus E_T and $\Delta\eta$.

Resolution Effects as a Function of E_T

An additional possible systematic error in the measurement of the color-singlet fraction as a function of E_T is due to jet resolution. The E_T^{min} cut imposed on both jets means that an E_T-dependent resolution could affect the measured color-singlet fraction. If the resolution is higher at high E_T than at low E_T, the effective E_T^{min} threshold changes; thus affecting the measured color-singlet fraction as a function of E_T. This effect has been investigated by imposing a shift in the second leading jet E_T^{min} relative to the leading jet E_T^{min}in various ways. Each method investigated shows that a reasonable E_T-dependent shift in the jet resolution does not change the shape of the measured color-singlet fraction as a function of E_T.

7.2.2 Interpretation of f_s versus E_T

Both methods for measuring f_s as a function of the dijet E_T show a slight rise with increasing E_T. Linear fits are performed on Fig. 7.7 to determine how f_s changes with E_T. Both a line with a slope and a line with zero slope are good fits to the data. The zero slope fit is $f_s = 0.80 \pm 0.05$, and has a $\chi^2/df = 6.5/7$, which gives a confidence level for the fit of 50%. When the fit is allowed to have a slope, it becomes $f_s = (0.38 \pm 0.22) + (0.12 \pm 0.06)E_T$. This fit has a χ^2/df of 2.8/6, which gives a confidence level of 83%. A line with a slope seems to fit the data best, but a model which predicts a flat dependence on E_T cannot be ruled out with this data. This behavior can be compared to the expectations from the different color-singlet models. Recall, however, that all of the models are at leading order and next-to-leading order effects could significantly alter their predictions.

The simple two gluon model and the BFKL calculation both predict that the color-singlet fraction falls as a function of E_T. The fall in the BFKL calculation is very steep at low E_T due to the dynamics present in the model. The soft color model predicts either a flat or rising color-singlet fraction with increasing E_T. This behavior is due entirely to changes in the initial parton distribution functions as a function of E_T. The U(1) Gauge Boson model also predicts a rise in the color-singlet fraction as a function of E_T which is due to dynamics in the model as well as changes in the initial parton distribution functions.

Table 7.7 shows how the leading order parton x changes in the data as a function of E_T. This measurement alone is inconsistent with the two gluon models. It is consistent with the soft color model and possibly consistent with the U(1) Gauge Boson model. In the U(1) Gauge boson model, if the mass m_B is taken to be 30 GeV and the coupling α_B is taken to be 0.1 and is roughly independent of E_T in the relevant range, then the model predicts a rise in the color-singlet fraction of about a factor of seven between 20 and 70 GeV. However, if the mass is taken to be $m_B = 20$ GeV, the rise in the color-singlet fraction in the same E_T range is only a factor of three. Therefore, the U(1) Gauge boson model is only consistent with our data with certain parameter combinations. When the final low E_T data is ready, the dependence of the color-singlet fraction on jet E_T can be used to place limits on the allowable parameters of the U(1) Gauge Boson model.

E_T bin (GeV)	$< E_T >$	$< \|\eta\| >$	x	$\%Q$
$15-25$	21.0	2.41	0.130	60
$25-30$	29.8	2.39	0.181	68
$30-35$	35.1	2.35	0.204	71
$35-40$	40.3	2.32	0.228	74
$40-45$	45.5	2.29	0.250	77
$45-50$	50.7	2.28	0.275	79
$50-60$	57.4	2.25	0.303	81
> 60	70.5	2.22	0.361	86

Table 7.7: For each E_T bin, the average E_T, η, parton x, and the percent of initial quarks are given.

Kinematic Biases in the Measurement of f_s versus E_T

Because of the E_T^{min} cut on the dijets, the shape of the measured color-singlet fraction as a function of E_T could be biased. This bias in the shape may arise due to the method used to bin the dijet E_T. It arises from the fact that the dijets are closely balanced in E_T when both jets are near the triggered E_T^{min} threshold, but may have a large ΔE_T when one jet's E_T is much greater than the E_T^{min} threshold. We eliminate this potential bias by binning the second leading jet E_T: since the jets are E_T-ordered, this means that there is an E_T^{min} cut on the dijets in every bin. Another possible bias is introduced by the configuration of the events we study: the color-singlet exchange events consist of largely back-to-back two jet events with a small η_{boost} and little additional radiation. The binning bias is eliminated by the binning method we choose. The configuration bias is small - at most a 20% effect between E_T of 30 and 70 GeV. These biases are not a large concern because they can be modelled in the theory. More detail is given in Appendix B.

7.3 Color-Singlet Fraction versus $\Delta\eta$

The color-singlet fraction as a function of the separation of the dijets, $\Delta\eta$, is measured for the high E_T data sample. Figure 7.8 shows the color-singlet fraction found with the fitting method as a function of $\Delta\eta$. The inner error bars are statistical and the outer error bars are statistical plus systematic added in quadrature. The error bars become quite large at high $\Delta\eta$ due to the limited statistics available to fit. Using the N_{00} method extends the measurement to higher $\Delta\eta$. Figure 7.9 shows the fraction using the N_{00} method as a function of $\Delta\eta$. The fraction found with the fitting method appears flat as a function of $\Delta\eta$ while the fraction from the N_{00} method rises slightly as $\Delta\eta$ increases.

$\Delta\eta$ bin	$<\Delta\eta>$	$<E_T>$	$<n_{cal}>$	f_s fit	f_s N_{00}
$4-4.2$	4.12	40.1	21.4	1.03 ± 0.19	$0.76^{+0.10}_{-0.11}$
$4.2-4.3$	4.25	39.5	21.4	0.78 ± 0.18	$0.68^{+0.11}_{-0.12}$
$4.3-4.4$	4.35	39.4	21.4	0.86 ± 0.21	0.70 ± 0.12
$4.4-4.5$	4.45	39.1	21.1	0.64 ± 0.17	0.60 ± 0.11
$4.5-4.65$	4.57	38.7	21.0	0.76 ± 0.17	$0.74^{+0.11}_{-0.14}$
$4.65-4.8$	4.72	38.4	20.9	0.66 ± 0.18	0.81 ± 0.13
$4.8-5.0$	4.89	38.0	20.7	0.65 ± 0.18	$0.89^{+0.13}_{-0.19}$
>5.0	5.29	37.0	19.8	1.09 ± 0.23	—
$5.0-5.2$	5.09	37.6	20.3	—	$1.11^{+0.18}_{-0.26}$
$5.2-5.5$	5.33	37.0	19.4	—	$1.17^{+0.22}_{-0.29}$
$5.5-5.9$	5.65	36.0	19.3	—	$1.21^{+0.33}_{-0.46}$
>5.9	6.06	34.6	18.1	—	$1.74^{+0.88}_{-0.95}$

Table 7.8: f_s versus $\Delta\eta$ for the fit method and the N_{00} method

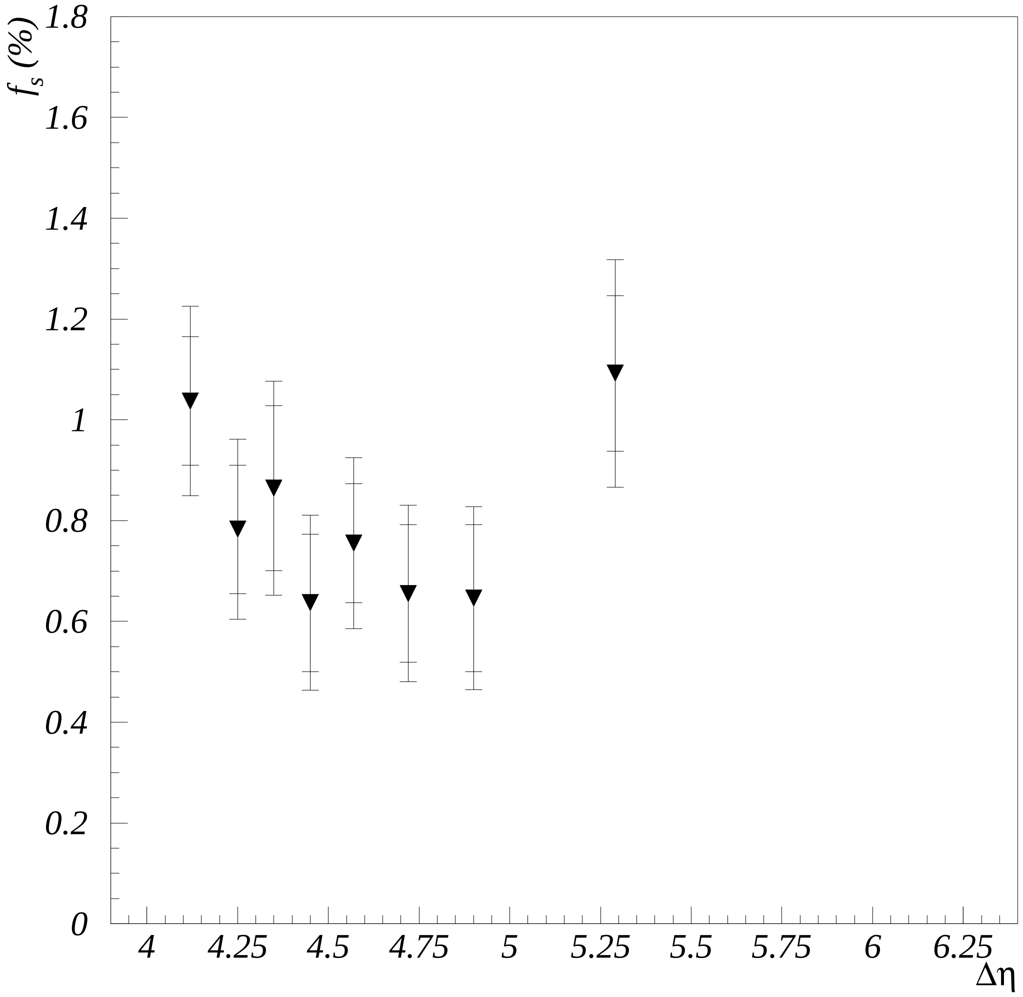

Figure 7.8: f_s as a function of $\Delta\eta$: fit method. Inner error bars are statistical; outer error bars are statistical plus systematic added in quadrature.

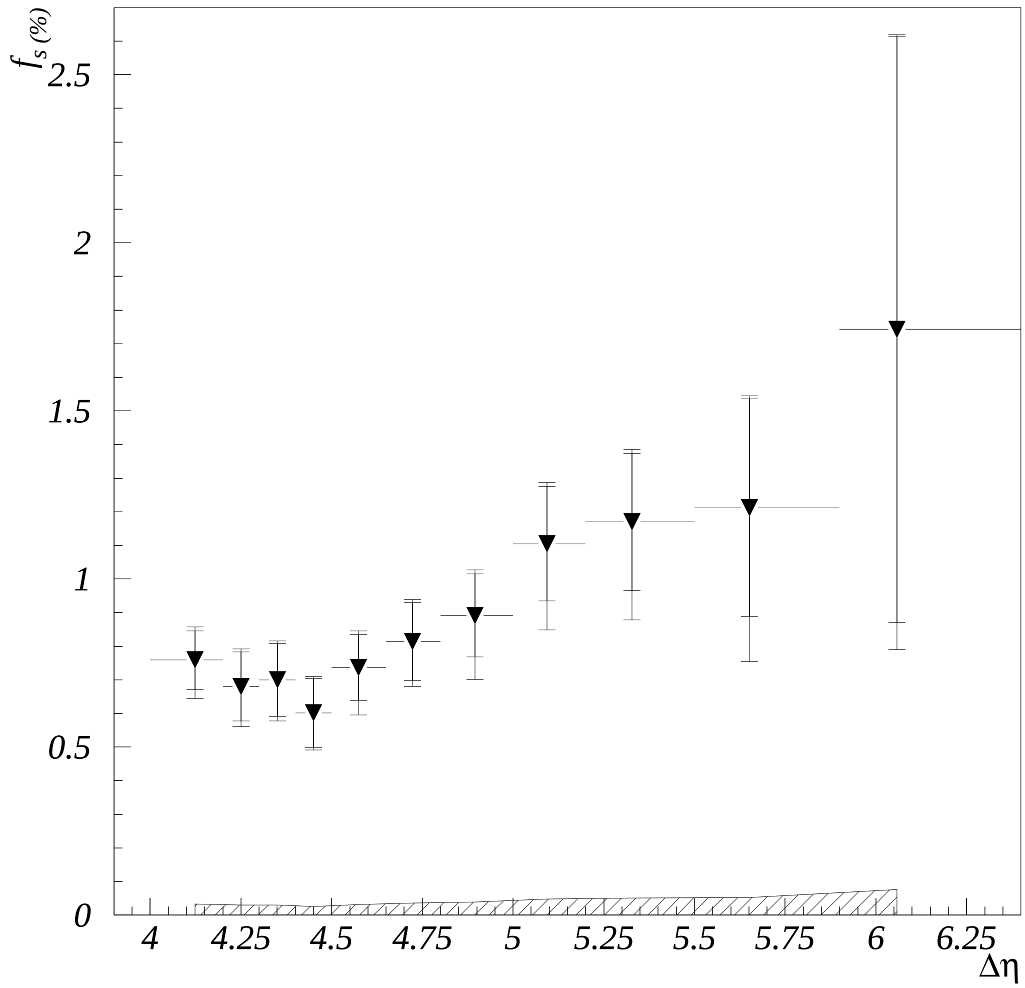

Figure 7.9: f_s as a function of $\Delta\eta$: N_{00} method. Inner error bars are statistical; outer error bars are statistical plus systematic added in quadrature. The shaded band is the normalization error which results from normalizing the fraction of events with zero towers and zero tracks to the color-singlet fraction found from fitting.

7.3.1 Interpretation of f_s versus $\Delta\eta$

The measured color-singlet fraction as a function of $\Delta\eta$ is approximately flat for the fit method and shows a slight rise when extended with the N_{00} method. As with the E_T measurement, linear fits are performed on Fig. 7.9 to determine how f_s changes with $\Delta\eta$. Again, both a line with a slope and a line with zero slope are good fits to the data. When the slope is constrained to be zero, the resulting is fit $f_s = 0.75 \pm 0.05$, and has a $\chi^2/df = 9.2/10$, which gives a confidence level for the fit of 53%. When the fit is allowed to have a slope, it becomes $f_s = (-0.78 \pm 0.66) + (0.34 \pm 0.15)\Delta\eta$. This fit has a χ^2/df of 3.8/9, which gives a confidence level of 91%. As was the case for the color-singlet fraction as a function of E_T, a line with a slope seems to fit the data best, but a model which predicts a flat dependence on $\Delta\eta$ cannot be ruled out with this measurement.

This behavior can also be compared to the different leading order color-singlet models. The simple two gluon model predicts a falling color-singlet fraction with $\Delta\eta$ due entirely to the changes in the initial parton distribution functions. The BFKL two gluon calculation predicts a rise in f_s at high, but unspecified, $\Delta\eta$. The soft color model predicts either flat or rising color-singlet fraction as a function of $\Delta\eta$ while the U(1) Gauge Boson model predicts a rise. The behavior of both of these models as a function of $\Delta\eta$ is due entirely to how they react to changes in the initial parton distribution functions.

Table 7.9 shows the leading order calculation for the parton x as a function of $\Delta\eta$ for our data. This measurement of f_s as a function of $\Delta\eta$ is consistent with the BFKL calculation, the soft color model, and the U(1) Gauge Boson model. It is inconsistent with the simple two gluon model.

$\Delta\eta$ bin	$<\lvert\eta\rvert>$	$<E_T>$	x	$\%Q$
$4-4.2$	2.06	40.1	0.175	68
$4.2-4.3$	2.13	39.5	0.185	70
$4.3-4.4$	2.18	39.4	0.194	71
$4.4-4.5$	2.23	39.1	0.202	72
$4.5-4.65$	2.29	38.7	0.212	73
$4.65-4.8$	2.36	38.4	0.226	74
$4.8-5.0$	2.45	38.0	0.245	76
$5.0-5.2$	2.55	37.6	0.268	78
$5.2-5.5$	2.67	37.0	0.297	80
$5.5-5.9$	2.83	36.0	0.339	83
>5.9	3.03	34.6	0.398	86

Table 7.9: For each $\Delta\eta$ bin, average η, average E_T, parton x, and the percent of initial quarks are given.

7.4 f_s for Unbalanced Jet Configuration

An additional interesting measurement is to determine the color-singlet fraction for dijet events in which the jets have different E_T^{min} cuts. The high E_T dijet trigger had an E_T^{min} cut of 25 GeV on both jets. This cut prevents a measurement of f_s for which one jet has E_T lower than 25 GeV. There are sufficient statistics, however, to make the measurement for one jet with $E_T > 55$ GeV and the second jet with $E_T > 25$ GeV. This configuration should give a different result than the nominal result for both jets with $E_T > 30$ GeV since the unbalanced E_T^{min} cut will allow more radiation in the event.

To make this measurement, we require two jets with $|\eta| > 1.9$, one jet with $E_T > 55$ GeV, and one jet with $E_T > 25$ GeV. The color-singlet fraction for this configuration is $f_s = (0.66 \pm 0.19)\%$. In contrast, when both jets have $E_T > 50$ GeV (the 50-60 GeV bin of Table 7.5), the color-singlet fraction is $f_s = (1.19^{+0.35}_{-0.37})\%$. Although the errors are large at this high E_T, the color-singlet fraction for the unbalanced configuration is clearly lower than for the balanced E_T^{min} cut.

This measurement cannot be compared to any of the existing color-singlet models. The possibility of jets which are unbalanced in E_T is not present in any of the models because they are all at leading order. However, it is instructive to examine the effect of additional radiation. This measurement may be perhaps compared to future color-singlet models.

Chapter 8

Conclusions

Strongly-interacting color-singlet exchange has been observed at a rate consistent with Bjorken's original prediction. However, the prediction he made was only for quark-quark initiated processes. While quark-quark processes dominate for high E_T and large $\Delta\eta$, in the two gluon model, the gap fraction due to quark-gluon and gluon-gluon processes is higher. Measurements in which the initial quark and gluon fraction vary help discriminate between competing color-singlet models. This thesis presents two of those measurements: the color-singlet fraction as a function of jet E_T and $\Delta\eta$. A third measurement of the ratio of color-singlet fractions at $\sqrt{s}$ = 630 and 1800 GeV (Jill Perkins' thesis [15]) completes the picture in the data.

Comparing the behavior of the measured color-singlet fraction as a function of E_T and $\Delta\eta$ to the existing color-singlet models shows that the two gluon models (both the simple model and the BKFL extension) are inconsistent with the data. The soft color and U(1) Gauge Boson models are qualitatively consistent with the measurement of f_s as a function of E_T and $\Delta\eta$. These measurements combined with the measurement of $R(630)/(1800)$ are a large step in the understanding of strongly-interacting color-singlet exchange. The general consensus in the high energy physics community is that the two gluon picture is the most likely scenario for strongly-interacting color-singlet exchange. Our measurements show that this picture, at least at leading order, is not correct.

Which picture of color-singlet exchange is correct? Each model is currently at leading order. Next-to-leading order effects of additional radiation may have a large effect of the behavior of the models. Therefore, we can only state that the leading order two gluon models are not complete enough to describe our data. This does not mean that the basic ideas of the two gluon models are wrong. Work is in progress [35] to extend the two gluon model by allowing additional gluons to be exchanged. A Next-to-Leading Log Approximation (NLLA) BFKL calculation is also in progress [36].

The soft color model qualitatively describes our data. Work is also in progress to make the comparison more quantitative [37]. The U(1) Gauge Boson model also qualitatively describes our data. However, depending on the choice of the mass and coupling constant, the predicted gap fraction may rise more steeply in E_T than the data. When the low E_T data is finalized, this measurement may be used to place additional bounds on the allowed mass and coupling in this model.

Measurement of the features of color-singlet exchange have stimulated much theoretical interest. We look forward to the next generation of color-singlet models and the results of their comparisons with our data.

Appendix A

Model for Total Energy

The total scalar energy in the calorimeter is a variable which might be used to determine the amount of multiple interaction contamination remaining in the data sample after the single interaction cuts. If the total energy is greater than 1800 GeV, then there is a positive identification of a multiple interaction. Figure A.1 shows the distribution of total energy for the high E_T trigger. Part of the single interaction cut is a hard cut at 1800 GeV; this cut is evident as an abrupt cutoff in the E_{tot} distribution. The shoulder at high E_{tot} is most likely due to multiple interactions. To determine how much contamination remains in the data sample, we will model the E_{tot} distribution for a single interaction and compare it to the distribution from the data.

The model for the total energy must take into account every source of energy deposition in the calorimeter. The sources of energy deposition are those due to noise, minimum bias activity, pileup (negative energy as discussed in Section 2.3.4), and hard physics processes. To get at all of these pieces of the total energy, we use

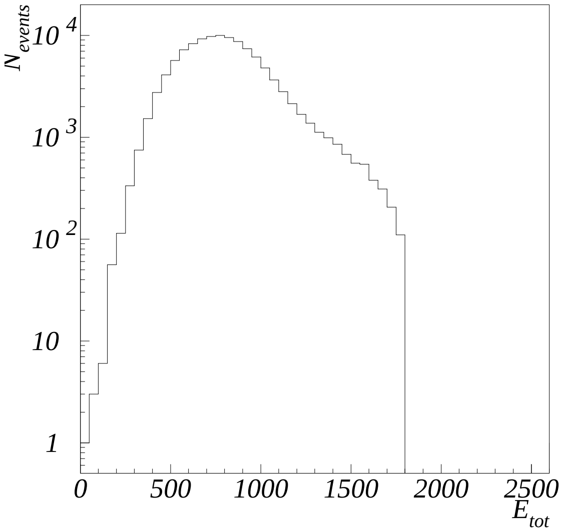

Figure A.1: Total calorimeter energy for High E_T sample.

three sources of data: jet data, Zero Bias(ZB) data and Minimum Bias(MB) data.

Zero Bias data is data recorded at specified time intervals during a physics run, but is not triggered by any physics event. This data is called Zero Bias because it should be free from any biases imposed by triggering on physics objects. This data is often used to study the pedestal offsets, or noise, in the calorimeter. To ensure that there are no hard physics processes which would overwhelm the pedestals, cuts are applies to the Zero Bias data to eliminate interactions.

Minimum Bias data is also collected during physics runs, but requires an interaction in the detector. A Minimum Bias interaction is a 'soft' interaction among a proton and antiproton which can underly the hard interaction of interest. The number of additional Minimum Bias interactions in addition to the hard interaction is a function of luminosity. Single interaction cuts have been applied to the

Minimum Bias sample to simulate adding one MB interaction at a time to the data.

The data used is from the low and high E_T data samples. The low E_T data sample is used at $\mathcal{L} \cong 0$, but there were not enough statistics for the high E_T trigger in this special low luminosity data, so the high E_T data is at a luminosity of $\sim$ 2E30 (which we will approximate as zero). Minimum Bias data at a luminosity of 5E30 will be added to the low $\mathcal{L}$ data and compared to data at $\mathcal{L}$ = 5E30. This comparison should yield the contamination present in the data. The different data samples and the terms they contribute to the total energy are tabulated in Table A.1 and shown in Fig. A.2.

Sample	Contributions to E_{tot}
Data($\mathcal{L} = 0$)	noise + hard physics process
ZB($\mathcal{L} = 5E30$)	noise + pileup
MB($\mathcal{L} = 5E30$)	noise + pileup + MB
ZB($\mathcal{L} = 0$)	noise

Table A.1: Data going into E_{tot} model

The total energy is modeled by:

$$\begin{aligned} E_{tot} = \quad & P(0)[Data \oplus ZB(\mathcal{L}=5) \ominus ZB(\mathcal{L}=0)] \quad + \\ & P(1)[Data \oplus MB \ominus ZB(\mathcal{L}=0)] \quad + \\ & P(2)[Data \oplus MB \oplus MB \ominus ZB(\mathcal{L}=0) \ominus ZB(\mathcal{L}=0)] \quad + \ldots \end{aligned} \tag{A.1}$$

$P_{\overline{n}}(k) = \frac{e^{-\overline{n}}\overline{n}^k}{k!}$ is the Poisson probability for k additional interactions when there are $\overline{n}$ interactions per crossing. The mean number of interactions per crossing is given by $\overline{n} = \mathcal{L}\tau\sigma_{\text{LØ}}$ where $\mathcal{L}$ is the instantaneous luminosity, τ is the crossing time

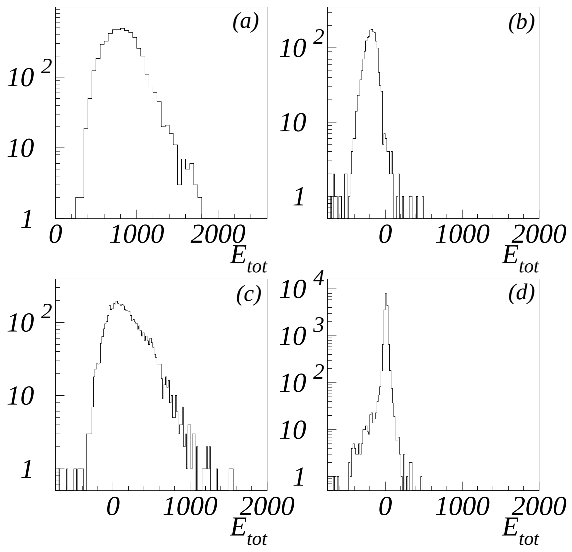

Figure A.2: Total calorimeter energy for (a) High E_T data at $\mathcal{L} = 0$, (b) Zero Bias data at $\mathcal{L} = 5\text{E}30$, (c)Minimum Bias Data at $\mathcal{L} = 5\text{E}30$, and (d) Zero Bias data at $\mathcal{L} = 0$.

(3.5 μs), and $\sigma_{\mathrm{L}\emptyset}$ is the cross section subtended by the Level Ø counters (46.7 mb). The average $\mathcal{L}$ of our data sample is approximately 5E30. This average value was used in place of the instantaneous $\mathcal{L}$ in the model. The first term in Equation A.1 is a model of the E_{tot} for a single interaction, the first two terms are a model for a single interaction plus one MB interaction, and so on. The full series represents all $\overline{n}$ MB interactions added to the data.

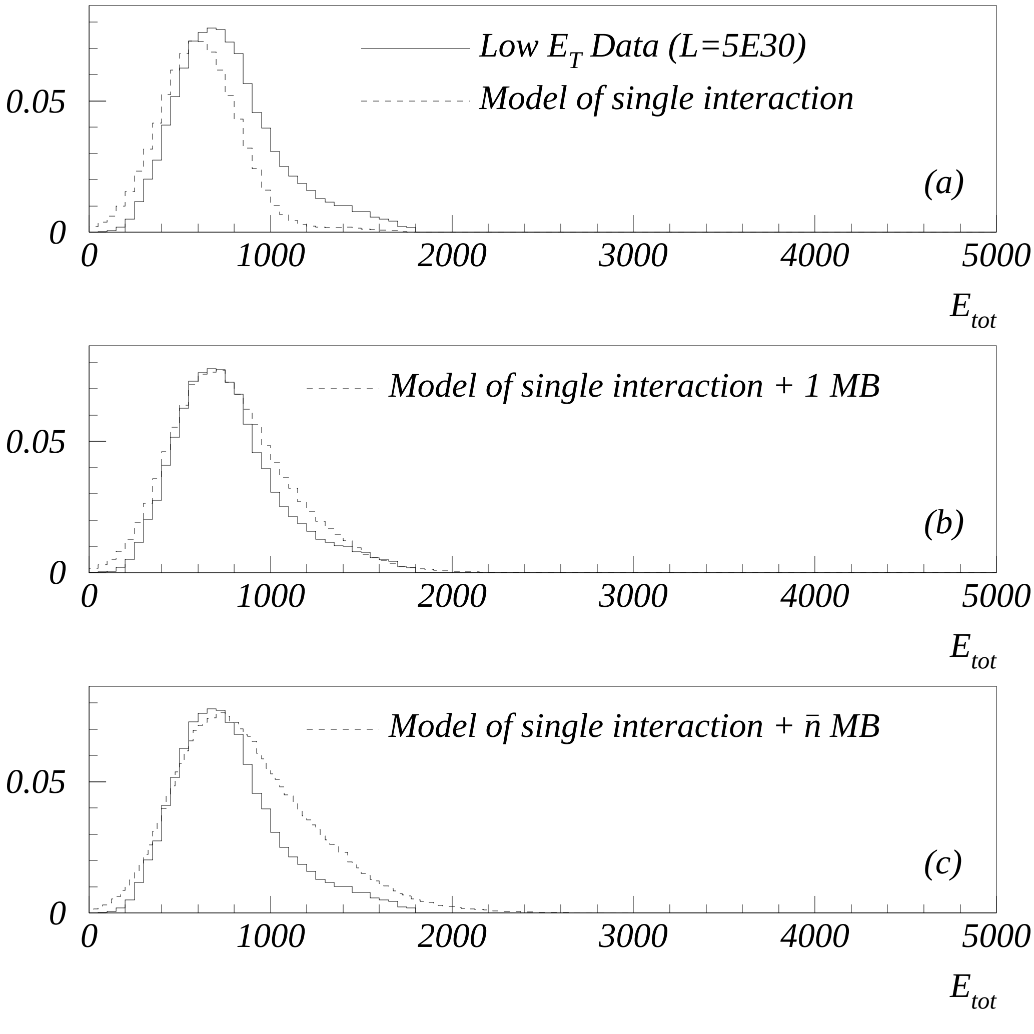

Figure A.3: E_{tot} for Low E_T data sample compared to model for (a) a single interaction, (b) a single interaction plus one Minimum Bias event, and (c) a single interaction plus $\overline{n}$ Minimum Bias events. The solid line is the data; dashed lines are the models.

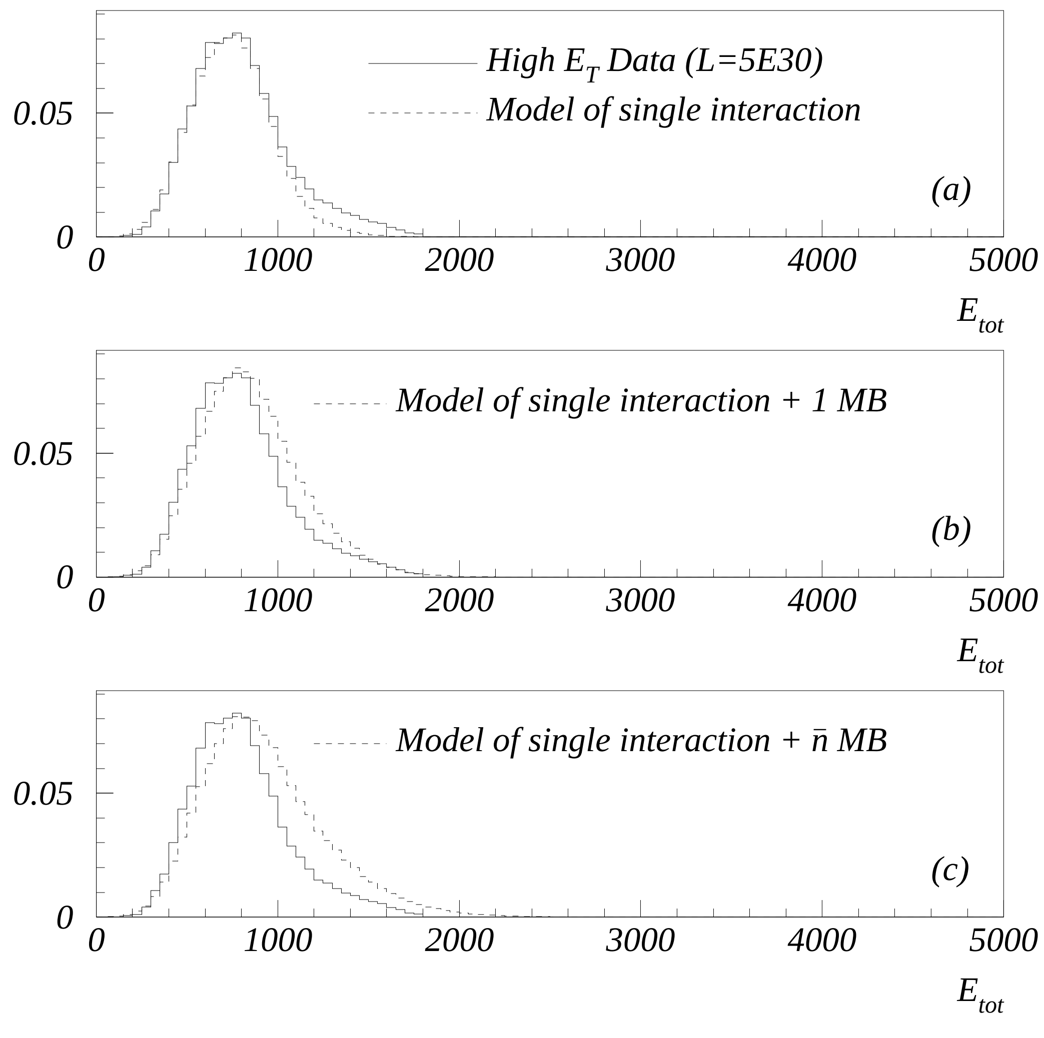

Figure A.4: E_{tot} for High E_T data sample compared to model for (a) a single interaction, (b) a single interaction plus one Minimum Bias event, and (c) a single interaction plus $\overline{n}$ Minimum Bias events. The solid line is the data; dashed lines are the models.

The E_{tot} distribution from the low E_T and high E_T data samples (at a luminosity of 5E30) are shown in Figs. A.3 and A.4 respectively. The model for the E_{tot} in a single interaction, a single interaction plus one MB interaction, and for a single interaction plus $\overline{n}$ MB interactions are also shown. The comparison with the model for a single interaction shows that there is indeed some contamination in the low E_T and high E_T data samples. The amount of contamination can be determined by including an acceptance for minimum bias in the model. One hundred percent acceptance for minimum bias means that all $\overline{n}$ MB interactions are present along with the hard interaction. Figure A.5 shows that about 30% acceptance for the low E_T sample and 20% acceptance for the high E_T sample give E_{tot} distributions which match the data. The contamination is then, given an acceptance $\mathcal{A}$:

$$C = \mathcal{A}(1 - e^{-\overline{n}}) \tag{A.2}$$

The contamination at $\mathcal{L} = 5E30$ is about 18% in the low E_T data and 12% in the high E_T data.

There are several factors, however, which prevent us from using this model as a method to correct for the multiple interaction contamination. One reason is that the shape of the E_{tot} distribution in the model does not completely match that of the data, especially at low values of E_{tot}. If the shape of the model is not correct, then there is no way to know if the resulting contamination is correct. Another factor is that the derived contamination is only for a specific luminosity. A correction would need to be determined for several value of luminosity and applied to data at that luminosity. To do this, however, would require fitting to find the

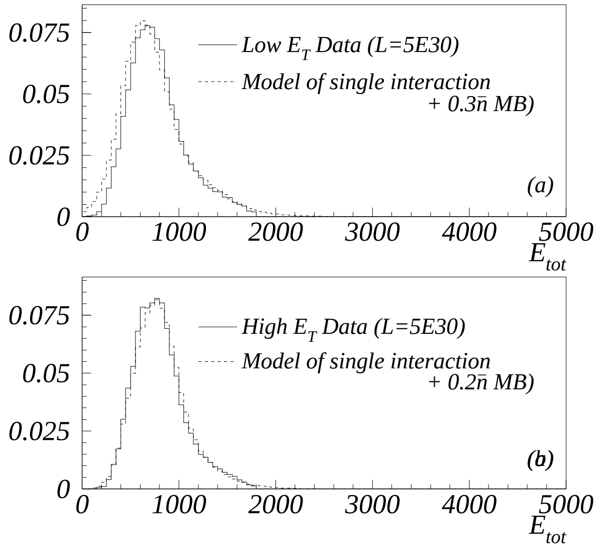

Figure A.5: (a) E_{tot} for Low E_T data sample with comparison to model of single interaction plus 30% acceptance for Minimum Bias, and (b) E_{tot} for High E_T data sample with comparison to model of single interaction plus 20% acceptance for Minimum Bias. The solid line is the data; the dashed line is the model.

fraction separately at different luminosities, this breaking the data sample up into smaller statistical samples. Section 6.4.1 shows that the fit error becomes quite large when the statistics are low, so this method would introduce large errors into the fraction. Although we will not use this model to apply a correction, it does provide qualitative information on the amount of multiple interaction contamination present in the data.

Appendix B

Kinematic Biases in the Measurement of f_S vs E_T

B.1 Bias from Binning Method

The physics bias imposed by the dijet E_T^{min} cut is a concern when considering the color-singlet fraction as a function of E_T. As the jet E_T is varied, the configuration of the events can change. Recall that the jets are E_T ordered, so the second jet E_T is always less than the leading jet E_T. When the leading jet E_T is near the imposed threshold, the second jet E_T must also be near the threshold. The two jets are almost back-to-back with very little additional jet activity in the event. When the leading jet E_T is much higher than the threshold, the second jet can have any E_T from the threshold value up to the leading jet E_T. There is more phase space for additional radiation and jet production in this configuration. Therefore, events with jet E_T near the trigger threshold will be more back-to-back than events with jet E_T far above the threshold.

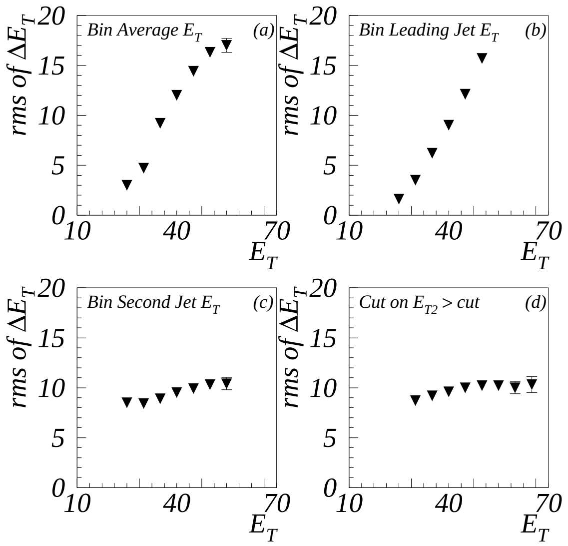

Figure B.1: Rms of ΔE_T between two leading jets for binning (a) average jet E_T (b) leading jet E_T (c) Second leading jet E_T (d) cut on $E_{T2} > threshold$.

How to define the E_T scale when there are are least two jets in the event is not clear. Some obvious choices are to bin the average E_T of the two jets, or to bin the leading jet E_T. Because of the 'configuration bias' introduced by triggering on dijets, there is an E_T dependent effect in both of these binning methods. This effect is shown (Fig. B.1(a) and (b)) by the variation in the width of ΔE_T between the two leading jets as a function of E_T. This variation means that the amount of radiation changes with E_T. Figure B.2 (a) and (b) contrasts the behavior of the dijet trigger to that of an inclusive single jet trigger. The inclusive trigger requires one jet with $E_T > 30$ GeV; the second jet can have an E_T as low as 8 GeV. The inclusive trigger shows that the E_T dependent effects in Fig. B.1 are largely induced by triggering on E_T balanced dijets. These binning methods which

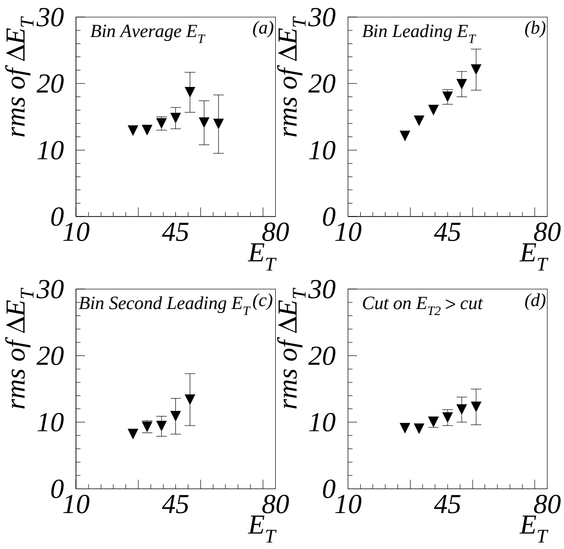

Figure B.2: Rms of ΔE_T between two leading jets for single jet inclusive trigger for binning (a) average jet E_T (b) leading jet E_T (c) Second leading jet E_T (d) cut on $E_{T2} > threshold$.

change the shape of the color-singlet fraction as a function of E_T are not wrong, but they make the measurement more difficult to interpret because the bias imposed by the trigger is convoluted with any real effects. Figure B.3 shows the measured color-singlet fraction (without luminosity or vertex correction) as a function of E_T for both of these binning methods for the high E_T trigger. As expected from the above discussion, the color-singlet fraction initially falls with E_T due to the 'configuration bias'.

Two additional methods for binning the color-singlet fraction as a function of E_T are to bin the second leading jet E_T or to impose an E_T^{min} cut on both jets. Both methods eliminate the 'configuration bias' as a function of E_T because each

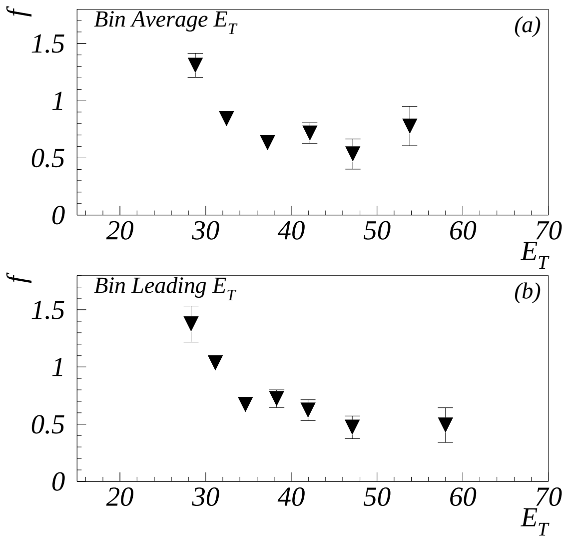

Figure B.3: Color-singlet fraction vs. E_T (high E_T trigger; uncorrected) for binning (a) average jet E_T and (b) leading jet E_T.

bin requires both jets to have an E_T greater than some threshold (thus mimicking the trigger requirement). The overall color-singlet fraction is still affected by the selection cuts, but the variation of this bias with E_T is eliminated. Figure B.1 shows that the width of ΔE_T shows little variation as a function of E_T for both of these binning methods. Therefore, binning the second leading jet E_T was chosen for the measurement of f_s versus E_T to facilitate comparisons to theoretical predictions.

B.2 Bias from Kinematics

Another possible effect on the shape of f_s as a function of E_T is introduced by the configuration of the events used for the measurement. The color-singlet fraction is the background-subtracted fraction of rapidity gap events to all events.

Color-singlet events are predominantly two jet events, whereas most events in the inclusive sample contain more than two jets. Because it is kinematically easier to produce high E_T in an event with only two jets, two jet events will preferentially populate the high E_T bins over events with more than two jets. This kinematic effect could introduce a shape in the color-singlet fraction as a function of jet E_T independent of any effects from color-singlet exchange.

To investigate the magnitude of the kinematic effect, opposite-side jet events are chosen from a single jet inclusive trigger by requiring two jets with $|\eta| > 1.7$ and $\eta_1 \cdot \eta_2 < 0$. The ratio of events with no central jets (no jets with $|\eta| < 1.7$) to the ratio of all events (R_{0jet}) is shown in Fig. B.4. This figure shows that the kinematic effect on the shape of f_s versus E_T is at most 20%. Therefore, the observed rise in f_s as a function of E_T is not due entirely to kinematic effects. While it is important to understand the presence of a kinematic effect, the effect itself is not a large concern because it can be modelled in the theory predictions.

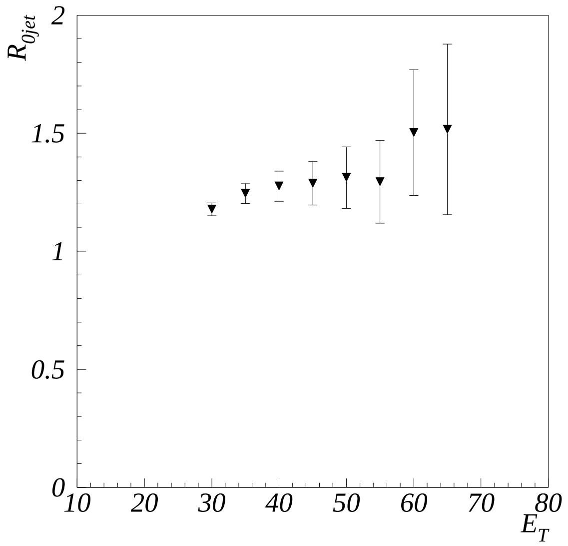

Figure B.4: Kinematic bias in f_s versus E_T: the ratio of the number of opposite-side events with no central jets to the number of all opposite-side jet events.

Appendix C

Multiple Interaction Tool

C.1 Introduction

A study has been made to optimize the multiple interaction tool . The multiple interaction tool for Run 1a was based on a study of runs with luminosities up to $8\text{x}10^{30}\text{cm}^{-2}\text{s}^{-1}$ (8E30) [38, 39]. The higher luminosities of Run 1b lead to the need to re-tune the multiple interaction tool definition. This study was done with ALL stream data from global runs spanning a wide range of luminosities and time (spread over the duration of Run 1b). Triggers which require multiple interaction flag equal to one at the trigger level are excluded. Runs with luminosity up to 28E30 were used in this study.

The detector elements used to evaluate the multiple interaction tool are the Level Ø luminosity counters, the calorimeter, and the central detector. Information from these detectors is combined in either a decision criteria or a weighting scheme to determine the value of the multiple interaction tool. In the weighting scheme, individual weights are assigned to certain variables by studying their behavior at high and low luminosity. These weights are then combined to a total weight which determines the value of the multiple interaction tool.

The ability of the multiple interaction tool to identify a single interaction varies with luminosity and the tool misidentifies more multiple interactions as single interactions as the luminosity increases. The multiple interaction tool returns a value of -1 to 5 based on the probability of a single interaction with the values having the following meanings.

- A tool value of 0 means there was no interaction.
- A tool value of -1 indicates that the Central Detector did not find any vertices, but other information points to a multiple interaction tool value of one.
- A tool value of one means 'most likely' a single interaction,
- two means 'likely' a single interaction,
- three means 'likely' a multiple interaction and
- four means 'most likely' a multiple interaction.
- A tool value of five means there are likely three or more interactions. Events with tool of five are a subset of events with tool equal to four.

C.2 Level Ø Contribution

Timing information from the Level Ø luminosity counters is used to determine an online multiple interaction flag. The multiple interaction flag has values from one to four analogous to those of the multiple interaction tool. The value of the multiple interaction flag is one element in the weighting scheme for the multiple interaction tool. Figures C.1(a)-(d) show the fraction of all events with multiple interaction flag values of one, two, three, or four. The solid line is the expected fraction of events for a single (Fig. C.1(a),(b)) or multiple (Fig. C.1(c),(d)) interaction. The expected fraction of events is derived from Poisson statistics. The average number of interactions per crossing is:

$$\overline{n} = \mathcal{L}\tau\sigma_{\mathrm{L}\emptyset}$$

where $\mathcal{L}$ is the instantaneous luminosity, τ is the crossing time (3.5 μs), and $\sigma_{L\emptyset}$ is the cross section subtended by the Level Ø counters (46.7 mb) [40]. The expected number of single interactions is then $P(1) = e^{-\overline{n}}$ and the expected number of multiple interactions is $P(>1) = 1 - e^{-\overline{n}}$. Figure C.2(a) shows the fraction of all events with flag 1+2 (events flagged as single interactions) and Fig. C.2(b) shows the fraction of all events with flag 3+4 (events flagged as multiple interactions). The fraction of events with flag 1+2 is higher than the expected fraction for single interactions even at low luminosities. This excess is probably due to the fact that multiple interaction flag equal to one is a looser cut in Run 1b than it was in Run 1a. The looser cut was desired by physics groups who wanted to use the multiple interaction flag equal to one requirement at the trigger level. From these figures, one can clearly see the saturation of flag = 1 at high luminosities, thus pointing to the need for other inputs to the tool.

The Level Ø SLOWZ vertex is also used as an input to the multiple interaction tool. The difference between the SLOWZ and the Central Detector (CD) vertex positions can be useful in the determination of single vs multiple interaction. Figures C.3(a)-(d) show a two dimensional plot of the CD Z vertex position vs the difference between the Z of the Level Ø and CD for several luminosity values. There is a clear horizontal band at low luminosity (more predominantly single interactions) which shifts to a diagonal band at high luminosity. This shows that the Level Ø and CD vertex positions agree well for single interactions, but do not agree when there are predominantly multiple interactions. These plots are used to assign a weight based on the difference between the CD and Level Ø Z vertex positions.

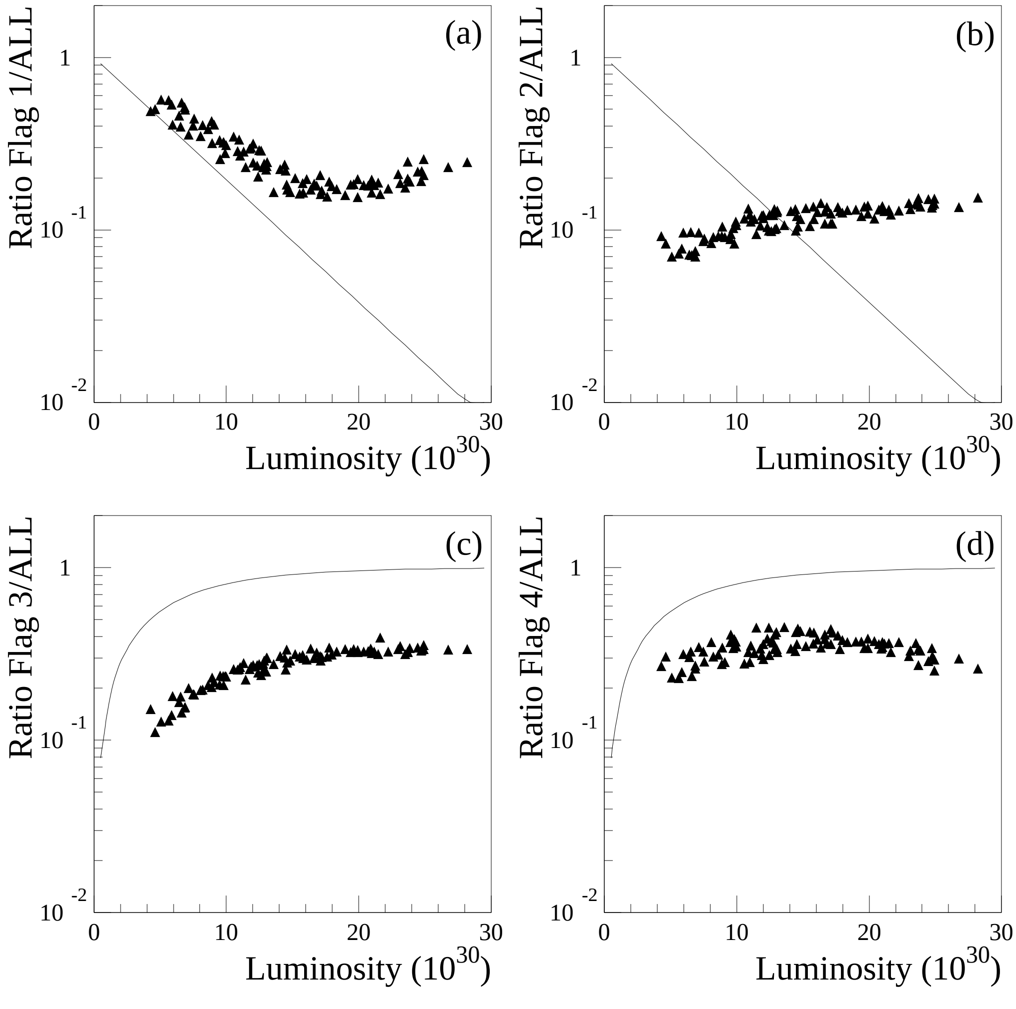

Figure C.1: Fraction of events with multiple interaction flag (a)one (b)two (c)three (d)four.

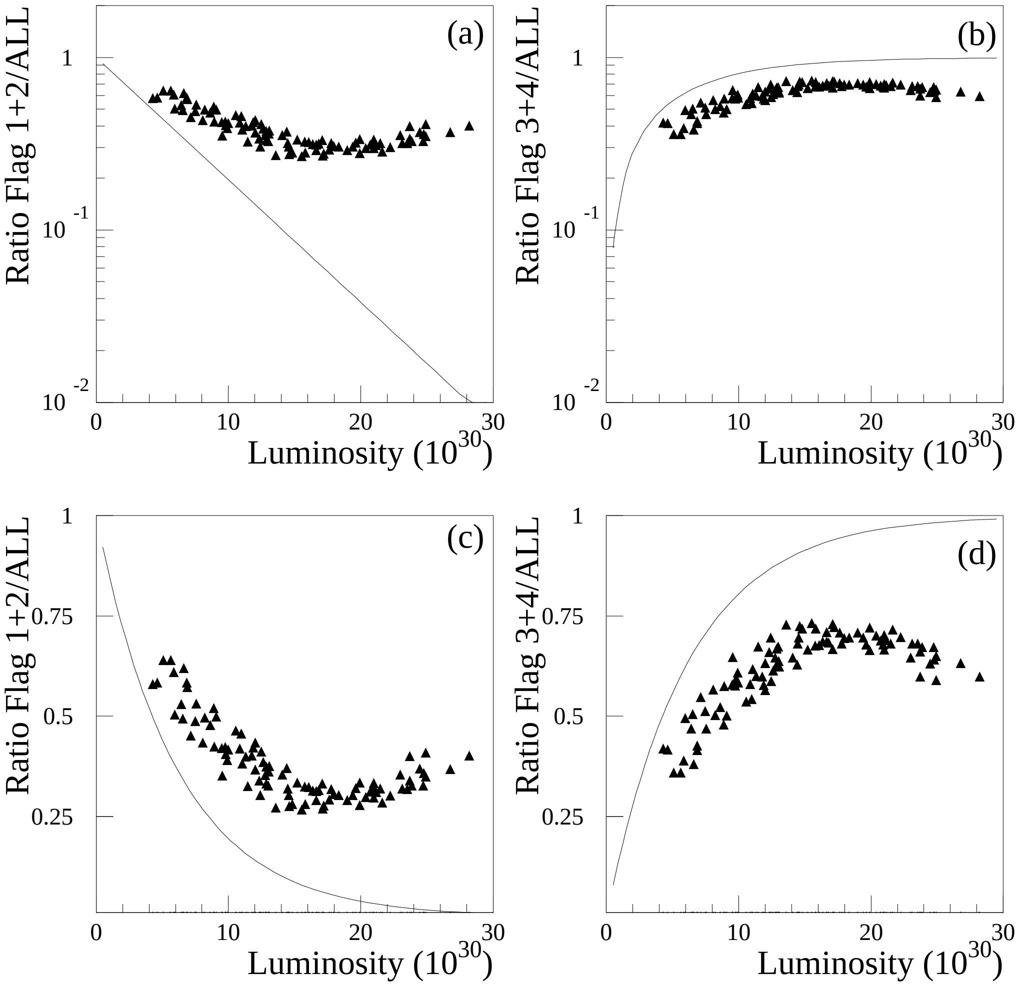

Figure C.2: Fraction of events with: multiple interaction flag 1+2 on (a) log scale, (c) linear scale and multiple interaction flag 3+4 on (b) log scale, (d) linear scale.

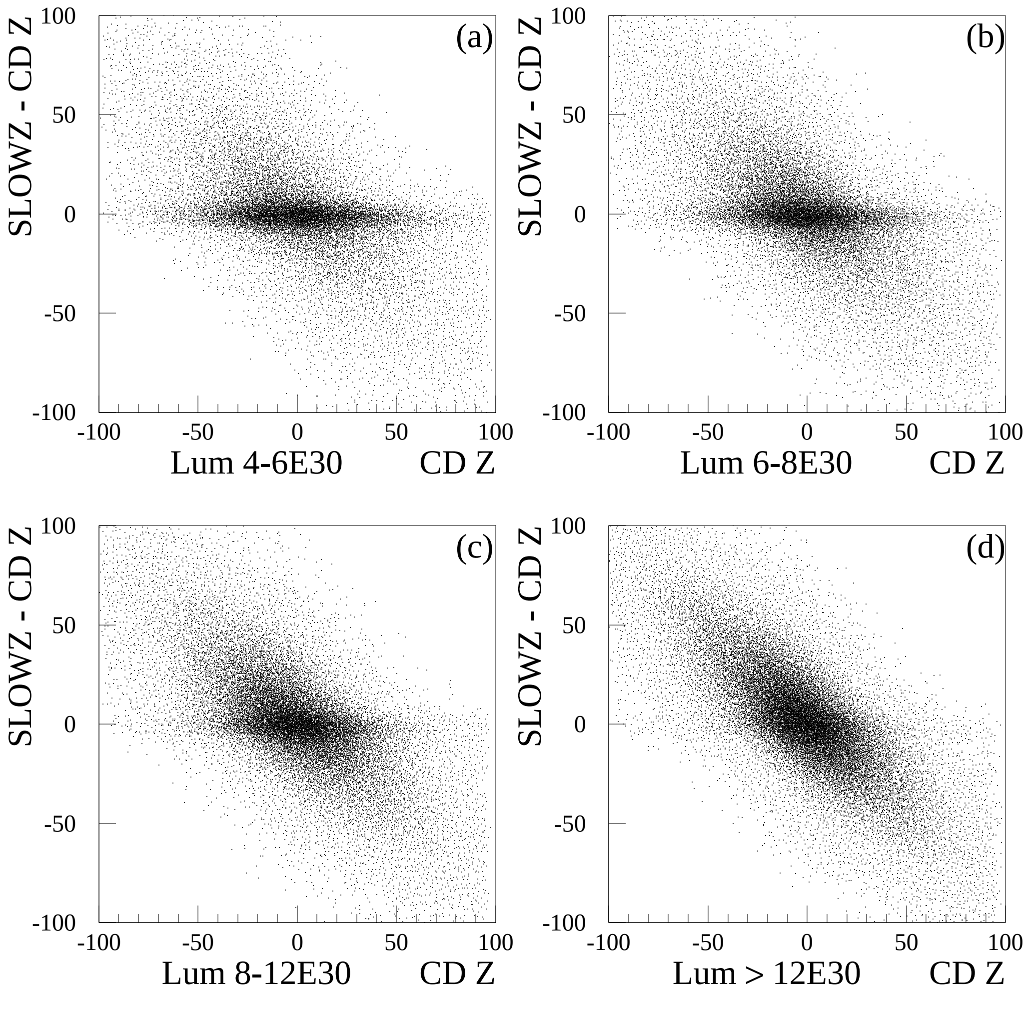

Figure C.3: CD Z vs Level Ø - CD Z for luminosities of approximately (a) 4-6E30 (b) 6-8E30 (c) 8-12E30 (d) greater than 12E30.

C.3 The Calorimeter Contribution

The energy in the calorimeter is used as an input to the tool. If the total scalar energy in the calorimeter is greater than 1800 GeV or if the scalar energy in either of the end calorimeters is greater than 900 GeV, the tool is automatically set to indicate a multiple interaction. The magnitude of the total and end calorimeter energies is shown in Fig. C.4. The discontinuity at 1800 GeV in the total energy distribution is caused by triggers which require a scalar energy sum of less than 1800 GeV at the trigger level. The magnitude of the total and end calorimeter energies is used to assign a weight to contribute to the tool.

C.4 Central Detector Contribution

The Central Detector contributes to a large part of the multiple interaction tool definition. The difference between the CD and Level Ø vertex positions was discussed in section C.2. Another input to the tool is the number of vertices found by the CD. If the Central Drift Chamber is used to find the vertex, a maximum of three vertices can be found with the current tracking code. If the Forward Drift Chambers are used for the vertex finding, a maximum of one vertex can be found. Therefore, the vertex multiplicity distribution is only used as an input to the tool if the vertex was found by the Central Drift Chamber (which is true in most cases). The distribution for the number of vertices for four luminosity ranges is shown in Figs. C.5(a)-(d). These figures show that the CD finds more multiple vertices as the luminosity increases, but still finds a substantial fraction of single vertex events even at the highest luminosities.

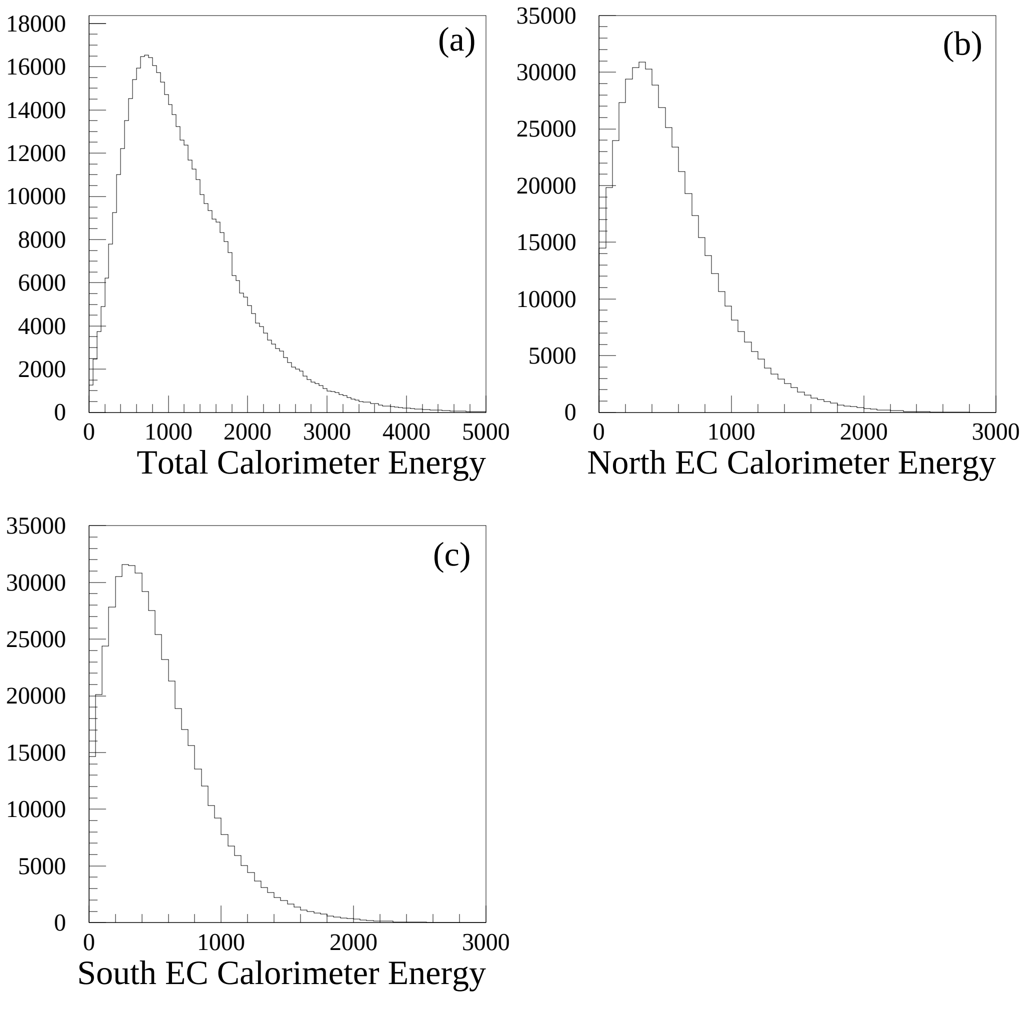

Figure C.4: Total scalar energy in (a) the entire calorimeter (b)the north end calorimeter and (c) the south end calorimeter.

The number of tracks used to find the primary vertex is also used in the weighting scheme. Figure C.6 shows the percentage of total tracks used for the primary vertex (weight) versus the number of tracks used to find that vertex. This plot shows a difference in concentration at low and high luminosities. We make a topological cut as indicated by the line and assign a weight accordingly. This cut is also supported by a Monte Carlo study discussed in section C.7.

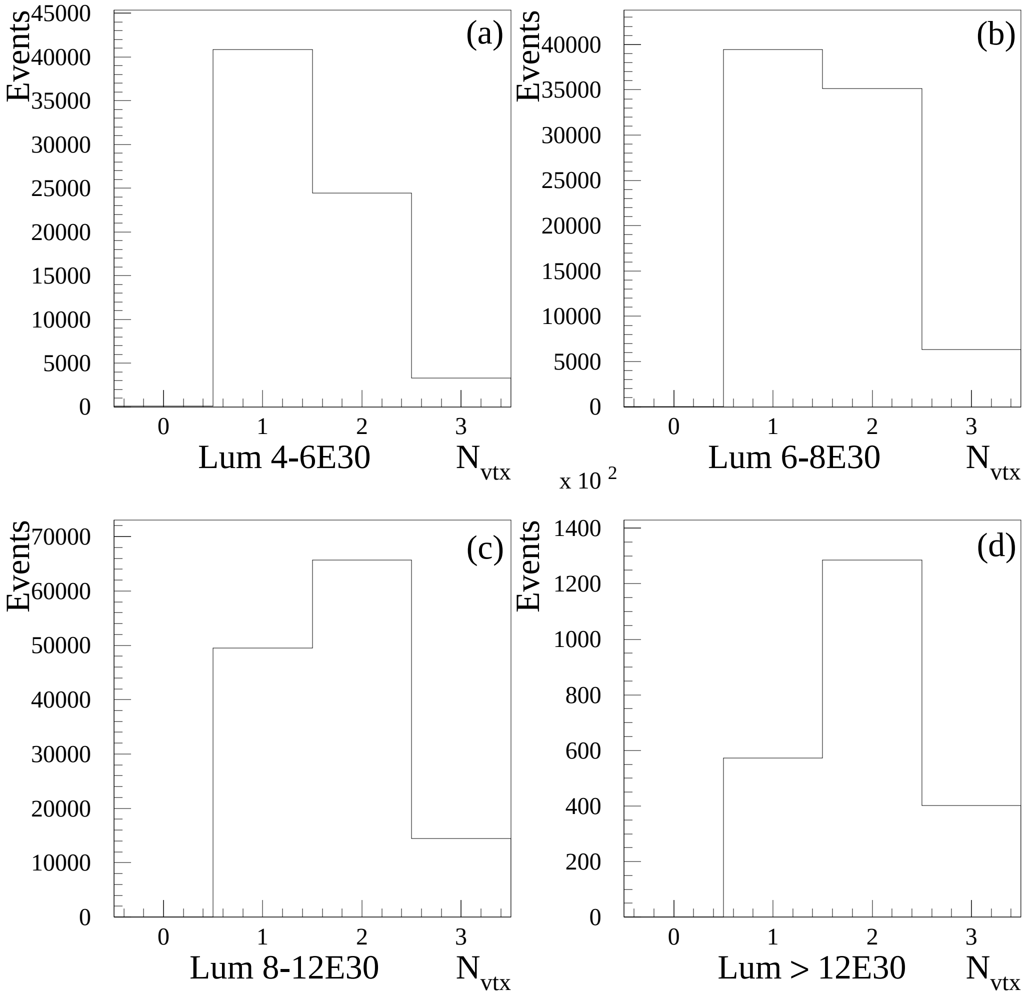

Figure C.5: Number of CD vertices.

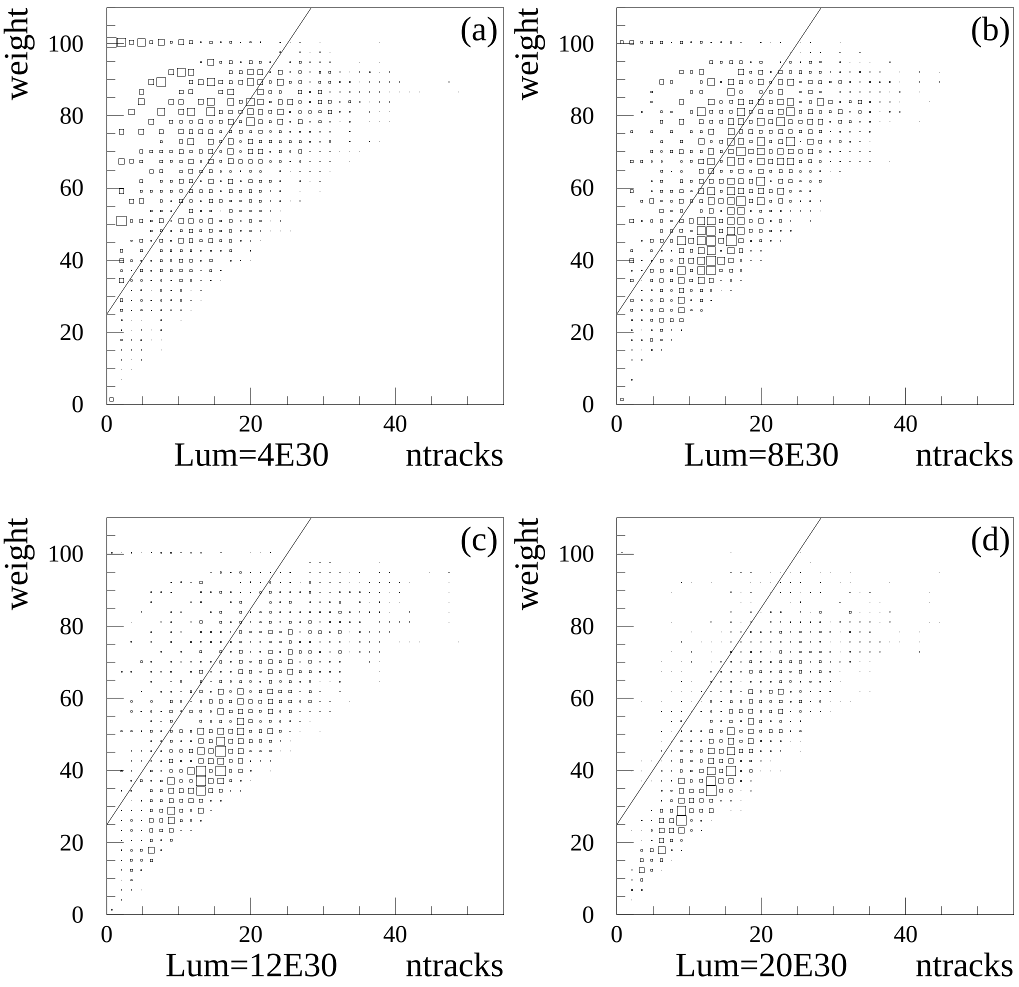

Figure C.6: Weight(percentage of tracks used) vs. number of tracks in primary vertex for luminosities of approximately (a) 4E30 (b) 8E30 (c) 12E30 and (d) 20E30.

C.5 Multiple Interaction Tool

The value of the multiple interaction tool is determined by either a standard decision criteria or by a weighting scheme (see Fig. 20). The standard decision is used to assign multiple interaction tool values of minus one, zero, one, four, and five. Figure 21 shows the criteria which go into the standard decision. If none of the conditions in the standard criteria are met, then the weighting scheme is employed. The weighting scheme is used to assign multiple interaction tool values of two, three, and four. In the weighting scheme, the value of the tool is determined by assigning a weight to each of the quantities mentioned in the previous sections, then adding up the individual weights. The final distribution of weights is then broken up to assign tool values of two, three, and four. The multiple interaction flag and the number of CD vertices each contribute 1/3 to the total weight. The remaining 1/3 is split equally between the contributions from the weights assigned to the difference in CD and Level Ø vertex positions, the calorimeter energies, and the topological cut on the number of tracks in the primary vertex. A detailed discussion of the multiple interaction tool algorithm is given in Section C.8. The saturation of the multiple interaction flag at high luminosities (as seen in Figs. C.1 and C.2) leads one to question giving it 1/3 of the total weight. However, varying the weight contributed by the multiple interaction flag with luminosity has little effect on the multiple interaction tool result, so the flag weight was kept as a constant.

Figure C.7 shows the distribution of values for the multiple interaction tool. Note that the relatively high number of multiple interaction tool four and five events is due to including a disproportionate amount of high luminosity data in our study sample. The distribution of luminosities for the data used in the study is shown in Fig. C.8. Figures C.9(a)-(d) show the fraction of events with multiple interaction tool equal to one, two, three, or four. Figures C.10(a) and (b) show the fraction of events with multiple interaction tool of 1+2 (singles) and 3+4 (multiples). The solid lines are again the expectation for a single or multiple interaction.

As the luminosity increases beyond approximately 20E30, the multiple interaction tool identifies increasingly more multiple interactions as single interactions. This contamination is clear in Fig. C.10(a) in which, at high luminosity, the fraction of events identified as single interactions is greater than the fraction of events expected for a single interaction. The contamination of multiple interactions into the tool result for a single interaction is a function of luminosity and can vary for different data sets. Figures C.11(a)-(d) show the deviation from expectation for a single interaction as a function of luminosity for both the multiple interaction flag and multiple interaction tool. Notice that the vertical scale for the multiple interaction flag plots is different from that for the multiple interaction tool plots. A line is fit to the tool ratios to show the deviation from expectation as a function of luminosity. Analyses which use the multiple interaction tool at the highest luminosities from Run 1b should beware of this contamination.

C.5.1 Comparison to Run 1a Definition

A comparison of the Run 1a definition of the tool to the new tool is shown in Figures C.12 and C.13. The new and old definition of the multiple interaction tool were run on the same global data set used in the previous discussion. The new tool does much better at high luminosities. Figure C.13(a) shows that the Run 1a definition of the multiple interaction tool used on Run 1b data produces more tool 1+2 events than expected for single interactions at all luminosities. This excess is again due to the different cuts used for the multiple interaction flag equal to one in Run 1a and Run 1b. The greater number of multiple interaction flag equal to one events in Run 1b is compensated for in the new multiple interaction tool by tuning it so there are less tool equal to two events. This difference is apparent in Fig. C.12(b). The greatest difference in the two definitions can be seen in the multiple interaction tool values of two and three – the values to which the new weighting scheme contributes the most.

Although the new multiple interaction tool is tuned to the Run 1b data conditions, it can be used on Run 1a data as well. Figures C.14 and C.15 show both the new and old multiple interaction tools run on a sample of Run 1a global data. The new tool definition is more inefficient at identifying single interactions than the Run 1a tool for the lowest luminosities (2-4E30) of the Run 1a data, but does slightly better than the old tool as the luminosity increases. These figures show that the new definition of the multiple interaction tool can be used for both Run 1a and Run 1b data.

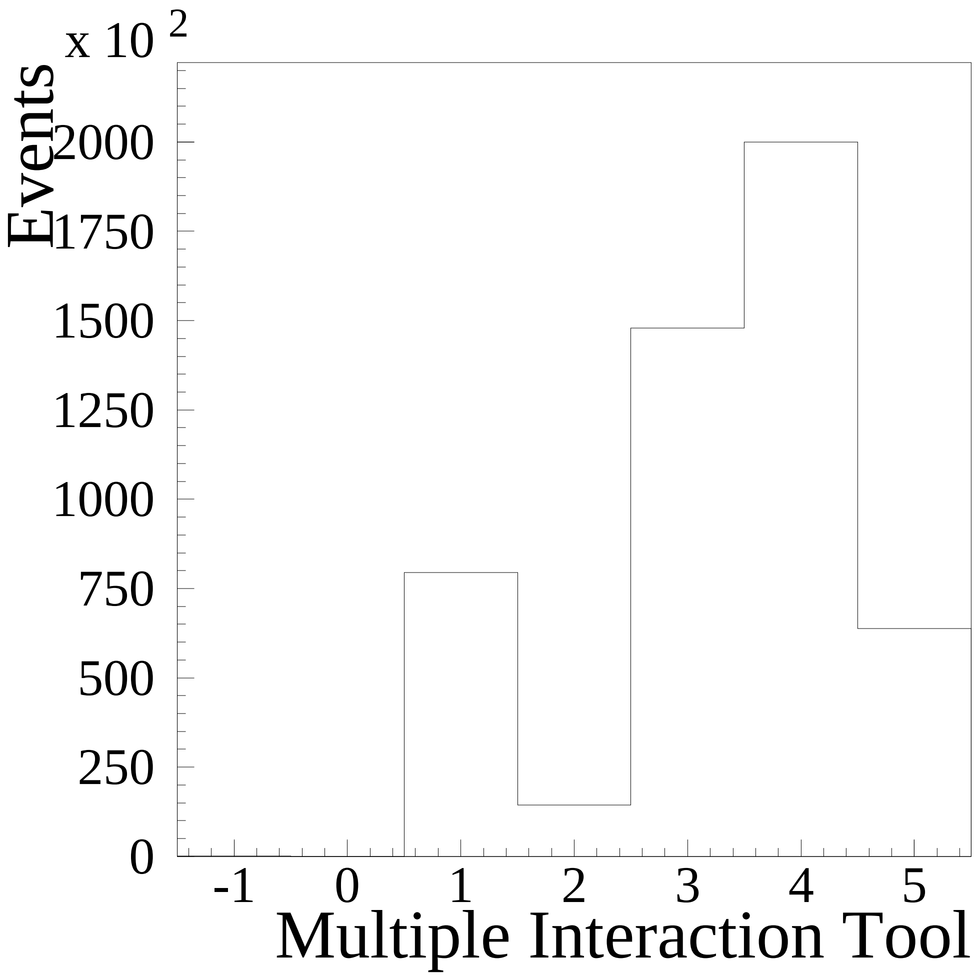

Figure C.7: Distribution of multiple interaction tool.

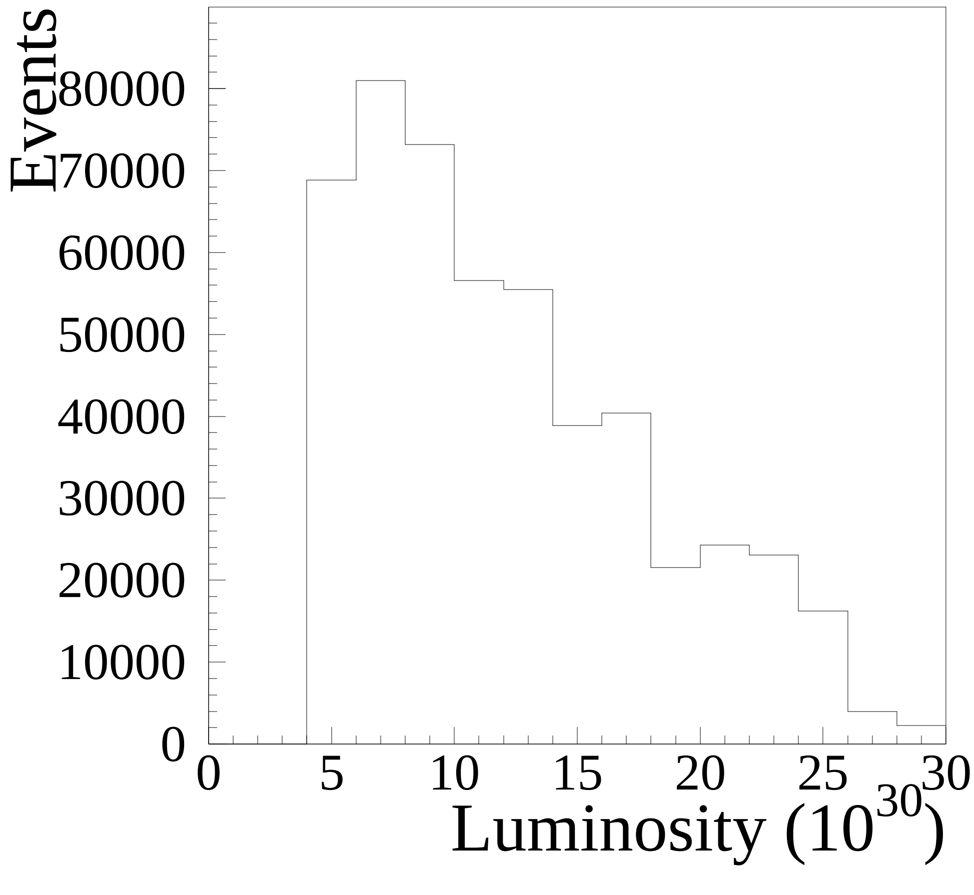

Figure C.8: Distribution of luminosity for data sample studied.

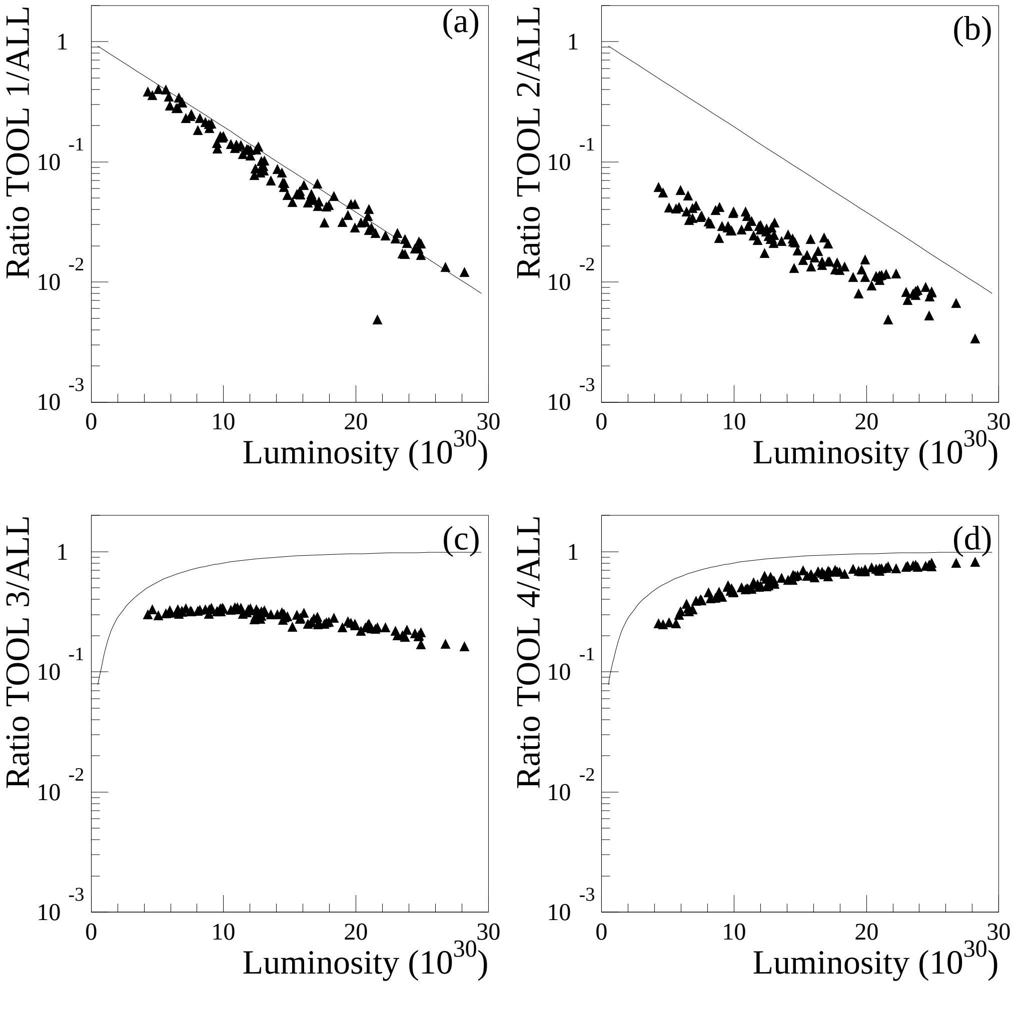

Figure C.9: Fraction of events with multiple interaction tool (a)one (b)two (c)three (d)four.

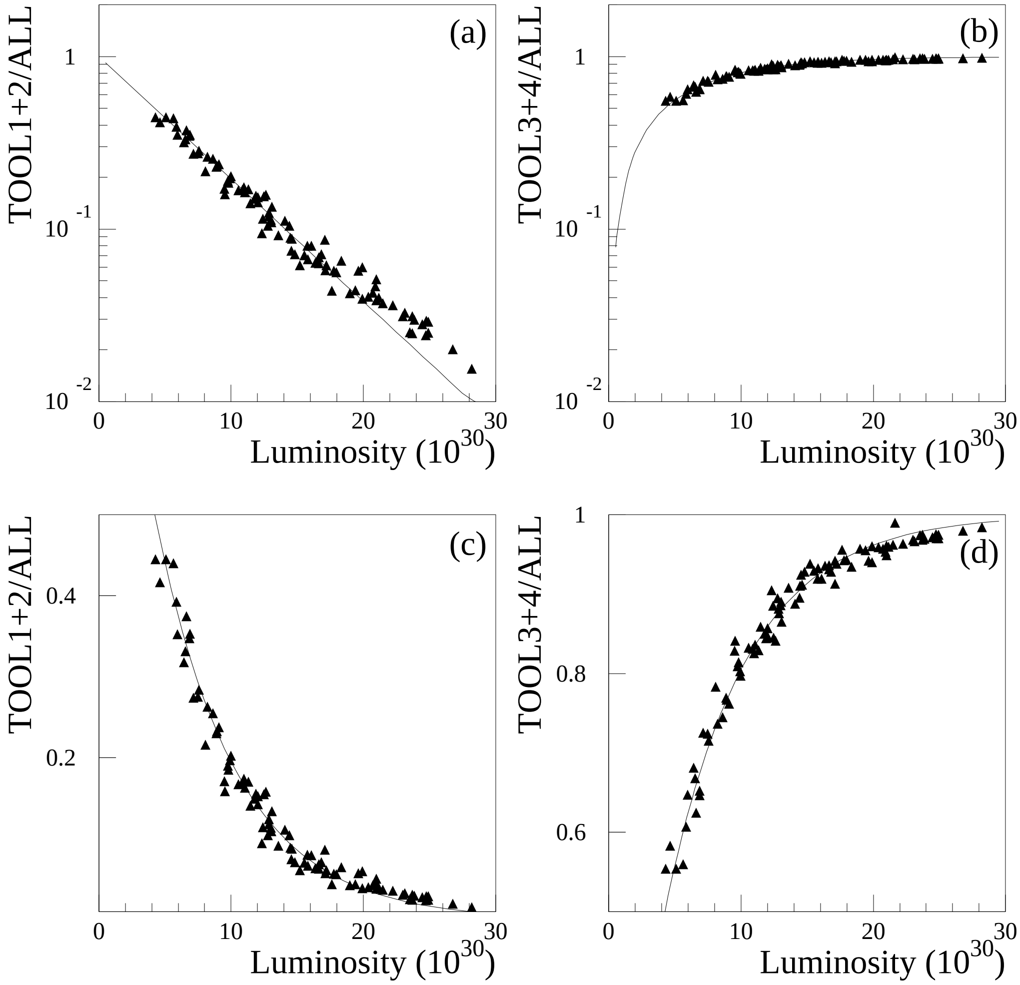

Figure C.10: Fraction of events with: multiple interaction tool 1+2 on (a) log scale, (c) linear scale and multiple interaction tool 3+4 on (b) log scale, (d) linear scale.

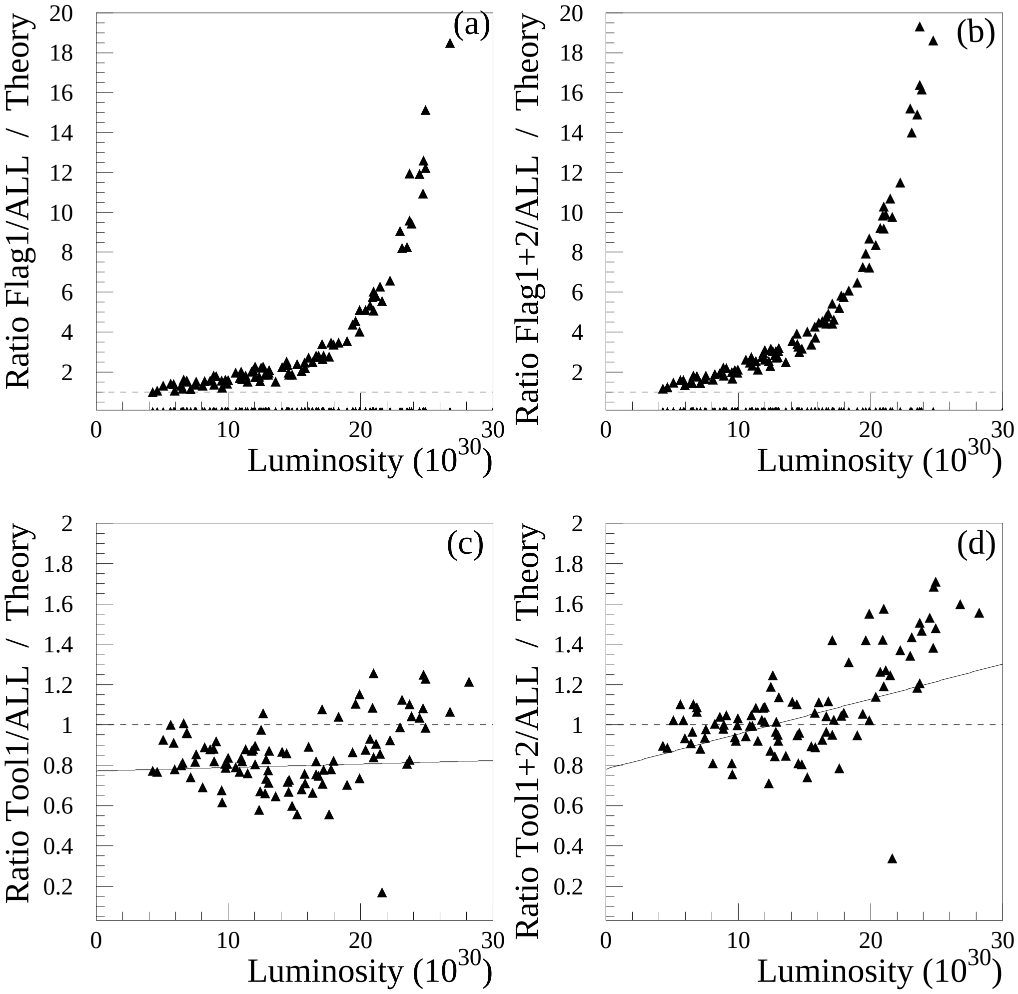

Figure C.11: Ratio of data/theory (expectation for single interaction) for (a) multiple interaction flag 1 (b) multiple interaction flag 1+2 (c) multiple interaction tool 1 (d) multiple interaction tool 1+2.

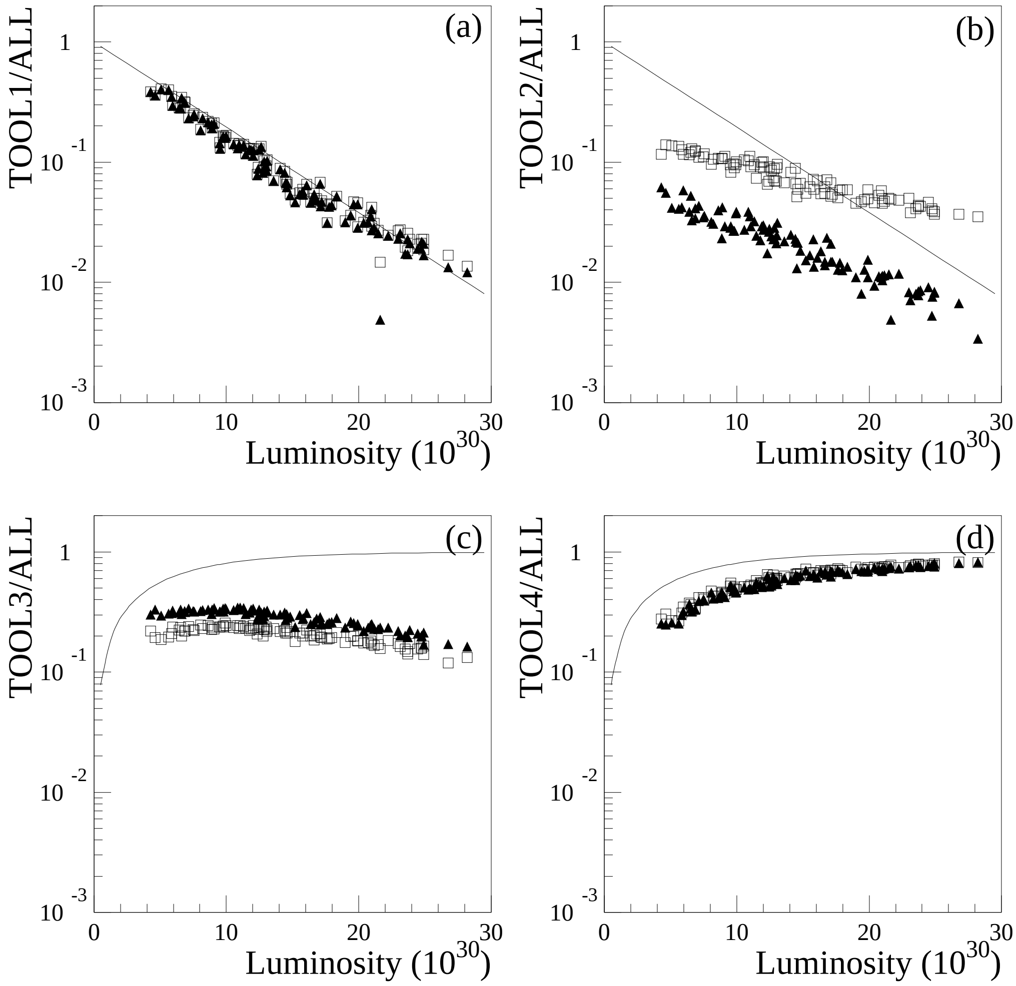

Figure C.12: Comparison of multiple interaction tool definitions. Open squares are the Run 1a multiple interaction tool definition and closed triangles are the new definition for multiple interaction tool of (a) one (b) two (c) three (d) four.

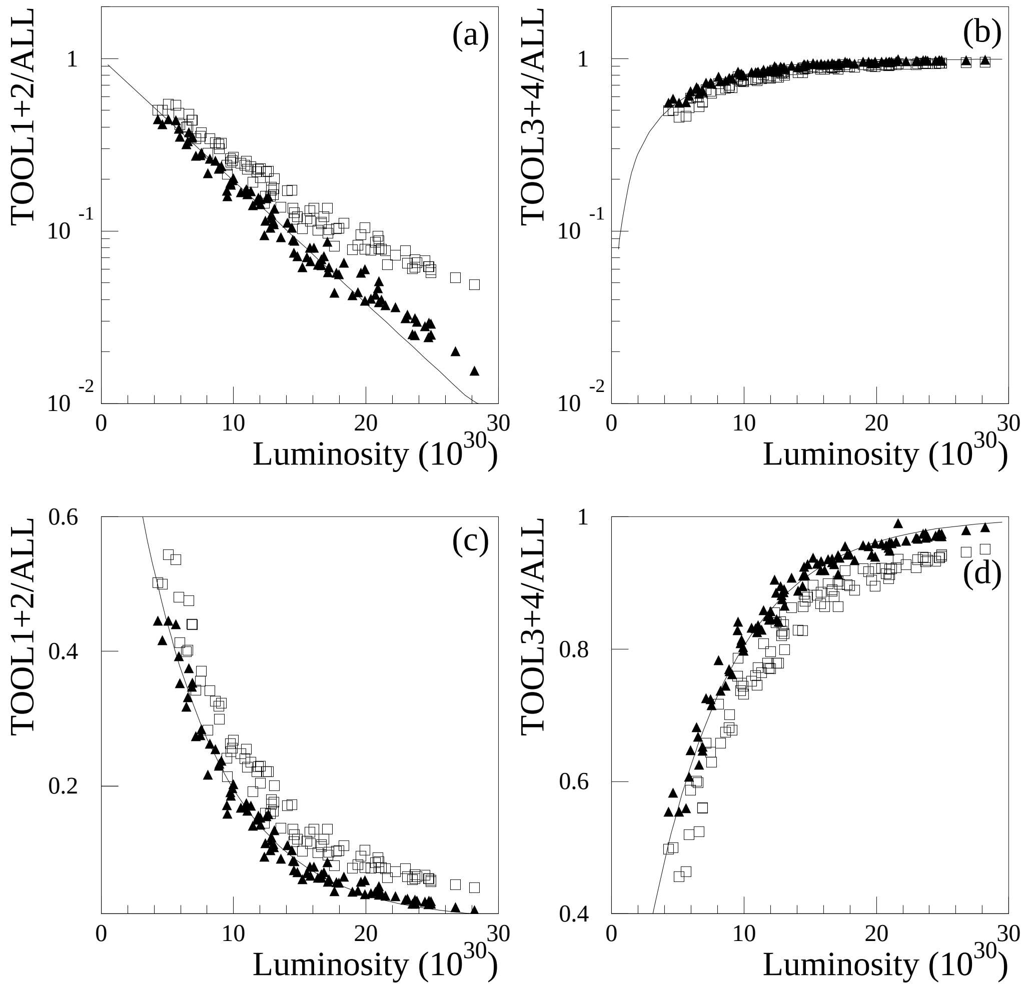

Figure C.13: Comparison of multiple interaction tool definitions. Open squares are the Run 1a multiple interaction tool definition and closed triangles are the new definition for: multiple interaction tool 1+2 on (a) log scale, (c) linear scale and multiple interaction tool 3+4 on (b) log scale, (d) linear scale.

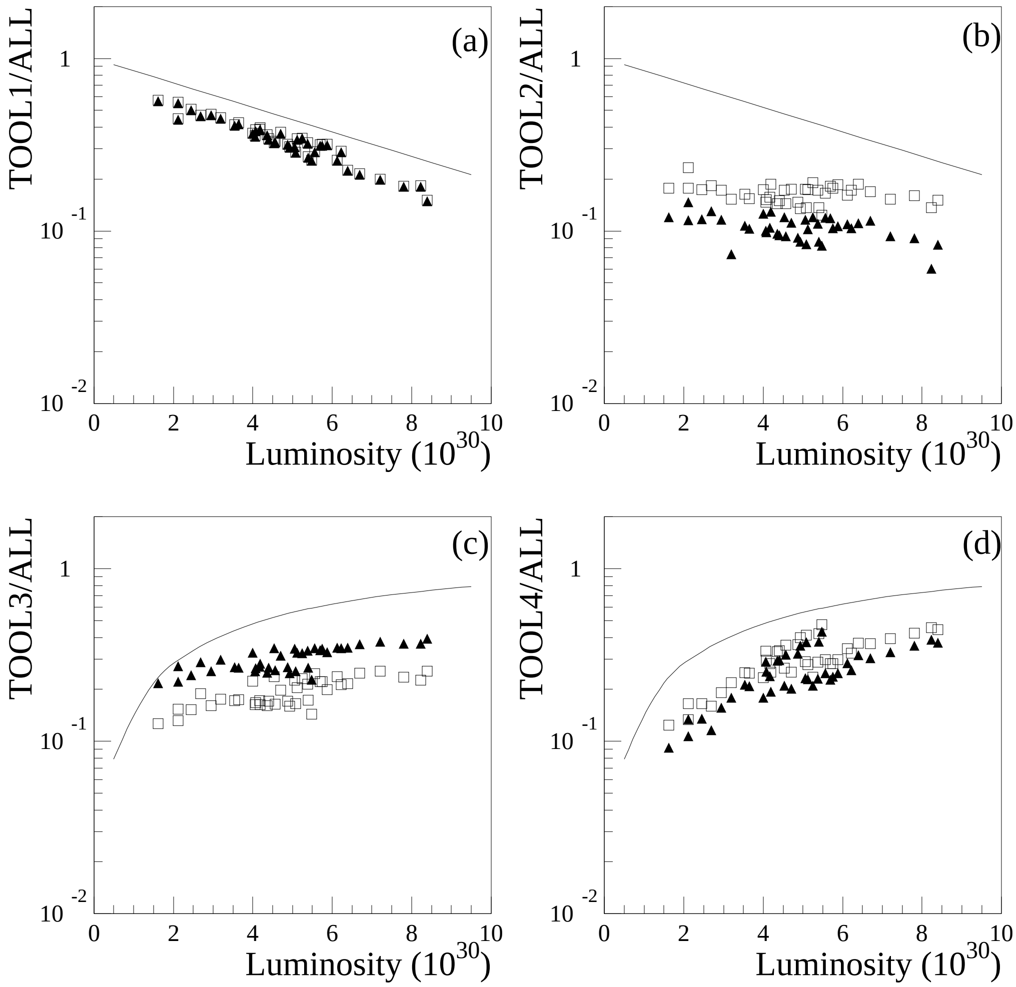

Figure C.14: Comparison of multiple interaction tool definitions on Run 1a data. Open squares are the Run 1a multiple interaction tool definition and closed triangles are the new definition for multiple interaction tool of (a) one (b) two (c) three (d) four.

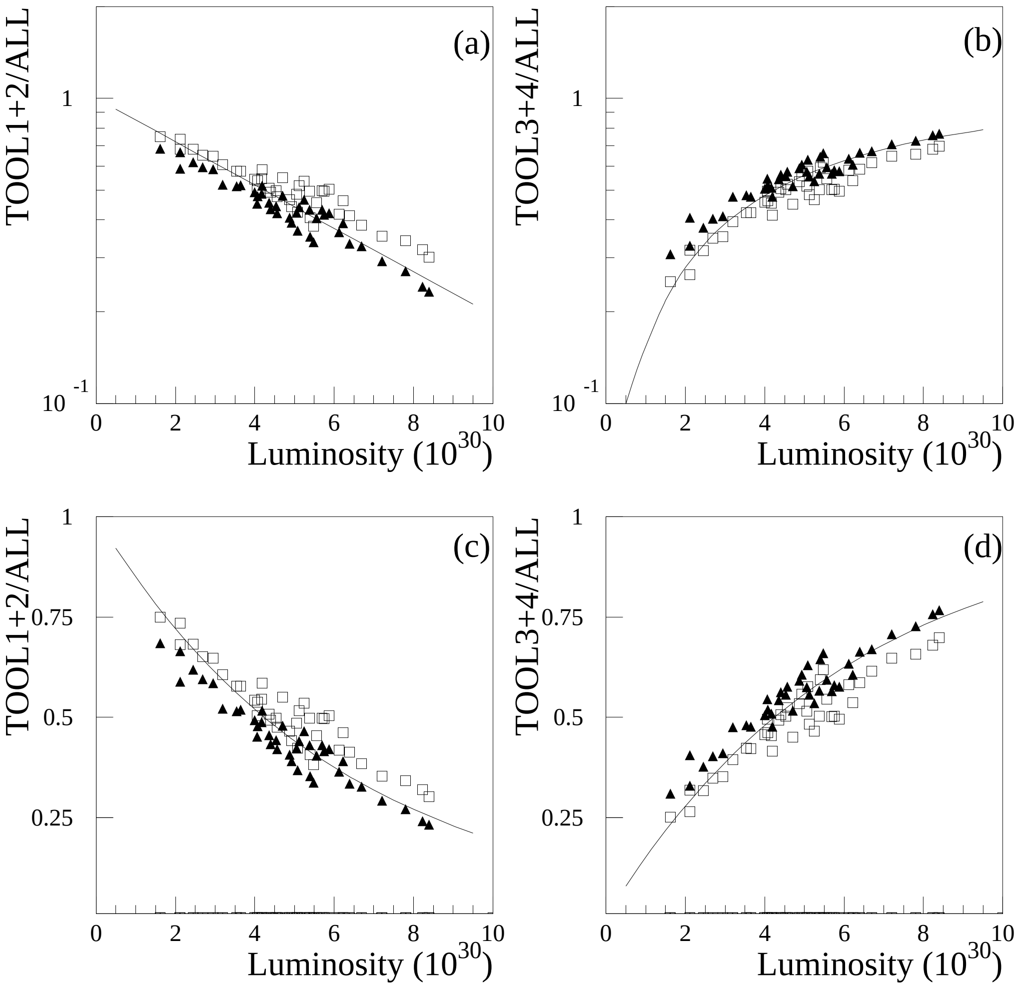

Figure C.15: Comparison of multiple interaction tool definitions on Run 1a data. Open squares are the Run 1a multiple interaction tool definition and closed triangles are the new definition for: multiple interaction tool 1+2 on (a) log scale, (c) linear scale and multiple interaction tool 3+4 on (b) log scale, (d) linear scale.

C.6 Test of the Multiple Interaction Tool

The new multiple interaction tool was tested on the same global sample of runs with only electron filters selected. The test was done with electron (W,Z, and top) triggers because these should be less sensitive to luminosity than some of the other global triggers. The results for the multiple interaction tool for this data sample are shown in Figures C.16 and C.17. Notice that the tool 1+2 result gives less single interactions than expected at low luminosities and more single interactions than expected at high luminosities. This test shows that the contamination from multiple interactions is not only a function of luminosity but also a function of the data studied.

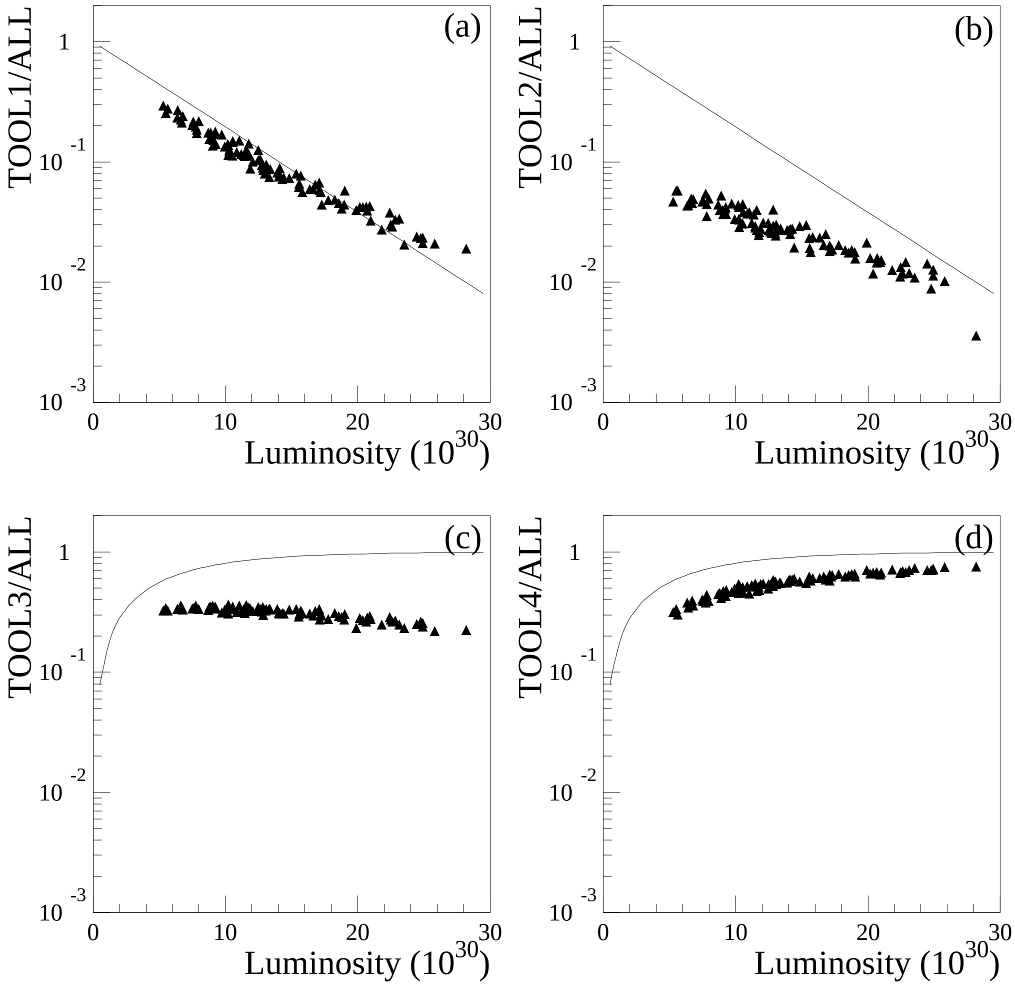

Figure C.16: Fraction of events (electron triggers) with multiple interaction tool (a)one (b)two (c)three (d)four.

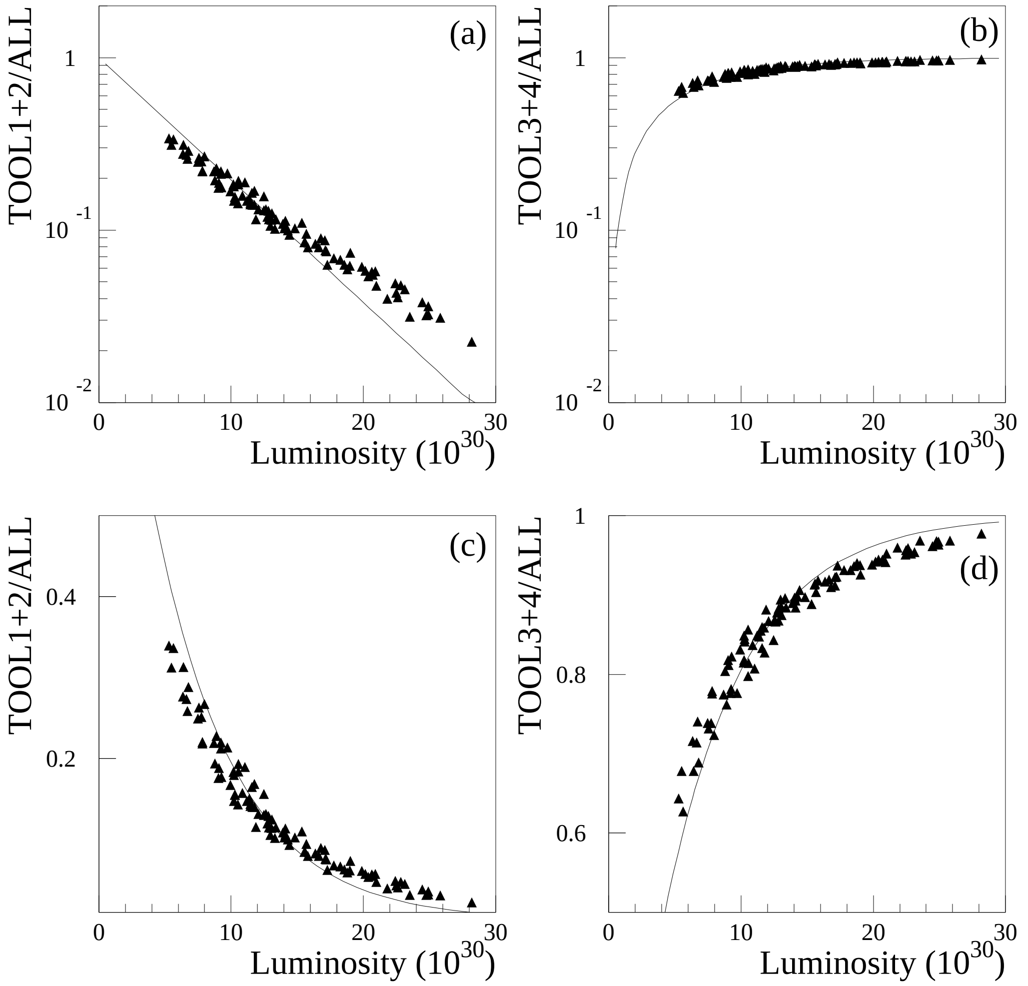

Figure C.17: Fraction of events (electron triggers) with: multiple interaction tool 1+2 on (a) log scale, (c) linear scale and multiple interaction tool 3+4 on (b) log scale, (d) linear scale.

C.7 Monte Carlo Studies

We used Monte Carlo generated event samples to test some of the distributions which go into the multiple interaction tool. Events were generated with ISAJET and a subset were passed through the MBR event generator to simulate an added minimum bias event. There are four event samples: $Z \rightarrow \mu\mu$ and top (180 GeV) both with and without a minimum bias event added. Figures C.18(a)-(d) show the vertex multiplicity for each of these samples. The CD predominantly finds one vertex in the events generated without a minimum bias event. When a minimum bias event is added to the ISAJET event, a second vertex is found only about half of the time. Figures C.19(a) -(d) show the plots used to make the topological cut in the multiple interaction tool definition. The solid line is the same cut used in the multiple interaction tool. Both figures C.18 and C.19 show the dependence of the vertex finding on the physics process involved.

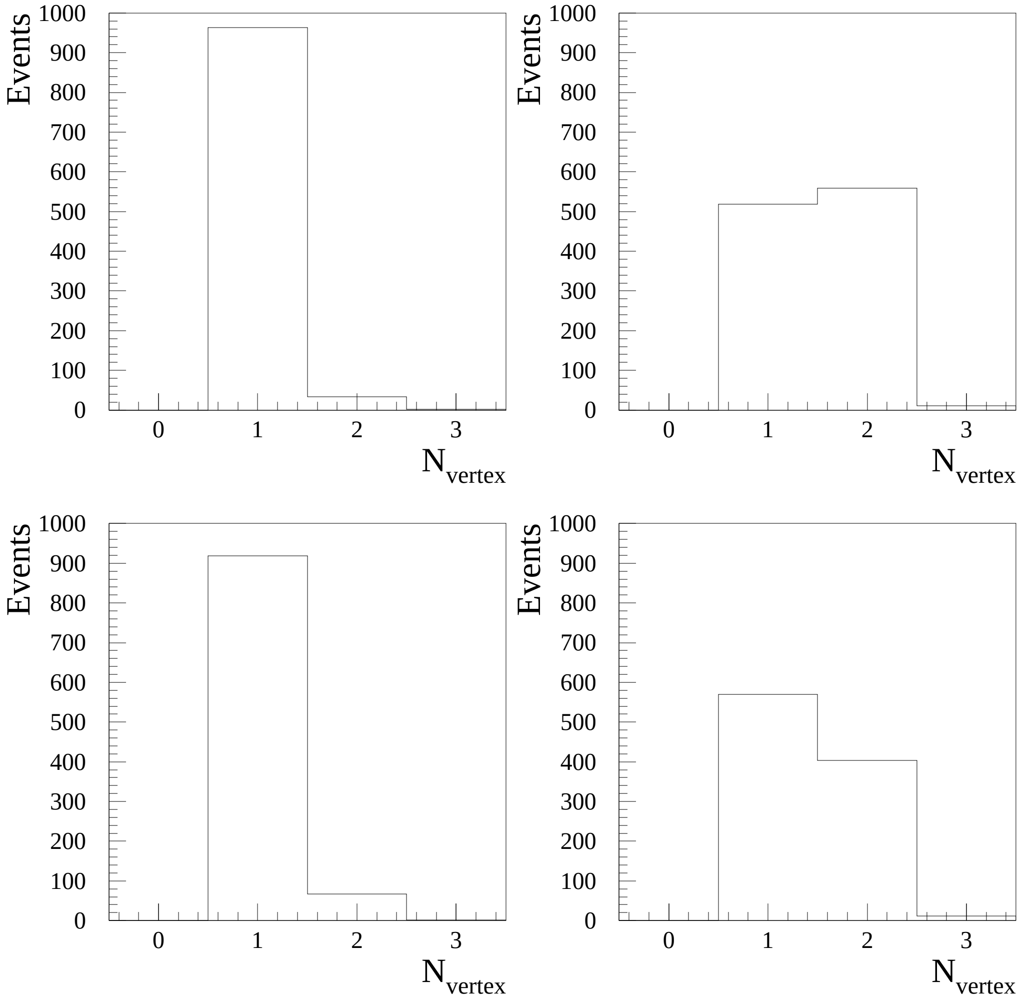

Figure C.18: Number of vertices for: (a)$Z \rightarrow \mu\mu$ (b)$Z \rightarrow \mu\mu + 1$ MBR (c) top(180 GeV) (d) top(180 GeV) + 1MBR.

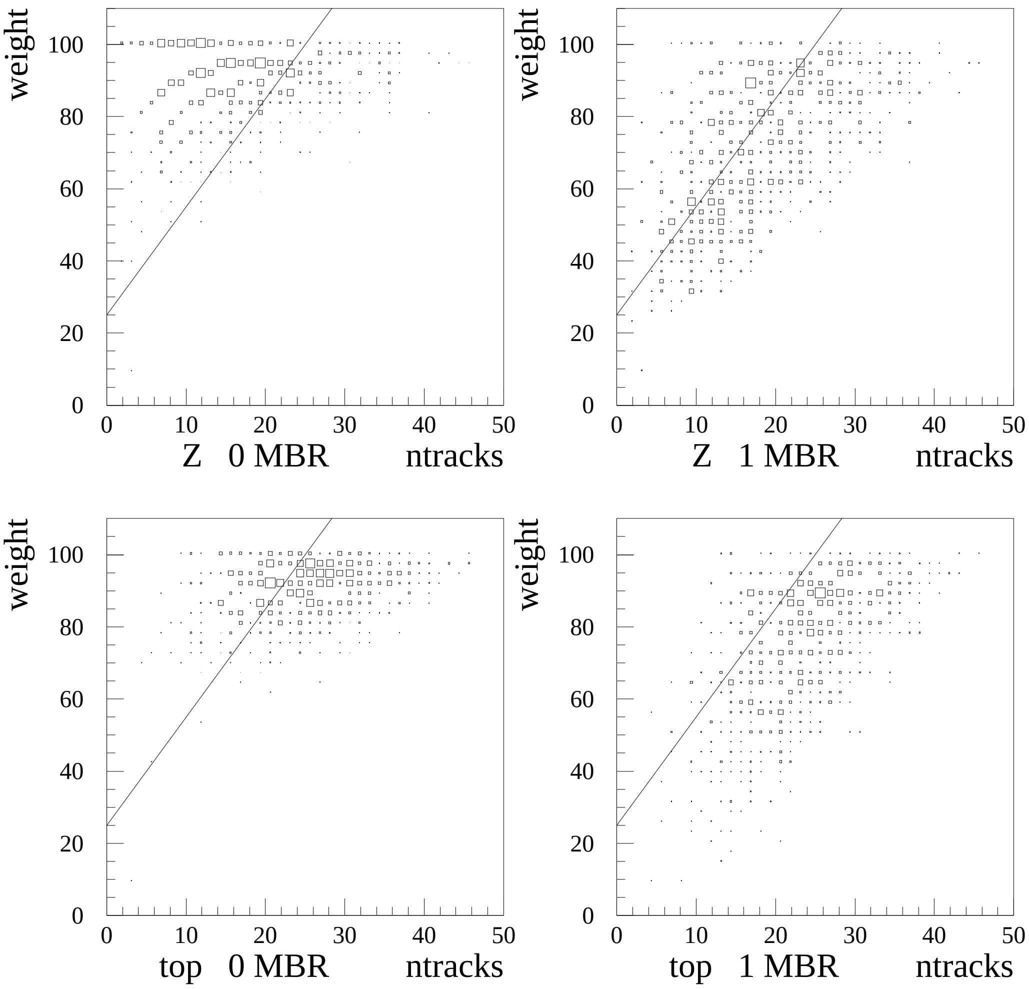

Figure C.19: Weight(percentage of tracks used) vs. number of tracks in primary vertex for : (a)$Z \rightarrow \mu\mu$ (b)$Z \rightarrow \mu\mu + 1$ MBR (c) top(180 GeV) (d) top(180 GeV) + 1MBR.

C.8 Multiple Interaction Tool Algorithm

The new multiple interaction tool uses a hybrid algorithm (see Fig. C.20) that bases its decision either on standard variable cutting criteria or on a weighting scheme. The latter is a new feature introduced in an attempt to improve the tool capability for distinguishing single and multiple interaction events, throughout the wide luminosity range covered by Run 1B data. This algorithm returns a value for the tool ranging from -1 to 5, as listed below:

- **-1** Likely a **S**ingle **I**nteraction but no vertex found by CD detector;
- **0** No interaction;
- **1** Most likely a **S**ingle **I**nteraction;
- **2** Likely a **S**ingle **I**nteraction;
- **3** Likely a **M**ultiple **I**nteraction;
- **4** Most likely a **M**ultiple **I**nteraction;
- **5** Likely three or more interactions.

Comparing to the old tool, two more values, namely -1 and 5, have been added.

In the sections C.2, C.3 and C.4, we discussed several variables which can help distinguish single and multiple interaction events. Weights are assigned to these variables as listed in Table C.8.1. These variables carry information from different pieces of detector and the way they are put together is described in the next paragraph and diagrams. However, there are a few terms used in the diagrams that need to be defined. GOOD INTERACTION requires the Level Ø interaction bit be set and requires a good SLOWZ flag. ETOT and EECN/EECS stand for the total energy measured by the calorimeter and the energy deposited in the end calorimeters(North/South), respectively. And in the Fig. C.22, VTX_WGT, FLG_WGT, ETOT_WGT, ECE_WGT, DELZ_WGT and

TRK_WGT represent the weights given to the variables in the diamonds.

As shown in Fig. C.20, the tool first tests the event for some requirements. These requirements can be seen in Fig. C.21, which represents the first part of the algorithm. If a condition inside the diamonds is satisfied, then a tool value is immediately assigned to the event. Nevertheless, if none of the conditions is matched, then the second part of the algorithm, shown in Fig. C.22, is reached. In Fig. C.22, the variables inside the diamonds are checked and given a weight, represented by the balloons just below. These weights are then summed up with relative weights ranging from 1/18 to 1/3, thus giving the total event weight, represented by the last balloon. Based on this weight, a tool value from 2 to 4 is assigned to the event. At this point, a tool value between -1 and 4 has already been given to the event. If tool is equal to four, one further test is performed in order to identify candidates which could be more than two interactions (tool=5). This is illustrated by the last diamonds in Figs. C.21 and C.22.

C.8.1 Weights

The weights were assigned in a nearly arbitrary way, only subject to the rule that variable values pointing to probably a single interaction were given a low weight while the values pointing to a multiple interaction were given high weights. Following this rule we divided the distributions in a few regions, assigning different weights(usually ranging from 0.2 to 0.9) for each region. The weights were tuned by studying the resulting behavior of the tool when each weight was set in a range of values. In such a scheme, the single interaction events are concentrated in the lower total weight values and the multiple ones in the upper values. The boundaries among the tool values are also chosen in a arbitrary way trying to represent as close as possible the theoretical expectations. The weights are listed in Table C.8.1.

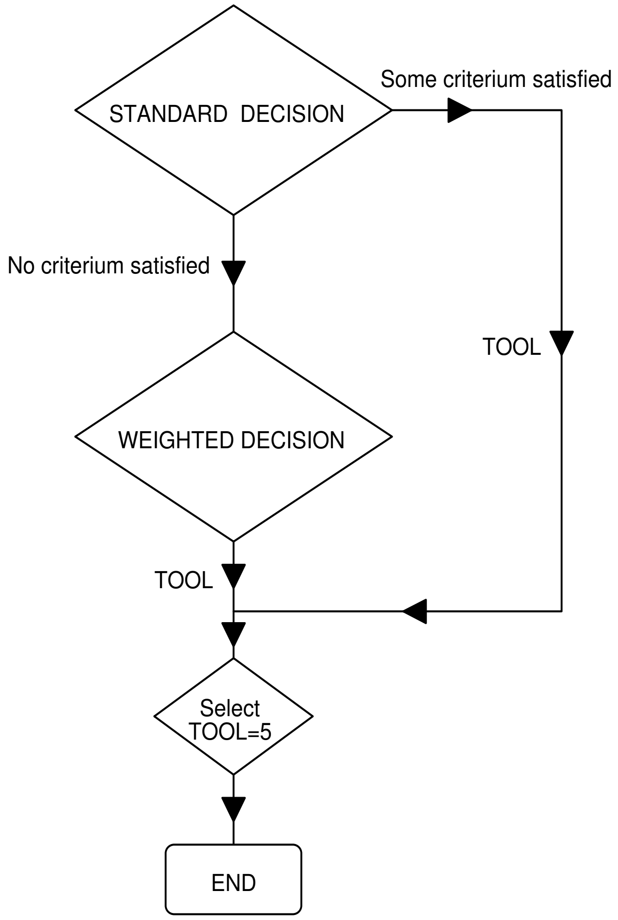

Figure C.20: New Multiple Interaction Tool Algorithm.

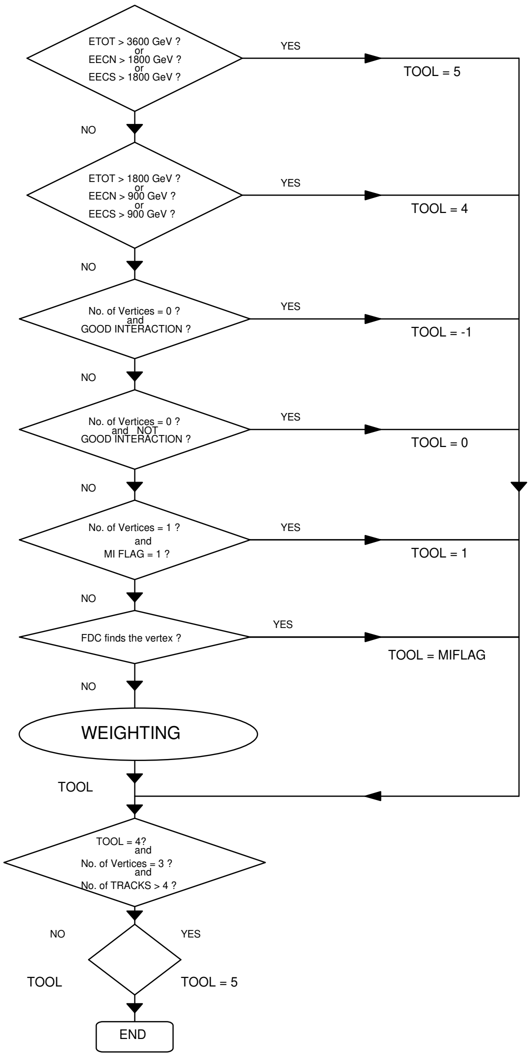

Figure C.21: The Standard Algorithm Part.

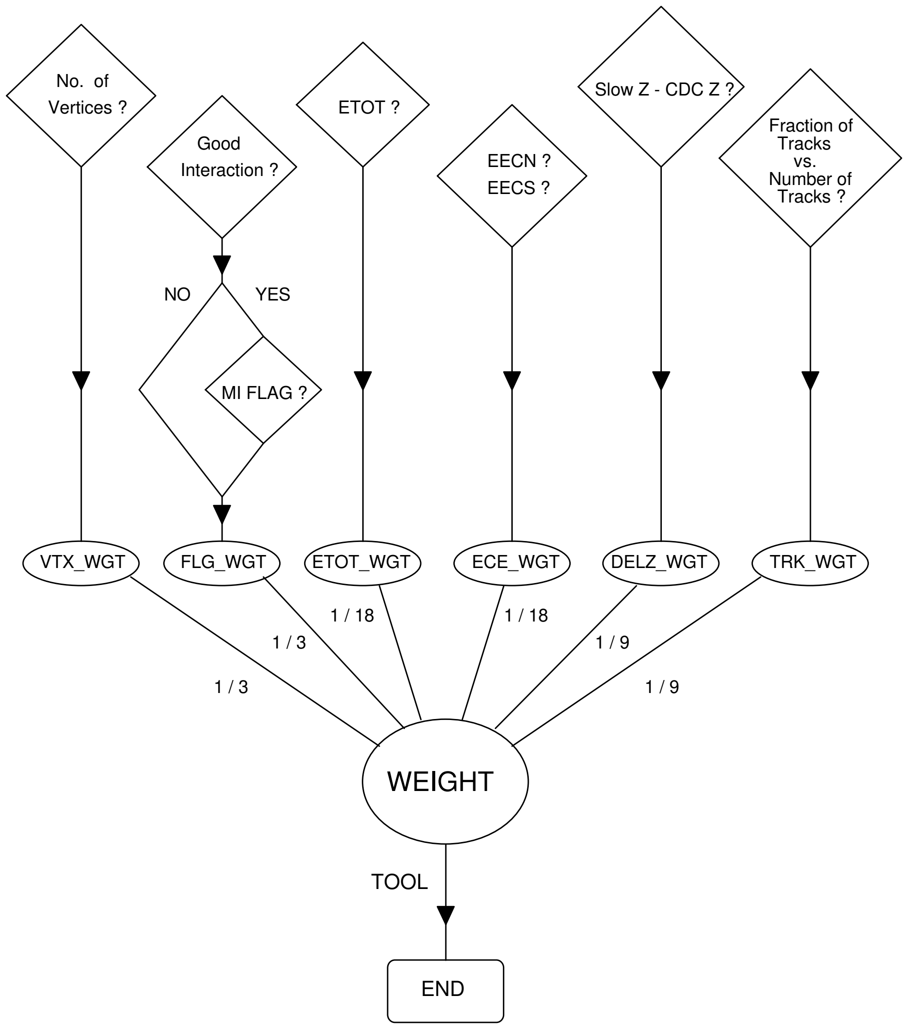

Figure C.22: The Weighting Algorithm Part.

VARIABLE	WEIGHT
Multiple Interaction Flag	
1	0.2
2	0.4
3	0.6
4	0.85
Number of Vertices	
1	0.2
2	0.6
3	0.9
others	0.99
| SLOWZ - CDZ | (cm)	
< 5	0.2
5 to 10	0.6
≥ 10	0.9
Total Energy (GeV)	
< 1000	0.2
1000 to 1500	0.65
≥ 1500	0.9
End Cap Energy (GeV)	
ECN < 600 or ECS < 600	0.2
$600 \leq$ ECN < 800 or $600 \leq$ ECS < 800	0.65
$800 \leq$ ECN < 900 or $800 \leq$ ECS < 900	0.9
Fraction of Tracks vs. Number of Tracks (1st Vertex)	
CDFRAC - 3 $\times$ CDNTRACK ≥ 25	0.2
CDFRAC - 3 $\times$ CDNTRACK < 25	0.8

Table C.1: Weights

C.9 Conclusions

The multiple interaction tool has been optimized for the higher luminosities of Run 1b. The new tool works well at identifying single interactions up to luminosities of about 20E30, then becomes increasingly contaminated as the luminosity increases. The efficiency for identifying single interactions and the contamination from multiple interactions appear to depend on the data set used. Therefore, analyses which use the multiple interaction tool should study its effect on their own filter selections. The new multiple interaction tool can be used equally well for all data from Run 1. The new multiple interaction tool will, however, produce different results than the Run 1a tool. The multiple interaction tool relies heavily on the Level Ø and central detector results. The Level Ø result is unchangeable, but the central detector result could theoretically be improved with better vertex finding.

Bibliography

[1] W.K.H. Panofsky, in proceedings of *14th International Conference on High Energy Physics*, edited by J. Prentki and J. Steinberger (CERN, Geneva, 1968).
E.D. Bloom *et al.*, *Phys. Rev. Lett.* **23**, 930 (1969).
M. Breidenbach *et al.*, *Phys. Rev. Lett.* **23**, 935 (1969).

[2] J.D. Bjorken, *Phys. Rev. Lett.* **16**, 408 (1966).
J.D. Bjorken, *Phys. Rev.* **163**, 1767 (1967).
K. Gottfried, *Phys. Rev. Lett.* **18**, 1174 (1967).
R.P. Feynman, *Phys. Rev. Lett.* **23**, 1415 (1969).

[3] H.L. Lai *et al.*, *Phys. Rev. D* **55**, 1280, (1997).

[4] A.D. Martin, R.G. Roberts, and W.J. Stirling, *Phys. Lett.* B **356**, 89 (1995).

[5] R.K. Ellis, W.J. Stirling, and B.R. Webber, *QCD and Collider Physics*, (Cambridge University Press, Cambridge, 1996).

[6] J.D. Bjorken, *Phys. Rev.* D **47**, 101 (1992).

[7] D. Zeppenfeld, MADPH-95-933 (1996).

[8] E. Gotsman, E.M. Levin, and U. Maor, *Phys. Lett.* B **309**, 199 (1993).

[9] R.S. Fletcher and T. Stelzer, *Phys. Rev.* D **48**, 5162 (1993).

[10] H.N. Chehime *et al.*, *Phys. Lett. B* **286**, 397 (1992).

[11] B. May, Ph. D. thesis, University of Arizona, 1994 (unpublished).

[12] DØ Collaboration, S. Abachi *et al.*, *Phys. Rev. Lett.* **72**, 2332 (1994).

[13] CDF Collaboration, F. Abe *et al.*, *Phys. Rev. Lett.* **74**, 855 (1995).

[14] DØ Collaboration, S. Abachi *et al.*, *Phys. Rev. Lett.* **76**, 734 (1996).

[15] J. Perkins, Ph.D. thesis, University of Texas, Arlington (in progress).

[16] G. Alner *et al.*, (UA5 Collaboration), *Phys. Rep.* **154**, 247 (1987).

[17] A. Giovannini and L. Van Hove, *Z. Phys.*C **30**, 391 (1986).

[18] P. Abreu *et al.*, (DELPHI Collaboration), *Z. Phys.*C **52**, 271 (1991); *Z. Phys.* C **56**, 63 (1992).

[19] A.D. Martin and C.J. Maxwell, *Z. Phys.*C 34, 71 (1987).

[20] C.S. Kim, *Z. Phys.* C **51**, 619 (1991).

[21] V. Del Duca and W.K. Tang, *Phys. Lett.* B **312**, 225 (1993).

[22] O.J.P. Éboli, E.M. Gregores, and F. Halzen, MADPH-96-965 (1996).

[23] C.D. Carone and H. Murayama, *Phys. Rev. Lett.* **74**, 3122 (1995).

[24] B.L. Combridge and C.J. Maxwell, *Nucl. Phys. B* **239**, 429 (1984).

[25] J. Perkins at *5th International Workshop on Deep Inelastic Scattering and QCD*, Chicago, Illinois, April 14-18, 1997, (proceedings to be published).

[26] DØ Collaboration, S. Abachi *et al.*, *Nucl. Instrum. Methods* A **338**, 185 (1994).

[27] J. Kotcher, Internal DØNote #2417 (1995).

[28] P. Rubinov, private communication.

[29] R. Kehoe and R. Astur, Internal DØNote #2908 (1996).

[30] B. Abbott *et al.*, Internal DØNote, in preparation.

[31] M. Bhattacharjee *et al.*, Internal DØNote #2887 (1996).

[32] C. Kim, T. Heuring, and S.Y. Jun, Internal DØNote #2477 (1995).

[33] K. Frame, K. Fatyga, Internal DØNote #3178 (1997).

[34] H.U. Bengtsoon and T. Sjöstrand, *Comput. Phys. Commun.* **46**, 43 (1987).

[35] D. Zeppenfeld, private communication.

[36] V. Del Duca, at the *Third Workshop on Small-x and Diffractive Physics*, Argonne, Illinois, September 26-29, 1996.

[37] F. Halzen, private communication.

[38] J. Bantly, Q. Li-Demarteau, 'Multiple Interaction Tool', Internal DØNote 1691.

[39] J. Bantly, 'Multiple Interaction Tool Study', Internal DØNote 1849.

[40] J. Bantly et al., 'Improvement to the DØ Luminosity Monitor Constant', Fermilab-TM-1930 (DØNote 2544).