RECONSTRUCTION AND ATTRIBUTES OF JETS OBSERVED IN $\sqrt{s}$ = 200 GEV PROTON-PROTON AND DEUTERON-GOLD COLLISIONS AT STAR

A Dissertation

by

THOMAS WILLIAM HENRY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2006

Major Subject: Physics

RECONSTRUCTION AND ATTRIBUTES OF JETS OBSERVED IN $\sqrt{s}$ = 200 GEV PROTON-PROTON AND DEUTERON-GOLD COLLISIONS AT STAR

A Dissertation

by

THOMAS WILLIAM HENRY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Carl Gagliardi
Committee Members,	Che-Ming Ko
	Saskia Mioduszewski
	Joseph Natowitz
	Robert Tribble
Head of Department,	Edward S. Fry

December 2006

Major Subject: Physics

ABSTRACT

Reconstruction and Attributes of Jets Observed in $\sqrt{s}$ = 200 GeV Proton-Proton and Deuteron-Gold Collisions at STAR. (December 2006)

Thomas William Henry, B.S., Colorado State University

Chair of Advisory Committee: Dr. Carl Gagliardi

The STAR collaboration recorded $\sqrt{s}$ =200 GeV per nucleon d+Au and p+p collision events during the year 2003 RHIC run. In the p+p and the d+Au data samples, it is possible to reconstruct jets and make comparisons between them. This dissertation describes the reconstructed jet sample from the p+p events, the measurements of the jet j_T distribution which quantifies the shape of the jet perpendicular to the jet direction, the jet fragmentation function (z), which quantifies the fraction of jet momentum carried by the jet particles, and the width of the parton momentum broadening distribution (k_T). This dissertation also describes the comparison of these results to the reconstructed jet sample from d+Au events. Measurements of jet j_T and jet fragmentation from p+p, d+Au, and PYTHIA are compared. The z and j_T distributions from p+p and also d+Au are found to be consistent with PYTHIA event simulation version 6.205. RMS(j_T) equals 612±12±30 MeV/c for p+p, and RMS(j_T) equals 630±13±30 MeV/c for d+Au. The p+p k_T Gaussian sigma (width) equals 2.08±0.12±0.13 GeV/c. This width is consistent with PYTHIA, k_T literature surveys, and other RHIC measurements. A limit on nuclear k_T broadening in the Au nucleus is also obtained, showing that nuclear k_T is consistent with zero to within 0.5 GeV/c. The present status of the R_{dAu} measurement from jets is also described. The nuclear k_T broadening, in particular, is a measurement which has been done at this energy for the first time at RHIC. These measurements, while interesting by themselves, also will contribute ultimately to the understanding of Au+Au collisions at RHIC, and the properties of the high-density matter which results from them.

DEDICATION

To Him who established the laws of nature and hid them for men to find, to those who do His will as they study and reveal the foundations of His universe, and to my parents who taught me how to learn.

ACKNOWLEDGEMENTS

Special thanks to Carl Gagliardi for all his help, guidance and support during my research. Thanks also to Bob Tribble and the other members of my committee for their help, suggestions, and sometimes difficult questions. Thanks to my parents for all their encouragement. Thanks also to the STAR collaboration and the EEMC group for giving me the opportunity help build the detector and obtain access to the wonderful storehouse of data recorded by STAR. Thanks to Michael Miller for all his work on the Jet Finder code, to Renee Fatemi for all her help in providing examples of how to execute programs on the farm, and to Akio Ogawa and Les Bland for pointers, input, and direction at the very beginning of my research. Thanks to everyone at the Cyclotron for making it a collegial and pleasant environment for myself and the other graduate students. Last but not least, thanks to Mary Anne Batson for helping me keep up with all the paperwork.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

I. INTRODUCTION

Physics, at its most basic level, is the struggle to understand the universe in terms of scientific observations. The simple and intuitive physical aspects of nature were understood long ago. Only in the last couple centuries did scientists succeed in describing nature's simple but unintuitive aspects. The purview of modern scientific endeavor is the study of the many remaining things that are not so simple.

Though at one time it was believed that protons and other nuclei were elementary, over the years we have learned that these systems are not only composite but very difficult to describe in closed theoretical formalisms. The energies of 200 GeV per nucleon used at the Relativistic Heavy-Ion Collider (RHIC) probe these systems at the level of partons, while the nuclear structures of the colliding particles enter in as initial boundary conditions. As the collision evolves, partons are modeled as being selected at random from a distribution which depends on the momentum of the parton and the energy transferred. The parton distribution functions have broad applicability.

It is not clear to what extent the parton distribution functions are more than just parameterizations. Each subatomic particle possesses unique parton distribution functions (pdfs), but the pdfs can be interpreted as parton probability densities[1]. It is therefore desirable to measure parton distribution functions as completely as possible, so that in the future connections can be made between different pdfs to advance our understanding. For well known pdfs, such as for nucleons, it may be even more important to obtain the perturbations of pdfs due to the presence of external potentials. The dependence of the pdf on external potentials will help determine the dependence of the parton distribution functions on parton binding energy. This will constrain theory in a fundamentally new way. RHIC provides the opportunity to characterize the dependence on the external nuclear potentials of the pdfs in the region accessed by 200 GeV per nucleon collisions.

Measuring parton properties is itself a complex endeavor. Free partons do not exist in nature on the scale of the detector, since at small scales other normally short-lived

---

This dissertation follows the style of Physical Review C.

vacuum quark pairs exist to bind with them. When the parton collides with a short-lived virtual quark pair, it is absorbed if it is a gluon, and if it is a quark it binds to form a comparatively stable composite particle. The virtual quarks bind or become free quarks, and these particles all share the parton's original momentum. Bound, comparatively stable particles continue on to reach the detector, but the free quarks bind and scatter with other virtual quark pairs. At each step, the free quarks have less momentum than before, and ultimately have so little momentum that they may bind with the remnants of the original nucleon from which the original parton was taken. This process is referred to as fragmentation. Although the partons are very interesting, experiments must instead measure the composite particles that arise during fragmentation. Since all the composite particles carry and share the momentum of the original parton, there are strong momentum correlations between them. All these composite particles created during fragmentation are called, collectively, a jet. The summed momentum of all the particles making up a jet, referred to as the jet momentum, is a satisfactory proxy for the momentum of the original parton.

Measuring the dependence of the pdf on the nuclear potential is complex. Jets are a valuable tool to accomplish this task. Therefore, it is important to understand jet properties in p+p and d+Au systems. That is the subject of this dissertation, and it forms the basis for measuring R_{dAu} for jets, which is defined as:

$$R_{dAu} \equiv \frac{\frac{\sigma_{d+Au}^{jet}}{\sigma_{d+Au}^{hadronic}}}{N_{bin}\frac{\sigma_{p+p}^{jet}}{\sigma_{p+p}^{inelastic}}}. \qquad (1)$$

Here, N_{bin} is the average number of nucleon-nucleon collisions given that a d+Au collision occurred. R_{dAu} will have to be measured to sufficient precision that the sub-processes (quark and gluon jets) can be distinguished. By distinguishing each sub-process, the fractional momentum distributions (x) of the partons will be directly related to the pdfs, the angle of the jet to the beam axis, and the jet transverse energy (E_T). The sub-process R_{dAu} measurement, associated with a specific x distribution defined by the angle to the

beam and E_T, will constrain the d+Au/p+p pdf ratios. The departure of the ratios from unity will arise because of the dependencies of the pdfs on the nuclear potential.

This dissertation also bears on the interpretation of di-hadron analyses. Di-hadron analyses are commonly done in the densely populated heavy-ion collisions, and also in cases where the detector has limited acceptance. These are cases where either too many or too few particles are present, preventing the jet finding algorithm from operating correctly. It is difficult to measure fragmentation properties of jets using two-particle correlations alone. Thus, it is useful to characterize the jet properties in order to calibrate the di-hadron measurements. Jet properties are typically characterized by two nearly independent variables, z and j_T. z is essentially a measurement of the longitudinal fragmentation in the jet, and is defined as $z \equiv \frac{p_h}{p_j}$, where p_h is the momentum of a hadron and p_j is the momentum of the jet. j_T is a complementary measurement of the transverse fragmentation of the jet, defined as $j_T \equiv \sqrt{p_h^2 - \left(\vec{p}_h \cdot \vec{p}_j / p_j\right)^2}$, where the arrows indicate momentum vectors. The z and j_T detector resolution effects will be shared by both full jet reconstruction and di-hadron analysis, with possible systematic uncertainties arising from jet reconstruction inefficiencies. Once the detector effects on z and j_T have been quantified, simulations can be used to predict the average z and j_T quantities for the various di-hadron analysis cuts.

Much recent progress has been made in measuring di-jets in Au+Au collisions at the Solenoidal Tracker At RHIC (STAR)[2]. Collider experiments are usually characterized by the center of mass energy, $\sqrt{s}$. Many experiments have been performed with p+p collisions both below and above the center of mass energy at which RHIC operates, $\sqrt{s} = 200$ GeV per nucleon[3]. RHIC is the first experiment to measure proton+nucleon (p+N), or deuteron+nuclon (d+N) collisions above $\sqrt{s} = 40$ GeV. It is important to quantify the effect of the gold nucleus on k_T, the transverse momentum of a jet pair. If the hard scattered partons collectively recoil against a gold nucleus differently

than they do against a proton, the interpretation of the Au+Au di-jet results become considerably more difficult. By measuring nuclear k_T, this dissertation provides important support for interpreting the k_T broadening measured in Au+Au as a "medium" effect[4].

A. Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the coupling between quarks and the SU(3) color field embodied by quanta called gluons. Quarks (antiquarks) carry a positive (negative) color charge, while gluons carry one positive and one negative color charge. Unlike the electric charge, which is a simple scalar, the color charge is a vector coming from the SU(3) group. Since gluons carry charge, the color fields constituted by them interact with other color fields. This makes it very difficult, in general, to calculate QCD wave-functions and expectation values.

The Lagrangian action formalism is one of the most advanced and abstract methods for obtaining equations of motion in classical mechanics and electromagnetism, but for QCD (and also QED) this formalism is the basic starting point. That such a powerful method (or its equivalent) is *required* for QCD is a measure of just how far theoretical physics has come since the days of Newton. The QCD Lagrangian, with color, flavor (quark), and spin indices suppressed, has the form:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \overline{\psi}\left(i\gamma_\mu D^\mu - m\right)\psi . \qquad (2)$$

F denotes the gluon field tensor, ψ the quark field, and D is the covariant derivative. To solve this equation, special boundary conditions, approximations, or both are required.

One of the most important properties of the QCD theory is color confinement. Qualitatively speaking, the strength of the color field does not fall as the distance between charges grows. Even at short distances, the potential energy of the color field becomes great enough to induce quark pair production in the vacuum. This makes it impossible to probe color fields or work with quarks directly. The difficulty of many QCD calculations reflects the difficulty in performing macroscopic experiments on quarks and gluons. QCD

describes an aspect of nature which is nonlinear, self interacting, difficult to picture and complex.

In later sections, quarks, gluons and partons will be discussed in the same context, but this dissertation makes no distinction between the partons used to explain experimental results and the quarks and gluons of QCD theory. The support for this assumption is necessarily indirect, but the identification is a very fruitful one. It is well beyond the scope of this dissertation to question the assumption that partons *are* quarks and gluons.

B. Perturbative Quantum Chromodynamics

For high-energy experiments, it is important to make theoretical predictions based on QCD. Within this restricted context, it is possible to use the Lagrangian equations with the specified boundary conditions; the initial conditions of the incident particles in the colliding beams are well defined. Under these circumstances, QCD can be approached perturbatively. If a parton from one nucleon collides with a parton from another nucleon with sufficient energy, the initial dynamics of the collision will follow the classical two-body elastic scattering to first order. On a short enough time scale, the magnitude of the momentum transfer is much greater than the magnitude of the change in potential energy of the parton as it escapes the nucleon.

Mathematically speaking, these circumstances are analogous to Quantum Electrodynamical calculations, indicating that a perturbative approach like that used to do calculations in QED should be effective. The Feynman diagrams, which can be deduced from the Lagrangian, can also be applied to QCD perturbative calculations given the correct vertex and propagator formulas.

1. Asymptotic Freedom

Vacuum color "polarization" results in an effective amplification of a color charge, sometimes called *antiscreening*. Thus, the color force becomes arbitrarily weak in the

limit of short distances or large energy scales. The asymptotic reduction in the value of the strong coupling constant in these limits is called asymptotic freedom [5].

Asymptotic Freedom leads to the formation of jets. Partons can scatter with a large momentum transverse to the collision axis while the remaining partons obtain little transverse momentum. Momentum conservation will make hard scattered partonic momentum observable macroscopically.

2. Factorization of Collision Dynamics

The cross section of obtaining a certain hadron is the natural quantity to compare with experiments. The physical process of beam nucleon + beam nucleon $\rightarrow$ hadrons can be factored into two pieces. First, beam nucleon + beam nucleon $\rightarrow$ parton + parton, and then parton $\rightarrow$ hadrons. The Feynman diagrams for the asymptotically free parton reaction are tractable, and this is included in the first step.

a) Parton Distribution Functions

The cross section of parton + parton scattering is also naturally factored into two pieces. One factor is the parton scattering cross section, which can be calculated by applying perturbative techniques to QCD (PQCD). The other factor is the probability of finding a specific parton with a fraction of the nucleon momentum, *x*. The formula can be written like so [6]:

$$\sigma_{ij\rightarrow kX} = \sum_{i,j} \int dx_1 dx_2 d\hat{t} f_i^1(x_1, Q^2) f_j^2(x_2, Q^2) \frac{d\hat{\sigma}_{ij->kX}}{d\hat{t}}, \qquad (3)$$

where Q^2 is the parton momentum transfer, and $\hat{t}$ is a Mandelstam variable defining the kinematics of the partonic cross section $\hat{\sigma}_{ij->k}$ producing the parton *k* given a scattering event between partons *i* and *j*. The parton distribution functions, $f_i(x, Q^2)$, express the probability that a parton *i* contains a fraction *x* of the nucleon momentum. Like fragmentation functions described in the next subsection, parton distribution function dependence on $\ln(Q^2)$ can be obtained from the DGLAP equations [7]:

$$Q^2 \frac{\partial}{\partial \ln Q^2} f_i\left(x, Q^2\right) = \int_x^1 dz \frac{\alpha_s}{2\pi} P_{ii}\left(z, Q^2\right) f_i\left(\frac{x}{z}, Q^2\right), \qquad (4)$$

where α_S is the strong coupling constant, and P_{ii} is the probability of scattering off a parton *i* with momentum transfer Q^2, given that a parton *i* has momentum fraction z.

The parton distribution functions themselves are very difficult to calculate from QCD first principles. No one has solved for them analytically with realistic approximations. In QED, or any theory where the coupling constant is small, more and more complicated interactions are more and more improbable, allowing calculations to neglect particle interaction graphs above a specific complexity and to achieve an approximation. In QCD, however, the coupling constant is strong, and in general more and more complicated parton interactions are not more improbable. This makes PQCD only applicable to specific situations where one particular set of particle interactions dominate, such as a hard parton-parton scatter. Another approach, called lattice QCD bounds the complexity of interactions spatially. Unfortunately, this is extremely computationally difficult, and computers have not yet been used to do thorough pdf calculations. The pdfs must therefore be experimentally determined. Major non-polarized data sets can be obtained from CTEQ [1], GRV [8], and MRST [9].

b) Fragmentation Functions

Equation 3 gives the cross section for obtaining a given parton from a nucleon collision. Due to color confinement, the parton will be forced to neutralize its color charge by combining with other partons to form color neutral hadrons. This process is called fragmentation. Like the probability of finding a parton in a nucleon, the probability of "finding a hadron in a parton" is also a very difficult calculation to perform. The fragmentation functions are very similar to the pdfs in many ways, and PQCD cannot be applied to deduce fragmentation functions either. Neither have lattice QCD calculations been successfully performed to date. In this case also, experiments are used to determine the shape of the fragmentation functions.

Like the pdf, the fragmentation function depends on both z and Q^2. Theoretically, the fragmentation function dependence on $\ln(Q^2)$ can also be deduced [7]:

$$\frac{\partial}{\partial \ln Q^2} D_i^h\left(z, Q^2\right) = \sum_j \int_x^1 \frac{dy}{y} \frac{\alpha_s}{4\pi} P_{ji}\left(\frac{z}{y}, Q^2\right) D_i^h\left(y, Q^2\right) \qquad (5)$$

Where α_S is the strong coupling constant, P_{ji} is the probability of finding a parton j "in" the parton i, and D_i^h is the probability of producing a hadron h from parton i. This dependence is important when comparing experiments done at different collision energies. Software simulation tools like PYTHIA[†] [10] utilize such DGLAP equations to produce consistent results for comparison with experiments at different energies.

Many experiments have been done at various $\sqrt{s}$ values in what are called deep inelastic scattering (DIS), hadron-hadron, and e^+-e^- experiments. DIS experiments involve the scattering of leptons off hadrons, and are in many ways simpler to interpret than the hadron-hadron collisions. e^+-e^- annihilation has been used to provide the cleanest measurements of fragmentation in a vacuum, but at higher $\sqrt{s}$ the DIS experiments have also provided valuable measurements. Figure 1 shows the trends in the fragmentation functions from DIS and e^+-e^- experiments as $\sqrt{s}$ changes [11]. The variable x shown in Figure 1 is equivalent to z as defined in this dissertation[‡].

[†] See http://www.thep.lu.se/~torbjorn/Pythia.html.

[‡] When discussing hadron-hadron collisions, x is commonly used to denote the momentum fraction of the hadron carried by a daughter parton. z is therefore used to denote the fragmentation function to avoid confusion.

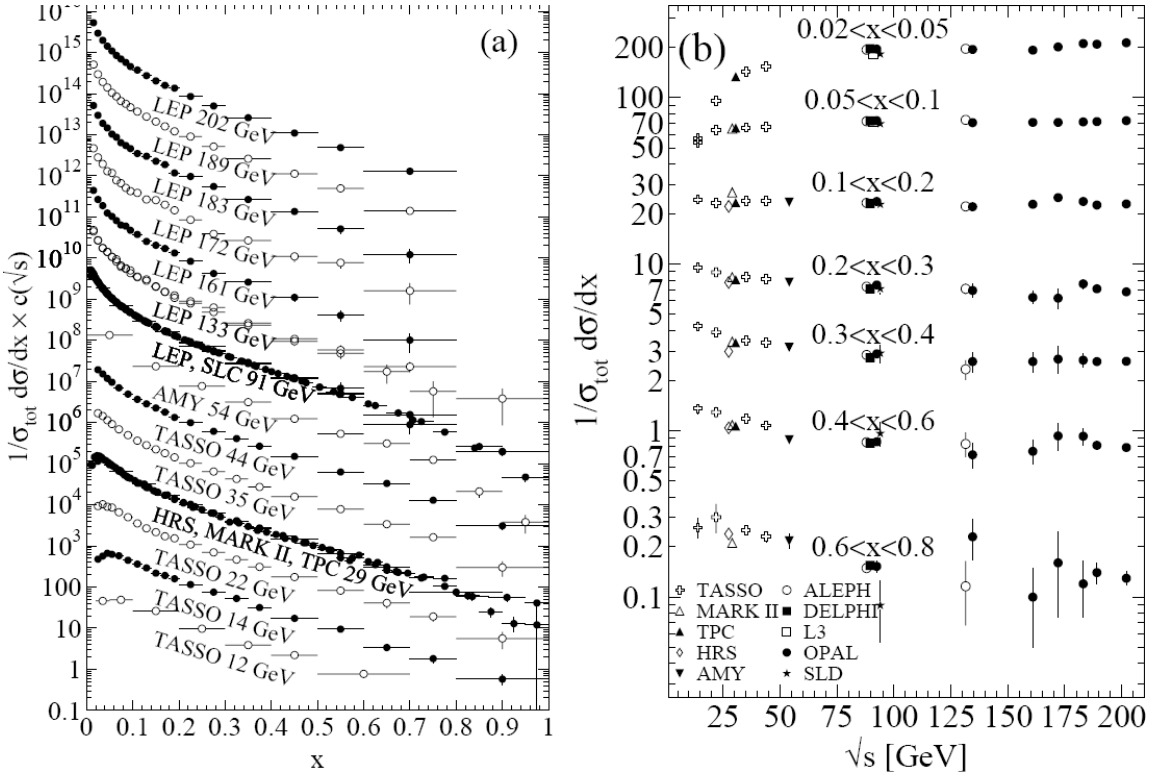

Figure 1. e+-e− fragmentation world data.
e^+-e^- fragmentation functions for all particles are shown (a) for different c.m. energies, $\sqrt{s}$, versus x and (b) for various ranges of x versus $\sqrt{s}$ [11]. For the purpose of plotting (a), the distributions were scaled by $c(\sqrt{s}) = 10^i$ where i is ranging from $i = 0$ ($\sqrt{s}$ = 12 GeV) to $i = 13$ ($\sqrt{s}$ = 202 GeV).

.

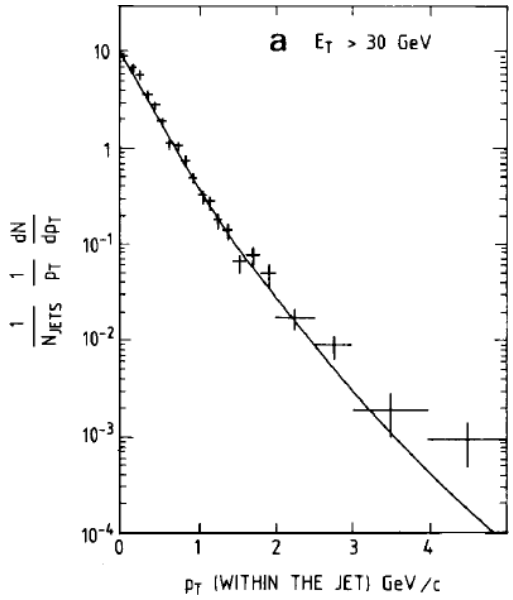

Figure 2. Jet j_T measured at CERN in 1983[12].

Even though the fragmentation of the parton is characterized by both z and j_T, only the z distribution is relevant for cross section measurements. As a consequence, references on the j_T distribution are not as common. The Review of Particle Physics has an entire section on the z distribution (the fragmentation function), but does not directly discuss the j_T distribution. Measurements of the j_T distribution, however, are certainly not new. Figure 2 shows jet j_T results from CERN produced in proton-antiproton collisions over 20 years ago [12], before this author had even reached the age of 10! Very recently, the PHENIX collaboration at RHIC has performed di-hadron j_T measurements which quantify the average jet j_T [13]. Experiments like these show that $\sqrt{\langle j_T^2 \rangle} \approx 600\,\text{MeV/c}$ at $\sqrt{s} \approx 200\,\text{GeV}$. Although there is a slight increase in j_T fluctuations as $\sqrt{s}$ increases, it is modest, and one expects to see something like 600±75 MeV/c for any DIS or hadron+hadron experiment.

C. Non-perturbative QCD and Intrinsic k_T

Perturbative QCD is an approximation, and differences due to non-perturbative effects are easily found. Initial and final state soft gluon radiation effects, for example, are not easily calculated by perturbative techniques. Gluon radiation is commonly modeled by taking a partonic Gaussian k_T smearing (with σ_{kT} as a parameter), and then neglecting the gluon radiation in detail [14]. To illustrate this, a parton cross section schematic equation can be written which includes a Gaussian k_T smearing model:

$$\frac{d\sigma_{ij\to kX}}{dyd\phi} = \sum_{i,j}\int_0^1 dx_1 \int_0^1 dx_2 \int d\vec{p}_{Tk} \int d\vec{k}_T \frac{1}{\sqrt{2\pi}\sigma_{kT}} \exp\left(-\frac{k_T^2}{2\sigma_{kT}^2}\right)$$
$$\delta\left(\vec{p}_{Tk} - \vec{p}_T(x_A, x_B, y, \phi) + \vec{k}_T\right) f_i^1(x_1, Q^2) f_j^2(x_2, Q^2) \frac{d\hat{\sigma}_{ij\to kX}}{dyd\phi}, \quad (6)$$

where $\vec{p}_{Tk}$ is the momentum of the outgoing parton, $\vec{p}_T$ is the final state transverse momentum of the two outgoing partons in the absence of k_T smearing, and $\vec{k}_T$ is the vector k_T kick. Here *y* and ϕ are coordinates specifying the direction of the outgoing parton *k*. ϕ measures the angle of the transverse (to the beam line) direction of the parton. The rapidity, *y,* is given by:

$$y = \frac{1}{2}\ln\left(\frac{E + p_L}{E - p_L}\right) \quad (7)$$

in which *E* is the energy of the particle, and p_L is the longitudinal momentum of the particle along the beam line. The rapidity is often more useful then the simple angle, because rapidity differences are boost invariant along the beam axis.

It is often the case that the mass of a particle is small compared to the total momentum, and in many cases is unknown. Throughout this dissertation it is useful to make an approximation that $y \cong \eta$, where η is called the pseudorapidity, and is defined as:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \quad (8)$$

where θ is the angle between the particle trajectory and the beam axis. The pseudorapidity equals the rapidity in the limit where the momentum is much greater than the mass. Pseudorapidity differences also possess the boost invariance property along the beam axis.

Experiments have shown that to make equation 6 work, the intrinsic k_T must be significantly greater than the value obtained from the quantum mechanical uncertainty principle via the radius of the proton, i.e. ~200 MeV/c [14]. World data shows that the k_T of the model appears to scale with $\sqrt{s}$ [14] (see Figure on pg. 88).

D. Outline

The format of the remainder of this dissertation attempts to produce a logical progression to the results. First, the STAR experiment at RHIC is described. The description of STAR event reconstruction follows which explains the methods used to obtain charged and neutral particle trajectories. The Jet Finding algorithms are then described, along with how the particle data is prepared for jet reconstruction. The discussion moves to the explicit database source used for jet reconstruction, and then the general cuts on the data to improve quality are explained. Everything will then be in place to describe the measurements and observations. Finally, the greater context of the results presented in this dissertation will be discussed, concluding with how this research can be extended into the future.

II. EXPERIMENT

The study of QCD systems is a major motivator for modern physics experimentation and theory. Since PQCD can be applied mainly at high energies, much of our knowledge of QCD has come from collider experiments. Many experiments in the last 20 years have collided hadrons or leptons at high energy, providing experimental measurements of pdfs, fragmentation functions, and the QCD parameters which enter into PQCD calculations. Relationships between various measurements are constrained by QCD, and such tests of QCD have included the running of the QCD coupling constant, rare processes (such as Drell-Yan), jet cross sections, three-jet events, and heavy-quark production. Before the construction of RHIC, this list did not include high-$\sqrt{s}$ combined with high-particle polarization experiments. Physicists also have a great deal of interest in high-$\sqrt{s}$ heavy-ion collisions for the purpose of measuring the hypothesized quark gluon plasma. Measuring such a state of matter can be very useful for further understanding of QCD, since it is outside the better explored results of perturbation theory.

A. RHIC

The Relativistic Heavy-Ion Collider (RHIC) at Brookhaven National Laboratory was commissioned to conduct measurements of highly polarized proton-proton collisions, and heavy-ion collisions at energies of $\sqrt{s}$ = 200 GeV per nucleon-nucleon pair. It began operation in 2000, following 10 years of development and construction. RHIC drives two-counter moving beams of ions, steered with superconducting magnets, in two- beam pipes which are built to intersect at six points around the large RHIC ring. Figure 3 shows an aerial photo of the RHIC ring.

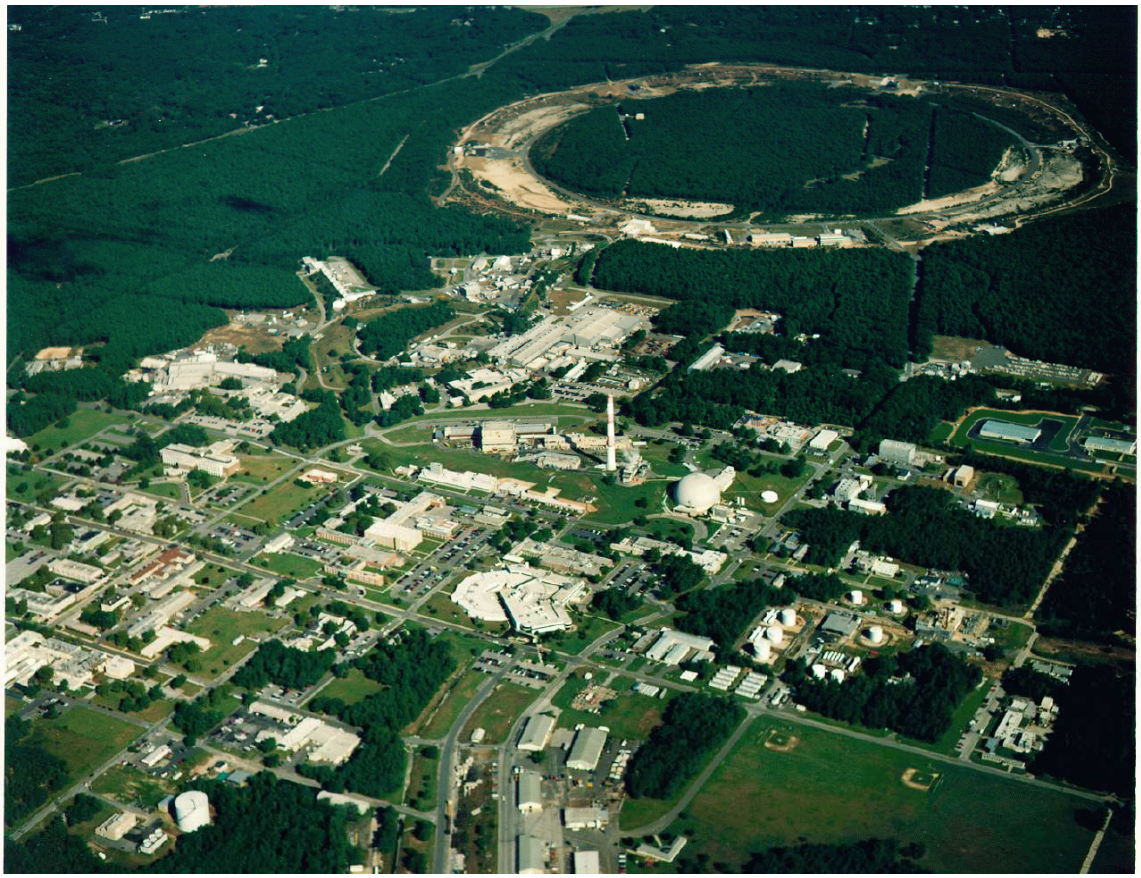

Figure 3. Aerial photo of the Brookhaven National Lab.
The RHIC ring is shown at the top of the photo[15].

B. The STAR Detector

The Solenoidal Tracker at RHIC (STAR) is one of the four original detectors constructed to take data from RHIC beams[16]. The others are PHENIX, BRAHMS, and PHOBOS. The STAR detector was designed to observe most charged and neutral particles produced in an event at mid-rapidity (particles with trajectories nearly perpendicular to the beam axis). Figure 4 illustrates several major components of STAR. STAR has a large conventional magnet which creates a uniform magnetic field of 0.5 Tesla. The magnetic field forces charged-particle trajectories to be helical with the radius

of the helix characteristic of the momentum of the particle. The charged-particle trajectories, and ultimately the particle momenta, are observed by the Time Projection Chamber (TPC) and Forward Time Projection Chamber (FTPC). After traversing the TPC, particles encounter the Time of Flight detector, which in combination with the momentum measurement can be used in many cases to identify the particles. The outermost detector of STAR is the Barrel Electromagnetic Calorimeter, designed to measure π^0, K_S, η, γ and other short lived neutral particles. STAR is capable of reconstructing the thousands of mid-rapidity particles produced in a single Au+Au collision event.

The primary goals of STAR are to search for signatures of quark-gluon plasma (QGP) formation in heavy-ion collisions, investigate the behavior of strongly interacting matter at high energy density, and measure gluon polarization using highly polarized beams of protons. STAR is able to correlate many observables on an event by event basis, making it suitable for measuring jets and other collective hadron signals. STAR consists of several types of detectors, each specializing in detecting certain types of particles or characterizing their motion. The data from these detectors are recorded simultaneously in order to obtain as much information as possible about each collision.

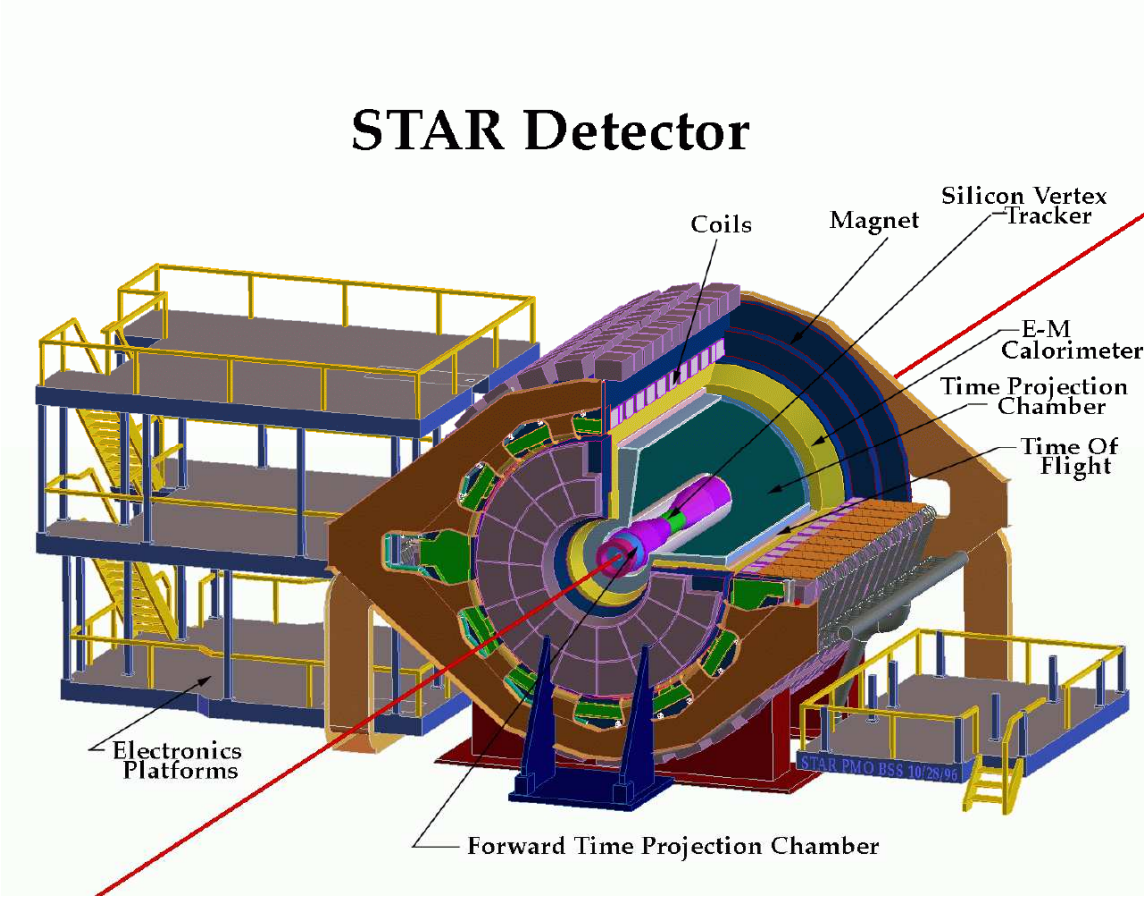

Figure 4. The STAR detector.
Several major subsystems are indicated in the diagram (see text for details)[15].

1. Trigger Detectors

The data acquisition system of STAR does not take data constantly, even when beams are present in the ring and physically interesting collisions are occurring. The rate at which information can be read and stored is limited, and so events must be prioritized. The common events are useful, but only a short time is required to obtain a sufficient sample. The more rare and interesting collisions are far more valuable, and a mechanism is needed to tell the data acquisition system when to read out and store data so that they are not lost. Several of the detectors built into STAR have the express purpose of detecting

the occurrence of an interesting collision. These detectors can be used to produce a trigger, a signal which is sent to the rest of STAR, which dictates that data be read out and recorded. In addition, trigger detectors can be used to monitor beam quality and collision rates, not only at STAR, but also at the main control of the RHIC ring. The primary trigger detectors used for this analysis are the Beam Beam Counter (BBC) and the Zero Degree Calorimeter (ZDC), used to produce minimum-bias triggers and monitor beam quality, and the Barrel Electromagnetic Calorimeter (BEMC) used to take what are called "high-tower" triggers. The minimum-bias trigger has a very high correlation with real ion collisions inside the STAR detector, and is designed to disfavor other events such as cosmic rays or collisions which occur elsewhere in the ring. The minimum-bias triggered samples are intended to make extrapolation from measured rates to cross sections straight forward. The minimum-bias sample is also valuable in that it is unbiased with respect to most physical processes. The "high-tower" trigger is useful for increasing the ratio of events with jets, since minimum-bias events do not contain large numbers of jet-like events. It is certainly biased with respect to physical processes, but the rarity of high-energy hadrons makes this trigger favor real ion collisions and select particularly interesting events.

a) Beam Beam Counter

The STAR Beam Beam Counter is a pair of scintillating detectors for counting charged-particle solid angle multiplicity. These detectors are mounted on the outside of the east and west poletips of the STAR magnet. Each BBC is made up of an inner and outer ring of scintillating material. The inner ring covers $3.9 < \eta < 5.0$, while the outer ring covers a pseudorapidity of $3.4 < \eta < 3.9$. Here the pseudorapidity, $\eta \equiv -\ln(\tan(\theta/2))$ where θ is the familiar polar angle defined with respect to the center of STAR. The inner ring and outer ring are tiled by large and small hexagonal scintillators, respectively, as shown in Figure 5. The large cells have exactly four times the dimensions of the small cells, and the small regular hexagons may be inscribed in a circle of diameter 9.64 cm. Like many scintillator based detectors, the response of the

detector is very fast, allowing it to count the total number of particles crossing the BBC plane very rapidly. Since the number of particles produced in a p+p collision is comparatively low, an east and west coincidence of the BBC detectors is used to produce a p+p minimum-bias trigger.

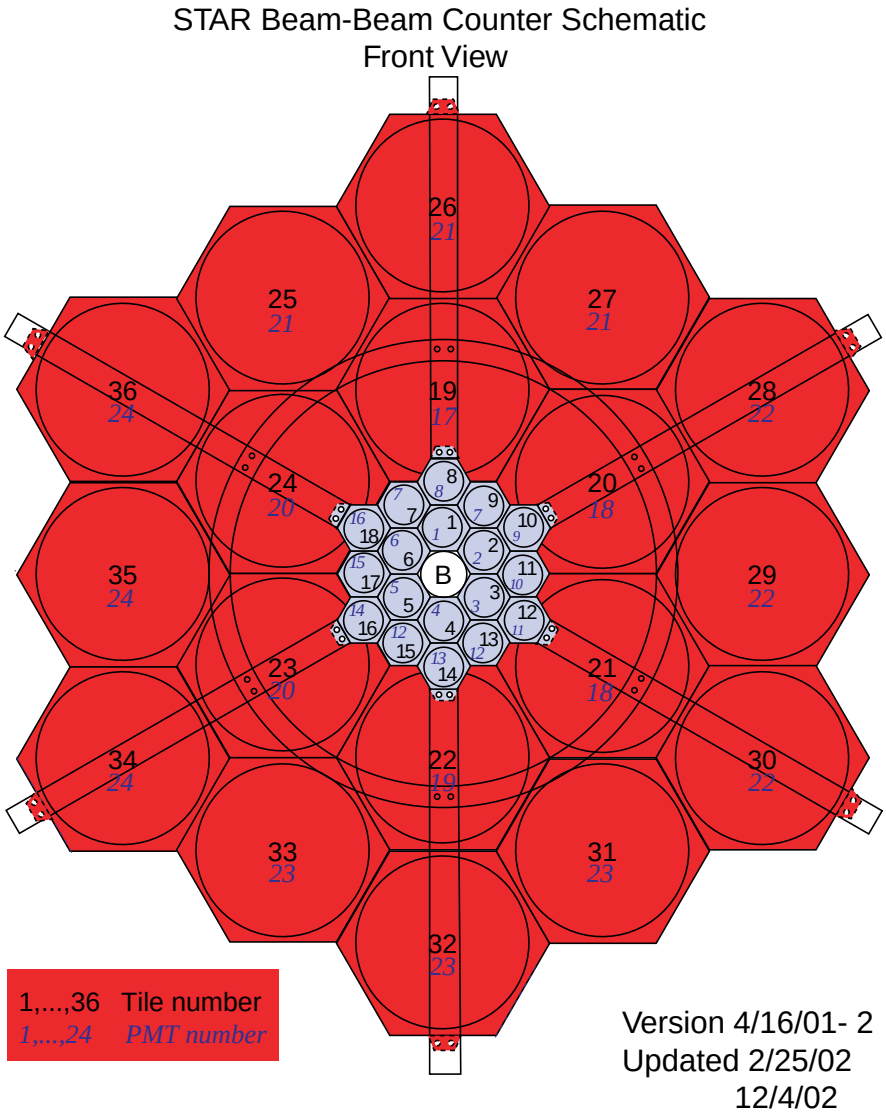

Figure 5. Diagram of the Beam Beam Counter.
The small hexagon marked `B' is reserved for the beam pipe[15].

b) Zero Degree Calorimeter

The Zero Degree Calorimeters are small hadron calorimeters at STAR, located downstream (east and west) of the region where the beams collide, usually called the interaction region (IR)[17]. The other experiments at RHIC use nearly identical detectors for essentially the same purpose. Because of strong magnetic fields between the STAR IR (interaction region) and the ZDC, charged particles are diverted before they can reach the detector. Thus, only neutral particles have a significant probability of reaching the ZDC. The ZDC is located roughly 18 meters from the center of STAR, and subtends a solid angle of approximately 30 μsr. The ZDCs are centered on the mathematical axis of STAR, while the beam is diverted with magnets along the ring.

In heavy-ion collisions, ZDCs are used to count the number of spectator neutrons and to produce coincidence signals. The binding energy of the Au nucleus is measured in MeV, and this is very small compared to the 200 GeV per nucleon in a heavy-ion collision. After a hard collision, the nucleus completely dissociates in the 18 meters from the IR to the ZDC. Under these conditions, only neutrons which did not collide with the opposing nucleus continue onward. Thus, in hard Au+Au collisions, the number of spectator neutrons found can be related to the centrality of the collision. In d+Au collisions, the neutron from the deuteron is sometimes a spectator and strikes the West ZDC. When this occurs, it is of particular interest, since it happens in a unique class of collisions where the proton from the deuteron collides peripherally with the Au nucleus. The rest of the time, however, the East ZDC receives many neutrons from the nucleus, and this is used as a minimum-bias d+Au trigger condition.

2. STAR TPC

The STAR Time Projection Chamber (TPC) is currently the largest detector of its kind in world. The TPC, when reading out an event, creates an internal 3-D image containing approximately 70 million data points. Fortunately, most of this "image" is "black", and does not need to be recorded. The TPC is constructed in the shape of a hollow cylinder, with inner radius of 0.5m, outer radius of 2m and overall length of 4.2m.

The TPC relies on the principle of electron drift. When a high-energy hadron collides with molecules of TPC gas, it often liberates an electron, which is induced to drift longitudinally down the length of the TPC to be read out as a signal at the ends. To produce this drift, a longitudinal electric field is imposed. A central carbon coated annulus membrane is stretched perpendicularly to the beam in the center of the TPC. This membrane is set to a high voltage of 28 kV, and gating grids at the ends of the TPC are set to ground. 183 resistors and equipotential rings along inner and outer radii are used to maintain a uniform electric field, since deviations cause altered drift velocities and paths. Behind the gating grid are detector pads which are used to read in the electron signals. The TPC has 5692 pixels in each of 24 sectors for a total of 136,608 detector pads. For each event, 512 time samples are taken at a rate of 9.4 Mhz.

The TPC uses a special mixture of gas which has desirable properties for this application. This gas, called P10, contains 10% Methane and 90% Argon. It is kept at 2 mbar positive pressure in order to prevent ambient air from being introduced. Given the voltages and gas properties, the drift velocity of electrons in the gas is on average 5.45 cm/μs with typical time variation of about 6%. This drift velocity is calibrated with directed laser beams, turned on only during a special data acquisition mode, which induce ionization in the gas along line segments whose exact positions are known. The transverse diffusion of electrons in P10 gas works out to 3.3 mm over the 210 cm drift length of the TPC. The longitudinal diffusion is similar, and comes to 5.2 mm. Dividing 5.2 mm by the drift velocity gives the best possible time resolution, which is about 100 ns, making 10 Mhz the optimal rate for taking both p+p and d+Au data.

In addition to measuring ionization trails, the TPC can also measure the energy deposited per unit length (dE/dx) of a particle. This value is related to the charge of the particle and its velocity by the Bethe-Bloch formula for energy loss:

$$-\frac{dE}{dx} = \frac{4\pi}{m_e c^2}\frac{nz^2}{\beta^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\left[\ln\left(2m_e c^2\beta^2\gamma^2 / I\right) - \beta^2\right] \qquad (9)$$

Where z is the charge in units of electron charge of the particle, β, γ are normal relativistic definitions, n is the electron density of the target, and I is the mean excitation

potential of the target. Thus, particle identification is sometimes possible when p < ~1 GeV/c.

The electrons drift for the most part according to elementary electromagnetic equations, but certain corrections need to be applied. These corrections can usually be done after the data has been taken. The magnetic and electric fields are not exactly longitudinal due to various effects. The careful construction of the magnetic field and electric field is only accurate within the constructed tolerance. During running, ions build up near the inner surface of the TPC, and build up a radial potential due to a charge density that goes like $q(r) = \frac{1}{r}$. Large numbers of particles can even create space charge in the gas which will repel itself. The sum of these effects will cause secular deformations in the image read by the pad electronics. Nevertheless, since they are small perturbations, most of these deformations of the image can be corrected.

Due to this hollow cylindrical shape, the detector responds with full efficiency to charged particles with $|\eta| < 1$, since these particles may potentially fire all 45 pad rows. The limit of the TPC's acceptance is at $|\eta| = 2$, above which particles are completely lost. For most analyses, a particle is required to make an electron trail with sufficient definition to be read out in 15 distinct data points, requiring that the particle fall within $|\eta| < 1.4$. This is the limiting case, and in practice the acceptance of the TPC rolls off smoothly in the range $1 < |\eta| < 1.4$. Another limitation, given the inner radius of 0.5m coupled with the standard magnetic field, is that a track with p_T < 150 MeV/c is impossible to detect. To quantify this, the STAR TPC is occasionally run at half field so that the minimum p_T drops to only 75 MeV/c.

3. STAR Barrel EMC

The STAR Barrel Electromagnetic Calorimeter (BEMC) fits between the TPC and the outer magnet. During the 2003 RHIC run, it covered the region $0 < \eta < 1$ and had full ϕ coverage. The BEMC is composed of 60 modules. Each module extends over 6º in ϕ

and the length of half the barrel, $0 < \eta < 1$. Each module contains 40 towers stacked in a 2x20 configuration, and is 26 cm wide and 293 cm long. Each tower thus subtends 0.05 radians in ϕ and 0.05 units of η. The towers are 21 electromagnetic radiation lengths deep. Each tower consists of a stack of 41 alternating layers of scintillating material and lead, each 5 mm thick. When electrons traverse the scintillating material, they generate 5 eV photons which are then picked up by wavelength shifting fibers. Each of the 21 scintillating layers feeds to a fiber, and all the fibers feed into a single Photomultiplier Tube (PMT), since the timing of photons from the scintillating layers is unimportant. The magnitude of the PMT response is linearly proportional to the energy of the incoming electromagnetic shower. The towers of the BEMC are projective, meaning that the angular area of each layer in a tower is the same with respect to the center of STAR.

The behavior of the tower is strongly dependent on the type of particle impinging upon it. High-energy electrons predominantly lose energy in matter by bremsstrahlung, and high-energy photons by $e^+ e^-$ pair production. The characteristic amount of matter traversed for these related interactions is called the radiation length. It is both the mean distance over which a high-energy electron loses all but 1/e of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon. The towers are most sensitive to photons and electrons, which both generate electromagnetic showers. Electromagnetic showers are almost completely absorbed by the BEMC since it is 21 radiation lengths deep. The BEMC was designed to be most sensitive to photons and electrons, and to make possible the reconstruction of π^0 and η particles after they decay into photons. Most other particles leave only a fraction of their energy in the BEMC. Hadrons, in general, just begin to generate a shower by the time they punch through, since the tower is about one hadronic interaction length deep. Occasionally, an anti-proton or anti-neutron will annihilate and deposit most of its energy. Usually, charged particles (other than electrons) will leave only small amount of energy in the tower, essentially independent of particle species or energy. Such a particle, which leaves on average 300 to 350 MeV in the detector, is called a Minimum-Ionizing Particle (MIP).

Though the response of each PMT is linear, the relationships between tube response and energy vary from tube to tube. The energy relationship for a tube is:

$$E = G \cdot (ADC - P) \tag{10}$$

Where E is the energy, P is the quiescent average PMT ADC count, G is the PMT gain, and ADC is the digitized voltage of the tube. The PMT tubes can all be calibrated relative to one another, and the approximate absolute gain calculated, using the consistent average energy of the MIPs. This relative calibration makes it possible to conduct an absolute calibration using identified electrons, which are too rare to calibrate a single tower. Each identified electron measured in the TPC is tracked to the tower it strikes in the barrel, and that tower's response is factored into the overall gain of the entire barrel.

Unlike the TPC, which requires the electrons to drift up to 2.1 meters to be measured, the BEMC is a very fast detector. The PMTs respond as soon as the light propagates to them, and so the BEMC can be used to make decisions to store events with an interesting BEMC response. For this analysis, the "interesting" BEMC response was simply a tower above a specific ADC threshold. This threshold corresponded to 2.5 GeV. Other "interesting" BEMC responses which have been used as triggers, but not used by this analysis, include a higher tower threshold, sums of rectangular groups of towers above a threshold, and sums of all towers above a threshold. During the 2003 RHIC run, these other trigger types were mostly experimental, but they have since become better understood and are very important for the long term spin goals of STAR.

Though not used in this analysis, the Barrel is also equipped with unique pre-shower and shower maximum detector (SMD) layers. The pre-shower layer consists of 2 scintillating layers 6 mm thick on the inner surface of the towers. The SMD, about 5 radiation lengths deep into the EMC, consists of a wire chamber with high-position resolution that provides a much more accurate profile of the shower at its maximum intensity. Since electromagnetic showers and hadronic showers differ greatly in shower shape, the SMD is intended to help discriminate between these two- cases, and allow a cleaner sample of electrons and photons to be obtained.

III. STAR EVENT RECONSTRUCTION

After an event is recorded by the detector, and later moved to permanent storage, the raw data needs to be processed prior to analysis. In this section, the focus is primarily on the TPC and BEMC detectors, since these are the sources of data used in this analysis. In general, each event is ultimately stored in a giant structured block of information. This structure, literally called the StEvent in software programs, is the end result of much elaborate processing. Much of this involves populating trees, lists, and other structures from flat arrays recorded by the hardware. Other sequences in the data processing conduct calibrations and corrections. Still others compute and store numerous derived quantities. This data structure can on occasion be as large as a Megabyte for a single event. Recalling that tens of millions of events can be recorded in a run makes this all the more impressive. Such huge data files are quite clumsy, and so STAR processes the data one final time to produce much smaller "muDst"[§] files, containing all the relevant data for physics analysis. However, before the small files can be made, the collision event has to be reconstructed.

A. TPC Reconstruction

The process of reconstructing the event in the TPC, as mentioned in the last section, involves analyzing a sparse array of approximately 70 million ADC counts. The algorithms used to do this must be designed carefully to avoid inefficiencies. Even though it is a sparse array, there are still enough data points to make order n^2 algorithms absurd. The detector is intended to record a 3-D picture of the particle trails, and the software needs to impose helical trajectories, called tracks, on the 3-D trails, to reconstruct the particle momenta. It is also necessary to calculate the most likely location of the collision, called the primary vertex, which produced the particles. For primary tracks, tracks which intersect the vertex, it is desirable to tweak the helical trajectories and improve the

[§] The micro Data Storage Tape (muDst) file is not one millionth the size as the name suggests. In reality, it takes up roughly 1% of the space. That is *still* a whole lot of data.

momentum resolution. Other helices, called secondary tracks, do not intersect the vertex, and these are tagged accordingly.

1. TPC Hit Finding

The readout pads at the ends of the TPC record digital ADC values between 0 and 512, indicating the electron current at each time step. The 2-D pad location coordinate, the time bin, and ADC value can be used to reconstruct the ionization energy at a 3-D location internal to the TPC. This 3-D location and ADC value is termed a pixel. A single electron may fire several pixels in the same pad row by drifting and minor showering in the gas. This, coupled with the discrete pads, will produce a blurred "image" which extends over several pads and time bins.

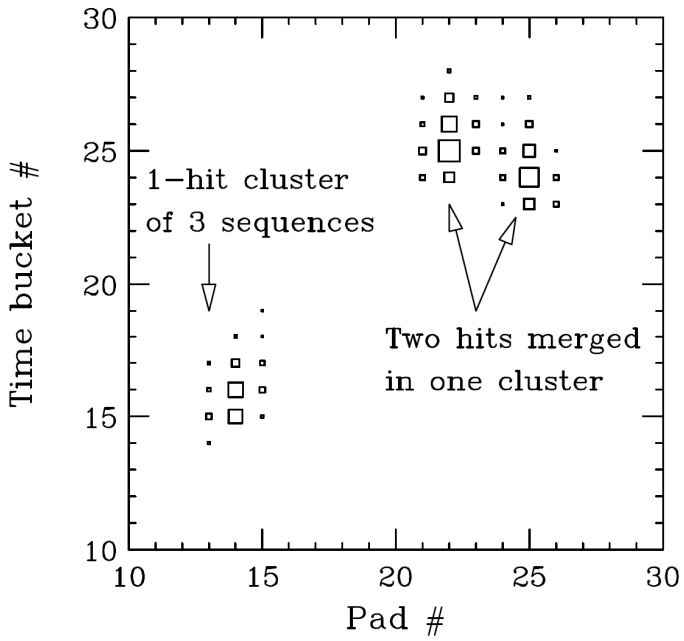

Figure 6. Clusters found by TPC clustering.
The size of the boxes represent ADC values. [18]

The first step in sharpening the image is to create clusters of all adjacent pixels above a certain threshold (all pixels below this threshold are ignored). While clustering, corrections are applied which depend on pad index and time bin, arising, for example, from field non-uniformity or space charge distributions. Each peak on the graph, corresponding to the physical collision of a particle with the gas of the TPC, is caused by what is termed a hit. Figure 6 shows how occasionally hit peaks will get combined into a single cluster. This can be detected by testing the RMS pad and time width of the cluster. If the RMS width is too large, a Gaussian fit in space and time is applied to split the cluster into its constituent groups. After clusters are split, the overall sector, pad row, pad, and time bin numbers within the cluster are used to reconstruct the STAR global (x, y, z) coordinates of each hit.

2. TPC Track Finding

After obtaining a list of (x, y, z) hits in global STAR coordinates, the stage is set for reconstructing the particle trajectories. The reconstructed helical trajectories are termed tracks. For this analysis, the Time Projection Chamber Tracker (TPT) software was used to reconstruct tracks. This program starts with seed hit triplets in the outermost pad rows, and progressively adds more and more hits to extend line segments from the outside inward. It then merges the line segments to produce full helices.

The TPT algorithm is very similar to the ones used for ALEPH and NA35 at CERN. The algorithm begins with the outermost TPC pad rows where occupancy is lowest. Lower occupancy helps to reduce the incidence of spurious hit triplets. Upon finding a set of three adjacent hits, close in both time and space, the algorithm extrapolates along a straight line moving inward one pad. If another hit is found at this location, these four hits are packaged into a helix segment, and the hits are removed from the search pool. This process is repeated until no additional helix segments can be found in the outermost pad row. The algorithm then moves one pad row inward, and repeats the search for 4 hit helix segments. Eventually, the entire TPC is searched for 4-hit helices in this way.

When every 4-hit helix has been found, each helical segment then gets extended in both directions up to the boundaries of the TPC. Whenever the helix closely passes another hit, that hit is removed from the search pool, and added to the helix defined by the four-hit segment being traced. After this is done for all helices, pairs of close helices are checked in order to avoid splitting a single helical trail into two- and measuring the particle twice. If the trails are similar enough, they are merged. Next, the TPT optimizes the helix fits using all the hits added to each helix. The optimization of the fit is done first in the (x, y) plane using a circular fit, and then in the (z, r) projection. The results of the algorithm at this point are placed in data structures called tracks.

An improvement of the track fits to the hits is then attempted using a Kalman filter routine. The Kalman program weights hits by measurement uncertainty, in each case, and defines a width for the hits in order to measure distance to the helix. This routine is an iterative process which begins by discarding hits which lie over 5 σ from the helix. Next, the fit is optimized, where the program takes account of time dependent momentum loss and scattering in the TPC gas using a GEANT [19] model based on detector materials and geometry. This process is iterated until the fit is stable. More information on the TPT algorithm can be found in STAR Note 281 [20]

3. Primary Vertex Finding

Once the tracks have been fully optimized, the TPT attempts to locate the collision point of the event, called the vertex. The location of the vertex, if included in the helix fit, improves the momentum resolution a great deal, especially for high- p_T tracks. In addition, knowing which tracks come from the real collision is very useful for background suppression.

Finding the vertex involves making successively better hypotheses of the vertex location. For the data used in this analysis, a single vertex was assumed, since it was very rare to have two- collisions in the same event during year 2003 data taking. Thus, all the tracks are initially considered as possibly coming from the vertex. The algorithm proceeds, during each iteration, by extrapolating all tracks to the beam line. The z values

of the tracks at the beam line are used to create a hypothesis for the location of the vertex. Every track that intersects the beam line at a location more than 5 cm from the vertex hypothesis is thereafter excluded and becomes a secondary track. This process is repeated until the vertex is stable. Finally, the primary tracks are again refit using the Kalman filter described above, with the addition of the vertex as a high-confidence data point.

B. Barrel EMC Particle Reconstruction and Jets

The reconstruction of π^0 and η particles, and their discrimination from electrons, MIPs, and hadronic annihilations, is a complex subject. Much work has been done at STAR, and a complete dissertation could be written on this subject alone. Jet analysis involves packaging hadrons in a bundle and measuring their collective properties. Since the purpose of this analysis is to measure parton properties, the fragmentation hadrons, such as π^0 and η particles are of secondary importance. Jet reconstruction sums all the particles together, independent of which hadrons deposit the energy. Up to the mass values of the particles, summing raw energy is equivalent to summing identified energy. Using the raw energy is the simplest procedure, and is a good approximation. The fractional change in the reconstructed energy of the jet due to neglected particle masses is quite small compared to the energy resolution effects introduced by missing particles or particles inappropriately included in a jet.

For this analysis, the raw BEMC hits are used directly for jet finding, with an intermediate reconstruction step which does not attempt true particle identification. This single step converts the hit energy to a four vector based on the location of the vertex as determined by the TPC track reconstruction. A vector is calculated from the vertex to the location of the tower hit, and a four momentum is assembled using this vector and a mass of zero. Sometimes, a track helix from the TPC intersects this tower. In that case, there is an effort made to suppress an average MIP response in the BEMC. The actual response of the BEMC to a charged particle (other than an electron) may vary widely from a constant MIP response. Antiprotons, in particular, may occasionally leave a normal MIP, but may

also annihilate with the scintillator material, dumping the full energy of the anti-proton into a non-standard BEMC shower. Since the energy of the particle producing the MIP has already been accounted for by the TPC track four momenta, the jet reconstruction summation would otherwise gain *on average* 300 MeV from the MIP. After subtracting some MIP energy, where appropriate, the four vector is ready for jet reconstruction.

It is very important to realize that the details of the subtraction, and the subtraction itself, are not critical. The magnitude of the MIP response of the detector is inherently random, and any constant, linear, or even arbitrary MIP suppression scheme cannot predict the specific MIP response of the detector in a given real event. MIP subtractions can cause only a modest reduction in MIP energy scale fluctuations. In fact, the energy scale fluctuations caused by missed particles is much larger than the fluctuations due to MIPs. MIP suppression also has an effect on the overall average energy scale of the jet, but this effect is irrelevant. The average energy scale must be obtained by comparing data with simulation, both of which undergo identical MIP suppression. The relatively small MIP energy scale effect will just get indiscriminately lumped in with all the other effects.

C. Acceptance, Efficiency, and Resolution of Tracks and Hits

In the next section, the jet finding algorithms are discussed. Before moving to that topic, a few words should be said about the effects of uncertainty in measurements of tracks and hits on the reconstruction of the jet. Energy, momentum, and angular resolutions of jets are dominated by missing or background tracks and hits. Individual track or hit uncertainties are much smaller, and even tend to cancel out when multiple track and hit momenta are added. Tracks in p+p and d+Au collisions are reconstructed between 85 and 90 percent of the time. About 10 percent of tracks are unrecoverable since they land in un-instrumented structural support regions of the TPC. Roughly 95% of electrons and photons incident on the BEMC are recovered in tower energy measurements. Though quite efficient, the BEMC also appears to be susceptible to correlated noise and background sources which require serious energy cuts. These cuts reduce the BEMC efficiency. The approach taken in this dissertation is to introduce deep cuts which

minimize the background. The detector response can then be cleanly predicted based on comparisons with simulation. Background and inefficiencies, corrected in this fashion, do result in larger energy scale and angular corrections.

IV. JET FINDING ALGORITHMS

Jets are the product of parton fragmentation. The hadrons in the jet are correlated with each other in momentum, since they form in the rest frame of the hard scattered parton. It is therefore possible to pick out which hadrons were likely produced by the same fragmenting parton. The only relevant detail of the hadron for this purpose is the momentum. Once four momenta have been reconstructed for each of the particles, the momentum correlations between them can be measured and quantified.

A. Detector and Jet Finder Decoupling

The task of jet finding in a collision can be decoupled into two- steps. The first step, which is intimately connected with the detector, is to interpret detector measurements as particles with four momenta. This has been discussed in the context of this analysis in the previous section. However, any detector which measures energies or momenta may be adapted to produce these four-vector momenta. In fact, as done in this analysis, four-vector momenta can come from simulation data and be fed into the jet finder. The second step is an abstract four-vector momentum analysis which bundles the four vectors into putative jets. The second step can be done using many algorithmic variations, and these are the topic of this section.

The first jet reconstruction programs to be run at STAR were written by Michael Miller and Konstantin Shestermanov. K. Shestermanov created a simple cone algorithm, and also adapted a well vetted Fortran k_T algorithm used for many years by collaborations at other experiments. M. Miller created a state of the art cone algorithm written in C++ based upon a literature review of best jet reconstruction practices at other experiments, in particular CDF [21]. He also wrote a simple k_T algorithm which was not as complex as the Fortran code adapted by Konstantin. The outputs of the cone codes and the k_T codes were to first order equivalent. In all cases, however, the two- decoupled steps of jet

finding were directly fused together. These codes directly read information from the muDst file TPC tracks, and had several different output formats.

This author wished to compare jet finder outputs, and to introduce data from the BEMC in addition to the TPC data into jet finding. Neither one of these tasks were possible given the architecture of the algorithms at that time. The necessary innovation was to first decouple the algorithms from the TPC, second to standardize the input four-vector format and the output jet format of the algorithms, and finally to introduce a module to build a four-vector list from the BEMC and combine it with the TPC list. This author was able to cleanly decouple the detector-based four-vector builder from the abstract four-vector based jet finding code. This also makes it straightforward to include data from additional detectors, such as the Endcap Electromagnetic Calorimeter, in the future. Most of the code written by this author to accomplish this is still being used in the current generation of jet finding software at STAR.

B. Charged Jets and Charged + Neutral Jets

Even though combining neutral and charged-energy information was an important goal of the decoupling work done on the jet finder, it is often useful to correlate the results of jet finding with and without the neutral energy component. This is relatively easy to accomplish. Given how jet finding is now factored into a four-vector builder and the abstract jet finding code, it is a simple matter to leave the neutral energy obtained from the BEMC (or any other detector) out of the four-vector list. For this analysis, there are important uses for charged-only jets. The BEMC at various times has had substantial anisotropies in ϕ and η due to power supply failures or regions which were not yet instrumented. Comparatively, charged-track acceptance in the TPC is very uniform in ϕ and η. As will be shown later, TPC only jets provide an excellent estimate of the jet thrust axis, though the energy scale is obviously less certain. It is especially useful for the di-jet analysis to obtain the thrust axis under uniform acceptance conditions.

C. Algorithms at STAR

The jet finding algorithms used at STAR can be broadly categorized into two-types. These two- types follow from the two- simplest correlations between four vectors. The first type, called cone algorithms, finds jets by correlating particles with similar trajectory *directions*. These algorithms bundle particles based upon inclusion in a geometric cone. The alternative type of algorithm is called a k_T algorithm, because the particle correlations are based upon *momentum* differences. It is described here for completeness, but k_T algorithm results are not used in this analysis. The reason for this will be explained in the k_T algorithm description.

1. Simple Cone Algorithm

STAR analyses can make use of a simple cone algorithm if desired. This algorithm starts by binning the four vectors into a grid, and setting the four vector of each bin to the summed momenta of the four vectors in it. Then the algorithm simply loops over the cells in the grid. Whenever it finds a cell with a momentum above a user specified seed threshold, it places a cone with a user specified angular radius centered on that grid cell. This cone is defined as a proto-jet. The proto-jet four-momentum is calculated by summing the four momenta of the cells falling within the cone. When the loop over grid cells is complete, the algorithm selects the best proto-jets. Every proto-jet which does not intersect another is considered final. Any proto-jet which does intersect another is compared. The most energetic proto-jet in each group of overlapping proto-jets is also considered final. All the final proto-jets are collected and returned as the output of the algorithm.

Because of the finite grid binning, this algorithm is $O(n)$, n being the number of particles. Given an appropriate seed threshold, the constant is small, and the algorithm is very fast. It is fast enough to run on any event, regardless of multiplicity, even Au+Au events, although the background would completely swamp the jet signal. Thus, this is the algorithm of choice for d+Au jet reconstruction.

2. Optimized Cone Algorithm

The optimized cone algorithm is M. Miller's state of the art cone algorithm based primarily on jet reconstruction experience gained at the CDF experiment. The optimized cone algorithm starts with the proto-jets as found by the simple cone algorithm just described. It then adds several further steps to find the optimum cone directions to produce the most energetic jets possible in the given four-vector lists and user specified cone radius.

a) Seedless Cone Algorithm

It is possible to scan the entire fiducial space of the detector for jets in an unbiased way. If the energy of a cone with user specified radius is calculated at every point in a discrete fiducial space, the maximal jet energies and momenta can be found. Unfortunately, this is very computationally intensive, even for low multiplicity p+p events, and while the results are better, it is impractical. The goal of the optimized cone algorithm is to obtain unbiased results in a computationally tractable manner. The midpoint cone algorithm and the splitting merging algorithm, when executed on the proto-jets found by the simple cone algorithm, produce very similar results to the seedless cone algorithm. The computation time is low enough to make it possible to run the optimized cone algorithm on p+p events.

b) Midpoint Cone Algorithm

The simplest optimization of a proto-jet is to align the cone with momentum vector of the proto-jet, include and exclude four vectors as the cone demands, and repeat until a stable cone alignment is achieved. For the sake of clarity, let this iterative process of realigning the proto-jet be called ***stabilization***, and a proto-jet which is the result be called a ***stabilized*** proto-jet.

The midpoint algorithm takes the proto-jets from the simple cone algorithm, and runs the *stabilization* on them. Next, the midpoint algorithm creates a proto-jet centered on the midpoint between every *stabilized* proto-jet pair. These midpoint proto-jets are

then *stabilized* as well. If they are not identical to another already *stabilized* proto-jet, they are added to the *stabilized* proto-jet list. Adding midpoint proto-jets and running the *stabilization* on them is repeated until no additional unique *stabilized* proto-jets are produced in the loop.

c) Splitting/Merging Algorithm

The midpoint algorithm can potentially produce a large number of stable proto-jets. Many of these will overlap, sometimes because they lock onto local maxima within a single jet. Though uncommon in p+p, if multiple jets exist in a single event, the jet finder will sometimes lock onto separate jets which happen to be close together. The splitting/merging algorithm is a process of selecting proto-jets as final reconstructed jets. The selection process begins with the highest E_T proto-jet found in the list. Next, this proto-jet is tested to determine if it shares cells with any other proto-jet. If there is no intersection, the proto-jet is accepted as a final jet. Otherwise, the following fraction is formed:

$$\frac{E_{Tshared}}{E_{Tneighbor}} = R \qquad (11)$$

Where $E_{T\ \text{shared}}$ is the energy in the cells shared, and $E_{T\ \text{neighbor}}$ is the energy of the highest E_T overlapping proto-jet. A choice is then made based upon the value of R. If R is less than the user specified parameter $f_{\text{split-merge}}$, the jets are split; otherwise the jets are merged. The parameter $f_{\text{split-merge}}$ is usually set to about 0.5. Merging is very simple; all the cells from the high-E_T proto-jet and the highest overlapping E_T proto-jet are collected with deletion of duplicates to create a new single proto-jet, and the two- constituent proto-jets are discarded. Splitting takes the overlapping cells away from both proto-jets and gives them to the nearest overlapping proto-jet, which may not be the highest E_T overlapping proto-jet. Duplicate cells are not copied. After one of either a splitting or merging has been performed, there is one less pair of overlapping proto-jets in the list. The highest E_T proto-jet is then selected again, and the process repeats, either accepting the proto-jet as

final, or reducing the number of intersections by one. The optimized cone algorithm completes when all proto-jets have been separated or merged into non-intersecting jets. Due to splitting and merging, final jet shapes may be more elliptical than circular, and may not be well characterized by the user specified cone radius.

3. j_T (kT) Cluster Algorithm

In the literature, this algorithm is usually referred to as the kT algorithm (without a subscripted "T"), but to keep this discussion clearly separate from the measurement of parton k_T, it is here referred to using the more accurate j_T nomenclature. The j_T cluster algorithm is controlled by a user supplied parameter analogous to a cone radius. This parameter is explicitly a relative momentum limit, and also functions as a minimum seed particle E_T . The algorithm proceeds by starting with the highest E_T four vector, called the seed. It is removed from the four-vector list. A partial jet momentum is then set equal to the momentum of the seed. Next, the four vector with the smallest relative momentum to the partial jet is removed from the list and added to the partial jet momentum. This transverse momentum to the jet thrust axis is called j_T. Four vectors are continually removed until the smallest relative momentum of a four vector to the partial jet is larger than the user specified momentum limit. The partial jet is then assigned final jet status, and a new seed particle is selected. This process repeats until there are no seed particle candidates remaining. Most implementations set the seed threshold sufficiently low that no four vectors remain unassociated with any jet.

In some ways, the j_T algorithm is more appealing then the cone algorithms. The process of combining particles with minimum momentum difference closely resembles physical jet fragmentation in reverse. One of the very first steps of this analysis, at least in time, was to compare the results of j_T and cone algorithms on the clean p+p system. The results showed that the physics could be obtained independent of the algorithm used, in essence permitting the work to be done using either path. Early on, it became clear that the j_T algorithms available for this analysis could not be easily tuned to filter out background, especially in high-multiplicity events at STAR. At $\sqrt{s} = 200$ GeV, portions

of the jet signal are difficult to distinguish from background, and the algorithm for jet reconstruction has to be tuned for best performance. In addition, the j_T algorithms available at STAR did not perform well on d+Au events. Since the algorithm is only a means to the physics, the decision was made to use the cone algorithms for this analysis.

V. SOURCES OF DATA ANALYZED

Starting late in 2002 the preparations for d+Au collisions were initiated. The RHIC ring is large, and cooling all the superconducting magnets takes several weeks. After the magnets had been cooled down to operating temperature, several more weeks were taken exercising the collider in order to tune it for collisions. On January 5, 2003 the d+Au beams were judged acceptable for physics. Data were taken using this configuration for 11 weeks.

On March 25, p+p collisions were introduced. RHIC commenced operations and ran an additional 10 weeks of physics. On May 31, the final p+p collisions were made, and the collider was transitioned back to its inactive state. This entire period of d+Au and p+p running, for the most part during the spring of 2003, is called RHIC Run 3. Data used for this analysis comes from data taken during RHIC Run 3.

A. Year 2003 Data

As discussed in section II, STAR uses various detectors to determine when to read out and store information about a collision event. To minimize the bias toward a particular physical outcome, STAR uses the BBC and ZDC to create a trigger which ideally has the smallest bias possible. During RHIC Run 3, the STAR detector collected nearly 22 million d+Au "minimum-bias" events, 1.2 million d+Au "high-tower" events, 1.6 million p+p "minimum-bias" events, and 1.9 million p+p "high-tower" events.

1. Year 2003 p+p Minimum-Bias Data

During the latter portion of Run 3, the p+p "minimum-bias" trigger data were recorded. A p+p "minimum-bias" trigger occurs after a coincidence between STAR's two- BBCs. This triggers on $87 \pm 8\%$ of the total Non-Singly Diffractive event cross section [22]. The amount of data stored was significantly less than the maximum possible, because the STAR collaboration made the decision to reduce the "minimum-bias" trigger bandwidth in favor of more rare and interesting triggered events. For this analysis, the

"high-tower" triggered data is quite valuable, since it contains a greater fraction of events with reconstructible jets.

2. Year 2003 d+Au Minimum-Bias Data

During the first portion of Run 3, d+Au "minimum-bias" trigger data was recorded. This "minimum-bias" trigger occurs after the east ZDC catches a beam rapidity neutron fragment from the Au nucleus. This triggers on $95 \pm 3\%$ of the d+Au hadronic cross section [23]. The STAR collaboration placed a fairly high priority on this trigger, and consequently many times more d+Au "minimum-bias" data were recorded than for p+p. The "minimum-bias" triggers for p+p and d+Au are quite different, and these differences must be taken into account when doing cross section comparisons. For jet reconstruction purposes, the "minimum-bias" d+Au data are difficult to use. This unpleasant fact comes from the presence of numerous particles produced in the d+Au collision which come from soft interactions. The ubiquity of "soft" particles combined with the rarity of a high- p_T jet in the "minimum-bias" d+Au sample decreases the jet signal to noise ratio, creating an increase in the energy scale fluctuations especially at modest jet p_T.

3. High-Tower Trigger Data

Due to bandwidth and storage constraints, only a small fraction of the total minimum-bias events encountered during the run can be permanently recorded by STAR. Certain measurements, for example observables which compare two- jets in a single event, require a sample of jets larger than that contained in the recorded minimum-bias sample. Therefore, to measure high-statistic jet observables, STAR also records data when it encounters a "high-tower" trigger.

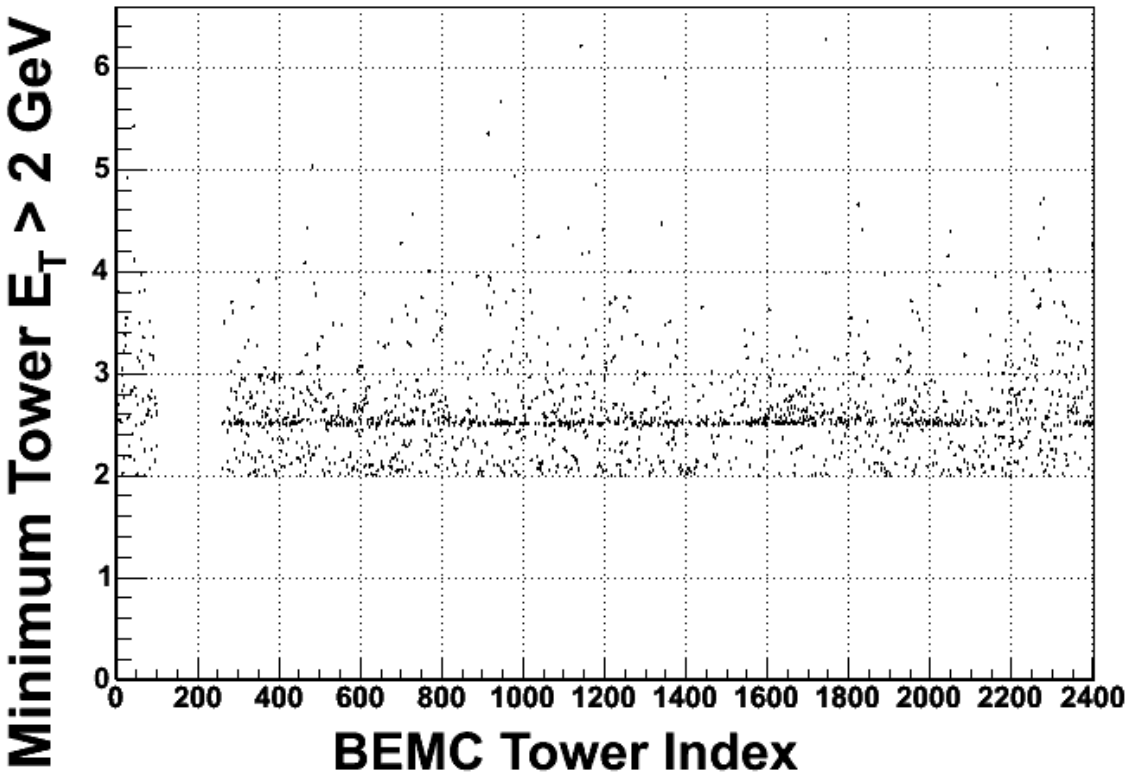

Figure 7. p+p "high-tower" triggered minimum tower E_T.
This figure shows the minimum tower E_T (assuming vertex at z=0) found in any jet for a given BEMC tower index assuming that the tower E_T > 2 GeV.

A "high-tower" trigger occurs when one of the Barrel EMC towers goes above a certain threshold and is in coincidence with a "minimum-bias" trigger. For Run 3, a threshold of approximately 2.5 GeV was used. The initial calibration used by STAR during Run 3 was based upon a small amount of data, and this differs from the final calibration based upon the entire data set. Some towers triggered the detector even though they had slightly less than 2.5 GeV of energy. Likewise, some towers were not taken as triggers even when they had 2.5 GeV of energy or slightly more. Figure 7 shows the minimum triggered energy of each tower after applying the final Barrel EMC calibration.

Figure 7 shows that the initial calibration varied significantly from the final calibration. By placing a cut on the jets requiring them to have a tower above 2.8 GeV in

the final calibration, the non-uniformity due to the calibration used by the triggering system can be reduced significantly.

The "high-tower" triggered data are richer in jets than the minimum-bias samples, and are also similar in the p+p and d+Au environments. This makes it possible to directly compare the two- systems. The minimum-bias data sets differ between p+p and d+Au due to the high-multiplicity inherent in d+Au collisions. This raises the question: do the "high-tower" triggered data sets also differ for the same reason? The answer is no, provided straightforward corrections are applied to the d+Au results, as the next section will discuss.

4. Year 2003 p+p Supplemented with d+Au Minimum Bias Particles

To measure the effect that the d+Au multiplicity background has on the reconstructed jet energy, the approach taken is to mock up the d+Au multiplicity background and add it to the p+p system. Define a p+p supplemented event as a p+p "high-tower" event with additional particles taken from a d+Au minimum-bias event such that the overall multiplicity distribution of the p+p supplemented events is equal to the overall multiplicity distribution of the d+Au "high-tower" events. Define a "2X" p+p supplemented event the same way, with the exception that the four momenta of the particles taken from the d+Au event are multiplied by a factor of 2.

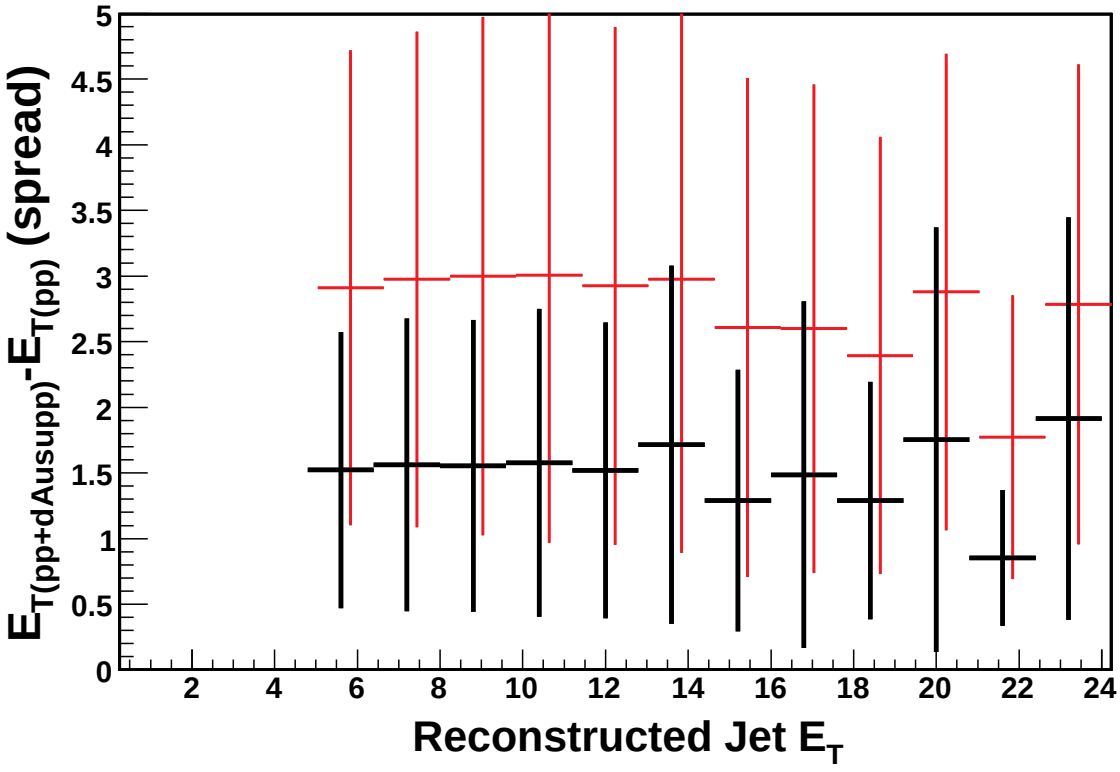

Figure 8. Spread of p+p supplemented minus p+p plain jet E_T vs. E_T . p+p supplemented is p+p data with the addition of d+Au minbias particles. Red indicates p+p supplemented energy multiplied by 2.

The background contribution to the jet in the d+Au "high-tower" event is probably about the same as the p+p supplemented event, and certainly less than the "2X" p+p supplemented event. A comparison of the jets obtained from p+p, p+p supplemented (p+p supp), and "2X" p+p supplemented reveals the magnitude of the systematic uncertainty in the reconstructed jet energy. Figure 8 shows the width of the E_T difference distribution (not the uncertainty in the average), which is called the spread. Figure 9 shows the approximately exponential distribution of these differences. Figure 8 shows that the "High-tower" triggered jets reconstructed in d+Au do have a contribution from the high-multiplicity d+Au background, and this is somewhere in the neighborhood of 1.5 GeV or about 15%-20% depending on jet E_T . In Figure 8 (and the Figure on pg. 45), the spreads

are a bit strange in the last 4 bins, and this is because the statistics are limited above ~19 GeV.

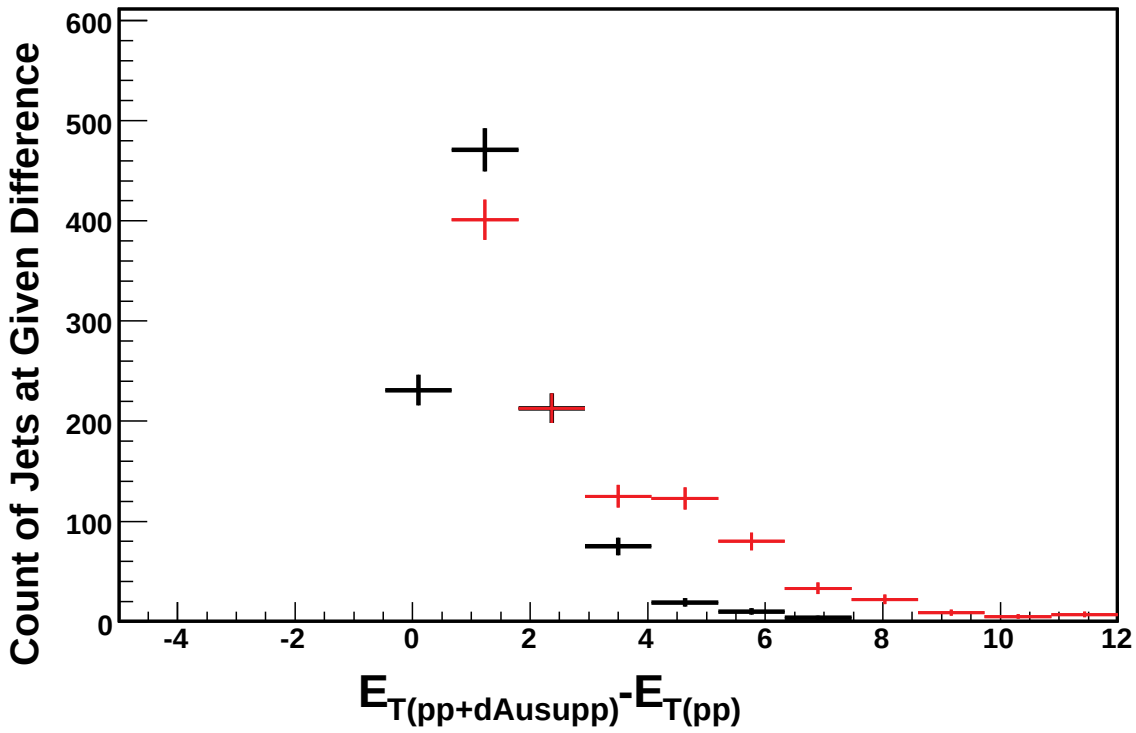

Figure 9. Difference of p+p supplemented and p+p plain jet E_T .
p+p supplemented is p+p data with the addition of d+Au minbias particles. Red indicates p+p supplemented energy multiplied by 2.

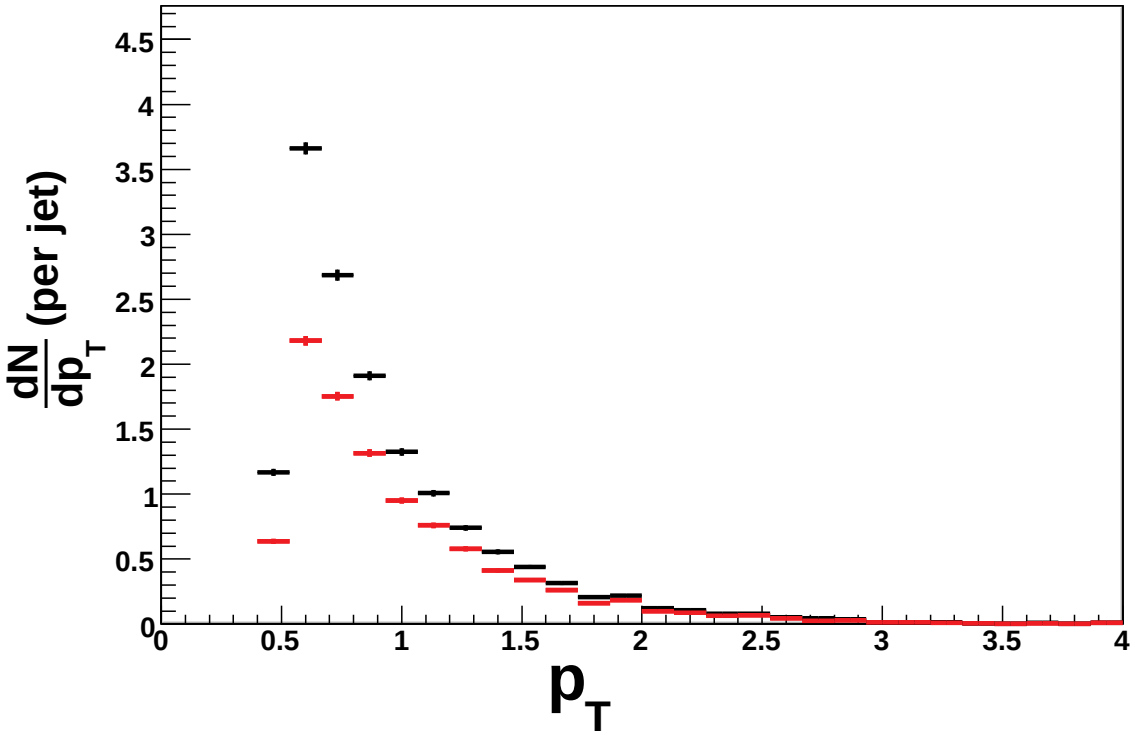

Figure 10. p_T of additional particles per p+p supplemented jet.
Black is all particles, where each BEMC hit is counted as a single "particle". Red is charged particles only.

The p+p+supp sample contains uncorrelated d+Au "minimum-bias" particles. Figure 10 shows the average number of these particles added to the jet as a function of p_T, for all particles and charged-only particles, as they contribute to the p+p+supp sample. The integral is roughly 1.5 particles per jet. The number of d+Au "minimum-bias" particles added to the jet has no discernable correlation to the energy of the jet. This is a simple consequence of the use of the simple cone algorithm and the way the p+p+supp sample is constructed. As long as the d+Au "minimum-bias" particles are not sufficiently abundant to allow the jet finder to produce a jet from the "minimum-bias" particles alone, the additional particles are quite isotropic.

Figure 11 shows the average difference in phi between the two- samples. The standard deviations of the distributions shown in Figure 11 are within a single BEMC tower width. In this case, the interpretation of the spread is straight forward since the

distribution is Gaussian centered at zero. The d+Au minimum-bias supplementation has very little effect on the thrust axis of the jet. Equivalent graphs for charged-only jets have identical behavior within uncertainties.

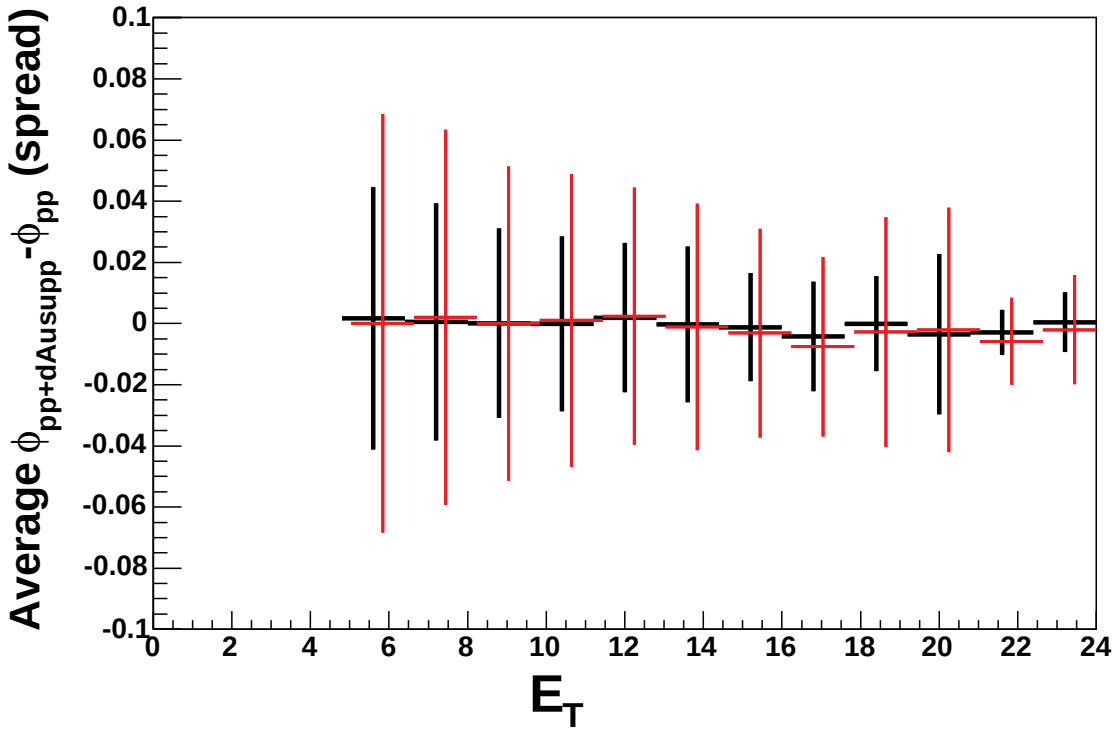

Figure 11. Average ϕ difference between plain and supplemented p+p jets. Bars indicate spread of distribution. Red indicates p+p supplement energy multiplied by 2.

B. PYTHIA

In order to interpret the data, the detector response to created particles must be understood. The STAR collaboration has chosen to simulate detector response using a software package called GEANT[19]. This is not enough by itself. Generating a set of input particles, as similar to the particles actually produced in a collision as possible, is

also required. For this purpose, STAR uses PYTHIA[10] simulations for comparison with p+p and d+Au data.

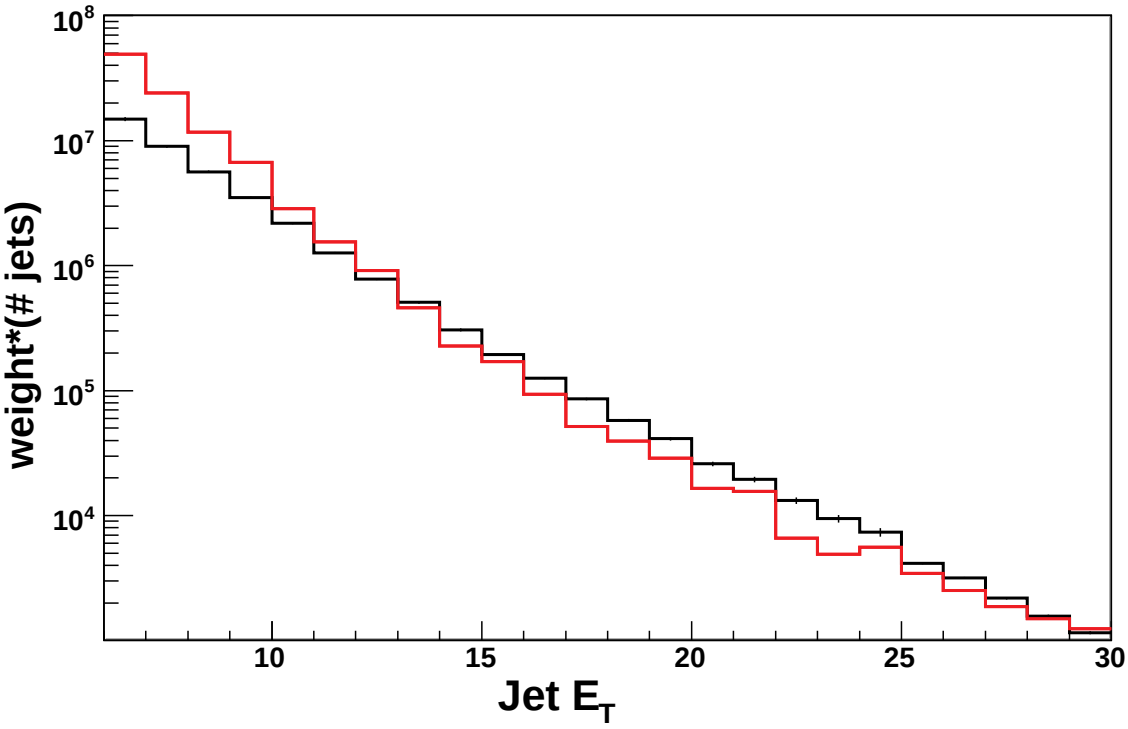

Figure 12. PYTHIA E_T distributions.
Jet reconstruction performed without detector simulation (BLACK), and Jet reconstruction performed after particles propagated through GEANT (RED).

The STAR Collaboration has created an array of PYTHIA data sets using PYTHIA version 6.205 and the CDF tune A[10]. These data sets are generated by specifying the momentum distribution of the initial jets. Each data set contains a plausible power law spectrum of initial jet momenta. Since the data sets must be finite, each data set is parameterized by a minimum-jet momentum threshold defining the bin. The data sets used for this analysis are defined in these bins: 2-3, 3-4, 4-5, 5-7, 7-9, 9-11, 11-15, 15-25, and 25-35 GeV. For many measurements, it is best to use jets from all these data sets for a more complete kinematic coverage. Since most of the samples have approximately the

same statistics, each data set must be given a specific weight to match an aggregate power law. The reconstructed jet p_T spectrum of the combined and appropriately weighted data sets is shown in Figure 12.

VI. CUTS ON RAW DATA

This section discusses cuts on data which are often applied to many different analyses in STAR, and possibly a few which are less universal but not necessarily restricted to jet reconstruction only. The cut on "high-tower" trigger data discussed in section V.B.3 to remove the systematic uncertainty arising from inconsistent "high-tower" triggers could technically go in this section, but it is so closely related to the "high-tower" trigger itself that it had to be placed in that context. Nonetheless, it is a good example of improving the data quality by removing parts of the data with a known bias. Jet specific cuts will be discussed as part of section VII, and details of all cuts will be given at the end if this section.

A. Sources of Systematic Uncertainty Motivating Cuts on Data

The STAR detector was designed around the assumption that the particle beams would collide within a certain region of space within the detector. STAR defines the center of this locus as (0,0,0). The beams are rather well confined in the x and y directions during normal RHIC operations. However, the location in z of the collision event, called the z vertex, may vary considerably. During normal physics running at RHIC, z vertex tends to follow a Gaussian distribution with a width of roughly half a meter, but this width can vary by a factor of two- depending on the quality of the fill. STAR was designed to ideally measure events with z vertex = 0, but moderate fluctuation of the z vertex is well within the detector design parameters. To properly reconstruct tracks, the STAR TPC needs to have a sufficient sample of hits. When the z vertex is too extreme, on the one side particles will leave very few hits in the TPC before exiting. The BEMC is also not designed for large z vertex positions. The BEMC towers are oriented toward (0,0,0), and so photons which come from other locations have sub-optimal detection efficiency.

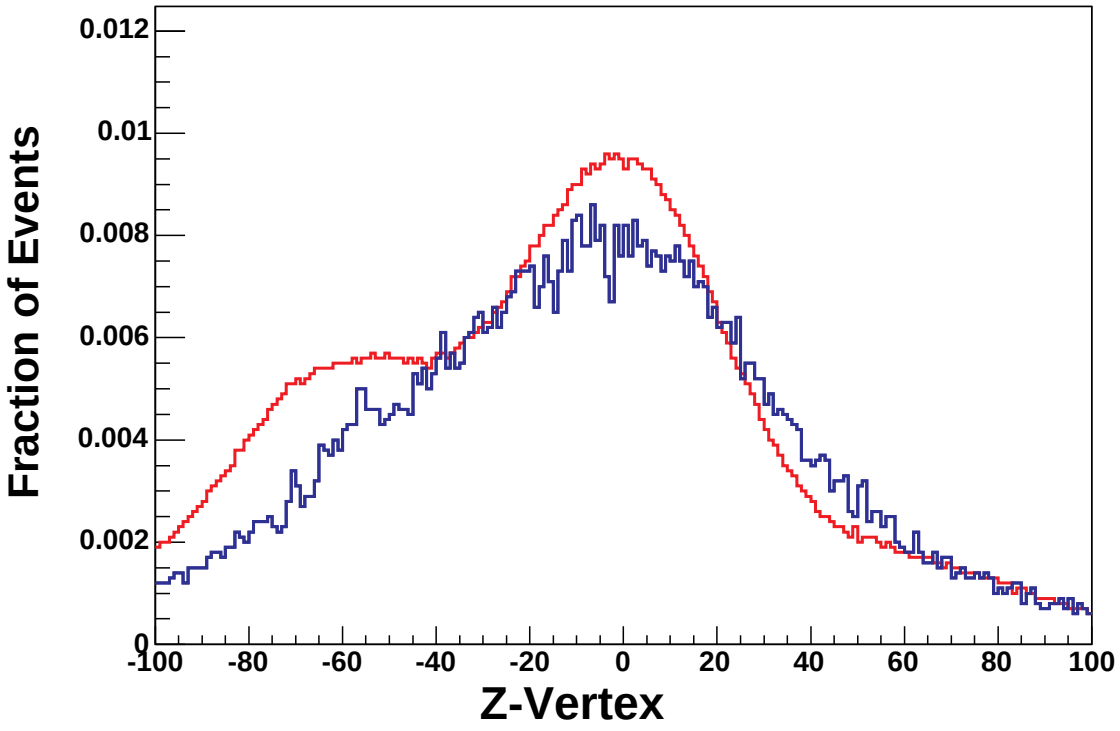

Figure 13. p+p z vertex distribution in observed events.
Blue is "minimum-bias" trigger, red is "high-tower" trigger. The events may or may not contain jets.

For this analysis, a z vertex cut of +/- 45 cm is applied. This is a software trigger, because the various STAR analyses are more or less sensitive to z-vertex location biases. STAR policy is to keep all rare triggers even if they are questionable until a complete analysis has been performed. Figure 13 shows the aggregate p+p minimum-bias z vertex distribution. As can be seen from Figure 14, the z vertex cut applied removes 40% of the total data. Figure 15 shows the aggregate d+Au minimum-bias z vertex distribution. Figure 16 shows the integral of Figure 15. For d+Au the z vertex cut removes 45% of the total data.

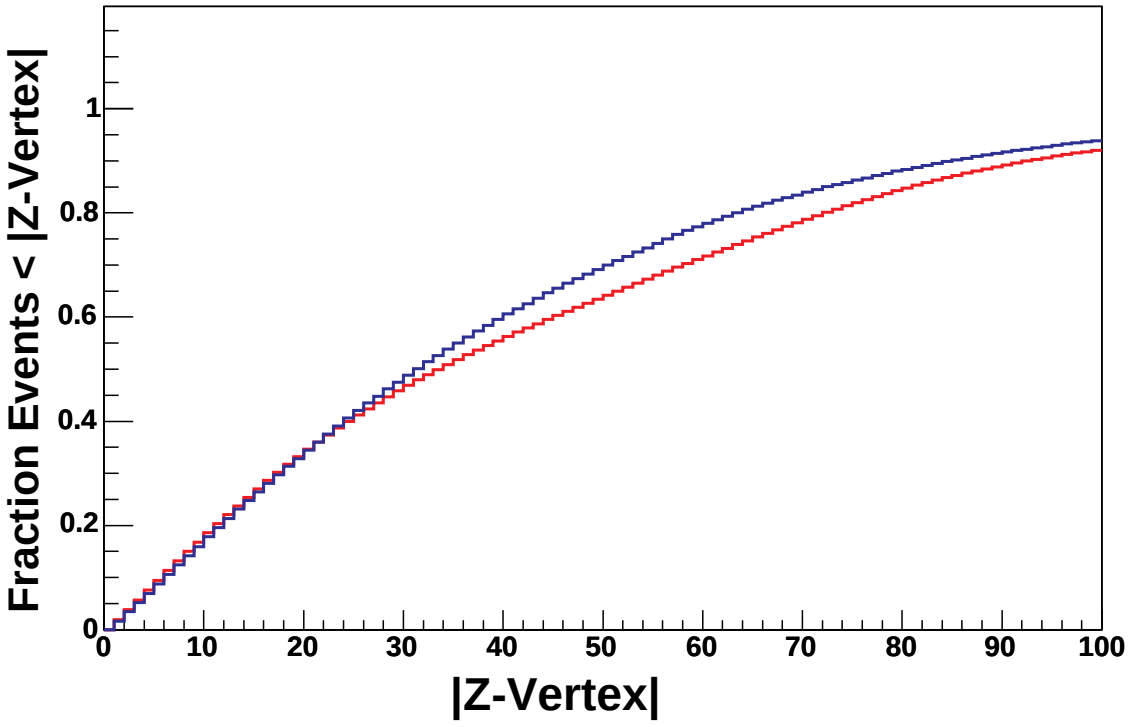

Figure 14. Fraction of events accepted in p+p.
The fraction of events accepted is shown as a function of the cut on the absolute value of the z vertex. Blue is "minimum-bias" trigger, red is "high-tower" trigger. Events may or may not contain jets.

The hits from the TPC can arise from other sources besides interesting collision hadrons. There are two- broad categories of spurious tracks. One category arises from collisions between the beam and other relatively stationary particles, and is termed "beam-gas" background. These sorts of collisions regularly occur at locations along the beam line a considerable distance from STAR, creating tracks which traverse the TPC or BEMC in a nearly horizontal fashion. These horizontal tracks have no relationship to a useful p+p or d+Au collision, but can nonetheless create hits within the detector. A reconstructed track which does not lead back to the main collision vertex clearly should not be used in the analysis. To eliminate these, STAR classifies as non-primary all tracks which do not pass within 3 cm of the collision vertex. This is called the DCA cut. The effectiveness of the DCA cut is weakened somewhat by the second source of background tracks, called "pile-

up" collisions. The TPC, having an extended readout time, commonly observes out-of-time tracks arising from collisions that occurred before or after the trigger fired. The vertex finding algorithm (ppLMV) described earlier simply finds "a" vertex. ppLMV does not perform optimizations. Given the fluctuations in p+p hadron production, especially in minimum-bias triggers, the vertex algorithm at times will snap on to a track nexus which does not correlate with the trigger condition. Due to the readout mechanics of the TPC, the out-of-time vertex moves away from the center of STAR as a function of how out-of-time it is, and so a z vertex cut is useful in suppressing them.

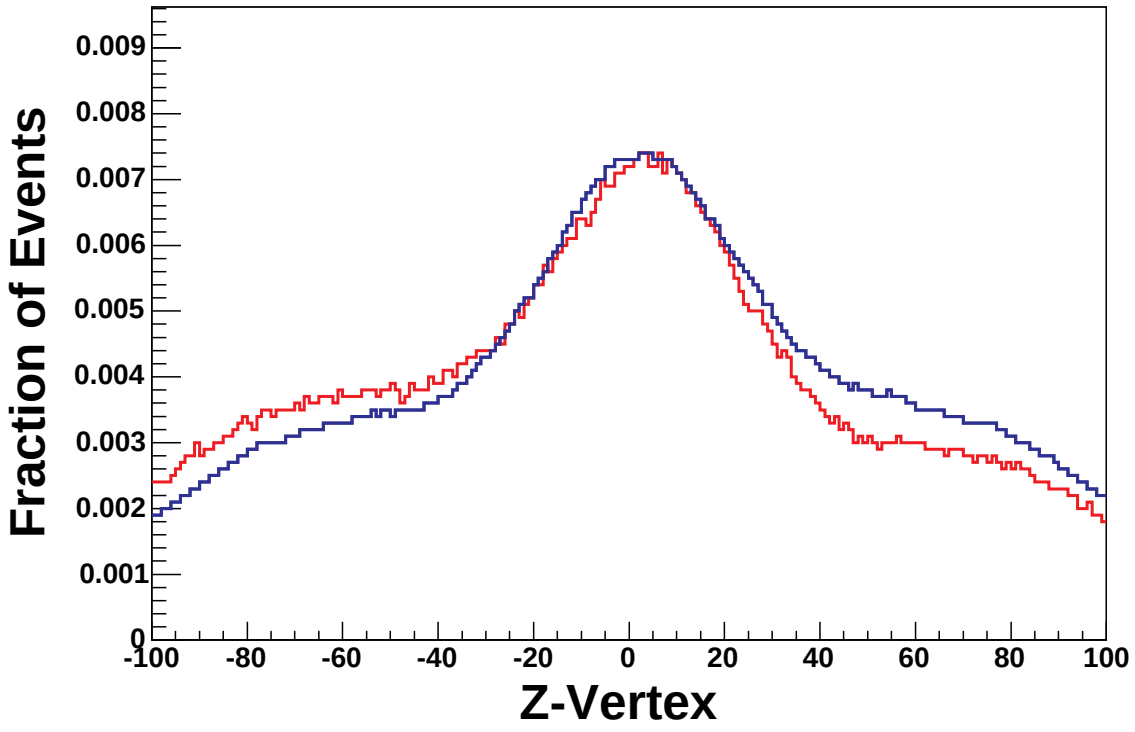

Figure 15. d+Au z vertex distribution.
Red is "high-tower" trigger, blue is "minimum-bias" trigger. Events may or may not contain jets.

Not all the background sources of hits even create reconstructible tracks. Spurious hits can distort the momentum measurement of tracked hadrons. STAR keeps a track flag

which tries to quantify the quality of the helix fit to the particle hits in the TPC. If this flag is not positive it means that the fitting algorithm ran into problems, for instance: the outlier removal eliminated too many points; there were not enough points to fit; there were too many fit iterations; there were too many outlier removal iterations; the outliers could not be identified; or perhaps there were not enough hits to start. All STAR physics results require this track fit flag to be positive.

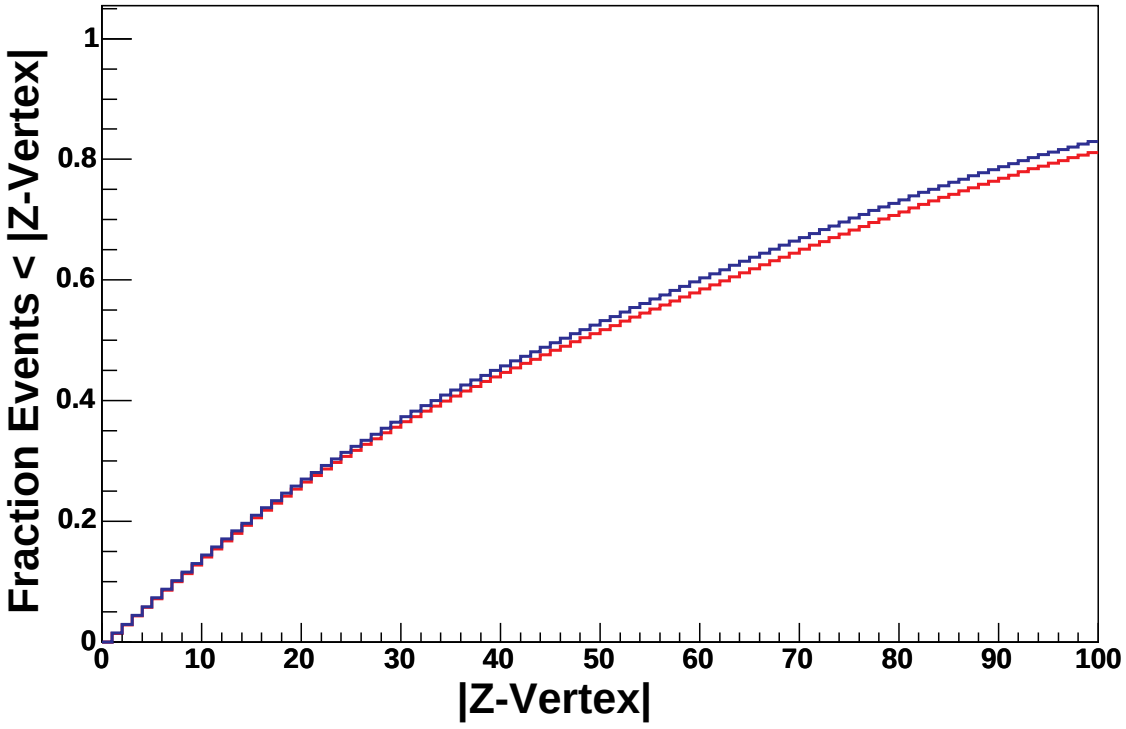

Figure 16. Fraction of events accepted in d+Au events.
Fraction of events accepted is shown as a function of the cut on absolute value of z vertex. Red is "high-tower" trigger, blue is "minimum-bias" trigger. Events may or may not contain jets.

Since the momentum determination of a hadron relies on accurately fitting a helix, an insufficient number of hits will result in large uncertainties. The greatest possible number of hits a track can leave is 45. Unfortunately, very few tracks leave that many

hits, so a more realistic requirement of 20 hits is used. This value of 20 helps ensure that tracks are the result of real hadrons in the TPC. In addition, the total number of possible hits for a given helical trajectory is well defined, and the ratio of the total number of hits in the helix to the number of hits possible for that helix is required to be greater than 0.51. This ensures that a single hadron will not create two- tracks and that it will only be counted once.

B. Systematic Effects of Background to Which Jet Signals Are Robust or Insensitive

The momentum resolution of tracks is optimal near a track p_T of about 0.6 GeV/c. However, the momentum resolution of tracks from 0.1 GeV/c all the way up to 3 GeV/c is less than a few percent [24]. For a reconstructed jet, these uncertainties even tend to average out. The largest uncertainty in the jet momentum comes from missing and or spurious particles, which each contribute roughly 100% of their respective momenta to the jet uncertainty. For jet reconstruction purposes, therefore, TPC track momentum resolution is insignificant.

The BEMC energy from each tower used by the jet finder is a function of the PMT ADC count for each event. In the process of calibrating each tower, scientists at STAR also generated a table of towers which had pathologies such as excessive ADC count rates or towers which could not be calibrated. This table is used in this analysis as a mask to block out towers whose responses are not understood. The number of towers in this mask varies from about 10% to as much as 20% near the end of Run 3.

The jet-energy scale is not strongly dependent on the number of towers in the status table. The "misunderstood" tower signals, if accidentally used in jet finding, would certainly throw off the jet-energy scale, since they commonly register unrealistically high energies. However, the rest of the towers do not introduce such instabilities. In an average unbiased jet, the neutral energy is about $\frac{1}{3}$ of the total energy. This is simply the fact that π^0, π^+, and π^- are more or less equally probable. In the "high-tower" triggered

data sample, the average neutral energy moves upward to about ½ of the total energy. The probability of losing a companion particle if it is neutral is less than 20%. Since the trigger particle is a given, a companion neutral particle would have on average less than half of half the total energy. The average energy scale effect of losing this neutral companion due to the status table works out to about 20%*25%=5%. The variation of the energy scale over time due to status table changes is even smaller, only (20%-10%)*25%=2.5%. Energy scale changes of this magnitude may have a significant effect on the cross section, but systematic uncertainties of the measurements done in this analysis are only modestly affected.

The distribution shapes undergo an even smaller bias due to the status table mask. For example, if a 6 GeV jet, made up of a 2 GeV/c π^0, π^+, and π^-, loses the π^0 to the status table mask, it will be reconstructed as a 4 GeV jet, and have little effect on the far more numerous complete 4 GeV jet sample. Meanwhile, the 6 GeV jet sample will only contain a trickle of more rare 8 GeV jets likewise reduced by the status table. The expectation, therefore, is that distribution shapes should have very little dependence on the status tables. Comparisons of the effects of status tables on simulated data, shown in the next section, are consistent with this expectation.

C. Summary of Data Cuts

This section lists the cuts actually applied to the data for this analysis. Since they differ, the "minimum-bias" and the "high-tower" data cuts are listed separately. The z-vertex cut, the total to neutral energy cut, and the cone radius parameter are described elsewhere in the text, but also mentioned here for completeness.

1. "Minimum-Bias" Data Cuts

Cone radius = 0.5 radians. For PYTHIA without GEANT cone radii used are 0.5 and 0.7.

Neutral over total energy ratio is between 0.1 and 0.8.

Jet pt > 6.0 GeV/c.

z-vertex is between -45 cm and 45 cm.

The projected location of a jet striking a cylinder of radius 2.0 m is between 40 and 130 cm when BEMC hit energy is included in jet reconstruction. The projected location of a jet striking a cylinder of radius 2.0 m is between -140 and 140 cm when no BEMC hit energy is included in jet reconstruction.

All events with more than 2 reconstructed jets are excluded.

Particles in jet are above 0.2 GeV/c for p+p, and 0.6 GeV/c for p+p+supp and d+Au. This applies to both tracks and BEMC hits.

Number hits per track > 20. Ratio hits per track to total possible hits per track > 0.5. Track flag > 0.

Minimum seed $E_T > 0.5$.

Split Fraction = 0.5.

DCA < 3.0 cm.

2. "High-Tower" Data Cuts

Only "high-tower" data cuts which are different from the "minimum-bias" data cuts are listed.

Cone radius = 0.7 radians for p+p and PYTHIA. Cone radius = 0.5 for p+p+supp and d+Au.

For charged plus neutral energy jets, jet pt > 9.0 GeV for cone radius 0.7, and jet pt > 8.0 for cone radius 0.5. For charged-only jets, jet pt > 7.0 for cone radius 0.7, and jet pt > 6.5 for cone radius 0.5.

For charged-only jets, jet η is required to be between -0.5 and 0.5.

Maximum high-tower in jet is greater than 2.8 GeV.

VII. DATA ANALYSIS AND RESULTS

The motivation for reconstructing jets has been described in previous sections, as well as issues relating to experimental setup, standard STAR event and particle reconstruction, the abstract jet finding algorithms which operate on particle four momenta, the data runs used as sources for events, and the cuts on the data. This section discusses the results of running the jet finder, cuts on the jets, comparison of jets reconstructed from the data sources, and calibration of jet results based on simulations. In particular, this section describes the jet j_T, z, and k_T measurements made during the course of this analysis. The next section will discuss the larger context in which these measurements belong.

A. General Jet Reconstruction Issues of Interest

1. Low Energy Jets

A cut on minimum-jet E_T, while very simple to apply, is very important. Contamination from extraneous particles affects low E_T jets the most. In addition, low E_T jets have greater fluctuation in particle number, making detector E_T resolution and ϕ resolution worse. As mentioned earlier, if particles are sufficiently low in p_T, they are eliminated from consideration before the jet finder is even run. The fractional energy loss from ignoring these particles is greatest for low E_T jets. Yet another complication arises when E_T is approximately equal to k_T. It is not clear whether, under these conditions, a jet can arise without a partner jet on the away-side, solely by recoiling against k_T. If so, it is certainly unclear what effect this would have on the jet fragmentation properties.

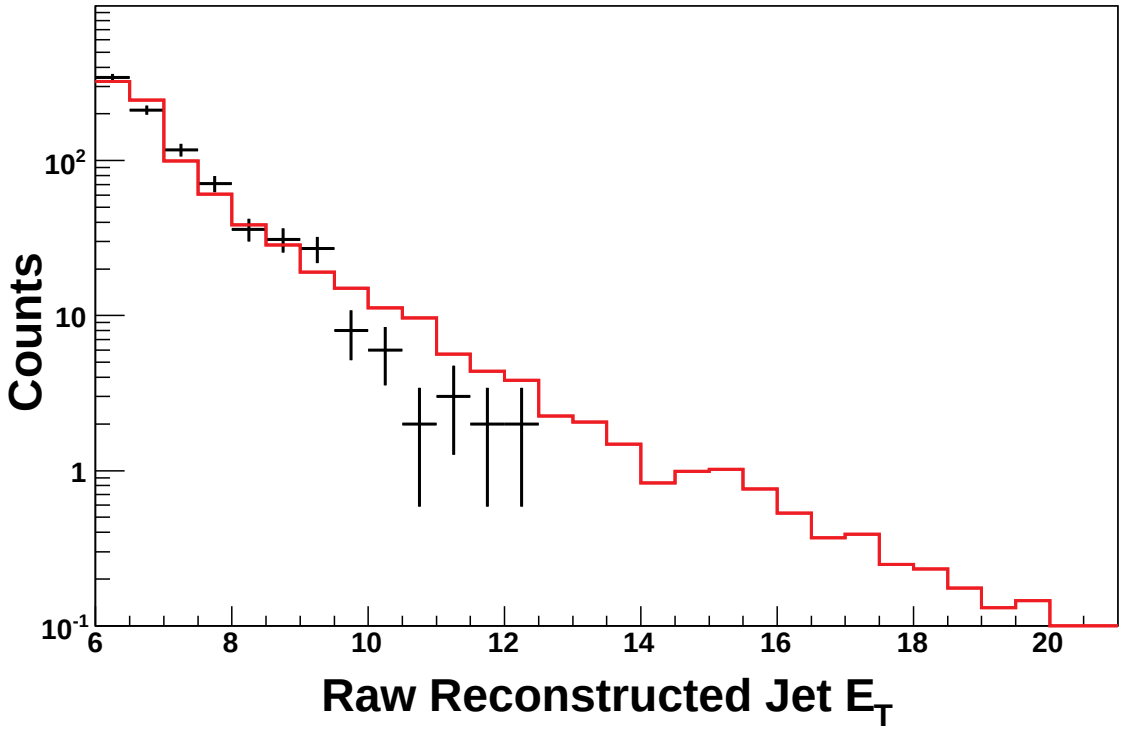

Figure 17. Raw reconstructed jet E_T for "minimum-bias" jets.
Black points are p+p data. Red is PYTHIA weighted according to thrown E_T.
PYTHIA is arbitrarily scaled for comparison.

Naturally, lower E_T jets are more numerous, granting better statistics, so it is useful to know the minimum-E_T jet which is safe to use for measurements. The analysis described here uses a minimum-E_T cut of 6 GeV for charged plus neutral energy jets. The results from jets reconstructed below 6 GeV have questionable properties. For example, the z distribution behaves quite differently below and above this E_T cut. Figure 17 and Figure 18 show the p+p data raw reconstructed jet E_T compared with PYTHIA simulation. The E_T agreement between PYTHIA and data is such that PYTHIA should be a good estimator for the detector energy resolution. "High-tower" triggered jets in this analysis are always 8 GeV or greater, and this is the range which Figure 18 displays.

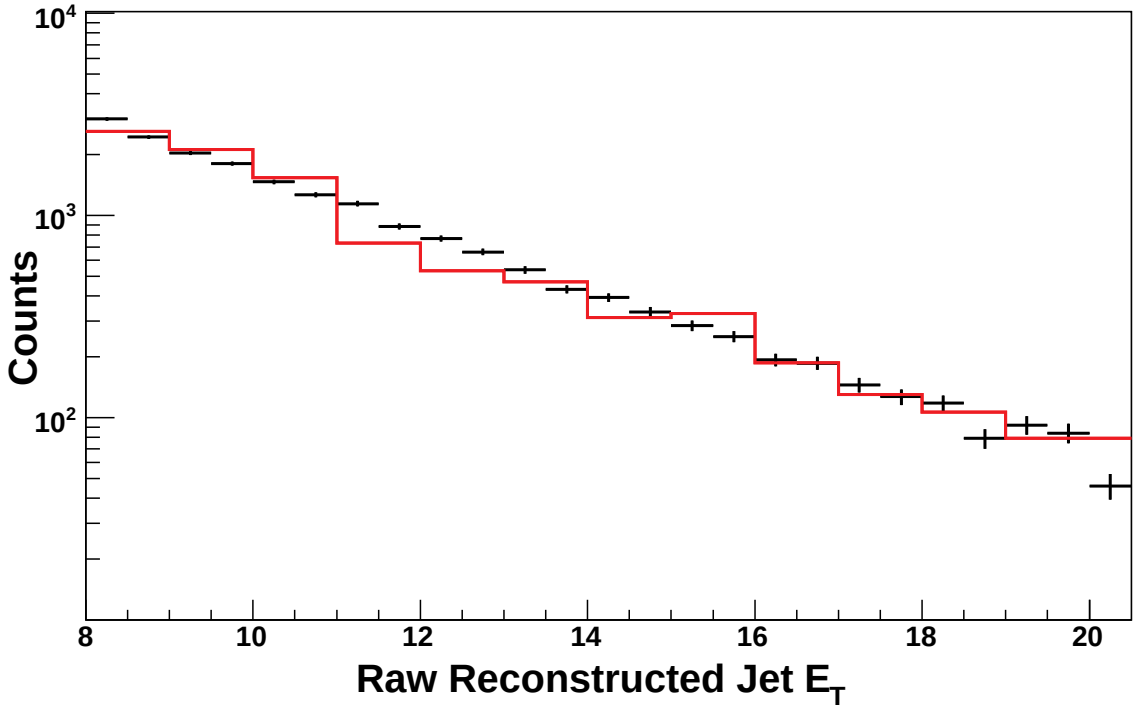

Figure 18. Raw reconstructed jet E_T for "high-tower" jets.
Black points are p+p data. Red is PYTHIA weighted according to thrown E_T.
PYTHIA is arbitrarily scaled for comparison.

2. Jet Level Cuts

Like all algorithms, the jet finder suffers from the "garbage in, garbage out" syndrome, which is one reason why the cuts from section VI are so important. Those cuts are standard cuts applied by most STAR analyses, but the jet finder is sensitive to types of "garbage" which charged-di-hadron analyses, for example, are not. Some of these forms of "garbage" are difficult to characterize at the particle level, and even when it is possible, it may be easier and more meaningful to introduce cuts on jet properties. In addition to cuts, this analysis uses standard parameters for jet reconstruction across the data sets, and these parameters follow a similar application framework to the cuts.

For all measurements in this analysis, one requirement is that no more than two-jets get reconstructed. Jets from events with three real jets can have very different properties when compared to real jet pairs. The events which do get reconstructed with three jets come from both physical tri-jets and events with excessive EMC energy. Since there are not enough three-jet events to produce statistically meaningful results, it is merely a nuisance that beam background commonly gets reconstructed by the jet finder into three or even more jets. This cut is most significant for two-jet relative measurements, since the number of events with single jets far exceeds the number of events with more. In effect, this is a requirement that the jet finder finds exactly two- jets when measuring di-jet quantities.

The cut on the minimum tower energy can be applied to the jets after jet reconstruction. It is usually most convenient to apply it after jet reconstruction to decrease the time it takes to fine tune the cut. Cutting jets is comparable to cutting events which do not have a tower above 2.8 GeV, but cutting jets makes the comparison between simulation and data more direct. Cutting jets or cutting events are not exactly equivalent; sometimes the jet finder will fail to reconstruct the jet which triggered the event, but successfully reconstruct that jet's partner. Some jets are therefore cut out because they failed to reconstruct properly. This is advantageous. Incidentally, the very interesting signal of the prompt photons would require a different policy, but no matter how interesting they are, prompt photons are not jets. It is better not to count prompt photons as "high-tower" triggered jets. The jet level minimum BEMC tower cut of 2.8 GeV is slightly superior to the event level version of the same cut.

The cone radius is a parameter that applies to all jets reconstructed with a cone algorithm. All jets used for this analysis are of this type. It is mentioned here because from a certain point of view it is a cut, and because this parameter is enforced at the same time that the other cuts mentioned in this section are enforced. The jets are initially reconstructed with several different radius parameters, but they all get stored afterward in the same pool. After jet finding is complete, and when the histograms and graphs are being created, the cone radius parameter is constrained to the value of interest.

The BEMC can at times exhibit large energy fluctuations which the jet algorithm will incorrectly interpret as jets. These energy fluctuations are believed to be caused by high-rapidity beam gas background particles traversing the BEMC. Particles, created in the region of STAR where collisions are designed to occur, normally only fire a single tower, or at most a few adjacent ones. It is believed, however, that high-rapidity particles can create showers in a series of towers, since they are moving almost parallel to the beam. The trajectories of such particles are believed to radiate from the focusing magnets closest to STAR where the products of beam gas collisions get kicked outward. The approach taken to weed out these background signals is to monitor the ratio of "neutral" particle energy to total particle energy. The word "neutral" is in quotes because this energy often contains a contribution from electrons and anti-proton annihilations, and in bad cases almost exclusively unphysical background.

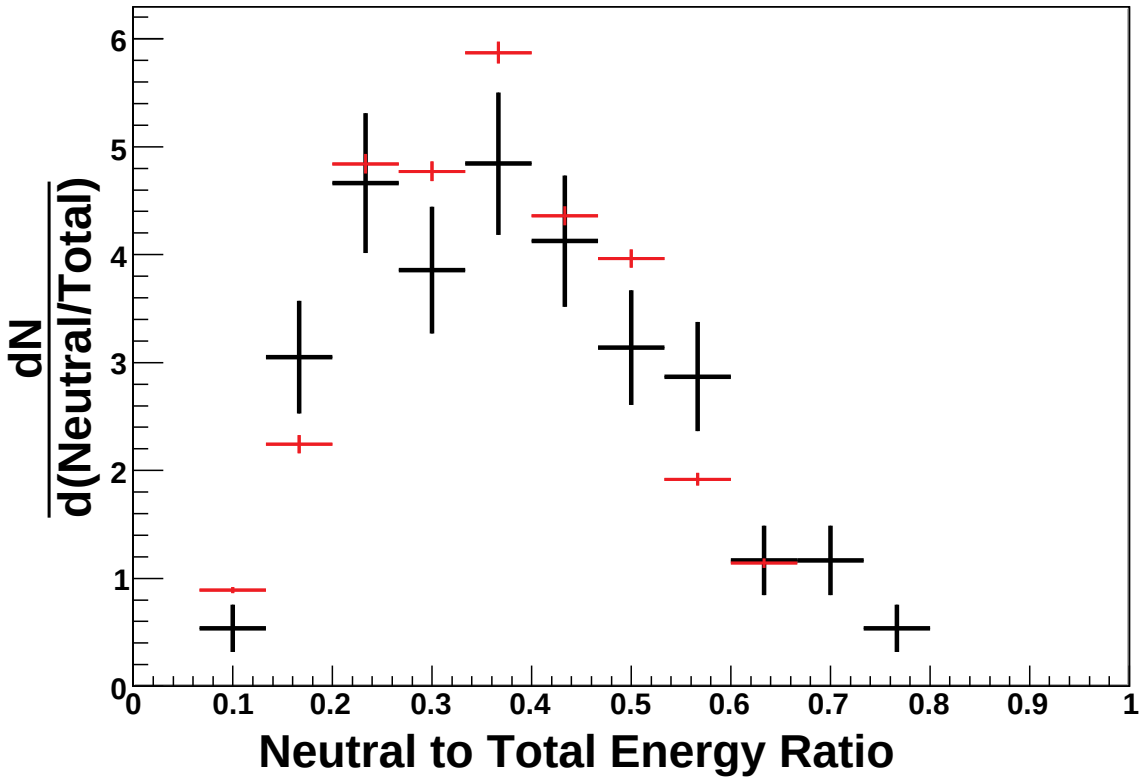

Figure 19. Neutral/total E_T ratio for "minimum-bias" jets.
All jets > 6 GeV. Black is data, Red is Simulation. Weight is 1 for data.

Single jet measurements, such as z and j_T discussed in part D, are made using both the "minimum-bias" triggered data sample, and the "high-tower" sample at a larger E_T. The neutral energy is required to be less than 80% of the total energy for these measurements. Figure 19 shows how the simulation and data compare in the ratio of neutral energy to total energy. Most of the jets pass this cut. For studies based on the "high-tower" triggered sample, such as the di-jet analysis, the neutral energy ratio is also required to be less than 80%. Figure 20 shows the di-jet neutral energy curves from simulation with a minimum tower requirement of 2.8 GeV and for "high-tower" triggered data.

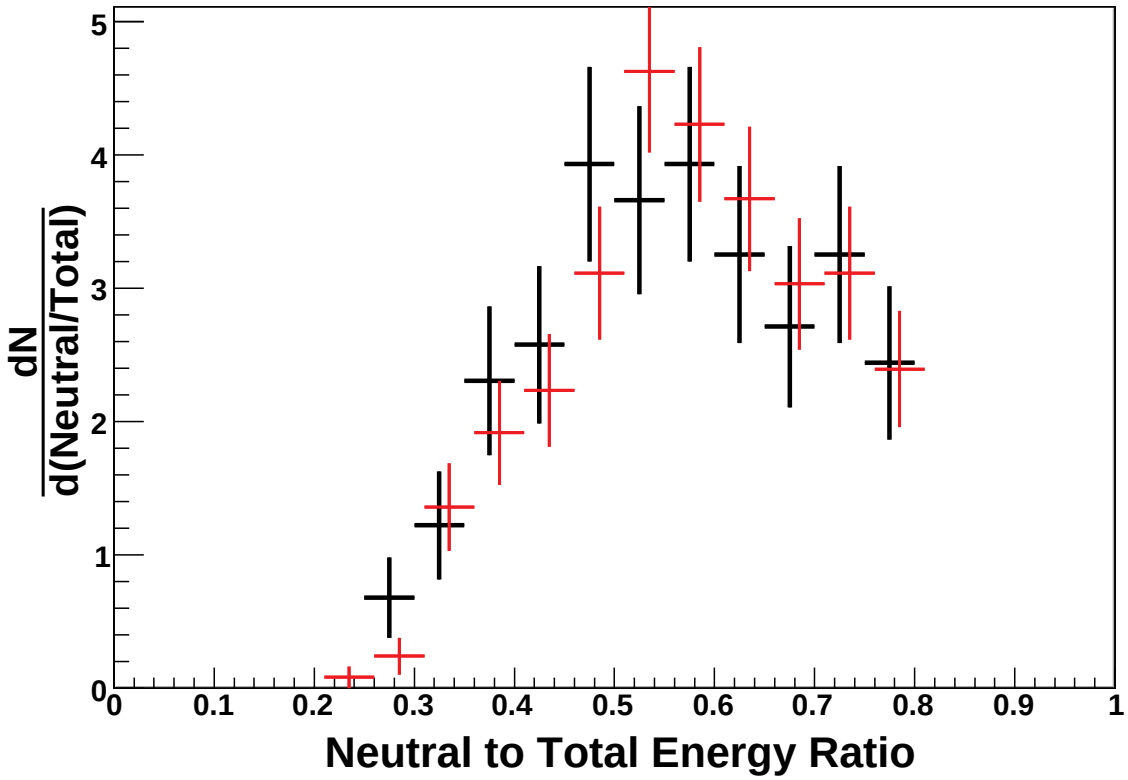

Figure 20. Neutral/total E_T ratio for di-jets.
"High-Tower" trigger used with one tower above 2.8 GeV. Black is data, Red is Simulation. Weights are 1 for data.

3. PYTHIA Simulated Jets

There are two steps in the generation of simulated events from PYTHIA. First, PYTHIA is used to generate a set of particles which is stored in a table according to the STDHEP standard[**]. In the second step, this information is propagated through GSTAR, the STAR adaptation of GEANT [16]. GSTAR is used to simulate the STAR detector response to the particles generated by PYTHIA.

The jet finder is run on the results of both of these steps. PYTHIA no-GEANT jets are reconstructed from four-momentum lists formed from the contents of the STDHEP

[**]STDHEP version 5.04 documentation, esp. pg. 5 at http://cepa.fnal.gov/psm/stdhep/

table. No-GEANT jets are as close to ideal jets as possible. PYTHIA with-GEANT jets are reconstructed from four-momentum lists formed from the simulated detector response to those same particles. PYTHIA with-GEANT jets are comparable to jets reconstructed from real data. By comparing no-GEANT and with-GEANT jets, the detector response and jet resolutions can be obtained. Several different detector level jet resolution quantities can be considered, such as E_T resolution, ϕ resolution, z resolution, and many other quantities.

4. Charged-Particle-Only Jets

The jet finder can search any list of four vectors given it for jets. One can easily limit these four vectors to charged particles found in the TPC. In fact, this was originally the only configuration possible before this author extended the software to accept BEMC energy. It is sometimes advantageous to compare jets with both neutral and charged energy to jets with only charged energy. For example, during Run 3 the BEMC had a few patches in ϕ with much reduced acceptance (due to dead PMT crates). Patches of low ϕ acceptance appear to "repel" jets, in that the jet thrust axis is much less likely to land in a region of low acceptance. This may cause reduced energy scale, increased ϕ fluctuations, and in some cases even change the average opening angle between two jets. The TPC, comparatively, is extremely uniform in ϕ. Measuring $\Delta\phi$ using a charged-only away-side jet helps to resolve problems with the non-uniformity in the BEMC. Figure 21 shows the reconstructed phi resolution, and Figure 22 shows the reconstructed charged-particle-only jet phi resolution.

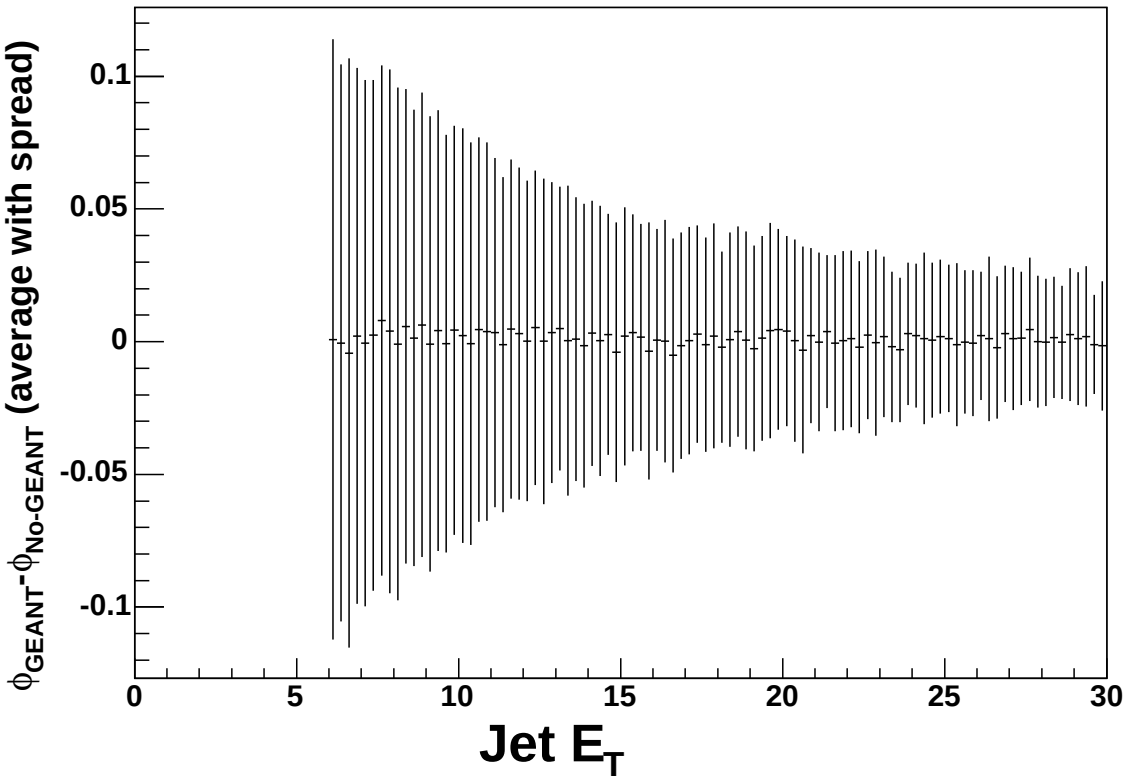

Figure 21. Spread of ϕ difference without and with-GEANT.
Difference is between PYTHIA simulated jets with and without detector response simulation.

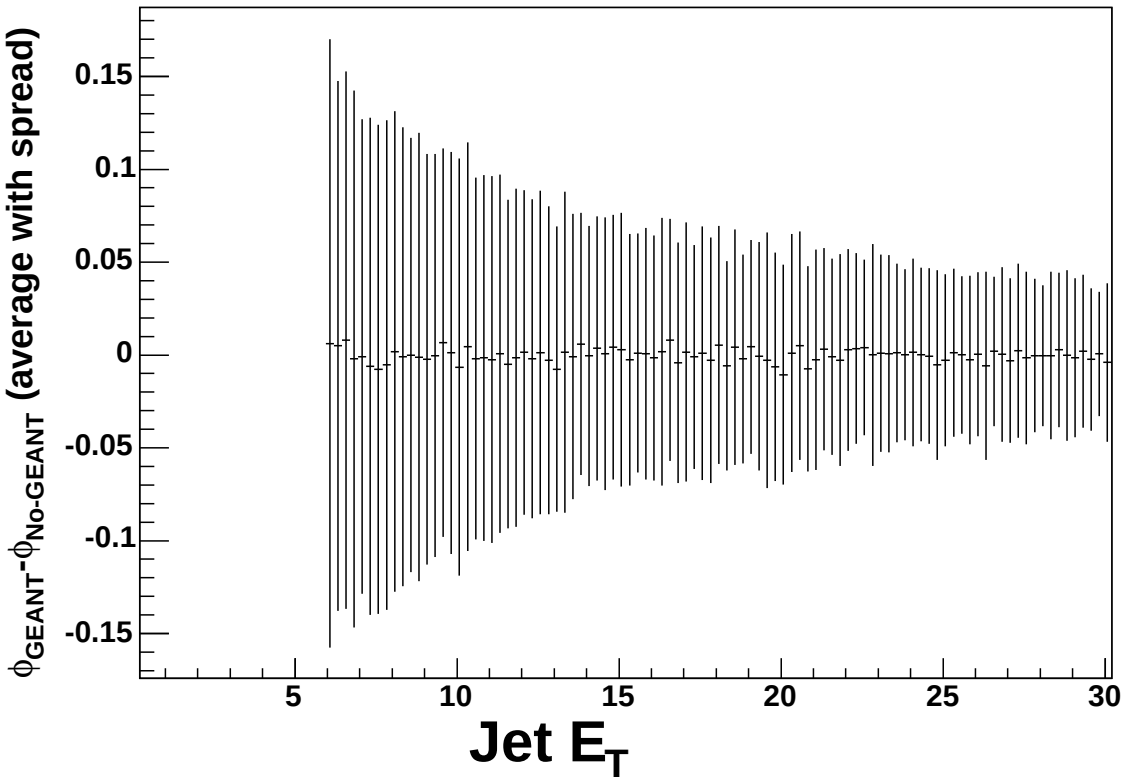

Figure 22. Spread of ϕ difference without and with c.p. only GEANT. Difference is between PYTHIA simulated jets with and without Detector response simulation. For PYTHIA no-GEANT all particles used, for PYTHIA with-GEANT only charged particles are used.

B. Jet E_T and Phi Resolution from PYTHIA Simulation

The E_T resolution is not easy to determine to high precision. This is due to two opposing tendencies in jet-energy reconstruction. One tendency is that as E_T increases, the absolute energy difference between true energy and reconstructed energy tends to increase proportionally. The other tendency is that as E_T increases, more high-energy particles are produced. Higher energy particles are more collimated and less likely to be missed in jet reconstruction, lowering the difference between true and reconstructed jet energy.

Figure 23 shows the average difference between PYTHIA no-GEANT jet E_T minus GEANT jet E_T as a function of the no-GEANT jet E_T value. For Figure 23, all jets are weighted evenly, regardless of the likelihood they will be produced in a collision. This means that if a single no-GEANT jet is chosen, then given the E_T of that jet one would get on average the energy difference shown in Figure 23. Figure 24 is identical with the difference that it is displayed as a function of the jet E_T with-GEANT. Figure 24 seems more useful for predicting the energy scale, since it is a function, ostensibly, of the energy reconstructed in the detector. However, Figure 23 and Figure 24 do not take into account the fact that lower energy jets are far more numerous. When one asks the question: "Given a PYTHIA only jet E_T, what is the average reconstructed jet E_T?", one is implicitly utilizing a PYTHIA no-GEANT E_T distribution that is *flat*.

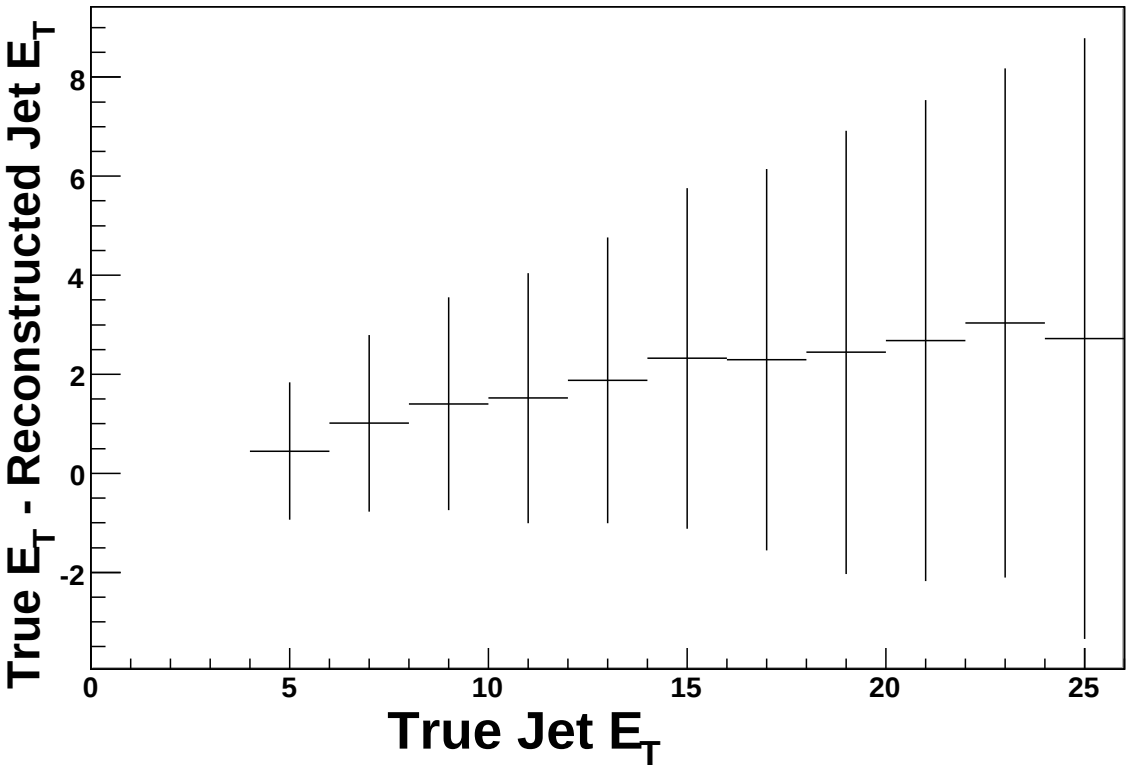

Figure 23. Spread of E_T detector response given true jet E_T vs. true E_T. Vertical axis shows average no-GEANT jet E_T – GEANT Jet E_T with spread given a specific no-GEANT jet E_T. The no-GEANT jet E_T distribution is *flat*.

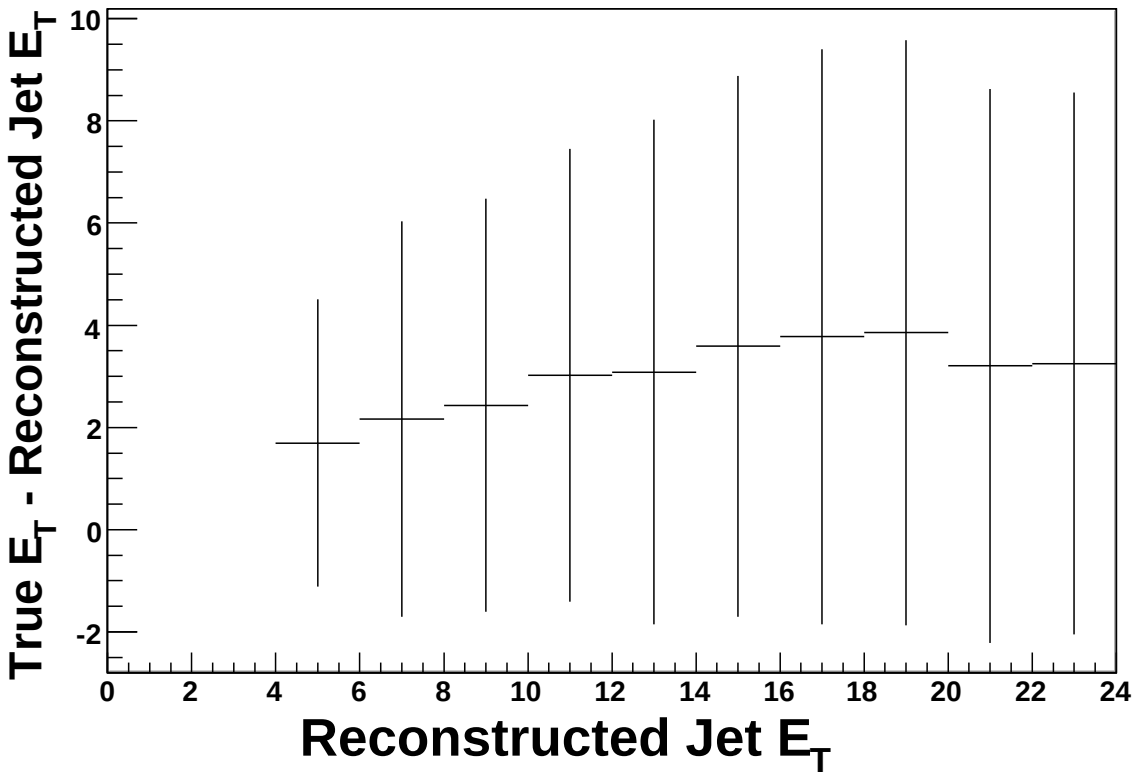

Figure 24. Spread of flat E_T detector response vs. reconstructed E_T .
Vertical axis shows average no-GEANT jet E_T – GEANT Jet E_T with spread versus the reconstructed jet E_T . The no-GEANT jet E_T distribution is *flat*.

Figure 25 shows PYTHIA no-GEANT jet E_T minus GEANT jet E_T as a function of E_T with the proper power law distribution weighting the no-GEANT jets. For most of the E_T range shown, nearly 2 GeV of excess energy is reconstructed *for a given reconstructed* E_T. This is possible because the energy response of the BEMC fluctuates essentially without bias both up and down. However, since the BEMC is not biased with respect to the direction of energy fluctuation, jet reconstruction is therefore biased against jets with low BEMC response. For example, the ratio of the number of jets in the intervals 7-9 and 9-11 GeV in the simulation used is about 8:3. As seen in Figure 23, the real jet E_T can easily fluctuate up or down 2 GeV. Imagining that each jet has equal probability of moving up 2 GeV, down 2 GeV, or unchanged, it would not seem strange for 1 unit of jets to remain in the 9-11 GeV bin, and for 2 units of jets to move up from 7-9 to the 9-11 GeV bin. Likewise, the number of jets dropping down to 9-11 GeV from 11-13 GeV would be small, so it is easy to believe that 2 out of three jets might be overestimated by 2 GeV in the 9-11 GeV bin.

Figure 26 and Figure 27 show PYTHIA no-GEANT jet ϕ minus PYTHIA with-GEANT jet ϕ for the same underlying jet for all jets above 6 GeV. These figures show that detector ϕ resolution from charged-only jets is almost as good as from charged + neutral jets. The detector ϕ resolution is a little more, and a little less, respectively, than the width of two BEMC towers.

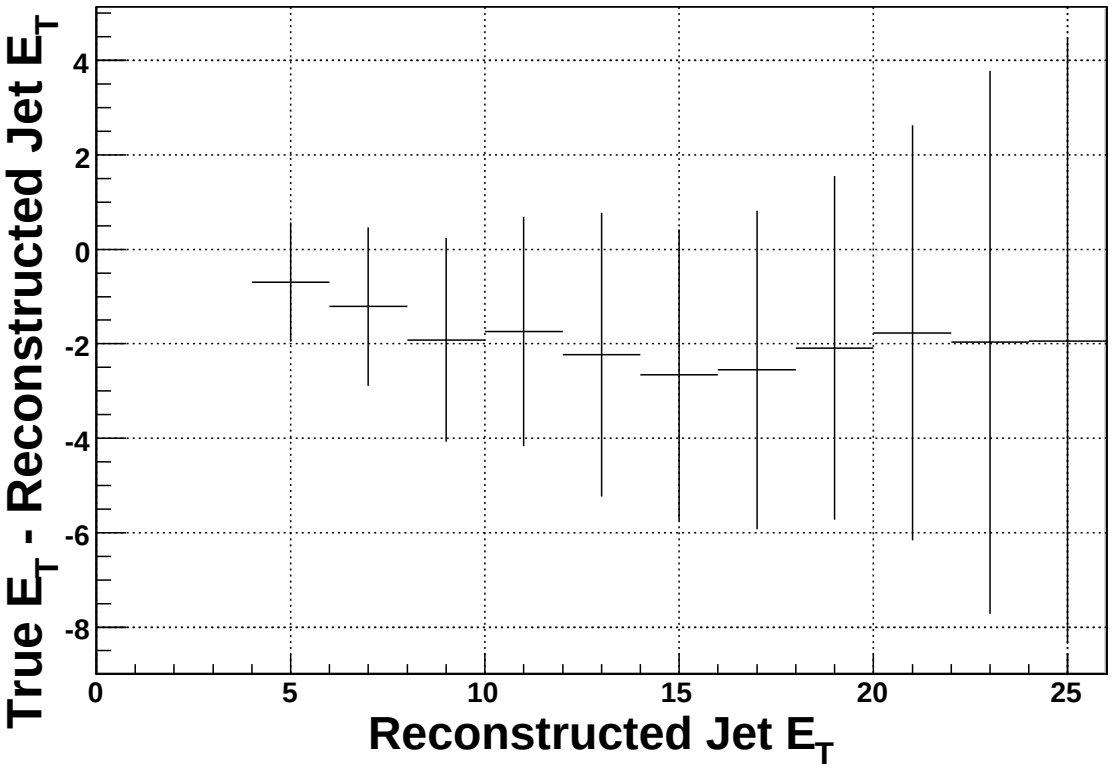

Figure 25. Spread of E_T detector response vs. true E_T .
Vertical axis shows average no-GEANT jet E_T – GEANT Jet E_T with spread given proper power law weight to no-GEANT jets.

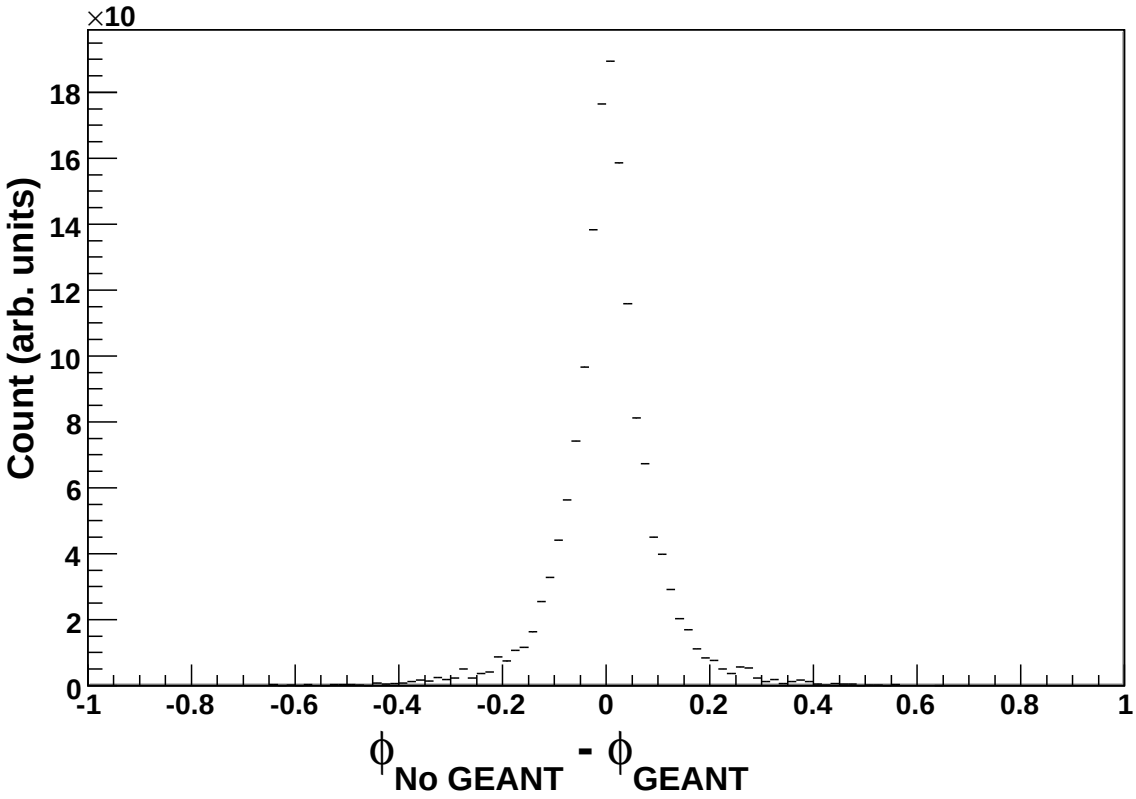

Figure 26. PYTHIA no-GEANT – with-GEANT ϕ difference.
ϕ difference between jets reconstructed with and without GEANT detector response simulation. RMS width is 0.094.

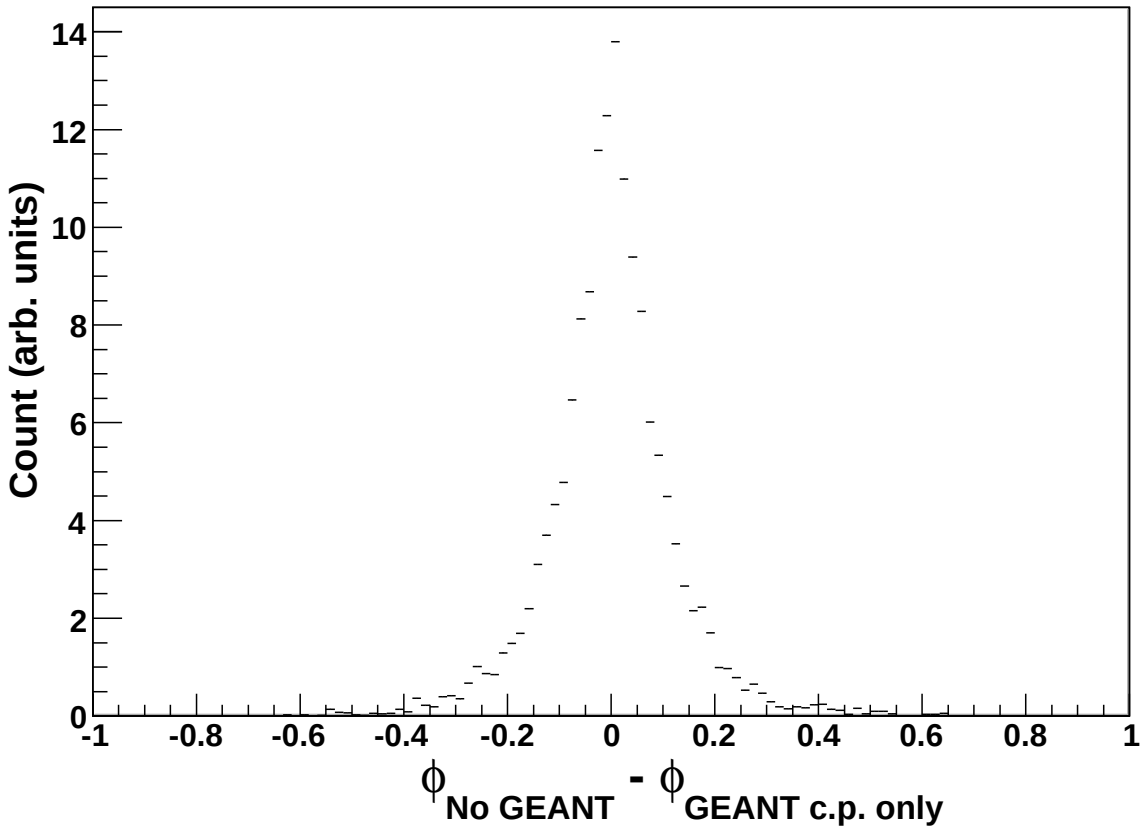

Figure 27. PYTHIA no-GEANT – with c.p. only GEANT ϕ difference. ϕ difference between jets reconstructed with and without GEANT detector response simulation. No-GEANT jets use all particles. GEANT jets use only charged particles. RMS width is 0.122.

C. Jet Fragmentation Data to PYTHIA Comparisons

The natural way to quantify the jet samples is to examine the "shape" of the jets obtained. The two main observables for jet fragmentation conceptually break the particle momentum distributions down into perpendicular and parallel components with respect to the jet thrust axis. Since the raw reconstructed jet sample is somewhat biased by non-ideal detector response and the jet reconstruction algorithm, it is important to correct the distortions introduced by the detector and software. By comparing direct PYTHIA jets (PYTHIA no-GEANT) with PYTHIA jets run through the GEANT detector simulator and then reconstructed by the software (PYTHIA with-GEANT), the detector response and algorithmic biases can be measured. Then, by comparing jets from data with PYTHIA with-GEANT jets, the true shapes of the jet fragmentation distributions can be extracted from the data.

1. Jet z

For the purpose of this dissertation, the variable z is defined as the ratio of the hadron momentum to the momentum of the jet. This z distribution is identical to the distribution commonly referred to as the fragmentation function. It quantifies the "shape" of the jet parallel to the jet thrust axis. Reconstructing neutral pions in the BEMC is a complex analysis that is not a part of this dissertation. This means that only the aggregate energy and momentum from the neutral particles is used for jet finding, but the individual properties of the neutral particles are not reconstructed. Since the momentum of individual neutral hadrons is not available, all the z distributions (and j_T distributions) shown here are charged-particle-only distributions. For clarity in the figures, charged-particle z and j_T distributions will be indicated with c.p. (charged particle), but in all cases the measured quantities discussed in the text refer to charged-particle-only distributions. It is not expected that neutral-particle distributions behave differently from their charged-particle counterparts, but no conclusion can be drawn from this analysis.

When measuring the fragmentation function, certain sources of systematic uncertainty need to be examined. Particles which do not come from mid-rapidity parton fragmentation can be generated by the usual beam-rapidity collision fragments. They can also come from particles entering the detector which have no intrinsic relationship to the collision event. Almost all of these particles are low energy, and when they are not discarded by the DCA cut, as a rule they will distort the fragmentation function only at low z. The unphysical "particles" produced in the BEMC by the high-rapidity beam gas background are the main exception. These "particles" create highly concentrated neutral energy, raising the neutral energy content of the reconstructed jet. By cutting out the jets with the most neutral energy, the impact of the background particles on the fragmentation function can be checked. Figure 28 shows how the z distribution is mostly insensitive to this neutral energy cut.

To assure that the data are directly comparable to the PYTHIA simulation, the data and the simulation are processed exactly the same whenever possible. The most significant difference in processing comes from the fact that the simulated BEMC is flawless: in theory every PMT works. To address this source of systematic uncertainty, the tables of unused PMTs are borrowed from the data and applied to the results of the simulation.

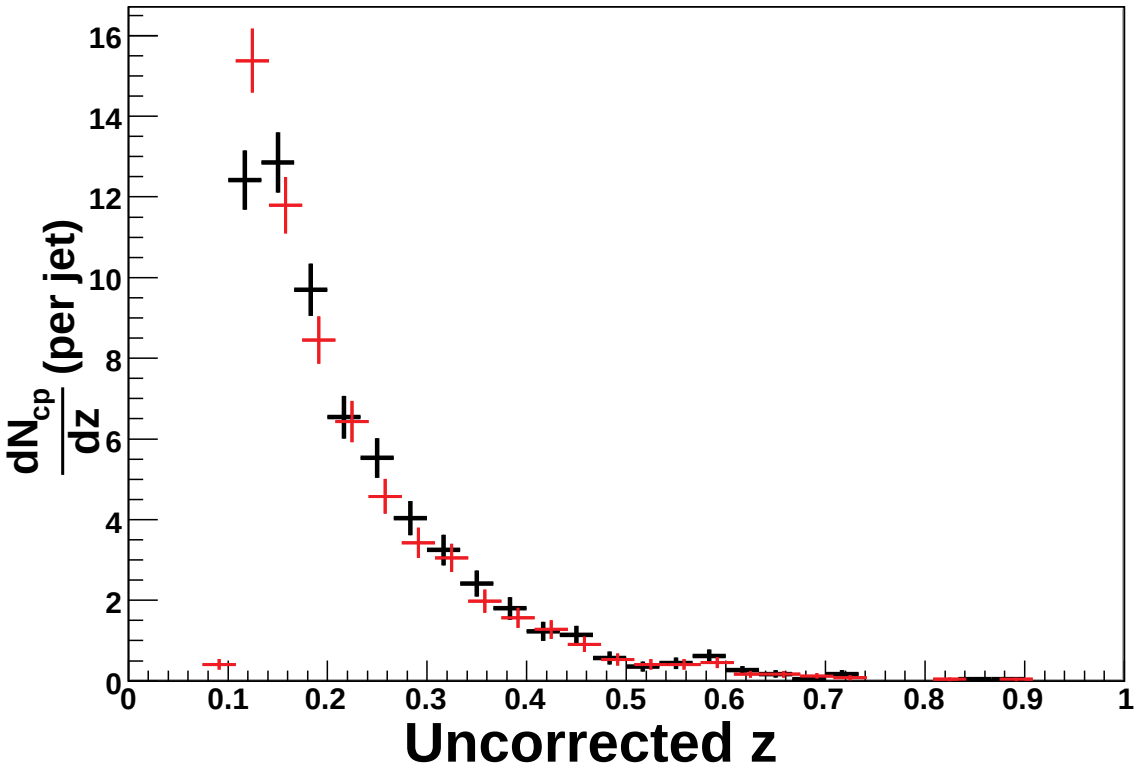

Figure 28. Jet c.p. z distribution for two neutral energy ratio cuts. The ratio is the neutral energy divided by the total energy. Black is with ratio < 0.65, red is with ratio < 0.8. Distributions are uncorrected, charged particles only.

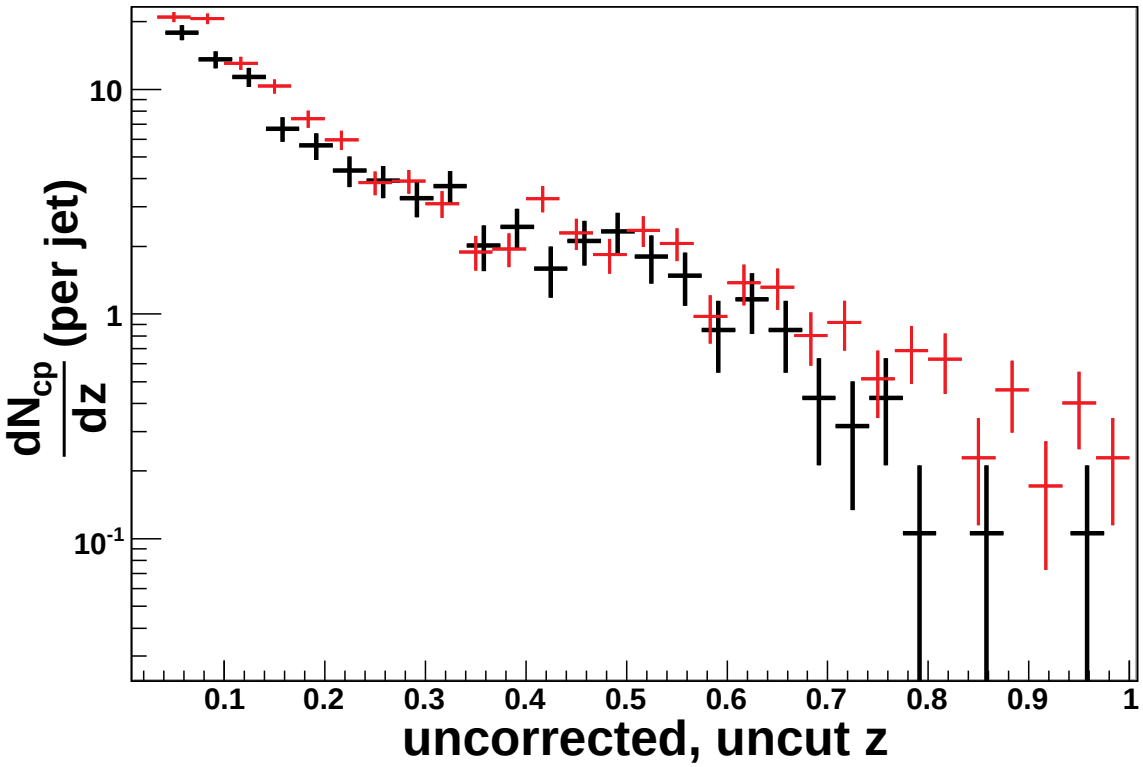

Figure 29. Raw "high-tower" c.p. *z* from p+p and d+Au.
c.p. *z* distributions are for jets which pass the "high-tower" condition, R=0.5, charged particles only, 15-20 GeV. Black is p+p, red is d+Au. No cut is made to remove high towers triggered by charged particles.

The "high-tower" triggered data sample contains a useful sample of jets between 15 and 20 GeV. The trigger condition imposes a bias into the fragmentation function due to the particle which caused the trigger. Figure 29 shows how the raw "high-tower" triggered jets contain a high-z bias. In fact, this z distribution bias in the 15-20 GeV "high-tower" triggered jets is due to *charged particles* causing the trigger by leaving over 2.8 GeV of energy in the BEMC. Figure 30 shows the angular difference between each charged particle with z greater than 0.3 and the location of the "high-tower".

As Figure 30 illustrates, by cutting out jets with charged-particle energy greater than 80% of the "high tower" and less than 0.07 radians away from the "high-tower", the "high-tower" trigger bias on the z distribution should be much reduced. A circle of 0.07 radians is just under 2% of the total 0.5 cone radius solid angle. Comparing the charged particle to the wrong "high-tower" is not a large concern, since only 1.07% of jets have more than one hit above 2.8 GeV. A cut of 0.07 radians removes 43% of the jets. Figure 31 shows how the z distribution compares very well with PYTHIA when jets that were triggered by charged particles are removed. Figure 32 shows how the uncorrected "minimum-bias" jet z from p+p and d+Au compares very well with PYTHIA.

To compare the jet distributions from PYTHIA no-GEANT and PYTHIA with-GEANT simulations, jets must be reconstructed at both levels. Using a cone algorithm, which implies a cone radius parameter, is exactly correct for PYTHIA with-GEANT, since those jets need to be completely comparable to data jets. Using a cone radius parameter without GEANT, however, is less desirable, since ideally the "true" or "complete" jet produced initially by the PYTHIA simulation should be reconstructed. PYTHIA no-GEANT jet properties possess cone radius dependence due to higher z particles clustering closer to the jet thrust axis. The differences between the cone radii are shown in Figure 33 and Figure 34. Moving from radius 0.35 to 0.5, and radius 0.5 to 0.7, the angular area is doubled each time.

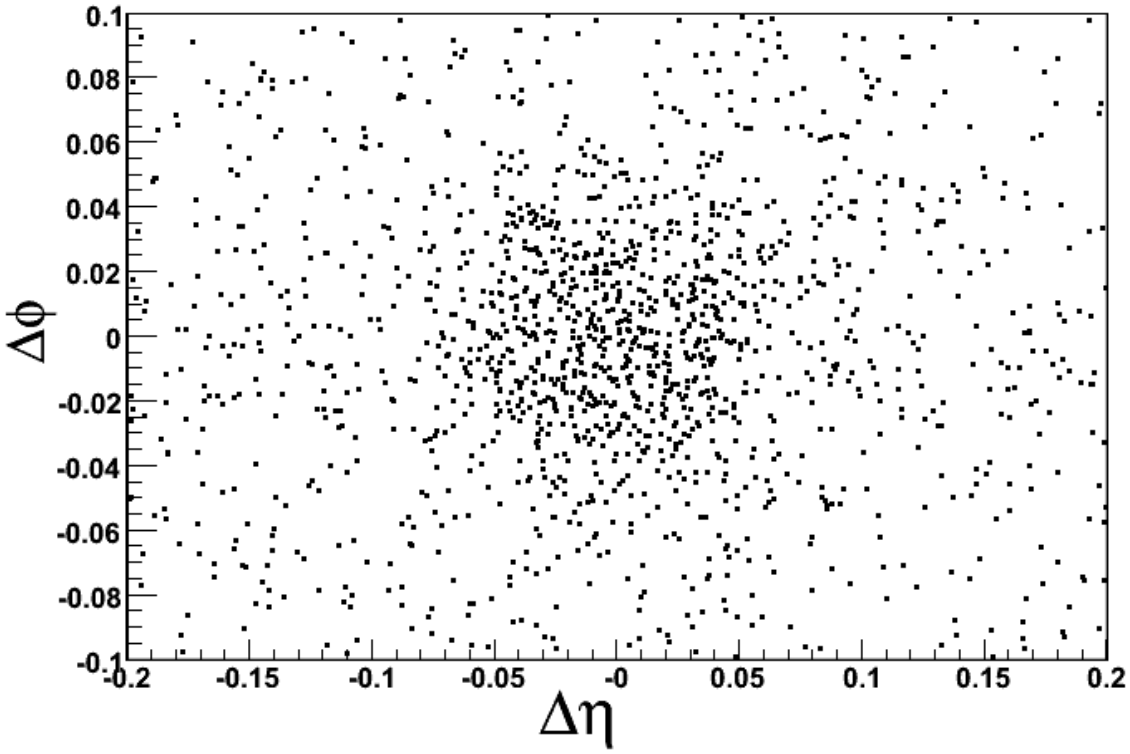

Figure 30. Angle between 15-20 GeV charged particle and "high-tower". All charged particles with z>0.3 are displayed relative to the "high tower" of the jet. Data is from p+p radius 0.5.

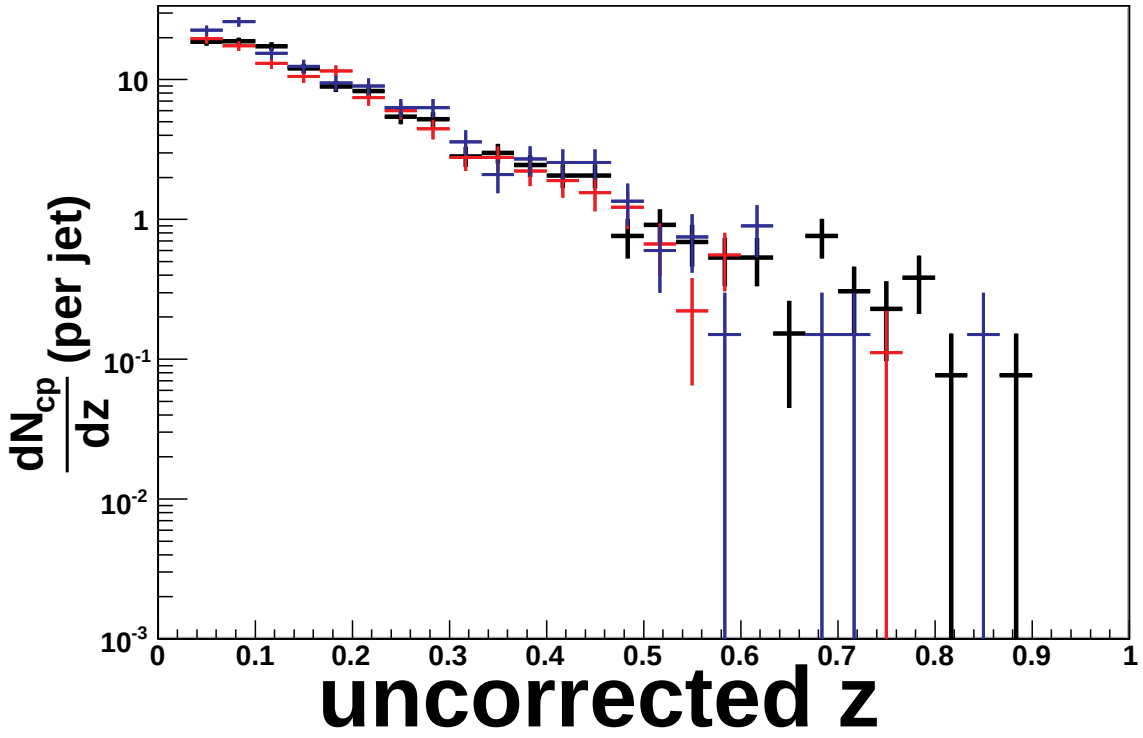

Figure 31. 15-20 GeV c.p. uncorrected z for p+p, d+Au, and PYTHIA. All distributions are for R=0.5, "high-tower"=2.8 GeV. Black is PYTHIA with-GEANT, Red is p+p with charged triggered jet cut > 0.07, and Blue is d+Au with charged triggered jet cut > 0.07.

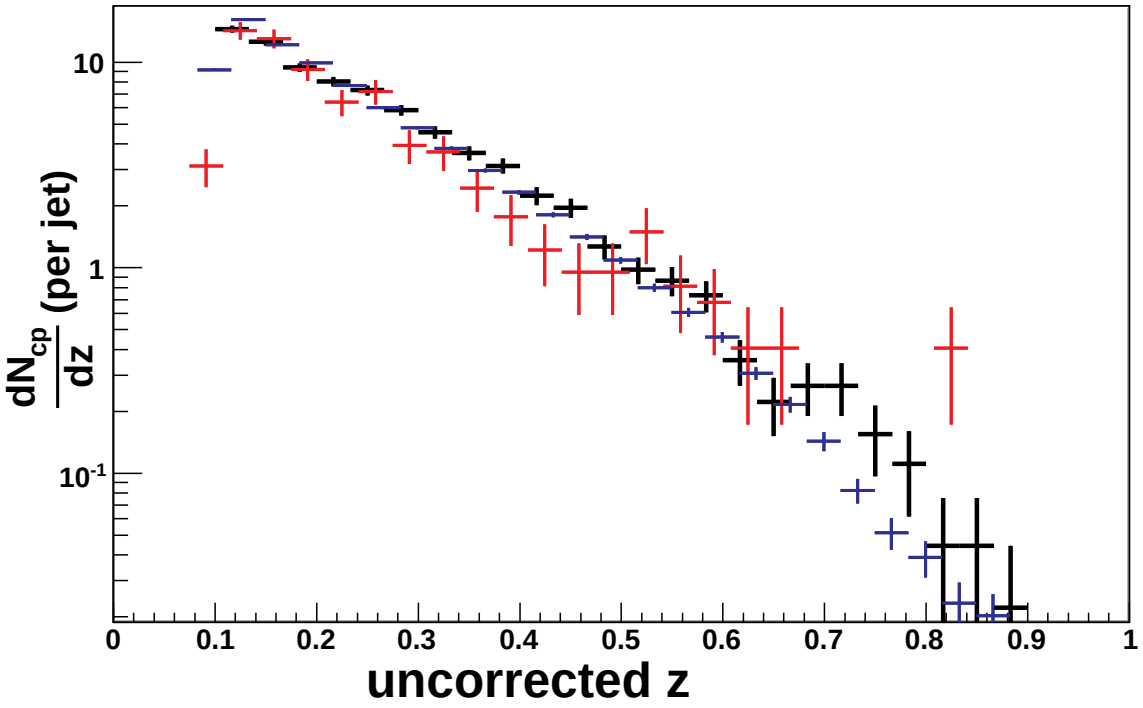

Figure 32. 6-7 GeV c.p. uncorrected z for p+p, d+Au, and PYTHIA.
All distributions are for R=0.5, "minimum-bias" triggered jets. Black is PYTHIA with-GEANT, Red is p+p, and Blue is d+Au.

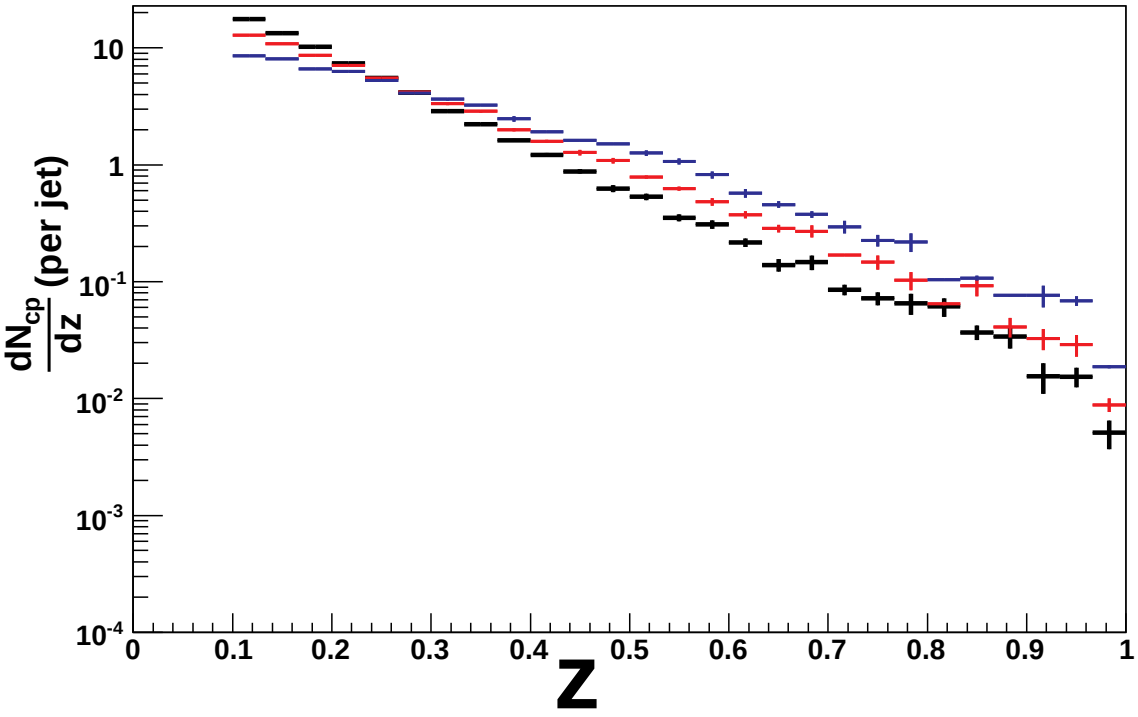

Figure 33. 6-7 GeV jet c.p. z distributions from cone radii 0.35, 0.5 and 0.7. c.p. z distributions are obtained from jets reconstructed from PYTHIA simulation without detector response simulation. Black is cone radius 0.7. Red is cone radius 0.5. Blue is cone radius 0.35.

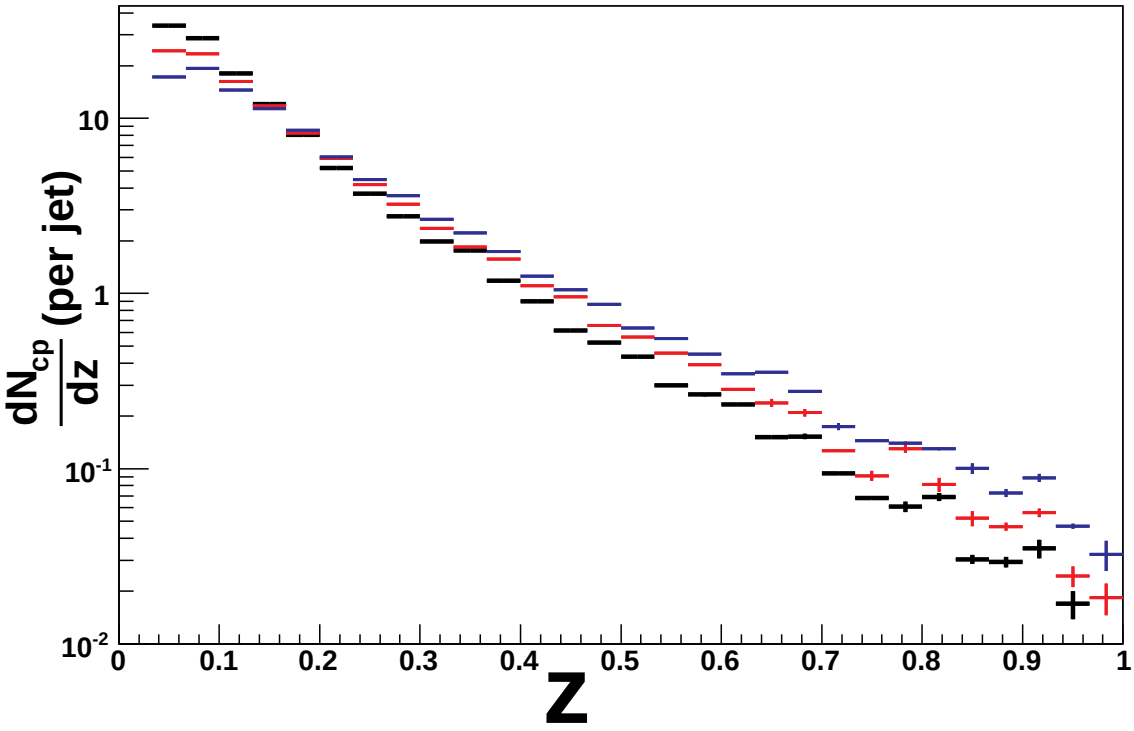

Figure 34. 15-20 GeV jet c.p. z distributions from cone radii 0.35, 0.5 and 0.7. c.p. z distributions are obtained from jets reconstructed from PYTHIA simulation without detector response simulation. Black is cone radius 0.7. Red is cone radius 0.5. Blue is cone radius 0.35.

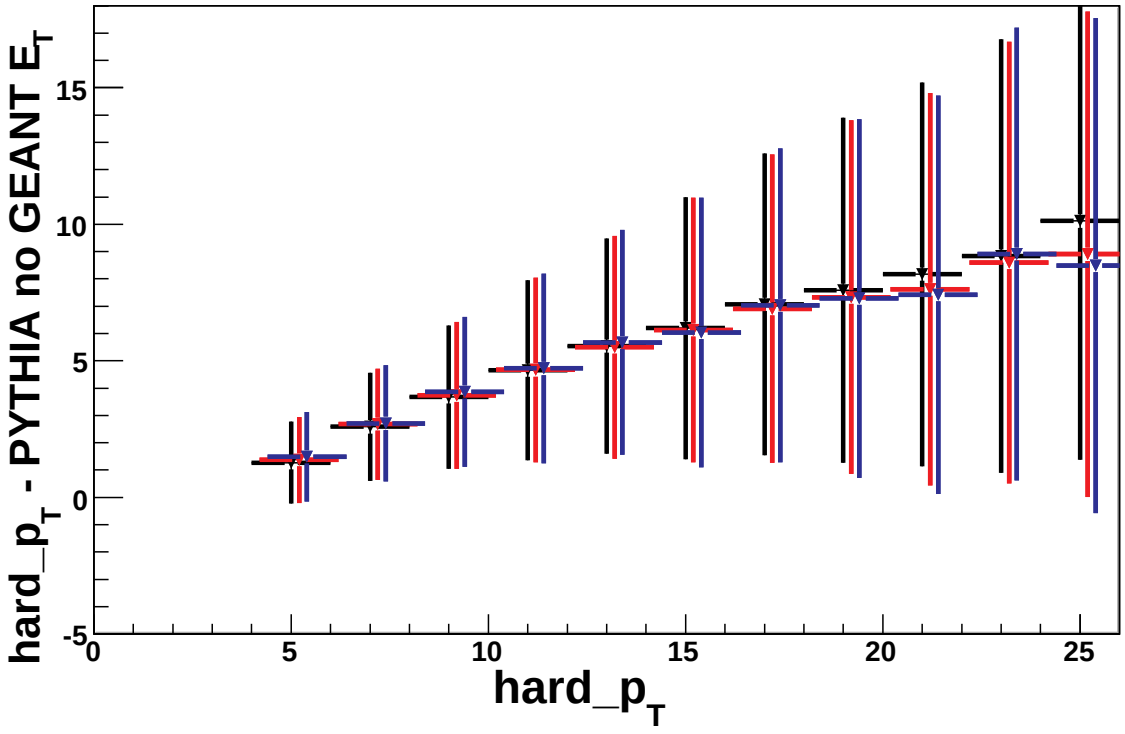

Figure 35. Spread of reconstructed E_T - parton E_T vs parton E_T .
Black is PYTHIA no GEANT radius 0.7, red is radius 0.5, blue is radius 0.35.

Figures 35 and 36 show the spread of the difference distribution between the PYTHIA no-GEANT jets and the hard parton scattering parameter. Figures 35 and 36 are plotted versus, respectively, the hard parton scattering parameter and the reconstructed jet E_T. The relationship between Figures 35 and 36 is similar to the case of the jet energy scale as discussed in section VII.B. The fluctuations in the reconstructed PYTHIA no-GEANT E_T, convoluted with the hard scattering parameter power law distribution, allows the average PYTHIA no-GEANT reconstructed E_T to exceed the hard scattering parameter for a given reconstructed E_T.

Figure 35 shows how PYTHIA simulates jet fragmentation using a string fragmentation model which can significantly reduce the E_T of the jet as compared to the initial hard parton scattering parameter. The hard parton scattering parameter may be analogous to the LO QCD Q^2 parameter, but it is not easily comparable to final state jets after the string "tension" reduces the initial parton momentum. The PYTHIA no GEANT reconstructed jet also includes gluon radiation and interactions with softer high rapidity collision fragments. Figures 35 and 36 also show that the PYTHIA no-GEANT jet energy scale is not strongly dependent on the cone radius. The larger cone radius of 0.7 will bias the z and j_T distributions the least while remaining stable with respect to E_T. Thus, it is reasonable to take the cone radius 0.7 as the most appropriate baseline with which to compare data to PYTHIA no-GEANT results.

The PYTHIA no-GEANT z distributions in these figures seem to have statistics clear out to $z = 1$. Examining these events in detail reveals that PYTHIA really does produce occasional jets with a single track above the minimum track threshold applied in the analysis.

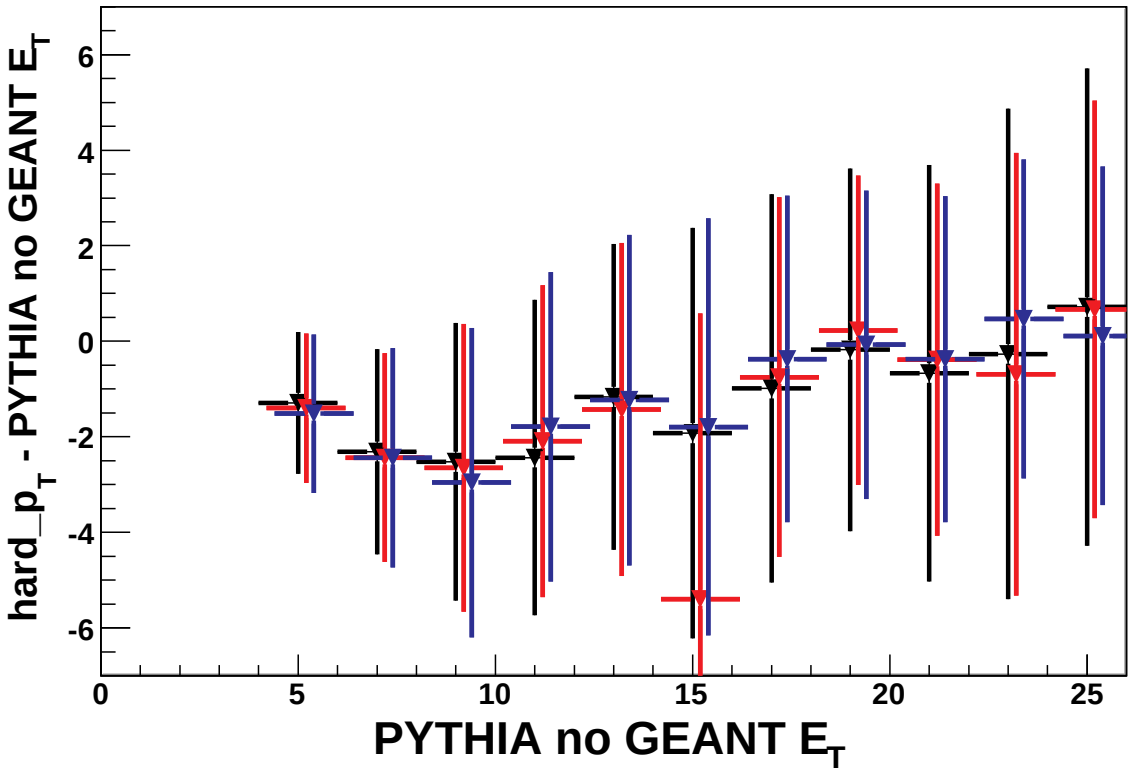

Figure 36. Spread of reconstructed E_T - parton E_T .
Black is PYTHIA no GEANT radius 0.7, red is radius 0.5, blue is radius 0.35.

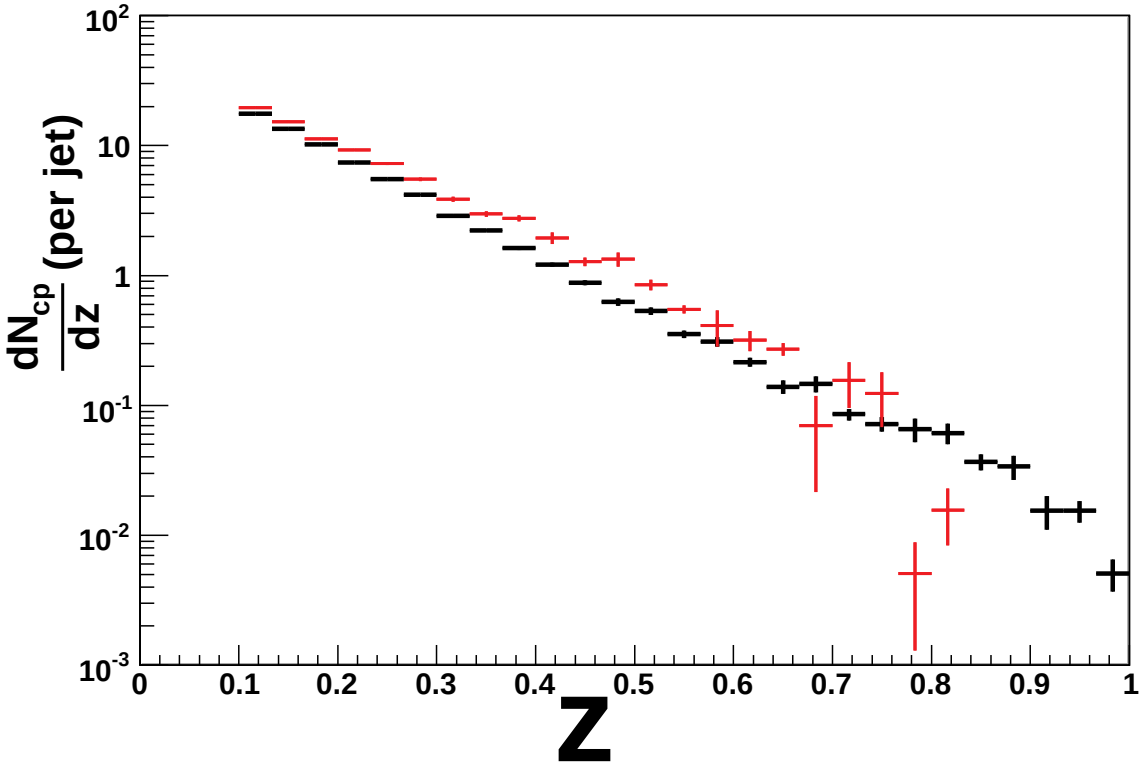

Figure 37. 6-7 GeV Jet c.p. z distribution with and without GEANT. The PYTHIA with detector response simulation is shown in red, and PYTHIA only in black. Cone radius is 0.7 for both.

Figures 37 and 38 show the comparison between the PYTHIA no GEANT jet z and the PYTHIA with GEANT jet z. The difference between the distributions comes more from the variable jet energy reconstructed by the detector, and less by the uncertainty in the individual particle momentum. It is therefore expected that the "shapes" of the two curves will not differ, but an average energy scale difference would change all z measurements. Due to the finite efficiency of the detector, about 15% of charged particles are lost, and an overall normalization difference is also expected.

There are two formulas used for correcting the z distribution: one is for the charged-particle z:

$$z' = \frac{p_{cp}}{p_j} * C_Z \qquad (11)$$

where z' is the corrected charged-particle z value, p_{cp} is the momentum of the charged-particle, p_j is the momentum of the jet, and C_z is a constant. The second equation is for the weight, w', of each z value in the histogram:

$$w' = w * C_w \qquad (12)$$

where w is the uncorrected weight (w=1.0 for data), and C_w is a constant. Table 1 shows the values of C_z and C_w used for correcting the z distribution. These values were obtained by iteratively adjusting C_z and C_w until the exponential fits as shown in Table 2 agreed to within the fit uncertainty. The uncertainties in Table 1 are therefore dominated by and calculated from the uncertainties of the PYTHIA with GEANT fits.

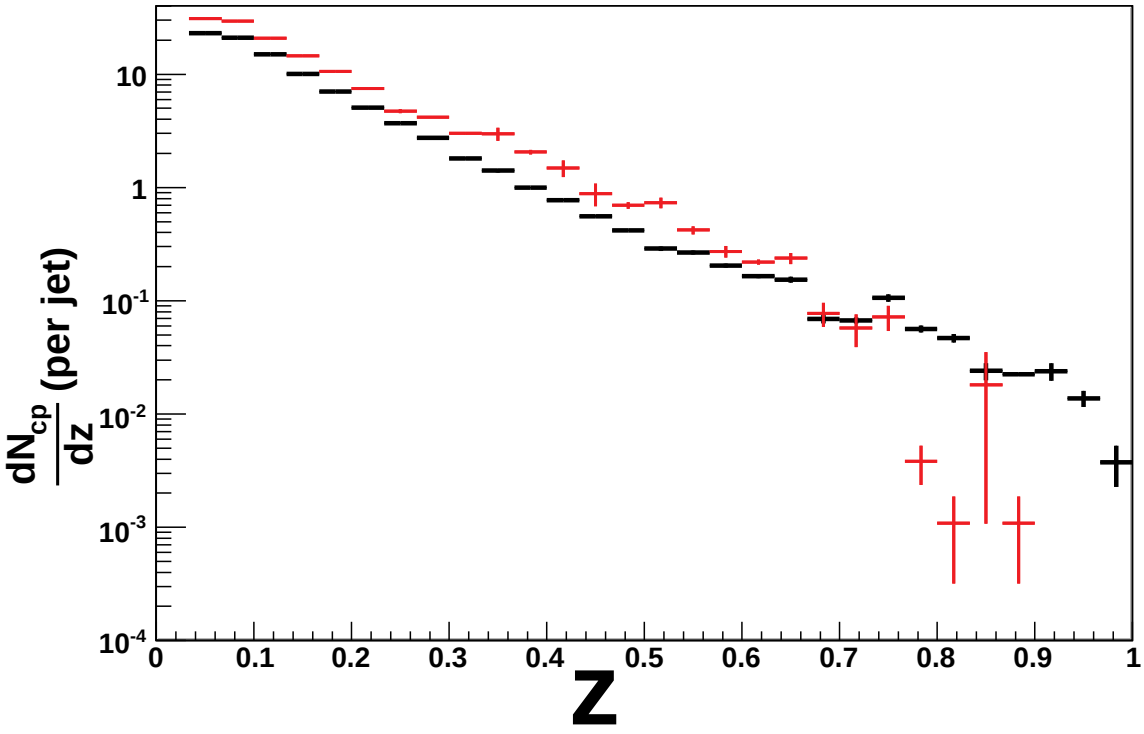

Figure 38. 15-20 GeV Jet c.p. *z* distribution with and without GEANT. The PYTHIA with detector response simulation is shown in red, and PYTHIA only in black. Cone radius is 0.7 for both.

Table 1. Values used in equations 11 and 12 to correct z distributions.

	C_z	C_w
15-20 GeV "high-tower" triggered jets	0.73±0.01	1.03±0.02
6-7 GeV "minimum-bias" triggered jets	0.78±0.02	1.07±0.03

Table 2. Basic exponential fit parameters for the curves in Figures 39, 40, 41, and 42.

	Constant	Slope
R=0.7 PYTHIA without GEANT 6-7 GeV	0.86±0.008	-8.952±0.003
R=0.5 PYTHIA with corrected GEANT 6-7 GeV	0.85±0.04	-9.0±0.17
R=0.5 p+p corrected 6-7 GeV	0.81±0.17	-9.4±0.8
R=0.5 d+Au corrected 6-7 GeV	0.75±0.02	-9.3±0.1
R=0.7 PYTHIA without GEANT 15-20 GeV	0.93±0.02	-10.1±0.01
R=0.5 PYTHIA with corrected GEANT 15-20 GeV	0.90±0.02	-9.9±0.11
R=0.5 p+p corrected 15-20 GeV	1.2±0.2	-8.9±1.7
R=0.5 d+Au corrected 15-20 GeV	0.78±0.13	-10.2±1.0

Figure 39 shows comparisons of PYTHIA no-GEANT at R=0.7 and PYTHIA with-GEANT at R=0.5 for no bias jets. These curves can be fit by a basic exponential over the range from 0.1 to 0.6. Table 2 shows the fit constants for Figures 39 and 40.

Figure 40 shows the same z comparison for jets from PYTHIA no-GEANT at R=0.7 no bias, and PYTHIA with-GEANT at R=0.5 "high-tower" with charged-particle triggered jets cut.

Figure 41 shows the "minimum-bias" z distribution from p+p collisions and d+Au collisions at STAR, corrected for detector resolution using the factors just described. The same correction factor obtained from comparing PYTHIA with and without GEANT is applied to the p+p and d+Au data in Figure 41. In addition, the d+Au energy scale is corrected by 20% due to background multiplicity contributions in the energy interval shown in this plot ($6 < E_T < 7$) as shown in Figure 8. It is compared in the figure to the

"ideal" distribution from PYTHIA no-GEANT. The d+Au curve has also been normalization corrected for background multiplicity effects by subtracting the weighted z distribution that the background multiplicity particles create. The distribution of particles used for this subtraction is shown in Figure 10.

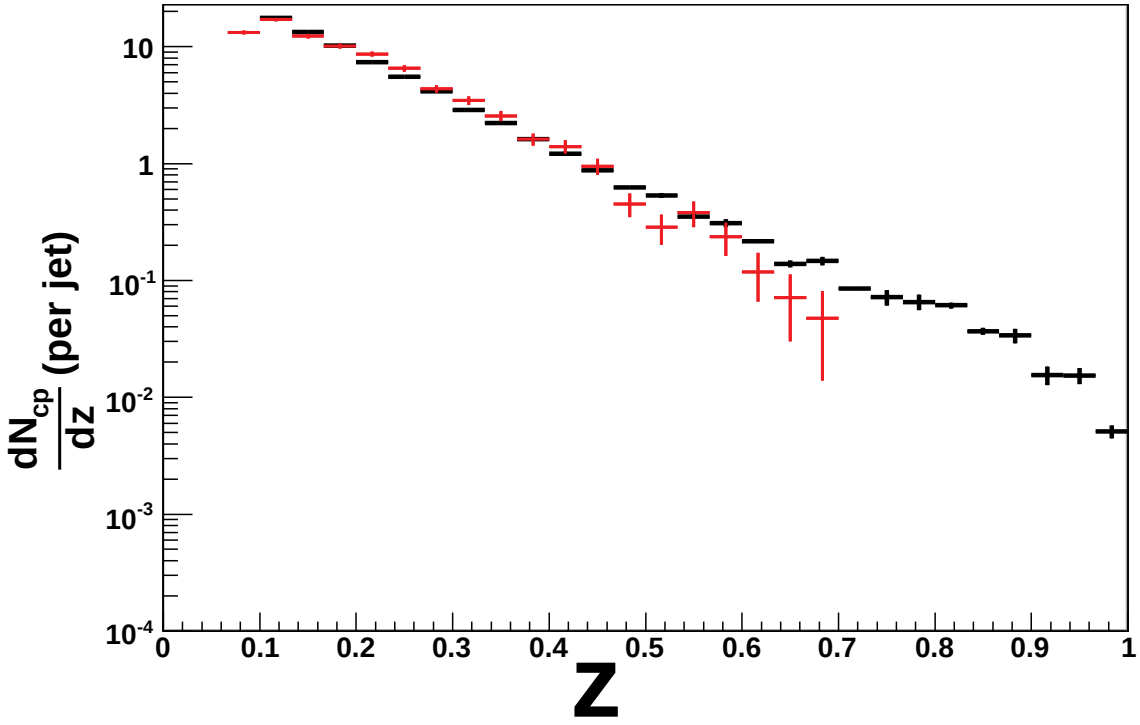

Figure 39. 6-7 GeV jet c.p. z distribution without and corrected with-GEANT. The PYTHIA with detector response simulation, after normalization correction, is shown in red, and PYTHIA only in black. Cone radius is 0.5 for red, and 0.7 for black.

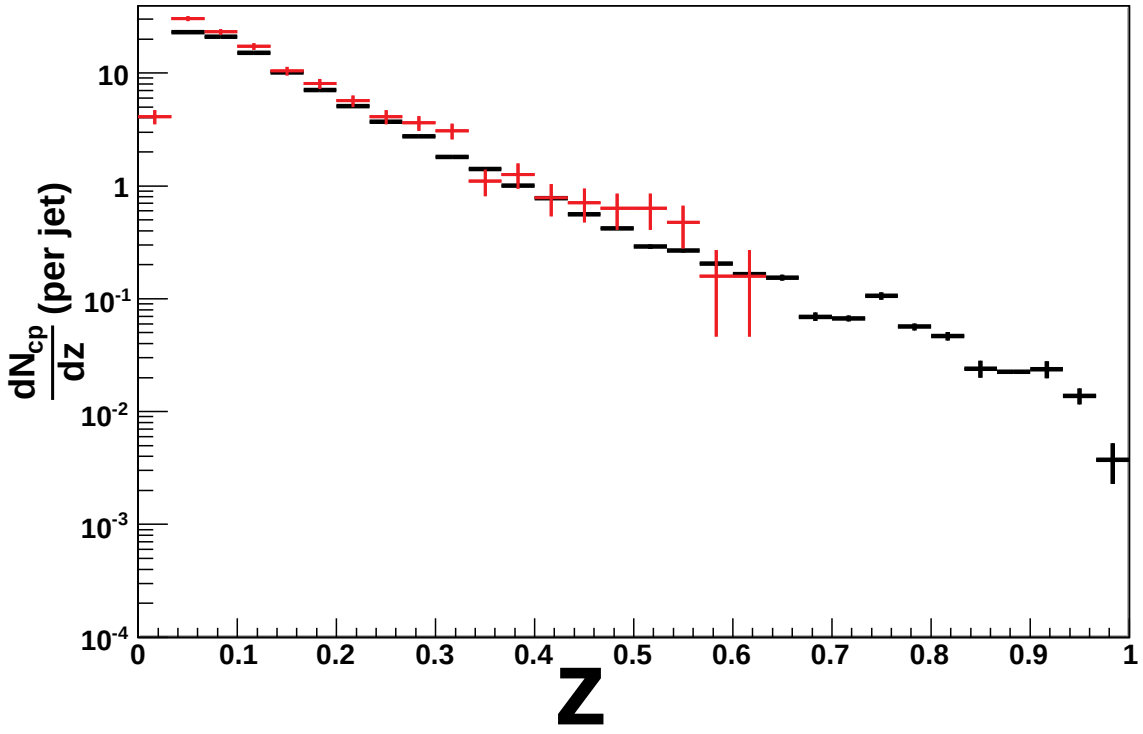

Figure 40. 15-20 GeV jet c.p. z distribution without and corrected with-GEANT. The PYTHIA with detector response simulation and corrected "high-tower" with trigger particle jet cut is shown in red, and PYTHIA only in black. Cone radius is 0.5 for red, and 0.7 for black.

Figure 42 shows the corrected p+p and d+Au z distributions for the "high-tower" jets with trigger particles removed compared with the PYTHIA no-GEANT and no-bias distribution. However, the d+Au energy scale is corrected by 1.5 GeV due to background multiplicity contributions in the energy interval shown in these plots ($15 < E_T < 20$). For the d+Au z distribution, the statistical uncertainty is about 10% for the first bin, and increases at higher z. On the other hand, the correction due to normalization is about 6% for the first bin, and drops off at higher z. Therefore, for this larger energy interval, the normalization correction that would be applied for background multiplicity effects has been neglected.

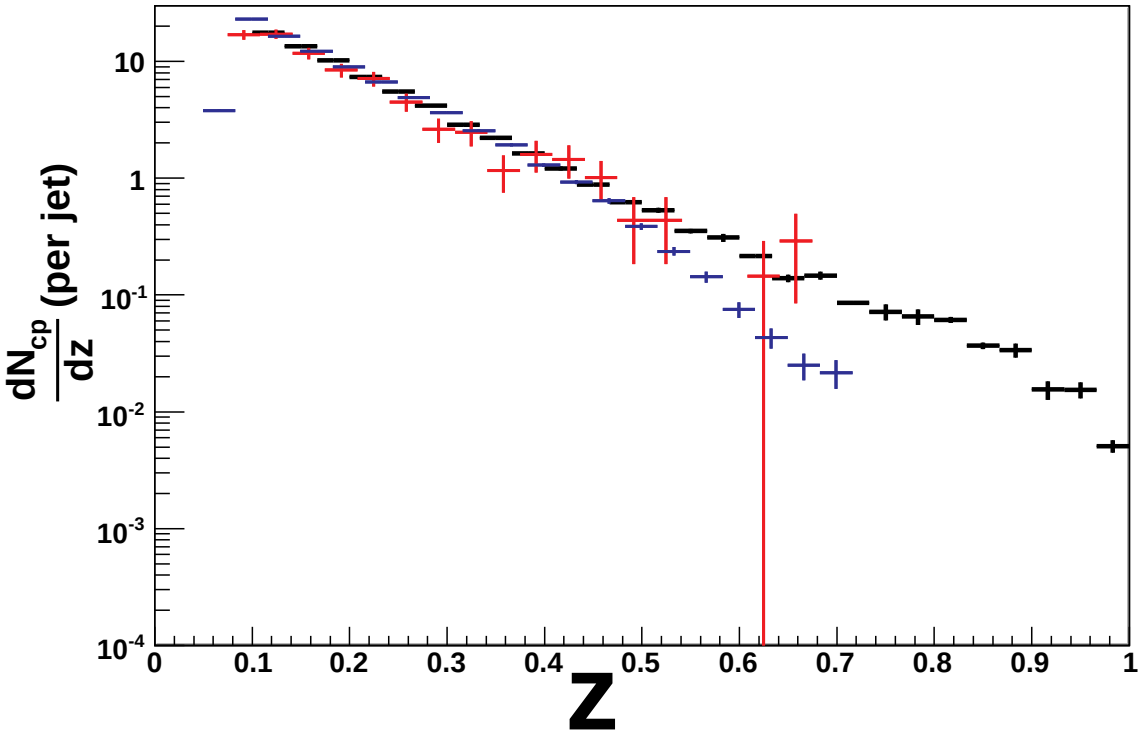

Figure 41. 6-7 GeV jet c.p. z distributions from PYTHIA, p+p and d+Au. PYTHIA no-GEANT c.p. z distribution in black. Red is corrected p+p data. Blue is d+Au data with d+Au uncorrelated background subtracted. Cone radius 0.5 for data, 0.7 for PYTHIA no-GEANT.

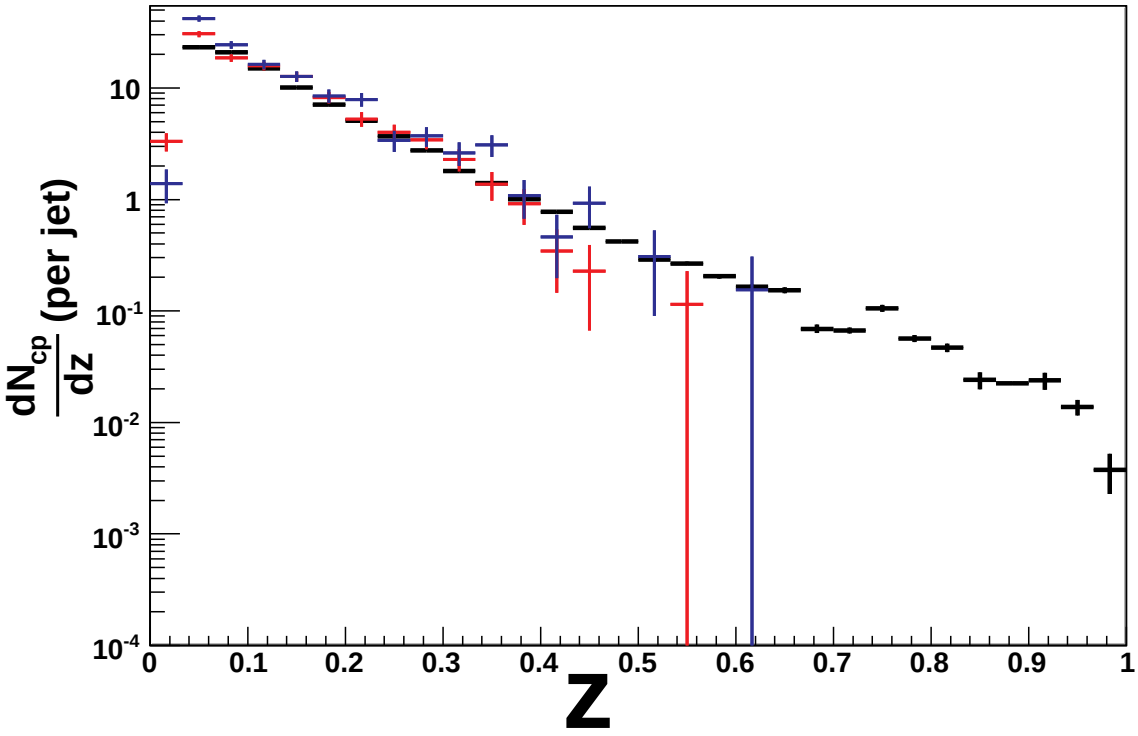

Figure 42. 15-20 GeV jet c.p. z distributions from PYTHIA, p+p and d+Au. PYTHIA no-GEANT c.p. z distribution is black. Red is p+p data. Blue is d+Au data with energy scale correction as described in the text. p+p and d+Au are cone radius 0.5 "high-tower" trigger particle removed, PYTHIA no-GEANT is 0.7 no bias.

2. Jet j_T

Jet j_T is the perpendicular component of the particle momentum with respect to the jet axis. Jet j_T is to the jet thrust axis as the p_T of a measured particle is to the beam axis. In fact, the remnants of the original particles that collide, dubbed beam jets since they continue in the direction of the beam, produce transverse momentum particles in an extremely similar physical process to particles produced by a typical jet. Jet j_T is not strongly dependent on a precise measurement of the energy scale of the jet, since only the direction of the jet is used to calculate j_T. Since the momenta of the particles used to construct the j_T distribution are accurately determined by the detector, uncertainties in the j_T are dominated by uncertainties in the jet axis.

Figures 43 and 44 show the uncorrected j_T from p+p, d+Au, and PYTHIA with-GEANT. Given the similarity between data and PYTHIA, it is possible to correct the j_T distributions for the detector response. The remainder of this section describes this correction.

In fact, the jet algorithm reconstructs the jet axis quite effectively. Since j_T and z are simply the two components of the same jet "shape", all systematic uncertainty issues that arise with z must be addressed with j_T also. For the sake of consistency, identical cuts are used when constructing the j_T and z distributions. Since the cuts are the same, the discussion of j_T is similar to the discussion of jet z in the previous subsection, and the following j_T figures have the same relative context as the previous z figures. Using a finite cone radius parameter prevents particles with sufficiently high-j_T from contributing to jet reconstruction. This means there is a point on the j_T curve where the effect of the cone parameter turns on, and in the following j_T figures, the j_T curve is truncated so that only the region unaffected by the cone parameter is shown.

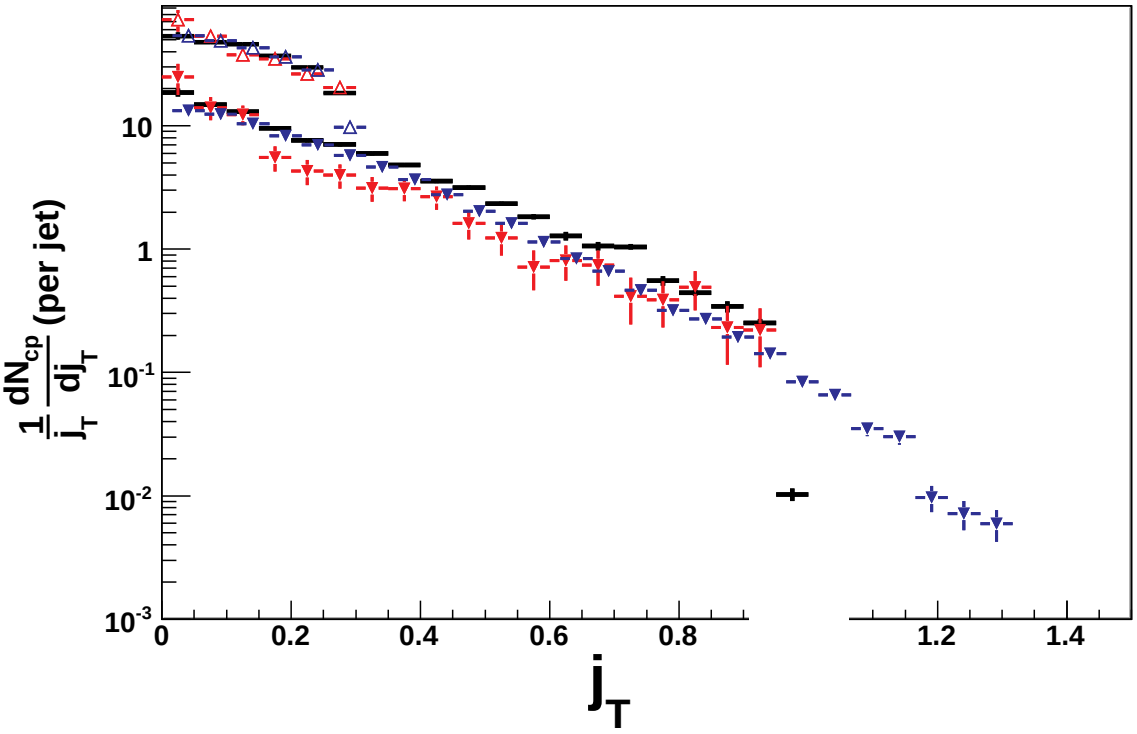

Figure 43. 6-7 GeV uncorrected jet c.p. j_T from data and PYTHIA.
Jets are reconstructed from PYTHIA with detector simulation (black), p+p data uncorrected (red), and d+Au data corrected for background effects only (blue). Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone Radius 0.5 in all cases.

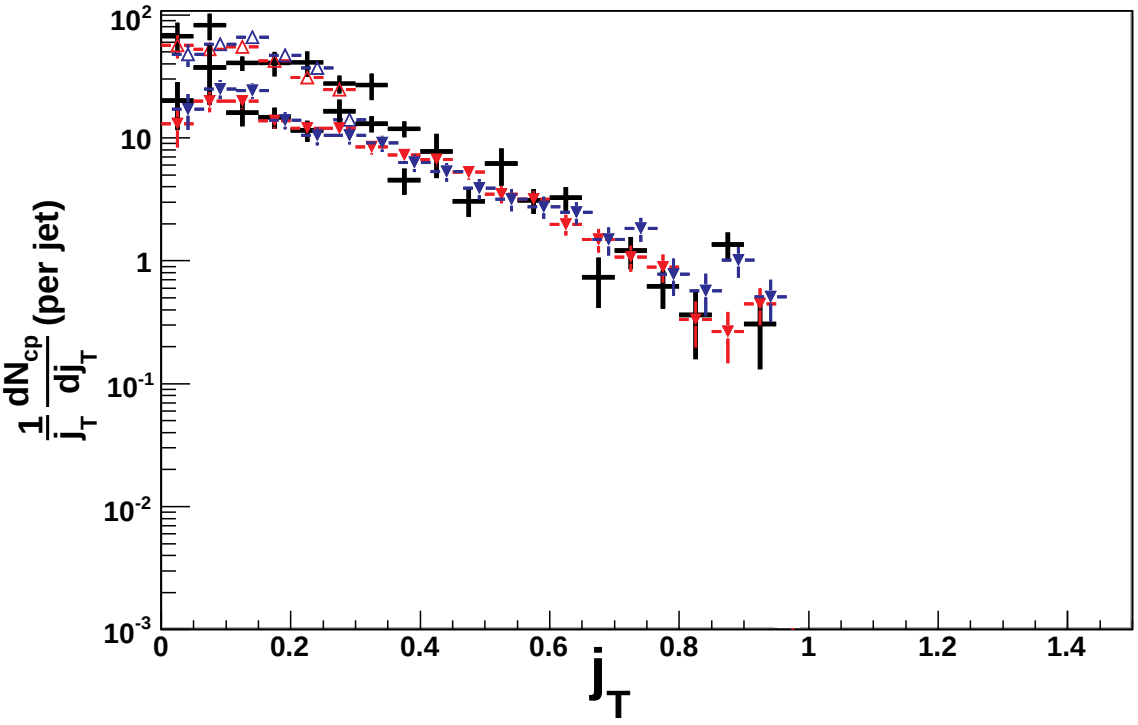

Figure 44. 15-20 GeV uncorrected jet c.p. j_T from data and PYTHIA. Jets are reconstructed from PYTHIA with detector simulation (black), p+p data uncorrected (red), and d+Au data corrected for background effects only (blue). Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius 0.5 in all cases.

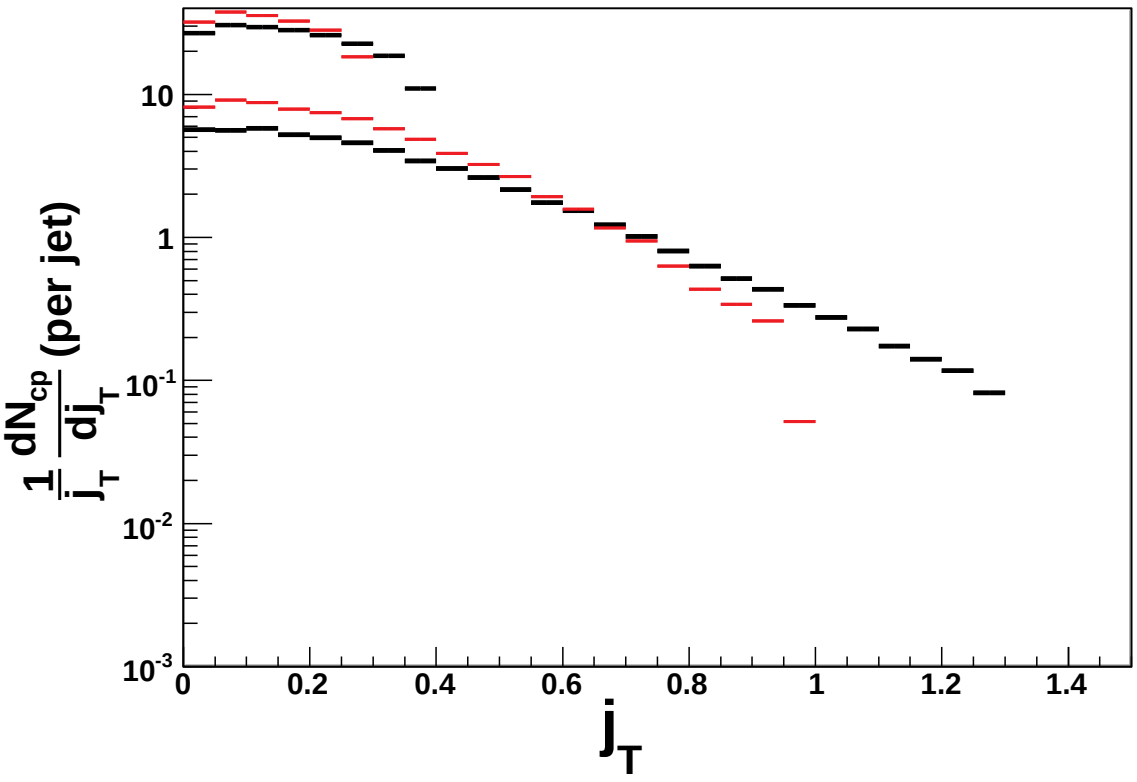

Figure 45. 6-7 GeV jet c.p. j_T from PYTHIA for cone radius 0.5 and 0.7. Jets are reconstructed without GEANT detector simulator. Black is cone radius 0.7, red is cone radius 0.5. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c.

Figure 45 shows the comparison between j_T from PYTHIA no-GEANT for different cone radii for jets at 6 to 7 GeV. Figure 46 shows the same for jets at 15 to 20 GeV. The cone-radius dependence is due to the same energy cut off for both cone radii, meaning that the true energy of the smaller cone radius jets is slightly higher. Since higher energy jets produce more particles at any given threshold than lower energy jets do, the overall multiplicity close to the thrust axis is higher for the same track energy.

Figure 47 shows the comparison between the PYTHIA no-GEANT jet j_T and the PYTHIA with-GEANT jet j_T for jets at 6 to 7 GeV. Figure 48 shows the same for jets at 15 to 20 GeV.

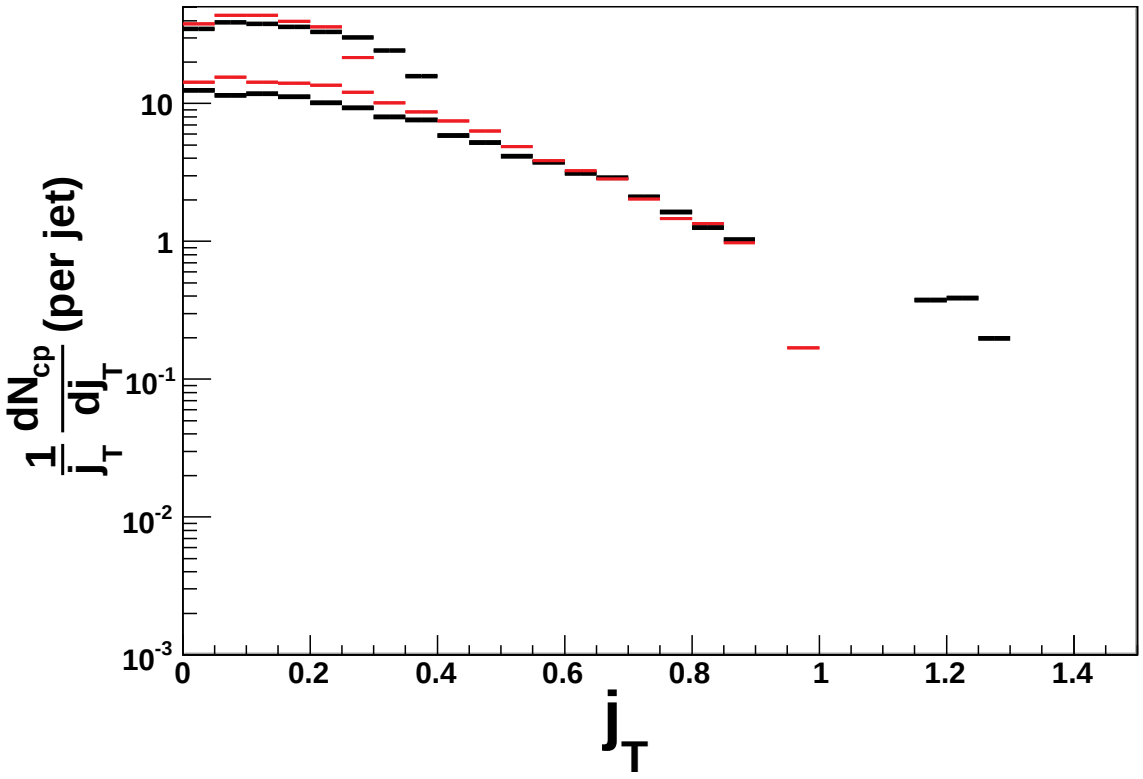

Figure 46. 15-20 GeV jet c.p. j_T from PYTHIA for cone radius 0.5 and 0.7. Jets are reconstructed without GEANT detector simulator. Black is cone radius 0.7, red is cone radius 0.5. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c.

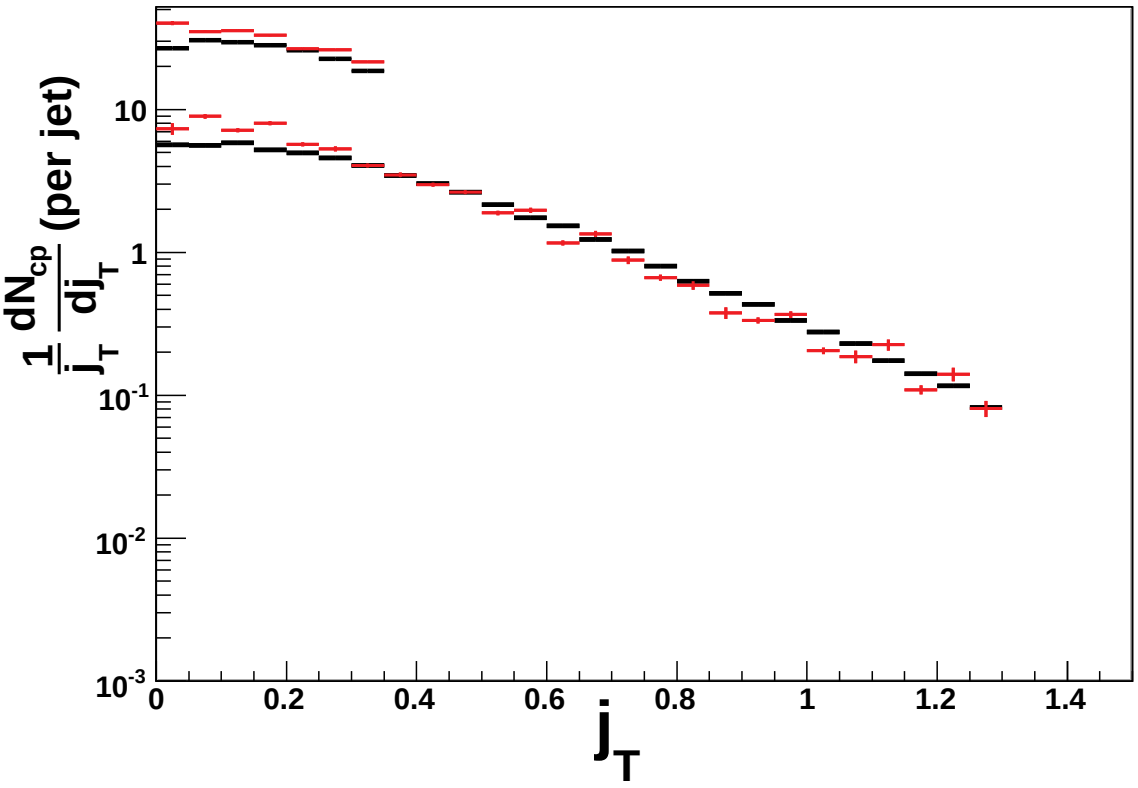

Figure 47. 6-7 GeV PYTHIA jet c.p. j_T with and without GEANT.
Jets from PYTHIA with detector simulation indicated by black, and without detector simulation in red. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.7.

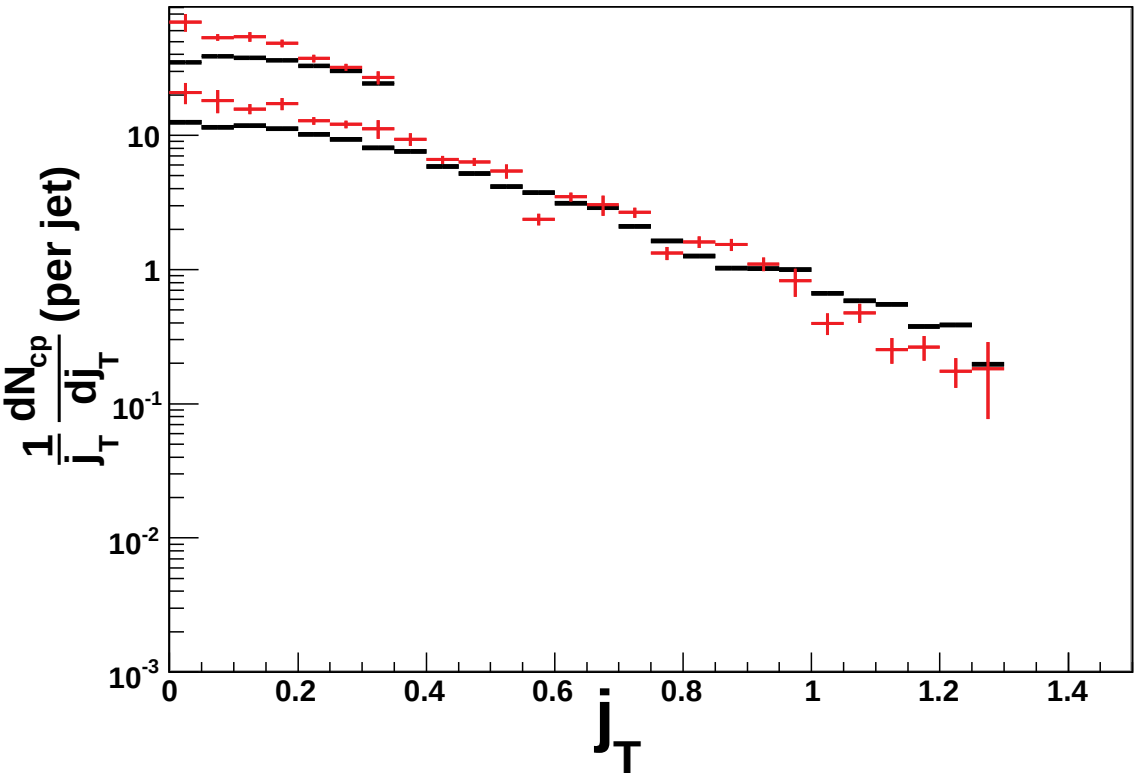

Figure 48. 15-20 GeV PYTHIA jet c.p. j_T with and without GEANT. Jets from PYTHIA with detector simulation indicated by black, and without detector simulation in red. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.7.

There are also two formulas used for correcting the j_T distribution: one is for the charged-particle j_T:

$$j_T' = j_T * C_{jT} \quad (13)$$

where j_T' is the corrected charged-particle j_T value, j_T is uncorrected charged-particle j_T, and C_{jT} is a constant. The smaller cone radius creates a bias toward smaller j_T, and this requires a C_{jT} constant greater than one. The second equation is for the weight, w', of each j_T value in the histogram:

$$w' = w * C_w \quad (14)$$

where w is the uncorrected weight (w=1.0 for data), and C_w is a constant. Table 3 shows the values of C_{jT} and C_w used for correcting the j_T distribution. These values were obtained by iteratively adjusting C_{jT} and C_w until exponential fits on the range 0.2 to 0.5 in j_T for PYTHIA without GEANT and corrected PYTHIA with GEANT curves agreed. The uncertainties in Table 3 are therefore dominated by and calculated from the uncertainties of the PYTHIA with GEANT fits.

Table 3. Values used in equations 13 and 14 to correct j_T distributions.

	C_{jT}	C_w
15-20 GeV "high-tower" triggered jets, c.p.>0.6 GeV/c	1.28±0.08	0.9±0.1
6-7 GeV "minimum-bias" triggered jets, c.p. > 0.6 GeV/c	1.5±0.02	0.9±0.014
15-20 GeV "high-tower" triggered jets, c.p.>2.0 GeV/c	1.28±0.05	0.7±0.09
6-7 GeV "minimum-bias" triggered jets, c.p. > 2.0 GeV/c	1.35±0.02	0.6±0.11
6-7 GeV "minimum-bias" d+Au jets only, c.p. > 2.0	1.35*1.35±0.05	0.6*0.8±0.15

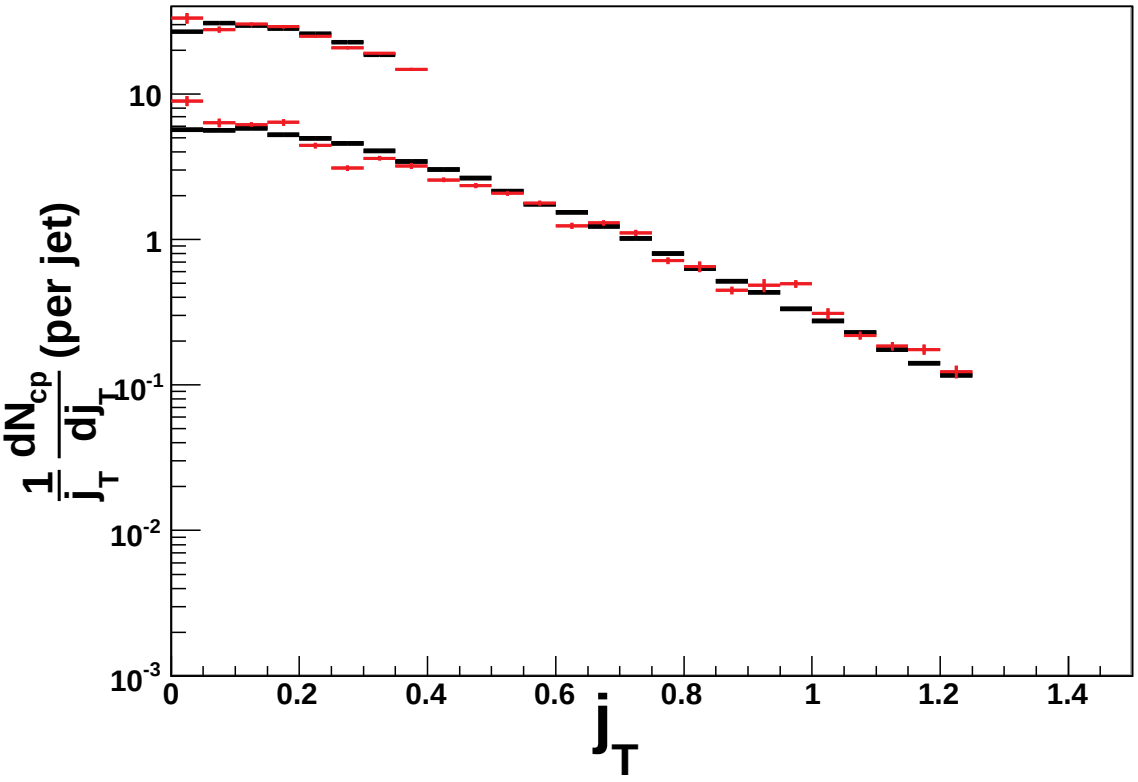

Figure 49. 6-7 GeV PYTHIA jet c.p. j_T corrected with and without GEANT. Black is Pythia no-GEANT cone radius 0.7. Red is Pythia GEANT cone radius 0.5 corrected.

From the comparison between PYTHIA with and without GEANT at radius 0.5 and 0.7, the constants in Table 3 can be obtained to correct for cone radius and detector resolution effects. For 6 to 7 GeV, the effect of the correction on the PYTHIA with-GEANT at radius 0.5, is shown in Figure 49 compared with the PYTHIA no-GEANT j_T radius 0.7 curves. Figure 50 shows this comparison for the 15 to 20 GeV "high-tower" triggered jets.

The d+Au curve for tracks above 2.0 GeV/c is also corrected by additional factors obtained by comparing the background supplemented p+p with the plain p+p. Table 3 shows the modified factor for 6-7 GeV d+Au jet j_T for the 2.0 GeV/c track threshold which takes into account the shifted jet thrust axis due to d+Au multiplicity background. Figure 51 shows the p+p and p+p+supp curves used to estimate the correction factors in Table 3. The d+Au curve extends past the p+p curve because it has been modified by these factors. The same comparison for the 0.6 GeV/c track threshold shows that for lower p_T tracks, j_T is less sensitive to the thrust axis, and no extra factor is required. Figure 52 compares the j_T distribution from p+p and d+Au collisions at STAR, corrected by the factors obtained from the PYTHIA comparisons, displayed with the PYTHIA no-GEANT results for 6 to 7 GeV. It appears that the PYTHIA simulation models the data well, except perhaps at very high j_T, and that after accounting for combinatorial d+Au background, the agreement between p+p and d+Au is also very good.

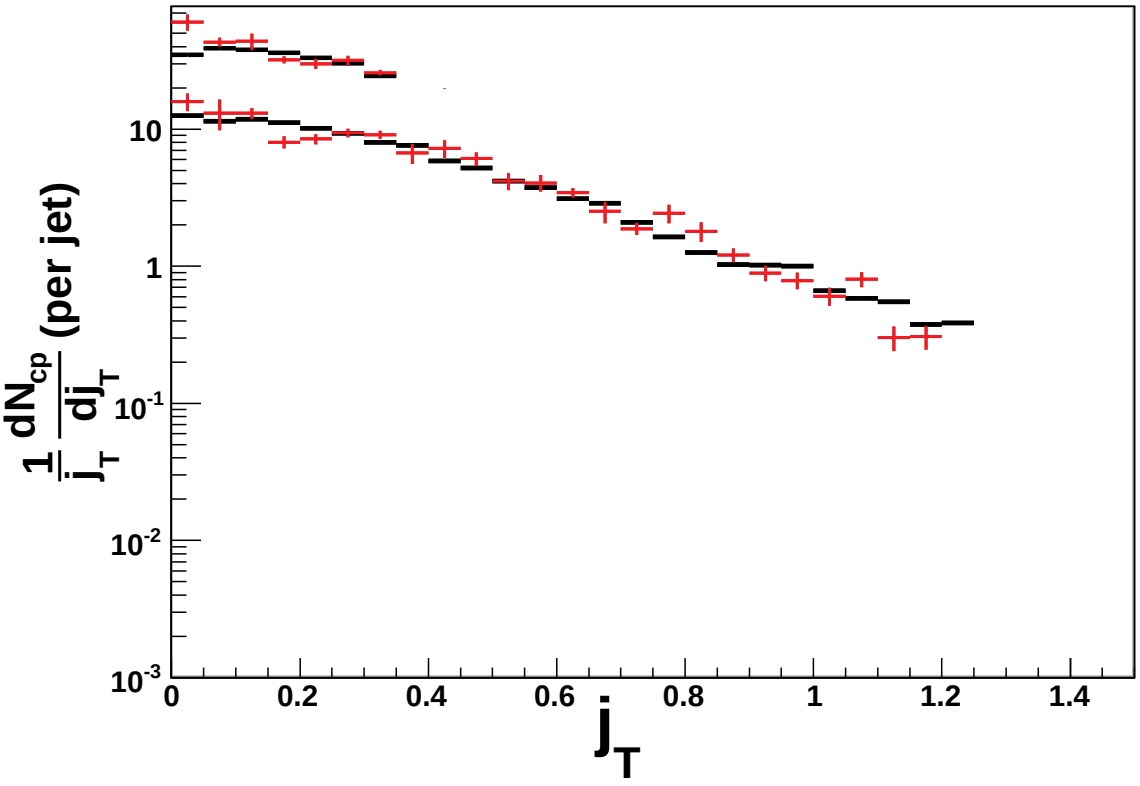

Figure 50. 15-20 GeV PYTHIA jet c.p. j_T corrected with and without GEANT. Black is Pythia no-GEANT cone radius 0.7. Red is Pythia GEANT cone radius 0.5 corrected.

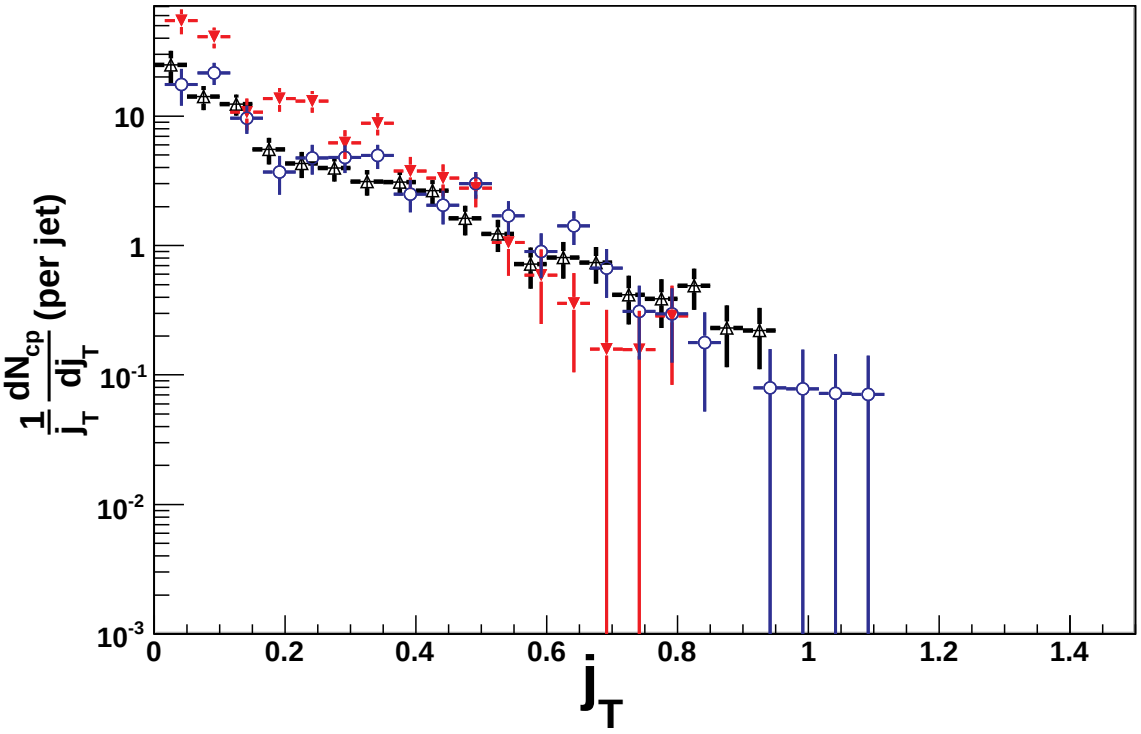

Figure 51. 6-7 GeV p+p and p+p+supp jet c.p. j_T .
Black is p+p cone radius 0.5. Red is p+p+supp cone radius 0.5 not corrected.
Blue is p+p+supp cone radius 0.5 with corrections from shifted jet thrust axis.

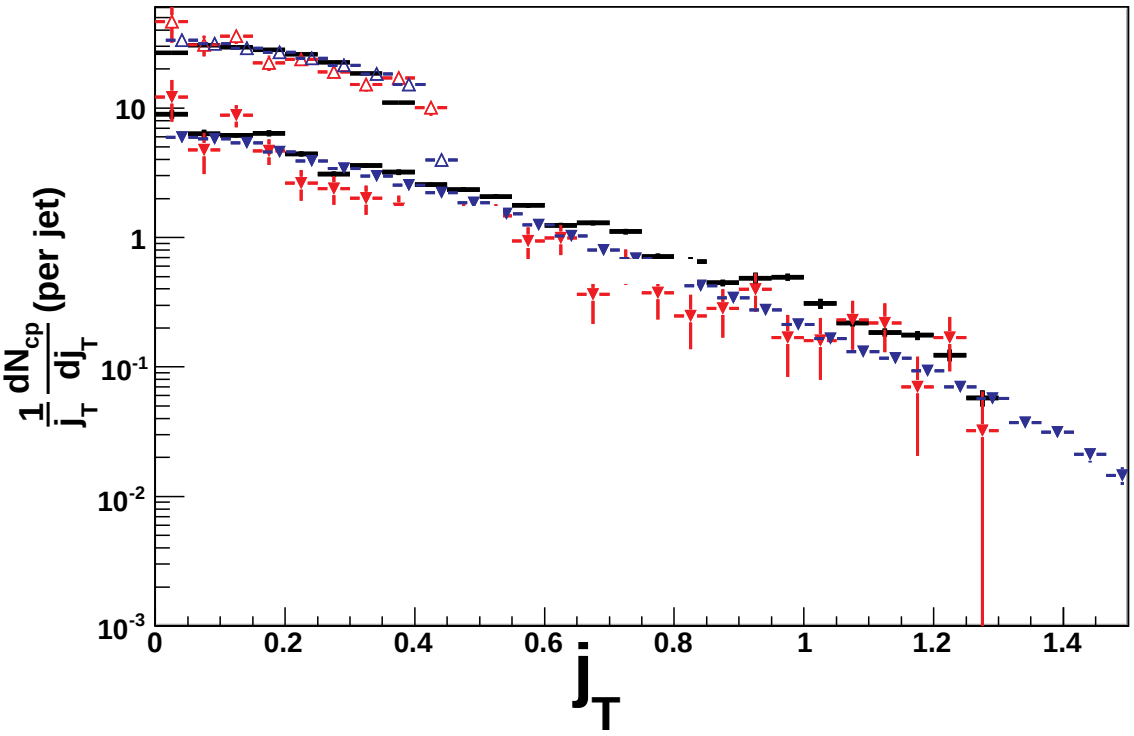

Figure 52. 6-7 GeV jet c.p. j_T from PYTHIA, p+p and d+Au data.
Jets are reconstructed from PYTHIA without detector simulation (black), p+p data corrected for detector and radius effects (red), and d+Au data corrected for detector, radius, and background effects (blue). Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5 for data, and 0.7 for PYTHIA.

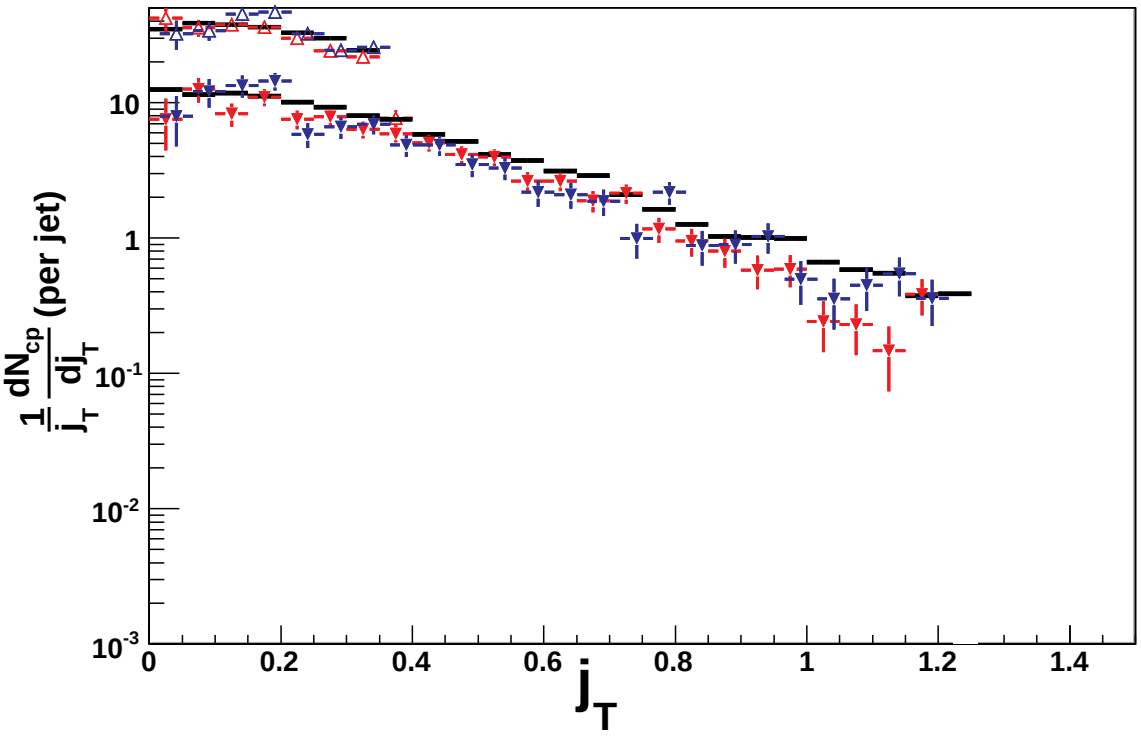

Figure 53. 15-20 GeV jet c.p. j_T from PYTHIA, p+p and d+Au data.
Jets are reconstructed from PYTHIA without detector simulation (black), p+p data corrected for detector and radius effects (red), and d+Au data corrected for detector, radius, and background effects (blue). Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5 for data, and 0.7 for PYTHIA.

Figure 53 shows the j_T distribution from p+p and d+Au collisions at STAR, corrected for by the two factors obtained from the PYTHIA comparisons, displayed with the PYTHIA no-GEANT results for 15 to 20 GeV. The same d+Au background multiplicity corrections are applied to the 0.6 GeV/c charged particles at the 15-20 GeV jet energy, even though the effect of the background multiplicity is somewhat reduced compared to the lower jet energies. The same correction for 2.0 GeV/c is neglected, since for each bin shown in the histogram, the correction would be on the order of 3.0×10^{-3}. Here also, the agreement between data and PYTHIA is quite good.

Jets are removed from the sample when a high-p_T charged particle is in coincidence with the high tower. These jets must be removed since a "high-tower" triggered on a charged particle has an abnormal energy response. For j_T in particular, charged particles which trigger the "high tower" have an artificially high weight on the jet thrust axis direction, since this triggering charged particle has some of its energy counted twice. The "high-tower"-triggered charged particles pull the jet thrust axis away from the other particles in the jet increasing their j_T, while simultaneously lowering the j_T of the triggering charged particle.

It is reasonable to assume that if a simple transformation can map PYTHIA with-GEANT at radius 0.5 onto PYTHIA no-GEANT at radius 0.7, the same transformation can be used to map the measurements onto the real j_T curve. Mathematically, what is done here is to apply a scale factor to the values along each axis. The shape of j_T is not modified. It is often assumed that $\frac{1}{j_T}\frac{dN}{dj_T}$ is Gaussian. In fact the j_T curves can be fitted with Gaussians, with one proviso. The 6-7 jet GeV c.p. > 2.0 GeV/c j_T distributions are not consistent with Gaussians centered at zero, but are consistent with Gaussians with negative mean. As Table 4 shows, both the p+p and PYTHIA distributions are consistent with d+Au in this range. All other curves are consistent with a Gaussian centered at zero.

Table 4. Gaussian fit parameters for the curves in Figures 54, 55, 56, 57, 58, and 59. Parentheses indicate fixed parameters. PYTHIA with GEANT fits are made to corrected distributions. All fits are radius 0.5. Gaussian fits with negative means are grayed.

6-7 GeV	Constant	Sigma	Mean
PYTHIA Gaussian, c.p. > 2.0 GeV/c	0.26±0.005	0.39±0.0	(0.0)
p+p Gaussian, c.p. > 2.0 GeV/c	0.15±0.02	0.40±0.02	(0.0)
d+Au Gaussian, c.p. > 2.0 GeV/c	0.11±0.01	0.41±0.01	(0.0)
PYTHIA Gaussian with offset, c.p. > 2.0 GeV/c	0.9±0.19	0.65±0.03	-0.85±0.1
p+p Gaussian with offset, c.p. > 2.0 GeV/c	(0.9)	0.66±0.05	-1.02±0.1
d+Au Gaussian with offset, c.p. > 2.0 GeV/c	1.3±0.2	0.67±0.02	-1.05±0.02
PYTHIA Gaussian, c.p. > 0.6 GeV/c	1.6±0.02	(0.32)	(0.0)
p+p Gaussian, c.p. > 0.6 GeV/c	1.4±0.1	(0.32)	(0.0)
d+Au Gaussian, c.p. > 0.6 GeV/c	1.62±0.01	0.32±0.01	(0.0)
15-20 GeV	Constant	Sigma	Mean
PYTHIA Gaussian, c.p. > 2.0 GeV/c	0.44±0.04	0.38±0.02	(0.0)
p+p Gaussian, c.p. > 2.0 GeV/c	0.47±0.03	0.39±0.01	(0.0)
d+Au Gaussian, c.p. > 2.0 GeV/c	0.42±0.03	0.41±0.01	(0.0)
PYTHIA Gaussian with offset, c.p. > 2.0 GeV/c	0.72±0.5	0.46±0.1	-0.28±0.4
p+p Gaussian with offset, c.p. > 2.0 GeV/c	0.53±0.1	0.41±0.03	-0.08±0.1
d+Au Gaussian with offset, c.p. > 2.0 GeV/c	1.5±1.2	0.61±0.1	-0.76±0.5
PYTHIA Gaussian, c.p. > 0.6 GeV/c	1.6±0.1	(0.41)	(0.0)
p+p Gaussian, c.p. > 0.6 GeV/c	1.75±0.1	(0.41)	(0.0)
D+Au Gaussian, c.p. > 0.6 GeV/c	1.9±0.1	(0.41)	(0.0)

Figures 54 and 55 show Gaussian fits to the PYTHIA with GEANT j_T curves. Figures 56 and 57 show Gaussian fits to the p+p j_T curves. Figures 58 and 59 show Gaussian fits to the d+Au j_T curves. Table 4 shows the fit values used for all these figures. The upper 0.6 GeV/c charged-particle j_T curves in Figures 54 and 56 are fitted with the same Gaussian shape as the upper 0.6 GeV/c charged-particle j_T curve in Figure 58, but with fitted constant term. The upper 0.6 GeV/c charged-particle j_T curves in Figures 55, 57, and 59 are fitted with the same Gaussian shape as for the d+Au 2.0 GeV/c charged-particle curve in Figure 59, but with fitted constant term since the width cannot be extracted from the 15-20 GeV jet 0.6 GeV/c charged-particle curves. While the Gaussian centered at zero fits well with the 0.6 GeV/c charged-particle j_T curves, these curves

cannot safely be extrapolated to higher j_T. The narrower width is due to a reduction of the phase space available to low p_T charged particles, often referred to as the seagull effect.

Mapping one Gaussian distribution to another can always be done by scaling the width and normalization constant, and this is what the corrections in Table 3 do. Since these distributions can be fitted with Gaussians, it makes sense that the corrections work so well.

It is easy to extract $\sqrt{\langle j_T^2 \rangle}$ from PYTHIA no-GEANT, p+p, and d+Au. For the 6-7 GeV range, $\sqrt{\langle j_T^2 \rangle}$ is $630 \pm 12 \pm 30$ MeV/c for d+Au, and $690 \pm 12 \pm 60$ MeV/c for p+p. For the 15-20 GeV range, $\sqrt{\langle j_T^2 \rangle}$ is $630 \pm 13 \pm 30$ MeV/c for d+Au, and $573 \pm 9 \pm 30$ MeV/c for p+p. These averages are dependent on the scale factors of Table 3. The systematic uncertainties of these $\sqrt{\langle j_T^2 \rangle}$ values come from uncertainties in the parameters of Table 3. In all cases, the data do not significantly differ from the PYTHIA no-GEANT-simulated $\sqrt{\langle j_T^2 \rangle}$ value of 627 MeV/c. The PYTHIA value is statistically very precise, and does not increase with jet E_T between 6 and 20 GeV. These values show that for p+p, the $\sqrt{\langle j_T^2 \rangle}$ differs by less than two sigma between 6 to 7 and 15 to 20 GeV jets, and for d+Au jets, there are no differences between energy ranges. Assuming that the differences between the two p+p measurements are due to the sample size, it seems plausible to take the average of the two p+p measurements weighted by systematic uncertainty: $612 \pm 9 \pm 30$ MeV/c.

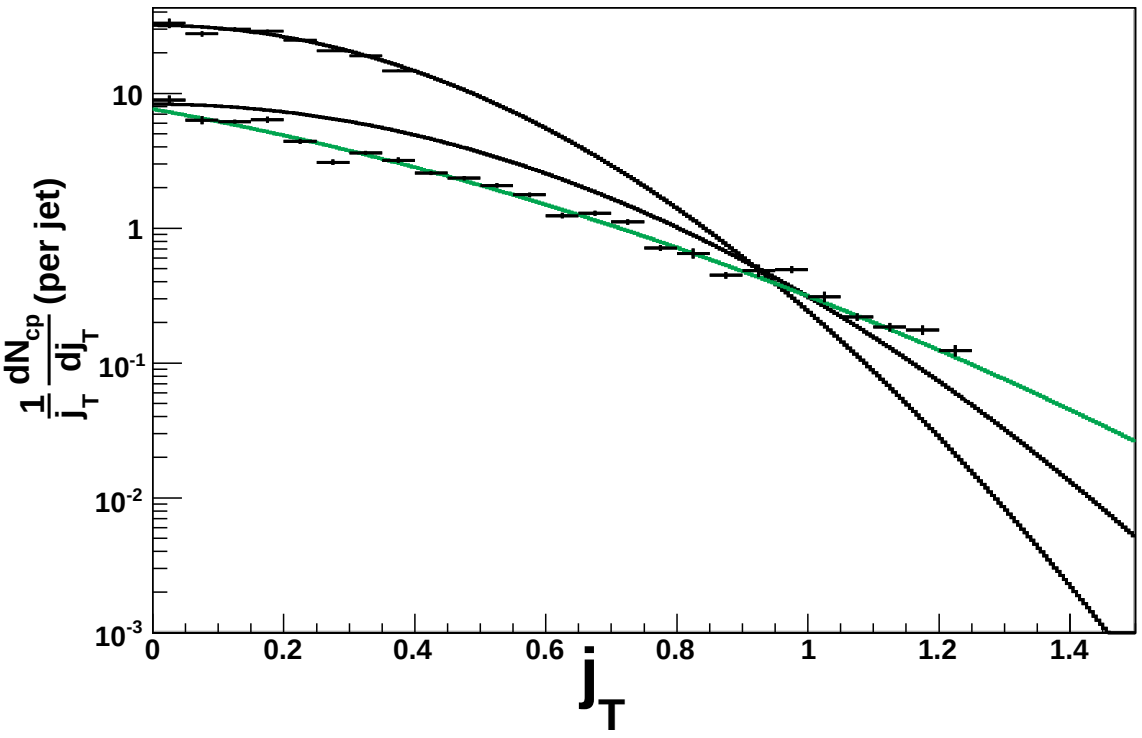

Figure 54. 6-7 GeV jet c.p. j_T from PYTHIA with GEANT with Gaussian fit. Black lines are Gaussians centered at zero, Green line is Gaussian with negative mean. PYTHIA with GEANT data corrected for detector and radius effects are the black data points. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5.

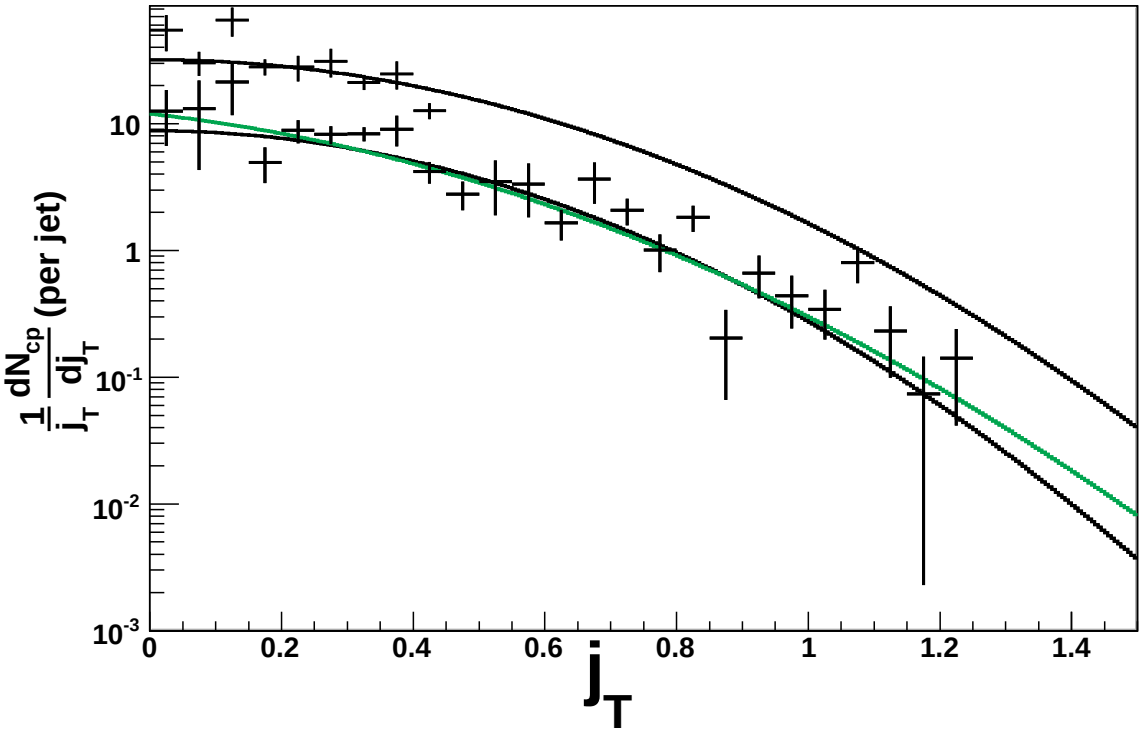

Figure 55. 15-20 GeV jet c.p. j_T from PYTHIA with GEANT with Gaussian fit. Black lines are Gaussians centered at zero, Green line is Gaussian with negative mean. PYTHIA with GEANT data corrected for detector and radius effects are the black data points. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5.

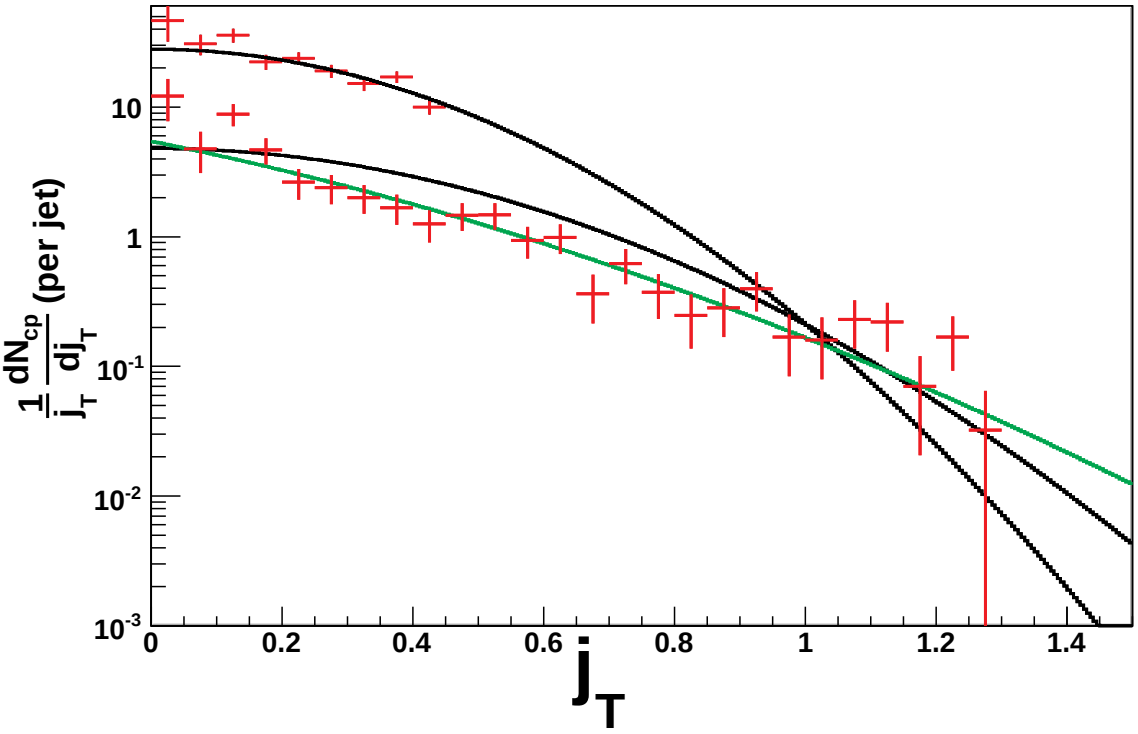

Figure 56. 6-7 GeV jet c.p. j_T from p+p with Gaussian fit.
Black lines are Gaussians centered at zero, Green line is Gaussian with negative mean. p+p data corrected for detector and radius effects are the red data points. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5.

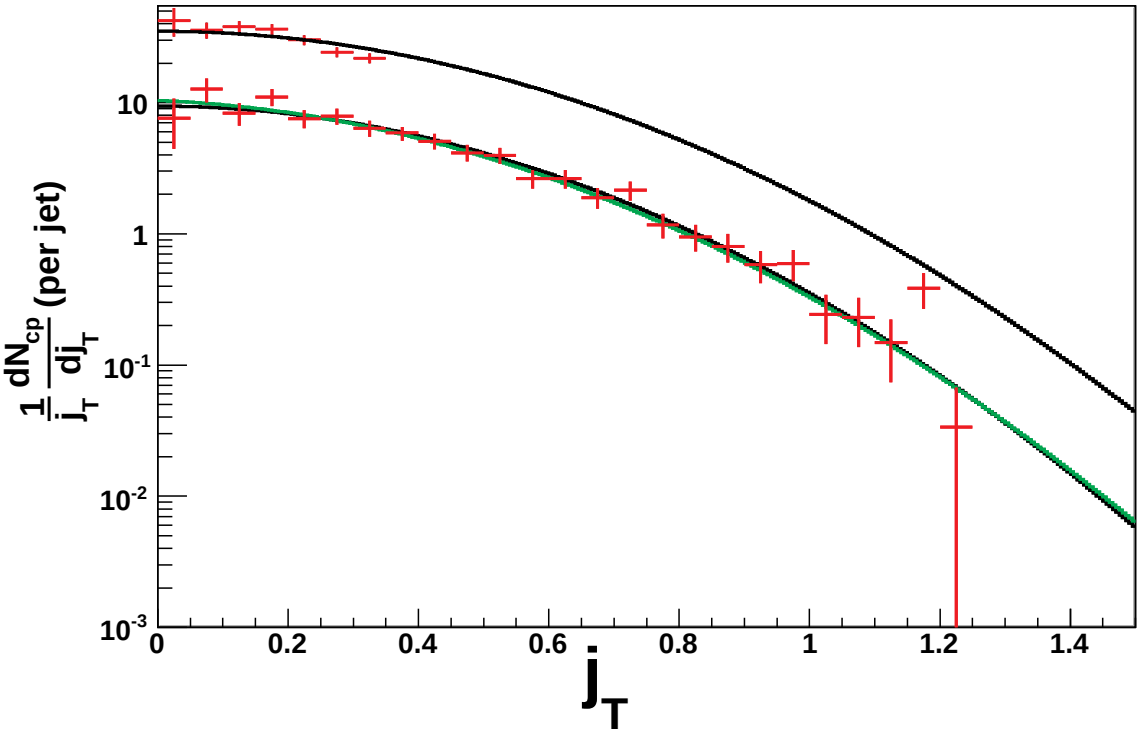

Figure 57. 15-20 GeV jet c.p. j_T from p+p with Gaussian fit.
Black lines are Gaussians centered at zero, Green line is Gaussian with negative mean. p+p data corrected for detector and radius effects are the red data points. Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5.

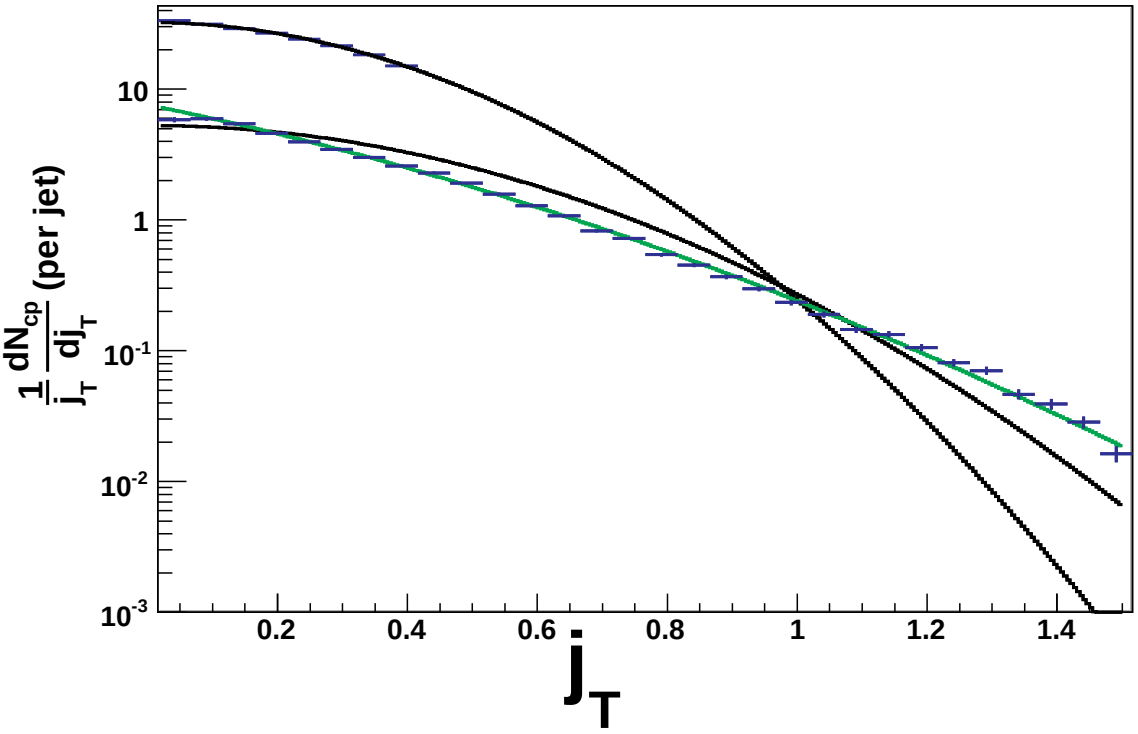

Figure 58. 6-7 GeV jet c.p. j_T from d+Au data with Gaussian fit.
Black is Gaussian centered at zero, Green is Gaussian fit with negative mean.
Blue data is d+Au corrected for detector, radius, and background effects (blue).
Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5 for data.

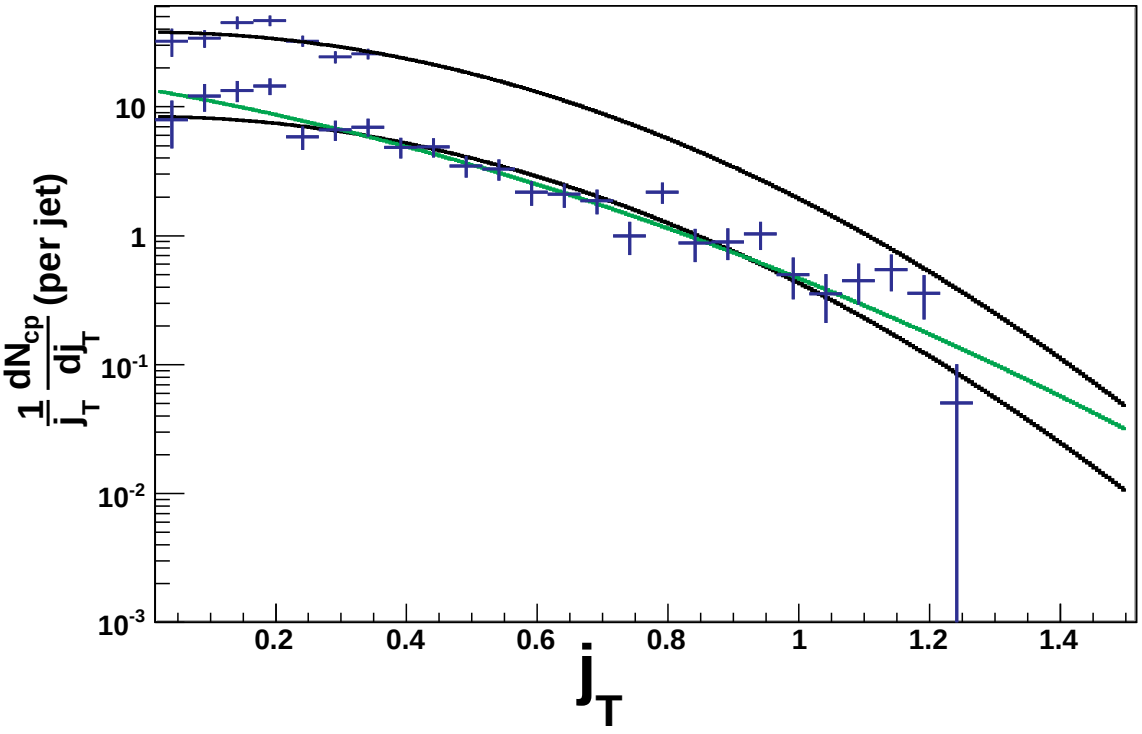

Figure 59. 15-20 GeV jet c.p. j_T from d+Au data with Gaussian fit.
Black is Gaussian centered at zero, Green is Gaussian fit with negative mean.
Blue data is d+Au corrected for detector, radius, and background effects (blue).
Upper band is for tracks > 0.6 GeV/c, lower > 2.0 GeV/c. Cone radius is 0.5 for data.

D. Back to Back Jet Reconstruction

The di-jet opening angle is an important observable obtained from jet events. For di-jets, the simple expectation from leading order QCD is that $\Delta\phi = \pi$. In a real collision, initial and final state gluon radiation shifts $\Delta\phi$ away from π. See Figure 60. The various non perturbative $\Delta\phi$ shift effects can be compiled into a phenomenological partonic parameter called k_T. This k_T model quantifies the transverse momentum of one parton

with respect to the beam. In this model, all the convolutions commute, since the model assumes Gaussian distributions for all k_T smearing.

Initial and final state k_T effects are completely ambiguous. Nonetheless, it is expected that each incoming parton gains a pre-scattering kick. Next, the final state partons start back to back in the center of mass frame, and proceed to undergo post-scatter broadening. Each final state parton therefore should have a total momentum kick equal to the sum of the post momentum kick, plus (or minus for the away parton) the pair momentum kick. The sum of the kicks of the incoming partons, assumed to be the same distribution for both, therefore contributes on average half of the center of mass transverse momentum to each outgoing parton.

Single jet (or hadron) inclusive cross sections are enhanced by this final-state k_T, which contains on average half the initial-state k_T sum and one unit of final-state broadening. For di-jet measurements, constructing distributions such as $\Delta\phi$ implicitly sums the k_T momentum of both partons before sampling. This includes all the initial-state and the sum of the final-state broadening.

When measuring $\vec{k}_T$ in this analysis, it is convenient to decompose the $\vec{k}_T$ vector into the three components most easily related to the detector. The total transverse momentum, $\vec{k}_T$, can be decomposed into detector (cylindrical) coordinates like so:

$$\vec{k}_T = k_{T\phi}\hat{\phi} + k_{T\eta}\hat{\eta} + k_{TE}\hat{E} \equiv k_{T\phi}\hat{\phi} + k_{TE}\hat{E} \qquad (15)$$

The $\hat{\eta}$ direction is defined as parallel to the beam axis pointing toward positive η. The $\hat{E}$ coordinate is then defined as a radial direction pointing in the direction of the charged-only jet. Finally, the $\hat{\phi}$ coordinate is defined perpendicular to the $\hat{E}$ axis, and is tangent to a circle centered on the beam axis. These coordinates are illustrated in Figure 60. Technically, k_T smearing effects also play a role in the $\hat{\eta}$ direction, but they are very difficult to quantify in hadron collisions since the momentum fraction carried by a hard-scattered parton cannot be measured on a per-event basis. Therefore the phenomenological k_T-smearing model is really 2 dimensional, and the magnitude of k_T

quoted in the literature assumes that $k_{T\eta} \equiv 0$. The easiest of the 2 nonzero components to measure is:

$$k_{T\phi} = E_T \sin(\Delta\phi), \tag{16}$$

which can be related, on average, to the per parton k_T quantity:

$$\left\langle k^2_{T\phi(parton)} \right\rangle = \frac{1}{2}\left\langle k^2_{T\phi(1)} + k^2_{T\phi(2)} \right\rangle = \frac{1}{2}\left\langle k^2_{T\phi} \right\rangle = \frac{1}{2}\left\langle E_T^2 \sin^2(\Delta\phi) \right\rangle \tag{17}$$

Where $k_{T\phi}$ is the $\hat{\phi}$ momentum component of k_T, $\Delta\phi$ is the angle between the jet and the negative projection of the away jet ($\Delta\phi$=0 when the jets are perfectly back to back), and E_T is the transverse momentum of the fully reconstructed jet.

It is difficult to measure the other component, $k_{TE} = E_1 - E_2 = \Delta E$, since this requires measuring the energy scale of both jets. The energy scale of charged-only jets is not well determined, while di-jets which are both detected in the BEMC require a complex acceptance correction which depends on the k_T distribution. However, assuming the k_T distribution is symmetric about the beam axis on average, $\left\langle k^2_{TE(parton)} \right\rangle = \left\langle k^2_{T\phi(parton)} \right\rangle$, and then $\left\langle k^2_{T(parton)} \right\rangle = 2 \cdot \left\langle k^2_{T\phi(parton)} \right\rangle = \left\langle E_T{}^2 \sin^2(\Delta\phi) \right\rangle$. The measurement of $k_T \equiv \sqrt{\left\langle k^2_{T(parton)} \right\rangle} = \sqrt{\left\langle E_T^2 \sin^2(\Delta\phi) \right\rangle}$ is described in the remainder of this section.

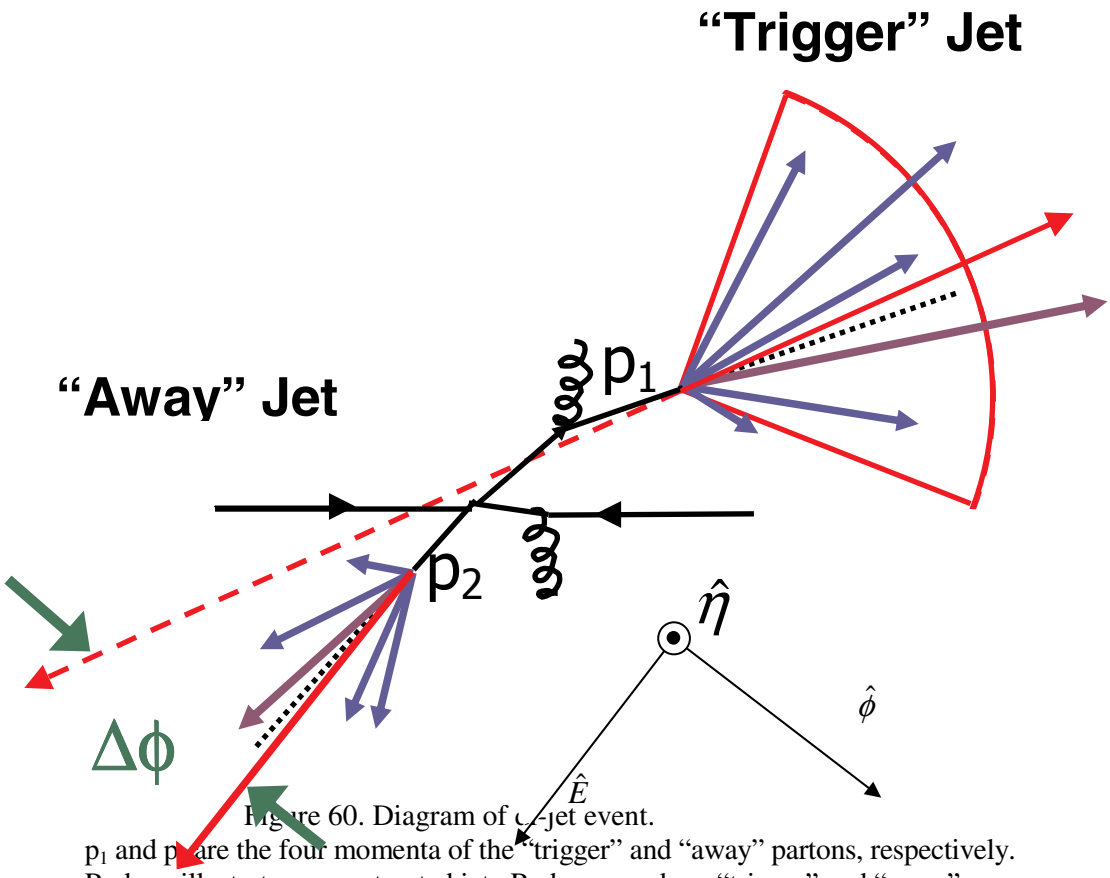

Figure 60. Diagram of c-jet event.
p_1 and p_2 are the four momenta of the “trigger” and “away” partons, respectively. Red arc illustrates reconstructed jet. Red arrows show “trigger” and “away” reconstructed jet axes. Green arrows show measured $\Delta\phi$.

When reconstructing di-jets, one “trigger” jet with observed $E_T > 7$ GeV is reconstructed from both neutral and charged hadrons. For this analysis, the “trigger” jet contains a high tower which triggered the event. This establishes the di-jet energy scale, but the BEMC acceptance during run 3 restricts this jet to positive η. The paired “away” jet direction is reconstructed using charged particles only. The “away” jet is therefore reconstructed in the less restricted range $-0.5 < \eta < 0.5$. Since the “away” jet by definition lacks neutral energy, it is not used to estimate jet-energy scale. Neutral particle

reconstruction from Run 3 was far from uniform in phi, and so a measurement of $\Delta\phi$ using neutral energy for both jets would require a large BEMC jet acceptance correction. A critical benefit of using charged-particle-only "away" jets is to avoid having to apply this complicated and large acceptance correction. However, this strategy also removes the incident parton distribution function bias which would exist if both jets were required to be contained within the Run 3 BEMC acceptance and thus have positive η.

1. p+p k_T

Like z and j_T, the p+p k_T measurement must be made in comparison with the PYTHIA simulations in a systematic manner. Unlike j_T, k_T is strongly dependent on the jet energy scale. While the definition of k_T shows that it is proportional to both the energy and $\sin(\Delta\phi)$, detector resolution effects convolute the measurements of both quantities. The energy and the $\sin(\Delta\phi)$ quantities are correlated, and therefore $\left\langle E_t^2 \right\rangle \left\langle \sin^2(\Delta\phi) \right\rangle$ does not necessarily equal $\left\langle E_T^2 \cdot \sin^2(\Delta\phi) \right\rangle$. It is possible to obtain $\left\langle E_t^2 \right\rangle$ and $\left\langle \sin^2(\Delta\phi) \right\rangle$ from their respective distributions, and correct these quantities for detector resolution effects, but these do not equal a measurement of the k_T distribution unless the correlation can be carefully taken into account. A clearer approach is to correct for the detector resolution effects on k_T directly by comparing the PYTHIA no-GEANT and PYTHIA with-GEANT simulations, and to quantify the detector resolution effect on $\left\langle E_T^2 \cdot \sin^2(\Delta\phi) \right\rangle$. Once the detector resolution has been quantified, the RMS(k_T) can be computed from the average of the measured quantity.

Let $\rho, R_E, \psi,$ and R_ψ be random variables. The random variable ρ models di-jet E_T, ψ models di-jet $\sin(\Delta\phi)$, $\overline{\rho} \equiv \rho + R_E$ models reconstructed di-jet E_T, and $\overline{\psi} \equiv \psi + R_\psi$ models the reconstructed di-jet $\sin(\Delta\phi)$. The subscript S where used indicates values derived from simulations, and the subscript D indicates values from data.

Ideally, the detector resolution effects modeled by the random variables R_E and R_ψ are the same for both simulation and data. The average values of ψ and R_ψ are zero since both true and reconstructed $\sin(\Delta\phi)$ distributions are symmetric about the origin.

Following these definitions, $\mathrm{RMS}(k_T) \equiv \sqrt{\langle E_T \sin(\Delta\phi)\rangle^2} = \sqrt{\langle \rho^2\psi^2\rangle}$. Neither the product $\rho^2\psi^2$, nor $\langle \rho^2\psi^2\rangle$ can be directly measured. It is possible, however, to obtain an expression for $\rho^2\psi^2$ in terms of other measured quantities. To obtain $\langle \rho^2\psi^2\rangle$ in terms of known quantities, expand $\langle \bar{\rho}^2\bar{\psi}^2\rangle$ and solve for $\langle \rho^2\psi^2\rangle$:

$$\begin{aligned}
&\langle \bar{\rho}^2\bar{\psi}^2\rangle = \langle (\rho^2 + 2\rho R_E + {R_E}^2)(\psi^2 + 2\psi R_\psi + {R_\psi}^2)\rangle \\
&= \langle \rho^2\psi^2\rangle + \langle 2\rho^2\psi R_\psi\rangle + \langle \rho^2 {R_\psi}^2\rangle + \langle 2\rho R_E\psi^2\rangle + \langle 4\rho R_E\psi R_\psi\rangle + \langle 2\rho R_E {R_\psi}^2\rangle \\
&+ \langle {R_E}^2\psi^2\rangle + \langle 2{R_E}^2\psi R_\psi\rangle + \langle {R_E}^2{R_\psi}^2\rangle \Rightarrow \\
&\langle \rho^2\psi^2\rangle = \langle \bar{\rho}^2\bar{\psi}^2\rangle - 2\langle \rho^2\psi R_\psi\rangle - \langle \rho^2{R_\psi}^2\rangle - 2\langle \rho\psi^2 R_E\rangle - 4\langle \rho R_E\psi R_\psi\rangle \\
&- 2\langle \rho R_E{R_\psi}^2\rangle - \langle {R_E}^2\psi^2\rangle - 2\langle {R_E}^2\psi R_\psi\rangle - \langle {R_E}^2{R_\psi}^2\rangle
\end{aligned} \tag{18}$$

It is useful to be able to substitute in for $\langle \rho\psi^2 R_E\rangle$. A similar strategy yields the following:

$$\begin{aligned}
&\langle \bar{\rho}\bar{\psi}^2 R_E\rangle = \langle (\rho + R_E)(\psi + R_\psi)^2 R_E\rangle = \langle \rho\psi^2 R_E\rangle + \langle {R_E}^2\psi^2\rangle + 2\langle \rho\psi R_\psi R_E\rangle \\
&+ 2\langle {R_E}^2\psi R_\psi\rangle + \langle \rho R_E{R_\psi}^2\rangle + \langle {R_E}^2{R_\psi}^2\rangle \Rightarrow \\
&\langle \rho\psi^2 R_E\rangle = \langle \bar{\rho}\bar{\psi}^2 R_E\rangle - \langle R_E^2\psi^2\rangle - 2\langle \rho\psi R_\psi R_E\rangle - 2\langle {R_E}^2\psi R_\psi\rangle - \langle \rho R_E{R_\psi}^2\rangle \\
&- \langle {R_E}^2{R_\psi}^2\rangle
\end{aligned} \tag{19}$$

Substituting then yields:

$$\begin{aligned}
&\langle \rho^2\psi^2\rangle = \langle \bar{\rho}^2\bar{\psi}^2\rangle - 2\langle \bar{\rho}\bar{\psi}^2 R_E\rangle + \langle R_E^2\psi^2\rangle + 2\langle {R_E}^2\psi R_\psi\rangle \\
&+ \langle {R_E}^2{R_\psi}^2\rangle - \langle \rho^2{R_\psi}^2\rangle - 2\langle \rho^2\psi R_\psi\rangle
\end{aligned} \tag{20}$$

Except for the first term, $\left\langle \overline{\rho}^2 \overline{\psi}^2 \right\rangle$, each average must be quantified using the simulation. Nevertheless, this expectation value provides the means to calculate the width of the k_T distribution:

$$RMS(k_T) \equiv \sqrt{\left(E_T \sin(\Delta\phi)\right)^2} = \sqrt{\left\langle \rho^2 \psi^2 \right\rangle} \qquad (21)$$

Figure 61 shows the $\overline{\psi}^2$ distribution for PYTHIA and data. The tail of this distribution has a large effect on the $\left\langle \overline{\psi}^2 \right\rangle$, and so a cut on $\overline{\psi}^2$ was imposed on all expectation values. This means that $RMS(k_T)$ is not taken from the entire k_T distribution. Therefore particular care is taken in interpreting this $RMS(k_T)$ value in terms of the true σ_{kT}. Figure 62 shows the $\overline{\rho}$ distribution for PYTHIA and data.

Table 5 shows the values and uncertainties of the known quantities. In each case, a histogram of the product is produced, and the mean is read off as the expectation value and listed in the value column. The statistical uncertainty column is then populated by reading off the RMS value and dividing by the square root of the number of entries. Several rows contain two entries. The top entry is with a $\overline{\psi}^2$ cut of 0.25, and the bottom entry is with a cut of 0.30. The row labeled $\left\langle \overline{\rho}^2 \overline{\psi}^2 \right\rangle$ is obtained from the data, while $\left\langle \overline{\rho}_S^2 \overline{\psi}_S^2 \right\rangle$ and the other quantities listed are all obtained from the simulation, as denoted by the subscript "S".

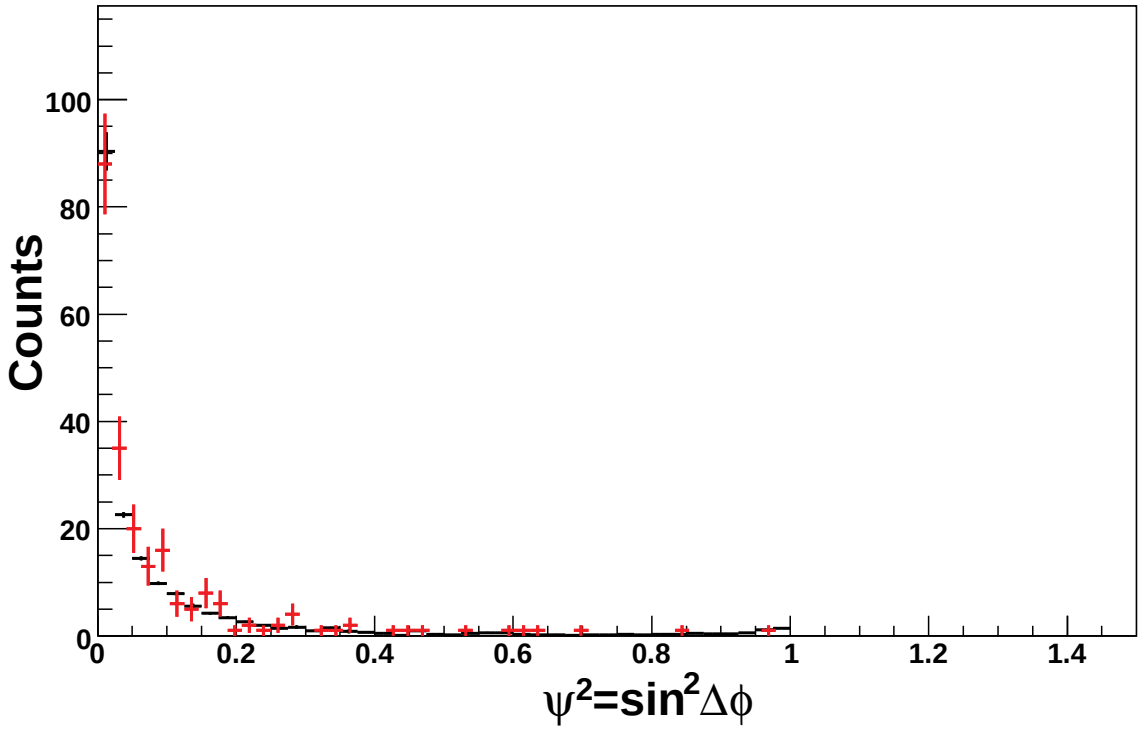

Figure 61. $\sin^2(\Delta\phi)$ distribution for PYTHIA and data.
PYTHIA is scaled arbitrarily for comparison, shown in black. p+p data is red.

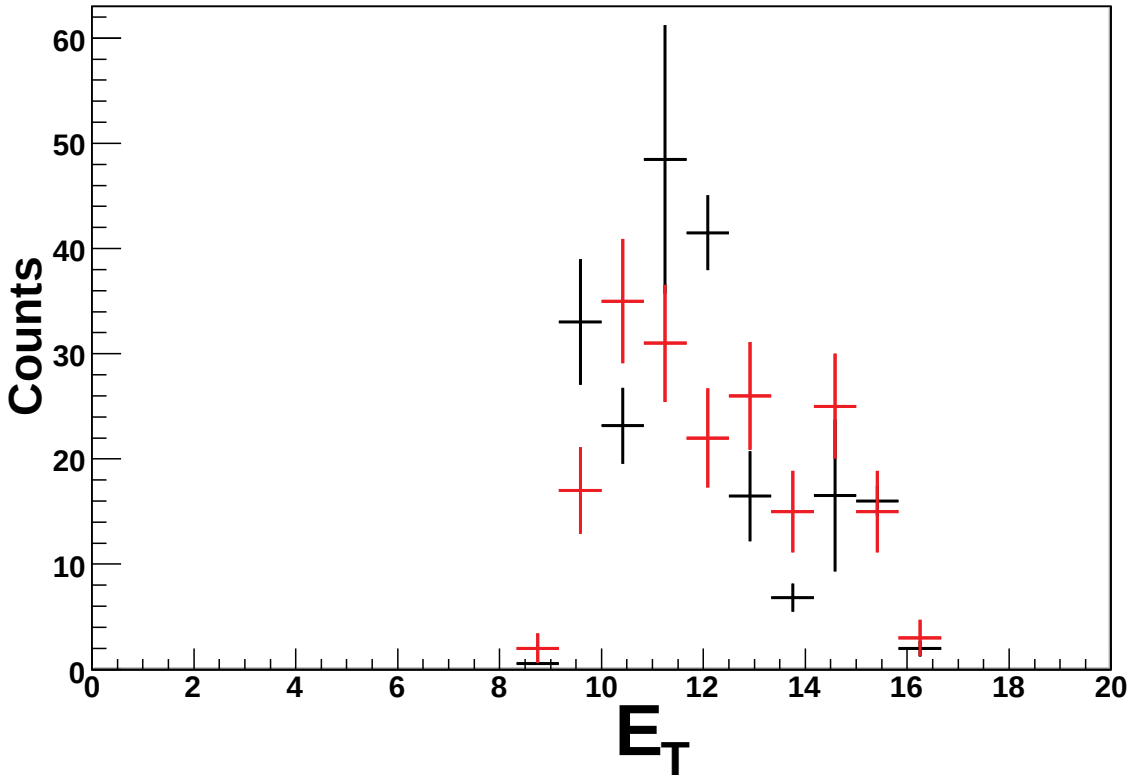

Figure 62. Di-jet E_T distribution for PYTHIA and data.
PYTHIA is scaled arbitrarily for comparison and shown in Black. p+p data is red.

Table 5. Values, statistical uncertainties, and systematic uncertainties of terms used to calculate RMS(k_T) $\equiv \sqrt{\langle \rho^2 \psi^2 \rangle}$. All quantities are in GeV^2/c^2, except RMS(k_T) which is GeV/c.

Quantity	Val	Stat	Sys	Quantity	Val	Stat	Sys
$\langle \overline{\rho}^2 \overline{\psi}^2 \rangle$	5.8 6.2	0.44 0.49	0.4				
$\langle \overline{\rho}_S^2 \overline{\psi}_S^2 \rangle$	5.51 5.64	0.34 0.35	0.4				
$-2\langle \overline{\rho}_S \overline{\psi}_S{}^2 R_E \rangle$	- 0.004	0.14					
$\langle R_E{}^2 \psi_S^2 \rangle$	0.19	0.021					
$2\langle R_E{}^2 \psi_S R_\psi \rangle$	-0.22	0.08					
$\langle R_E{}^2 R_\psi{}^2 \rangle$	0.185	0.05	0.08				
$-\langle \rho_S{}^2 R_\psi{}^2 \rangle$ ††	-1.00	0.1	1.15				
$-2\langle \rho_S{}^2 \psi_S R_\psi \rangle$	0.40	0.23					
$=\langle \rho^2 \psi^2 \rangle$	5.35 5.75	0.53	1.22	RMS(k_T)	2.31 2.40	0.11	0.27
$=\langle \rho_S^2 \psi_S^2 \rangle$	5.06 5.19	0.46	1.22	RMS(k_T)	2.25 2.28	0.10	0.27

The values of the average quantities used to calculate $\langle \rho^2 \psi^2 \rangle$ and extracted from PYTHIA for di-jets with identical cuts to that of the data are shown in Table 5. While the full PYTHIA simulation contains a large number of events, the requirement that the PYTHIA reconstructed jets possess a "high tower" is quite strict. By the time all the cuts from the data are applied to the simulation, the simulation and the data values are equivalent given the uncertainties.

†† A cut was applied to remove outliers as discussed in the text

The distribution used to find $-\left\langle \rho_S{}^2 R_\psi{}^2 \right\rangle$ possesses a few data points with very high values. Figure 63 shows this distribution up to 100 GeV^2/c^2. Three cuts were applied, one at 15 GeV^2/c^2 (giving 0.98), one at 30 GeV^2/c^2 (giving 1.03), and no cut (giving 2.15 GeV^2/c^2). The values from the first two cuts were used to estimate the value, and the difference between that value and the no-cut value to estimate the systematic error.

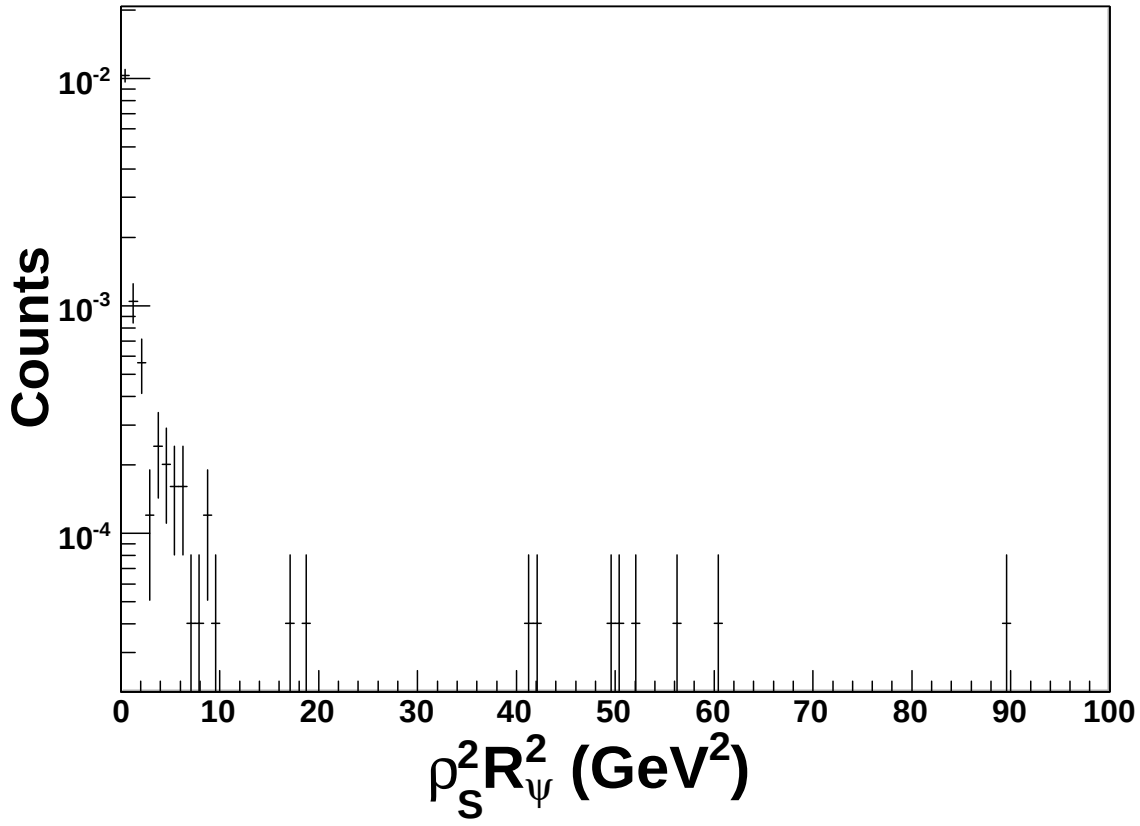

Figure 63. $\rho_S^2 R_\psi{}^2$ distribution obtained from simulation.

The partial RMS(k_T) measured above is not exactly comparable to other experiments. The value desired is the Gaussian width of the k_T distribution. Since the partial RMS(k_T) is proportional to the Gaussian width, the two quantities can be related

using the simulation. Assuming that the shape of the k_T distribution is the same in both the data and simulation, the relative width ratios will be the same for data and simulation as well:

$$\frac{RMS(k_{TD})^2\Big|(C_D = x\cdot\sigma_{kTD})}{\sigma^2{}_{kTD}} = \frac{RMS(k_{TS})^2\Big|(C_S = x\cdot\sigma_{kTS})}{\sigma^2{}_{kTS}} \qquad (22)$$

Where C_D and C_S are the cuts imposed on the data and simulation respectively. To interpret the partial p+p RMS(k_T) measured thus far, it is necessary to fit the PYTHIA k_T distribution and measure a conventional width. The partial RMS(k_T) is dependent on the $\overline{\psi}^2 < 0.25$ cut, but the conventional width from the fit, σ_{kT}, is not. Using a Gaussian assumption for the shape, Figure 64 shows the fit to the PYTHIA k_T distribution with a $\overline{\psi}^2 < 0.25$ cut. The PYTHIA fit gives $\sigma_{kT} = 1.99$ GeV/c, while using the same values as in the table, the PYTHIA RMS(k_T)=2.25 GeV/c. It appears that the PYTHIA partial RMS(k_T) is larger than the fitted width by 12%.

Using the ratios assumption, the p+p width from the table is $\sigma_{kT} = 2.08 \pm 0.12 \pm 0.12$ GeV/c, quite similar to the PYTHIA derived $\sigma_{kT} = 1.99 \pm 0.02$ GeV/c. The PYTHIA σ_{kT} is independent of the $\overline{\psi}^2$ cut to within statistical uncertainty, and so the average of the fitted σ_{kT} values (1.97±0.02 GeV/c for $\overline{\psi}^2 < 0.3$ and 2.01±0.02 GeV/c for $\overline{\psi}^2 < 0.25$) is used. Even though the wider $\overline{\psi}^2 < 0.3$ cut produces an RMS(k_T) slightly larger than with the smaller $\overline{\psi}^2 < 0.25$ cut, the p+p width nonetheless works out to $\sigma_{kT} = 2.13$ GeV/c when the ratio assumption is taken into account. The values derived from the two cuts are consistent, showing the robustness of the overall derivation and the insensitivity to the $\overline{\psi}^2$ cut chosen. The systematic uncertainty on the p+p σ_{kT} has been reduced from the RMS(k_T) systematic uncertainty because the large systematic uncertainty introduced by the $-\left\langle \rho_S{}^2 R_\psi{}^2 \right\rangle$ term is correlated between the p+p

and PYTHIA RMS(k_T) values. If the $-\left\langle \rho_S{}^2 R_\psi{}^2 \right\rangle$ is assumed equal to the extreme -2.15 GeV^2/c^2, the p+p $\sigma_{kT} = 2.10$ GeV/c.

The raw $\left\langle \overline{\rho}^2 \overline{\psi}^2 \right\rangle$ terms from p+p and PYTHIA were not a priori expected to be proportional to the width of the respective k_T distributions. Nonetheless, it is useful to note that the raw PYTHIA value for $\left\langle \overline{\rho}^2 \overline{\psi}^2 \right\rangle$, 5.5 GeV^2/c^2, is very similar to the p+p value. If $\left\langle \overline{\rho}^2 \overline{\psi}^2 \right\rangle$ was proportional to the width of k_T, the calculation would suggest that p+p $\sigma_{kT} = 2.09$ GeV/c. Taking all effects into account, and systematic uncertaintites just discussed, $\sigma_{kT} = 2.08 \pm 0.12 \pm 0.13$ GeV/c.

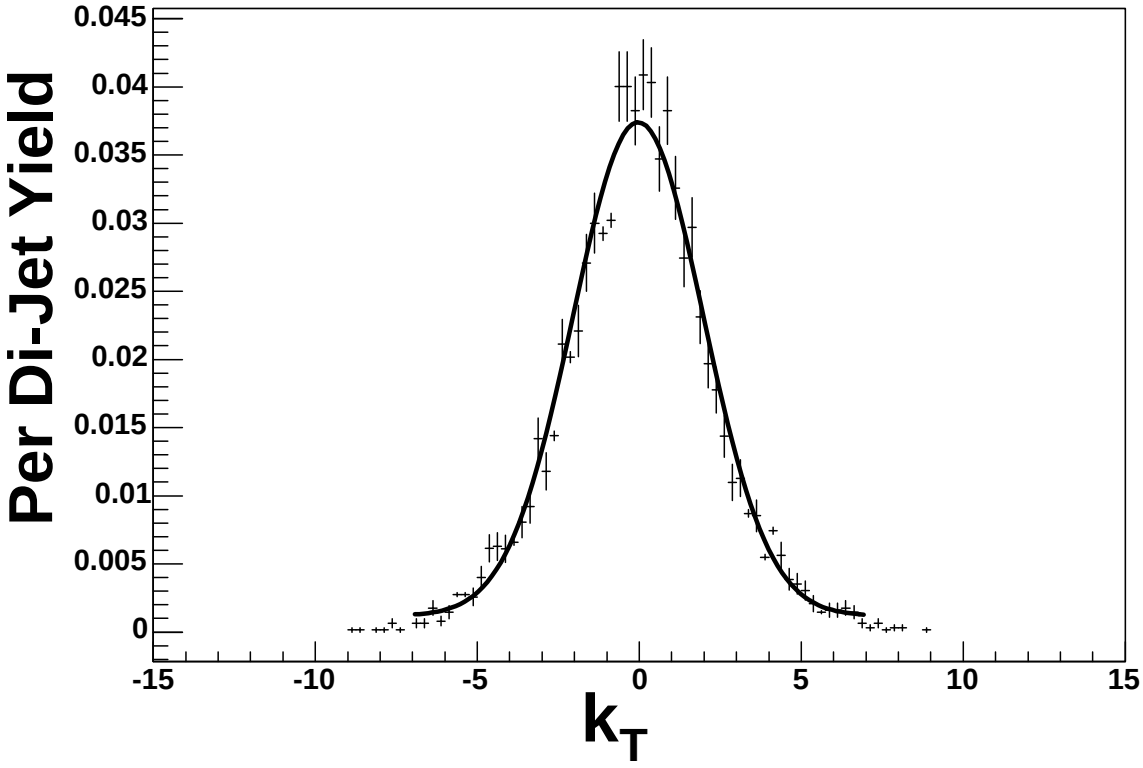

Figure 64. PYTHIA di-jet k_T.
k_T is reconstructed from PYTHIA without detector simulation. Fit is with a Gaussian + constant. Average fitted width is 1.99±0.02 GeV/c (see discussion).

2. d+Au Nuclear k_T

The nuclear component of the d+Au partonic k_T is in some ways easier to measure than the p+p partonic k_T. The nuclear component of the d+Au partonic k_T which arises from the presence of the Au nucleus can be measured relative to the p+p baseline. Unlike the p+p partonic k_T measured in the last section, many systematic uncertainties roughly cancel in taking the difference between d+Au and p+p. The nuclear k_T component can be represented using ρ and ψ as defined in the last section, where several of the terms completely cancel:

$$\sigma^2_{kTNuc} \equiv \left\langle \rho^2_{dAu}\psi^2_{dAu} \right\rangle - \left\langle \rho^2_{pps}\psi^2_{pps} \right\rangle \qquad (23)$$

Where σ_{kTNuc} is the component of the d+Au partonic k_T contributed by the nucleus, $\left\langle \rho^2_{dAu}\psi^2_{dAu} \right\rangle$ is the average $\rho^2\psi^2$ product for d+Au, and $\left\langle \rho^2_{pps}\psi^2_{pps} \right\rangle$ is the average $\rho^2\psi^2$ product for p+p supplemented by d+Au background particles as described in section V.A.4. Table 6 shows the values of the differences between d+Au and p+p supplemented di-jet quantities.

For most of the quantities in Table 6, except the basic difference $\Delta\left\langle \overline{\rho}^2\overline{\psi}^2 \right\rangle$, the value must be estimated from the simulation. For this dissertation, the energy and angular detector resolutions are assumed to be the same for both systems, which implies that the detector resolution factors in the table are also the same. An effort was made, by imposing

equivalent cuts, to make the p+p+supp and d+Au distributions have the same average energy. Since the actual distributions have fluctuations, there are still small differences in the average energy of the two samples. Each instance of ρ, whenever it appears, should be proportional to the ratio of the d+Au E_T to the p+p supplemented E_T. The value of the E_T ratio is 0.99 ± 0.03, obtained by taking the ratio of the average E_T of d+Au and supplemented p+p, i.e. 12.87 GeV/ 13.02 GeV. Similarly, each instance of ρ^2 should be proportional to the ratio of d+Au to p+p+supp E_T^2. This ratio is 0.94 ± 0.05, i.e. 126.9 GeV^2/ 134.6 GeV^2, which is evidently a little smaller than 0.99^2 because of slight differences in the distributions. Each ψ^2 term should be proportional to the ratio of d+Au ψ^2 to the supplemented p+p ψ^2. The value of the ψ^2 ratio is 1.1 ± 0.1, i.e. $(0.0354/0.0322)$. The d+Au term of the difference should be increased by these ratios for each factor of ρ, ρ^2 and ψ^2 found in the term. The p+p+supp term of the difference should remain unmodified. Thus, an estimate of each difference can be formed, for example, $\Delta\left\langle \overline{\rho}_S \overline{\psi}_S{}^2 R_E \right\rangle = 0.002 * 0.99 * 1.1 - 0.002 = 0.00018$. As the table shows, the nuclear k_T component of d+Au k_T is consistent with zero to a precision of about 0.5 GeV/c.

Table 6. Values, statistical uncertainties, and systematic uncertainties of terms used to calculate RMS(Δk_{TdAu}) $\equiv \sqrt{\Delta\langle \rho^2 \psi^2 \rangle}$. All values are in GeV^2/c^2 except RMS(Δk_{TdAu}) which is GeV/c.

Quantity	Val	Stat	Sys	Term	Val	Stat	Sys
$\Delta\langle \overline{\rho}^2 \overline{\psi}^2 \rangle$	0.13 -0.27	0.53	0.4				
$-2\Delta\langle \overline{\rho}_S \overline{\psi}_S{}^2 R_E \rangle$	-0.00036	0.0006					
$\Delta\langle R_E{}^2 \psi_S^2 \rangle$	0.023	0.02					
$2\Delta\langle R_E{}^2 \psi_S R_\psi \rangle$	0.0	0.0					
$\Delta\langle R_E{}^2 R_\psi{}^2 \rangle$	0.0	0.0					
$-\Delta\langle \rho_S{}^2 R_\psi{}^2 \rangle$ ‡‡	0.06	0.21					
$-2\Delta\langle \rho^2{}_S \psi_S R_\psi \rangle$	0.12	0.18					
$=\Delta\langle \rho^2 \psi^2 \rangle$	0.33 -0.07	0.6	0.4	RMS(Δk_{TdAu})	0.57 -0.26	0.5 0.34	0.4 0.4

E. The Challenge of R_{dAu} via Jets

To measure the dependence of R_{dAu} on p_T to the highest possible energies, it is desirable to obtain a sample with many high-energy jets. Measuring a relative cross section of jets in d+Au "high-tower" triggered events is the next step in obtaining R_{dAu}. The d+Au "minimum-bias" sample is also substantial, and reconstructing jets from "minimum bias" may even be an improvement over R_{dAu} derived from inclusive hadron measurements. Unfortunately, this cannot be done with year 2003 p+p data, since the year 2003 p+p "minimum-bias" data is very limited. The p+p "minimum-bias" picture improves in year 2005 data, but that is beyond the scope of this analysis. Looking

‡‡ The same cut applied in Table 4 is applied here also.

forward, it should be fairly easy to utilize more recent p+p "minimum-bias" data in combination with the d+Au "minimum-bias" data to produce an interesting R_{dAu} measurement with a noticeable improvement over previous hadron R_{dAu} observations. Utilizing year 2003, 2004, 2005, and 2006 p+p "high-tower" data combined with the d+Au "high-tower" data from 2003, while more challenging than for "minimum-bias" jets, is even more promising for measuring R_{dAu} at the highest p_T so far. Unfortunately, there was a different "high-tower" threshold for each of these four years, and this could complicate an analysis unless strict "high-tower" threshold cuts are imposed.

No one has yet obtained the jet cross section in d+Au "high-tower" triggered events. This will be a challenge. Figure 65 shows the non-uniformity in d+Au reconstructed jet ϕ. It will be necessary to understand in greater depth the detector acceptance issues involved. For the measurements mentioned in this dissertation, this non-uniformity causes few problems. For R_{dAu}, however, the acceptance needs to be well defined. Another issue can be observed from Figure 8. This estimate of the energy scale in d+Au works well for z, j_T, and k_T. j_T is relatively insensitive to the energy scale, while the percentage uncertainty of z and k_T from the energy scale is roughly equal to the percentage uncertainty of the energy scale. In contrast, the R_{dAu} uncertainty is 10%[23] due to a 2.7%[23][25][26] uncertainty in charged particle p_T. Extrapolating from 14% energy uncertainty from Figure 8, and using this relationship between energy uncertainty and R_{dAu} uncertainty, yields an R_{dAu} uncertainty of 50%. Understanding the energy scale and the ϕ acceptance will be significant challenges when measuring R_{dAu} in the "high-tower"-triggered events.

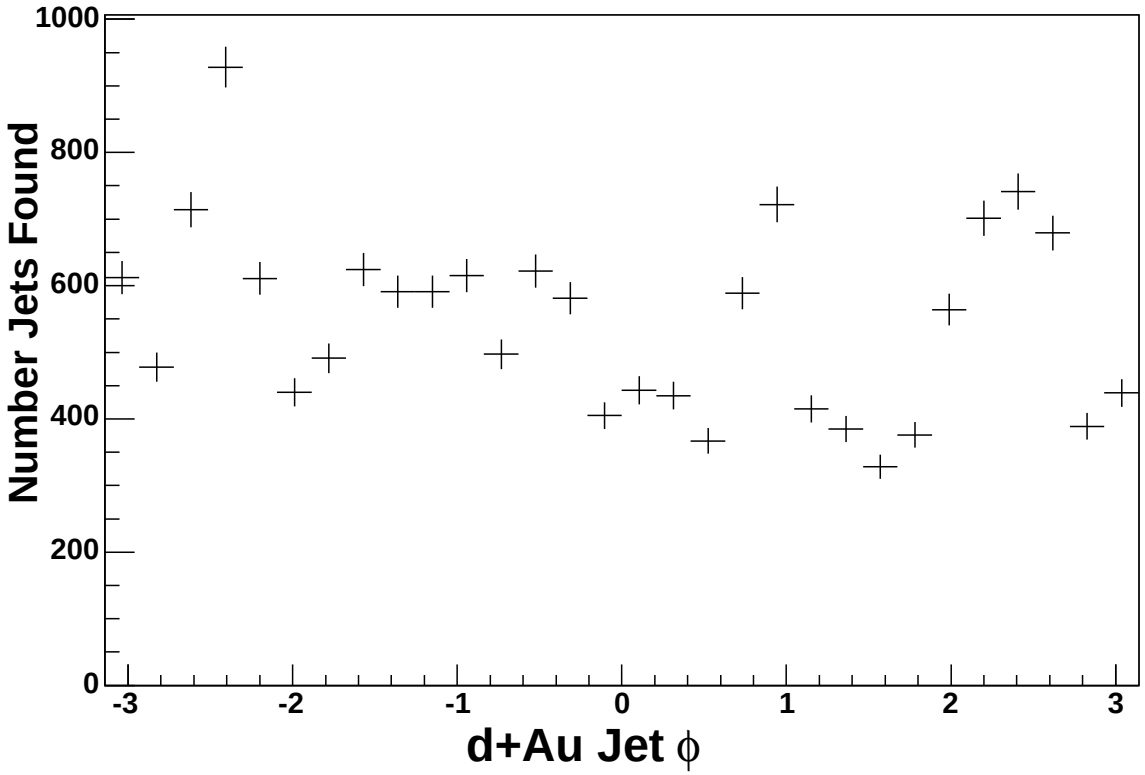

Figure 65. Jet ϕ reconstructed from d+Au "high-tower" triggered events.

VIII. CONCLUSIONS

A. Data vs PYTHIA as Phenomenological Comparison

PYTHIA has been designed as a parameterization of real data. Some of the models used to parameterize particle production in PYTHIA are based upon theoretical physics, such as the string fragmentation. However, the parameters used to guide the output of PYTHIA are tuned in order to obtain close agreement with the data, with less emphasis placed on the theoretical sensibilities of the parameters.

Comparing data to PYTHIA is therefore in a sense comparing data to data, or more accurately, comparing data to trends in other data. A great deal of data has come from the Tevatron at Fermilab, and the most significant recent tunings of PYTHIA were done relative to those experimental results[10]. PYTHIA, in its current incarnation at STAR, produces simulations that are essentially an extrapolation to lower jet energy from higher jet p_T data at other experiments. Any significant deviations between data and simulation would be very interesting. If the data deviate from trends at higher energy, it implies the existence of a physical process not parameterized by PYTHIA. This section places the jet results from STAR into that context.

1. Jet Fragmentation Phenomenological Agreement

As seen in Figures 41 and 42, above a certain minimum z value, the p+p data and the PYTHIA simulation are in very good agreement. When the simulation is plotted for cone radius 0.5 as in Figures 31 and 32, the differences are even less apparent. As Table 2 shows, for 6-7 GeV jets the slope of an exponential fit is about -9, and for 15-20 GeV it is about -10. A fit to Figure 1 gives 8.8±0.2, and is consistent with the 6-7 GeV jet z slopes. All evidence suggests that the extrapolation of the longitudinal fragmentation physics from higher to lower jet energy works well, and that any physics not modeled by PYTHIA at higher energy has little effect even at jet energies as low as 6 GeV.

At higher $\sqrt{s} = 900$ GeV experiments, the typical 50-60 GeV reconstructed jets come from partons with a fraction of total nucleon momentum (x_T) of around 0.06 to 0.07. The $6 < E_T < 7$ GeV jets reconstructed in this analysis have a similar x_T with dramatically different Q^2. However, the $15 < E_T < 20$ GeV jets reconstructed from the "high-tower" data sample come from a higher x_T range of about 0.15, and while they also agree with PYTHIA, they have a slightly softer fragmentation physics. The p+p fragmentation physics was in some sense anticipated, but this is the first time a z distribution has been obtained from d+Au collisions. The agreement between d+Au and p+p is an important result.

2. Jet j_T

$\sqrt{\langle j_T^2 \rangle}$ =612±12±30 MeV/c for p+p and $\sqrt{\langle j_T^2 \rangle}$ =630±13±30 MeV/c for d+Au. The d+Au appears to track extremely well with the p+p data. Figures 52 and 53 show minimal differences between simulation, p+p, and d+Au j_T. The $\frac{1}{j_T}\frac{dN}{dj_T}$ distribution does not appear to be a Gaussian centered at zero, as the values in Table 4 demonstrate. The PYTHIA simulation also agrees better with a Gaussian with a significant negative mean.

Recent di-hadron analyses at RHIC, notably PHENIX, have measured $\sqrt{\langle j_T^2 \rangle}$ [13]. They find for p+p collisions that $\sqrt{\langle j_T^2 \rangle} = 585 \pm 6 \pm 15$ MeV/c. Since this is a di-hadron analysis, one arguably expects $\sqrt{\langle j_T^2 \rangle}$ to be lower. Taking all particles into account, PYTHIA calculations produce $\sqrt{\langle j_T^2 \rangle} = 630$ MeV/c. The same PYTHIA simulation shows that if the j_T measurement is restricted to the leading particle, which is the single most energetic particle in the jet, $\sqrt{\langle j_T^2 \rangle} = 557$ MeV/c. The leading particle plays a larger role then other particles in defining the thrust axis, and thus falls closer to the thrust axis

on average than the other particles. It is therefore reasonable to assume that a di-hadron measurement would fall somewhere in between the leading and next to leading particle $\sqrt{\langle j_T^2 \rangle}$, and well below the all particle p+p $\sqrt{\langle j_T^2 \rangle}$ measured here of $612 \pm 12 \pm 30$ MeV/c. Furthermore, PHENIX finds no significant difference between p+p and d+Au $\sqrt{\langle j_T^2 \rangle}$ [27], which is also consistent with this analysis.

B. p+p k_T

The k_T Gaussian sigma (width) measured here equals 2.08±0.12±0.13 GeV/c. The good agreement between simulation and p+p k_T is expected to some extent. k_T evolves as the log of the center of mass energy, as literature surveys have shown [14]. It comes as little surprise that PYTHIA correctly interpolates the exponential evolution of k_T to an intermediate $\sqrt{s}$. The k_T width measured in this dissertation is effectively a leading order quantity that incorporates initial and final state gluon radiation (which appear as next to leading order in a pQCD calculation) in addition to "intrinsic" k_T effects.

1. Theoretical k_T

Theoretical calculations have been performed for the value of k_T by various authors, such as I. Vitev [28]. That author, for instance, calculates that $\sqrt{\langle k_T^2 \rangle_{pp}} = 1.92$ GeV/c at $\sqrt{s} = 200$ GeV as shown in Table 5 of his paper. His motivation is to use this value to estimate the effect that the k_T will have on single hadronic cross sections when lower energy hadrons get kicked to higher energy by the k_T momentum. This is in the context of single hadrons, and it is also a theoretical quantity not necessarily comparable to the experimental k_T measurement done in this dissertation. Nonetheless, it is interesting to note the similarity of these experimental and theoretical numbers.

2. Other Experimental Measurements

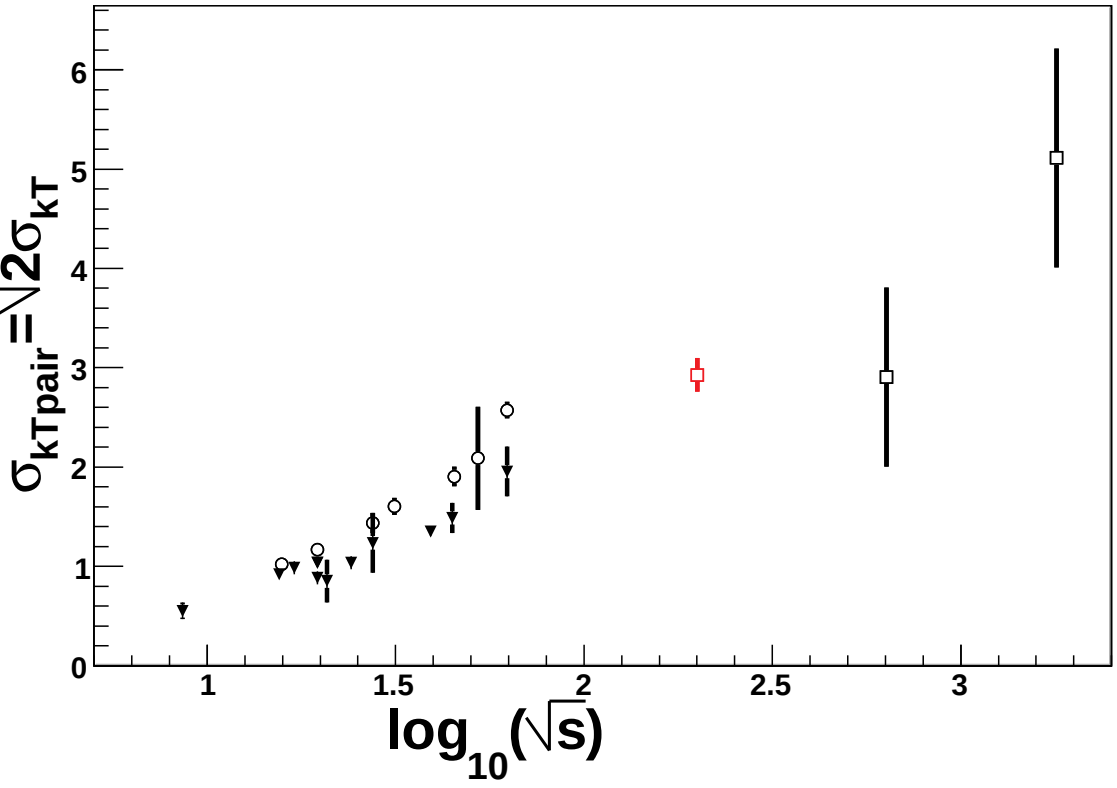

Figure 66. Dependence of σ_{kT} on $\sqrt{s}$.
Triangles are for di-muon, circles for di-photon, and squares di-jets. Red data point is 200 GeV σ_{kT} as described in section VII. Red error bar is statistical and systematic uncertainties added in quadrature.

Apanasevich *et. al.* conducted a broad survey of experimental k_T measurements [14]. Figure 66 shows that the measurement of k_T done in section VII is completely consistent with world data. PHENIX also has measured $\sigma_{kT} = 2.37 \pm 0.06 \pm 0.3$ GeV/c[13], which is consistent with the measurement reported here.

C. d+Au Nuclear k_T

Only experiments at RHIC have attempted to measure d+Au nuclear k_T at energies of $\sqrt{s} = 200\,\text{GeV}$. Using di-hadron correlations, PHENIX has estimated nuclear k_T to be very small[27]. The measurement described for this analysis in section VII is also consistent with zero. In contrast, a measurement at much lower $\sqrt{s} = 28.2\,\text{GeV}$ fixed target has apparently obtained a large nuclear k_T value of 1.78 ± 0.2 GeV/c[29]. These results are not necessarily contradictory, since they come from very different $\sqrt{s}$ values. It would be interesting to know at what $\sqrt{s}$ the nuclear k_T appears to vanish, and if the change is gradual or sudden.

1. k_T Broadening in Cold Nuclear Matter Theoretical Estimates

The author of [28] does not make a direct attempt to estimate $\sqrt{\left\langle \Delta k_T^2 \right\rangle_{dAu}}$, the nuclear component of k_T in d+Au. However, the following equations appear in that paper and are used to implement a theoretical estimate of the Cronin Effect. They can be used for an estimate of $\sqrt{\left\langle \Delta k_T^2 \right\rangle_{dAu}}$:

$$\left\langle \Delta k_T^2 \right\rangle_{dAu} = \frac{\mu^2}{\lambda} \langle L \rangle \xi \tag{24}$$

$$\frac{\mu^2}{\lambda_q} = 0.06 GeV^2 \Big/ fm \tag{25}$$

$$\frac{\mu^2}{\lambda_g} = 0.14 GeV^2 \Big/ fm \tag{26}$$

$$\langle L \rangle = 3/4 R_A = 3/4 \cdot 1.2 A^{\frac{1}{3}} = 5.9\, fm \tag{27}$$

$$\xi = \ln(1 + c_T \cdot E_T^2) \tag{28}$$

Where $E_T = 11.0$ GeV, and $c_T = 0.14$ GeV^{-2}. Calculating gives that $\sqrt{\langle \Delta k_T^2 \rangle_{dAu\ q}} = 1.01$ GeV/c, and $\sqrt{\langle \Delta k_T^2 \rangle_{dAu\ g}} = 1.54$ GeV/c, for quarks and gluons respectively. The theory that the Cronin Effect is generated by a nuclear $\sqrt{\langle \Delta k_T^2 \rangle_{dAu}}$ of between 1.0-1.5 GeV/c appears to be ruled out by this measurement of $\sqrt{\langle \Delta k_T^2 \rangle_{dAu}} \leq 0.5$ GeV/c. Perhaps the Cronin Effect is in reality generated by parton distributions which have been altered by the nuclear potentials.

2. Implications for Au+Au Collisions

The absence of cold nuclear broadening is a necessary measurement in order to interpret Au+Au results properly. Much work is currently being done to interpret di-jet signals obtained from Au+Au collisions[2], and the null measurement described in section VII.E is an important implicit assumption. The analysis in [2] quantifies the observed (lack of) back-to-back broadening through direct measurements of the back-to-back correlation width in d+Au, peripheral Au+Au, and central Au+Au. If the cold nuclear broadening k_T component had been large, the back-to-back jet correlations in Au+Au would be much more difficult to interpret. This analysis proves an important fact – that the nucleus normally has little effect on k_T broadening of di-jets at RHIC energies.

D. Future Directions

The Spin Physics program at RHIC calls for several more years of p+p running. The sources of background, which were not well understood in year 2003, can now be handled. In addition, as of the 2006 p+p, the STAR EMCs are fully instrumented, and shielding has been added upstream of the STAR detector to prevent beam background particles from reaching the detector. The procedure developed for obtaining p+p k_T shown in this dissertation is surprisingly effective, and can potentially be applied to measuring the Sivers effect[30].

With the large body of recent and future data, and the large coverage of STAR, it will also be interesting to examine the dependence of k_T on a broad range of parton *x*. Since average *x* can be calculated for a given sample of jets, it may be possible to see a correlation between longitudinal and transverse parton momentum. This could be especially valuable for spin physics, where the polarization of the proton is expected to modify these parton momentum distributions.

As mentioned at the end of the section VII, it will also be interesting to construct the jet R_{dAu} measurement and compare it with R_{dAu} obtained from hadron cross sections. Perhaps RHIC will even do another d+Au run in the future, and, coupled with a careful trigger implementation, this would allow an R_{dAu} measurement to be done to even higher p_T.

REFERENCES

[1] CTEQ Collaboration, S. Kretzer *et al.*, Phys. Rev. **D69**, 114005 (2004).

[2] STAR Collaboration, J. Adams *et al.*, *arXiv*: nucl-ex/0604018.

[3] M. Begel, Ph.D. Dissertation, University of Rochester, (1998).

[4] M. Gyulassy *et al.*, *arXiv*: nucl-th/0302077.

[5] D. Gross, Nucl.Phys.Proc.Suppl. **74** 426-446 (1999), *arXiv*: hep-th/9809060.

[6] M. E. Peskin, D. V. Schroeder, "An Introduction to Quantum Field Theory" (Westview, Boulder, CO, 1995).

[7] Yu L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977); V.N. Gribov L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972); G. Altarelli G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[8] M. Glück, E. Reya, A. Vogt, Eur. Phys. J. **C5**, 461–470 (1998).

[9] MRST Collaboration, A. D. Martin *et al.*, Eur. Phys. J. **C39**, 155 161 (2005).

[10] CDF Collaboration, R.D. Field *et al.*, *arXiv*: hep-ph/0510198, CDF-ANAL-CDF-PUBLIC-7822

[11] Particle Data Group, S. Eidelman *et al.*, Physics Letters **B592**, 1 (2004, 2005, and partial update for edition 2006).

[12] UA1 Collaboration, G. Arnison *et al.*, Phys. Lett. **B123**, 115 (1983).

[13] PHENIX Collaboration, S.S. Adler *et al.*, *arXiv*: hep-ex/0605039.

[14] E706 Collaboration, L. Apanasevich *et al.*, Phys. Rev. **D59,** 074007 (1999).

[15] STAR Collaboration, C. Adler *et al.*, Nucl. Instr. And Meth. **A499**, Issues 2-3, (2003)

[16] STAR Collaboration, C. Adler *et al.*, Nucl. Instr. And Meth. **A499**, Issues 2-3, 624-632 (2003)

[17] STAR Collaboration, C. Adler *et al.*, Nucl. Instr. And Meth. **A470**, 488 (01), *arXiv*: nucl-ex/0008005.

[18] M.A. Lisa. Technical Report, STAR Note **0238** (1996).

[19] R. Brun, R. Hagelberg, M. Hansroul, and J.C. Lassalle. Technical Report CERN-DD-78-2-REV, CERN.

[20] R. Bossingham et al. Technical Report, STAR Note **281** (1997).

[21] Michael Miller, Yale University, Ph.D. Dissertation, (2003).

[22] STAR Collaboration, J. Adams *et al.*, Phys. Rev. Lett. **91**, 172302 (2003), *arXiv*: nucl-ex/0305015.

[23] STAR Collaboration, J. Adams *et al.*, Phys. Rev. Lett. **91**, 072304 (2003), *arXiv*: nucl-ex/0306024.

[24] STAR Collaboration, A. Rose, *arXiv*: nucl-ex/0307015.

[25] STAR Collaboration, C. Adler *et al*., Phys. Rev. Lett. **90,** 032301 (2003).

[26] STAR Collaboration, J. Adams *et al.*, Phys. Rev. Lett. **91**, 072304 (2003).

[27] PHENIX Collaboration, S.S. Adler *et al.*, Phys. Rev. **C73**, 054903 (2006)

[28] I. Vitev, Phys. Lett. **B562**, 36 (2003), *arXiv*: nucl-th/0302002.

[29] E609 Collaboration, M.D. Corcoran *et al.*, Phys. Lett. **B259**, 209 (1991).

[30] D. Boer and W. Vogelsang, Phys.Rev. **D69**:094025 (2004), *arXiv*: hep-ph/0312320.

VITA

Name:	Thomas William Henry
Address:	18505 Clydesdale Rd, Colorado Springs, CO 80908
Email Address:	thomas-w-henry@neo.tamu.edu
Education:	B.S., Double Major Physics and Computer Science, Colorado State University, 1997