Abstract

Semi-Inclusive Measurement of the Shared Groomed Momentum Fraction of Jets in Au+Au Collisions at STAR

Daniel B. Nemes

2023

Jet quenching is one of the main signals used to investigate the properties of a quark-gluon plasma (QGP). Besides energy loss, jet quenching can also manifest itself in the modification of jet substructure. This thesis focuses on using STAR data to measure the substructure observable z_{g}, a result of SoftDrop grooming, which probes the physics of the first hard splitting of a hard-scattered parton. As opposed to previous measurements, this thesis explores employing a semi-inclusive approach to measure z_{g}, selecting candidate jets found within the recoil region of a high transverse momentum trigger particle. Requiring a high transverse momentum trigger object is expected to induce a surface bias on the event selection, potentially causing selected candidate jets in the recoil region to be biased towards having a longer path length within the medium. Consequently, these jets are expected to be more quenched and thus are good candidates to probe for modification of z_{g} at RHIC energies. Combinatorial jet contribution, arising from the large fluctuating background in heavy-ion collisions, is subtracted at the ensemble level, as opposed to employing jet-by-jet discrimination, in order to avoid imposing biases on the selected jet sample.

Semi-Inclusive Measurement of the Shared Groomed Momentum Fraction of Jets in Au+Au Collisions at STAR

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Daniel B. Nemes

Dissertation Director: Helen Caines

December 2023

Acknowledgments

First I would like to express my utmost gratitude to my advisor, Professor Helen Caines. Without Helen's guidance and support, this thesis work would not have been possible. I could not have asked for a better mentor, Helen encouraged me every step of the way. I will always be grateful for Helen and John letting me join Yale's RHI group and granting me the opportunity to learn and grow in such an open and welcoming environment. I would like to thank the RHI group members, from whom I learned so much from over the years. I would also like to thank the collaborators at STAR, particularly those in STAR's jet-related physics working group who provided valuable feedback and helped form my analysis for this thesis. And finally I would like to thank my family for all the love and support they have given during my time here at Yale, especially my two sisters Coleen and Erin. Coleen, thank you for being one of my closest friends, keeping me sane and well fed during the COVID lock downs. Erin, thank you for being there for me when I needed you the most, helping push me through stress and poor health to finish this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory that governs the strong interactions between quarks and gluons, the constituents of baryons and mesons. Two important characteristics of QCD are color confinement and asymptotic freedom. The coupling strength of QCD (α_s) describes the relative strength of the strong force, and increases at low energy scales (Q^2) and large distances. If quarks from a hadron, such as a proton, are pulled apart and ejected from each other, it becomes energetically favorable to spawn quark/anti-quark pairs because of the magnitude of the strong force. A consequence is that under normal conditions quarks are confined in color neutral states, baryons and mesons. However as the energy scale increases, the relative strength of the strong force decreases, often referred to as the running of α_s, which can be seen in Figure 1.1. Only at asymptotically high energy scales do quarks and gluons interact weakly enough for the strong force to be calculable by perturbative QCD (pQCD). One avenue of studying QCD is thus going to very high energy scales to test pQCD calculations, which can be done by colliding matter, such as electron/positrons or protons, at relativistic energies and measuring the byproducts. When a high Q^2 collision occurs (collision with high energy transfer) in a proton-proton collision,

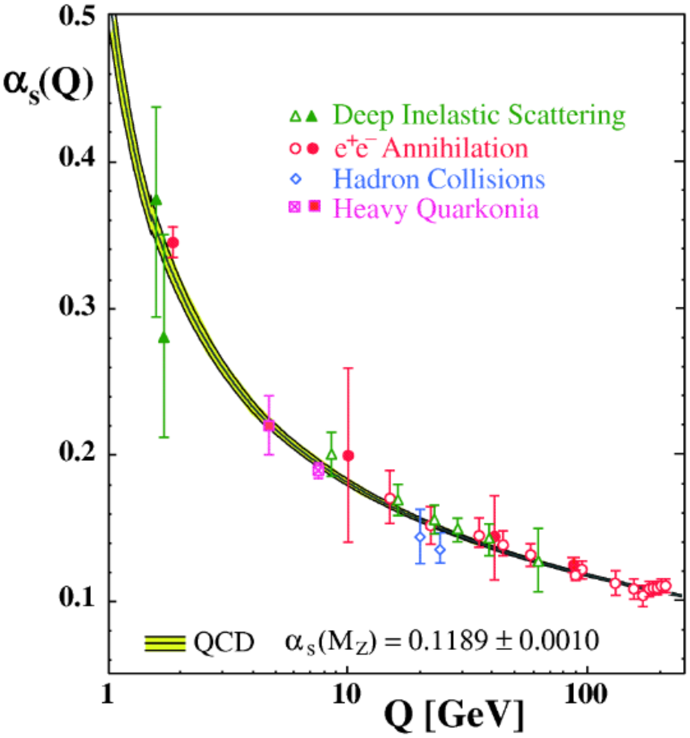


Figure 1.1: Running α_s as measured by multiple types of experiments, compared with QCD calculations [1]. The value of $\alpha_s(M_z)$ reported is the world average at Q = M_z, used as an established reference point for QCD calculations.

.

back-to-back partons are ejected with, to first order, equal and high transverse momentum (p_{T}). Since the strong coupling α_s becomes stronger as the distance between the color charges increases, it becomes energetically favorable to create quark and anti-quark pairs. This leads to a collimated spray of particles known as a jet. Because of color confinement, the partons which are byproducts of jets eventually hadronize into color neutral states that are measured by experimental detectors. Measuring these particles allows researchers to study the properties of QCD and to test the accuracy of models and pQCD calculations.

1.2 Quark-Gluon Plasma

While quarks and gluons are normally confined into hadronic matter, an area of active research is the novel state of matter in which the effective temperature and the density of gluons and quarks is so great that the quarks and gluons are effectively deconfined, known as the Quark-Gluon Plasma (QGP). The conditions for QGP formation can be described using a phase diagram of QCD, such as in Figure 1.2, in which the axes chosen are often the effective temperature and the baryon chemical potential μ_B. A QGP medium is not a commonly occurring phenomena that is easily measurable; the natural conditions for a QGP are expected to have been satisfied for a short time period briefly after the Big Bang. However the QGP is believed to be briefly achieved in the moments after colliding heavy-ions at relativistic speeds, creating a high density fireball of QCD matter with sufficiently low baryon-chemical potential and high enough temperature to allow the partons to be deconfined. The QGP medium from a heavy-ion collision forms quickly and becomes

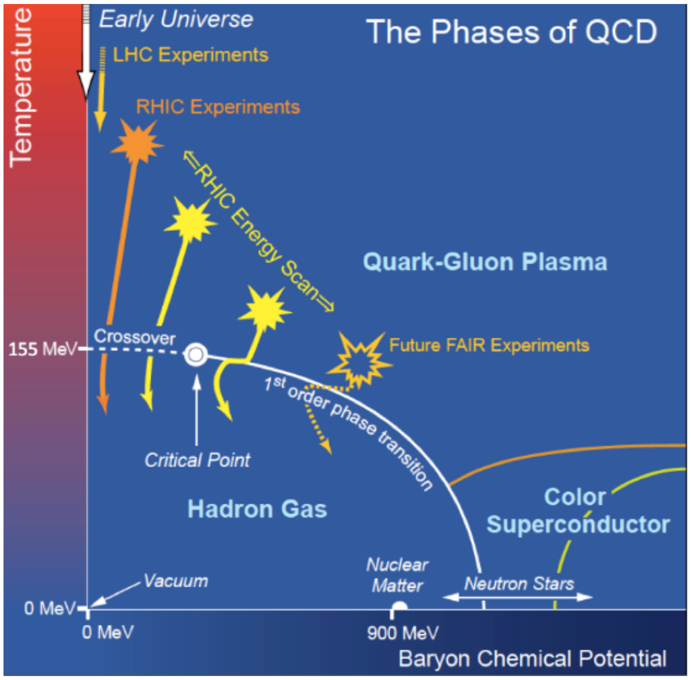


Figure 1.2: QCD phase diagram [2].

thermalized at about ~ 1 fm/c [12], and as it evolves in time it quickly falls in temperature until it transitions into a hadron gas with confined partons. The QCD phase diagram and an approximate evolution of heavy-ion collisions at multiple experiments can be seen in Figure 1.2. Collisions at the LHC exist at a smaller μ_B than other experiments due to its higher collision energies, which results in an increased amount of quark-antiquark pairs, lowering μ_B. At near zero μ_B the crossover from the QGP to a hadron gas is expected to be a smooth transition as predicted by lattice QCD [13]. However, at increased μ_B a first order phase transition and a possible critical point is conjectured to exist. The QGP is subject to ongoing research of heavy-ion collider facilities because it provides an opportunity to study a strongly-coupled QCD medium.

1.3 Heavy-ion Collisions

While we cannot directly measure the deconfined partons, experimentalists can measure remnant signals after the partons hadronize and reach their detectors. Heavy-ion collisions are complex systems, and the measured final state particles are the consequence of many complicated physical processes. A simplified outline of the stages of a heavy-ion collision can be seen in Figure 1.3. The colliding heavy ions are disc-shaped due to being highly Lorentz contracted as they collide at relativistic speeds. Moments after they collide, the collision region can reach extreme energy densities, and at this time the collision's fireball is in a pre-equilibrium state. After about 1 fm/c, the fireball expands and begins to enter thermal equilibrium, forming a QGP medium [14]. In this stage many partons are thermally produced. As time evolves and the QGP medium expands and cools, the medium enters into a cross-over phase where the QGP falls below the necessary temperature and density conditions to be maintained. Here the free partons begin to hadronize and form color neutral bound states. The time when all the particles have undergone hadronization and inelastic interactions stop is referred to as chemical freeze-out. After chemical freeze-out, the medium behaves as a hadron gas and many elastic collisions can still occur. As the medium continues to further expand, the elastic collisions cease and the collision enters kinetic freeze-out, which results in the momenta of the particles no longer changing. Following kinetic freeze-out, the particles then stream towards an experiment's detectors to be measured.

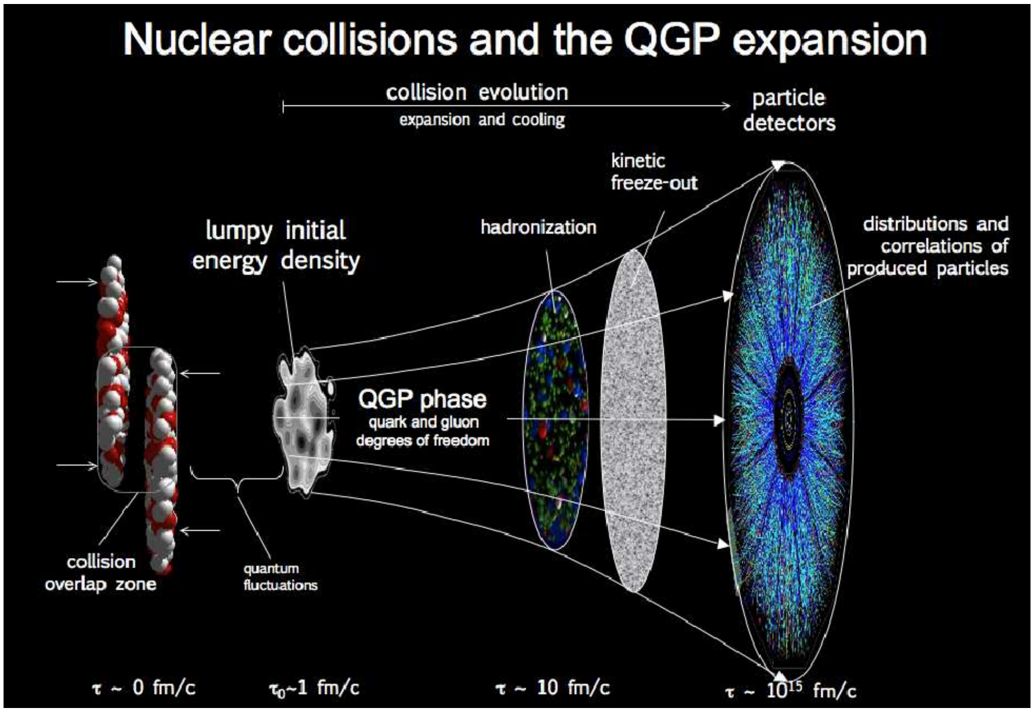


Figure 1.3: Stages of a heavy-ion collision [3].

1.3.1 Impact Parameter

Not all heavy-ion collisions are direct impacts of the colliding ions, they can collide with a range of impact parameters. Collisions with different impact parameters are expected to exhibit different event characteristics and sensitivities to different physical phenomena. Depending on the size of the overlap region, the amount of nucleons participating in the collision (N_{part}) and the amount of proton-proton like binary collisions (N_{coll}) can differ greatly. Collisions with a low impact parameter and large overlap region, referred to as "central" collisions, will have the most participating nucleons in the collision, creating the hottest and largest fireball. With decreasing impact parameter the size and initial temperature of the collision decreases, however the initial collision overlap geometry becomes more asymmetric, allowing for increased sensitivities to certain correlation measurements. Lastly, in glancing collisions with a large impact parameter, referred to as "peripheral" collisions, one expects significantly less sensitivity to signatures of the QGP.

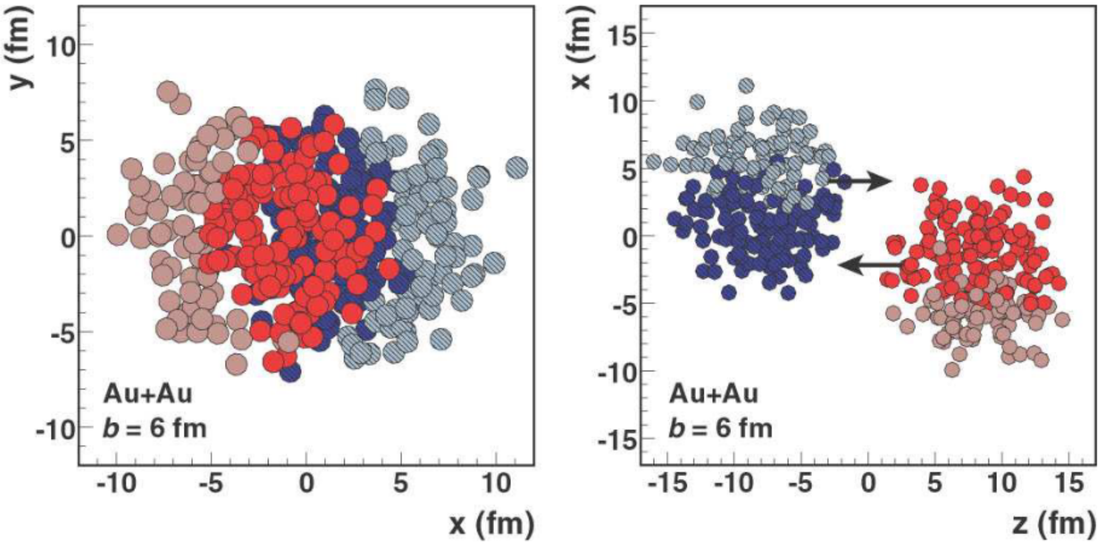


Figure 1.4: Glauber Monte Carlo event of two colliding gold nuclei with impact parameter b = 6 fm, shown in the transverse plane and along the beam direction [4].

However, the impact parameter of the collision isn't experimentally accessible and instead events are classified into centralities based on some form of measured event activity, usually charged particle multiplicity, which is correlated with how many nucleons participated in the collision. In order to relate measured event quantities back to important physical quantities such as N_{part}, N_{coll}, and the impact parameter, the Glauber model is often used [4]. The Glauber model is a simple model which treats the colliding nuclei as a series of independent binary nucleon-nucleon collisions if the nucleons geometrically overlap. An example of a Glauber Monte Carlo event can be seen in Figure 1.4, with the collision in the transverse plane in the left image and the same incoming nuclei along the beam direction on the right. In the figures each circle represents individual nucleons, with the darker colored circles corresponding to nucleons participating in the collision. It is important to note that positions of the nucleons within each colliding nuclei can fluctuate, which results in collisions with the same impact parameter having a variance in N_{part} and N_{coll}. Using many Glauber Monte Carlo events one can generate the same distribution used to define the centrality bins in data, and then extract the corresponding mean $\langle N_{\mathrm{part}} \rangle$, $\langle N_{\mathrm{coll}} \rangle$, and $\langle b \rangle$ for each bin. A mock example of using charged particle multiplicity to define centrality bins and their corresponding physical quantities extracted with the Glauber model is shown in Figure 1.5. Conventionally centrality is reported as a percentile, with smaller values of centrality relating to collisions with a smaller impact parameter and higher values of centrality having a larger impact parameter.

1.4 Signatures of the QGP

1.4.1 Collective Motion

One important signature that the QGP is formed in heavy-ion collisions comes from measurements at RHIC and the LHC that have shown that final state particles exhibit signs of collective motion that can be described by hydrodynamics. If there was no QGP medium formed and the colliding nucleons behaved as many independent collisions, then a mostly isotropic distribution of soft particles is to be expected. However it has been shown that correlations of low-p_{T} particles have anisotropic distributions in the plane transverse to the collision [15][16]. This final state particle anisotropy can be expressed as a Fourier decomposition in the form of equation 1.1, where ϕ is the angle in azimuth, ψ_{RP} is the

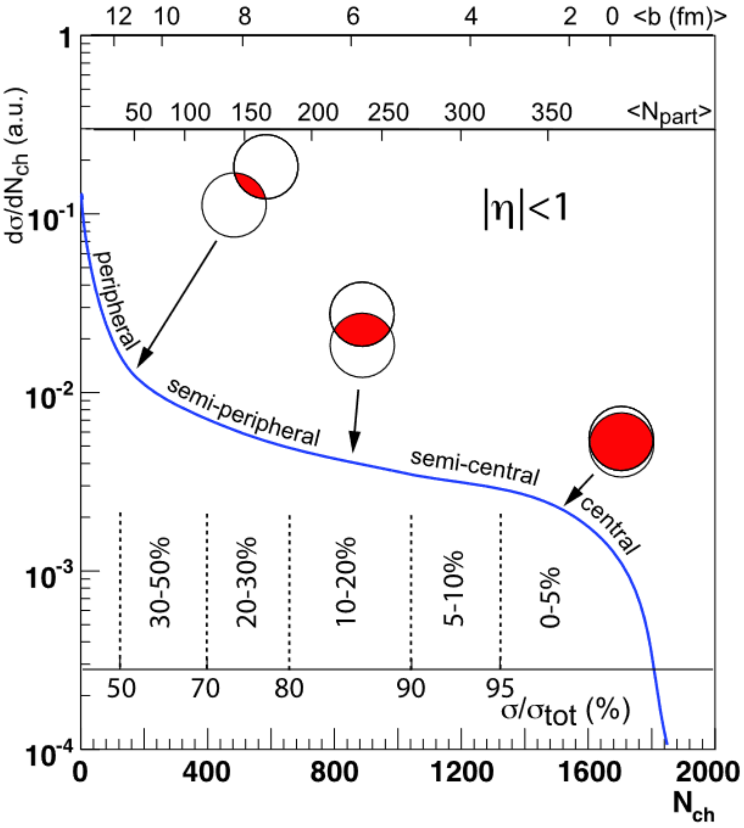


Figure 1.5: Example of using the Glauber model to define centralities using number of charged tracks [4].

reaction plane angle[1], and v_n is the n^{th} order flow coefficient.

$$\frac{\mathrm{d}N}{\mathrm{d}\phi} \propto \left[1 + \sum_{n=1}^{\infty} 2v_n \cos(n(\phi - \psi_{\mathrm{RP}}))\right] \tag{1.1}$$

These transverse anisotropies are signs of collective expansion of a strongly-coupled medium, and many measurements of the flow coefficients v_n have been performed, providing important information about the early expansion of the medium. The most dominant contribution is v_2, named elliptic flow. This is because the momentum anisotropies are

---

[1] The reaction plane in experiment is not directly accessible, and often instead of the reaction plane ψ_{RP} a different symmetry plane is used in equation 1.1, such as the participant plane or event plane

heavily driven by the pressure gradients of the initial geometry of the heavy-ion collision, with the most dominant contribution being elliptical in nature. A simple illustration of the initial geometry of a semi-central collision can be seen in Figure 1.6.

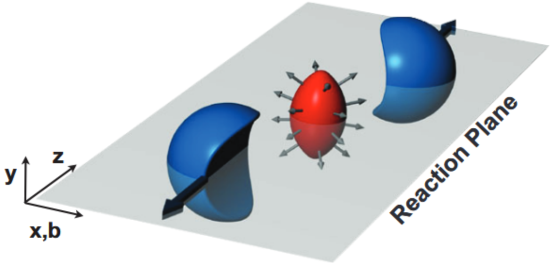


Figure 1.6: Initial collision geometry with a nonzero impact parameter [5].

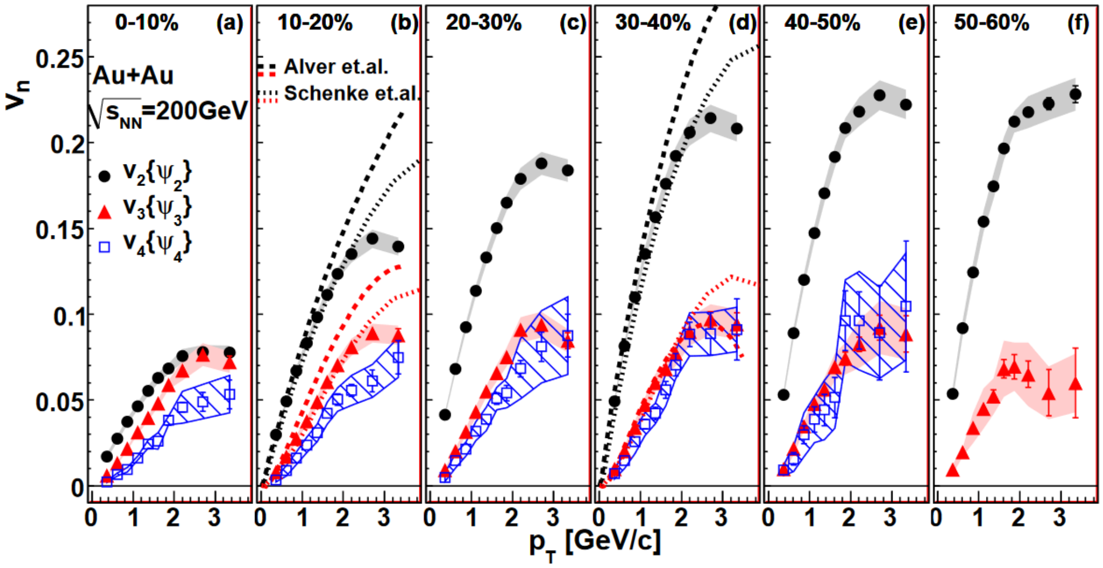


Figure 1.7: v_2, v_3, and v_4 measured by PHENIX using the event plane method as a function of p_T for multiple centrality bins in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. Model comparisons are shown for 10-20% and 30-40% centrality bins, using $\eta/s = 1/4\pi$ [6].

Flow measurements have been made by multiple experiments at both RHIC and the LHC. Figure 1.7 shows an example of a flow measurement made by PHENIX, where they measured v_2, v_3 and v_4 compared to a couple of hydrodynamic models. Note that as the collision becomes less central the v_2 increases, due to the increase in eccentricity of the collision geometry. Higher order v_n are also nonzero but do not scale as strongly with centrality, as they are in large part driven by fluctuations of the initial state. The ability of hydrodynamic models to reproduce the data, as seen in the figure, supports the hydrodynamic description of the collision's fireball. Using v_n from data and hydrodynamic models, one can extract hydrodynamic transport coefficients, such as the value of the shear viscosity to entropy density, η/s. The value of η/s has a predicted theoretical limit of $\eta/s \geq 1/4\pi$ from AdS/CFT [17], and the QGP is sometimes referred to as a "near perfect liquid" because it is a medium with an η/s value close to this lower bound [18].

1.4.2 Hard Probes

In relativistic heavy-ion collisions there are many binary nucleon-nucleon collisions, in which constituents of the nucleons in the nucleus can interact and scatter to produce a jet, like in the case of proton-proton collisions. In proton-proton collisions jets are well understood theoretically, with their cross section being able to be calculated in pQCD and having been shown to agree with data at both RHIC[7] and the LHC[19]. This can be seen in an inclusive jet cross section measurement in p+p collisions at STAR in Figure1.8.

In A+A collisions, a high Q^2 collision occurs at a time scale much shorter than the formation of the QGP, allowing for the scattered partons to interact with the strongly-coupled medium [20]. Because jets are well understood in p+p collisions and have an early formation time, they are excellent probes of the QGP [21]. As the scattered partons traverse the QGP, they interact with the strongly-coupled medium and lose energy, mostly via induced gluon radiation and recoiling off of partons in the medium. This loss of energy due to interactions with the medium is known as jet quenching. Jet quenching has been observed at RHIC[22] and the LHC[23] through multiple observables, most notably the

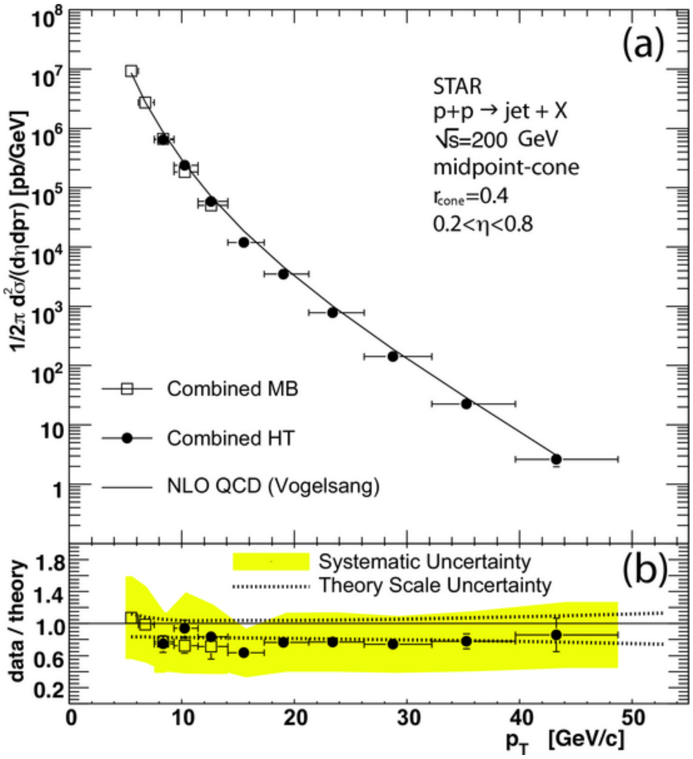


Figure 1.8: Inclusive jet cross section for p+p collisions at $\sqrt{s}$ = 200 GeV as a function of p_{T} compared to NLO QCD calculations [7]

single hadron R_{AA}, referred to as the nuclear modification factor,

$$R_{\mathrm{AA}} = \frac{dN^{\mathrm{AA}}/dp_{\mathrm{T}}}{T_{\mathrm{AA}}d\sigma^{pp}/dp_{\mathrm{T}}} \tag{1.2}$$

R_{AA} is the ratio of the p_T spectra of reconstructed hadrons in A+A collisions over cross section in inelastic p+p collisions scaled by T_{AA}, the nuclear overlap function, which is usually extracted from simulations using the Glauber model [4]. While jets are usually comprised of multiple final state particles, high-p_{T} particle spectra can serve as a proxy to jets, because particles reconstructed with high enough p_T almost exclusively originate from high Q^2 scatters. The single hadron R_{AA} has been measured by many experiments[8, 23, 24], and one such measurement made by STAR can be seen in Figure 1.9. An R_{AA} at unity would indicate that there is no modification of the expected amount

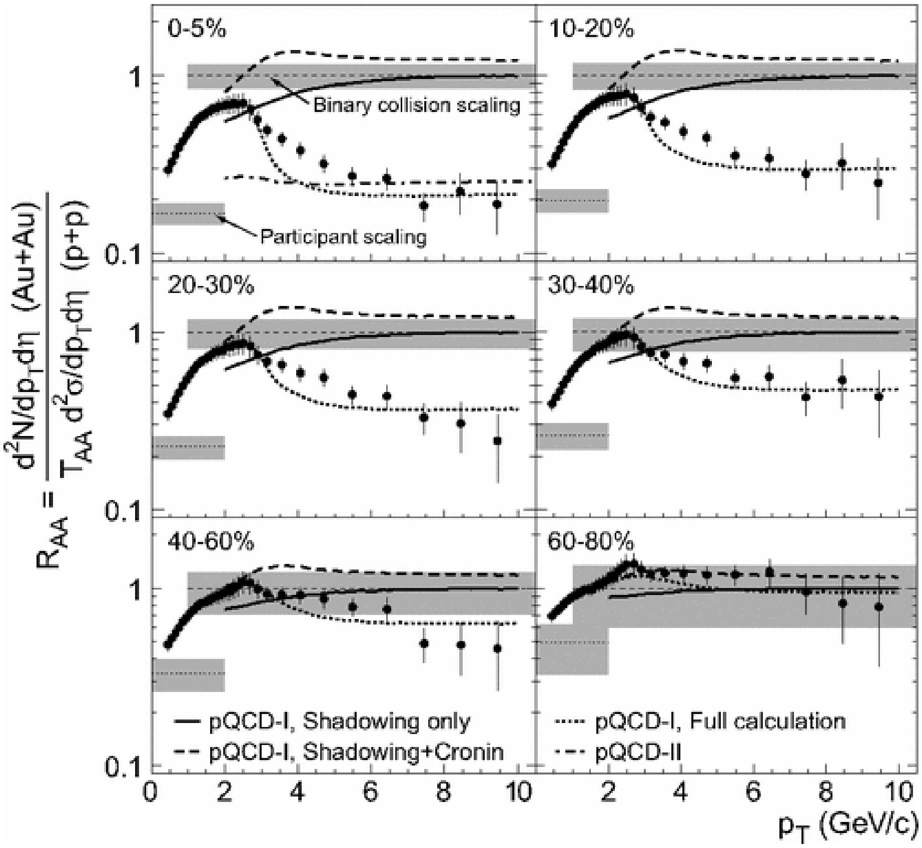


Figure 1.9: Single hadron R_{AA} measurement in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV by STAR for different centralities [8].

of yield in A+A collisions compared to p+p collisions. An R_{AA} above unity would signify enhancement, and below unity would imply a suppression. As can be seen in the STAR measurement, high-p_{T} hadrons were most suppressed in central collisions, where the QGP medium being formed is expected to be hottest. The high-p_{T} hadrons measured in the most peripheral bin, 60-80% central, agree with unity within uncertainties, which is expected as the overlap region of the collision is believed to be too small for a medium being formed capable of jet quenching. While the single hadron R_{AA} sitting below unity would imply suppression, one can posit a possible problem with the expected scaling. The accuracy of the scaling can be checked by looking at direct high-p_{T} photons, which originate from hard scatters but are not coupled to QCD, and thus should escape from any strongly-coupled medium with minimal modification. The direct photon R_{AA} has been

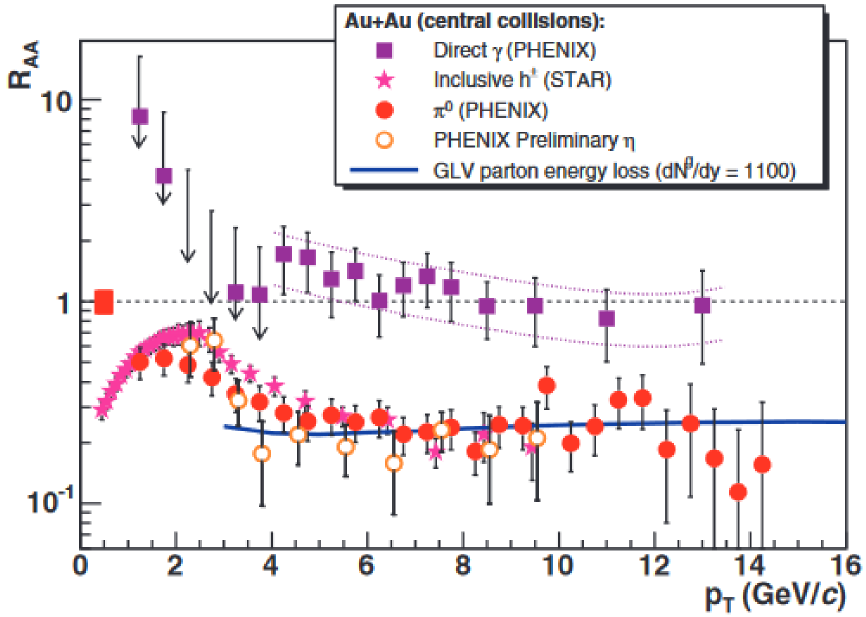


Figure 1.10: Direct photon R_{AA} measured by PHENIX compared to hadron R_{AA} in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV [9].

measured by multiple collaborations, and has been shown to agree with unity at high-p_{T} [9, 25]. One such measurement by PHENIX can be seen in Figure 1.10. The agreement with unity at high p_{T} for photons supports the accuracy of the scaling used in the single hadron R_{AA}, validating R_{AA} as evidence of jet quenching.

1.5 Jet Reconstruction

Measuring jets experimentally is not trivial, as experimenters are only able to measure final state particles found in detectors, which include particles which may not originate from jets. Ideally, experimentally reconstructed jets capture the kinematics of the hard scattered parton, but it is practically not the case. Particle momenta in hadron collider experiments are usually measured in the coordinates of the azimuthal angle ϕ, pseudorapidity η, and transverse momentum p_{T}. Pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the angle between the particle momentum and the beam axis. Using η is preferred over the simple geometric angle θ due to the distance metric between two particles i and j in these coordinates, $\Delta R_{i,j} = \sqrt{(\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2}$, being Lorentz invariant for massless particles.[2] Historically jet finding was done with cone finding algorithms, drawing cones in $\phi - \eta$ space around final state particles. However, many of the original cone algorithms were not easily calculable in pQCD, due to the measured jet sample being sensitive to collinear and soft emissions. In order to make meaningful comparisons to theory, more modern experimental jet definitions are chosen to satisfy infrared and collinear safety (IRC). This is to say that the observables are insensitive to soft gluon emissions and collinear splittings so their cross sections are calculable in pQCD. While there does exist cone algorithms which satisfy IRC safety, such as SISCone [26], most modern jet analyses utilize jet clustering algorithms implemented with the FastJet package [27]. While theoretically constituents of jets may extend to different distances in $\phi - \eta$ space, experimental definitions of jets require a fixed jet resolution parameter R, which limits the size of the reconstructed jet. There are three commonly used clustering algorithms, and all three can be described generally as follows:

1. For each pair of particles *i* and *j*, calculate the distance metric $d_{ij} = min(p_{T,i}^{2k}, p_{T,i}^{2k})\Delta R_{i,j}^2/R^2$ and for each particle calculate the beam distance $d_{iB} = p_{T,i}^{2k}$. The parameter k can be 1,-1 or 0 and selects the algorithms k_t, anti-k_t and Cambridge-Aachen (C/A) respectively.

2. Find the minimal value of the distance parameters d. If it is a pair of particles, combine them into a composite object[3]. If it is a single object then it is considered a clustered "jet" and is removed from the record.

3. Repeat steps 1 and 2 until all objects are removed from the event record and are

---

[2] Using rapidity y instead of η would make the distance metric invariant for particles with mass, but due to the difficulty of reconstructing every particle's mass in experiment, η is used instead.

[3] There exist several recombination schemes so this step has some nuance, but for many analyses this is simply the 4-vector addition of the two momentum vectors

considered clustered into jets.

The anti-k_t algorithm ($k = -1$) is often used in reconstructing jets from heavy-ion collisions as it is the most insensitive to the underlying event found in heavy-ion collisions [28]. The anti-k_t algorithm prioritizes clustering the high-p_T particles first, resulting in a clustering algorithm which is robust against the back reaction of the plentiful low p_T particles found in an A+A event. The product of anti-k_t clustering is mostly circular jets with an effective area of πR^2 in $\phi - \eta$ space, which is why the jet resolution parameter R is often labelled the jet radius. The Cambridge/Aachen (C/A) algorithm, resulting from k=0, is a purely geometric clustering scheme which clusters jets with an angular-ordered clustering history. While C/A clustering is not suitable for finding jets in an A+A environment, it is used for certain jet grooming/tagging techniques and for the calculation of some jet substructure observables. Finally, $k = 1$ is the k_t-algorithm, which preferably clusters soft low-p_T particles first. Like C/A clustering, the k_t-algorithm is sensitive to the soft background in A+A collisions and is not often used primarily for the purpose of jet reconstruction in heavy-ion analyses. However, the k_t scheme does have a use in heavy-ion analyses, being used for estimating the amount of soft background from non-jet processes (see Section 4.1).

1.6 Jet Substructure

Jet quenching is more than just energy loss, it can result in acoplanarity, the deflection of dijets away from being back-to-back in azimuth [29], and also the modification of jet substructure, the way the particles inside jets are distributed. Measuring modification of jet substructure in A+A collisions can provide experimentalists with information about how energy is lost, not simply how much. For example, the jet fragmentation function $D(z)$ measures the longitudinal momentum fraction $z \equiv \frac{p_\mathrm{T}}{p_\mathrm{T}^\mathrm{jet}} \cos(\Delta R)$ that each particle within a jet carries. By comparing the ratio of the A+A to the p+p distribution, $R_{D(z)}$, information regarding how the distribution of particle momenta within a jet changes due to jet quenching can be obtained. Such a measurement was performed by ATLAS, and the distribution $R_{D(z)}$ for 0-10% central events can be found in Figure 1.11. A depletion

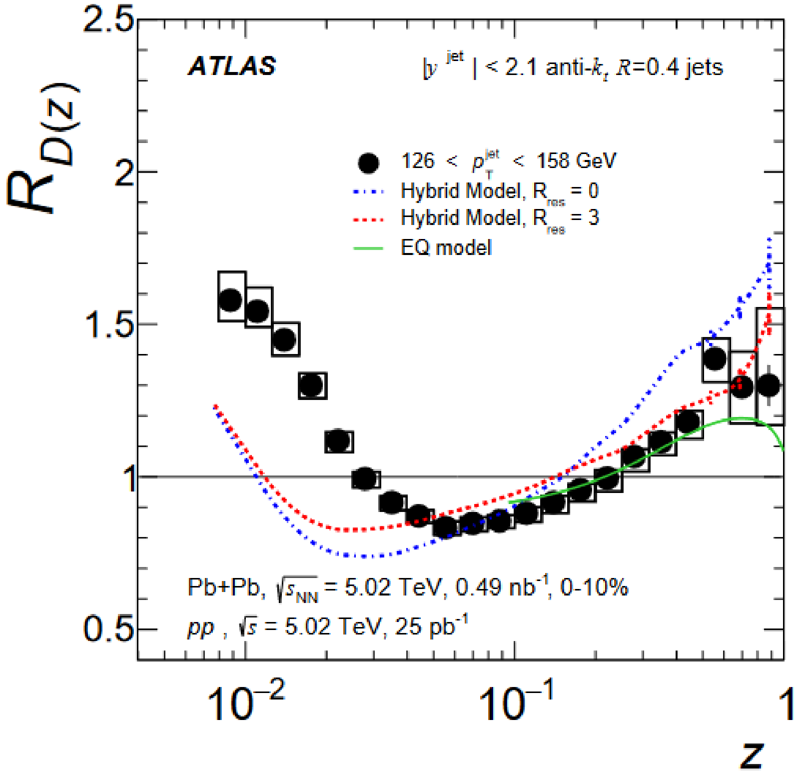


Figure 1.11: Measured $R_{D(z)}$ by ATLAS in Pb+Pb 0-10% central collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV compared with model calculations [10].

of particles with intermediate z and an enhancement of particles with low z is seen. This behavior is expected due to jet quenching effects, where the energy loss of intermediate partons is transferred to softer particles. The CMS collaboration made a similar measurement of the jet fragmentation function, and came to similar conclusions [30]. The high-z enhancement in the ATLAS measurement can be explained from the increase in the fraction of quark jets measured in Pb+Pb collisions within a given p_{T} interval [31]. Gluon jets are expected to be quenched more than quark jets, due to their higher color charge and wider fragmentation pattern [32]. The selected jet p_{T} bin in Pb+Pb collisions contains quenched jets, which due to differences in energy loss is expected to have a higher quark fraction than in p+p collisions. Quark jets are known to be more collimated and fragment harder, causing an enhancement at high z with an enhanced quark fraction. The high z enhancement was found to disappear when measuring the fragmentation function of Pb+Pb and p+p collisions with isolated photon-tagged jets [33]. The photon p_{T} does not interact with the QGP medium and thus selects jets with identical initial conditions in both systems. While the model calculations in Figure 1.11 capture some of the behavior of the distribution, there is still discrepancy which can be used to constrain models. The fragmentation function is just one of many jet substructure observables that can be studied to learn how jet quenching affects the distribution of energy within a jet.

Jet substructure measurements are constantly evolving and new observables are enabled by the advancements of more sophisticated experimental techniques, which allow more intricate jet observables to be measured. Such techniques include, but are not limited to, jet trimming [34], pruning [35], tagging [36], and grooming [37]. Jets are complicated objects which contain physics sensitive to a large range of momentum scales, and these experimental techniques can be used to isolate or reduce sensitivity to certain physics. For example, SoftDrop is a jet grooming technique which removes soft and wide-angled radiation from the jet, resulting in a jet with less sensitivity to non-perturbative effects such as hadronization [38]. The phase space of radiation which SoftDrop grooms away is regulated by two free parameters, z_{cut} and β. SoftDrop is a recursive algorithm that can be understood as follows.

1. Recluster the jet using the C/A algorithm, a purely geometric clustering scheme to generate an angular ordered clustering history.

2. Undo the last step of C/A clustering, obtaining two parent subjets with momenta $p_{T,1}$

and $p_{T,2}$. Check if the subjets pass the SoftDrop condition, such that $\frac{min(p_{\mathrm{T},1},p_{\mathrm{T},2})}{p_{\mathrm{T},1}+p_{\mathrm{T},2}} > z_{\mathrm{cut}}(\frac{\Delta R}{R})^{\beta}$, where ΔR is the distance in $\phi - \eta$ space between the two subjets, and R is the jet resolution parameter.

3. If the branches pass the SoftDrop condition, the jet which is a composite of the two subjets is considered groomed and the algorithm terminates. If the branches fail the condition, the subleading subjet is dropped and the leading subjet is passed back to Step 2. If the jet at this point only constitutes a single particle then the algorithm terminates.

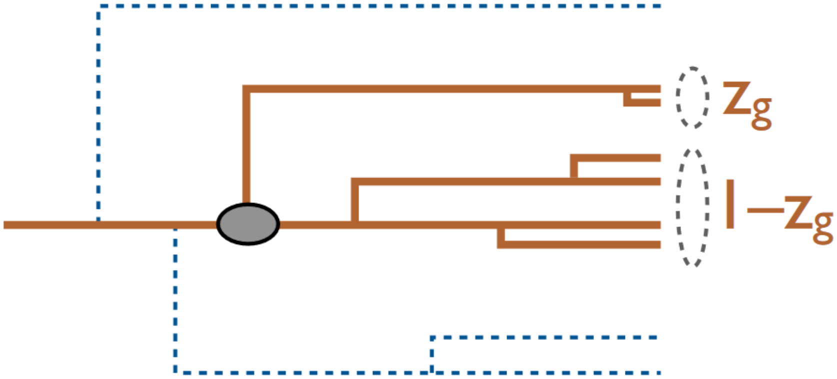


Figure 1.12: Illustration of SoftDrop grooming of an angular ordered clustering history. The dash lines represent branches of the clustering history groomed away for not passing the SoftDrop criterion.

After this grooming procedure, jet observables can be measured with an expected lower sensitivity to non-perturbative effects. Additionally the grooming procedure itself can be used to define jet observables which probe the phase space of jet physics.

Just like how jet definitions are chosen to satisfy IRC safety, jet observables also are usually chosen for satisfying IRC safety in order to be easily calculable in pQCD. However, IRC safety is not a strict requirement, as there exist IRC unsafe observables which can be calculated using certain pQCD techniques. The jet fragmentation function is one example, while it is collinearly unsafe it is still able to be predicted with pQCD. Another example is a class of observables that are known as Sudakov Safe observables [11] [39]. While IRC safe observables have finite cross sections order-by-order in pQCD, Sudakov safe observables are not well defined at any fixed perturbative order, but the singularities are regulated using an all-orders resummation. The observable of interest to this thesis, the shared groomed momentum fraction (z_{g}), belongs to this class of observables. The observable z_{g} is the measured momentum fraction of the groomed jet that the subleading subjet carries after SoftDrop grooming, $z_{\mathrm{g}}= \frac{min(p_{\mathrm{T},1},p_{\mathrm{T},2})}{p_{\mathrm{T},1}+p_{\mathrm{T},2}}$. An illustration can be seen in Figure 1.12. For differing selections of β in the SoftDrop grooming algorithm, z_{g} varies in regards to IRC safety. For the selection of $\beta = 0$, which removes the angular dependence of SoftDrop grooming, z_{g} is not IRC safe but is Sudakov Safe. For the rest of this thesis, the observable z_{g} is implied to correspond to this Sudakov Safe selection of SoftDrop parameters. The shared groomed momentum fraction is a physically interesting observable

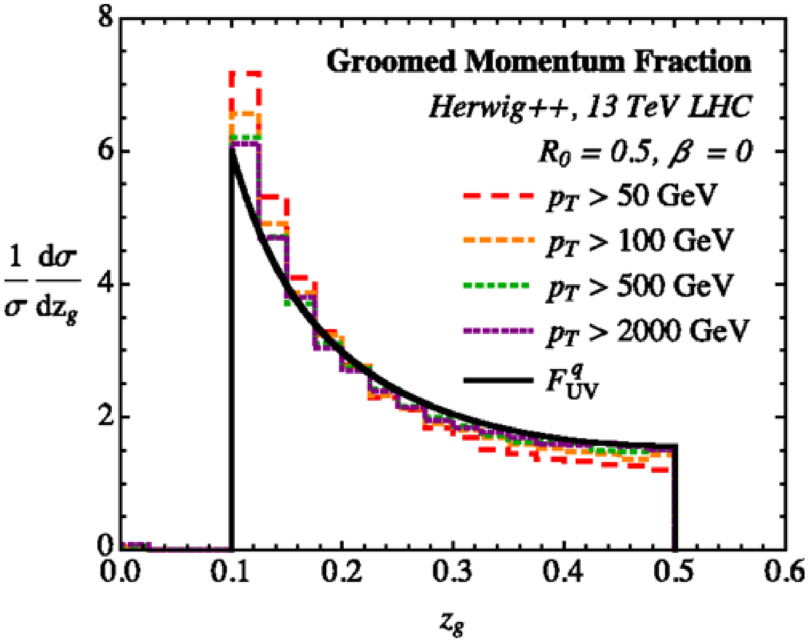


Figure 1.13: Distribution of z_{g}, using z_{cut} = 0.1 and β = 0, from using HERWIG++ simulated p+p data at $\sqrt{s_{\mathrm{NN}}} = 13$ TeV compared to the quark splitting function in the UV limit [11].

.

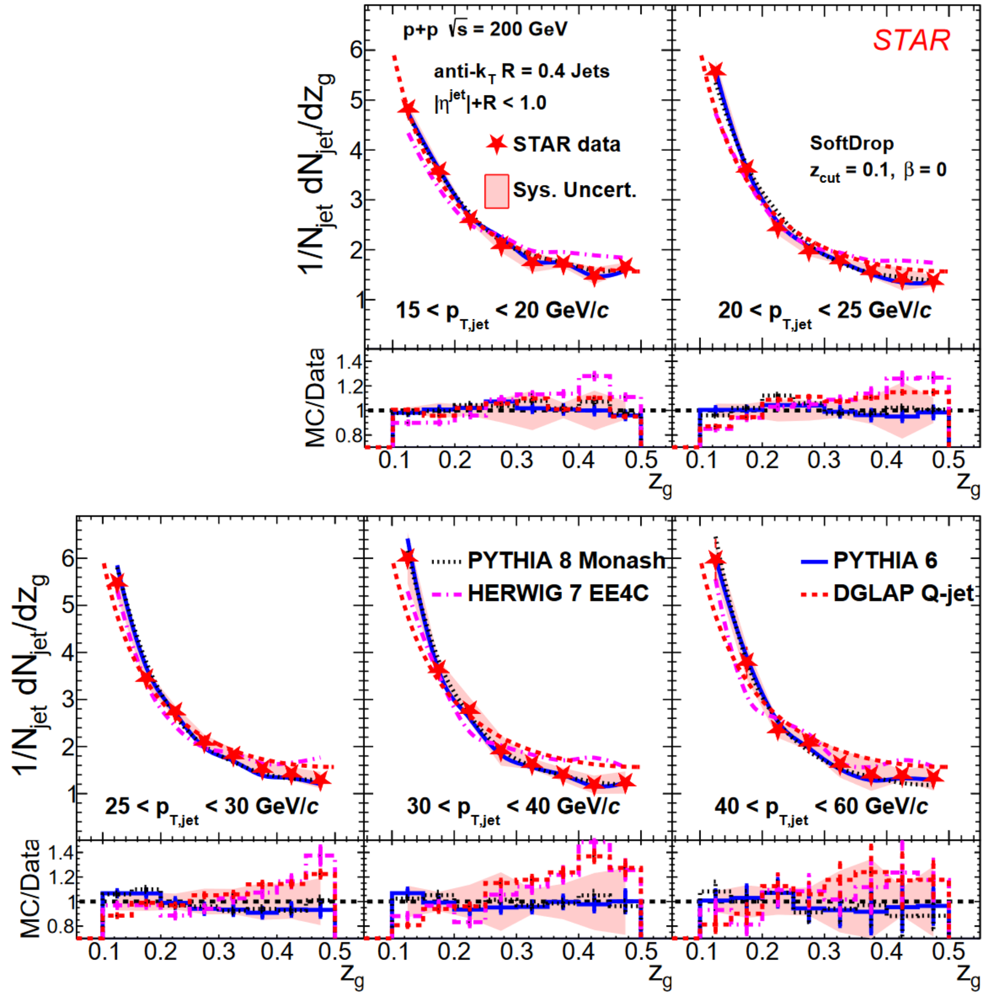


Figure 1.14: Measurement of z_g by STAR in p+p collisions at $\sqrt{s}$ = 200 GeV for different $p_\mathrm{T,jet}$ bins. Data is compared to MC generators and the DGLAP splitting function for a quark jet.

.

because it probes the first hard splitting of the jet. In vacuum it is closely related to the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) splitting functions [40, 41], which describe the fragmentation of jets in pQCD. This relation can be seen in Figure 1.13 with the quark splitting function compared to Monte Carlo event generated p+p data. The correspondence of z_g to the splitting functions makes it an attractive observable to study, and has been measured at both RHIC and LHC energies[42][43]. The STAR measurement of z_g in p+p collisions can be found in Figure 1.14, and agrees well with MC generators and the quark splitting function[4].

In A+A collisions, measuring z_{g} can be a valuable way to give insight into how jets lose energy within the QGP. By comparing the measured distributions in p+p to A+A collisions one can search for the modification of the DGLAP splitting functions due to the presence of a QGP medium. Another question that measuring z_{g} in A+A collisions can probe is whether the two prongs of the first jet splitting are resolved coherently as a single color charge or are incoherently resolved as two independent color charges. If the two prongs of the first split independently radiate gluons, the jets are expected to lose more energy and the subjets would be less balanced than if the two prongs are coherently interacting with the medium [44]. By comparing to models which implement different choices of coherent and incoherent energy loss, measuring z_{g} could potentially provide information on how jets lose energy as they traverse the QGP. Experimentally, measuring z_{g} in A+A collisions is not a trivial task, as the heavy-ion background complicates such a measurement.

The shared groomed momentum fraction in A+A collisions has been measured in multiple experiments, each with differing analysis details and kinematic ranges. The CMS collaboration measured z_{g} in Pb+Pb collisions for jets in p_{T} bins above 140 GeV/c, and found modification relative to p+p smeared data [45]. The ALICE collaboration also made a measurement of z_{g} in central Pb+Pb collisions, but with z_{cut} = 0.2 for $R = 0.2$ jets, using the tighter z_{cut} and small jet radius such that they could fully unfold and directly compare to p+p data [43]. They found that the z_{g} distribution was unmodified relative to p+p, however the groomed radius θ_{g}, the angular distance between the subjets of the groomed jet, was found to become more narrow in Pb+Pb compared to p+p. Also at RHIC energies, STAR made an early measurement of z_{g} [46] for a specific dijet sample, selecting jets with a hard core requirement (discussed in Section 4.1.1) and found no modification of z_{g}. Selecting jets with a hard core or high-p_{T} requirement is known to impose a surface

---

[4] Jets at RHIC kinematics are dominated by quark jets

bias, selecting jets which are biased towards losing less energy in the medium, reducing the sensitivity to measure jet quenching signals. The work of this thesis presents another measurement of z_g in Au+Au collisions at STAR, avoiding the use of a hard core requirement and instead using a novel technique to deal with contributions from the heavy-ion background, discussed in detail in Chapter 4.

Chapter 2

Experiment

2.1 RHIC and the STAR Detector

The Relativistic Heavy Ion Collider (RHIC)[47] is a versatile ion collider located at Brookhaven National Laboratory. RHIC has the ability to collide both protons and heavy ions over a large range of collision energies. RHIC is capable of colliding hadrons with energies as low as $\sqrt{s_{\mathrm{NN}}}$ =3 GeV in fixed target mode to as high as $\sqrt{s}$ = 500 GeV for p+p collisions, and colliding a wide range of particles species including p+p, p+Au, d+Au, He+Au, Au+Au, Cu+Cu, Cu+Au, U+U, Ru+Ru, and Zr+Zr. Such versatility and range makes RHIC extremely suitable for studying the QGP and for searching for critical point phenomena. When it began operations, there were four experiments located at four interaction points of the beams. The STAR experiment located at 6 o'clock, PHENIX at 8 o'clock, PHOBOS at 10 o'clock, and BRAHMS at 2 o'clock. At the time of this thesis, only the STAR experiment is operational, with the new sPHENIX detector soon to begin running.

The RHIC collider accelerates particles to top energies in a multi-step process. The particles to be accelerated, protons from the LINAC accelerator or heavy ions from the Electron Beam Ion Source (EBIS) first enter a Booster synchrotron, that then feeds into the Alternating Gradient Synchrotron (AGS), which boosts their energies up to 9.75 GeV/c

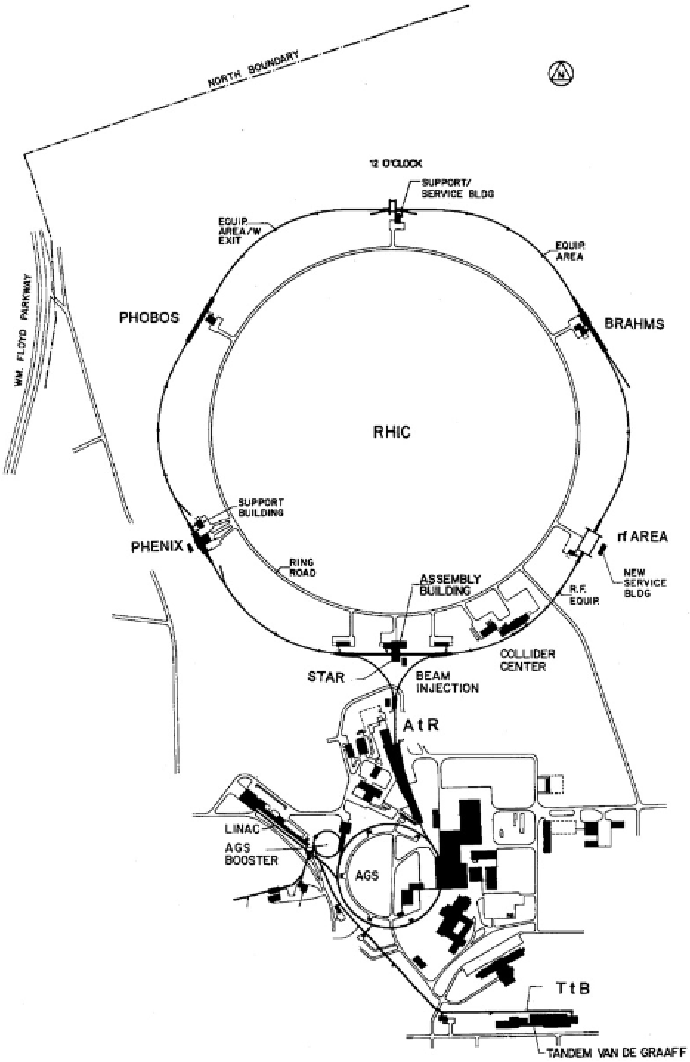


Figure 2.1: Schematic of the RHIC complex at BNL, before the existence of the EBIS.

per nucleon for gold ions. From the AGS, the particle beam enters into one of the two rings of the main RHIC collider. The beams of particles in the collider rings are then accelerated to the target energies, up to a momentum of 100 GeV/c per nucleon for gold ions. The two rings intersect at up to four different interaction points, where experimental detectors can be located. A schematic of the RHIC collider complex is found in Figure 2.1.

The STAR experiment is a general purpose detector, with many subsystems designed to take a wide range of measurements[48]. Over the years STAR has received many upgrades and changes in detector subsystems, and the STAR detector is constantly evolving

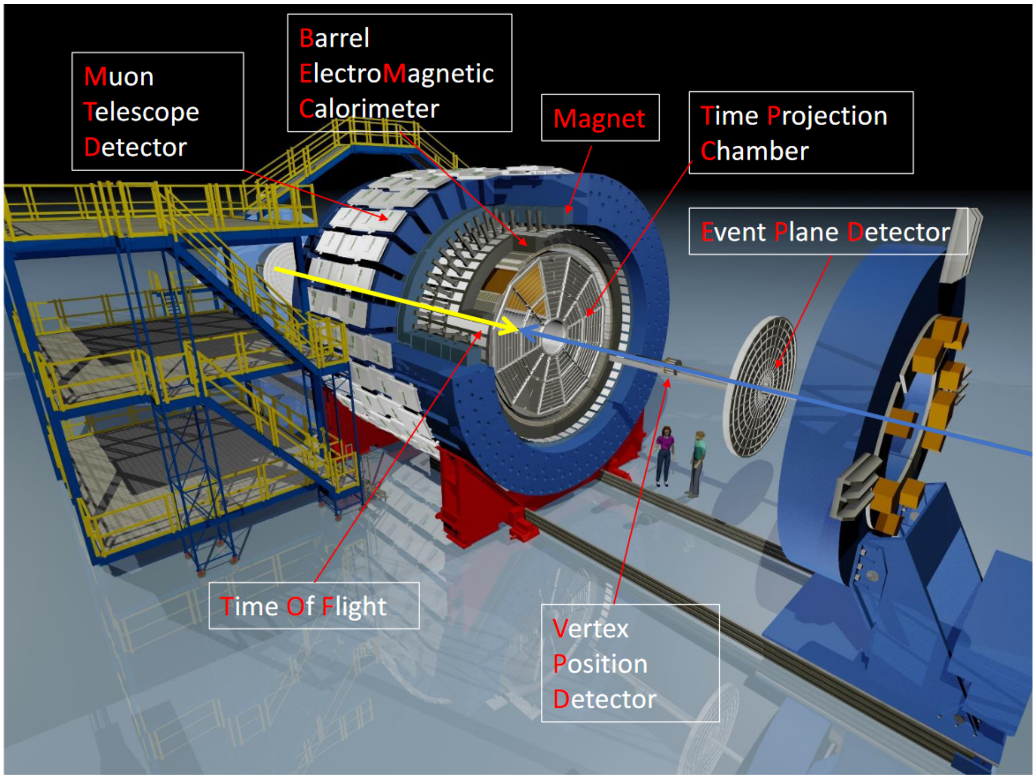


Figure 2.2: The STAR detector and some of its subsystems.

to make new and improved measurements. The STAR detector and some of its subsystems is shown in Figure 2.2. One of the key features of the STAR detector is the central barrel magnet[49], which provides a uniform magnetic field of 0.5 T along the beam direction for reconstructing charged particles in STAR's Time Project Chamber (TPC). The detectors most important to this thesis work are covered in the following subsections, however a brief mention of the detectors labelled in Figure 2.2 can be found in Table 2.1.

2.1.1 Time Projection Chamber

STAR's TPC is one of the main detectors of the experiment and is responsible for reconstructing charged particles at mid-rapidity[50]. The TPC is a cylindrical detector that sits inside STAR's large solenoidal magnet with an inner radius of 50 cm and an outer radius of 200 cm. The active volume of the TPC is filled with a gaseous mixture of 90% argon

Detector	Key Feature
TPC	Reconstructs charged tracks
BEMC	Reconstructs π^0 and γ
TOF	Used for particle identification of charged particles
EPD	Measures high-rapidity charged particles
MTD	Used to detect and trigger on muons, particularly from heavy-flavor decays
VPD	Reconstructs the collision vertex position along the beam-line

Table 2.1: The detectors labelled in Figure 2.2 and a brief note of their key features of importance.

and 10% methane, and is kept at 2 mbar above atmospheric pressure. At the center of the TPC is the Central Membrane, which separates the TPC in half and, with the two endcaps at each side of the detector, provides a uniform electric field of $\sim$ 135 V/cm across the active volume. The overall structure of the TPC is shown in Figure 2.3.

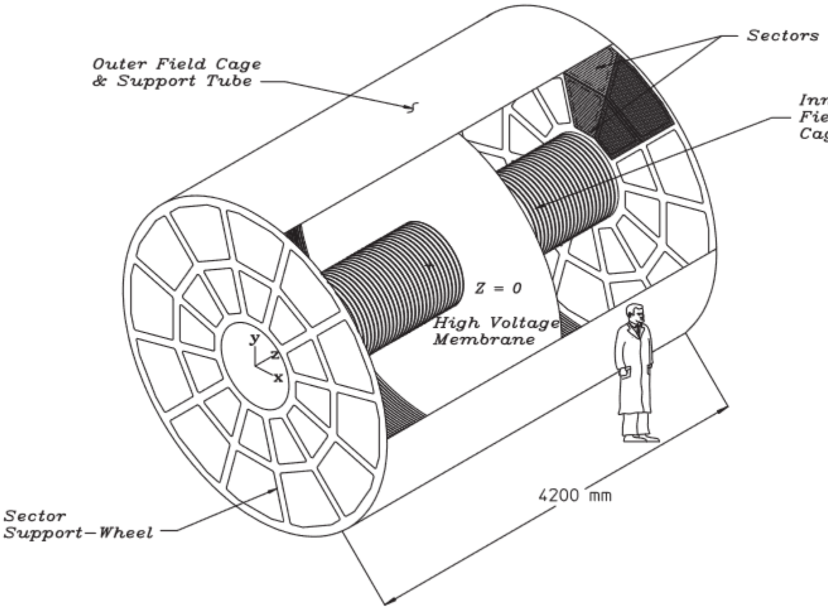


Figure 2.3: Design of the STAR TPC, showing the central membrane and the endcaps divided into sectors.

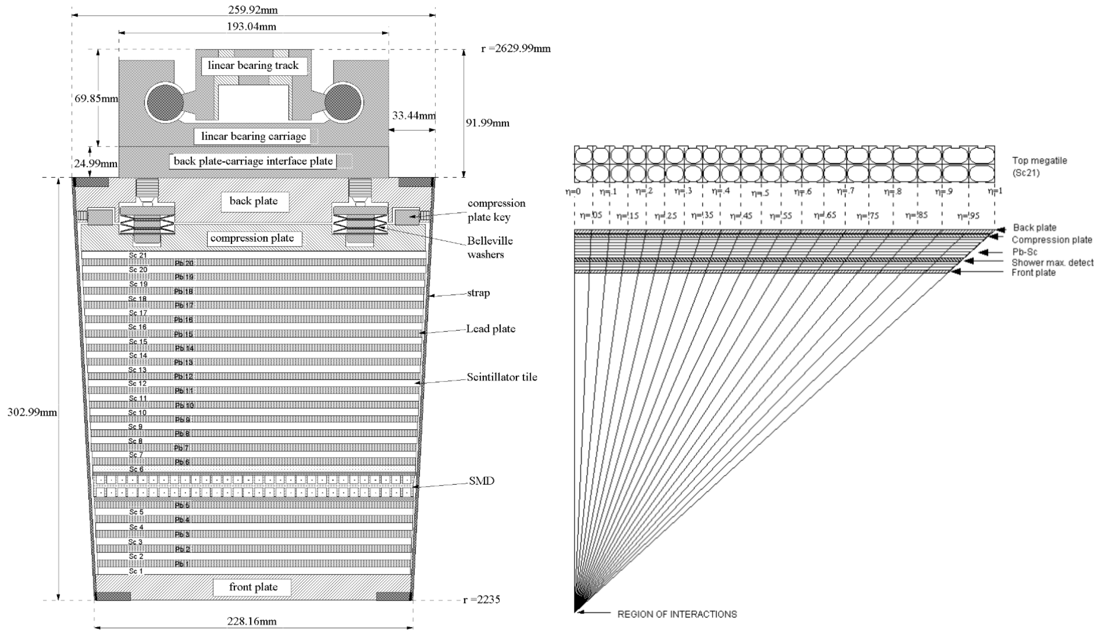
As charged particles pass through the TPC, they ionize the gas and showers of electrons drift in the electric field towards pad planes located at the endcaps of the TPC. The location of the collected charge on the pad planes is used to reconstruct the position of the hit points in the transverse plane, while the time it takes for the electrons to drift to the pad rows on the endcaps is used to reconstruct the hit point in the z-direction (along the beam direction). The gas mixture in the TPC is chosen because of its fast and stable drift velocity of ~ 5.45 $\mathrm{cm}/\mu\mathrm{s}$, allowing for the timing information to be used in this way. The TPC endcaps each consist of 12 sectors (for a total of 24), with each containing 45 pad rows, covering a range in pseudorapidity of $|\eta| < 1.0$ and full azimuthal coverage. If a high-p_{T} charged particle passes through the TPC, a maximum of 45 hit points can be determined, one for each of the pad rows. Using the hit points determined by the TPC, the trajectory of the charged particles are reconstructed by STAR's track reconstruction software, discussed in Section 3.1.1. It is important to note that the trajectory of charged particles passing through the TPC are helical, due to the presence of STAR's magnetic field. While the maximum number of hit points is 45, the maximum possible hit points for a charged particle can vary due to being at high rapidity or at low momentum such that their curvature does not allow them to escape the TPC.

Using the curvature of a particle's reconstructed helical track, one can measure the ratio of the particle's momentum to charge. From this ratio the particle's momentum and sign can be inferred. The TPC is capable of reconstructing particle momenta from as low as 100 MeV/c (limited by track finding efficiency) to as high as 30 GeV/c (limited by track momenta resolution). The TPC can also be used for particle identification. As charged particles traverse the active volume of the TPC, they lose energy and the measured energy loss, dE/dx, can be used to identify particles at low and intermediate p_{T}.

2.1.2 Barrel Electromagnetic Calorimeter

The Barrel Electromagnetic Calorimeter (BEMC)[51] is a detector which sits in the central barrel outside of the TPC and whose main purpose is to reconstruct certain neutral particles, namely photons and neutral pions, which are not reconstructed as tracks in the TPC. The BEMC consists of 120 calorimeter modules, each containing 40 individual towers in a 2x20 layout in ϕ and η. In total there are 4800 calorimeter towers each covering a square approximately 0.05x0.05 in $\Delta\phi$-$\Delta\eta$ space. In order for the towers to cover equal phase space even at different detector η's, the towers are projective towards the interaction point, which results in the towers differing in size. A tower at mid-rapidity ($\eta \sim 0$) covers an area approximately the size of about $10 \times 10\ \mathrm{cm}^2$ and becomes increasing larger as one goes to higher rapidity. The entire BEMC has a similar coverage to the TPC, with 2π azimuthal coverage in ϕ and $|\eta| < 1.0$. A schematic of a BEMC tower and module can be seen in Figure 2.4. The BEMC is a sampling calorimeter which is comprised of alternating layers of lead and plastic scintillating material. When photons or electrons traverse the lead material in the tower, they initiate a shower of electromagnetic radiation which, as it passes through the scintillating material, creates light signals that are then measured by photomultiplier tubes (PMTs). The measured ADC values are converted into energy values based on calibrations of the individual towers.

As an electromagnetic calorimeter, the BEMC is highly efficient at detecting photons and electrons, because they deposit a significant amount of their energy into the towers. However most hadrons are also capable of hadronically depositing energy into the BEMC, but to a lesser degree. Only about 30-40% of charged hadrons deposit a significant amount of their energy into the towers through nuclear interactions. Instead charged hadrons as they pass through the BEMC usually deposit small amounts of energy, about 20-30 MeV, due to electromagnetic ionization. This energy deposition is referred to as “Minimum



(a) (b)

Figure 2.4: Schematic of the BEMC. (a) Side view of a BEMC tower. (b) Side view of a BEMC module, showing the projective nature of the towers.

Ionizing Particles" (MIP), and is used to calibrate the towers' energy scale [52].

Within the lead and scintillating material layers, at about 5 radiation lengths deep where the electromagnetic shower is expected to be at its maximum density, exists the Shower Maximum Detector (SMD). The SMD can be seen within Figure 2.4a. While the BEMC is capable of measuring the energy of photons and π^0s, at high p_T the individual tower sizes are too small relative to the opening angle of the π^0 decay to resolve the difference between the two. The SMD consists of two wire planes in ϕ and η which provides the necessary spatial resolution for the distinction between high-p_T photons and π^0s.

2.1.3 Forward Trigger Detectors

The beam provided by RHIC is not constant but divided into bunches of ions or protons, with the collision rate of these bunches being $\sim$ 9.37 MHz, occurring approximately every 107 ns. Detectors such as the TPC are too slow to measure every bunch crossing, due to the drift time of electrons and the time it takes to process the hits into tracks ($\sim$ 100 ms). In order to take meaningful data, STAR uses multiple fast detectors as triggers to select collision events that are of interest. Some of these fast detectors used for data triggering are discussed below.

The Zero Degree Calorimeters (ZDCs) [53] are Cerenkov detectors located nearly along the beam direction to detect beam fragmentation neutrons. Charged remnants fall out of the acceptance of the ZDCs due to RHIC's magnetic field, leaving mainly neutrally-charged constituents to be detected by the ZDCs at $|\eta| > 6.0$. The acceptance of the ZDCs is shown in Figure 2.5. Coincidence of the east and west ZDCs signifies collision activity in the detector and is often a requirement for most triggered events. The ZDC coincidence rate is also used as a measure of the luminosity[54], being highly correlated with the collision rate.

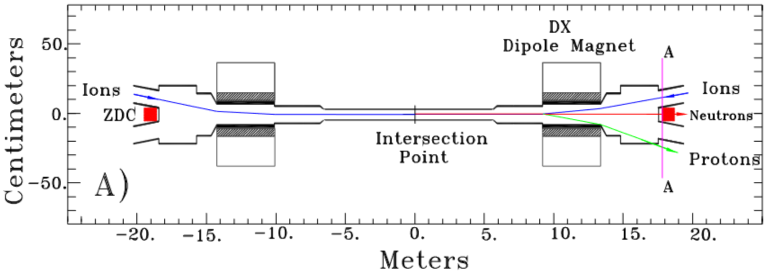


Figure 2.5: Schematic of the ZDCs and their acceptance.

.

The Vertex Position Detector (VPD) [55] is a pair of detectors located at high rapidity, $4.24 < |\eta| < 5.1$, at ±5.75 m in both beam directions. Each VPD detector is comprised of 19 modules of scintillating material which detect signals from mostly photons from π^0 decays and charged pions in Au+Au collisions. It is primarily used to measure the z-position of the interaction point of a collision using the timing difference between the signals of the east and west VPDs. This information, in tandem with track-level information reconstructed using the TPC, is used to determine the interaction vertex position.

The VPD's are highly efficient and are better at resolving the z-position of the collision vertex in heavy-ion collisions than in p+p collisions, due to the larger number of particles produced. The VPD is also used for triggering purposes, using the coincidence rate as a sign of event activity.

The Beam-Beam Counter, (BBC) [56] is a pair of scintillating detectors located around the beam pipe at ±3.75 m away from the interaction point in both east and west directions. Each BBC detector is comprised of an inner ring of 6 hexagonal scintillating tiles and an outer ring of 12 hexagonal scintillating tiles. The BBC covers a rapidity range of $2.2 < |\eta| < 5$. Like the ZDC, the BBC coincidence rate can be used as a measure of luminosity. The BBC was added as a STAR detector primarily for the purpose of being used as a trigger for proton collisions. While the BBC was used during the data taken relevant to the work of this thesis, the BBC has since been replaced by the newer Event Plane Detector (EPD) [57] in 2018.

2.1.4 DAQ and Triggering

The STAR trigger [58] system is constantly evolving depending on the subsystems available and is tailored to the physics goals of the run year. The information provided by the detectors to the STAR trigger system is broken up into four different levels. The Level 0 trigger is synced to the rate of bunch crossing in the beam (∼9.37 MHz) and utilizes low-level information recorded from the fastest detectors to initially select events of interest. Some of the data provided to Level 0 includes hits and ADC values in the ZDC, VPD, BBC, and the BEMC. The Level 0 trigger is only allotted ∼1.5 μs of processing time to select triggered events. An event selected by Level 0 then begins to process data from the slow detectors and passes the event onto higher levels of the trigger system. At these higher levels, the trigger system receives higher-level information from detectors, such as reconstructed BEMC and TPC data, which require time to process and are used to abort or accept the triggered events.

The most common triggered event is “Minimum Bias”, which is meant to select collision events with light trigger conditions to sample events as inclusively as possible. The minimum bias trigger definition changes depending on the collision energy and particle species, but for heavy-ion collisions generally requires coincidence of the ZDCs and VPD. Not only does the STAR trigger system select generic collision events, it can also select rarer events with specific event properties, such as central events or events with jets. Selecting on jets is achieved by utilizing the low-level information from the BEMC to select events in which a tower has sufficiently high ADC counts, resulting in what is called a high tower trigger. Alternatively, patches of BEMC tower ADC counts can be summed up and selected if the patch has a sufficient ADC sum, known as a “jet patch” trigger. In heavy-ion collisions high tower triggers are much more commonly used due to the large background from the high multiplicities of the collisions.

If a triggered event is accepted, then the data from all the relevant detectors is streamed to STAR’s DAQ system [59]. The STAR DAQ is responsible for receiving the data from all the detector subsystems, reducing their rate, and streaming them to the RHIC computing facility (RCF) [60] to be stored on tape using the High Performance Storage System (HPSS). The data at this stage consists of mostly raw detector information, which then is processed into higher-level data used by analyzers.

Due to the very high rate of minimum bias triggers compared to high tower triggers, combined with the limited bandwidth of the DAQ, not every minimum bias event is recorded. Such high rate triggers are pre-scaled, meaning they only sample a fraction of the events in order to leave bandwidth for other rarer triggers to be recorded.

Chapter 3

Data and Event Reconstruction

3.1 Run 14 Dataset

For Run14 in 2014, STAR took a large amount of Au+Au events at collision energy $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV over a large range of luminosity. This dataset served to replace STAR's Run11 Au+Au $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV dataset, by virtue of having significantly increased statistics. More planned Au+Au data at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV is expected to be taken in 2023 and 2025 which will eventually replace Run14 as the flagship Au+Au dataset. Run14 is notable for the presence of the Heavy Flavor Tracker (HFT), a silicon detector sitting inside the TPC with the purpose of resolving secondary vertices from short-lived particles [61]. While the HFT was an valuable tool for heavy flavor analyses, it was found that including the HFT detector in track reconstruction caused issues with non-uniformity of efficiency and momentum resolution for jet analyses. The Run14 data was then reproduced without the HFT in track reconstruction specifically for the jet physics working group at STAR for a few select triggers, most notably the minimum bias VPDMB30 trigger (total of $\sim$44 million events) and the BHT2*VPDMB30 trigger (total of $\sim$216 million events) used for this thesis, discussed in Section 3.1.3.

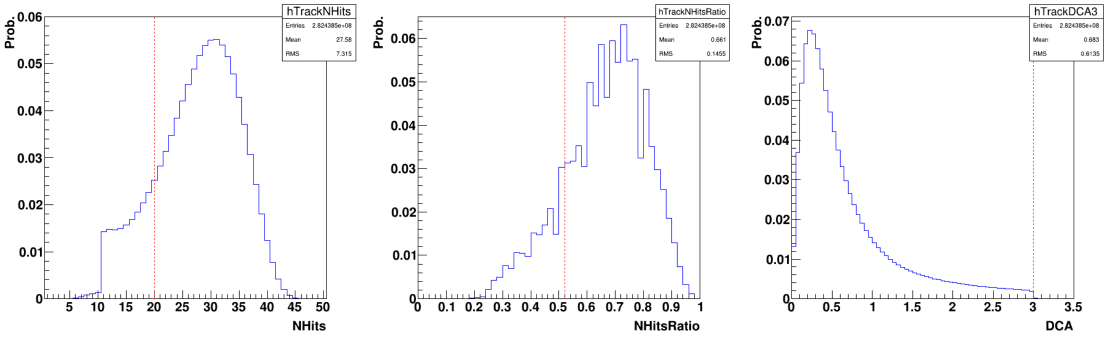
3.1.1 Particle Reconstruction

Primary Tracks

The raw data from the TPC is in the form of hit points in space. STAR's track reconstruction software uses the recorded hit points to form tracks, the helical trajectory of charged particles which are reconstructed in the TPC. Run14 was one of the last runs with its data mainly reconstructed using the Sti track reconstruction algorithm. Following Run14 most datasets utilized StiCA (Cellular Automaton) track reconstruction, which boasts higher track finding efficiency and improved timing performance [62]. The tracks reconstructed with just the TPC information alone are labelled "global tracks". Not all global tracks originate from the collision of interest that the detector triggered on; many of these tracks originate from "pileup", or other secondary collisions which can contaminate the event. Additionally, daughters of weakly decaying particles and secondary interactions with detector material and the beam pipe contribute to global tracks.

The global tracks reconstructed in the TPC are projected back towards the interaction point and used to reconstrcut the location of the collision's vertex. Global tracks which can be matched to fast detectors, such as the BEMC, are less likely to originate from pileup from out-of-bunch crossings and are weighted more heavily for determining the primary vertex. The VPD provides a high resolution measurement of the vertex position in the z-direction along the beamline, and the primary vertex is chosen to be the highest ranked vertex that falls within ± 3 cm of the z-vertex position as determined by the VPD. After determining the primary vertex, the global tracks with a sufficiently small distance of closest approach (DCA) in three dimensional space to the primary vertex are then refitted using the vertex as one additional fit point, yielding what is called a "primary track". This small DCA requirement greatly reduces the amount of pileup tracks, by eliminating the contribution from pileup collisions with vertices located at different positions in the detector. The DCA requirement also cuts global tracks from secondary decays which do not point back to the collision's vertex.

The initial set of primary tracks is not the final constituents used by analyzers; final analysis-dependent cuts are imposed to remove remaining pileup and poor quality tracks due to certain detector effects. One such quality cut is requiring the number of hit points, called "nhits", used in track reconstruction to be greater than 20 (out of a possible max of 45 hit points), to remove low quality tracks which will have poor momentum resolution. Another quality cut requires the ratio of hit points used to the possible amount of hit points (the possible amount can vary depending on the trajectory of the track) to be greater than 0.52. This cut removes contributions from track splitting, where one particle has its track split into two different tracks and is double counted. In some analyses a stricter DCA cut than 3 cm primary track definition is used, however for Run14 Au+Au the standard set of cuts, which were also used for this thesis, no stricter cut was imposed. It was found for Run14 that a tighter DCA cut would have reduced the track finding efficiency below desired levels. Tracks are also required to have $p_{\mathrm{T}} > 0.2$ GeV/c, due to very low tracking efficiency at smaller values of p_{T}. Likewise, tracks are required to have $p_{\mathrm{T}} < 30$ GeV/c due to the poor momentum resolution at high-p_{T} values. If a track was found to have p_{T}



(a) (b) (c)

Figure 3.1: Primary track QA from minimum bias Run14 low production 0-10% centraility. The red dashed line signifies the value of the cut used. Note that for the track DCA, the primary tracks after official reconstruction already have the less than 3 cm cut imposed.

> 30 GeV/c, instead of simply cutting the track, the event as a whole was determined to be bad and thrown out. Histograms of some of the primary track quality cuts can be seen in Figure 3.1.

Towers

The hits in the towers of the BEMC are used primarily to reconstruct neutral pions and photons which are not reconstructed as tracks in the TPC. While there are 4800 towers with full coverage in ϕ, not all towers are used in analysis as some towers are excluded as a "bad" tower. A tower can be considered bad for numerous reasons, and the determination of a list of bad towers can be analysis dependent. First, some towers have anomalous MIP peak distributions during calibration, and will have their status deemed bad during data production. Following these initially set bad towers, towers are added to the bad tower list if they have abnormally low firing rate (cold or dead towers), have abnormally high firing rate (hot towers), or have their mean energy $\langle \mathrm{E} \rangle$ deviate more than $\pm 3\sigma$ from the average. It should be noted that a tower that is deemed bad might not be bad for the entirety of the year's run. One such example are trips in a BEMC HV PMT box, which occur during data taking and can result in large strips of towers (a total of 80 towers per PMT box) all being temporarily cold for multiple runs of data. By adding these runs to the bad run list (list of data taking runs considered bad, see Section 3.1.3), one can recover the affected towers and have a smaller bad tower list. For Run14, it was decided that two standardized bad tower lists be used, one minimizing the amount of bad towers (total of 403 bad towers) by having a more restrictive bad run list and one maximizing the amount of good runs by having a more restricting bad tower list (total of 812 bad towers). For this thesis work the more restrictive bad tower list was used to increase the amount of statistics available. This choice is enabled by only using charged tracks for jet reconstruction in this analysis instead of full charged and neutral constituent jet reconstruction, in which the additional

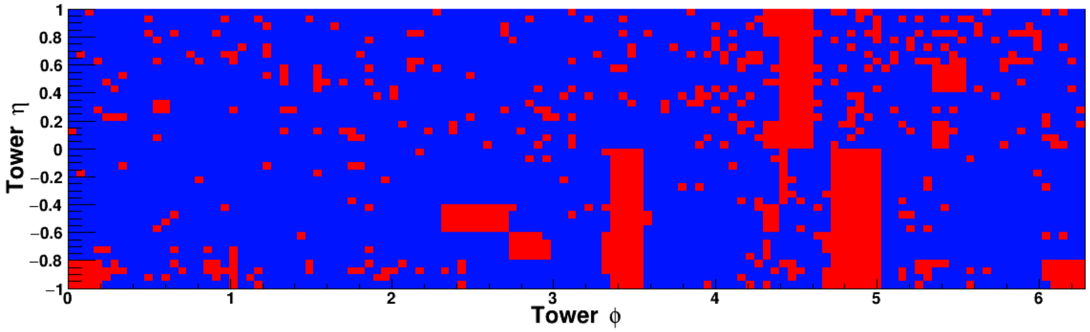


Figure 3.2: Tower status map in ϕ and η used for Run14. Red entries are towers deemed "Bad" and blue entries are towers used for analysis in this thesis.

gaps in the reconstruction of neutral constituents would be more problematic. A tower status map in ϕ and η[1] for the chosen restrictive bad tower list can be seen in Figure 3.2. Entries which are blue correspond to "good" towers and entries which are red correspond to "bad" towers. The long strips of bad towers seen in the status map are from problematic HV PMT boxes during some runs of data taking.

While the main purpose of the BEMC is to detect photons and neutral pions, hadrons can also hadronically deposit a fraction of their energy into the towers they pass through. This deposition can cause double counting when including both information from the BEMC and the TPC in jet analyses. To account for such an effect, the helical tracks from the TPC are projected to the surface of the BEMC and the energies of the tracks are subtracted from the towers the tracks points to. While hadrons usually only deposit a fraction of their energy into the towers, 100% of their energy is subtracted which improves the jet energy resolution for unfolding of jets using both track and tower information. In this thesis work, the towers aren't used in jet reconstruction and are only used to select on high-p_{T} trigger objects, in which the hadronic subtraction's main purpose is to remove the contribution of charged tracks to the high energy tower signals.

The energy and ADC information for hits in the BEMC are stored based on the ID of

---

[1] The location of a tower in ϕ and η is dependent on the vertex position in the event. Shown in the Figure are the tower locations for an event with $\mathrm{V_z} \sim 0$ cm.

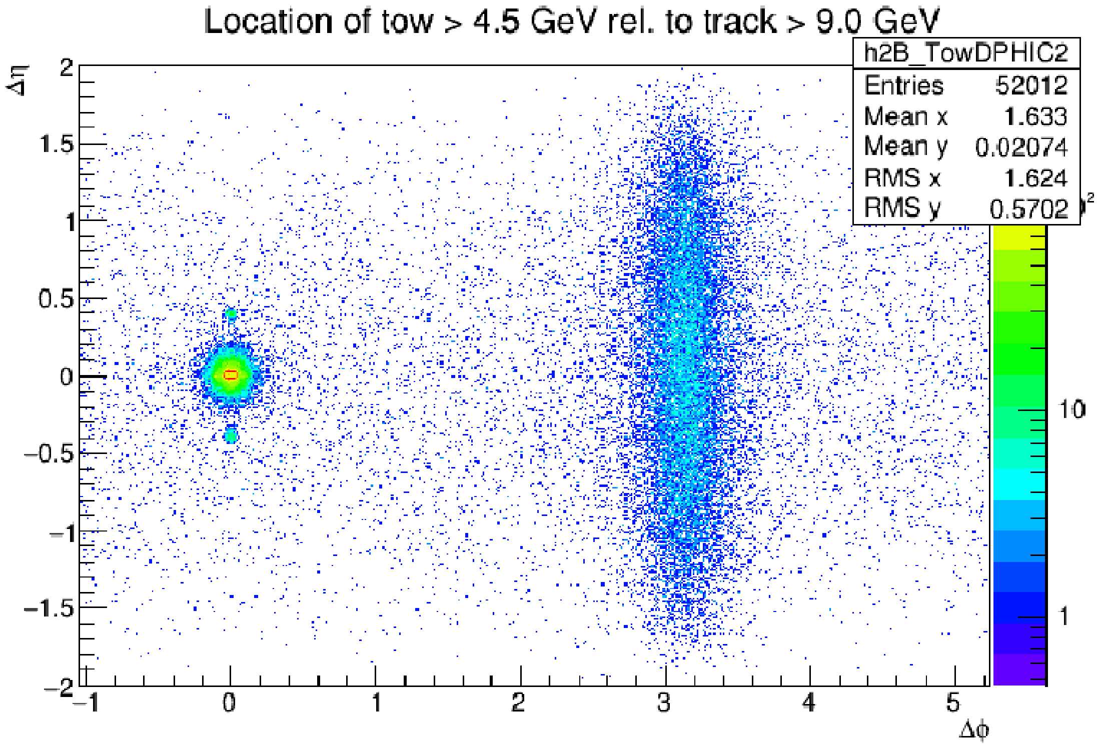
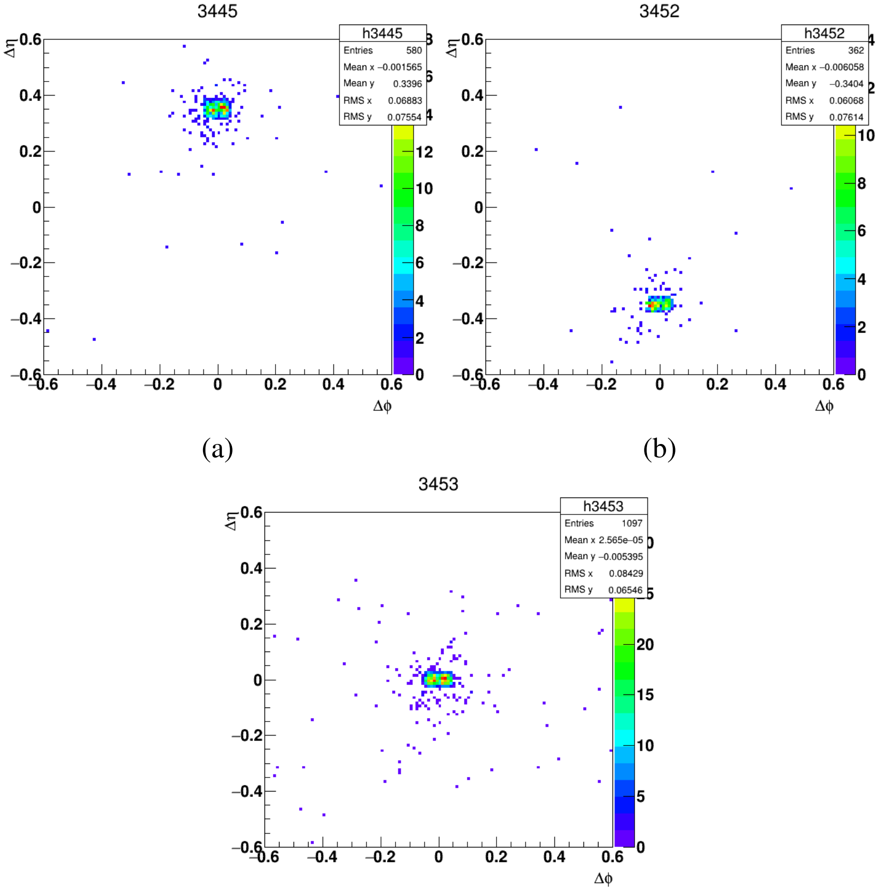


Figure 3.3: Correlation of towers with $E_{\mathrm{T}} > 4.5$ GeV (after hadronic correction) and tracks with $p_{\mathrm{T}}> 9.0$ GeV/c. Hot spots at $\Delta\phi = 0$ and $\Delta\eta = \pm 0.4$ correspond to swapped towerIDs, resulting in their energy not being hadronically corrected.

the towers (towerID), and analyzers convert the towerID into position in ϕ-η space based on the mapping of the towerIDs and the position of the event vertex. The towerIDs as seen by the analyzer, however, are already corrected for known swapped towers, towers known to not be located at their expected location due to reasons such as miswiring and incorrect PMT assignments during assembly. Because swapped towers can be corrected at the data reconstruction level, this effect is usually invisible to analyzers and causes no issues. However for Run14, during this thesis work additional unknown and uncorrected tower swaps were found. This was initially discovered by looking at correlations between high-p_{T} tracks and high-E_{T} towers as seen in Figure 3.3. The normal features of the near-side jet peak and the away-side jet ridge can be seen, but additional hot spots at $\Delta\phi = 0$ and $\Delta\eta = \pm 0.4$ are present. These hot spots result from towers reconstructed at locations in software different to their locations in reality. By looking at the towerIDs of towers

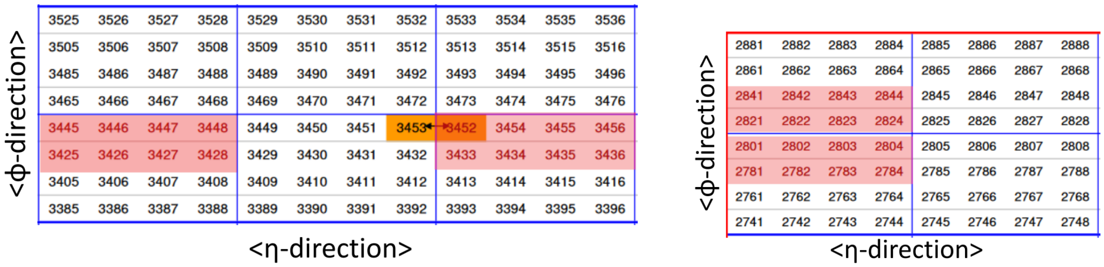


(a) (b)

(c)

Figure 3.4: Example of using hadronic deposition high-p_{T} tracks to create a projection of tower location for specific towerIDs. Plotted is the relative locations of tracks with $p_{\mathrm{T}} > 6$ GeV/c when the selected tower registers $E_{\mathrm{T}} > 3$ GeV. (a) Tower 3445 which was swapped with (b) tower 3452, separated by $\eta \sim 0.4$. (c) Shows an example of a non-swapped tower, being correctly projected at (0,0).

contributing to these specific hot spots, it was found that they were caused by 8 towers being swapped with another 8 towers. To further investigate, for each suspected tower affected, when they registered $E_{\mathrm{T}} > 3.0$ GeV, the relative location in ϕ-η of any tracks nearby with $p_{\mathrm{T}} > 6$ GeV/c was recorded. This effectively takes advantage of hadronic deposition of high-p_{T} tracks to create a projection of the tower location, as seen in Figure 3.4. The selected track p_{T} was chosen such that the tracks are relatively straight, as seen in the figure, where the small difference in the peaks in $\Delta\phi$ are due to the different sign of the curvature of positive and negative tracks. Interestingly, while performing the detailed search into the towers swapped in η, an additional 8 pairs of towers were found to be swapped in ϕ at a distance of about 0.1 rad. The effect of these additionally swapped towers is not clearly visible in Figure 3.3, and they were only swapped for half of Run14, beginning on run day 106 (April 16, 2014) and lasted until the end of the run. The total of 32 swapped towers can be seen in Figure 3.5. The nature of both found swaps being towers in a 4 by 2 grid in ϕ-η points to a cabling error resulting in swapping the inputs to the PMT boxes, with the swap of the towers in ϕ occurring during an a no-beam access day for maintenance during Run14. Swapping the towers at the event reconstruction level would require a total reproduction of the Run14 data, so instead the swapped towers were additionally added to the bad tower list.



3525	3526	3527	3528	3529	3530	3531	3532	3533	3534	3535	3536
3505	3506	3507	3508	3509	3510	3511	3512	3513	3514	3515	3516
3485	3486	3487	3488	3489	3490	3491	3492	3493	3494	3495	3496
3465	3466	3467	3468	3469	3470	3471	3472	3473	3474	3475	3476
3445	3446	3447	3448	3449	3450	3451	3453	3452	3454	3455	3456
3425	3426	3427	3428	3429	3430	3431	3432	3433	3434	3435	3436
3405	3406	3407	3408	3409	3410	3411	3412	3413	3414	3415	3416
3385	3386	3387	3388	3389	3390	3391	3392	3393	3394	3395	3396

(a)

2881	2882	2883	2884	2885	2886	2887	2888
2861	2862	2863	2864	2865	2866	2867	2868
2841	2842	2843	2844	2845	2846	2847	2848
2821	2822	2823	2824	2825	2826	2827	2828
2801	2802	2803	2804	2805	2806	2807	2808
2781	2782	2783	2784	2785	2786	2787	2788
2761	2762	2763	2764	2765	2766	2767	2768
2741	2742	2743	2744	2745	2746	2747	2748

(b)

Figure 3.5: Snippets of the TowerID map showing the towers which are affected by swapped tower IDs highlighted in red. (a) Towers swapped in η. (b) Towers swapped in ϕ. The orange highlighted towers 3453 and 3452 correspond to towers with already known swapped wiring before the additional tower swap.

3.1.2 Centrality

For the Run14 data set, the number of primary tracks found within the region $|\eta| < 0.5$, named the reference multiplicity ($refmult$), is used to determine the centrality event classes, as described in section 1.3.1. However the raw reference multiplicity alone is not sufficient for the final centrality determination, as the $refmult$ of events is not only a function of the impact parameter but also is sensitive to event properties such as luminosity and the vertex position in the z-direction (V_z). As the luminosity increases, the in-time pileup also increases which enhances the yield of global tracks within the TPC. However due to the increase in amount of activity in the TPC, the quality of tracks decreases and less primary tracks are found, resulting in a decrease in $refmult$. This decrease in primary track finding efficiency is referred to as being caused by "occupancy effects". The z-vertex position affects $refmult$ due to the small change of the η acceptance and the non-uniformity of the TPC. To account for these effects, the $refmult$ is corrected for these dependencies such that the corrected $refmult$ distribution is independent of luminosity and vertex position. The corrected $refmult$ is then used to determine the centrality of any given event. The work in this thesis focuses on the centrality bins 0-10% and 10-20%, which corresponds to a corrected $refmult$ of 364+ and 257-364 respectively.

3.1.3 Event and Triggers

For this thesis work, two types of triggered events were used. First is the STAR VPDMB30 minbias trigger, which requires coincidence of hits in both east and west ZDCs, and for the vertex position determined by the VPD to fall within $\pm$30cm in the z-direction. As mentioned in Section 2.1.4, the purpose of the minimum-bias trigger is to select Au+Au collisions as inclusively as possible, resulting in events that likely do not contain high-Q^2 interactions. These minimum bias events are used primarily for creating mixed events, which is described in Section 4.2.1. The VPDMB30 trigger is not the only minimum-bias trigger for Run14, but it used in this study due to it being the complementary trigger to the BHT2*VPDMB30 trigger. The BHT2*VPDMB30 trigger has the same requirements as the minimum bias trigger, but imposes the additional requirement of having a tower in the BEMC with sufficiently high ADC value, corresponding to approximately $E_{\mathrm{T}} \sim$

5.4 GeV. Figure 3.6 shows a comparison of the tower E_T distribution for both VPDMB30 and BHT2*VPDMB30 events in 0-10% centrality, with the effect of the high tower trigger clearly visible. This high tower trigger selects rare events with a high-Q^2 scatter which are of interest for measuring jets in the analysis. There exists a BHT3 trigger for Run14, with a higher tower ADC threshold which would be ideal for the analysis work of this thesis, however it does not include the same VPDMB30 requirement that the corresponding minimum bias triggered events have. The z-vertex position of BHT3 events range from $-100 < V_z < 100$ cm, and without a matching minimum bias event sample for such a large range in z-vertex the trigger is unsuitable for the work in this thesis.

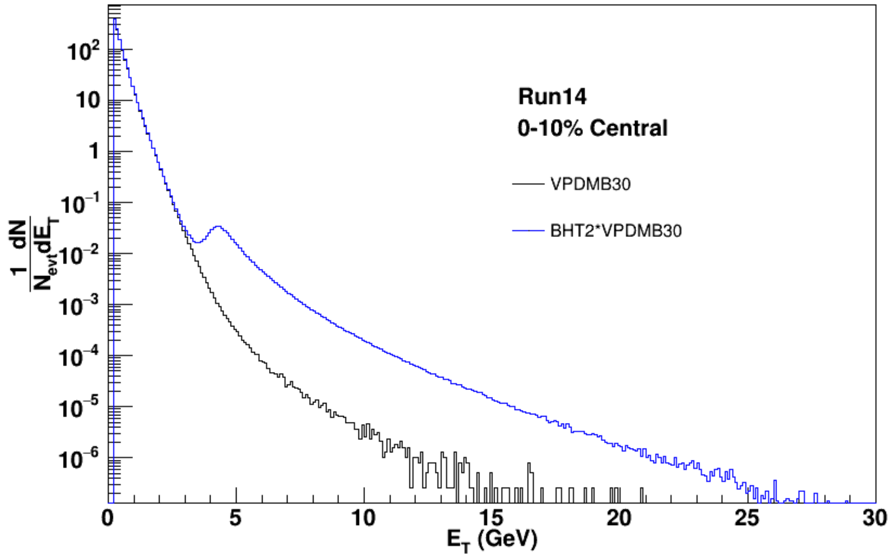
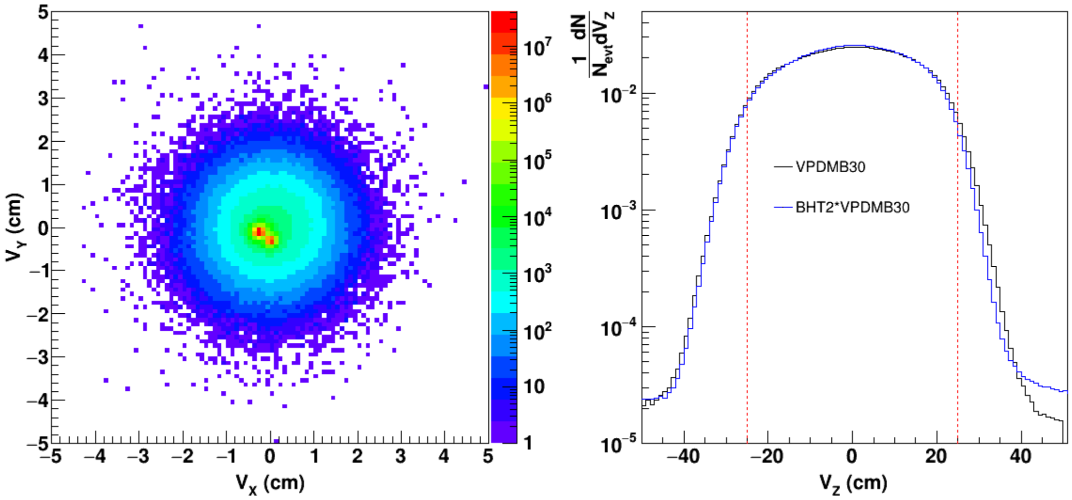


Figure 3.6: Tower E_T spectra for both VPDMB30 and BHT2*VPDMB30 0-10% central events. The effect due to the HT2 trigger turn on at E_T ~5.4 GeV can be seen.

Not all runs of data are of good quality and make it into the analysis. Some runs have abnormal event-qualities and are deemed to be unsuitable for analysis. During data taking, if runs are known to be bad due to trips in detectors or were spotted to have poor QA information, they are explicitly marked as bad and are not produced for analysis. However sometimes runs slip by and make it into data production, and the analyzer has to clean the run list. A bad run list is constructed individually for each trigger by looking at event-level quantities. The criteria used for the standard Run14 bad run list includes $\langle p_{\mathrm{T,Track}} \rangle$, the mean vertex location in the plane transverse to the beam $\langle V_z \rangle$, $\langle V_R \rangle$, $\langle \mathrm{ZDC}_{\mathrm{coin}} \rangle$ and $\langle refmult \rangle$. As mentioned previously with the bad tower list, there are two standardized bad run lists for Run14 jet analyses, depending on the choice of having a more or less restrictive bad tower list. The bad run list used, corresponding to the previously mentioned bad tower list, resulted in a total of 360 runs deemed bad, out of a total 2779 runs.



(a) (b)

Figure 3.7: (a) Collision vertex position for HT2*VPDMB30 events in the plane transverse to the beam direction before cuts. (b) Z-vertex position for VPDMB30 and BHT2*VPDMB30 events. The red lines denote V_Z cuts used in the analysis.

After removing bad runs, additional event-level cuts are imposed. While the triggers for both VPDMB30 and BHT2*VPDMB30 are required to have their z-vertex position V_z be within ±30 cm of the center of the detector, a more strict requirement of $|V_z| < 25$ cm was imposed to improve the binning used in the analysis (described in Section 4.2.1). The events are also subject to a cut requiring their radial vertex location to be $|V_R| < 1$ cm, in order to avoid background contributions from collisions with the beam pipe, which has a radius of 2.5 cm. The distributions of the vertex position before cuts can be found in

Figure 3.7. In the figure, the V_Z distributions between VPDMB30 and BHT2*VPDMB30 agree well within the range selected for analysis. Events are also rejected if the ZDC coincidence rate is greater than 100 kHz or less than 20 kHz to avoid low statistics causing issues with binning in the analysis method.

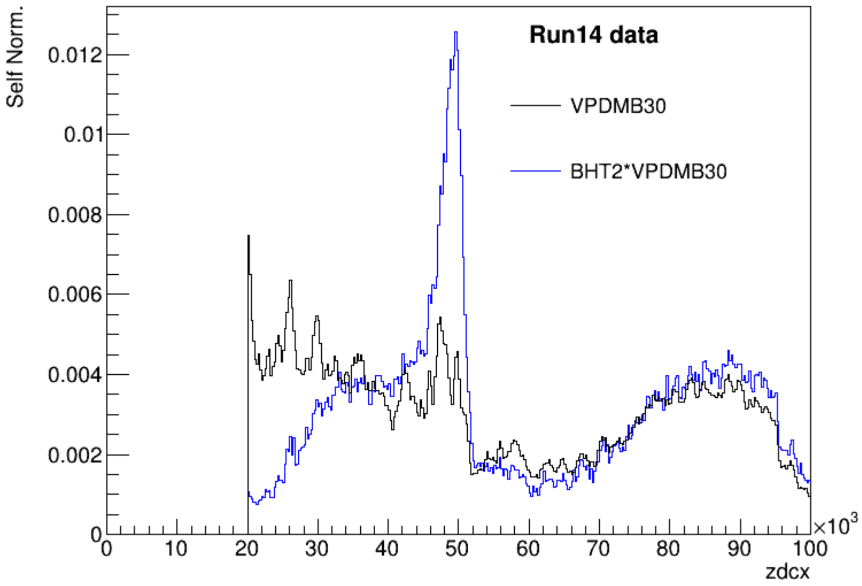


Figure 3.8: Run14 ZDC coincidence rate for the minium-bias trigger VPDMB30 and the high tower trigger BHT2*VPDMB30. Cut requiring ZDC coincidence to be greater than 20 kHz is already imposed.

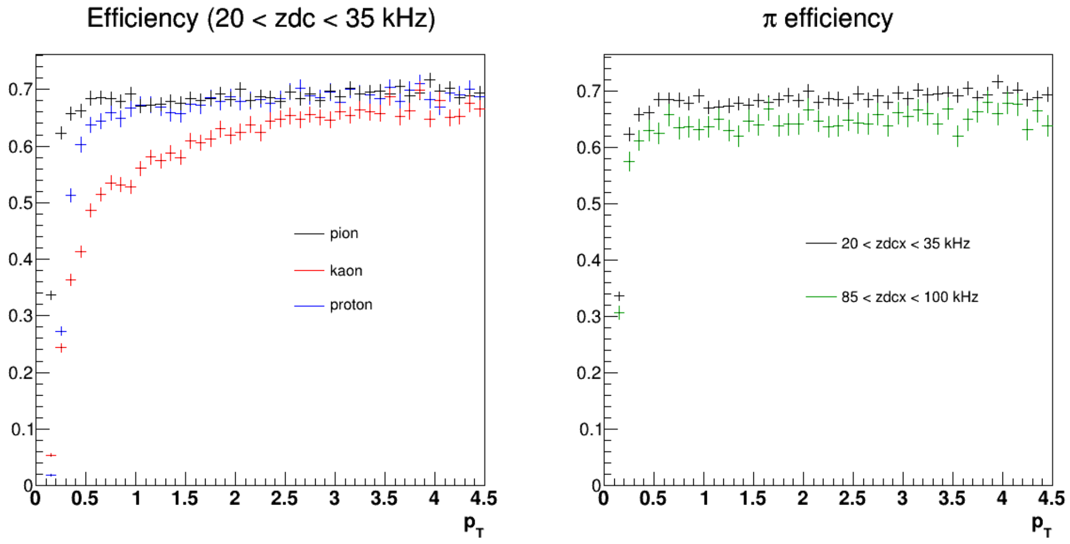
While the high-tower and minimum-bias triggered data are very similar, one important difference is the raw relative rate of high-tower triggers to minimum-bias triggers increases with increasing luminosity, measured by the ZDC coincidence rate. Due to this fact, the prescale of the VPDMB30 trigger increases and decreases depending on the amount of high-tower triggers. However, the changing of the prescale does not keep the relative rate of VPDMB30 and BHT2*VPDMB30 triggers the same. The self-normalized distributions of the ZDC coincidence rate for the VPDMB30 and BHT2*VPDMB30 trigger are shown in Figure 3.8, where the differences in the relative rates of minimum-bias and high-tower triggers are visible. During data taking, the data were broken into three different production streams (labelled low, mid, and high) based on the ZDC coincidence rate. The trigger setup for the peak in ZDC coincidence rate of ~ 50 kHz for BHT2*VPDMB30 events corresponds to data in the mid-luminosity production. For these runs the VPDMB30 prescale was noticeably higher, causing a relative spike in BHT2*VPDMB30 events compared to VPDMB30. The increase at low ZDC coincidence for VPDMB30 mostly corresponds to early data in Run14 before the data were split into different productions, where the prescale of VPDMB30 was significantly lower. There exists a sizeable amount of Run14 VPDMB30 data below a ZDC coincidence rate of 20 kHz but due to the low amount of BHT2*VPDMB30 triggers it was removed from the analysis. Because of the shape difference between the two triggers with respect to ZDC coincidence rate, comparisons using VPDMB30 and BHT2*VPDMB30 triggers have to be appropriately made in bins of ZDC coincidence rate.

3.2 Embedding

Measurements made using the STAR detector are convoluted by detector effects, such as momentum smearing of particles and particle reconstruction efficiency. Such effects need to be corrected for using a method called unfolding. In order to understand and correct for the detector effects, they are modeled with a simulation of the STAR detector using GEANT[63]. This is usually done in two ways. First, at the single particle level, where real data events have a simulated particle, generated using a Monte Carlo generator such as PYTHIA [64], embedded into them and the reconstructed values are compared to the embedded particle. Using the information learned from the single track embedding, analyzers can then smear PYTHIA events and embed them into real events, a technique called “Fast Simulation”. The other method, which is usually preferred, embeds the entire high-Q^2 events into real events and runs the GEANT simulation, directly yielding a sample of dijet embedded events. This process is referred to as "Dijet Embedding" and more directly allows users to model detector effects. The analysis in this thesis primarily uses Fast Simulation, with its details described below.

3.2.1 Fast Simulation

Fast Simulation in STAR is mainly used when there is no official dijet embedding done for the data set. It has the benefit of being flexible and requiring a much smaller embedding sample produced by the STAR embedding team. The accuracy of Fast Simulation relies on the particle finding efficiency and track momentum resolution not significantly differing inside the cone of a jet vs. outside a jet. Assuming this is the case, the analyzer uses the single track efficiency and momentum resolution to smear their own sample of produced PYTHIA events, and embeds them into real minimum-bias events to end up with a result similar to that of dijet embedding.



(a) (b)

Figure 3.9: Efficiency curves for Run14 single track embedding for (a) three different particles species and (b) differing ZDC coincidence rate.

The single track embedding performed by the STAR embedding team is done by using a Monte Carlo generator to generate particles and then simulating the effects of the STAR detector using GEANT[63]. The simulated reconstructed hit points in the TPC are then embedded into real minimum bias events and event reconstruction takes place as normal, maintaining the information of the original MC particle input. While the embedding is done by the embedding team at STAR, the extraction of efficiency and momentum resolution curves are left to the analyzer. The single track embedding only provides MC tracks with p_{T} up to 5 GeV/c, where the track finding efficiency plateaus and is unchanging as you increase track p_{T}.

The extrapolation to higher track p_{T} is trivial, as the efficiency after the plateau can be used. Some examples of track efficiency curves can be seen in Figure 3.9. The efficiency for the Kaon is noticeably lower than pions and protons at low-p_{T} due to the probability to decay and not be reconstructed by the TPC. Another important feature is that the tracking efficiency drops as the luminosity (or ZDC coincidence rate) increases, due to occupancy effects as described in 3.1.2. The tracking efficiency is also dependent on track η and ϕ, due to non-uniformities of the STAR TPC, as seen in an example two dimension efficiency histogram in Figure 3.10. Since the track finding efficiency is a function of particle species, ZDC coincidence rate, track η, and track ϕ, sampling the efficiency for Fast Simulation purposes is done in bins of all of these quantities in order to best emulate real data.

For $p_{\mathrm{T}} < 5$ GeV/c, like in the case of the efficiency, the track momentum resolution is directly sampled from the single track embedding. The momentum resolution of pions, kaons, and protons from the single track embedding can be found in 3.11. For non-pion particles, at very low p_{T} the reconstructed momentum differs from the generated p_{T} values due to the default pion mass assignment in track reconstruction. However unlike the track finding efficiency, the momentum resolution of tracks is not flat and is highly p_{T} dependent. The momentum resolution of tracks worsens as the p_{T} increases, due to the straightening of tracks within the TPC making the determination of the radius of curvature

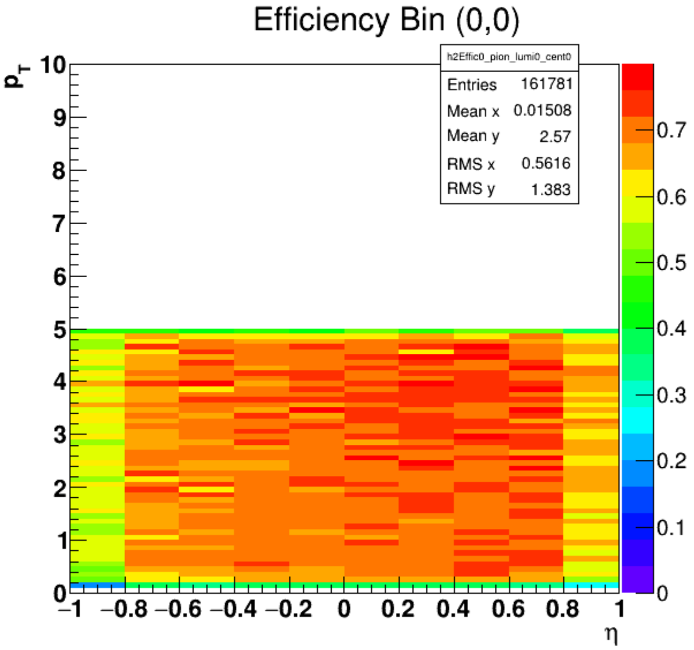
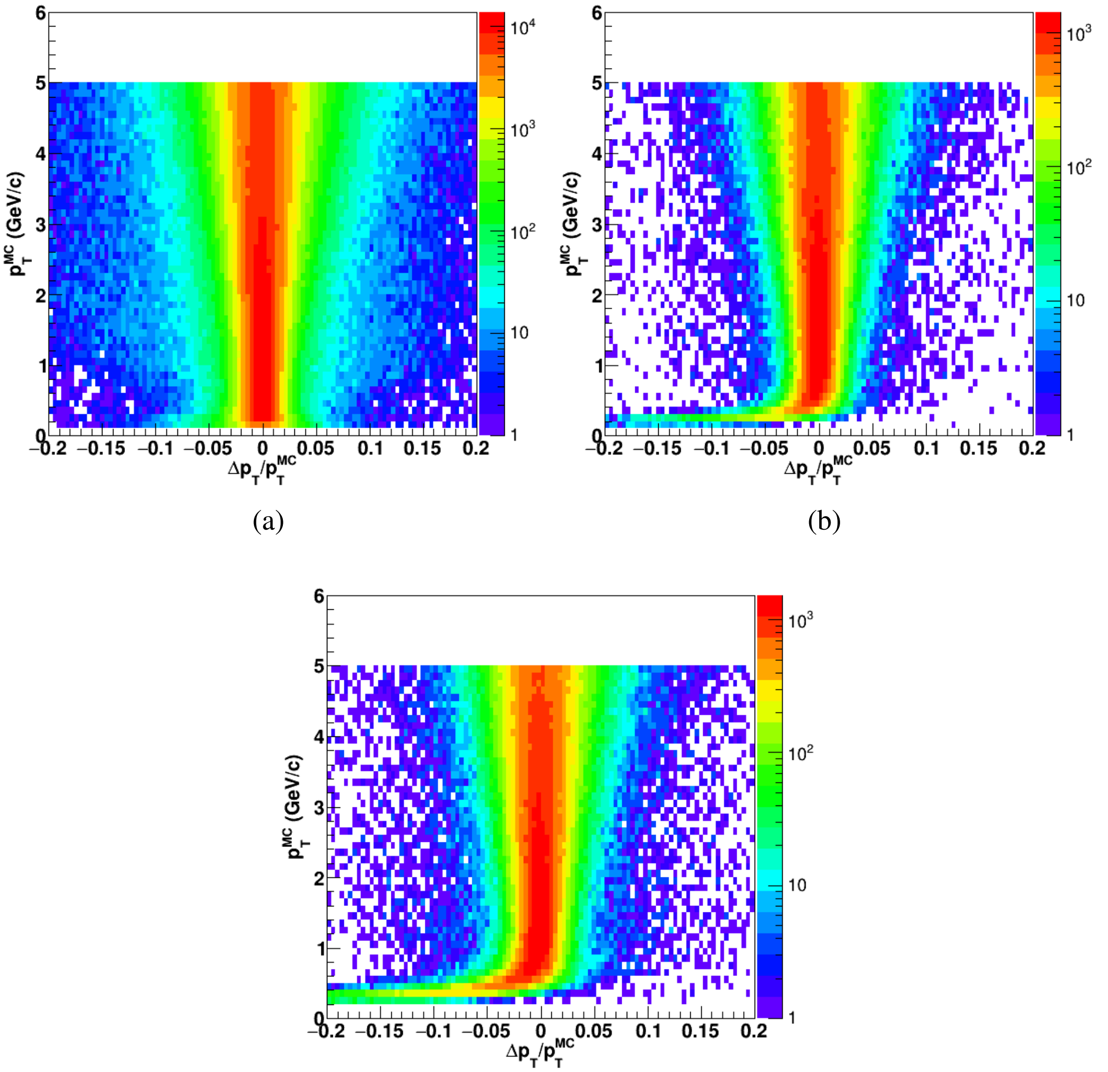


Figure 3.10: Example 2-D efficiency histogram determined using single track embedding of pions.

less precise.

In order to emulate the momentum smearing of particles above $p_{\mathrm{T}} > 5$ GeV/c, the momentum resolution needs to be parameterized as a function of p_{T}. This is accomplished by measuring $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{MC}}$ in p_{T} bins of width 200 MeV/c and fitting each bin with a Gaussian. After extracting the widths of the Gaussian distributions for each bin, the σ is plotted against $p_{\mathrm{T}}^{\mathrm{MC}}$ and a linear fit is performed to extract the intercept and slope (p0 and p1) fit values for the parameterization. This procedure was done for each particle species and in bins of ZDC coincidence rate. Figure 3.12 shows the Gaussian fit and the extracted p_{T} scaling for pions in a single ZDCx bin. For pions it was found to that $\sigma/p_{\mathrm{T}} \sim (0.66 + 0.32 * p_{\mathrm{T}})\%$.

While historically the momentum resolution for tracks was fit with only a single Gaussian, it was found that a double Gaussian fit (with a small pedestal) captured the behavior of the tails of the distribution better. The same procedure was applied to pions as seen



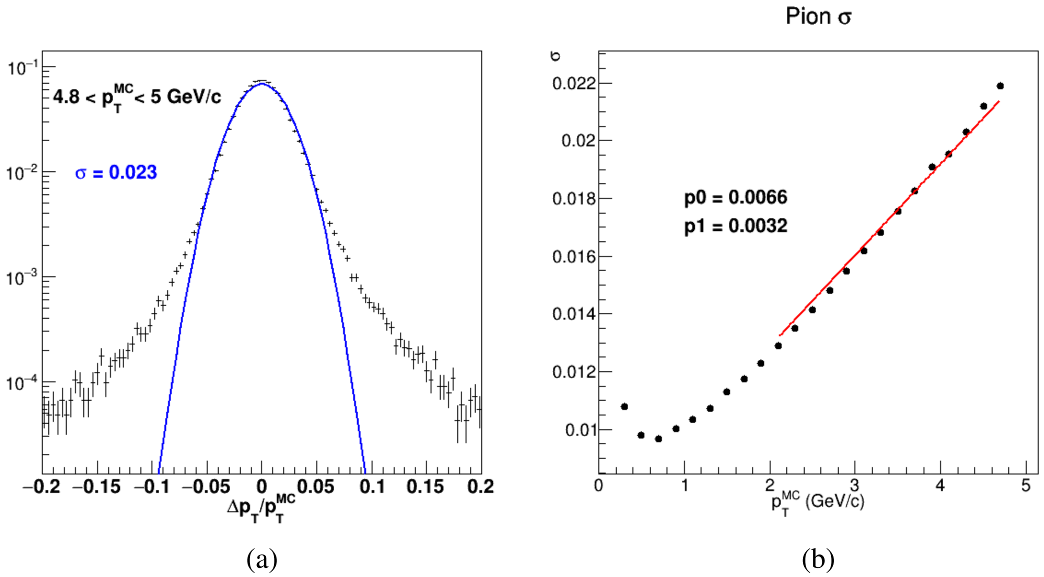
(a)

(b)

(c)

Figure 3.11: Momentum resolution for (a) pions, (b) kaons, and (c) protons.

in Figure 3.13. While both Gaussians were found to have p_{T} scaling, the first Gaussian captures the peak of the distribution and corresponds to the normally reported momentum resolution. However, while the second Gaussian's contribution is subleading and captures the tails of the distribution, it also captures information under the peak. The second Gaussian widths were of lower quality due to lower statistics at the tails, and were found to have different slopes with respect to p_{T}. Both single and double Gaussian fits were tested



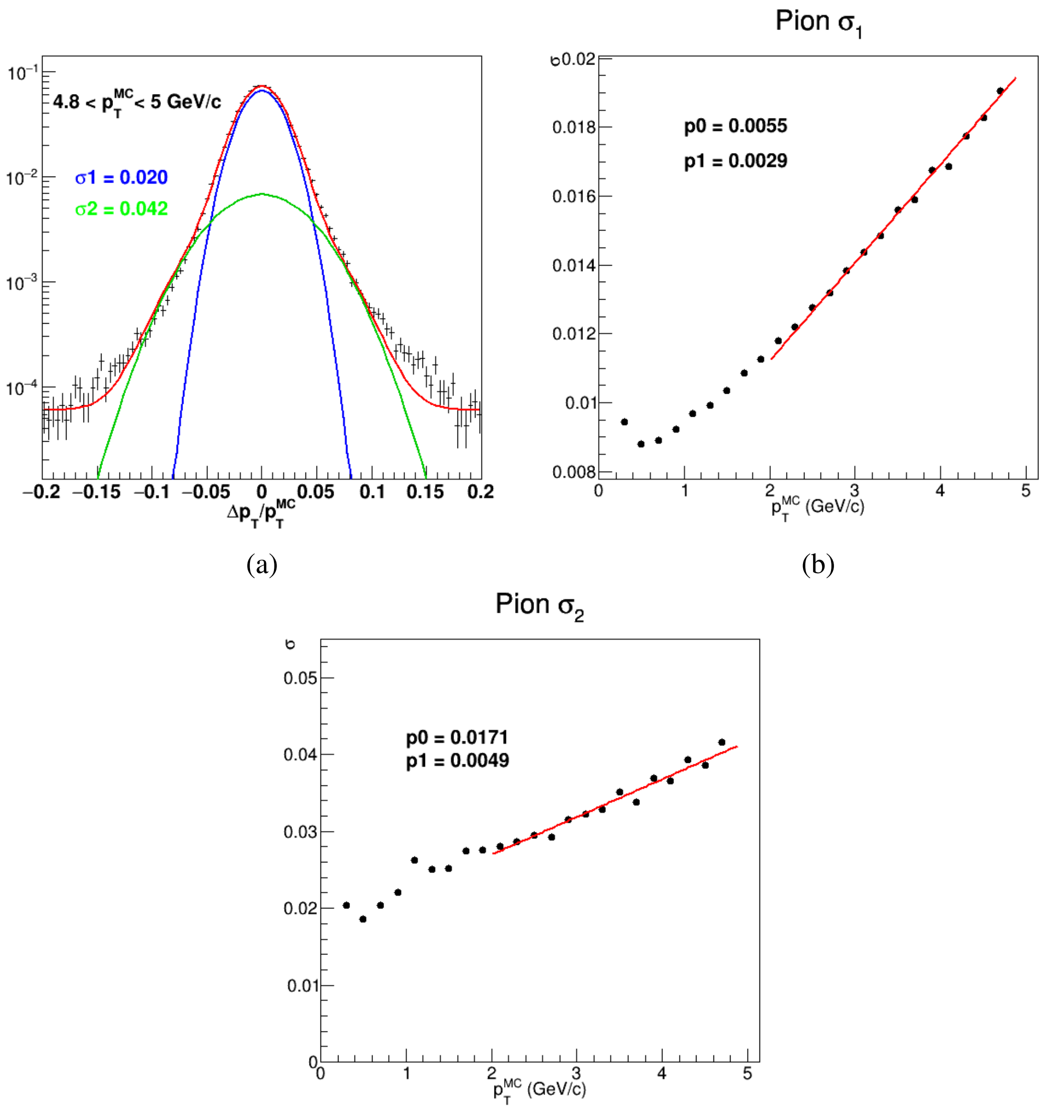
(a)

(b)

Figure 3.12: (a) Gaussian fit to the momentum resolution for pions for $4.8 < p_{\mathrm{T}}^{\mathrm{MC}} < 5$ GeV/c. (b) Fits to extracted σ values to extract the p_{T} dependence.

and the choice was found to have little impact on the final result. The double Gaussian fit was used primarily and the difference between the results using the double and single Gaussian fits was taken as a systematic and was found to be small.

After obtaining the parameterizations of track efficiency and momentum resolution, PYTHIA8 p+p events were generated with the same kinematic requirements as used in the analysis steps, requiring a photon or π^0 with $E_{\mathrm{T}} > 9$ GeV as a trigger (see Section 4.2 for details). The generated p+p events are embedded into minimum bias VPDMB30 events, sampled weighted according to the ZDC coincidence rate distribution of the selected events used in the analysis. The particle-level input is smeared according to the properties of the selected event. At the particle-level, detector-level (smeared but not embedded), and fully embedded-level, anti-k_t jets with R=0.4 are reconstructed within the same recoil range $|\phi_{\mathrm{trig}} - \phi_{\mathrm{jet}}| > \frac{3\pi}{4}$ as is done in the analysis. The response matrices going from particle-level to detector-level (instrumental response), detector-level to reconstructed level (background effects), and from particle-level to reconstructed level (instrumental and background effects) for 0-10% central events are shown in Figure 3.14.

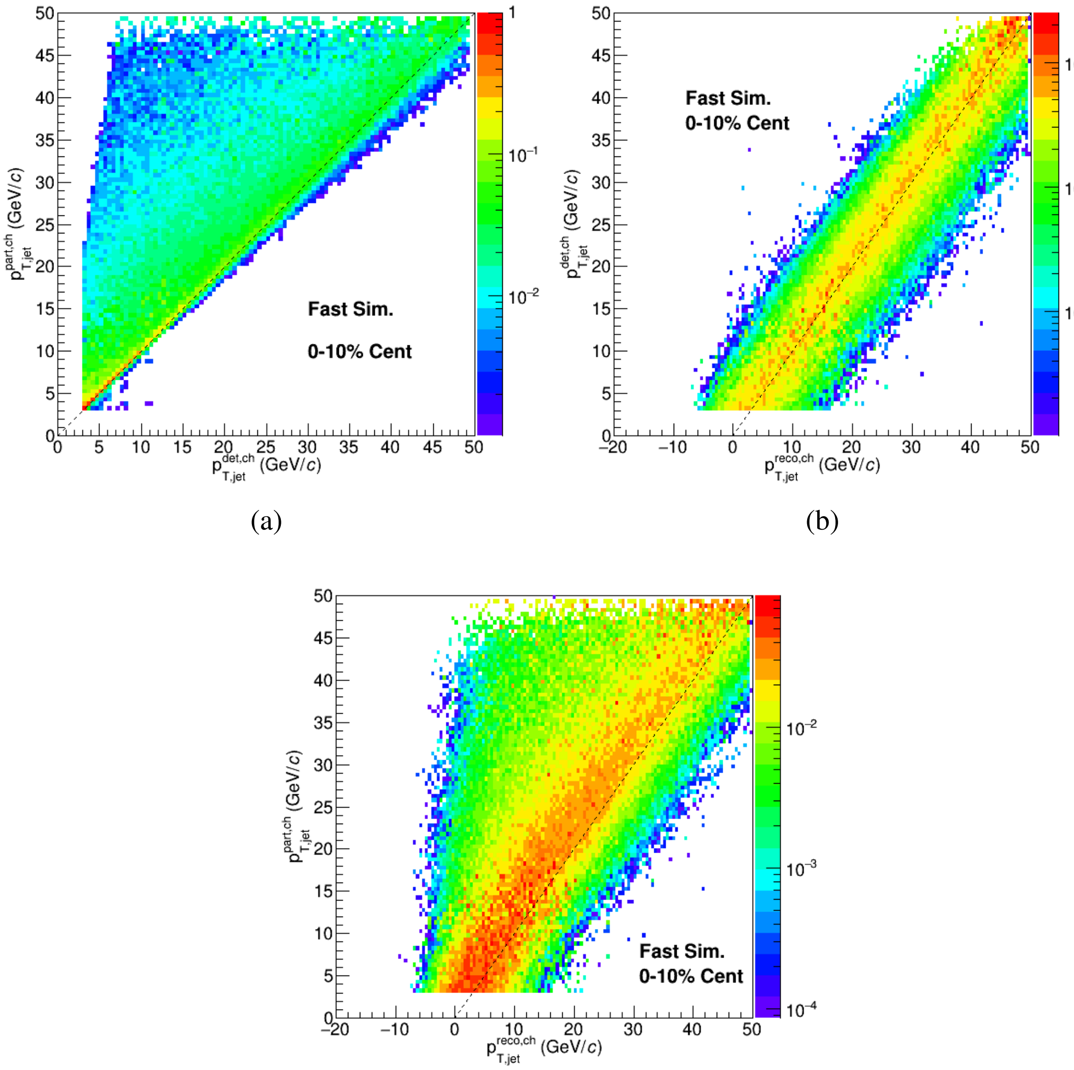


(a) (b)

(c)

Figure 3.13: (a) Double Gaussian fit of the momentum resolution for pions for $4.8 < p_{\mathrm{T}}^{\mathrm{MC}} < 5$ GeV/c. (b, c) Fits to extracted σ_1 and σ_2 values to extract the p_{T} dependence.

Note that the x-axis scale changes to include jets with reconstructed $p_{\mathrm{T,jet}} < 0$ GeV/c when including background effects (Figures 3.14b and 3.14c), due to the first order correction of the uncorrelated background, discussed in Section 4. The detector-level jets are matched to particle-level jets if they contain tracks matched to at least 15% of the particle-level jets' p_{T}. The same matching criteria is used to match reconstructed- to detector-level jets and reconstructed- to particle-level jets. The effects due to smearing and efficiency, the UE background, and their combination on the reconstructed jets can seen in the figure.



(a)

(b)

(c)

Figure 3.14: Response matrices for Fast Simulation for 0-10% central events. (a) From particle-level to detector-level, (b) detector level to reconstructed, and (c) from particle-level to reconstructed. The rows in each response matrix are normalized to unity.

Chapter 4

Analysis

4.1 Jet Reconstruction in A+A

In heavy-ion collisions it is a nontrivial task to isolate particles which originate from a high Q^2 process. In central A+A collisions there exists a large background from the numerous low Q^2 collisions in each event, often referred to as the underlying event (UE). The amount of p_{T} originating from the UE contained in a jet can be on the order of or greater than the amount p_{T} which originates from the hard scatter. In order to measure jet observables, such as the jet p_{T}, one needs to correct for this UE contribution. The sensitivity of a jet to soft and diffuse background particles is measured by its catchment area (A_{jet}), the area in ϕ-η phase space that a jet will capture soft background particles around it.[1] There are multiple methods of calculating a jet's area, but the most common method is using an active area determination, adding infinitesimally soft "ghost" particles diffusely in ϕ-η space. After jet clustering (or cone finding), the ghost particles clustered within the jet are used to determine the jets effective catchment area.[2] Anti-k_t jets, which are used to

---

[1] The area of a jet, A_{jet}, is not truly a geometrical area, but is actually the sensitivity of a jet algorithm to soft radiation. It is also not a value inherent to a single jet, but is dependent on the environment surrounding the jet.

[2] When using E-scheme recombination for jets, the appropriate area calculation is not the scalar sum of the ghost areas, it is the perpendicular component of the 4vector sum of the ghosts.

reconstruct jet candidates in A+A collisions, are often roughly circular with $A_{\text{jet}} \sim \pi R^2$, which yields a sensitivity to background particles that scales quadratically in jet resolution parameter R. This R-dependent sensitivity is one of the main reasons why jet analyses in A+A collisions commonly use a jet resolution parameter of $R = 0.4$ or lower.

While the amount of p_{T} from the UE within a given area fluctuates greatly, the first step is to estimate and correct for an average p_{T} density (ρ) in the event. This is accomplished by breaking up the phase space of the detector into patches, calculating the p_{T} density of particles within each patch and taking the median. While originally this was done with predefined grids in ϕ-η, modern analysis use jet clustering as a way to randomly define patches of area for calculating ρ. Both k_t and C/A clustering schemes are suitable for calculating ρ, as they both result in random patches of non-zero area. The anti-k_t algorithm however is a poor choice due to the presence of zero and near-zero jet areas, resulting in some unstable and undefined patches of p_{T} density. Ultimately the convention has been settled to use k_t jets, resulting in the ρ definition,

$$\rho = \underset{k_t-\text{jet}}{\text{median}} \left\{ \frac{p_{\text{T,jet}}}{A_{\text{jet}}} \right\} \tag{4.1}$$

Using the median makes the definition of ρ more robust against contributions from signal jets to the background estimation. However, even when taking the median, the presence of signal jets still can bias the value of ρ, so it is customary to exclude the two hardest k_t-jets from the median calculation. While excluding the two hardest is often motivated by the idea that it removes the contribution from signal dijets, one has to keep in mind that k_t-jets are highly sensitive to the background and do not correspond one-to-one with the candidate signal jets measured using the anti-k_t algorithm. The distributions of ρ determined for this analysis work is shown in later subsections, in Figures 4.6, 4.8, and 4.17. Using the measured value of ρ, the raw jet p_{T} of anti-k_t jets reconstructed for analysis are then corrected at first order with a pedestal-style correction,

$$p_{\mathrm{T,jet}}^{\mathrm{reco}} = p_{\mathrm{T,jet}}^{\mathrm{raw}} - \rho * A_{\mathrm{jet}} \tag{4.2}$$

This area-based subtraction only results in shifting the reconstructed jet p_{T} by the average background contribution, it does not account for background fluctuations. Due to the fluctuating background, jet candidates can have $p_{\mathrm{T,jet}}^{\mathrm{reco}} < 0$. Removal of the effect of background fluctuations on the reconstructed jet p_{T} is often done at the statistical ensemble level using unfolding techniques [65].

Constituent Subtraction

The area-based subtraction method is used to, at first order, correct the jet p_{T} for the background, however more complicated jet subtraction techniques are needed to measure more sophisticated jet observables. There are various background subtraction techniques available, but the one of interest for this thesis is Constituent Subtraction [66]. Constituent Subtraction is a general purpose jet subtraction technique which modifies jets at the constituent level. Constituent Subtraction makes use of the ghost particles used for area determination to subtract p_{T} from the constituents within the jets. This is done by first creating a list of all matches of physical particles i and ghost particles k ordered by a distance measure,

$$D_{i,k} = p_{\mathrm{T},i}^{\alpha} * \Delta R_{i,k} \tag{4.3}$$

from lowest to highest. The parameter α is a free parameter and $\Delta R_{i,k}$ is the distance between the ghost and physical particle in ϕ-η space. The ghost particles are all evenly assigned p_{T} such that the sum of all the ghost p_{T} is equal to $\rho * A_{\mathrm{jet}}$. Then iteratively over all pairs of ghosts and particles, from smallest to large $D_{i,k}$, the amount of p_{T} assigned to the ghost is subtracted from the constituent. If the ghost was able to subtract all of its p_{T} from the constituent, then it is removed and the process continues. The algorithm terminates when all the ghosts are removed or when the algorithm reaches $\Delta R_{i,k} > \Delta R^{\mathrm{max}}$, where ΔR^{max} is another free parameter.

An important feature of Constituent Subtraction is that it subtracts practically the same amount of p_{T} as the area-based method, assuming ΔR^{max} is set to a high enough value that the algorithm does not terminate early.[3] However, by the nature of Constituent Subtraction, it cannot yield jets with $p_{\mathrm{T,jet}}^{\mathrm{reco}} < 0$. This is important because this thesis work uses these negative p_{T} jets, as explored in Section 4.2.3. It should be noted that Constituent Subtraction as described above is done on a jet-by-jet basis. However, Constituent Subtraction can be done at the event level through the same process but instead of using ghost particles within the area of a jet, using ghosts within the entire event and modifying the constituents at the event-level before jet clustering [67]. While this method may improve the jet energy resolution, it is not suitable for this analysis, because the jets reconstructed after event-level constituent subtraction do not correspond directly to area-based subtracted jets.

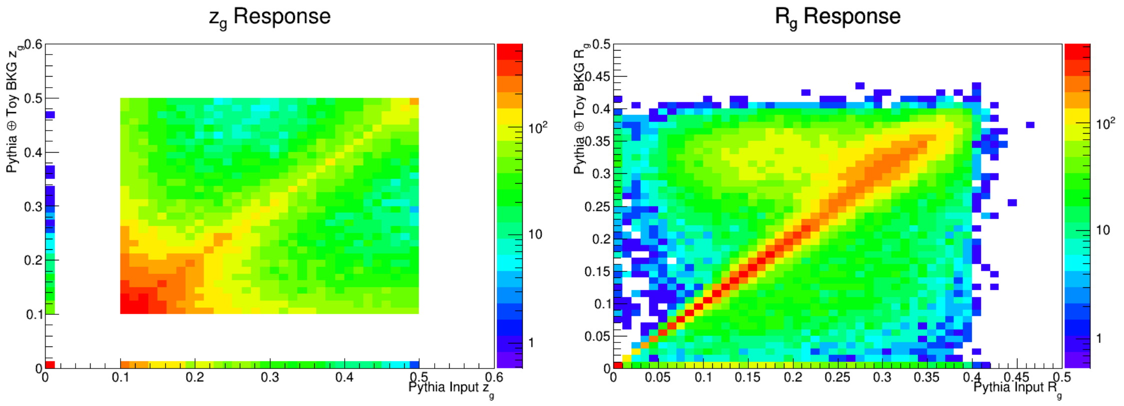
Subtraction Performance Study

The performance of Constituent Subtraction for z_{g} was studied by embedding PYTHIA8 events into a toy MC background, consisting of tracks generated randomly in ϕ-η according to a thermal distribution. The study was originally done for the expected jet p_{T} ranges and background densities for full jet reconstruction, embedding R = 0.4 anti-k_t PYTHIA jets with 20 $< p_{\mathrm{T,jet}} <$ 30 GeV/c into a background with $\langle\rho\rangle \sim$ 62 GeV/(c Area). While the kinematics are different than in the analysis, the generic features of the performance of Constituent Subtraction still apply. Figure 4.1 shows the response for z_{g} and R_{g} using Constituent Subtraction with the free parameter α = 2. Setting α = 2, which tunes

---

[3] Identically the same when using pt-scheme recombination.

Constituent Subtraction to preferably subtract p_{T} from low-p_{T} constituents in the jet, was found to have better performance than smaller values of α. Note that the bins at z_{g} and R_{g}



(a) (b)

Figure 4.1: The (a) z_{g} and (b) R_{g} response matrices for PYTHIA8 jets with $20 < p_{\mathrm{T,jet}} < 30$ GeV/c using Constituent Subtraction at the jet-level with α = 2.

= 0 (on both axes) correspond to jets that do not pass SoftDrop grooming. In the response matrices non-diagonal elements can be seen. This non-diagonal behavior is easily understood by looking at the horizontal band at reconstructed $R_{\mathrm{g}} \sim 0.35$ in the R_{g} response. This structure is due to the presence of "Fake Subjets", subjets that would not normally pass the SoftDrop criterion $\left(\frac{min(p_{\mathrm{T},1},p_{\mathrm{T},2})}{p_{\mathrm{T},1}+p_{\mathrm{T},2}} > z_{\mathrm{cut}}\right)$ but due to the background get promoted above z_{cut}. The opposite effect in which subjets which would normally pass SoftDrop but do not due to the background and get demoted to a dropped subjet occurs, though to a lesser degree. These effects of promoting and demoting branches in the clustering history above and below z_{cut} can most easily be seen in the z_{g} = 0 bins along each axis. These non-diagonal features are problematic because they make the measurement difficult to correct back to the particle level distributions via unfolding. In order to reduce the effect of these non-diagonaly elements, one would have to go to a smaller jet resolution parameter and/or to higher z_{cut}, which was done in the ALICE measurement of z_g [43]. For this analysis, the original definition of z_{g} was kept and not unfolded, instead making a comparison to smeared PYTHIA embedded into Au+Au events.

4.1.1 Combinatorial Jets

Not only do fluctuations of the soft background exist within signal jets, but upwards fluctuations can result in purely combinatorial jets, sometimes referred to as “fake jets”. These jets are reconstructed with particles only originating from many low Q^2 processes, and thus are considered contamination of the jet signal. Historically a common way of discriminating between signal jets and combinatorial jets is by applying a leading hadron cut, requiring jets to contain a sufficiently high-p_{T} particle such that it must have originated from a high Q^2 process[22]. While this method succeeds in separating signal from background, it does so at the high cost of imposing a bias on the measured jet sample. First, requiring a high-p_{T} hadron selectively chooses jets which have a harder fragmentation. This bias gets reduced as one increases jet p_{T}, causing the minimum jet p_{T} for a substructure measurement to be increased. Second, requiring a high-p_{T} hadron in A+A collisions induces a surface bias on the measured jet sample. Jet energy loss is expected to be path length dependent, the shorter path the jet-initiating parton travels in the medium the less energy loss the jet experiences. By selecting a jet sample with a high-p_{T} particle, we are selecting a jet sample which is expected to be biased to having a shorter path length in the medium, and thus our signal of jet modification is potentially reduced. A similar method is to select jets which are clustered with only high-p_{T} particles, such that the final jet sample has a “hard core”. This is done by several analyses at STAR and imposes a similar bias to that of requiring a single high-p_{T} particle[68].

4.2 Semi-inclusive Approach

Restricting a jet sample by cutting on jet p_{T} directly or indirectly can cause a bias and effectively limit the lower bound range of jet p_{T} one can meaningfully measure. An alternate method to remove the effect of combinatorial jets from the measured jet sample is to make a semi-inclusive measurement. Here, minimal jet-by-jet cuts are applied and instead combinatorial jets are subtracted at the level of ensemble distributions using a mixed event technique[69, 70]. This method begins with taking events with a high-p_{T} trigger object, sufficiently high p_{T} such that the trigger must have come from a high-Q^2 process, and measuring all jet candidates within the recoil range, such that $|\phi_{\mathrm{trig}} - \phi_{\mathrm{jet}}| > \frac{3\pi}{4}$. The jet candidates at this point are comprised of both jets which are correlated and uncorrelated with the trigger particle. In order to capture the behavior of the uncorrelated yields, a mixed event sample is created from real events such that any jets reconstructed are entirely combinatorial in nature. Knowing the per-trigger uncorrelated yields from the mixed events allows one to extract the signal yields after careful consideration of details outlined in the sections below.

4.2.1 Event Classes

Same Event

The events containing the high-p_{T} trigger object with the physically correlated jets belong to the class of events named "Same Event" (SE). These events are selected from BHT2*VPDMB30 triggered events, with the additional requirement of having a tower in the BEMC with $E_{\mathrm{T}} > 9$ GeV, after hadronic correction. This additional higher E_{T} cut than the Run14 high tower trigger is made to have a sufficiently high energy trigger object. This leaves no doubt that the particle derives from a high-Q^2 process and also increases the

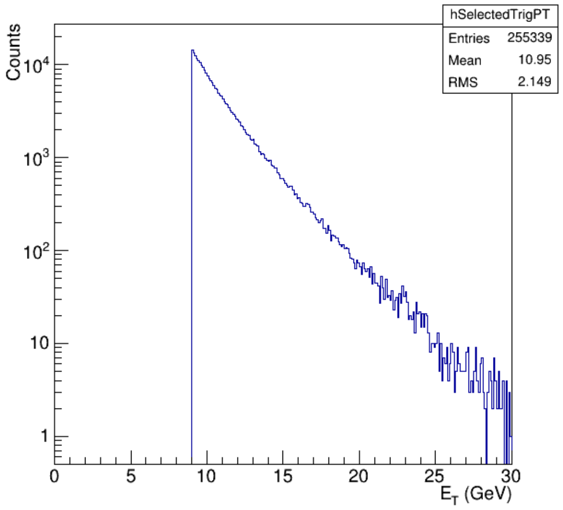


Figure 4.2: Selected trigger tower E_{T} found in 0-10% central events.

per-trigger correlated jet yields within lower measured jet p_{T} bins. In the case of having multiple towers with $E_{\mathrm{T}} > 9$ GeV, one of the trigger candidates is selected as the trigger object at random, in order to avoid a bias in the energy of the trigger spectra. Such a high energy requirement makes this analysis method very statistically limiting, yielding approximately 250 thousand events in the Same Event class, as seen in Figure 4.2. There is no discrimination between hadrons and direct photons which deposit sufficient energy into the tower, both are indiscriminately taken as trigger objects.

Mixed Event

Mixed Events (ME) are events which are designed to capture the behavior of the detector and jet reconstruction in order to emulate the combinatorial background without having physics correlations among the particles used to create them. Mixed Events are constructed by first sorting minimum bias events, in this case VPDM30 events, into bins of z-vertex position, ZDC coincidence rate, event-plane angle, and particle multiplicity. The bin axes are chosen such that events within each bin are sufficiently similar such that gross event properties are similar. Ideally the bin widths would be infinitesimally small, but practically they are chosen to be as small as possible given the statistics available. The limiting factor is that every bin has to contain at least as many events as the maximum amount of tracks for events within that bin. This constraint incentivizes having the statistics smoothly evened out among the bins. The bins in V_Z and Ψ_{EP} are defined with static widths, while the bins in ZDCx and track multiplicity are variable. The ZDCx bin widths are manually chosen such that there are roughly equal amounts of statistics in each, to account for the spiky structure (as already shown in Figure 3.8). The particle multiplicity bin widths are percentile based, with the multiplicity ranges for each (V_Z, ZDCx) bin are defined separately such that each multiplicity bin has uniform statistics. Figure 4.3 highlights the need for such bin-dependent widths, where the high ZDC coincident rate events have a noticeably lower average track multiplicity. The number of bins and widths for each axis can be found in Table 4.1, for a total of 960 mixing bins.

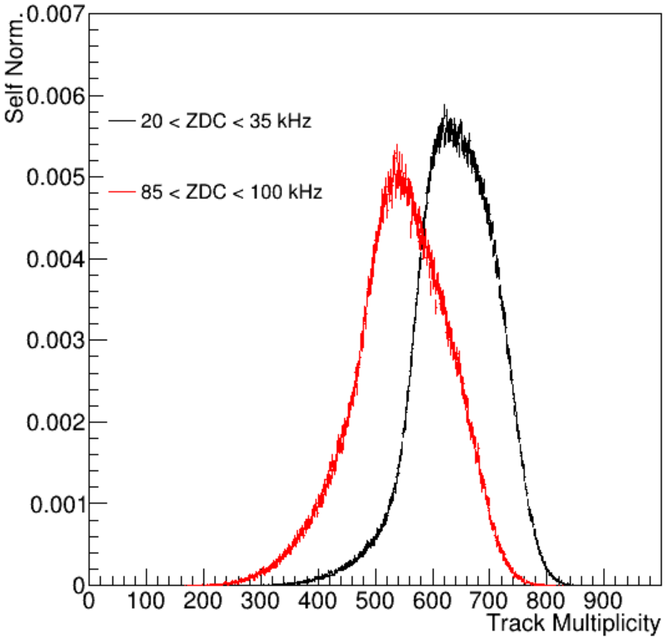


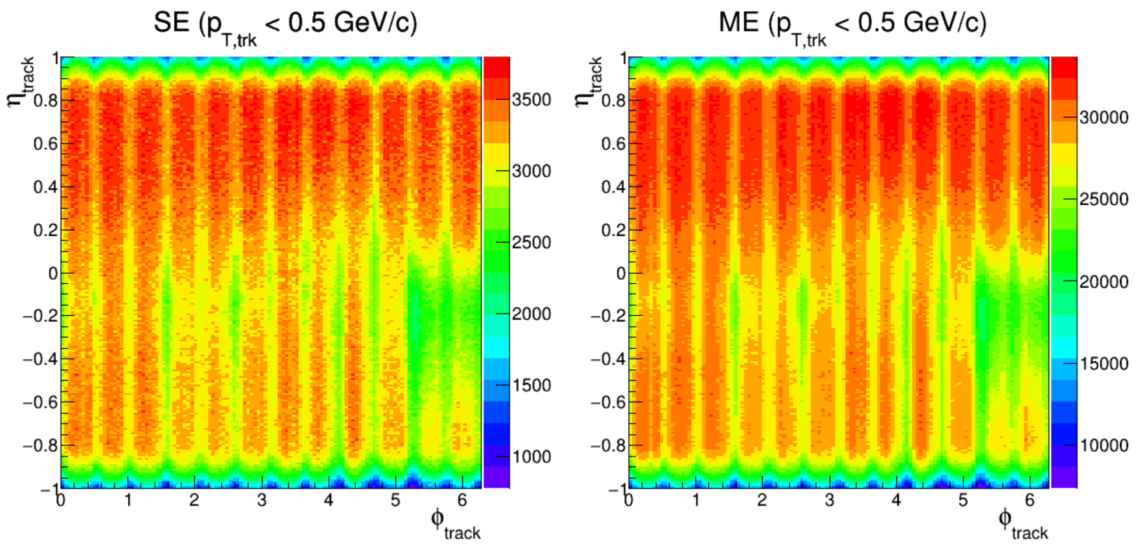
Figure 4.3: Track multiplicity for 0-10% central VPDMB30 events. The difference between the red and black histograms is primarily due to occupancy effects lowering the primary track efficiency at high ZDC coincidence rate, as explained in section 3.1.2.

Bin Type	No. of bins	Width
V_Z	5	10 cm
Ψ_{EP}	4	$\pi/2$
ZDCx	6	Variable
Multiplicity	8	Percentile

Table 4.1: Chosen bins and bin widths used for creating Mixed Events.

After sorting the VPDMB30 events into their respective bins, then Mixed Events are created by first randomly sampling a particle multiplicity of a VPDMB30 event within a selected bin. Next, single random tracks from separate unique real events are sampled until the target multiplicity is reached. The use of only one track per real event is implemented in order to satisfy the requirement of destroying any inter-particle correlations. After the Mixed Event is created, the used tracks are removed from the pool and the next Mixed Event is created. This process keeps occurring until any event in the pool runs out of tracks. While originally each track was used uniquely one time in the Mixed Events, it was found that the Mixed Events did not have enough statistics. In order to increase the statistics available, after the pool becomes too sparse, the pool is regenerated and remixed. To make sure the Mixed Events are sufficiently independent from each other, when generating any given Mixed Event a random event is selected and from that a random track is chosen.

In Figure 4.4 the distributions of tracks with p_{T} less than 0.5 GeV/c in $\phi - \eta$ space for ME and SE event classes are shown. Their projections into ϕ and η are found in Figure 4.5. The red bands in 4.4 are the centers of the TPC sectors, where the tracking efficiency is at its highest. From the figures one can see the Mixed Events capture the behavior of the detector, including the lower efficiency from the poor TPC sectors at $\phi >$ 5.4 and $\eta < 0$. While the Mixed Event class captures the behavior of the UE in the Same Event class, the key difference is that there are no physically correlated jets.

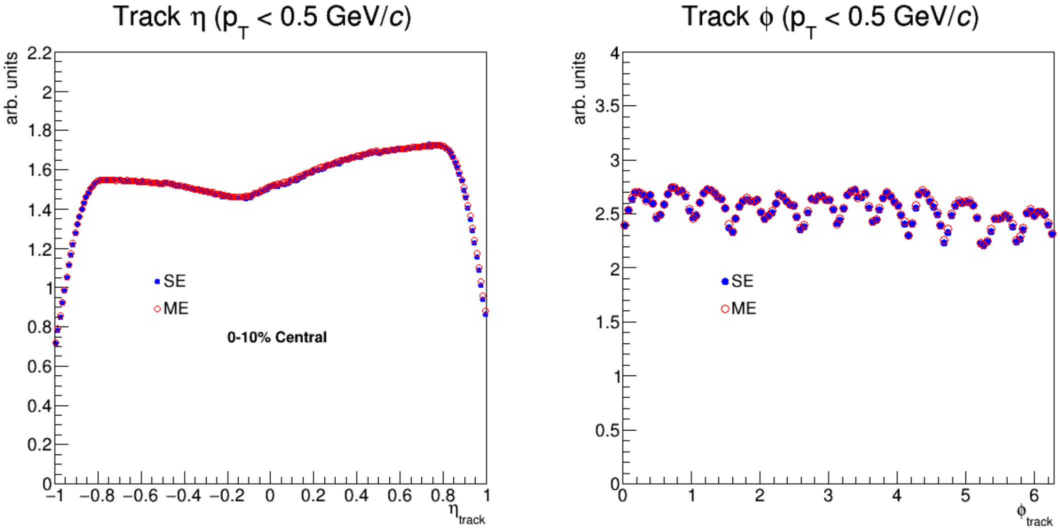


(a) (b)

Figure 4.4: Charged tracks in $\phi - \eta$ with $p_{\mathrm{T}} < 0.5$ GeV/c in both (a) SE and (b) ME event classes.

4.2.2 SE and ME Median p_{T}-Density

When iterating over all the events found in the SE class, for every SE event multiple ME events are sampled from the event mixing bin corresponding to the SE event. For example, if a SE event falls into the (0,2,2,6) bin, then ME events from that same (0,2,2,6) bin are sampled. This is done such that the events from the ME sample have correspondingly similar event qualities, even if the High Tower and Minimum bias events have different distributions for their event-level properties (i.e. difference in ZDCx distributions, as shown in Figure 3.8). Then for both SE and ME event samples, the median-p_{T} density ρ is calculated. Typically when calculating ρ, to eliminate the contribution from any hard scatter present, the 2 hardest k_t-clustered jets are excluded from the median calculation. The ME events do not have any correlated hard contributions, so if one calculates the SE and ME ρ in the same way, the ME ρ will be shifted to a lower value. However, the ρ of an event is designed to be a description of the underlying event and not dependent on any hard physics present. To account for such a discrepancy, a slightly different definition of ρ between the



(a) (b)

Figure 4.5: Charged tracks in (a) η and (b) ϕ for SE and ME event samples.

ME and SE is used. Ultimately the absolute definition of ρ is to some degree arbitrary, it is more a practical experimental contrivance. However, for the sake of this analysis the relative agreement between the ME and SE is important, and for this analysis excluding a single hardest jet from the ρ calculation for SE, and excluding no hard jets from the ME calculation was found to be optimal and was used. The initial ρ distributions for SE and ME central events can be found in Figure 4.6. The ratio of SE/ME is relatively flat and around unity near the peak of the distribution, however a structure can be seen in the ratio at $\rho \sim 20\,\mathrm{GeV}/(c\,\mathrm{Area})$. Upon investigation, it was discovered to be caused by events in the lowest multiplicity bin, which had a strange bias relating to selecting low multiplicity events within a centrality bin. The η distributions for three different multiplicity bins can be found in Figure 4.7. The $refmult$ definition of counting charged tracks within $|\eta| < 0.5$ has a clear and visible effect for the lowest multiplicity bin η distributions, and begins to disappear in the second multiplicity bin. Interestingly, the ME event sample is able to capture this behavior, yet there is disagreement between SE and ME in the shape of ρ for these events. Agreement between SE and ME is paramount, and an additional cut

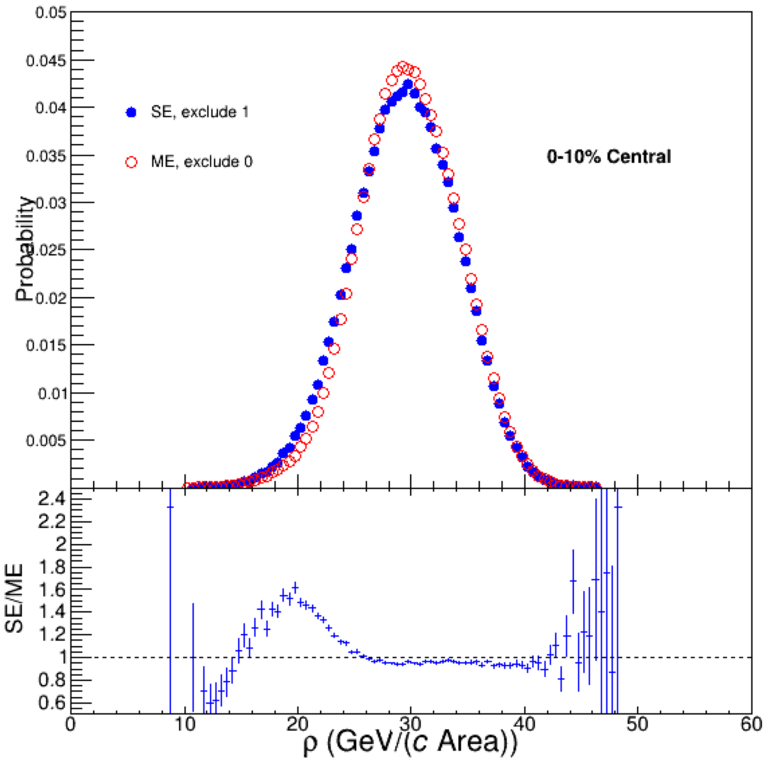
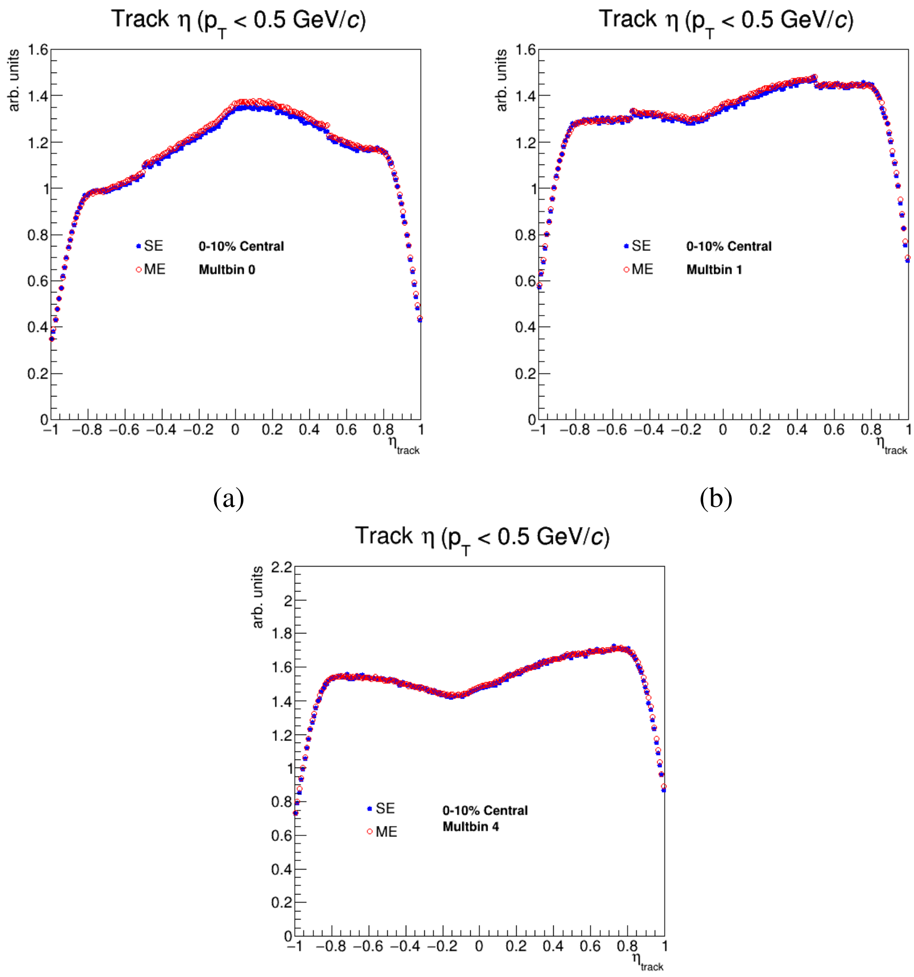


Figure 4.6: Median p_{T} density ρ in SE and ME event classes in 0-10% central events, calculated excluding the hardest jet in SE and excluding no jets in ME. Bottom panel shows the ratio of the two distributions.

on events was imposed that removed the bottom 8% multiplicity events, before defining the multiplicity bins for mixing. Following this additional cut, the ρ distributions can be seen to agree for SE and ME event classes in Figure 4.8.

Perfect agreement between the distributions is not possible because of the need for mixing in bins of nonzero width. The ME ρ distribution is characteristically more narrow than the SE ρ distribution, even in individual mixing bins. This causes an enhancement in the ratio at the tails of the distribution, which is more noticeable at the left edge of the distribution. The effect of the disagreement between SE and ME ρ is taken as a source of systematic uncertainty, by varying the choice of definition of ρ, which is similar to a shift of the distributions (but not identical to).



(a) (b)

(c)

Figure 4.7: Track η distributions for both SE and ME event classes within (a) the lowest multiplicity bin, (b) the second lowest multiplicity bin, and (c) one of the middle multiplicity bins.

4.2.3 Recoil Jet Spectra

For Same Events, jets are reconstructed using the anti-k_t algorithm with jet resolution parameter R=0.4 and are required to be located in the recoil window of the trigger object, $|\phi_{\rm trig} - \phi_{\rm jet}| > \frac{3\pi}{4}$. The jet candidates' reconstructed $p_{\rm T}$ are corrected at first order for the large underlying event using area-based subtraction as outlined in Section 4.1. The result is that the per-trigger recoil jet yield distribution extends below jet $p_{\rm T}$ = 0 GeV/c. These

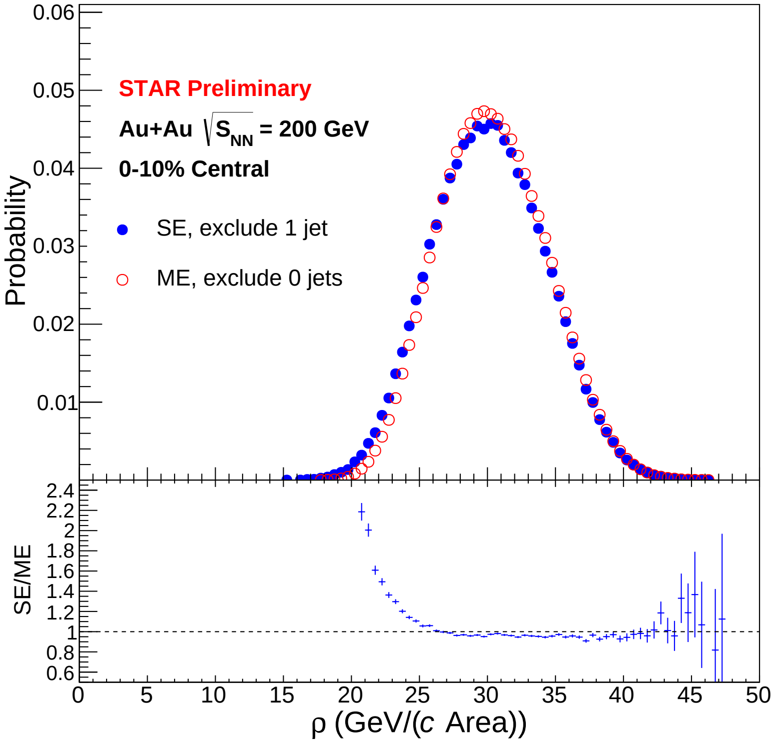


Figure 4.8: Median p_{T} density ρ in SE and ME event classes in 0-10% central events, calculated excluding the hardest jet in SE and excluding no jets in ME. The bottom 8% multiplicity events before binning are removed, which improves agreement between the SE and ME event classes. Bottom panel shows the ratio of the two distributions.

negative p_{T} jets are not necessarily combinatorial and are due to fluctuations in the UE; after the pedestal-like subtraction the correlated signal on a downward fluctuation can fall below p_{T} = 0 GeV/c. The same procedure is done in the Mixed Event class, however due to the nonexistence of a trigger object, the trigger axis is sampled from Same Events in order to avoid effects from the non-uniformity of the detector.

The only jet-by-jet discrimination is a jet area cut, $A_{\mathrm{jet}} > 0.35$, removing jets with low jet area which are artifacts of anti-k_t clustering. This cut serves the purpose of reducing the artificially enhanced jet p_{T} around zero from the per-trigger jet yield distributions due to the large amount of jets with near zero jet area from anti-k_t clustering. The reconstructed jet p_{T} and jet area before the area cut is applied is shown in Figure 4.9 for SE and ME

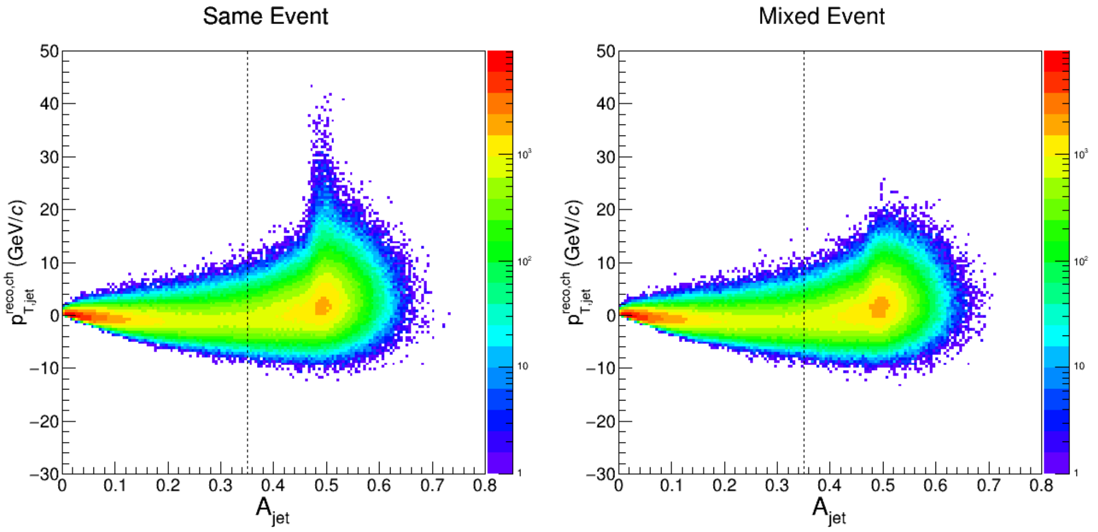


Figure 4.9: Reconstructed jet p_{T} after pedestal subtraction versus jet area for SE and ME event classes, before the area cut at $A_{jet} > 0.35$ (dashed line) is applied. The SE and ME distributions are similar below the area cut, with the difference between the SE and ME samples only being prominent above the area cut.

event classes. The distributions are mostly similar, with the only difference being the rising high-p_{T} tail at $A_{\mathrm{jet}} \sim 0.5$ in SE, which is driven by the correlated signal yield. The rest of the SE distribution, including the section below the area cut, matches the ME distribution, which is entirely driven by combinatorial jets.

The overlaid per-trigger recoil jet yields can be seen in Figure 4.10. The left most edge of the distribution is expected to be entirely dominated by combinatorial contributions, and the shape of the per-trigger yields between the Same Event and Mixed Event classes should agree. The magnitudes of the yields however don't agree due to the fact that the total integrated yield of jets in the recoil region is geometrically driven and roughly the same between the populations. Without requiring a jet p_{T} threshold, when you reconstruct jet candidates within a given acceptance the jet finder will always return jets, regardless if they are correlated or not. The values of the integrals of the entire SE and ME distributions agree within 1% for the 0-10% bin. Due to the conservation of number of reconstructed jets, the enhancement of per-trigger yield in the Same Event high jet p_{T} region, which is

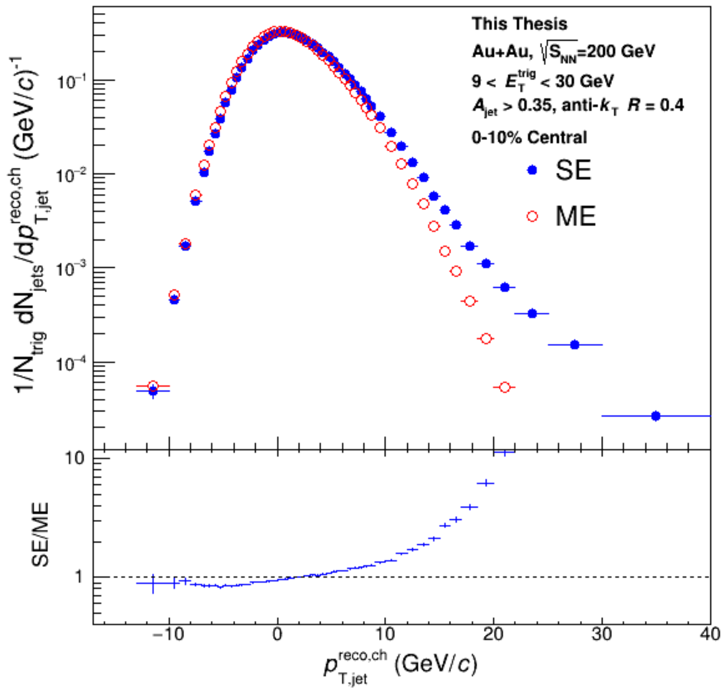


Figure 4.10: Per-trigger recoil jet yields for SE and ME in 0-10% central events. Combinatorial jet yield not yet subtracted, which dominates left edge of both distributions.

physically correlated to the trigger object, causes a smaller yield in the purely combinatorial region. To account for such an effect, the Mixed Event distribution is scaled down by a factor f_{ME}. This scale factor is found by integrating the left most edge of the distribution where the region of the ratio of SE/ME is relatively flat, in which the shape of the two distributions are the same. The scale factor f_{ME} is not strongly sensitive to the selection of the scaling region range, and shifting the scaling region is used as a systematic uncertainty. The per-trigger recoil yields for SE and the scaled ME distribution for 0-10% centrality can be found in Figure 4.11. Note that after the scaling, the left most region in the ratio is flat and falls on top of unity. The values of the SE integral, ME integral, the scaling region upper-bound, and the f_{ME} values can be found in Table 4.2 for different centrality bins. The 40-60% peripheral bin was not used in the analysis but is included for instructive purposes. The agreement between the SE and ME integrated per-trigger yields is best in

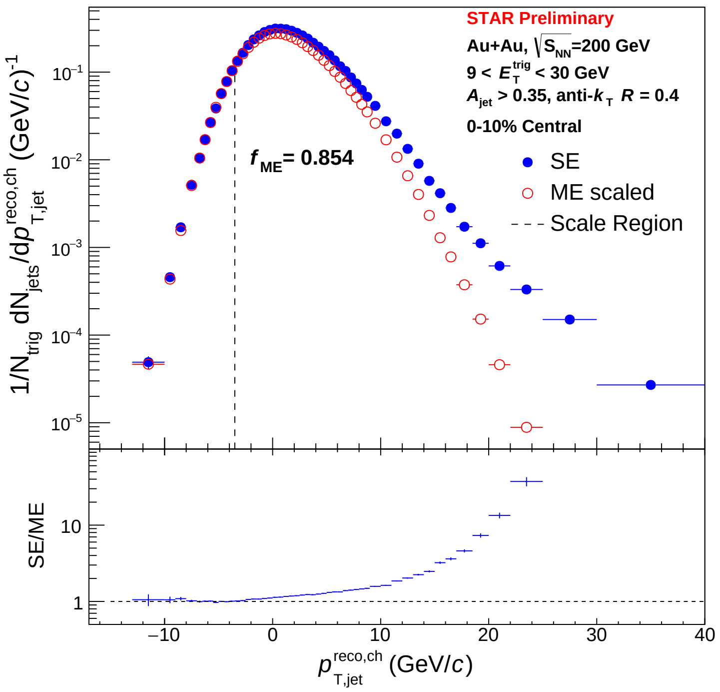


Figure 4.11: Per-trigger recoil jet yields for SE and ME scaled down by a factor f_{ME} in 0-10% central events. Combinatorial jet yield not yet subtracted, which dominates left edge of both distributions. Dashed line denotes right edge of scaling region used to determine f_{ME}. Bottom panel shows the ratio for SE/ME, in which the ratio increases at increasing jet p_{T}, where one expects to find a higher contribution of correlated yield compared to combinatorial yield.

the most central bin and decreases for more peripheral bins.

Using the SE and scaled ME recoil jet yield, one can then extract the per-trigger signal yield by subtracting the two distributions, as seen in Figure 4.12 for 0-10% events. The left edge of the distribution where correlated yield is not expected is mostly consistent with zero. This per-trigger recoil jet yield has the combinatorial contributions subtracted out, but still is not corrected for fluctuations of the UE within signal jets.

Centrality	SE Integral	ME Integral	Scale Region Upper-bound	f_{ME}
0-10%	5.45	5.48	-3.5	0.854
10-20%	5.26	5.36	-3.5	0.851
40-60%	5.02	5.18	-1.0	0.776

Table 4.2: Integrals, scale region upper-bound, and f_{ME} for different centralities.

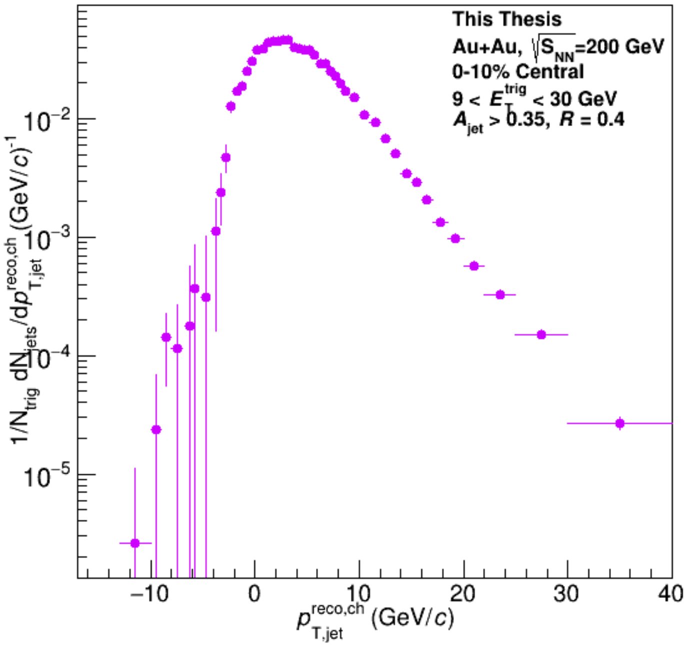
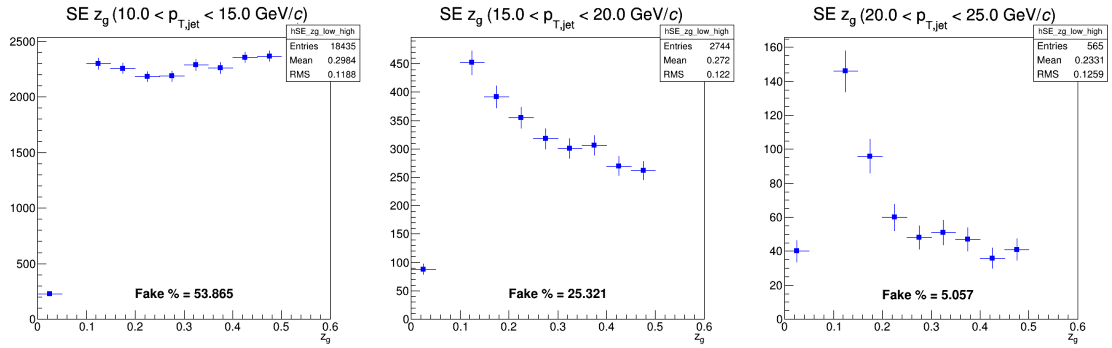


Figure 4.12: Raw combinatorial-subtracted per-trigger recoil jet yields in 0-10% central events. Only statistical uncertainties are shown.

4.3 Shared Groomed Momentum Fraction

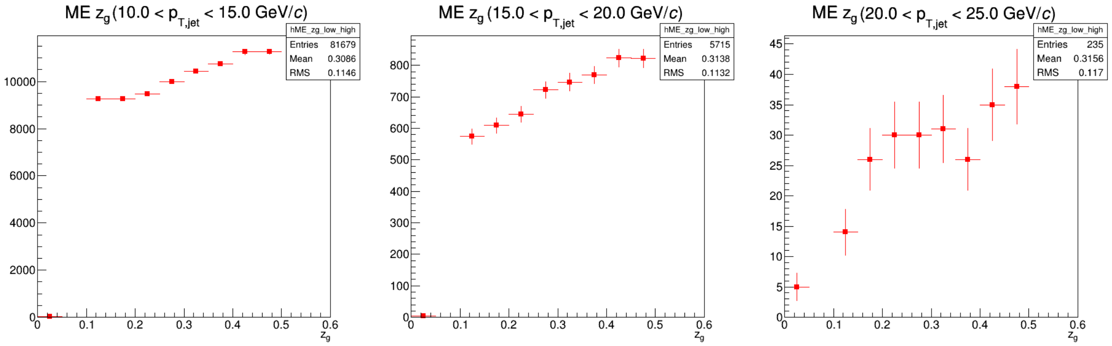
While the reconstructed p_{T} of the recoil jets are determined via area-based subtraction, in order to calculate z_{g} the jets are subtracted using Constituent Subtraction at the jet-level, as outlined in section 4.1. After subtracting, by comparing the ME and SE recoil jet p_{T} spectra, within any given jet p_{T} bin one can extract the expected combinatorial contribution to the SE distribution. The z_g distribution for SE in 0-10% central events can be seen in Figure 4.13, with the expected percentage contribution of combinatorial jets labelled as

"Fake %" for three jet $p_{\rm T}$ bins. The lowest jet $p_{\rm T}$ bin's $z_{\rm g}$ distribution is mostly flat, due to the large contamination of combinatorial jets. As the jet $p_{\rm T}$ increases, the fraction of uncorrelated signal drops drastically, with the highest jet $p_{\rm T}$ bin overwhelmingly dominated by signal jets.



(a) (b) (c)

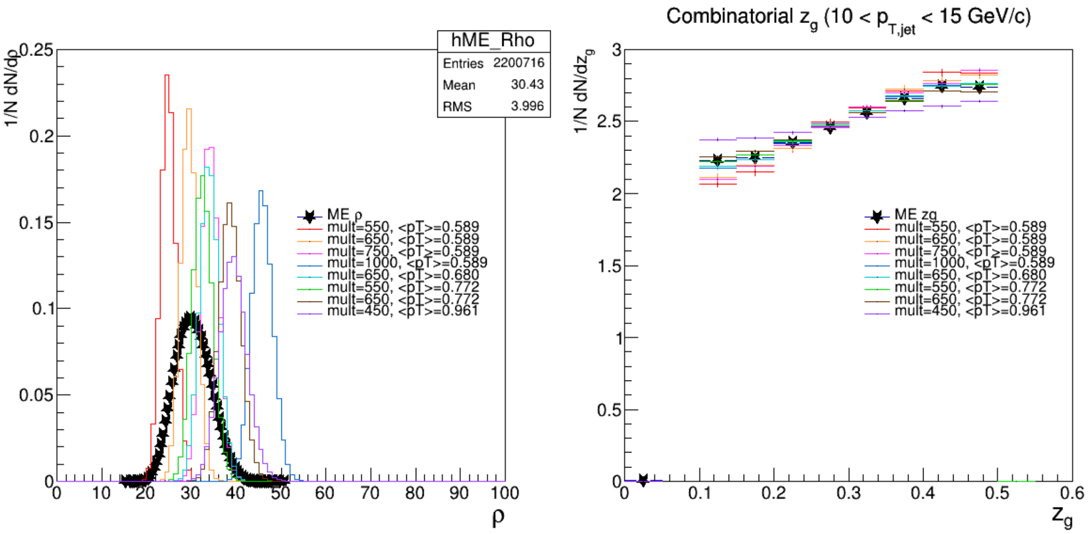
Figure 4.13: Measured $z_{\rm g}$ distributions for Same Event sample in 0-10% central events for three different jet $p_{\rm T}$ bins. Expected combinatorial contribution labelled for each bin.



(a) (b) (c)

Figure 4.14: Measured $z_{\rm g}$ distributions for Mixed Event sample in 0-10% central events for three different jet $p_{\rm T}$ bins.

The ME $z_{\rm g}$ distribution is shown in Figure 4.14 and has a drastically different shape than SE and the usual 1/z shape one expects measuring $z_{\rm g}$. The ME jets are entirely combinatorial, and applying SoftDrop and measuring $z_{\rm g}$ one finds that the reconstructed jets are largely symmetric. It is important to note the zero bin represents the amount of jets which never pass the SoftDrop criterion. A jet which never passes SoftDrop grooming is likely to have its p_{T} dominated by a single track, because to fail SoftDrop every single comparison between subjets within the clustering history must be highly asymmetric. While this occurs with some low probability for signal jets, it is exceedingly rare and practically zero for combinatorial jets to fail SoftDrop grooming. Asking for a combinatorial jet to not pass SoftDrop is in some way similar to asking for a combinatorial jet to have a high p_{T} leading track; it is unlikely to occur.



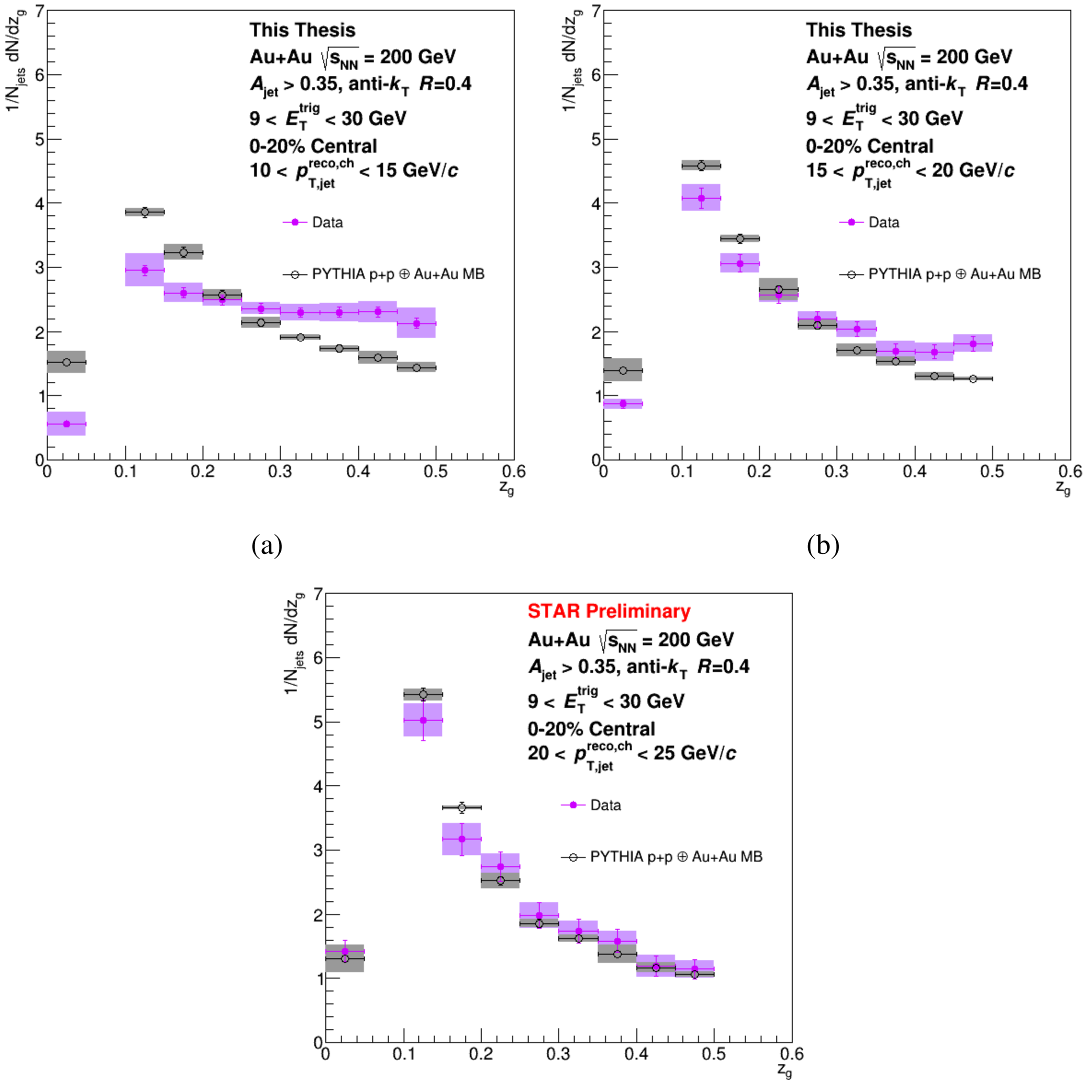
(a) (b)

Figure 4.15: Monte Carlo study investigating the dependence of the combinatorial z_{g} distribution on background characteristics. **(a)** Value of ρ for different multiplicity and mean p_{T} inputs. **(b)** Measured z_{g} distributions. Both are compared to the Mixed Event distributions shown as starred points.

The ME sample was designed to capture the behavior of the combinatorial jets contaminating the SE sample, however any possible differences in the shape of z_{g} of these combinatorial jets would be a source of systematic uncertainty, and thus was studied. To study the robustness of the combinatorial z_{g} distribution, a simple toy MC model was made by throwing particles randomly in ϕ and η, performing anti-k_t jet reconstruction, performing jet-level constituent subtraction, and measuring z_{g} of the upward fluctuations of the background. The p_{T} spectra and multiplicity of the randomly thrown particles were both varied to determine the sensitivity of the measured z_g distribution to the characteristics of the heavy-ion background. Sampled results of this study can be seen in Figure 4.15. While varying the multiplicity of the background had negligible effect on the measured z_g, the p_T spectra of the background particles had a minor but not significant effect. Because the tracks in ME are completely uncorrelated, the shape of the ME z_g distribution is driven largely by geometry, and the steepness of the slope has a slight sensitivity to the $\langle p_\mathrm{T} \rangle$ of the background. The ME z_g points for 0-10% central events exhibit the same shape as the distribution for the MC study, assuring that z_g of combinatorial jets is rather robust, and the ME distribution is suitable to model the underlying combinatorial distribution within the SE sample.

4.4 Results

Using this information we can apply an ensemble level subtraction to measure the correlated signal distribution of z_g. Within each jet p_T bin for 0-10% and 10-20% centrality, the ME z_g distribution was scaled by the extracted combinatorial contribution and then subtracted from the SE z_g distribution. While the subtraction is done in bins of centrality of width 10%, the final measured distribution is reported for 0-20% for statistical reasons, as a semi-inclusive measurement is extremely statistically limiting. In Figure 4.16, the combinatorial subtracted z_g distributions are shown for three jet p_T ranges and are compared to fast simulated PYTHIA embedded into minimum-bias Run14 data. Only the highest p_T bin is reported as a STAR preliminary, which is the $p_\mathrm{T,jet}$ least sensitive to combinatorial jets and the subtraction method. Before combinatorial subtraction, this $p_\mathrm{T,jet}$ bin contained 5% contamination for 0-10% and sub 1% for 10-20% events. When compared to smeared PYTHIA p+p there was found to be no modification of the z_g distribution in this bin. The two lower $p_\mathrm{T,jet}$ bins do show a change in distribution when comparing to the smeared



(a)

(b)

(c)

Figure 4.16: Measured combinatorial subtracted z_g distributions compared to fast simulated PYTHIA embedded into minimum-bias Au+Au Collisions for 0-20% centrality. Only highest jet p_T bin has been shown as a STAR preliminary.

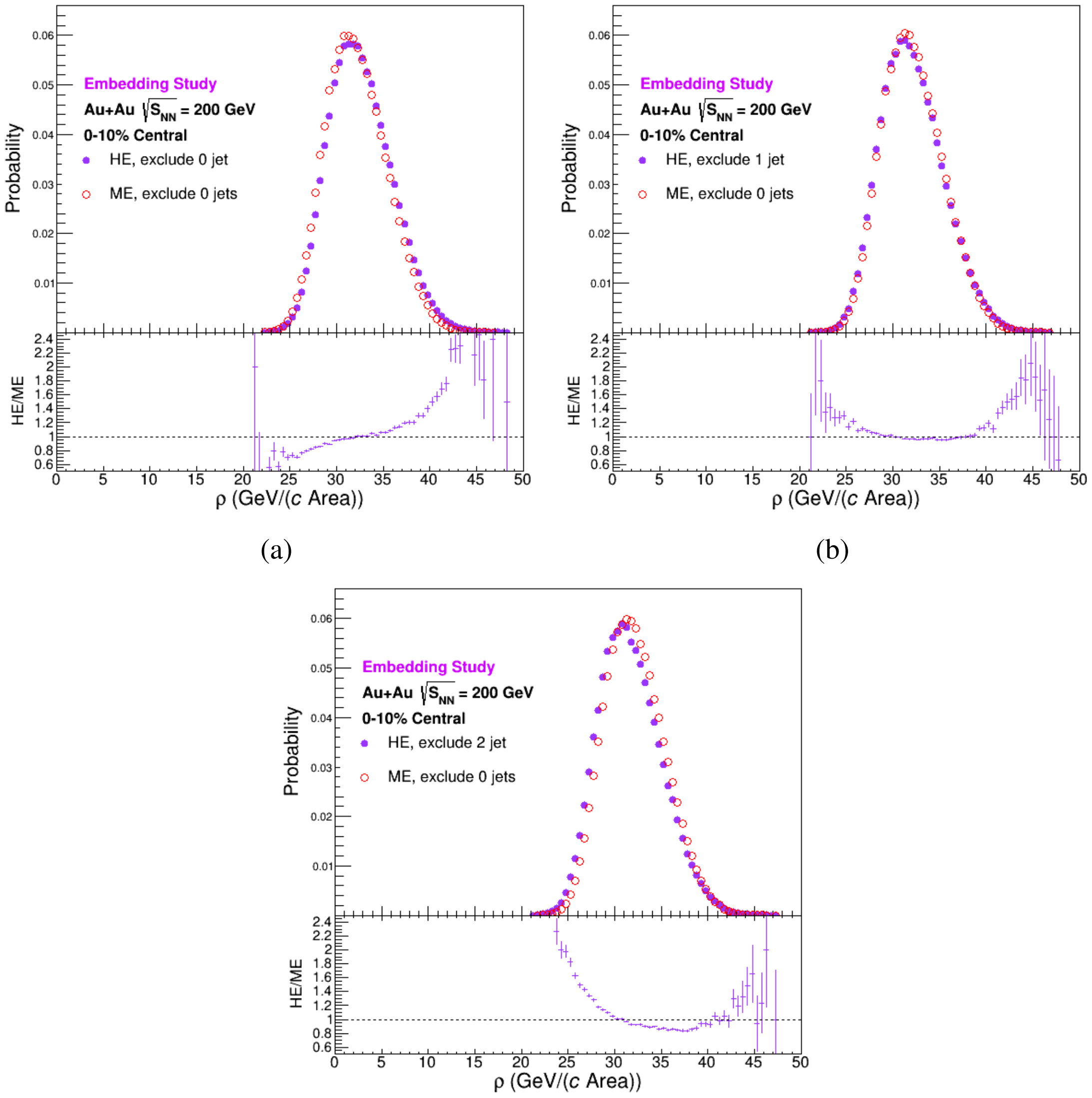
PYTHIA embedding, however it should be noted that these two bins are sensitive to the details of the ensemble-level combinatorial jet subtraction. While the systematic uncertainties due to variations of ρ and f_{ME} are expected to capture possible misestimations of the combinatorial jet contribution, the modification within these two bins is seemingly consistent with contamination due to combinatorial jets. Most notably, while the highest

$p_{\mathrm{T,jet}}$ bin has agreement with the z_{g} = 0 bin, there is a large difference for this bin in 10 < $p_{\mathrm{T,jet}}$< 15 GeV/c jets. As previously mentioned, combinatorial jets rarely fail Soft Drop, and contamination leftover from the ensemble-level subtraction would yield such a suppression in this bin. This behavior warranted extra investigation, and thus only the data for the high $p_{\mathrm{T,jet}}$ bin was requested and approved as a preliminary for STAR.

4.5 Further Study

In order to further investigate the performance of the combinatorial subtraction, a study was performed by embedding PYTHIA events into Mixed Events. This was done to probe the rate of contamination of combinatorial jets within a given $p_{\mathrm{T,jet}}$ bin in order to compare to. PYTHIA8 events were generated with the same high-p_{T} neutral requirement as in the Same Event class, requiring a photon or π^0 with $p_{\mathrm{T}} > 9$ GeV/c. These particle-level PYTHIA events were then embedded into Mixed Events (recreated with a different seed than the ME in the analysis), sampling the SE distributions in ZDC coincidence rate and event V_Z. No smearing was applied to the PYTHIA events because the study is testing the performance of the subtraction method as a whole and is not necessarily concerned with the details of Fast Simulation. These embedded events were named “Hybrid Events” and in the study replaced the Same Event class for the analysis steps. The same procedure for the analysis is followed, with each HE event sampling the corresponding ME bin.

First, the median-p_{T} density ρ was calculated for HE and ME events, with the choice of the number of hard jets excluded in the calculation explored. A comparison of the ρ distributions for 3 choices can be found in Figure 4.17. Just like in the Same Event case, the best agreement can be found with excluding a single hard jet in the ρ calculation for Hybrid Events. This is determined when the ratio under the peak of the distribution being flat and approximately agreeing with unity. The change in the choice of the number of



(a)

(b)

(c)

Figure 4.17: Median p_T-density ρ in Hybrid and Mixed events for different choices of the number of hardest jets excluding in the calculation for HE.

excluded jets in the ρ calculation has similar behavior to what was found in the actual analysis with the SE and ME ρ comparison. It should be noted that the structure of the enhancement at high ρ is in part due to the effect of embedding additional particles into the ME to create the HE, causing an increase into the max multiplicity of the events. Ultimately the same definition of ρ in the SE class was used for HE, excluding a single hard jet from the ρ calculation.

Following the ρ calculation, jets were reconstructed for HE and ME within the recoil region $|\phi_{\text{trig}} - \phi_{\text{jet}}| > \frac{3\pi}{4}$, where the trigger axis was not a high-E_{T} tower but instead the trigger photon or π^0 within the PYTHIA event. The jets were area-based subtracted, and the ME jets were scaled down by the f_{ME} scale factor as determined by the region dominated by combinatorial jets. The comparison between the HE and scaled ME per-trigger recoil jet yields can be found in Figure 4.18. The scale factor was found to be

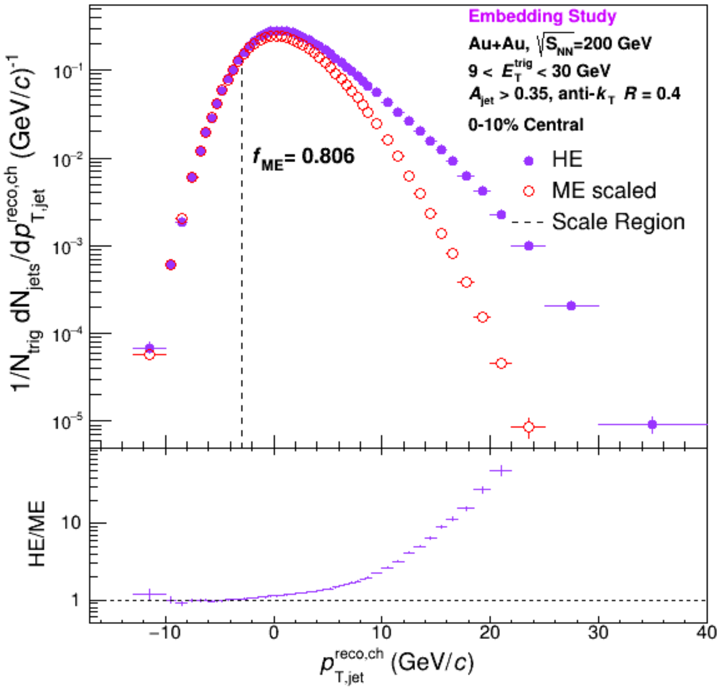


Figure 4.18: Per-trigger recoil jet yields for Hybrid and Mixed Events for 0-10% centrality. Mixed Event distribution scaled down by f_{ME}. The ratio of the distributions is shown in the bottom panel.

$f_{\text{ME}} = 0.806$, which is noticeably lower than f_{ME} as found in the analysis. This smaller f_{ME} is expected because no efficiency or smearing was applied to the PYTHIA events, causing an enhanced per-trigger yield of high-p_{T} jets than one would expect in actual data. Due to the integral of both the HE and ME (before scaling) being the same, an increase in per-trigger yields at the high $p_{\text{T,jet}}$ region for HE events would cause a decrease in the per-trigger yields at the scaling region. Even with the difference in the scaling factor, the gross features of the distributions remain the same, with the ratio in the combinatorial dominated region being flat at unity.

Using the per-trigger recoil yields for HE and the scaled ME, the combinatorial subtracted HE yields can be obtained in the same manner as in the analysis. While in the analysis there is no ability to check the purity of the subtracted yields, with this embedding study a comparison to the expected actual yields within a given $p_{\mathrm{T,jet}}$ bin can be explored. The jets reconstructed in the HE events are checked if they are matched to a PYTHIA particle-level reconstructed jet. The HE jets are considered matched if their jet axes are within $\Delta R < 0.4$ of a PYTHIA jet. However, it was found that there is a sensitivity to the minimum jet p_{T} used to cluster the PYTHIA jets. A comparison of the combinatorial subtracted per-trigger yields in the HE study to the amount of jets matched to PYTHIA with varying minimum PYTHIA $p_{\mathrm{T,jet}}$ can be found in Figure 4.19. When

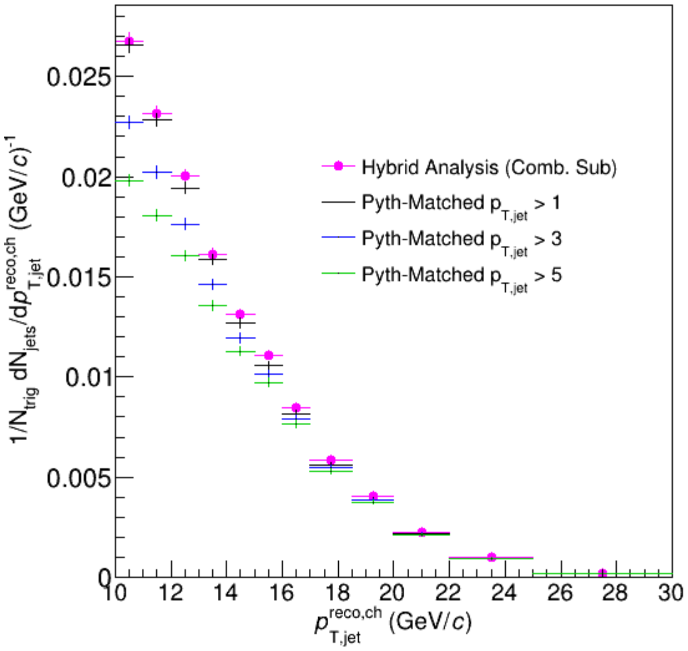


Figure 4.19: Comparison of the combinatorial subtracted per-trigger yields found in the Hybrid analysis to the per-trigger yields geometrically matched to PYTHIA jets, with varying minimum PYTHIA jet p_{T}.

matching to PYTHIA jets with $p_{\rm T} > 5$ GeV/c, it would seem as if a lot of the jets at low $p_{\rm T,jet}$ after the combinatorial subtraction are unmatched. A naive interpretation based on this matching requirement alone would be that the combinatorial subtraction is not completely subtracting the fake jets within these bins. However, when including matches to lower $p_{\rm T}$ PYTHIA jets, the fraction of matched jets increases. While such a low reconstructed $p_{\rm T,jet}$ is not very jet-like, it is still correlated to the trigger object and shows up as being not-combinatorial. It therefore seems possible that the per-trigger correlated yields in the $10 < p_{\rm T,jet}^{\rm reco} < 15$ GeV/c and $15 < p_{\rm T,jet}^{\rm reco} < 20$ GeV/c bins in Figure 4.16 likely have contributions from low $p_{\rm T}$ correlated signals, which behave like combinatorial jets when calculating their $z_{\rm g}$ because their $p_{\rm T}$ is dominated by the background. In other words, per-trigger yields after combinatorial subtraction may contain jets which are signal because they are correlated with the high-$p_{\rm T}$ trigger, but are of such low-$p_{\rm T,jet}$ that they are not meaningful to measure their $z_{\rm g}$ due to being dominated by the background fluctuations. The difference between the embedded PYTHIA and the data in the two lowest bins of Figure 4.16 is at least in part driven by the details of the background and the jet matching in Fast Simulation, and not necessarily the properties of the jets themselves. At higher $p_{\rm T,jet}$ this effect becomes diminished, the amount of correlated $p_{\rm T}$ required to reach higher $p_{\rm T,jet}$ bins increases. Only the highest $p_{\rm T,jet}$ bin, $20 < p_{\rm T,jet}^{\rm reco} < 25$ GeV/c, is high enough $p_{\rm T}$ such that it is insensitive to low-$p_{\rm T}$ signal contributions and a meaningful comparison to Fast Simulation can be made.

Chapter 5

Discussion and Outlook

One of the strengths of the semi-inclusive approach is minimally discriminating on a jet-by-jet basis which allows reconstructing jets to a near-zero $p_{\mathrm{T,jet}}$. However care has to be taken when interpreting the results of such low-p_{T} objects. Normally a measurement is unfolded to account for the fluctuation of the background, and jets are analyzed at meaningful $p_{\mathrm{T,jet}}$ values. While jet measurements can be made without unfolding and making comparisons to a smeared or embedded reference, they are usually done so with cuts that select biased but hard jets. The very lack of discrimination on a jet-by-jet basis, while ideal for the method, causes issues making comparisons without unfolding. The two low-$p_{\mathrm{T,jet}}$ bins in this z_{g} analysis likely have low-p_{T} contributions below the p_{T} threshold one would normally consider to be a jet. These contributions are not a flaw, they are correlated signal and the method treats them as such. However, it is not meaningful to measure substructure of such jets in $p_{\mathrm{T,jet}}^{\mathrm{reco}}$ bins that contain such contributions. In the end, only the highest $p_{\mathrm{T,jet}}^{\mathrm{reco}}$ bin in the analysis, $20 < p_{\mathrm{T,jet}}^{\mathrm{reco}} < 25$ GeV/c was released as a preliminary, which is high enough $p_{\mathrm{T,jet}}^{\mathrm{reco}}$ such that there is not a significant contribution of very low-p_{T} yield. In this bin, it was found there was no modification relative to the embedded and smeared PYHTIA p+p comparison. It is important to note that despite z_{g} not being modified, these jets are known to be quenched, as a previous measurement by STAR has shown that the semi-inclusive jet spectra in central Au+Au collisions to be highly suppressed [69]. This lack of modification of z_{g} is in agreement with the previous z_{g} measurement at STAR[42]. The measurement in this thesis presents a different selection of jets with different analysis steps and supports the same conclusion. It should be noted that both measurements are not unfolded and are compared to at the detector level, which leaves room for further measurements. However, as stated previously, z_{g} is extremely difficult to unfold due to the presence of non-diagonal elements in the response matrix. With current background subtraction techniques, it is not feasible to unfold without reducing the contribution from the background such as going to small jet R or increasing the z_{cut}, as was done in the ALICE measurement of z_{g}. While such a measurement would be interesting, restricting the phase space in order to be able to unfold causes the observable to deviate away from its intended purpose of trying to capture the kinematics of the first hard splitting.

The details of extending the semi-inclusive approach to jet substructure observables has been explored in this thesis. Using the semi-inclusive approach to measure a substructure observable is an attractive method, it selects a sample of jets while avoiding the normal surface and jet-biases that reduce the potential to find modification of jet substructure. While z_{g} is an interesting observable in its own right, for the sake of a semi-inclusive measurement, it is not ideal. The method requires an observable that can be unfolded back to the particle-level distributions in order to account for the minimal selection criteria imposed. It should be noted that a downside to the analysis method is that it is extremely statistics hungry, requiring a high-p_{T} trigger much greater than the online high tower trigger greatly reduces that amount of usable events. While in this thesis work there was no differentiation between photon and π^0 triggers, with sufficient statistics the difference between the triggers can be used to select different sample of jets with expected differences in path length in the medium. A possible future measurement using a semi-inclusive approach to measuring jet substructure will greatly benefit from the increased statistics expected from future STAR Au+Au runs in 2023 and 2025.

Appendix A

Supplementary Figures

In Chapter 4, most of the figures shown are for 0-10% central events. The same analysis steps were taken for 10-20% central events, and the corresponding figures can be found here.

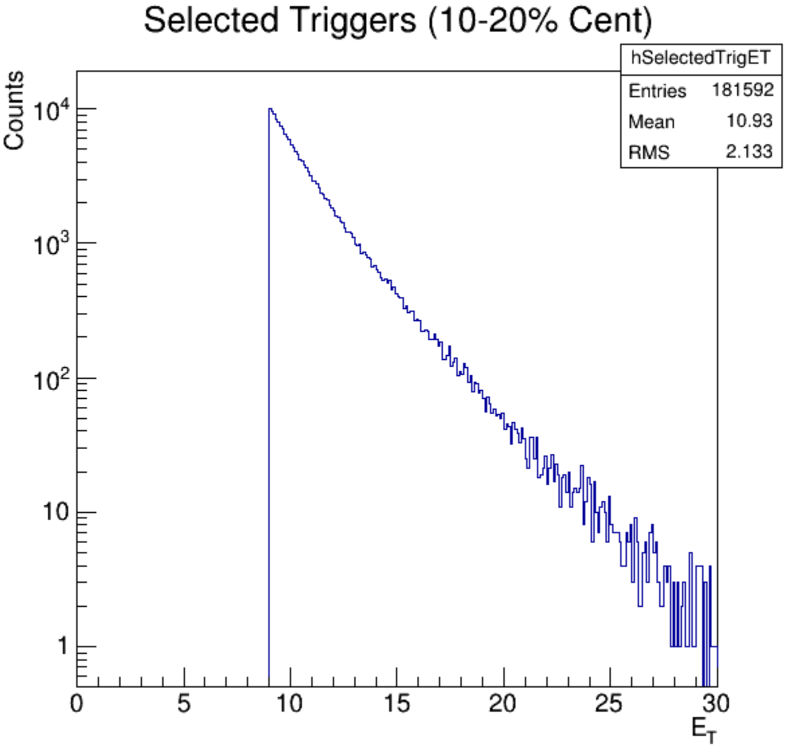


Figure A.1: Selected trigger tower $E_{\rm T}$ found in 10-20% central events.

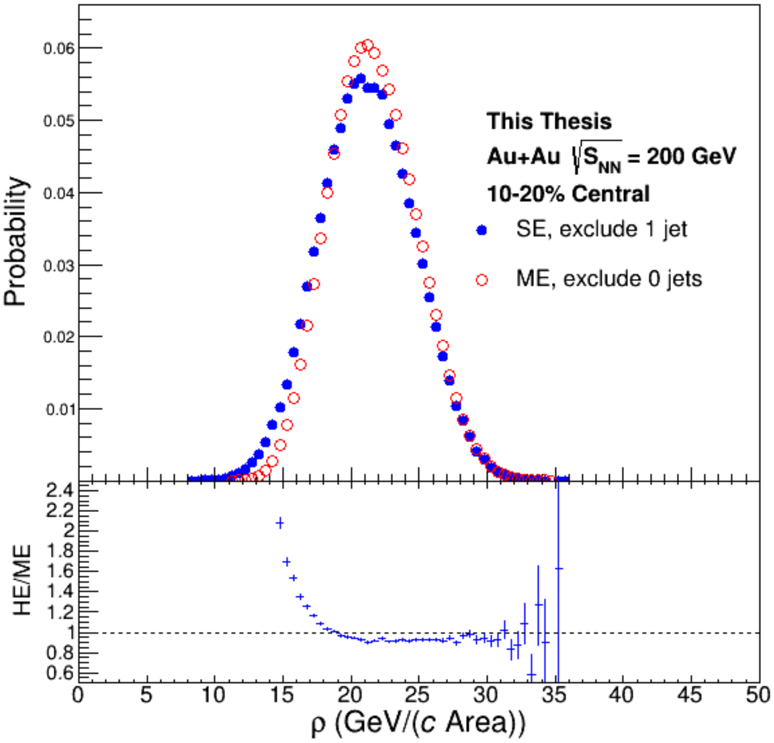


Figure A.2: Median p_{T} density ρ in SE and ME event classes in 10-20% central events, calculated excluding the hardest jet in SE and excluding no jets in ME. Bottom panel shows the ratio of the two distributions.

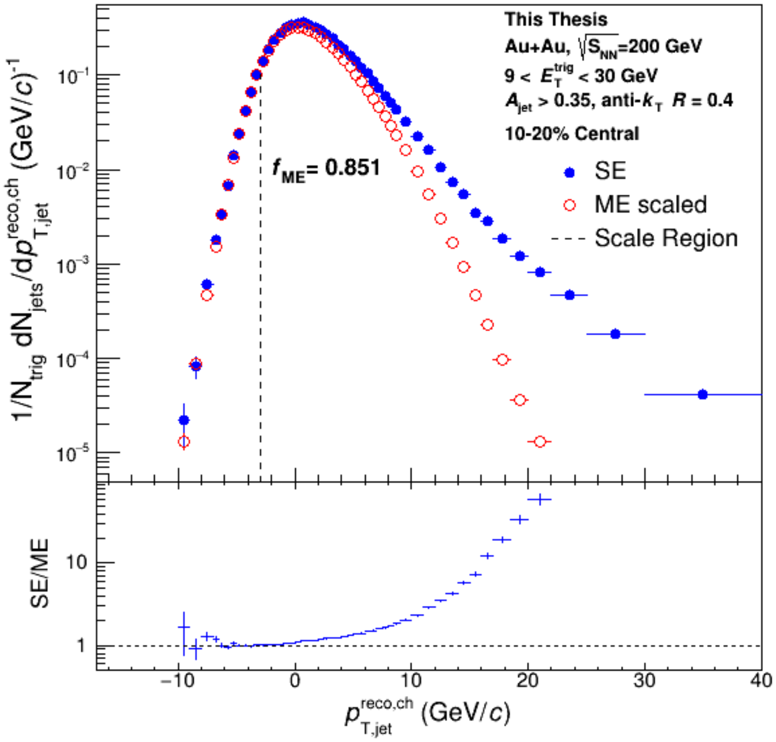


Figure A.3: Per-trigger recoil jet yields for SE and ME scaled down by a factor f_{ME} in 10-20% central events. Combinatorial jet yield not yet subtracted, which dominates left edge of both distributions. Dashed line denotes right edge of scaling region used to determine f_{ME} Bottom panel shows the ratio for SE/ME, in which the ratio increases at increasing jet p_{T}, where one expects to find a higher contribution of correlated yield compared to combinatorial yield.

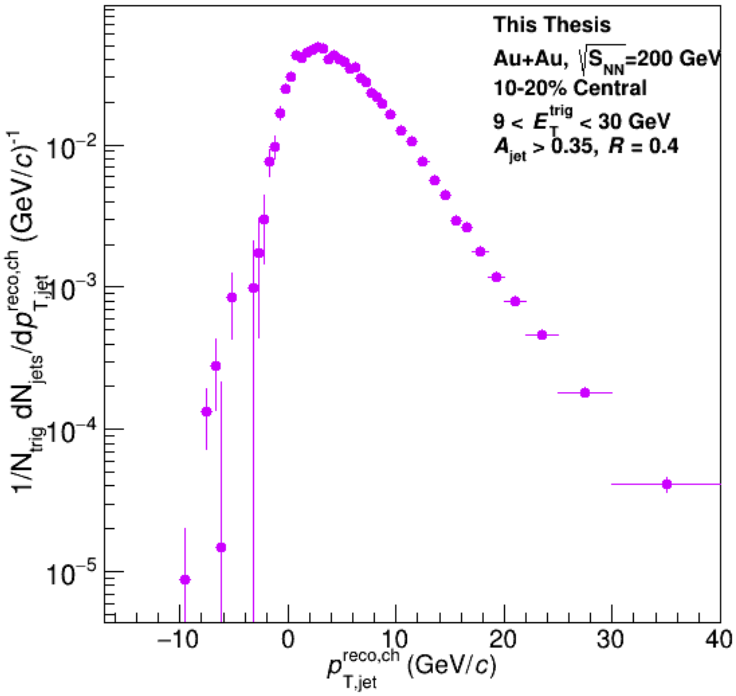
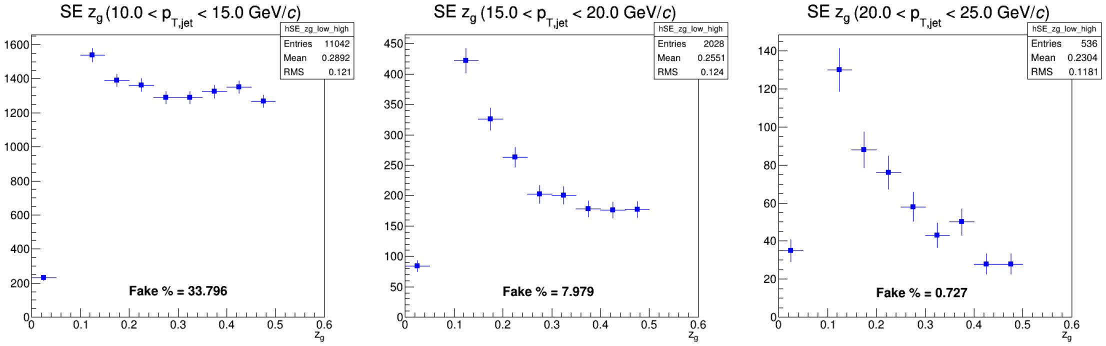
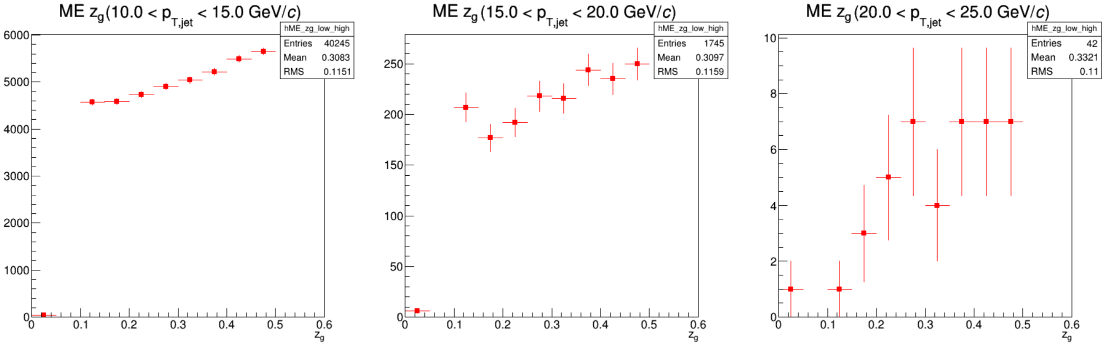


Figure A.4: Raw combinatorial-subtracted per-trigger recoil jet yields in 10-20% central events. Only statistical uncertainties are shown.



(a) (b) (c)

Figure A.5: Measured z_g distributions for Same Event sample in 10-20% central events for three different jet p_T bins. Expected combinatorial contribution labelled for each bin.



(a) (b) (c)

Figure A.6: Measured z_{g} distributions for Mixed Event sample in 10-20% central events for three different jet p_{T} bins.

Glossary

BBC Beam Beam Counter. 31

BEMC Barrel Electromagnetic Calorimeter. 28, 31, 32, 34, 36, 37, 41, 59

BHT2 Barrel High Tower Trigger 2. 33, 41, 42

HFT Heavy Flavor Tracker. 33

LHC Large Hadron Collider. 4, 7, 9, 10

ME Mixed Event, a class of artificial events created by mixing real minimum bias events. x–xii, 63–65, 67–74, 86

QCD Quantum Chromodynamics. 1, 4

QGP Quark Gluon Plasma. 4, 6, 7, 10, 12, 17, 21, 23

RHIC Relativistic Heavy Ion Collider, a heavy-ion collider located at Brookhaven National Lab. 7, 9, 10, 23, 24, 29, 32

SE Same Event, class of real events with a high-p_{T} trigger particle. x–xii, 63–65, 67–71, 73, 74, 86

STAR Solenoidal Tracker At RHIC. 11, 12, 23–25, 27, 31–34, 46, 47, 58, 74, 76, 81, 83

TPC Time Projection Chamber, STAR experiment's primary detector capable of reconstructed charged particle tracks. 25–29, 34, 36, 37, 41, 47

UE Underlying Event, the large background of particles from a heavy-ion collision which is not directly correlated with high-Q^2 processes. 53, 62, 70

VPD Vertex Position Detector. 30–32, 34

VPDMB30 STAR trigger component requiring the reconstructed primary vertex by the VPD detector to fall within ±30 cm of the center of the detector. 33, 41–45

ZDC Zero Degree Calorimeter. 30–32, 44, 45, 47

Bibliography

[1] S. Bethke. Experimental tests of asymptotic freedom. *Progress in Particle and Nuclear Physics*, 58(2):351–386, apr 2007. vi, 2

[2] U. Heinz and P. et al. Sorensen. Exploring the properties of the phases of QCD matter - research opportunities and priorities for the next decade, feb 2015. vi, 3

[3] P. R. Sorensen. Elliptic flow: A study of space-momentum correlations in relativistic nuclear collisions. *arXiv: Nuclear Experiment*, pages 323–374, jun 2009. vi, 5

[4] M.L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg. Glauber modeling in high-energy nuclear collisions. *Annual Review of Nuclear and Particle Science*, 57(1):205–243, nov 2007. vi, 6, 8, 11

[5] R. Snellings. Elliptic flow: a brief review. *New Journal of Physics*, 13(5):055008, may 2011. vi, 9

[6] PHENIX Collaboration. Measurements of higher order flow harmonics in $\mathrm{Au}+\mathrm{Au}$ collisions at $\sqrt{s_{NN}} = 200\ \mathrm{GeV}$. *Phys. Rev. Lett.*, 107:252301, dec 2011. vi, 9

[7] B. I. Abelev and Aggarwal et al. Longitudinal double-spin asymmetry and cross section for inclusive jet production in polarized proton collisions at $\sqrt{s}$=200 GeV. *Phys. Rev. Lett.*, 97:252001, dec 2006. vi, 10, 11

[8] J. Adams et al. Transverse-momentum and collision-energy dependence of high-p_T hadron suppression in $\mathrm{Au}+\mathrm{Au}$ collisions at ultrarelativistic energies. *Phys. Rev. Lett.*, 91:172302, oct 2003. vi, 11, 12

[9] K. Reygers (PHENIX Collaboration). High-pT direct-photon results from PHENIX. dec 2005. vii, 13

[10] ATLAS Collaboration. Measurement of jet fragmentation in Pb+Pb and pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV with the ATLAS detector. *Phys. Rev. C*, 98:024908, aug 2018. vii, 16

[11] A. J. Larkoski, S. Marzani, and J.Thaler. Sudakov safety in perturbative QCD. *Physical Review D*, 91(11), jun 2015. vii, 19

[12] Fu-Ming Liu and Sheng-Xu Liu. Quark-gluon plasma formation time and direct photons from heavy ion collisions. *Phys. Rev. C*, 89:034906, mar 2014. 4

[13] Laszlo P. Csernai, Joseph I. Kapusta, and Larry D. McLerran. Strongly interacting low-viscosity matter created in relativistic nuclear collisions. *Phys. Rev. Lett.*, 97:152303, oct 2006. 4

[14] W. Busza, K. Rajagopal, and W. van der Schee. Heavy ion collisions: The big picture and the big questions. *Annual Review of Nuclear and Particle Science*, 68(1):339–376, oct 2018. 4

[15] B. I. Abelev et al. (STAR Collaboration). Centrality dependence of charged hadron and strange hadron elliptic flow from $\sqrt{s_{NN}} = 200$ GeV Au+Au collisions. *Phys. Rev. C*, 77:054901, may 2008. 7

[16] K. Aamodt et al. (ALICE Collaboration). Elliptic flow of charged particles in pb-pb collisions at $\sqrt{s_{NN}} = 2.76\ \text{TeV}$. *Phys. Rev. Lett.*, 105:252302, dec 2010. 7

[17] G. Policastro, D. T. Son, and A. O. Starinets. Shear viscosity of strongly coupled $n = 4$ supersymmetric yang-mills plasma. *Phys. Rev. Lett.*, 87:081601, aug 2001. 10

[18] K. J. Eskola. Nearly perfect quark–gluon fluid. *Nature Physics*, 15(11):1111–1112, nov 2019. 10

[19] ATLAS Collaboration. Measurement of the inclusive jet cross-section in proton-proton collisions at 7 TeV using 4.5 fb-1 of data with the ATLAS detector. *JHEP*, 2015(2):153, feb 2015. 10

[20] Miklos Gyulassy and Michael Plümer. Jet quenching in dense matter. *Physics Letters B*, 243(4):432–438, jul 1990. 10

[21] K. C. Zapp, F. Krauss, and U. A. Wiedemann. A perturbative framework for jet quenching. *JHEP*, 2013(3), mar 2013. 10

[22] J. Adam et al. (STAR Collaboration). Measurement of inclusive charged-particle jet production in Au+Au collisions at 200 GeV. *Physical Review C*, 102(5), nov 2020. 10, 58

[23] V. Khachatryan et al. Charged-particle nuclear modification factors in PbPb and pPb collisions at 5.02 TeV. *JHEP*, 2017(4), apr 2017. 10, 11

[24] B. Abelev and J. Adam et al. Centrality dependence of charged particle production at large transverse momentum in Pb–Pb collisions at 2.76 TeV. *Physics Letters B*, 720(1):52–62, mar 2013. 11

[25] S. Chatrchyan and V. Khachatryan et al. Measurement of isolated photon production in pp and PbPb collisions at 2.76 TeV. *Physics Letters B*, 710(2):256–277, apr 2012. 13

[26] G. P. Salam and G. Soyez. A practical seedless infrared-safe cone jet algorithm. *JHEP*, 2007(05):086, may 2007. 14

[27] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *The European Physical Journal C*, 72(3), mar 2012. 14

[28] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 2008(04):063, apr 2008. 15

[29] Jorge Casalderrey-Solana and Doga Can Gulhan et al. Angular structure of jet quenching within a hybrid strong/weak coupling model. *JHEP*, 2017(3), mar 2017. 15

[30] S. Chatrchyan and V. Khachatryan et al. Measurement of jet fragmentation in PbPb and pp collisions at 2.76 TeV. *Physical Review C*, 90(2), aug 2014. 16

[31] M. Spousta and B. Cole. Interpreting single jet measurements in Pb+Pb collisions at the LHC. *The European Physical Journal C*, 76(2):50, jan 2016. 16

[32] ATLAS Collaboration. Comparison of inclusive and photon-tagged jet suppression in 5.02 TeV Pb+Pb collisions with ATLAS, mar 2023. 17

[33] A. M. Sirunyan et al. (CMS Collaboration). Observation of medium-induced modifications of jet fragmentation in pb-pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02 \text{ TeV}$ using isolated photon-tagged jets. *Phys. Rev. Lett.*, 121:242301, dec 2018. 17

[34] D. Krohn, J.Thaler, and L. Wang. Jet trimming. *JHEP*, 2010(2), feb 2010. 17

[35] S. D. Ellis, C. K. Vermilion, and J. R .Walsh. Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches. *Phys. Rev. D*, 81:094023, may 2010. 17

[36] S. Acharya et al. (Alice Collaboration). First measurements of N-subjettiness in central Pb-Pb collisions at 2.76 TeV. *JHEP*, 2021(10), oct 2021. 17

[37] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam. Towards an understanding of jet substructure. *JHEP*, 2013(9), sep 2013. 17

[38] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler. Soft drop. *JHEP*, 2014(5), may 2014. 17

[39] A.J. Larkoski and J. Thaler. Unsafe but calculable: ratios of angularities in perturbative QCD. *JHEP*, 2013(9):137, sep 2013. 19

[40] Vladimir Naumovich Gribov and L N Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15(4):438–450, 1972. 20

[41] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298–318, aug 1977. 20

[42] J. Adam and L. Adamczyk et al. Measurement of groomed jet substructure observables in p+p collisions at 200 GeV with STAR. *Physics Letters B*, 811:135846, dec 2020. 20, 82

[43] S. Acharya and Adamov et al. Measurement of the groomed jet radius and momentum splitting fraction in pp and Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02\ \text{TeV}$. *Phys. Rev. Lett.*, 128:102001, mar 2022. 20, 21, 57

[44] Y. Mehtar-Tani and K. Tywoniuk. Groomed jets in heavy-ion collisions: sensitivity to medium-induced bremsstrahlung. *JHEP*, 2017(4), apr 2017. 21

[45] A. M. Sirunyan and A. et al Tumasyan. Measurement of the splitting function in pp and pb-pb collisions at $\sqrt{s_{NN}} = 5.02\ \text{TeV}$. *Phys. Rev. Lett.*, 120:142302, apr 2018. 21

[46] K. Kauder. Measurement of the shared momentum fraction zg using jet reconstruction in p+p and Au+Au collisions with STAR. *Nuclear Physics A*, 967:516–519, 2017. The 26th International Conference on Ultra-relativistic Nucleus-Nucleus Collisions: Quark Matter 2017. 21

[47] H. Hahn et al. The RHIC design overview. *Nucl. Instrum. and Meth. A*, 499(2):245–263, mar 2003. 23

[48] J. Adams et al. (STAR Collaboration). Experimental and theoretical challenges in the search for the quark–gluon plasma: The STAR collaboration's critical assessment of the evidence from RHIC collisions. *Nuclear Physics A*, 757(1-2):102–183, aug 2005. 24

[49] F. Bergsma and C.O. Blyth et al. The STAR detector magnet subsystem. *Nucl. Instrum. and Meth. A*, 499(2):633–639, mar 2003. 25

[50] M. Anderson et al. The STAR time projection chamber: a unique tool for studying high multiplicity events at RHIC. *Nucl. Instrum. and Meth. A*, 499(2-3):659–678, mar 2003. 25

[51] M. Beddo et al. The STAR barrel electromagnetic calorimeter. *Nucl. Instrum. and Meth. A*, 499(2):725–739, 2003. 28

[52] T.M Cormier, A.I Pavlinov, M.V Rykov, V.L Rykov, and K.E Shestermanov. STAR barrel electromagnetic calorimeter absolute calibration using "minimum ionizing particles" from collisions at RHIC. *Nucl. Instrum. and Meth. A*, 483(3):734–746, may 2002. 29

[53] C. Adler and A. Denisov et al. The RHIC zero degree calorimeters. *Nucl. Instrum. and Meth. A*, 470(3):488–499, sep 2001. 30

[54] A. J. Baltz, C. Chasman, and S. N. White. Correlated forward–backward dissociation and neutron spectra as a luminosity monitor in heavy-ion colliders. *Nucl. Instrum. and Meth. A*, 417(1):1–8, nov 1998. 30

[55] W.J. Llope et al. The STAR vertex position detector. *Nucl. Instrum.and Meth. A*, 759:23–28, sep 2014. 30

[56] C. A. Whitten. The Beam-Beam Counter: A local polarimeter at STAR. *AIP Conference Proceedings*, 980(1):390–396, feb 2008. 31

[57] J. Adams and A. Ewigleben et al. The STAR event plane detector. *Nucl. Instrum. and Meth. A*, 968:163970, jul 2020. 31

[58] F.S. Bieser and H.J. Crawford et al. The STAR trigger. *Nucl. Instrum. and Meth. A*, 499(2):766–777, mar 2003. 31

[59] J.M. Landgraf and M.J. LeVine et. al. An overview of the STAR DAQ system. *Nucl. Instrum.and Meth. A*, 499(2):762–765, mar 2003. 32

[60] B. G. Gibbard and T. G. Throwe. The RHIC computing facility. *Nucl. Instrum. and Meth. A*, 499(2):814–818, mar 2003. 32

[61] G. Contin and L. Greiner et al. The STAR MAPS-based PiXeL detector. *Nucl. Instrum. and Meth. A*, 907:60–80, nov 2018. 33

[62] Fisyak Yuri. Tracking in STAR, feb 2000. Third Tracking Workshop, FIAS. 34

[63] S. Agostinelli et al. Geant4—a simulation toolkit. *Nucl. Instrum.and Meth. A*, 506(3):250–303, jul 2003. 45, 47

[64] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 2006(05):026–026, may 2006. 45

[65] A. Höcker and V. Kartvelishvili. SVD approach to data unfolding. *Nucl. Instrum. and Meth. A*, 372(3):469–481, apr 1996. 55

[66] D. W. Miller P. Berta, M. Spousta and R. Leitner. Particle-level pileup subtraction for jets and jet shapes. *JHEP*, 2014(6), jun 2014. 55

[67] P. Berta, L. Masetti, D. W. Miller, and M. Spousta. Pileup and underlying event mitigation with iterative constituent subtraction. *JHEP*, 2019(8):175, aug 2019. 56

[68] L. Adamczyk et al. (STAR Collaboration). Dijet imbalance measurements in STAR. *Physical Review Letters*, 119(6), aug 2017. 58

[69] L. Adamczyk et al. (STAR Collaboration). Measurements of jet quenching with semi-inclusive hadron + jet distributions in Au+Au collisions at 200 GeV. *Physical Review C*, 96(2), 2017. 59, 82

[70] J. Adam et al. (Alice Collaboration). Measurement of jet quenching with semi-inclusive hadron-jet distributions in central Pb-Pb collisions at 2.76 TeV. *JHEP*, 2015(9), sep 2015. 59