Inclusive jet production in ultrarelativistic proton-nucleus collisions

Dennis V. Perepelitsa

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Inclusive jet production in ultrarelativistic proton-nucleus collisions

Dennis V. Perepelitsa

High-p_{T} processes in proton- and deuteron-nucleus collisions at TeV energies are the best presently available way to study the partonic structure of the nucleus in a high-density regime. Jet production over a wide range of phase space can significantly constrain the current knowledge of nuclear parton distribution functions (nPDFs), which are substantially less well understood than the corresponding PDFs in protons and which have only recently begun to be treated in a spatially-dependent way. An accurate knowledge of nPDFs is crucial for a definitive control of perturbative processes in a cold nuclear environment, since high-p_{T} probes are used to quantitatively investigate the hot QCD matter created in ultrarelativistic nucleus-nucleus collisions. Furthermore, jets from low Bjorken-x partons can probe the transition from the dilute to saturated nuclear regimes.

Jet production is investigated in d+Au collisions at $\sqrt{s} = 200$ GeV with the PHENIX detector at the Relativistic Heavy Ion Collider (RHIC), and in p+Pb collisions at $\sqrt{s} = 5.02$ TeV with the ATLAS detector at the Large Hadron Collider (LHC). The measurements shown here utilize $\int \mathcal{L}dt = 23$ nb^{-1} and 0.2 pb^{-1} of 200 GeV d+Au and pp data, respectively, recorded in 2007-8 at RHIC and $\int \mathcal{L}dt = 31$ nb^{-1} and 4.1 pb^{-1} of 5.02 TeV p+Pb and 2.76 TeV pp data, respectively, recorded in 2013 at the LHC. Jets are reconstructed using the $\sigma = 0.3$ Gaussian filter and $R = 0.4, 0.6$ anti-k_{T} algorithms.

Inclusive, centrality-dependent jet yields within $|\eta| < 0.35$ and 10 GeV $< p_{\mathrm{T}} <$ 40 GeV in 200 GeV d+Au and pp collisions are presented. The jet yield in d+Au collisions relative to the geometric expectation is found to be slightly suppressed (≈ 0.9) in central events and moderately enhanced (≈ 1.3) in peripheral events, with no modification when averaged over all d+Au events. Separately, inclusive, centrality-dependent jet yields within $|y^*| < 4.4$ and 25 GeV $< p_{\mathrm{T}} <$ 800 GeV in 5.02 TeV p+Pb and 2.76 TeV pp collisions are presented. The event centrality in p+Pb

collisions is determined by the sum of the transverse energy in the Pb-going forward calorimeter, $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, and the mean number of participating nucleons $\langle N_{\mathrm{part}} \rangle$ is estimated using the Glauber and Glauber-Gribov models of semiclassical p+A collisions. The jet yield in p+Pb collisions relative to the geometric expectation is found to be suppressed in central events and enhanced in peripheral events. The modifications are found to be stronger at higher-p_{T} and at more forward (downstream proton) rapidities. Furthermore, it is observed that for each centrality selection, the modification at all rapidities is consistent with a simple function of the total jet energy only. The implications of the results are discussed, including a comparison of the modifications between RHIC and LHC energies.

Table of Contents

List of Figures

List of Tables

Acknowledgments

The sin of my ingratitude even now
Was heavy on me. Thou art so far before
That swiftest wing of recompense is slow
To overtake thee. Would thou hadst less deserved,
That the proportion both of thanks and payment
Might have been mine! Only I have left to say,
More is thy due than more than all can pay.

Macbeth, Act I, Scene IV

First and foremost I must thank my advisor Brian Cole, who taught me to think clearly and precisely about experimental methodology, to always keep the motivation for a measurement in the back of my mind, and to press on when circumstances seem difficult or even insurmountable. Brian leads by example. I greatly appreciate his enthusiasm for the physics, his indefatiguable energy, and his patience for me as a student.

I am very grateful to the current and former members Columbia experimental heavy ion group, including Aaron Angerami, Bill Zajc, Eric Vazquez, Ali Hanks, Nathan Grau and David Winter, as well as to the Columbia University Physics Department staff.

During my work on PHENIX and ATLAS, I have found myself benefiting from the physics intuition, experimental wisdom and personal advice of many scientists I hold in the highest esteem, including Barbara Jacak, Justin Frantz, Gabor David, Anne Sickles, Mike Tanenbaum, Mike McCumber, Jamie Nagle and Alan Dion on PHENIX and Peter Steinberg, Sasha Milov, Witold Kozanecki, Zvi Citron, Oldrich Kepka and Will Brooks on ATLAS. Thank you.

My time in graduate school was brightened by the presence of many colleagues, fellow students, and new and old friends. I offer my deep and sincere gratitude –

to Luke Goetzke, Chris Osborn, AAron Veicht as well as Arthur Franke and Pontus Alqhvist, for their advice, friendship, and help with homework and qualifying exams; to Martin Rybar and Martin Spousta, who welcomed me to the ATLAS jet group; to Martin Purschke, who was one of the first friendly faces on PHENIX I met when I started working on the Event Builder as one of my first projects in graduate school; to Javier Duarte, for showing me around parts of Geneva;

to Alex, Maddie and Frank for putting up with me during the most intensive writing period; to Grace, for so much more than just my promise to include her here; to Andrew, who I hope will not be surprised to learn that he is an inspiration to me; to Sonali, who helped me navigate my early life in NYC in several respects; to Fred, Halley, and Lauren, as well as Kennelly and Sam, for their open and unassuming (in the best sense) friendship; to Alice, Birdy, Michelle, and the many interesting people I met in the Columbia University Players and King's Crowne Theater Troupe; to David and Kevin; to Lauren and Jonathan, who I was lucky to meet when I still had the time and energy for theater;

to Brian (and our friend Duck-Cat) – we made it, buddy, we sure showed those spin echoes!; to my dear friends back in Boston; to Clayton, Erik, Jim, Ternus, Yelena, Joy, Ken, Eli, Jacky, Ash and Diandra, in no particular order, who are close to me even though we are all increasingly scattered and busy; you all deserve paragraphs of your own; to Kathleen, who understands me like few others do; to Greg, who brightens any room he is in; to the new generation of MIT students I had the privilege to get to know but never overlapped with, including Lauren ("Canada"), Karen, Ducky Ben, and others; to Dr. Anna, of course; to Charles, who has been a constant in my life; to my roommate Chris, who was probably a little more lenient with me than he should have been; to Charlie from York, and Tom, both of whom are kind and gracious souls;

to the many colorful people who made my life a little more interesting who are only through oversight not included here; and to everybody who showed me kindness and patience and support, especially during the last several months.

But I must especially acknowledge my family. I am owe a deep gratitude to my mother, who is a moral center in my life and has been my unconditional supporter; to my brother, who is one of the sharper, more interesting and life-affirming people I know; to my sisters, who are bright and creative young women; and to my stepdad Chris, who has been a stalwart in my family.

Thank you all. You have affected me in ways that are beyond my ability to express. Every day

I am aware of what your warmth, courage and humanity have done for me. I hope that I, at least in part, have done the same for you.

"Te amo como se aman ciertas cosas oscuras,
secretamente, entre la sombra y el alma."

– *Pablo Neruda*

for Christina

Chapter 1

Introduction

Once more unto the breach, dear friends, once more;
Or close the wall up with our English dead.
In peace there's nothing so becomes a man
As modest stillness and humility:
But when the blast of war blows in our ears,
Then imitate the action of the tiger;
Stiffen the sinews, summon up the blood,
Disguise fair nature with hard-favour'd rage;

Henry V, Act III, Scene I

Quantum chromodynamics (QCD) is a theory based on a remarkably small number of principles such as local gauge invariance and the preservation of a few key symmetries. Despite this, the emergent phenomena of the theory, such as confinement and dynamical chiral symmetry breaking, are rich, complex and not obvious from a first-principles reading of the QCD Lagrangian. Compellingly, QCD can be studied under the extreme conditions of high temperature and density through the use of ultrarelativistic hadronic and nuclear collisions. In the present era, proton-nucleus (p+A) collisions at RHIC and the LHC are a promising laboratory in which to study conventional and novel QCD effects in a high parton density environment.

At high momentum transfer (Q^2), measurements of jets in p+A (and, similarly, deuteron-nucleus, or d+A) collisions serve as a test of perturbative QCD (pQCD) approaches based on

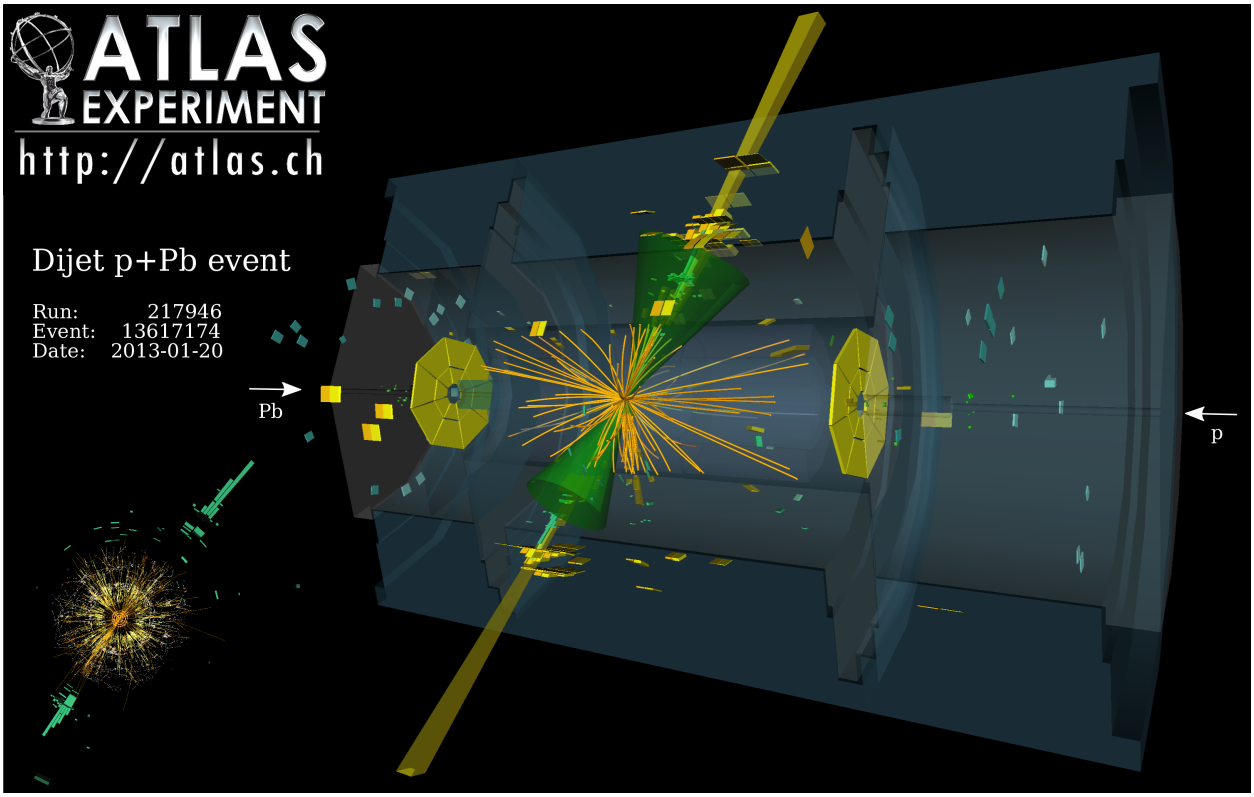


Figure 1.1: Event display of a high-p_{T} dijet event recorded by ATLAS during the 2013 $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV p+Pb data taking at the LHC.

collinear factorization. In this framework, measurements of inclusive jets in a large rapidity and transverse momentum range can significantly improve knowledge of nuclear parton distribution functions (nPDFs) over a wide kinematic (x, Q^2) space. nPDF's are known with much less accuracy than their counterpart in protons, with large uncertainties in the gluon PDFs and tension between nPDF sets in several places. A careful study of jet production as a function of the colliding proton's position in the nucleus can even provide information about the impact-parameter dependence of nPDF modification.

At lower, more intermediate Q^2, jets are sensitive to the higher twist corrections to hard processes which are enhanced by the presence of the nuclear medium. These cold nuclear matter effects include initial and final state energy loss of the hard scattered partons, nuclear shadowing of low-x partons and transverse momentum broadening of the parton towards mid-rapidity. Furthermore,

hard probes in p+A collisions provide information about the initial state in heavy ion (A+A) collisions, which needs to be understood separately and systematically to give proper context to the dramatic final state effects attributed to the formation of the deconfined quark-gluon plasma (QGP). As the phenomenon of jet quenching moves from the discovery to the measurement stage at the LHC, quantitative constraints on the cold nuclear effects before formation of the QGP become crucial.

At sufficiently low-Q^2 (and Bjorken-x), jets may even be sensitive to beginning of nonlinear changes in the partonic structure of the nucleus. At low x, the proton wavefunction is characterized by a steep rise in the gluon parton distribution function (PDF). Since a naive application of the QCD evolution equations implies that the cross-section at high-energy should grow so large as to violate unitarity bounds, it is generally thought that the gluon densities cannot increase without limit and non-linear dynamics must become important. When probed at higher and higher energy (lower x), the number density of gluons increases until they begin to overlap in phase space. Here, recombination effects are expected to become important, resulting in a saturation of the gluon density at some scale $Q^2_{\rm sat}$. In heavy nuclei, the longitudinal overlap of gluons in neighboring nucleons is expected to increase the saturation scale by a factor $\approx A^{1/3}$.

These nuclear effects can be best explored in a thorough and systematic way through measurements of fully reconstructed jets. Only the large kinematic reach of full jets can probe the transition from a dilute nuclear regime to a high-density regime of possibly novel QCD effects in such detail. Since the development of background subtraction procedures in hadronic jet reconstruction and related jet technology, modern jet algorithms have become a standard tool in relativistic nuclear physics. Thus, the subject of this dissertation is an exploratory measurement of jet production intended to map out the partonic structure of the heavy nucleus. It is the first centrality-dependent measurement of jet yields in d+A collisions at RHIC and the first such triple-differential measurement (in centrality, rapidity and transverse momentum) in p+A collisions at the LHC.

This dissertation is organized as follows. Chapter 2 describes the theoretical and experimental background on QCD, jets and the role of the cold nuclear medium. Chapter 3 describes the modern design of hadron colliders, including RHIC and the LHC. Section 4 describes the instrumentation of the PHENIX detector at RHIC. Chapter 5 describes the instrumentation of the ATLAS detector at the LHC. Section 6 describes the measurement of full jets in d+Au collisions at RHIC. Section 7

describes the determination of the p+Pb event centrality and extraction of geometric quantities. Section 8 describes the measurement of full jets in p+Pb collisions at the LHC. Section 9 discusses the physics implications of the results.

Chapter 2

High-energy Nuclear Collisions

Either forbear,
Quit presently the chapel, or resolve you
For more amazement. If you can behold it,
I'll make the statue move indeed, descend
And take you by the hand; but then you'll think–
Which I protest against–I am assisted
By wicked powers.

Winter's Tale, Act V, Scene III

2.1 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the Yang-Mills quantum field theory with gauge group SU(3)[1; 2; 3; 4], with Lagrangian density given by

$$\mathcal{L}_{QCD} = -\frac{1}{4} F^C_{\mu\nu} F^{\mu\nu}_C + \sum_f \bar{\psi}_f \left(i\gamma^\mu D_\mu - m_f \right) \psi_f \tag{2.1}$$

where C is the color index that runs from $C = 1$ to $N_C^2 - 1 = 8$, corresponding to the dimensionality of SU(3), the index f runs over the fermions (in the Standard Model, these are the six quark flavors) with bare mass m_f and Dirac spinor ψ_f, and γ^μ are the four Dirac matrices. The fermion fields have a color index $\psi_f^{(A)}$ that runs from $A = 1$ to $N_C = 3$ but it is suppressed in the

following notation for brevity. The field tensor $F^C_{\mu\nu}$ is related to the eight gauge fields A^C_μ, called *gluon* fields, through

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g f_{ABC} A^B_\mu A^C_\nu \tag{2.2}$$

where f_{ABC} are the structure constants of SU(3) defined by $[t_A, t_B] = i f_{ABC} t_C$ for the eight 3×3 generators t_A of SU(3), and g is the QCD coupling constant (more commonly referred to by $\alpha_s = g^2/4\pi$. The covariant derivative is defined as

$$D_\mu = \partial_\mu - igt^C A^C_\mu \tag{2.3}$$

The gluon field in the covariant derivative reflects local gauge symmetry. Consider an SU(3) color transformation

$$U = e^{i\epsilon} = \exp\left(i \sum_C \epsilon_C t^C\right) \tag{2.4}$$

The fermion fields are fundamental representations of SU(3) and transform like

$$\psi_f \to e^{i\epsilon}\psi_f \tag{2.5}$$

$$\bar{\psi}_f \to \bar{\psi}_f e^{-i\epsilon} \tag{2.6}$$

Then the quantity $\bar{\psi}_f \left(\partial_\mu - m\right) \psi_f$ (the non-covariant derivative) is not gauge invariant by itself, since

$$\bar{\psi}_f \left(\partial_\mu - m\right) \psi_f \to \bar{\psi}_f \left(\partial_\mu - m\right) \psi_f + i\bar{\psi}_f (\partial_\mu \epsilon)\psi_f \tag{2.7}$$

But the gauge field transforms as

$$A_\mu \to A_\mu + \frac{i}{g}\left(\partial_\mu \epsilon\right) \tag{2.8}$$

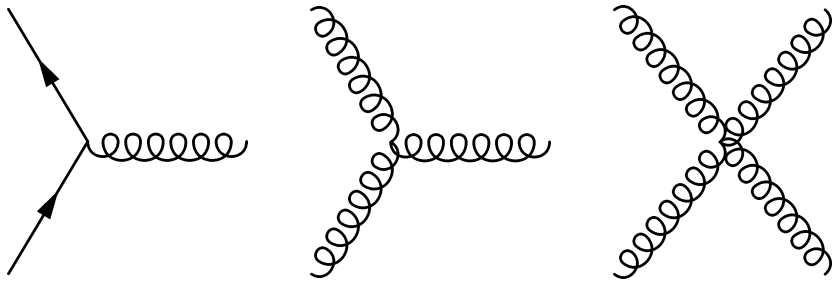


Figure 2.1: Diagrams corresponding to the three terms in the QCD $\mathcal{L}_{int}$.

and thus the quantity $\bar{\psi}_f D_\mu \psi_f$ is gauge invariant. Since a mass term for the gauge bosons ($m_g^2 A_C^\mu A_\mu^C$) would violate gauge invariance, the gluons are massless. This is true for the U(1) theory of QED as well and is reflected in the fact that photons are massless. (It would be true for the full SU(2)×U(1) electroweak Lagrangian as well, but for the presence of the Higgs field and the resulting spontaneous symmetry breaking at low temperatures, which gives masses to the $W^\pm$ and Z bosons.)

However, it is the non-abelian nature of the SU(3) gauge group that will prove to have important consequences for the theory and distinguish it from the U(1) theory of QED in a number of ways, as we will see when renormalizing the theory in Section 2.1.1.

Writing out the terms in Equation 2.1, we can decompose $\mathcal{L}_{QCD} = \mathcal{L}_0 + \mathcal{L}_{int}$, where the free field Lagrangian is

$$\mathcal{L}_0 = \sum_f \bar{\psi}_f \left(i\gamma^\mu \partial_\mu - m_f\right) \psi_f - \frac{1}{2} \sum_C (\partial_\mu A_\nu^C)(\partial^\mu A_C^\nu) - (\partial_\mu A_\nu^C)(\partial^\nu A_C^\mu) \tag{2.9}$$

where the first term gives rise to the $N_f = 6$ fermion propagators and the second term gives rise to the $N_C^2 - 1 = 8$ gluon propagators. The interaction Lagrangian is

$$\mathcal{L}_{int} = \sum_f g A_\mu^C \bar{\psi}_f \gamma^\mu t^C \psi_f - g f_{ABC} A_B^\mu A_C^\nu \left(\partial_\mu A_\nu^A\right) - \frac{1}{4} g^2 \left(f^{ABC} A_B^\mu A_C^\nu\right) \left(f_{ADE} A_\mu^D A_\nu^E\right) \tag{2.10}$$

where the first $gA\bar{\psi}\psi$ term is a fermion-gauge boson vertex, the second $gAA\partial A$ term is apparently a three gauge boson vertex and the third g^2AAAA term is a four gauge boson vertex. The Feynman diagrams for these are shown in Figure 2.1.

Actually, there is one more term which must be introduced into the Lagrangian as a consequence of gauge fixing. Since the path integral formulation does not implicitly know about SU(3) gauge

symmetry, a Fadeev-Popov ghost term $\mathcal{L}_{ghost}$ is added to the Lagrangian as a computational tool that ensures that physically equivalent solutions related by a gauge transformation are not double-counted. The ghost term is

$$\mathcal{L}_{ghost} = \partial_\mu \bar{c}^C D^\mu c^C \tag{2.11}$$

where c^A is the ghost field. Ghosts only show up as virtual particles in loops, and the ghost term gives the Feynman rules for a ghost propagator and a ghost-ghost-gluon vertex.

By contrast, consider the theory of quantum electrodynamics (QED), an abelian gauge field theory with symmetry group U(1) and Lagrangian density

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu\left(\partial_\mu + ieA^\mu_{QED}\right) - m\right)\psi - \frac{1}{4}\left(F_{QED}\right)^{\mu\nu}\left(F_{QED}\right)_{\mu\nu} \tag{2.12}$$

the structure constants f in Equation 2.2 vanish since the generators of U(1) commute trivially. In QED, only the fermion-fermion-boson ($gA\bar{\psi}\psi$) interaction vertex exists and there are no pure boson interaction terms. Unlike gluons, photons themselves do not carry the charge that they couple to and do not interact with each other.

2.1.1 Asymptotic freedom and confinement

It is a general feature of renormalizable quantum field theories that the effective strength of the coupling changes as a function of the momentum scale Q^2 with which they are probed. The fields and physical quantities which appear in the QCD Lagrangian in Equation 2.1 are "bare" quantities, which do not include the corrections from higher-order diagrams containing virtual particle loops. In order for the coupling constant to have a meaningful value at some scale Q^2, it is necessary reformulate the Lagrangian in terms of a physical part and a "counter-term" Lagrangian which is chosen to exactly cancel the divergences in the physical parameters order by order. As a consequence of this, the strength of the coupling constant g (or equivalently, α_s) changes with Q^2 in a way dictated by the counterterms.

The running of the coupling with energy is encoded in the renormalization group (RG) equation for α_s, in which the β function is typically expanded in powers of α_s,

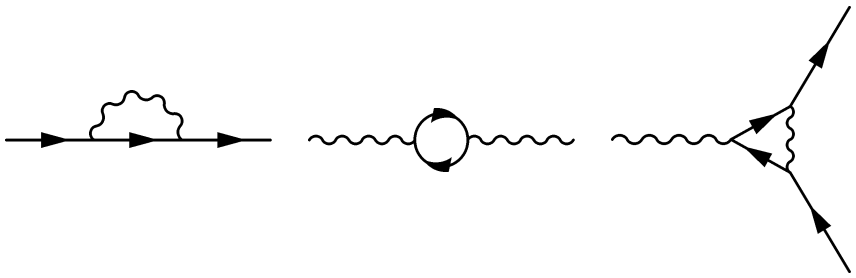


Figure 2.2: One-loop diagrams contributing to the running of the QED coupling constant.

$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \frac{\partial \alpha_s}{\partial \log Q^2} = \beta(\alpha_s) = -\alpha_s^2 \left(\beta_0 + \alpha_s \beta_1 + \dots\right) \tag{2.13}$$

Once $\beta(\alpha_s)$ is known, then Equation 2.13 can be solved to find the coupling at a scale Q^2 given the coupling at a lower scale μ^2:

$$\alpha_s(Q^2) = \alpha_s(\mu^2) \frac{1}{1 + b_0 \alpha(\mu^2) \log(Q^2/\mu^2) + \mathcal{O}(\alpha_s^2)} \tag{2.14}$$

That is, Equation 2.14 describes how the coupling "runs" from μ^2 to Q^2. The β function can be determined to lowest order by considering the divergences in one-loop corrections to the tree diagrams for each term in the Lagrangian.

By analogy, we begin with QED. In QED at the one-loop level, there are three one-loop diagrams which contribute to the running of $\alpha_{QED} = e^2/4\pi$, shown in Figure 2.2. These are the electron self-energy, the QED vertex function and the photon propagator. It turns out that the first two cancel exactly to all orders due to the Ward identity, which reflects the underlying gauge invariance of QED. Thus, the QED one-loop beta function is supplied only by the photon vacuum polarization,

$$\beta_0^{QED}\left(\alpha_{QED}^2\right) = -1/3\pi \tag{2.15}$$

Combining this with Equation 2.14, the QED coupling constant increases with increasing Q^2. This is consistent with the explanation of electric charge screening: the closer you probe an electric charge, the less it is "screened" by fermion-antifermion pairs with opposite charge.

In contrast, there are seven such diagrams in QCD. The diagrams with a QED analogue (albeit with the complication of color factors) are shown in Figure 2.3 while those which include the 3-boson

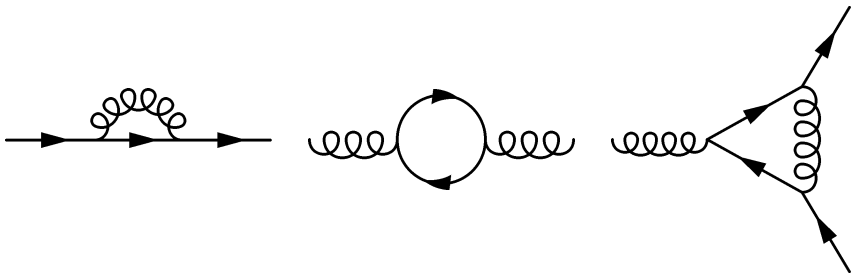


Figure 2.3: One-loop diagrams contributing to the running of the QCD coupling constant, which have QED analogues.

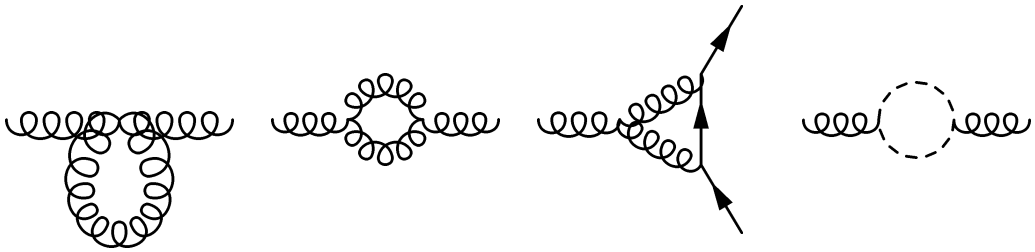


Figure 2.4: One-loop diagrams contributing to the running of the QCD coupling constant, which have no analogues diagrams in QED.

and 4-boson vertices and thus have no QED analogue are shown in Figure 2.4. Taken together, the one-loop beta function is

$$\beta_0(\alpha_s^2) = (33 - 2n_f)/12\pi \tag{2.16}$$

where n_f is the number of fermion flavors, the positive term in Equation 2.16 comes mainly from the non-Abelian diagrams and the negative term comes from fermion loop diagrams. This result, which can be readily calculated from the first principles of quantum field theory, was first done by David Gross, David Politzer and Frank Wilczek in 1973[5; 6], which earned them the Nobel Prize in 2004[1]. In our universe, there are few enough fermion flavors that $\beta_0 > 0$, which has phenomenal implications for the physical consequences of QCD. At high Q^2, α_s tends asymptotically to 0. This is called *asymptotic freedom* and implies that at sufficiently high Q^2, perturbative calculations of QCD processes should be well behaved. This can be understood as *color anti-screening*: since

[1] Actually, a sign error in the first calculation of the β function almost convinced Wilczek et. al.[7] that Yang-Mills theories were not asymptotically free, as expected at the time!

gluons carry the color charge, probing at long distance scales (small Q^2) actually increases the effective coupling from the gluon cloud surrounding the color source being probed. Interestingly, it can be shown that only non-Abelian gauge theories can have $\beta > 0$ in Equation 2.14 and be asymptotically free.

On the other hand, the choice of scale μ^2 above is quite arbitrary. We can define a scale Λ^2 by $1 = \alpha_s(\mu^2) b_0 \log\left(\mu^2/\Lambda^2\right)$, such that Equation 2.14 can be rewritten as

$$\alpha_s(Q^2) = \frac{1}{b_0 \log(Q^2/\Lambda^2)} \tag{2.17}$$

which implies that there is a momentum scale Λ_{QCD}, which $\approx 200 - 400$ MeV depending on the renormalization scheme and number of available quark flavors, at which $\alpha_s\left(Q^2 \sim \Lambda^2_{QCD}\right)$ begins to become large. Below or even near this scale, perturbative expansions of QCD are not possible and the Feynman diagram approach does not apply. The only first principles calculations in this regime come from *lattice QCD*, in which the action of the QCD Lagrangian is numerically computed on a discretized lattice.

This also sets the scale for the sizes of light hadrons, $\hbar c/\Lambda_{QCD} \approx 1$ fm, roughly the distance free partons in the hadron can move away from each other before the attractive potential brings them back together. Calculations show that this potential rises linearly with distance.

A summary of the Q^2 dependence of α_s from experiments and lattice calculations is shown in Figure 2.5. The world average for the value of the strong coupling constant at the Z boson mass scale[1] is

$$\alpha_s\left(M_Z^2\right) = 0.1184 \pm 0.0007 \tag{2.18}$$

2.1.2 Particle content

In the Standard Model of particle physics, there are six spin-1/2 fermions and six anti-fermions which carry color charge and transform as **3** and $\bar{\mathbf{3}}$ representations of SU(3), respectively. The lightest three of these, the *up*, *down* and *strange* quarks share an approximate SU(3) flavor symmetry due to their small masses.

After the discovery of the muon in 1936[8], a large number of strongly-interacting particles were produced in laborities throughout the following decades. There were too many of them

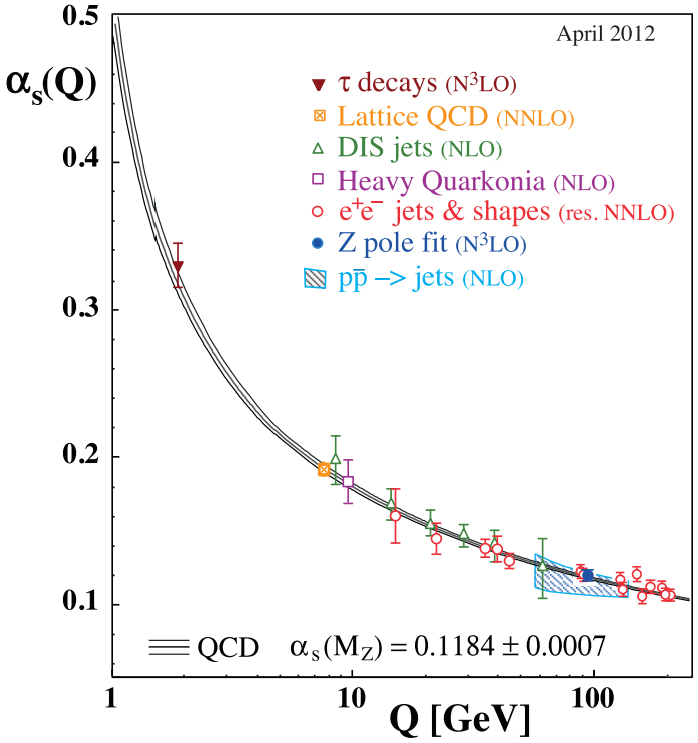


Figure 2.5: Summary of experimental data on $\alpha_s(Q^2)$, from [1].

for all to be fundamental, so it was thought that their mass and other properties were caused by a much smaller number of elementary particles. In 1961, Gell-Mann and Ne'eman argued[9; 10] that all mesons and baryons of the same spin and parity could be arranged in a (Y, T_3)-plot (where Y is the hypercharge and T_3 is the isospin) in a way analagous to representations of SU(3). The baryons could be arranged according to $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$ and the mesons according to $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$. One baryon in the **10** decuplet – the Ω^-, with spin-3/2, charge -1, strangeness $+3$ particle – had not yet been observed. It was finally observed in a bubble chamber at Brookhaven National Laboratory[11], with an experimentally measured mass, quantum numbers and even decay modes corresponding very neartly with Gell-Mann's prediction.

In 1964, Gell-Mann[12] and Zweig[13] proposed that was an SU(3) *flavor* symmetry and the observed hadrons were composed of three quarks, elementary spin-1/2 fermions which are fundamental representations of the symmetry. This SU(3) symmetry was not exact (since the light quark masses are non-vanishing) as can be seen from the mass splitting in the decuplet, $(m_\Sigma - m_N) / (m_\Sigma + m_N) \approx$

12%. The SU(2) isospin symmetry is a much stronger, but still inexact one, with $(m_n - m_p) / (m_n + m_p) \approx$ 0.2%. Murray Gell-Mann was awarded the 1969 Nobel Prize, but it was for his contributions to the theory of the strong interaction generally, since the quark model and its ramifications had not yet been fully experimentally ratified.

One major puzzle remained. There were spin-3/2 baryons which in the quark model were composed of the same flavor quark with a symmetric spin configuration, such as the Ω^- (sss), the N^{*-} (ddd) and the N^{*++} (uuu). In these ground state baryons, the quark wavefunctions had zero angular momentum and were thus totally symmetric. Thus, the overall wavefunction appeared to be completely symmetric, violating the connection between fermion spin and antisymmetric statistics. It was proposed that quarks were fundamental representations of a separate SU(3) symmetry, and carried a quantum number called *color*. Thus, with "red", "green" and "blue" quarks, an antisymmetric color wavefunction (e.g. $|rgb\rangle - |rbg\rangle + |gbr\rangle - |grb\rangle + |brg\rangle - |bgr\rangle$) keeps the overall wavefunction antisymmetric. Unlike the approximate SU(3) flavor symmetry used to phenomenologically explain the observed hadron spectrum, the SU(3) color symmetry was an exact symmetry of the underlying strong nuclear interaction. This solution actually opened another puzzle – why only color singlet states were observed in Nature – answered only with the discovery of confinement and the identification of QCD as the theory of the strong interaction.

Five years later the fourth quark, called *charm*, was discovered. It had been predicted to exist on a number of theoretical grounds, such as lepton-quark generation symmetry[14]. The charm quark was observed in the form of the J/ψ particle, which is the first excited $c\bar{c}$ bound state with $m_{J/\psi} = 3.1$ GeV/c^2, through the $J/\psi \to e^+ + e^-$ decay mode simultaneously at the alternating gradient synchotron (AGS) at Brookhaven National Laboratory (BNL) [15] and at the Stanford Positron Electron Accelerating Ring (SPEAR) at the Stanford Linear Accelerator Center (SLAC)[16] in 1974. Samuel C.C. Ting and Burton Richter shared the 1976 Nobel Prize for this discovery. A number of charmed mesons were soon discovered ($c\bar{u}$, $\bar{c}u$, $c\bar{d}$, $d\bar{c}$, etc.), with properties well described by the quark model (albeit within the increasingly approximate SU(4) flavor symmetry).

In 1973, two more quarks were predicted to exist[17] as a way of explaining the observed indirect CP violation in neutral kaon decays[18] in 1964. The authors reasoned that no realistic models of CP violation existed within a "quartet" (four-quark) model but showed that with a third quark

generation it is possible to introduce a CP-violating phase in the 3×3 unitary matrix which described the mixing between weak force interaction eigenstates and the mass eigenstates. For predicting the third generation of quarks, Makoto Kobayashi and Toshihide Maskawa shared the 2008 Nobel Prize (shared with Y. Nambu, above).

The bottom quark b was discovered in the form of the Υ meson, the lowest energy $b\bar{b}$ state with $m_\Upsilon = 9.6$ GeV/c^2, through the $\Upsilon \rightarrow \mu^+ + \mu^-$ decay mode by the E288 Collaboration at Fermilab[19] in 1977. Over the next 18 years, long searches at SLAC, DESY and the SPS failed to find the top quark t, indicating that it must be subtantially more massive than the b. It was finally discovered in 1995 through the decay of the top-antitop quark pair $t\bar{t} \rightarrow W^+bW^-\bar{b}$, with particular emphasis on the subsequent $W^\pm \rightarrow l^\pm\nu_l$ decay channel, at the Collider Detector at Fermilab (CDF) [20] and D0 [21] experiments. Unlike the other heavy quarks, the mass of the top is so large ($m_t = 174$ GeV/c^2[1]) that it decays, almost always $t \rightarrow W^+b$ with total width $\Gamma = 2.0$ GeV, before it can hadronize and form any $t\bar{q}$ meson or t-containing baryon.

Conclusive evidence for the existence of gluon was found in the form of 3-jet events ($e^+ + e^- \rightarrow q\bar{q}g$, where each of the three partons fragments into a jet) at the Positron-Electron Tandem Ring Accelerator (PETRA) at the Deutsches Elektronen-Synchrotron (German Electron Synchotron; DESY), with results from the MARK-J[22], PETRA[23] and PLUTO[24] experiments. Since the scattered electron and positron have zero baryon number, the quarks must be produced in quark-antiquark pairs. The third jet is necessarily caused by very hard final state radiation in the form of a gluon. A more detailed analysis[25] of the angular correlations between the produced jets confirmed that the third jet is consistend with a spin-1 particle, exactly as required for the QCD gluon.

No further quarks have been discovered within the energies available to experiments. It is thought, however, that the stringent limits on the number of neutrino types[1] from e^+e^- colliders and cosmology, along with the parsimony of requiring equal numbers of lepton and quark generations may imply that there are only six flavors. There were efforts to name the third generation of quarks *beauty* and *truth*, respectively, but the terms *bottom* and *top*, first introduced in [26], have become the most popular usage.

2.1.3 Factorization and Deep Inelastic Scattering

Factorization[27] is the separation of hadronic cross-sections into short-scale (high-Q^2) partonic processes and long-scale (non-perturbative) processes such as hadronization. Factorization allows us to meaningfully apply calculations involving *partonic* cross-sections to physically measurable *hadronic* cross-sections. Under the factorization scheme, the single particle inclusive cross-section in proton-proton collisions ($pp \to h + X$, where h is the particle of interest and X denotes a sum over the remainder of the final state) can be written symbolically at leading order as

$$d\sigma^{pp\to h+X} = \sum_{abcd} \int dx_a \int dx_b \int dz f_{a/p}(x_a, Q^2) f_{b/p}(x_b, Q^2) d\hat{\sigma}^{a+b\to c+d}(Q^2) D_{c/h}(z) \tag{2.19}$$

where $f_{a/p}$ is the probability of finding a parton of type a inside the proton p, c is the parton produced in the hard-scattering that eventually fragments into final-state hadron h carrying z of its energy, and $D_{c/h}$ is the probability for c to fragment in this manner. $\hat{\sigma}$ is the partonic $a+b \to c+d$ cross-section for producing a parton c with momentum p_h/z, at a hard scale Q^2. $\sum_{abcd}$ sums over all incoming and outgoing parton species in the $2 \to 2$ QCD scattering (obviously, some choices of parton flavors, such as $gg \to qq$, do not occur at leading order and thus do not contribute to the total sum).

The individual elements in Equation 2.19 also implicitly depend on the the factorization scale μ_f (not the same as the renormalization scale μ) which serves as a collinear cutoff below which all QCD behavior is collected into f and D. Roughly speaking, corrections with internal lines of the order μ_f^2 should be included in the "hard" part of the factorized cross-section, while those below it will be grouped into f and D. The partonic cross-section does not depend on μ_f to leading order, but does depend on it logarithmically at higher-orders. For real calculations, the factorization scale is typically taken to be $\mu_f \sim Q$ and is varied by a factor of 2 in each direction $\left(\frac{\mu_f}{2}, 2\mu_f\right)$ to test the sensitivity of the calculation to the factorization scale. One of the earliest successful uses of the factorization theorem was to compute the cross-section of inclusive π^0 production in pp collisions and find a good agreement with measurements taken at the CERN ISR[28].

As an illustrative example of the factorization theorems, consider a high energy lepton-hadron (e.g. e^+p) scattering in the center of mass frame. The internal configuration of the partons (including any normally short-lived virtual partonic states) in the hadron are time dilated and

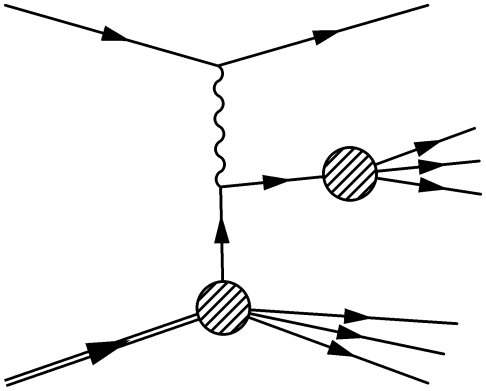


Figure 2.6: Lepton-hadron scattering diagram.

"frozen" for the duration of the collision. We can think of each parton as having some definite momentum fraction x of the hadron's total momentum in this frame. In this regime, as the lepton traverses the hadron, we can neglect the parton-parton interactions which occur before and after, but nor during, the collision. Thus, it makes sense to describe the collision as a lepton interacting with a (particular) free parton of definite momentum.

In fact, high-Q^2 lepton-hadron scattering, called *deep inelastic scattering* (DIS), was historically used to probe the partonic structure of hadrons. Lepton-hadron scattering is shown diagrammatically in Figure 2.6 and is the reaction

$$l(k) + h(p) \to l'(k') + X \tag{2.20}$$

where l and l' are the ingoing and outgoing lepton with four-momenta k and k', respectively, h is the hadron with four-momentum p and X is the final hadronic state. Because of the clean initial and final state of the lepton, the kinematics of the underlying lepton-parton interaction can be reconstructed exactly, via

$$q^2 = (k - k')^2 \approx -Q^2 \text{ as } m_l/Q \to 0 \tag{2.21}$$

To calculate the structure of the cross-section, we can perform an integral over all final hadronic states, as long as we are only interested in measuring the scattered lepton and not in the X system. The hadronic states cannot be calculated perturbatively, but this is a theoretically sound procedure since we are interested in an *inclusive* cross-section – the details of the hadronization are irrelevant

as long as is happens with probability 1. In this case, the inclusive differential cross-section can be generally written as

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \left(2W_1 \sin^2 \frac{\theta}{2} + W_2 \cos^2 \frac{\theta}{2} \right) \tag{2.22}$$

where $E = k^0$, $E' = (k')^0$ and α is the QED coupling constant. All details of hadron structure are encoded in $W_{1,2}(p, k)$. We can rewrite this in a way that collects the common energy scale by introducing the dimensionless variables

$$y = 1 - \frac{E'}{E} \tag{2.23}$$

$$x = \frac{Q^2}{2M(E - E')} \tag{2.24}$$

where x is the fraction of the hadron's momentum carried by the parton and y is the fraction of the lepton's energy lost in the hadron rest frame. The cross-section is then

$$\frac{d^2\sigma}{dxdy} = \frac{2\pi\alpha^2}{MEx^2y^2} \left[xy^2 F_1 + \left(1 - y - \frac{M}{2E} xy \right) F_2 \right] \tag{2.25}$$

where the description of the internal hadron structure is now encoded in the *structure functions* $F_{1/2}\left(x, Q^2\right)$.

In 1968, $e^- + p$ experiments were performed for the first time at the Stanford Linear Accelerator with 7-10 GeV electrons on a hydrogen target, probing F_1 and F_2 as a function of x for a range of Q^2 values[29; 30]. Since then, $\mu + p$ and $e + p$ experiments at CERN and DESY have significantly expanded the explored phase space. A representative summary of world data on the proton structure function is shown in Figure 2.7. For an intermediate range of x values, the structure functions are largely independent of Q^2, a phenomenon referred to as *Bjorken scaling*[31]. This means that changing the energy of the probe does not resolve any more detail about the structure of the proton, and that it is composed of point-like particles with no further substructure. Bjorken scaling validates the parton model and introduces the notion that, for sufficiently high-Q^2 probes, the partons are asymptotically free inside the proton. In this regime, the "inelastic" lepton-hadron collisions are really incoherent elastic lepton-parton collisions. This discovery means that for a wide range of Q^2, the lepton-hadron scattering cross-section can be written as

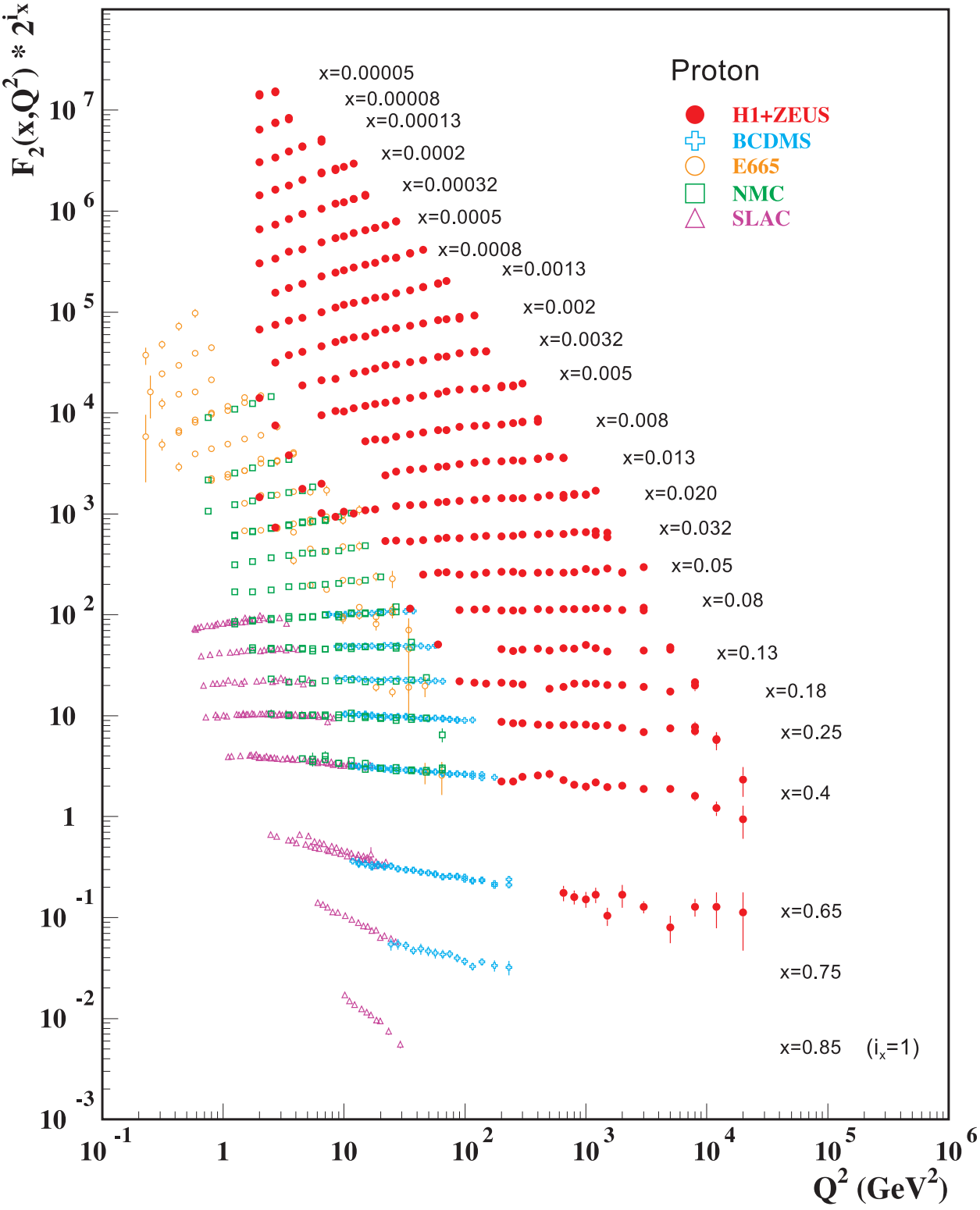


Figure 2.7: Representative data on the proton structure function F_2 measured in DIS experiments, from [1].

$$\sigma^{l+h}(x,Q^2) = \sum_a \int_x^1 d\xi f_{q/h}(\xi)\sigma_B^{l+q}(x/\xi,Q^2) \tag{2.26}$$

where σ_B is the Born cross-section for lepton-quark scattering. Equation 2.26 is the inspiration for the more complicated form of the factorization theorem for hadronic collisions shown in Equation 2.19.

Surprisingly, DIS experiments also revealed that a relatively small ($\sim 50\%$) fraction of the total proton momentum was carried by the partons visible to lepton scattering. (The rest, as it turned out, is carried by the electrically neutral gluons.) Furthermore, the data confirm the Callan-Gross relation,

$$F_2(x,Q^2) = 2xF_1(x,Q^2) \tag{2.27}$$

which is a consequence of the spin-1/2 nature of partons (on the other hands, if partons had spin-0, $F_1 = 0$). The structure functions can be constructed at a more basic level from the *parton distribution functions* (PDFs), $f_i(x,Q^2)$, which encode the probability, at scale Q^2, of finding a parton of type i with momentum fraction x, via

$$F_2 = \sum_i Q^2 x f_i \tag{2.28}$$

Various momentum and quantum number sum rules give relationships between the f_i such that the total momentum, valence quark number, strangeness, isospin, and other properties of the quark, antiquark and gluon distribution functions add up to the total value in the hadron in question. An example application of this is the ratio F_2^{ep}/F_2^{en} of the structure functions in the proton and neutron, respectively, which can be extracted from proton and deuteron DIS experiments. At low x, sea quarks dominate the distributions, making protons and neutrons "look" the same to the electromagnetic probe and resulting in $F_2^{ep}/F_2^{en} \to 1$. At high x, valence quarks dominate. Since the up and down quark pdfs in protons and neutrons are related by an isospin transformation (e.g. $u_n = d_p$ and vice versa), $F_2^{ep}/F_2^{en} \to 4$ essentially from the fact that $Q_u^2 = 4Q_d^2$. In fact, the data show exactly this[32].

On the other hand, for very higher-Q^2 or at low-x, Bjorken scaling is broken. That is, the proton has more (low-x) constituents as it is probed on finer and finer distance scales (higher-

Q^2). As it turns out, this evolution of the parton distribution functions with Q^2 can be described quantitatively in QCD, as discussed in the next section.

2.1.4 Parton distribution functions

The parton distribution functions defined in Equation 2.19 evolve as a function of Q^2, according to a set of equations that encode the probabilities for quarks and gluons to split or merge into other partons (for example, through a quark radiating a soft gluon or a gluon turning into a quark-antiquark pair). These are called the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP)[33; 34; 35] equations and are applicable in the regime where Q^2 and x are relatively large. For quark, antiquark and gluon densities $q_i(x,Q^2)$, $\bar{q}_i(x,Q^2)$ and $g(x,Q^2)$, the full coupled set of equations at leading order (LO) is

$$Q^2\frac{\partial}{\partial Q^2}\begin{pmatrix} q_i(x,Q^2) \\ \bar{q}_i(x,Q^2) \\ g(x,Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\begin{pmatrix} P_{q_iq_j}\left(\frac{x}{\xi}\right) & 0 & P_{q_ig}\left(\frac{x}{\xi}\right) \\ 0 & P_{\bar{q}_i\bar{q}_j}\left(\frac{x}{\xi}\right) & P_{\bar{q}_ig}\left(\frac{x}{\xi}\right) \\ P_{gq_j}\left(\frac{x}{\xi}\right) & P_{g\bar{q}_j}\left(\frac{x}{\xi}\right) & P_{gg}\left(\frac{x}{\xi}\right) \end{pmatrix}\begin{pmatrix} q_j(\xi,Q^2) \\ \bar{q}_j(\xi,Q^2) \\ g(\xi,Q^2) \end{pmatrix} \tag{2.29}$$

where the $P_{\alpha\beta}(\frac{x}{\xi})$ functions describe the probability density to find a parton of type α and momentum fraction x in a parton of type β and momentum fraction ξ. Equation 2.29 are actually renormalization group equations, and are used to calculate $f(x,Q^2)$ given some initial starting point $f(x,Q_0^2)$.

A good knowledge of PDFs is critical for precision tests of QCD. There exist several PDF sets which are constructed from global fits to data at various (x,Q^2) and using next-to-leading or even next-to-next-to-leading order solutions to the DGLAP equations to connect the results at different Q^2 scales. They have been experimentally measured approximately within the phase space of $2\times10^{-5} < x < 0.9$ and $1.69 < Q^2 < 3\times10^5$ GeV^2. Two recent comprehensive set of functions are the NLO CTEQ-Jefferson Lab 2012 (CJ12)[36; 37] or the NNLO CTEQ 2010[38] sets.

Parton distribution functions have a few general features. At fixed Q^2 they decrease with increasing x (it is harder to find a parton with a larger momentum fraction than it is a smaller one). Figure 2.8 shows the CJ12 light quark and gluon PDFs at an example Q^2 along with their uncertainties. Notice that the gluon PDF is divided by a factor of 10 on the right plot, and the gluon

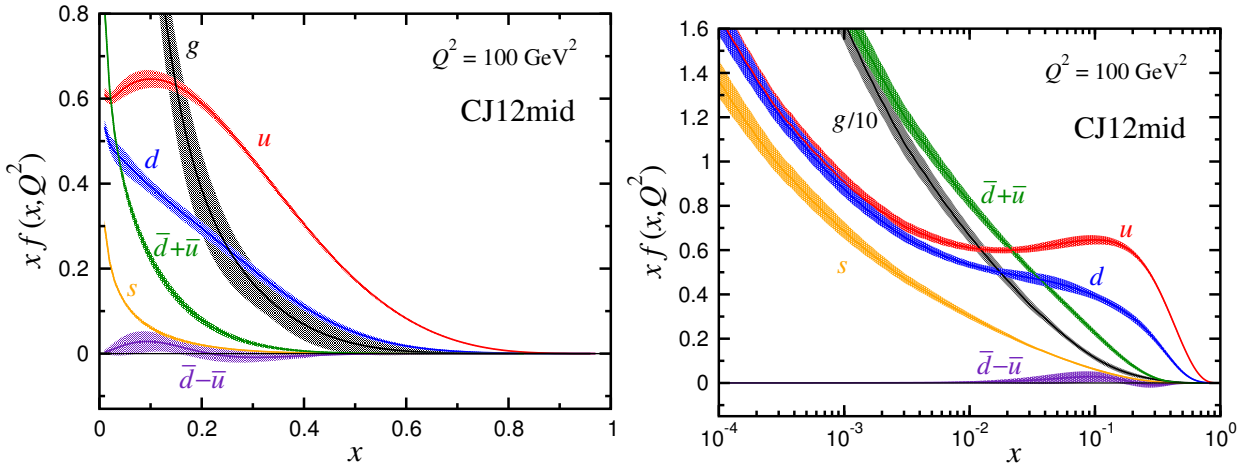


Figure 2.8: CJ12 light quark and gluon PDFs at $Q^2 = 100$ GeV^2, with linear (logarithmic) x-scale on the left (right), from [36].

PDF dominates over all quark PDF's below $x \lesssim 0.2$ (at least, at $Q^2 = 10^2$ GeV^2). Furthermore, as Q^2 increases, the low-x parts of the PDF grow faster than the high-x parts.

Experiments at HERA[39; 40] first observed a large increase in the structure function F_2 for decreasing x below $< 10^{-2}$, prompting speculation as to whether the DGLAP formalism could successfully describe this behavior in a high gluon density regime. Naively, at leading order the gluon splitting function $P_{gg} \propto 1/x$ becomes singular as $x \to 0$. Thus it is thought that at very high energies, a separate treatment must be applied. The DGLAP formalism resums contributions to loop diagrams that are powers of $\log\left(Q^2/\Lambda^2_{QCD}\right)$. However, at large $\sqrt{s}$ (or, equivalently, small $x = Q^2/s$), contributions of the form $\left(\alpha_s \log\left(\frac{1}{x}\right)\right)^n$ become the dominant ones and linear DGLAP evolution may not be appropriate. In this regime, it is thought that the Balitsky-Fadin-Kuraev-Lipatov (BFKL)[41; 42] equations must be used to describe the evolution in x instead. A more detailed discussion of the nonlinear QCD behavior of partons at very low-x is deferred to Section Section 2.2.

When appropriate, the parton distribution function formalism has been generalized in ways that are relevant to other physics scenarios. For example, the factorization picture does not apply in the case of multiple parton interactions (if > 2 partons in the hadron-hadron collision paticipate in hard scatterings), and a new formalism using double parton distribution functions would have to be introducted. For example, these distribution functions would depend on the scales Q_1^2, Q_2^2 and

momentum fractions x_1 and x_2 and even, in principle, the transverse distance between the partons b[43]. Another example is the generalized parton distribution functions (GPDFs)[44], which include the full phase space description of partons in the nucleon instead of just the Bjorken-x. Integrating out the relevant coordinates results in the transverse-momentum-dependent (TMD) PDFs, which are important for the understanding of the spin structure of nucleons.

2.1.5 QCD jets

One of the most straightforward QCD observables is the inclusive spectrum of *jets*, collimated streams of hadrons from the end stage of a parton shower. This concept of a jet, which is loosely associated with the shower of an originating hard parton that has undergone hadronization, must be made more specific before it is a useful experimental or theoretical concept. Jets are basic QCD objects as well as the dominant final state object in hadronic collisions. Thus, developing the technology for measuring jets in a variety of collision systems has been a subject of much interest.

In a leading order picture, a jet is a fragmenting parton. *Jet reconstruction* is a procedure by which the momenta of the fragments are resummed back to that of the original parton, in effect undoing the fragmentation process. Unfortunately, QCD does not allow an unambiguous separation of the final state hadrons into those originating from the hard-scattered parton and those that do not. Thus the definition of a jet requires a resolution parameter, typically denoted R or σ, which defines the angular scale of what radiation is included in the jet. This resolves the ambiguity, in the case of a high-energy quark radiating a hard gluon, between when the gluon should be considered its own jet and when it should be included as part of the quark jet (e.g. when the j_{T} of the gluon is $< p_{\mathrm{T}} \sin \Delta R$, it stays in the jet). Furthermore, observables in QCD are only well-defined at all orders of perturbation theory if they are insensitive to the collinear splitting of partons ("collinear safety") and the emission of very soft radiation at large angles ("infrared safety").

Although jets can be picked out by the human eye in an event display, it is complicated and not necessarily straightforward to develop a good procedure for identifying where the collimated energy flow in an event goes that is subject to the requirements of infrared and collinear safety. For example, early methods of quantifying the presence of a jet involved calculating the quadrupole moment tensor of hadron momenta in the event, the "sphericity" parameter (which categorizes events on a continuum between purely isotropic and dijet-like) or the "thrust" of the event which

further characterized the topology of the energy flow. In modern usage, jet reconstruction is the process of turning a set of measured hadrons, clusters or other energy deposits $\{p^\mu\}_{hadron}$ into a unique set of jets $\{p^\mu\}_{jet}$ through the use of a jet reconstruction algorithm. The requirements of collinear and infrared safety can be defined in this notation. Let f be the jet reconstruction procedure which maps $f\left(\{p^\mu\}_{hadron}\right) \to \{p^\mu\}_{jet}$.

Collinear safety is the requirement that if $p_n^\mu \mid\mid p_{n+1}^\mu$ and $(p^\mu)'_n = p_n^\mu + p_{n+1}^\mu$, then

$$f(\{p_1^\mu, \ldots, p_{n-1}^\mu, p_n^\mu, p_{n+1}^\mu\}) = f(\{p_1^\mu, \ldots, p_{n-1}^\mu, (p^\mu)'_n\}) \tag{2.30}$$

And infrared safety is the statement that in the limit $p_{ghost}^\mu \to 0$,

$$f(\left\{p_1^\mu, \ldots, p_n^\mu, p_{ghost}^\mu\right\}) = f(\{p_1^\mu, \ldots, p_n^\mu\}) \tag{2.31}$$

The history of modern algorithms begins with the Snowmass Accord[45], which motivated the need to standardize jet algorithms and described the desirable experimental and theoretical features of a useful algorithm. One of the earliest jet definitions is the cone algorithm with radius R. The set of n hadrons in the event is described by $\{E_{\mathrm{T}}, \eta, \phi\}$. For a given cone position (η_C, ϕ_C), the total jet energy is given by

$$E_{\mathrm{T}J} = \sum_{i=1}^{n} \theta(R - \Delta R) E_{\mathrm{T}i} \tag{2.32}$$

where $\Delta R^2 = (\eta_i - \eta_J)^2 + (\phi_i - \phi_J)^2$, that is, only hadrons within R of the jet axis are included in the calculation. Then, the jet position is recalculated according to

$$\eta_J = \frac{1}{E_{\mathrm{T}J}} \sum_{i=1}^{n} \theta(R - \Delta R) E_{\mathrm{T}i} \eta_i \tag{2.33}$$

$$\phi_J = \frac{1}{E_{\mathrm{T}J}} \sum_{i=1}^{n} \theta(R - \Delta R) E_{\mathrm{T}i} \phi_i \tag{2.34}$$

The procedure is repeated with the updated cone position at $(\eta_c, \phi_c) = (\eta_J, \phi_J)$ and continues until (η_J, ϕ_J) converges, resulting in a final jet. This algorithm, while serving as a prototype for later developments, suffered from numerous technical deficiencies, including the need to "seed" an initial guess (η_c, ϕ_c), instability with the addition of an infrared particle near the edge of the jet axis, and non-convergence of the iterative procedure.

Since then, many improved and more sophisticated algorithms have been proposed and used in experiments. In this work, two modern jet algorithms are used, the *Gaussian filter*, a seedless cone-like filtering algorithm, and the *anti-k_{T}* algorithm, a sequential recombination algorithm.

In Gaussian filter[46] formalism, let the event energy density $\rho(\eta, \phi)$ be

$$\rho(\eta, \phi) = \sum_{i \in \text{ hadrons}} \delta(\eta - \eta_i)\delta(\phi - \phi_i)(p_{\mathrm{T}})_i \tag{2.35}$$

then for a possible jet position (η_J, ϕ_J), the jet p_{T} is given by

$$(p_{\mathrm{T}})_J(\eta_J, \phi_J) = \int d\eta \int d\phi \rho(\eta, \phi) h(\eta - \eta_J, \phi - \phi_J) \tag{2.36}$$

where $h(\Delta\eta, \Delta\phi)$ is the kernel which controls the contribution the jet energy from the energy flow in the event as a function of angular separation. For example, in the traditional cone algorithms, $h(\Delta\eta, \Delta\phi) = \theta(R^2 - \Delta\eta^2 - \Delta\phi^2)$, e.g. only particles within $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2 < R^2$ are included, but contribute 100% of their energy when they do. In the Gaussian filter, the kernel is chosen so that the energy weighing is a smoothly decreasing function of ΔR withour the hard angular cutoff that has historically caused problems for earlier algorithms. Thus, for the Gaussian filter algorithm, $h = \exp(-\Delta R^2/2\sigma^2)$ where σ is the parameter that controls the angular-dependent energy weighing (and thus, the angular resolution).

The algorithm finds a jet whenever $(p_{\mathrm{T}})_J$ is a local maximum in (η_J, ϕ_J)-space, with the jet's energy given by

$$p_{\mathrm{T}jet} = \max_{(\eta_J, \phi_J)} (p_{\mathrm{T}})_J(\eta_J, \phi_J) \tag{2.37}$$

Notably, this avoids the problem of needing to seed any initial starting location. The algorithm searches the entire phase space and simply finds the points (η_J, ϕ_J) at which the convolution $(p_{\mathrm{T}})_J$ is a local maximum.

The Gaussian filter algorithm was originally developed for use in heavy-ion collisions and in the small-acceptance PHENIX detector, where several key properties of the algorithm were thought to give it an advantage over other algorithms in an environment with large underlying event and with respect to pathological effects introduced near the edges of particle acceptance.

Additionally, since the algorithm selects the jet direction to align with the peak energy flow, it may be less sensitive to energy from the periphery of the jet lying outside the acceptance.

The Gaussian filter algorithm was shown to have good experimental behavior and the ability to successfully reproduce NLO predictions for the pp cross-section in 200 GeV collisions at RHIC and then used to measure jet suppression in central Cu+Cu collisions at $\sqrt{s_{NN}} = 200$ GeV[47]. The present work extends the use of this algorithm to d+Au collisions at 200 GeV, the third collision system in which it has been used.

The anti-k_{T} algorithm[48] is a member of a larger family of sequential recombination algorithms. Sequential recombination algorithms work by operating on a list of four-momenta $\{p^\mu\}$, iteratively combining pairs of them in a specified way until a certain point where the procedure terminates and returns the set of jets. A sequential recombination algorithm is fully specified when the method for deciding which particles to combine, the scheme to be used to combine the momenta of the particle pair, and the criterion for stopping the iteration are given. In anti-k_{T} terminology, the working list of four-vectors are all termed "protojets" until the algorithm terminates, at which point they are "jets".

For each step of the iteration, construct the protojet-protojet distance d_{ij} and protojet-beam distance d_{iB} as follows.

$$d_{ij} = \min\left({p_{\mathrm{T}}}_i^{2p}, {p_{\mathrm{T}}}_j^{2p}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{2.38}$$

$$d_{iB} = {p_{\mathrm{T}}}_i^{2p} \tag{2.39}$$

where p is a parameter defined by the algorithm and $\Delta R_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i + \eta_j)^2$. Find the lowest d. If it is a d_{iB}, remove this protojet from the list and call it a final jet. If it is a d_{ij}, combine these two protojets into one protojet, typically via $p_i^\mu + p_j^\mu = p_k^\mu$, and insert it back in the list. Since every iteration decreases the list of protojets by one, the procedure will terminate after $|\{p_i^\mu\}|$ iterations.

For $p = 1$, this algorithm is known as the k_{T} algorithm, first developed for use in $e^+ + e^-$ collisions. The behavior of the k_{T} algorithm can be intuited somewhat by the form of Equation 2.39. Because $p > 0$, the minimum function in d_{ij} will cause the algorithm to cluster very soft particles first. In fact, in the small-angle limit where $\sin\Delta R \approx \Delta R$, $d_{ij} \approx \min\left(j_{\mathrm{T}i}^2, j_{\mathrm{T}j}^2\right)$ where the j_{T}

are the transverse momenta with respect to the other fragment. Thus, the k_{T} algorithm collects fragments with successively larger j_{T} with respect to one another. Although the k_{T} algorithm has been successfully used in experiments at $e^+ - e^-$ colliders, it is not well-behaved in the presence of a hadronic underlying event.

On the other hand, the counterintuitive choice of $p = -1$ leads to the anti-k_{T} algorithm which has a substantially different behavior. Consider the behavior of the algorithm in the presence of a hard particle with $p_{\mathrm{T}hard}$. The $1/{p_{\mathrm{T}}}^2$ weighing means that d_{ij} reduces to ${p_{\mathrm{T}}}_{hard}^{-2}\Delta R^2/R^2$. Thus, the algorithm will successively cluster the nearest fragment (regardless of p_{T}, as long as $p_{\mathrm{T}} < p_{\mathrm{T}hard}$) with the hard fragment. In the absence of other high-p_{T} particles, this will continue while $d_{ij} < d_{iB}$, or $\Delta R^2 < R^2$, meaning that all particles within a radius R will end up in the jet. This gives anti-k_{T} jets their famously cone-like shape.

This sequential recombination family of algorithms is implemented efficiency in the package `FASTJET` [49; 50]. Naïve implementations of the anti-k_{T} algorithm have a time complexity of $\mathcal{O}(N^3)$ for an input of N initial hadrons, since for each of the hadrons, every other hadron must be examined to form every possible d_{ij} (N^2 combinations) and these must be remade repeated at each step of the N iterations. Such a runtime is completely prohibitive for jet reconstruction in high-multiplicity pp or HI events. However, this can be reduced using nearest-neighbor algorithms from computational geometry[51] to an $\mathcal{O}(N^2)$ procedure which implements the algorithm exactly or an $\mathcal{O}(N \log N)$ procedure which finds a good approximate solution.

With the notion of a jet definition in hand, it is now possible to construct theoretical predictions about jet cross-sections. Following the factorization formula in Equation 2.19, we can expand the QCD predictions for cross-sections in proton-proton collisions order by order via

$$d\sigma^{p+p\to X} = \sum_{ij} \int dx_1 \int dx_2 f_{i/p}(x_1) f_{j/p}(x_2) \sum_n \alpha_s^n d\hat{\sigma}_{(n)}^{i,j\to X}(x_1, x_2) \tag{2.40}$$

where the renormalization and factorization scale dependence has been suppressed. Since the PDFs (and any further fragmentation functions, in the case of a non inclusive cross-section) are universal, all process dependence in Equation 2.40 is contained in the $d\hat{\sigma}_{(n)}$.

With a lot of bookkeeping and care, it is not too hard to write out the Feynman graphs which contribute to the leading order term $d\hat{\sigma}_2^{ij\to kl}$ for $2 \to 2$ QCD scattering. The master formula for generic $2 \to N$ processes in the center of mass frame is

$$d\sigma = \frac{1}{2s}\left(\prod_{i=1}^{N}\frac{d^3p_i}{(2\pi)^3}\frac{1}{2E_i}\right)\cdot(2\pi)^4\delta^4\left(p_A+p_B-\Sigma p_i\right)\cdot\left|\mathcal{M}\left(p_A,p_B\rightarrow\{p_i\}\right)\right|^2 \tag{2.41}$$

where the $(E_i, \vec{p}_i)$ are the outgoing 4-momenta of the N particles and $\mathcal{M}$ is the invariant matrix element. Specializing to $2 \rightarrow 2$ processes, in the high-energy limit this reduces to

$$d\sigma \xrightarrow{m\rightarrow 0} \frac{d\Omega}{16\pi^2 s}\left|\mathcal{M}\left(p_A,p_B\rightarrow p_C,p_D\right)\right|^2 \tag{2.42}$$

where $d\Omega$ points in the direction of $\vec{p}_C$ (which is constrained in this limit by four-momentum conversation to be $= -\vec{p}_D$). The diagrams for the four possibilities ($q_iq_j \rightarrow q_iq_j$ with distinct $i \neq j$, $q_iq_i \rightarrow q_iq_i$, $q_i\bar{q}_i \rightarrow gg$ and $gg \rightarrow gg$) are shown in Figure 2.9. The matrix element for the remaining processes (such as $qg \rightarrow qg$, $gg \rightarrow qq$ and $q\bar{q} \rightarrow q\bar{q}$) can be obtained by applying crossing symmetry to $q\bar{q} \rightarrow gg$ or $qq \rightarrow qq$. Averaging over incoming and summing over outgoing spins and colors, the results can be summarized as

$$\sum\overline{\left|\mathcal{M}^{q_jq_k\rightarrow q_jq_k}\right|}^2 = \frac{1}{4N_C^2}\left[a(s,t,u)\right] \tag{2.43}$$

$$\sum\overline{\left|\mathcal{M}^{q_jq_j\rightarrow q_jq_j}\right|}^2 = \frac{1}{4N_C^2}\left[a(s,t,u)+a(s,u,t)+b(s,t,u)\right] \tag{2.44}$$

$$\sum\overline{\left|\mathcal{M}^{q_i\bar{q}_i\rightarrow gg}\right|}^2 = \frac{1}{4N_c^2}\left[c(s,t,u)\right] \tag{2.45}$$

$$\sum\overline{\left|\mathcal{M}^{gg\rightarrow gg}\right|}^2 = \frac{1}{4(N_C^2-1)^2}\left[d(s,t,u)\right] \tag{2.46}$$

where at leading order the a, b, c and d functions are

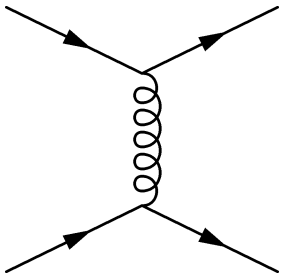
$$a(s,t,u) = 2(N_C^2-1)\frac{s^2+u^2}{t^2} \tag{2.47}$$

$$b(s,t,u) = -4\frac{N_C^2-1}{N_C}\frac{s^2}{ut} \tag{2.48}$$

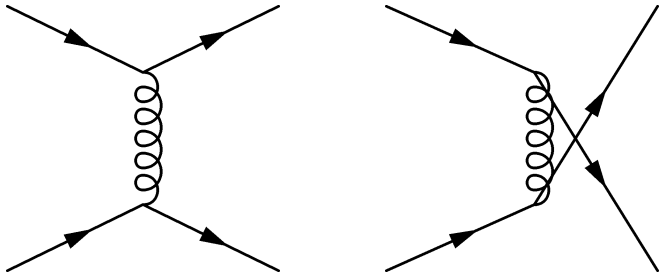
$$c(s,t,u) = 2\frac{N_C^2-1}{N_C}\left(\frac{N_C^2-1}{ut}-\frac{2N_C^2}{s^2}\right)\left(t^2+u^2\right) \tag{2.49}$$

$$d(s,t,u) = 16N_C^2(N_C^2-1)\left(3-\frac{ut}{s^2}-\frac{us}{t^2}-\frac{st}{u^2}\right) \tag{2.50}$$

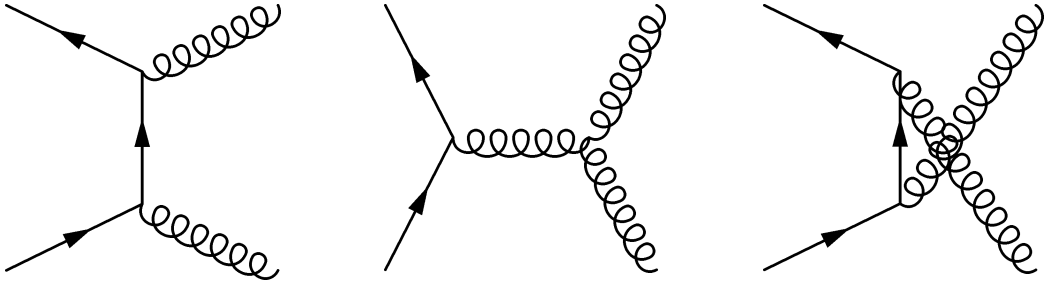
Calculating the NLO terms includes the addition of many more diagrams and was first performed in [52]. The general form of the next-order term is



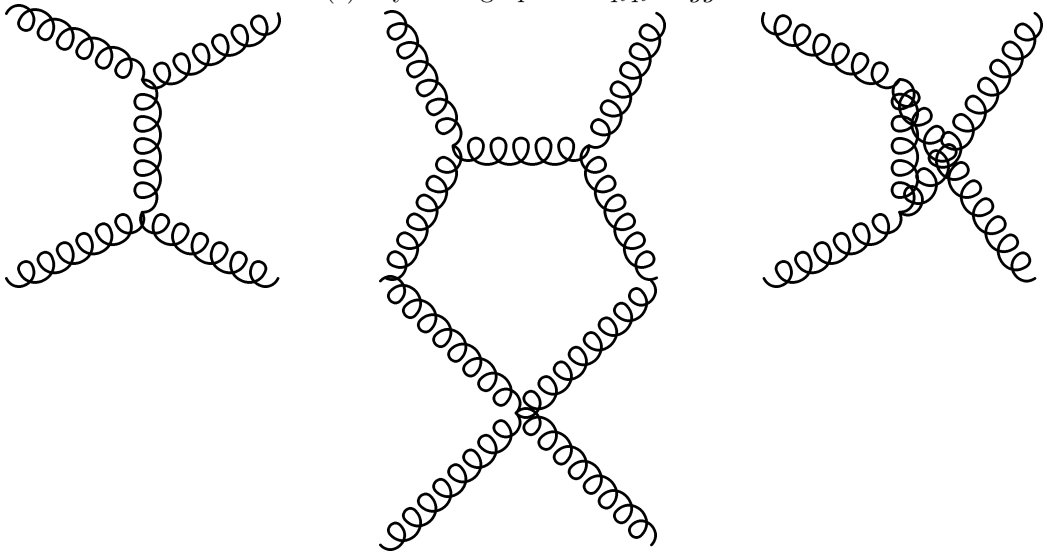
(a) Feynman graph for $q_i q_j \to q_i q_j$ for $i \neq j$



(b) Feynman graphs for $q_i q_i \to q_i q_i$ (identical quark flavors)



(c) Feynman graphs for $q_i \bar{q}_i \to gg$



(d) Feynman graphs for $gg \to gg$

Figure 2.9: Diagrams corresponding to the matrix elements needed to evaulate LO QCD partonic cross-sections.

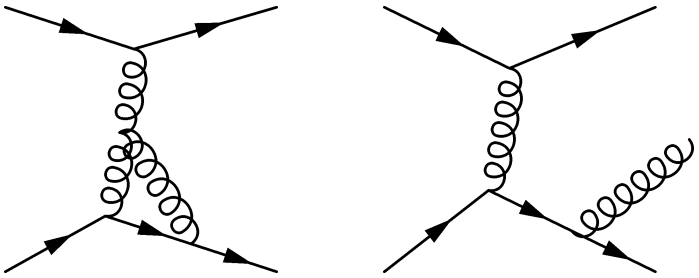


Figure 2.10: Examples of diagrams contributing to NLO QCD cross-sections.

$$\alpha_s^3 d\hat{\sigma}_3^{2\to X} = \alpha_s^2 \left(\alpha_s d\hat{\sigma}_3^{2\to 2,virtual}\right) + \alpha_s^3 \left(d\hat{\sigma}_3^{2\to 3,real}\right) \tag{2.51}$$

where $d\hat{\sigma}_3^{2\to 2,virtual}$ includes the virtual corrections to the $2 \to 2$ diagrams with one internal loop and $d\hat{\sigma}_3^{2\to 3,real}$ are the real $2 \to 3$ diagrams, of which one example each is shown in Figure 2.10.

The inclusive differential jet cross-section is then given by

$$\frac{d\sigma}{dE_{\mathrm{T}}dy}^{p+p\to\mathrm{jet}+X} = \sum_{ijk} \int dx_1 \int dx_2 f_{i/p}(x_1) f_{j/p}(x_2) \sum_n \alpha_s^n \frac{d\hat{\sigma}_n}{dE_{\mathrm{T}}dy}^{i,j\to k+X}(x_1, x_2) \tag{2.52}$$

where $\frac{d\hat{\sigma}_n}{dE_{\mathrm{T}}dy}^{i,j\to k+X}(x_1, x_2)$ is the partonic cross-section for producing a parton with (E_{T}, y) given incoming partons of energies $x_1\sqrt{s}/2$ and $x_2\sqrt{s}/2$, respectively. This was first computed to NLO order ($n = 2$ and $n = 3$) in [53].

Note that because the PDFs cannot be written in a closed form, neither can the cross-section. For an interesting comparison, consider the analogous calculation of $e^+e^- \to$ hadrons. There is only one diagram at leading order, which corresponds to the QED photoproduction of a quark-antiquark pair. In the high-energy, massless quark limit, the LO matrix element is almost identical to that for $e^+e^- \to \mu^+\mu^-$, save with the addition of SU(3) color factors, proper accounting for the fractional quark electric charges, and a sum over allowed quark flavors. Even the NLO cross-section, which is the first order to include QCD vertices, can be readily calculated[2] as

$$\sigma_{NLO}^{e^+e^-\to\mathrm{hadrons}}(Q^2) = \frac{4\pi\alpha_{QED}^2}{3Q^2} \sum_f (3Q_f^2) \left[1 + \frac{\alpha_s}{\pi} + \mathcal{O}\left(\alpha_s^2\right)\right] \tag{2.53}$$

where Q_f is the charge of quark flavor f. Clearly the initial hadronic state significantly complicates the situation. Although there are QCD processes for which the full NNLO cal-

culation is known, they are very specialized electroweak processes, such as Higgs or W-Higgs production[1], and NLO is the best order to which the inclusive jet (really, parton) cross-sections are known. A popular code for computing NLO cross-sections in hadronic collisions is `NLOJet++`[54; 55]. Even in a LO parton picture, the jet at the hadronic ("truth") level does not fully match the kinematics of the parton. In the literature[56; 57], the major reasons for the mismatch are sometimes separated into three categories.

The first is out of cone radiation (sometimes called "splash out"), which is the possibility that for a given resolution parameter R, some of the parton shower's energy may fall outside the area in which the jet energy is measured. This is a perturbative effect and calculations show that the energy difference in the hadronic jet with respect to the initial parton goes as $\sim \alpha_s p_{\mathrm{T}} \log(1/R)$, i.e. a larger jet radius recovers more of the energy. Although the values depend on the jet energy and resolution R, the typical energy difference for $R = 0.4$ is of the order 5-10%, with a larger effect for gluon jets (which have a softer, more diffuse fragmentation pattern) than for quark jets.

The second is the non-perturbative effects of hadronization, in which during the transition from parton shower to final-state hadrons the system loses some transverse momentum. This scale of this effect cannot be calculated from first principles, but ansatz calculations and MC implementations of hadronization models indicate that the energy loss is of the order $\sim$ 0.5-1 GeV $/R$, where the smaller (larger) value is for quark (gluon) jets.

The third effect only applies to hadronic collisions and is from the presence of the underlying event (UE) from the partonic remnants in the hadrons which did not participate in the hard scattering. Actually, due to the nature of QCD, it is impossible to separate this effect from the hadronization discussed above, since it is ambiguous whether any given hadron in the final state came from the UE or the parton shower, but they are often discussed separately. The susceptibility to the underlying event is directly proportional to the jet area $\sim R^2$. Very generally, this effect necessitates an experimental procedure to subtract the additional energy. Taken together, these imply that there are a range of value of R which are not so small that the out of cone and hadronization corrections are large but not so large that the effects of the underlying event cannot be controlled. For jets in pp collisions at LHC energies, some studies have shown that $R \approx 0.4$-0.6 is ideal[57], depending on the composition of quark and gluon jets. For high-μ (number of mean interactions per crossing) pp collisions such as that delivered by the LHC towards the end of Run

I, the UE is largely from uncorrelated additional collisions. The procedure in recent ATLAS jet results[58] is to subtract a mean transverse energy for each additional reconstructed vertex in the event taken from that observed in minimum bias pp collisions, but the details are $\langle\mu\rangle$, N_{PV} and jet radius dependent.

On the other hand, the UE in ion-ion collisions is a dramatically different phenomenon. In $\sqrt{s} = 2.76$ TeV Pb+Pb collisions, the mean E_{T} at mid-rapidity can reach ~ 250 GeV per unit $d\eta d\phi$[59]. Furthermore, the UE is ion-ion collisions is extremely correlated from point to point, making naive UE subtraction proposals unworkable. The systematic subtraction of the underlying event pedestal from jets in a way that respects the correlation and minimizes the bias on the reconstructed jets is a major issue in heavy ion physics. ATLAS has developed a procedure which subtracts an azimuthally-modulated background in a rapidity-dependent way while being careful to exclude the jet energy from the determination of the UE[60]. The procedure used in the p+Pb jet analysis implements this procedure, which is described in detail in Section 8.3.1.

Now consider the kinematics of dijets in the leading order $2 \to 2$ picture of a parton-parton scattering with momentum transfer Q^2, where the partons have momentum fraction x_1 (e.g. $p^{\mu}_{part} = xp^{\mu}_{had}$) and x_2 and the hadron-hadron center of mass energy is $\sqrt{s}$. In the laboratory frame (hadrons have $p^{\mu}_{had} = (\sqrt{s}/2, 0, 0, \pm\sqrt{s}/2)$), the parton-parton center of mass system is

$$p^{\mu}_{part,1} + p^{\mu}_{part,2} = x_1 p^{\mu}_{had,1} + x_2 p^{\mu}_{had,2} = \left((x_1 + x_2)\frac{\sqrt{s}}{2}, 0, 0, (x_1 - x_2)\frac{\sqrt{s}}{2}\right) \tag{2.54}$$

$$\Rightarrow (x_1 p^{\mu}_1 + x_2 p^{\mu}_2)^2 = x_1 x_2 s \tag{2.55}$$

At leading order, the partons exchange momentum Q^2 and come out azimuthally back-to-back with $E_{\mathrm{T}} = Q$. In the parton-parton center of mass frame, the partons have rapidity y^* and $-y^*$, where $E_{\mathrm{T}} \cosh(y^*) = \sqrt{x_1 x_2 s}$. This parton-parton COM frame is related to the hadron-hadron COM frame (also the lab frame) via a boost $\beta = (x_1 - x_2)/(x_1 + x_2)$ which can also be written as the rapidity of the dijet system $y^{\mathrm{dijet}} = \tanh^{-1}(\beta)$. It follows that in the lab frame, the dijets appear with $y_1 = +y^* + y^{\mathrm{dijet}}$ and $y_2 = -y^* + y^{\mathrm{dijet}}$. Finally, one can relate the rapidities and transverse energies of the dijet system to the original x_1 and x_2 of the hard-scattering,

$$x_1 = \frac{E_{\mathrm{T}}}{\sqrt{s}}\left(e^{+y_1} + e^{+y_2}\right) \tag{2.56}$$

$$x_2 = \frac{E_{\mathrm{T}}}{\sqrt{s}}\left(e^{-y_1} + e^{-y_2}\right) \tag{2.57}$$

In reality, as discussed above, higher order contributions as well as initial and final state radiation must be taken into account, which results in event-by-event variations in the angular and transverse momentum balance of the partonic jets.

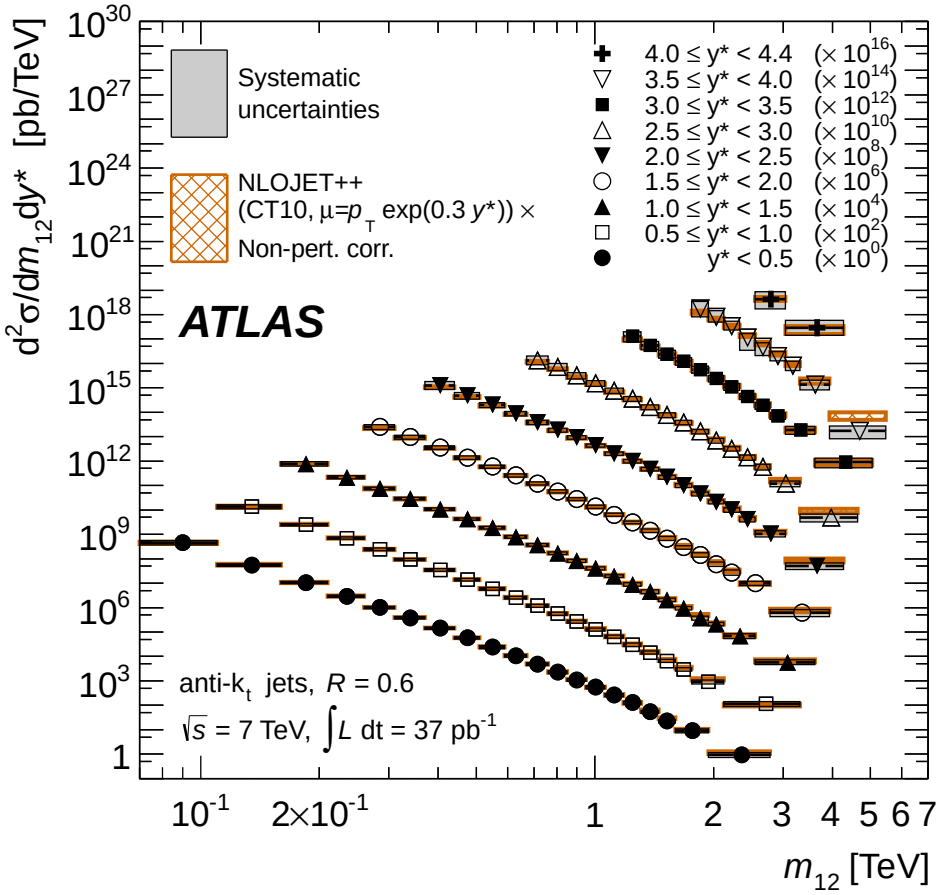


Figure 2.11: Double-differential dijet cross-section as a function of mass and y, as measured by ATLAS in 7 TeV pp collisions, from [58].

A recent measurement which compares the current best available NLO jet cross-sections and

experimental techniques in hadronic jet reconstruction is the measurement of $R = 0.4$ and $R = 0.6$ anti-k_{T} dijet pairs in 7 TeV pp collisions by ATLAS[58]. The dijet mass spectrum is shown in Figure 2.11, demonstrating excellent agreement with theory over almost 9 units of rapidity and nine orders of magnitude in the cross-section.

2.2 QCD at high densities and high temperatures

2.2.1 Geometry of nuclear collisions

Early p+A experiments were motivated, among other reasons, by the observation that hadronic showers in h+p collisions develop over a longer distance scale than the mean free path for h-nucleon interactions in a dense nucleus[61]. Thus, the showers could be studied in the early stages of their development by comparing h+A collisions to a h+p baseline. For hadron-nucleon (for definiteness, and in the case of QCD interactions where isospin is a good symmetry, we can take the nucleon to be a proton) cross-section σ_{hp}, and hadron-nucleus cross-section σ_{hA} on a nucleus of size A, the mean number of interactions $\bar{\nu}$ is assumed to given by

$$\bar{\nu} = A\sigma_{hp}/\sigma_{hA} \tag{2.58}$$

In early fixed target h+A experiments without the ability to trigger, only the mean number of interactions was seen. Varying $\bar{\nu}$ was typically performed by varying the probe hadron (for example, π^+, K^+ and p in some Fermilab experiments[62]) and the target nucleus A (ranging from light elements like Be (A = 9) to heavier targets including Au (A = 197) or Pb (A = 208)).

Very generally, the basic observable for investigating nuclear effects on particle production is the ratio of yields

$$R_A = \langle n \rangle_{hA} \,/\, \langle n \rangle_{hp} \tag{2.59}$$

The mean multiplicity $\langle n \rangle_{hA}$ was typically measured for different identified particles and also characterized in terms of its dispersion, $D = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}$.

Since then, much work has been done to determine the collision geometry on an event-by-event basis. Before RHIC, it was common to determine a relationship between ν and the observed number of *grey tracks* N_{gray}, which were consistent with slow forward protons and deuterons knocked out of

the nucleus[63; 64], and the number of *black tracks* N_{black}, consistent with even slower evaporated particles produced by the cooling nucleus but which were less correlated with ν. (The names *grey* and *black* were historically from emulsion experiments, where the difference in color came from the difference in grain density between tracks of different momenta.) For example, this was done at the E910 experiment at the BNL AGS[65]. Thus, selecting on N_{grey} would select different configurations of the p+A geometry.

Following this, heavy ion experiments at colliders (RHIC and the LHC) all determine the geometry of p+A and A+A collisions in each event by examining some measure of the soft underlying event which is correlated with the global geometry. This is done through a non-relativistic formulation of nucleus-nucleus collisions called the Glauber model.

The Glauber model[66; 67; 68] formulates nucleus-nucleus collisions as the superposition of many smaller nucleon-nucleon collisions between the individual nucleons in the colliding nuclei. It assumes that each nucleon continues on a straightline path through the opposing nucleus, is not substantially deflected by any interactions and has the same probability for interacting with a nucleon at the beginning of the nucleus as it does when it comes out the other end. The model needs only two external inputs to specify the geometry: the density profile of the relevant nuclei $\rho(\vec{r})$ as a function of the distance $\vec{r}$ from the center of mass and the inelastic nucleon-nucleon cross-section $\sigma_{\rm inel}^{NN}$. The density is typically taken to be a Woods-Saxon distribution (although there are other parameterizations),

$$\rho_A(\vec{r}) = 4\pi r^2 \rho_A(r) = 4\pi r^2 \rho_0 \left(\frac{1 + w(r/R)^2}{1 + \exp\left(\frac{r-R}{a}\right)} \right) \tag{2.60}$$

R is the nuclear radius parameter, a is the "skin depth" and w describes non-spherical distortions of the nucleus. The relevant heavy nuclei in this work, with densities measured by elastic electron scattering experiments, are ^{197}Au ($R = 6.38$ fm, $a = 0.54$ fm, $w = 0$) and ^{208}Pb ($R = 6.62$ fm, $a = 0.546$ fm, $w = 0$)[69]. The deuteron is parameterized with a Hulthen wavefunction, with the resulting probability density function

$$\rho_d(\vec{r}) = 4\pi r^2 \rho_d(r) = 4\pi r^2 \rho_0 \left(\frac{\exp(-\alpha r) + \exp(-\beta r)}{r} \right)^2 \tag{2.61}$$

where $\alpha = 0.228$ fm^{-1}, $\beta = 1.18$ fm^{-1}[70; 71], and ρ_0 is set by the normalization requirement on the probability distribution.

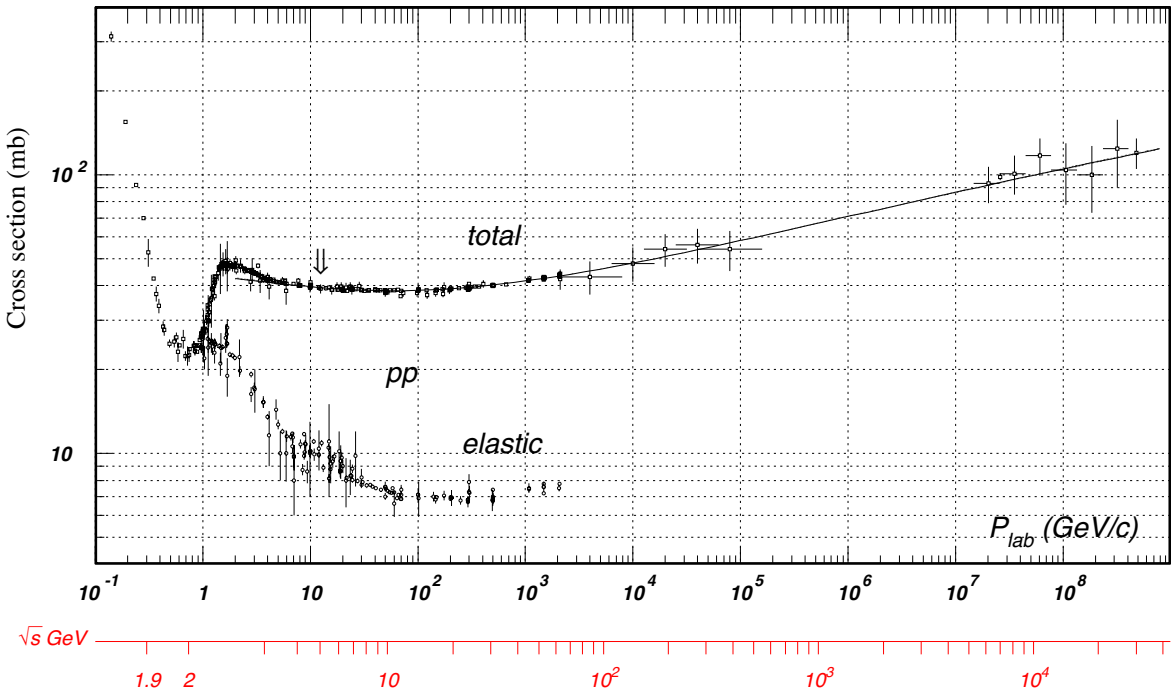


Figure 2.12: World data on p+p total and elastic cross-sections, from [1].

Obtaining an accurate value for the nucleon-nucleon inelastic cross-section $\sigma_{\text{inel}}^{NN}$ can be complicated by several factors. Direct measurements of the cross-section are complicated by the fact that the observed cross-section must be extrapolated to a region of diffractive mass below what a given detector can measure. For example, in the recent ATLAS measurement[72] of the proton-proton cross-section at $\sqrt{s} = 7$ TeV, $\sigma_{\text{inel}}^{pp} = 69.4 \pm 2.4(\text{exp}) \pm 6.9(\text{extr})$ mb, a non-trivial 10% error from the extrapolation, although a very recent TOTEM measurement[73] of $\sigma_{inel}^{pp} = 72.9 \pm 1.5$ mb at this energy was performed using a promising method based on the optical theorem. Unfortunately, the cross-section may not even be measured at the desired $\sqrt{s}$; for example, the widely assumed $\sigma_{\text{inel}}^{pp} = 42$ mb at 200 GeV in RHIC Glauber analyses has never been published. In this case, it must be extracted from fits to world data on the total and elastic cross-section via $\sigma_{\text{inel}} = \sigma_{\text{tot}} - \sigma_{\text{elastic}}$, shown in Figure 2.12. For example, one may use the methodology based on [74]. In any case, the $\sigma_{\text{inel}}^{NN}$ used in the Glauber model is varied to generate a systematic uncertainty which is typically the dominant source of uncertainty for many geometry-related observables.

With external information about ρ and σ_{inel}, we can describe the rest of the Glauber formalism. In the following derivations, we assume that ρ is normalized so that $\int d^3\vec{r}\rho(\vec{r}) = 1$. $\hat{T}_A(\vec{s})$ is the probability per unit area of finding a nucleon at transverse position $\vec{s}$, inclusive of z,

$$\hat{T}_A(\vec{s}) = \int dz\rho(\vec{s}, z) \tag{2.62}$$

Then $\hat{T}_{AB}(\vec{b})$ is the *nuclear overlap function* (or *thickness function*) for nuclei A and B separated by an impact parameter b, defined by

$$\hat{T}_{AB}(\vec{b}) = AB \int d^2\vec{s}\hat{T}_A(\vec{s})\hat{T}_B(\vec{s} - \vec{b}) \tag{2.63}$$

$\hat{T}_{AB}$ can be thought of as a nucleonic luminosity. (The notation above is slightly different than used in some other derivations, where $\hat{T}_{AB}$ may not have the AB factor and is instead a probability per unit area.) Given an inelastic nucleon-nucleon cross-section $\sigma_{\text{inel}}^{NN}$, the quantity

$$N_{\text{coll}}(\vec{b}) = \hat{T}_{AB}(\vec{b})\sigma_{\text{inel}}^{NN} \tag{2.64}$$

is the effective number of nucleon-nucleon collisions. Alternately, Equation 2.64 can be rewritten as the expectation value of AB throws of the probability $P_{1+1}(\vec{b})$ of a nucleon drawn from each distribution to collide,

$$N_{\text{coll}}(\vec{b}) = ABP_{1+1}(\vec{b}) = AB\left(\sigma_{\text{inel}}^{NN} \int d^2\vec{s}\hat{T}_A(\vec{s})\hat{T}_B(\vec{s} - \vec{b})\right) \tag{2.65}$$

Going further, we build the probability $P(n; \vec{b})$ of n collisions by considering a combinatoric separation of the AB possible collisions into n one-on-one collisions and $AB - n$ misses,

$$P_{A+B}(n; \vec{b}) = \binom{AB}{n} \left[P_{1+1}(\vec{b})\right]^n \left[1 - P_{1+1}(\vec{b})\right]^{AB-n} \tag{2.66}$$

Then the A+B inelastic cross-section $\sigma_{\text{inel}}^{A+B}$ at fixed impact parameter $\vec{b}$ is the probability of having at least one collision,

$$\sigma_{\text{inel}}^{A+B}(\vec{b}) = \sum_{n=1}^{n=AB} P_{A+B}(n; \vec{b}) = 1 - P(0, \vec{b}) = 1 - \left[1 - P_{1+1}(\vec{b})\right]^{AB} \tag{2.67}$$

and the total (b-integrated, or minimum bias) A+B inelastic cross-section is

$$\sigma_{\text{inel}}^{A+B} = \int d^2\vec{b}\sigma_{\text{inel}}^{A+B}(\vec{b}) = 2\pi \int_0^{+\infty} (bdb)\sigma_{\text{inel}}^{A+B}(|b|) \tag{2.68}$$

Finally, by analogy to Equations 2.64 and 2.66, we can construct the nucleon-nucleus inelastic cross-sections for a nucleon in A interacting with B at position $\vec{s}$ (or vice versa with $A \leftrightarrow B$),

$$\sigma_{\text{inel}}^{B}(\vec{s}) = \sum_{n=1}^{n=B} \binom{B}{n} \left[\sigma_{inel}^{NN}\hat{T}_B(\vec{s})\right]^n \left[1 - \sigma_{inel}^{NN}\hat{T}_B\right]^{B-n} \tag{2.69}$$

Then the effective number of participating nucleons N_{part} is

$$N_{\text{part}}(\vec{b}) = A \int d^2\vec{s}T_A(\vec{s})\sigma_{\text{inel}}^{B}(\vec{s}) + B \int d^2\vec{s}T_B(\vec{s}-\vec{b})\sigma_{\text{inel}}^{A}(\vec{s}-\vec{b}) \tag{2.70}$$

In principle, the integrals defined above can be evaluated analytically or numerically in what is called the *optical form* of the Glauber model. The optical Glauber model thus gives (not necessarily integer) values of the impact parameter dependent and mean N_{part} and N_{coll}, as well as the total inelastic cross-section $\sigma_{\text{inel}}^{AB}$. A limitation of the optical model is that it does not encode any information about the correlations of nucleon positions (e.g. their locations are independent of one another) or their fluctuations event by event (e.g. N_{part} is always the same for a given $|b|$). This is actually a larger deficiency in the model than might at first appear, since the optical Glauber model always gives perfectly symmetric collisions along the plane perpendicular to $\vec{b}$ and thus has no way to generate odd anisotropy moments v_3, v_5, etc.

To include all of these effects, it is much more common to use an MC procedure[75; 76] to simulate a large number of individual A+B collisions, each of which has its own N_{part}, N_{coll} and $\vec{b}$, and then look at the distributions of these values and the correlations between them. (On the other hand, every event with impact parameter b looks exactly the same in the optical model.) In the Glauber MC, instead of a smooth density distribution, the nuclei are populated event-by-event with discrete nucleons whose positions are drawn randomly according to $P(x,y,z) = 4\pi(x^2+y^2+z^2)\rho(r=\sqrt{x^2+y^2+z^2})$. (In practice, a nucleon position is only populated if it some minimum distance d_{min} away from all other nucleons.) After the nucleons have been placed, the nuclei centers are separated by impact parameter b in the transverse plane, with b drawn from $P(b) \approx b$. The straight-line trajectories of all nucleons are considered. Inelastic nucleon-nucleon

collisions are treated semiclassically, in that a pair of nucleons from the different nuclei are said to collide if the distance r_{AB} between then satisfies

$$r_{AB} < \sqrt{\frac{\sigma_{inel}^{NN}}{\pi}} \tag{2.71}$$

Then, $N_{\rm part}$ is total number of nucleons (on either side) which underwent ≥ 1 collision, and $N_{\rm coll}$ is the number of nucleon-nucleon collisions. For an A+B collision with at least one interaction, $2 < N_{\rm part} < A + B$ and $1 < N_{\rm coll} < AB$ (in practice, $N_{\rm coll}$ is far smaller than this limit, since only in an extremely pathological geometry is it possible for a nucleon in A to collide with *every* nucleon in B). Thus, by statistically sampling many simulated events, $N_{\rm part}$ and $N_{\rm coll}$ distributions can be determined as a function of b. In addition, many other important geometric quantities can be calculated, such as the eccentricity of the struck nucleon positions in the transverse plane as a function of impact parameter, but these are not discussed here.

Figure 2.13 shows an example of an MC Glauber simulation of a Au+Au collision. The nuclei are represented as a superposition of nucleons with an effective transverse area $\pi r_{AB}^2 = \sigma_{inel}^{NN}$, and the nuclei are displaced via an impact parameter b between their centers of mass. The struck nucleons (those contributing to $N_{\rm part}$ and $N_{\rm coll}$ for this event) are shaded more darkly than the non-participating (spectrator) nucleons. Generally, simple geometric arguments show that for A+B collisions,

$$\langle N_{\rm coll} \rangle \sim AB/(R_A + R_B)^{1/3} \tag{2.72}$$

where $R_A \sim A^{1/3}$ is the typical dimension of the nucleus. Thus for p+A collisions, $\langle N_{\rm coll} \rangle \sim A^{1/3}$ while for A+A collisions, $\langle N_{\rm coll} \rangle \sim A^{4/3}$.

Experimentally, $N_{\rm part}$, $N_{\rm coll}$ and b cannot be measured directly. Instead, events are ranked by some measurement of the soft underlying event (hereafter generically referred to as $E_{\rm T}$, though it could also be a multiplicity) which characterizes the size of the underlying event (generally related to higher-$N_{\rm part}$ collisions). The Glauber description of the collision geometries must then be associated with the $E_{\rm T}$ distribution observed in data. Let $P(E_{\rm T}; N_{\rm part})$ be the probability density function of an event with $N_{\rm part}$ participants resulting in a signal of $E_{\rm T}$. Then the total observed distribution $dN/dE_{\rm T}$ can be related to the Glauber distribution $dN/dN_{\rm part}$ via

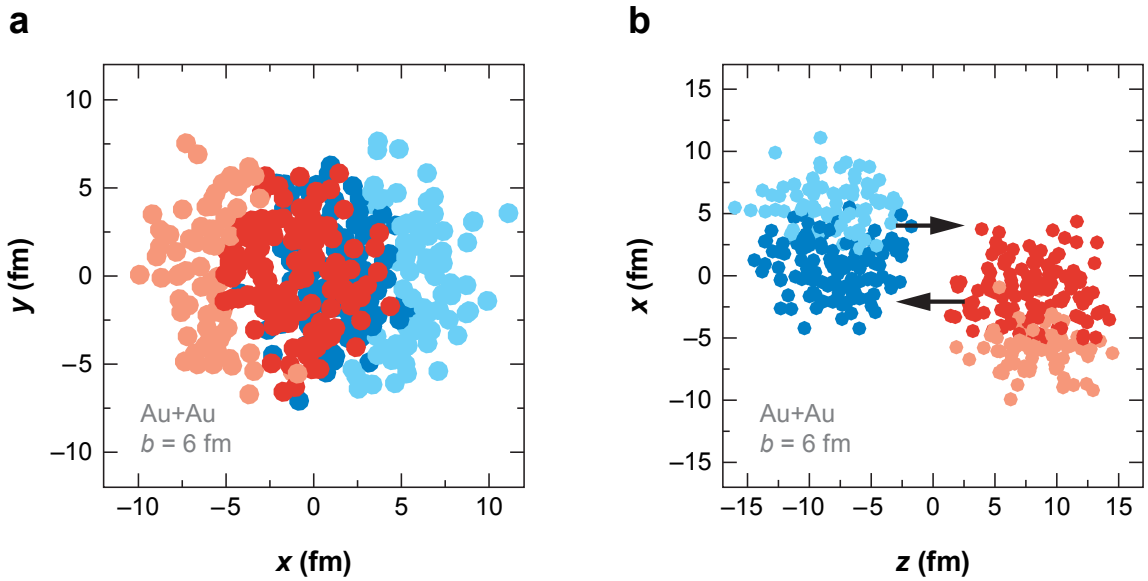


Figure 2.13: Example of an MC Au+Au Glauber event with $b = 6$ fm, shown in the plane perpendicular to the beam direction (left) and in a plane which includes the beam direction (right). The circles represent the positions of nucleons with effective radius r_{AB}. Different colored nucleons are from different nuclei, with the participating nucleons drawn with a darker color, from [68].

$$\frac{dN}{dE_{\mathrm{T}}} \propto \epsilon(E_{\mathrm{T}}) \sum_{N_{\mathrm{part}}} P(E_{\mathrm{T}}; N_{\mathrm{part}}) \frac{dN}{dN_{\mathrm{part}}} \tag{2.73}$$

where $\epsilon(E_{\mathrm{T}})$ is the detector trigger efficiency, which is necessary to properly deal with possible inefficiencies in the peripheral part of the inelastic A+B cross-section. The challenge then rests on finding a reasonable description of soft particle production $P(E_{\mathrm{T}}; N_{\mathrm{part}})$ and efficiency ϵ that Equation 2.73 matches the data well. In the past, it has been common to describe P with a negative binomial distribution (which is discrete but usable since E_{T} will very likely use discrete binning) with mean μ and parameter k,

$$P_{NBD}(x;k,\mu) = \left(\frac{k}{k+\mu}\right)^k \left(\frac{(x+k-1)!}{(k-1)!x!}\right) \left(\frac{\mu}{k+\mu}\right)^x \tag{2.74}$$

$$\tag{2.75}$$

or with a continuous Gamma distribution described by a shape parameter k and scale parameter θ,

$$P_{Gamma}(x;k,\theta) = \frac{x^{k-1}\exp\left(-x/\theta\right)}{\Gamma(k)\theta^k} \tag{2.76}$$

where the mean of the distribution is $k\theta$ and variance is $k\theta^2$. Both the NBD and the Gamma distribution have the useful property that the N-fold convolution of the function are the functions with a simple scaling of their parameters,

$$P_{NBD}(x;k,\mu)^{*N} \rightarrow P_{NBD}(x;Nk,N\mu) \tag{2.77}$$

$$P_{Gamma}(x;k,\theta)^{*N} \rightarrow P_{Gamma}(x;Nk,\theta) \tag{2.78}$$

From here on out, we will use the Gamma distribution as an example.

To model the changing signal distribution with increasing $N_{\rm part}$, the function parameters (k, μ or k, θ) are themselves allowed to vary with $N_{\rm part}$, and the model of particle production typically makes the assumption that the distribution of $E_{\rm T}$ scales in some simple way with $N_{\rm part}$.

The simplest possibility is the *wounded nucleon* (WN) model[77], in which the mean particle multiplicity scales directly with the number of participants,

$$k^{WN}(N_{\rm part}) = \frac{N_{\rm part}}{2}k_0 \tag{2.79}$$

$$\theta^{WN}(N_{\rm part}) = \theta_0 \tag{2.80}$$

where the Gamma distribution with k_0 and θ_0 describes the soft particle production in pp collisions.

Generally, the application of the Glauber model to data requires choosing $P(E_{\rm T};N_{\rm part})$ (for the Gamma distribution case, $k(N_{\rm part})$ and $\theta(N_{\rm part})$) in a way that is well motivated and satisfies

Equation 2.73. The details of how this is performed in the present work is deferred to the relevant experimental sections (Section 6.2.3 for d+Au collisions in PHENIX and all of Section 7 for p+Pb collisions in ATLAS).

After this, events in data are assigned a *centrality*, which is their E_{T} percentile among all events. For example, 0-10% centrality events are those with E_{T} in the highest tenth of all events. Then the mean N_{part} (or, similarly, N_{coll}) can be calculated for events in that centrality class via

$$\langle N_{\mathrm{part}} \rangle^{cent} = \int_{E_{\mathrm{T}}^{\mathrm{low}}}^{E_{\mathrm{T}}^{\mathrm{high}}} dE_{\mathrm{T}} \sum_{N_{\mathrm{part}}} N_{\mathrm{part}} \times \frac{dN}{dN_{\mathrm{part}}} \times P(E_{\mathrm{T}}; N_{\mathrm{part}}) \tag{2.81}$$

where $(E_{\mathrm{T}}^{\mathrm{low}}, E_{\mathrm{T}}^{\mathrm{high}})$ define the limits of the centrality range. Similarly, events in a given centrality category are also assigned a mean N_{coll} and $T_{\mathrm{AB}} = N_{\mathrm{coll}}/\sigma_{inel}^{NN}$.

The cartoon in Figure 2.14 illustrates the relationship between the geometric Glauber quantities and the signal observed in data. The left side of the axes corresponds to events with a larger impact parameter b, a smaller number of participant nucleons N_{part}, a smaller value of the soft underlying event observable (here denoted N_{ch}) and a larger event centrality (e.g. a higher centile when events are sorted in order of decreasing N_{ch}). On the other hand, the right side of the axes correspond to events with a smaller b, a larger N_{part}, larger typical values of N_{cm} and small event centrality.

With this formalism it is possible to define a more sophisticated version of Equation 2.59 which compares the yield of particles in a given centrality of A+A (or A+B, or p+A) collisions to the yield observed in pp collisions, scaled by the number of nucleon-nucleon collisions ${N_{\mathrm{coll}}}^{cent}$ in that centrality bin. This is called the *nuclear modification factor*, styled R_{AA} for symmetric A+A collisions (or R_{dAu} and R_{pA} for d+Au and p+A, respectively),

$$R_{\mathrm{AA}}(p_{\mathrm{T}}, \eta) = \frac{(1/N_{evt}^{cent}) d^2N^{cent}/dp_{\mathrm{T}}d\eta(p_{\mathrm{T}}, \eta)}{\langle T_{\mathrm{AB}} \rangle^{cent} d^2\sigma^{pp}/dp_{\mathrm{T}}d\eta(p_{\mathrm{T}}, \eta)} \tag{2.82}$$

where $(1/N_{evt}^{cent})(d^2N^{cent}/dp_{\mathrm{T}}d\eta)$ is the per-event yield of particles in the centrality bin, $\langle T_{\mathrm{AB}} \rangle^{cent}$ is the mean value of nuclear overlap function in the bin, and $d^2\sigma^{pp}/dp_{\mathrm{T}}d\eta$ is the cross-section in pp collisions. Since $\langle T_{\mathrm{AB}} \rangle$ can be thought of as a nucleonic luminosity, $1/\langle T_{\mathrm{AB}} \rangle \times d^2N/dp_{\mathrm{T}}d\eta$ is effectively the per-nucleon cross-section for events in the given centrality class. Generally, $R_{\mathrm{AA}} < 1$ is referred to as *suppression*, indicating fewer produced particles at the given (p_{T}, η) than would be expected in comparison to naive geometric-scaled cross-section in pp collisions, while $R_{\mathrm{AA}} > 1$

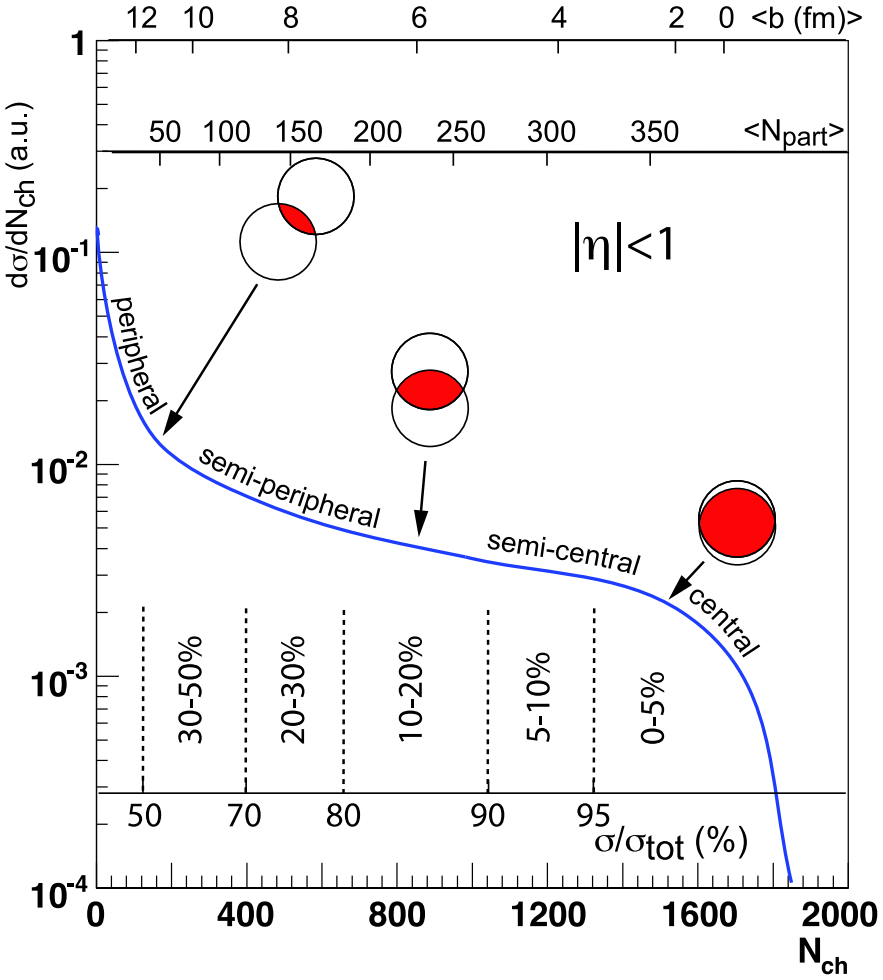


Figure 2.14: Cartoon illustrating the relationship between the impact parameter b, the number of participants $N_{\rm part}$, the geometric orientation of the nuclei, the distribution of the soft underlying event observable and the event centrality cuts, in A+A collisions, from [68].

is referred to as *enhancement*, and indicates the opposite.

In lieu of a pp cross-section reference, or to examine the differences in particle production between the different centralities more directly, the nuclear modification factor can be constructed as the central to peripheral ratio $R_{\rm CP}$,

$$R_{\rm CP}(p_{\rm T},\eta) = \frac{(1/\langle T_{\rm AB}\rangle^{central})(1/N_{evt}^{central})d^2N^{central}/dp_{\rm T}d\eta(p_{\rm T},\eta)}{(1/\langle T_{\rm AB}\rangle^{peripheral})(1/N_{evt}^{peripheral})d^2N^{peripheral}/dp_{\rm T}d\eta(p_{\rm T},\eta)} \tag{2.83}$$

The methodology above describes the most basic version of the application of the Glauber

model to data. Other variants have been investigated, such as modeling the particle production as a function of $N_{\rm coll}$ instead of $N_{\rm part}$. Another popular variant is the *two-component model*[78; 79], in which the mean multiplicity is taken to scale with

$$(1-x)N_{\rm part}/2 + xN_{\rm coll} \tag{2.84}$$

where $0 < x \ll 1$ describes the contribution to the total multiplicity from hard scattering events. Additional extensions of the model have been proposed in which the participating particles are the constituent quarks of the nucleons instead of the nucleons themselves[80; 81]. In this model, three quarks from each nucleon are populated around the nucleon position according to a probability density distribution

$$\rho(r) \propto \exp\left(-\sqrt{12}r/r_m\right) \tag{2.85}$$

where $r_m = 0.811$ fm is the RMS charge radius of the proton. In an analogue of Equation 2.71, two quarks interact if they pass within a distance

$$r_{qq} < \sqrt{\frac{\sigma_{qq}^{inel}}{\pi}} \tag{2.86}$$

of one another, where the σ_{qq}^{inel} is the "inelastic constituent quark-quark cross-section" and is set by requiring Glauber simulations of pp collisions to reproduce the correct total $\sigma_{\rm NN}^{inel}$.

Furthermore, there are extensions to the Glauber model called the *Glauber-Gribov model* which are based on the principle that the nucleon-nucleon cross-section right at the moment of interaction must surely fluctuate from event to event[82; 83; 84]. In proton-proton collisions, this effect is already baked into the definition of the average cross-section. However, in p+A collisions, the particular configuration of quarks and gluons present in the proton when it reaches the nucleus is frozen for the duration of the proton's trajectory through the nucleus. Thus, a particularly "strong" proton configuration would have a higher than average cross-section for collisions with *all* the nucleons in its path, and vice versa. The effect is likely not pronounced in A+A collisions, where the color fluctuations among the ≈ 200 participants / nucleus are likely to cancel each other out on average. But in p+A collisions, where the cross-section is controlled by the configuration of the proton, the effect is potentially large. A simple way to model this phenomenon within the

Glauber MC framework is to introduce a probability distribution for the nucleon-nucleon inelastic cross-section $P(\sigma_{\mathrm{NN}})$, which is sampled every event.

In this work, the centrality determination in p+Pb collisions is performed using the standard Glauber procedure as well as the Glauber-Gribov model with two different parameterizations of the size of the event-by-event fluctuations.

2.2.2 Cold Nuclear Matter effects and pQCD

Perturbative QCD (pQCD) calculations should be applicable to hard processes in collisions involving nuclei. However, the distributions of nucleons inside nuclei and their interactions with one another do not appear explicitly in QCD. Naively, one could imagine the extreme model in which p+A collisions as a simple superposition of individual proton-nucleon collisions. In this view, nucleons don't "know" that they are in a nucleus and the short-range physics of the partonic hard scattering and long range physics of the parton shower and hadronization are not affected by the presence of additional nuclear material.

In that case, it is possible to imagine that the partonic structure of the nucleus is simply the equivalent to the partonic structure of its nucleon constituents, convoluted by the Fermi motion of the nucleons inside the nucleus[85]. Thus, the nuclear parton distribution function $f_{q/A}(x)$ can be written as the convolution over the standard parton distribution function $f_{q/p}(y)$ and the distribution of nucleon momenta z inside the nucleus $f_{p/A}(z)$,

$$f_{q/A}(x) = \int dy \int dz f_{q/p}(y) f_{p/A}(z) \delta(x - yz) \tag{2.87}$$

where the delta function picks out partons with energy $x = yz$. Thus, the only effect is that parton momentum fractions in nuclei are "smeared" by the relative motion of the nucleons. Experiment has shown that at least for low and moderate values of Q^2, this is not the case.

Modifications to the rates and details of hard scattering processes due to the nuclear environment are generally called *cold nuclear matter* (CNM) effects. There are many CNM effects observed or predicted to be present in data, including partonic energy loss, shadowing, the transverse momentum broadening of partons, the decorrelation of back to back jet pairs and even the modification of fragmentation functions. These are discussed below. There are also the so-called "isospin effects", which are the trivial effect that the up and down quark content in nuclei (which

have Z protons and $A-Z$ nucleons) is different than that in protons. Thus, calculations of the the production of direct photons in nuclear collisions (at leading order, through the $qg \to q\gamma$ diagram, but with a different magnitude of the matrix element for $Q = +2/3$ up quarks and $Q = -1/3$ down quarks) must take this into account. In the present analysis, where isospin-blind QCD diagrams dominate the production of jets, these are ignored.

There are several theoretical approaches in trying to understand the origins and observable effects of cold nuclear matter effects. First, there are approaches to casting cold nuclear matter effects in a factorization and pQCD formalism, in which the various effects arise from the multiple scattering of hard partons on the background partons inside the nucleus[86]. Second, there are approaches in which the nuclear modification is assumed to be universal in nuclear-x and Q^2 and are folded into a set of nuclear parton distribution functions (nPDFs). Third, there are approaches which focus on the non-linear evolution of QCD at low-x and attempt to explicitly construct the partonic wavefunction in the nucleus. The latter two are discussed in the subsequent subsections.

At high-Q^2, it is possible to characterize the contributions to the QCD scattering diagrams by the *twist* of the operator[87; 88], where the twist $t = d-m$ is equal to the dimension of the operator minus its spin. The leading term is $t = 2$ (*leading twist*) which describes the contributions from single scatterings. However, there are corrections from multiple scatterings of the partons involved with other (soft) partons in the hadrons. These are called the *higher twist* terms, $t = 4, 6, \ldots$, are generally suppressed by powers of Q^{2-t}. In nuclei of mass number A, the leading twist matrix elements are naturally proportional to $\approx A$, since there are more partons available for scattering. However, the next higher-twist term includes not only the effects of multiple scatterings within the same nucleon, but of multiple scatterings on partons in neighboring nuclei. Thus, the higher twist term has a different dependence and grows as $\approx A^{4/3}$[89; 90; 91]. Thus, in nuclei the relative contribution of the $t = 4$ elements is enhanced and the ratio to the $t = 2$ leading twist elements is

$$\approx \alpha_s A^{1/3}/Q^2 \tag{2.88}$$

The greater role of these terms in collisions involving nuclei can be tested experimentally in p+A collisions. However, it is important to note that all higher twist effects should decrease with increasing Q^2. The phenomenon of very energetic partons from high-Q^2 scatterings being unaffected by the nuclear environment is called *color transparency*. In this regime, only the effects of the Fermi

motion (in Equation 2.87) should remain. The broadening of a parton distribution function which is steeply falling in x would therefore result in an *increase* in the rate of hard processes at fixed x.

Historically, one of the most recurring and visible CNM effects is the *Cronin effect*, after a series of fixed target experiments at Fermilab[92; 93]. Cronin et al. compared the invariant cross-section $Ed^3\sigma_{hA}/d^3p$ in h+A collisions for a variety of different A, to the invariant cross-section for a hydrogen target ($A = 1$), and modeled the ratio as a power of the nuclear number A^α. In the language of Equation 2.58, if the cross-sections exhibited $N_{\rm coll}$-scaling, we would expect $\frac{d\sigma_{hA}/dp^3}{d\sigma_{hp}/dp^3} = \bar{\nu}\frac{\sigma_{hA}}{\sigma_{pA}} = A$, e.g. $\alpha = 1$. On the contrary, $\alpha > 1$ for $p_{\rm T} > 2$ GeV charged particles, reaching as high as $\alpha \approx 1.3$ for protons and antiprotons. There competing explanations for the origin of the effect[94], but a popular one is that it is caused by the soft multiple rescattering of a high-$p_{\rm T}$ particle with other soft particles in the nucleus[95]. This results in a broadening of the *transverse* momentum, in which the partons are pushed closer to mid-rapidity. Jet broadening can be well explained within a higher twist formalism[96].

Recent data[97; 98] seem to show that at fixed $p_{\rm T}$ the Cronin enhancement is larger at smaller collision energies, which could support this idea, since the particle spectra are steeper at lower $\sqrt{s}$. Historically, the stronger Cronin enhancement in baryons as opposed to mesons (for example, as measured by PHENIX[99]) has also been a topic of much interest.

Another historical idea is that of *energy loss*, where a high-energy parton escapes the nucleus with less energy than it would have in a simple pp collisions. At a basic level, a colored parton traversing the nuclear medium should be expected to lose energy similar to the way in which an (electrically) charged particle passing through ordinary matter loses energy through collisional modes (where the charged particle ionizes or interacts with the atomic medium) and radiative modes (bremmstrahlung radiation in the vicinity of atoms). This was first observed in the $\alpha < 1$ dependence of fixed target p+A collisions with multiple nuclear targets[100]. Since then, there has been a large body of work attempting to quantitatively describe energy loss in the initial or final state[101; 102; 103], even relating it to medium-induced changes in the fragmentation function[104] and using effective theories to describe the effects[105]. Predictions of energy loss in p+A collisions at the LHC[86], such as in Figure 2.15 suggest there could be modifications of the jet yield at the 25% level at mid-rapidity and as much as 30-40% at very forward rapidities.

Depending on the Q^2 of the hard scattering, the fragmenting parton could be sensitive to the

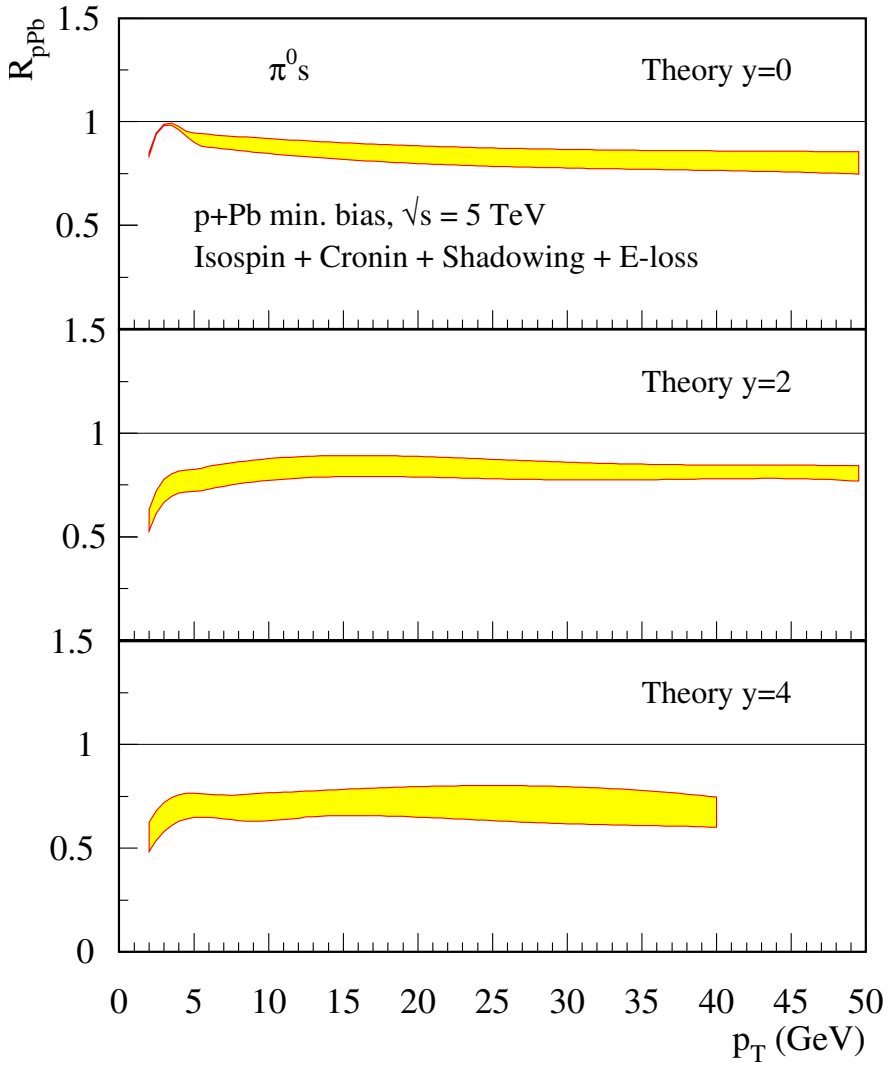


Figure 2.15: Predictions for the $R_{\rm pPb}$ at high-$p_{\rm T}$ from energy loss in a pQCD calculation, from [86].

nuclear medium in different ways. At high-Q^2, where the parton hadronizes outside the nuclear medium, it may be that only initial state energy loss effects are important. On the other hand, parton showers undergoing the hadronization process in the nuclear medium could be subject to a variety of interesting effects[106] for which the A dependence is not wel known[107]. Thus, information on jet production over a wide range of Q^2 is important.

At lower Q^2, as the distance scale of probe becomes larger and more of the nucleus is seen by the

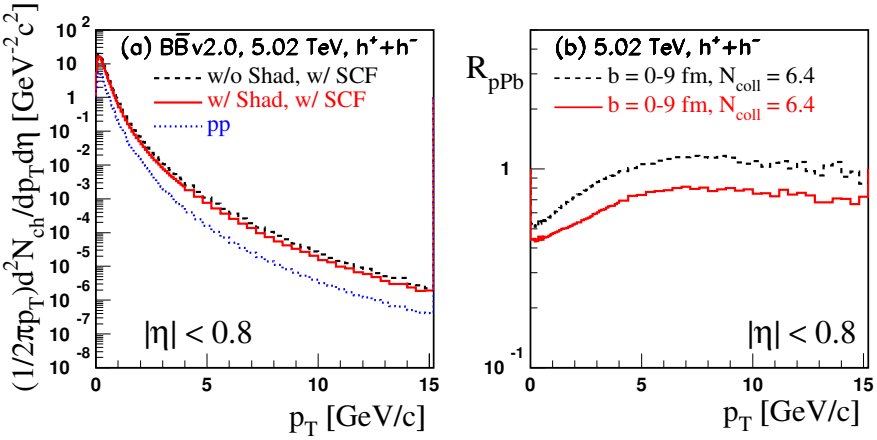


Figure 2.16: Predictions for the R_{pPb} at the LHC, variants of the HIJING Monte Carlo, from [108].

parton, the role of coherence effects between multiple soft scatterings should become important[109; 110]. Desctructive interference between competing terms results in a suppression of the cross-section at small values of nuclear x (also called *shadowing*), small constructive interference might be manifest at somewhat higher x (*anti-shadowing*). Additionally, this could result in the azimuthal decorrelation of back to back jets[111], as the $\Delta\phi \approx \pi$ balance of dijets becomes broadened by successive rescatterings on soft gluons.

Data at LHC energies is badly needed to distinguish between models. For example, Figure 2.16 shows a prediction for the R_{pPb} [108] using the HIJING MC[112] event generator with variations on pQCD-inspired models of shadowing.

2.2.3 Nuclear parton distribution functions

As described above, the partonic structure of the nucleus can be significantly affected by the high-density environment. This section describes the *nuclear parton distribution functions*, which parameterize the modifications observed in experiments. In the perturbative regime, the modification of parton i density in a nucleus of size A are quantified through *nuclear parton distribution functions* (nPDFs) f_i^A and their modification factors relative to free nucleons R_i^A,

$$R_i^A(x, Q^2) = f_i^A(x, Q^2)/A \cdot f_i^{proton}(x, Q^2) \tag{2.89}$$

Importantly, the nPDF's assume that all modifications in the jet (or photon, heavy quark, etc.) rate are universal – that is, they are only functions of x and Q^2 in nuclei of a given A. Unlike in simpler hadron-hadron collisions, the factorization theorems have not been rigorously proven in collisions involving nuclei and are assumed to hold. Any modification of the parton structure in nuclei are encoded in the nPDF's. (Indeed, if factorization were broken, the effects would be formally "parameterized" by the nPDFs.) The accurate determination of the nPDFs is absolutely critical for the interpretation of all measurements of QGP effects on hard probes in A+A collisions[113; 114]. In particular, understanding the b dependence of nPDF modification can allow a better evaluation of the real centrality dependence of high-p_{T} jet or heavy flavor suppression.

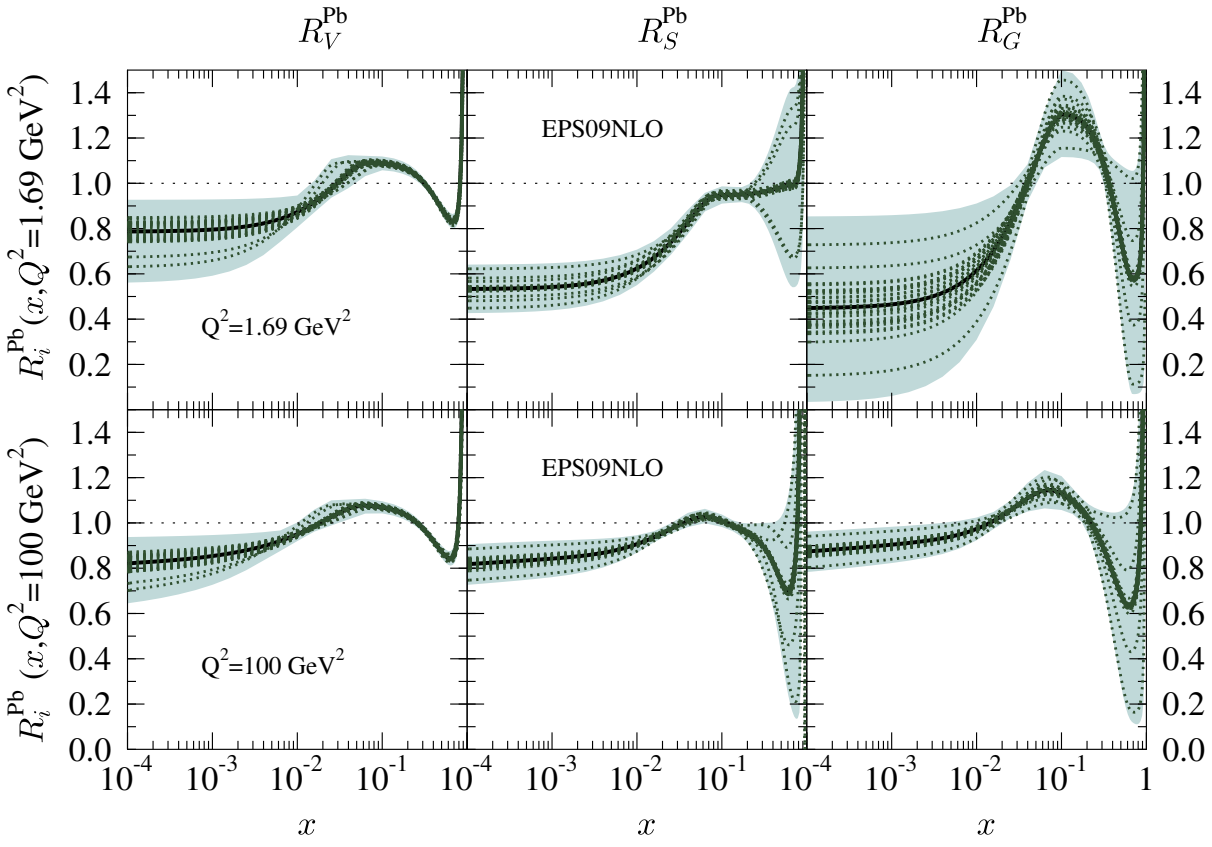


Figure 2.17: EPS09 results for the nuclear parton distribution functions in Pb, from [115].

Several nPDF sets exist, with some of the more popular ones including the HKN[116] and

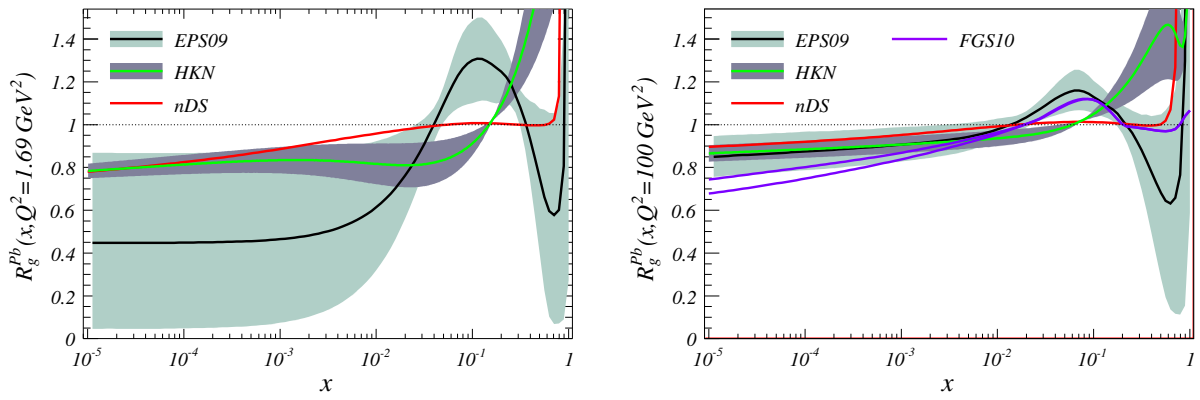


Figure 2.18: Gluon nPDF and uncertainty as described by different nPDF sets, for $Q^2 = 1.69$ GeV^2 (left) and $Q^2 = 100$ GeV^2 (right) in Pb nuclei, from [114].

EPS09[115] parameterizations. The nPDF sets describe the R_i^A with some functional form determined using DGLAP evolution and a global fit to a combination of DIS l+A, fixed-target p+A and RHIC d+Au data. The phase space of available data for nPDFs is significantly smaller than that for PDFs in nucleons, meaning that nPDF's are known much more poorly in some regions than their proton pdf equivalents. Figure 2.17 shows the next-to-leading order EPS09 nPDF modification factors for valence quarks, sea quarks and gluons, at two values of Q^2. The gluon nPDF at low Q^2 is particularly underconstrainted. Additionally, there are tensions between different nPDF's sets, particularly in the nuclear modification of the gluon distribution function, in Figure 2.18. Clearly, more information is needed.

The general pattern of nPDF modification vs. x are typically categorized into four distinct regions. Note that unlike the previous section on higher-twist pQCD effects on the production of hard probes in p+A collisions, the nPDF formalism is agnostic about the physical origin of the modification – it just seeks to describe it self-consistently within a DGLAP framework.

- *Shadowing region.* The dominant effect at low-x ($\lesssim 0.1$) is the suppression of the nuclear structure functions relative to their A-scaled free nucleon equivalents. This is possibly the most interesting of the four regions. Experimental data on shadowing first came from μ+A DIS experiments with the E665 detector at Fermilab[117] and the New Muon Collaboration (NMC) spectrometer at CERN[118]. The E665 results show that the strength of shadowing

generally increases with increasing nuclear size and with decreasing x. On the other hand, it appears to have a weak Q^2 dependence, decreasing slowly with $x < 0.1$ and no noticable dependence for $x > 0.1$. More recently, it has also been observed in d+Au collisions at RHIC[119; 120], although the precise extraction of the kinematics is complicated by the hadronic initial state. Shadowing is not yet comprehensively understood, but is generally thought to arise from effects related to the changes in partonic density in the nucleus[121] (but see the sections before and after this one).

- *Antishadowing region.* The NMC experiment has demonstrated a small but significant excess ($\approx 2\%$) in the bound to free nucleon structure function in the region $0.1 \lesssim < x < \lesssim 0.2$, with the most convincing demonstration of this effect coming from μ+Sn and μ+C[118] DIS data. This effect has been interpreted as the consequence of multiple scattering of anti-quarks[122] or also as the trivial consequence of the momentum sum rule, which requires that a suppression in the shadowing region be compensated by an enhancement somewhere else[123].

- *EMC effect region.* It was first observed in μ+Fe scattering by the European Muon Collaboration (hence the name of the "EMC" effect)[124] that at intermediate $0.2 \lesssim x \lesssim 0.8$ the nuclear PDFs are actually suppressed relative to the free nucleon PDFs. This effect has been verified down to very small nuclei[125] and is at most a weak function of Q^2. The origin of the EMC effect is not yet generally understood, and remains a topic of active research (including, for example, new ideas relating the strength of the effect to short range nucleon-nucleon correlations[126]).

- *Fermi motion region.* The dominant effect at high-x ($\gtrsim 0.8$) comes from the Fermi motion of the nucleons inside the nucleus[127]. The additional motion smears out the kinematics of the partons, and since the parton distribution functions fall rapidly at high-x, this results in an apparent enhancement in the bound to free nucleon ratio.

The d+Au system at RHIC was crucial for understanding the jet suppression observed in Au+Au collisions. One of the most studied results in this vein (shown in Figure 2.19) is the R_{dAu} for high-p_{T} π^0 and η mesons measured in PHENIX[97]. At the time, these data were interpreted as showing a very moderate p_{T} dependence and all points consistent with $R_{\mathrm{dAu}} = 1$ within statistical and systematic uncertainties. Thus, nPDF effects were judged to be small and the data is regularly

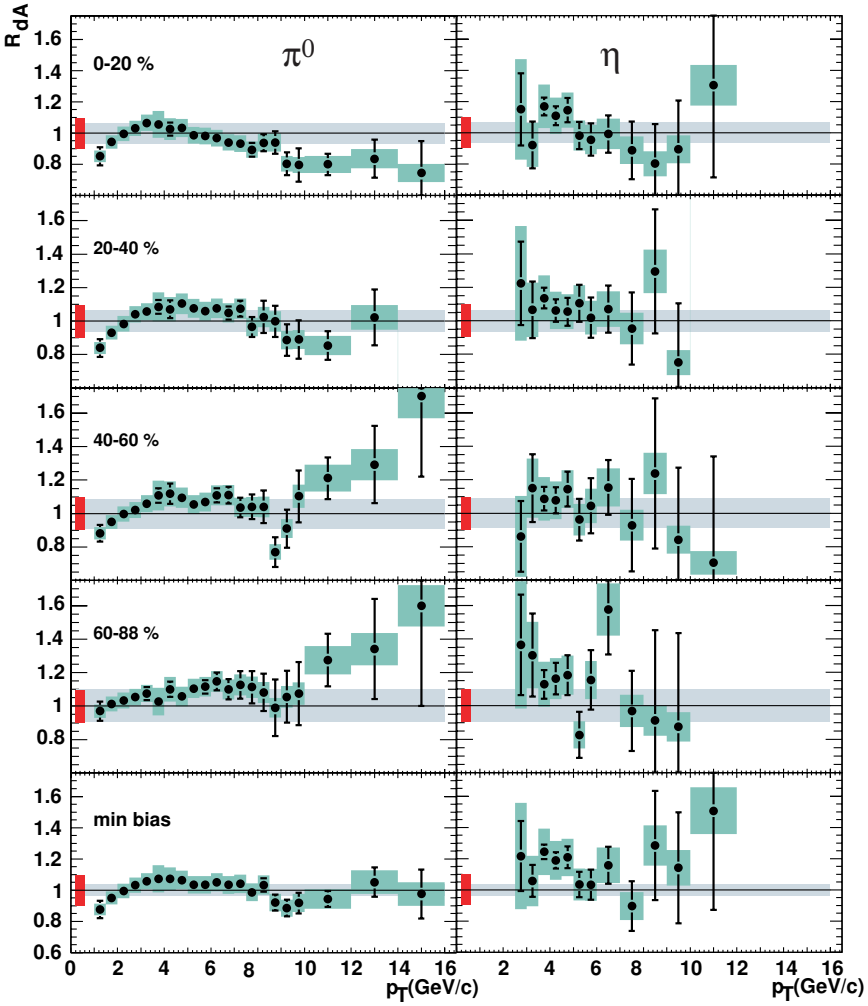


Figure 2.19: Centrality-dependent nuclear modification factor for neutral pions and eta mesons at mid-rapidity in d+Au collisions, from [97].

included in nPDF global fits. Recently, there has been progress in developing an impact-parameter dependent extension to the EPS09 sets[128]. The b dependence is mostly derived by considering the A dependence in data (since different nuclei will have a different $\bar{b}$, roughly according to $\bar{b} \approx \int_0^R b d^2b / \int_0^R d^2b \sim A^{1/3}$) but also considers the centrality-dependent PHENIX mid-rapidity π^0 data[97]. The b dependence is judged to be very small in the kinematic range shown. In fact, it can be seen that the two most peripheral bins trend to $R_{\rm dAu} > 1$ at high-$p_{\rm T}$ and $R_{\rm dAu} < 1$ in the most central collisions, but the statistical errors are too large to make a strong claim. By

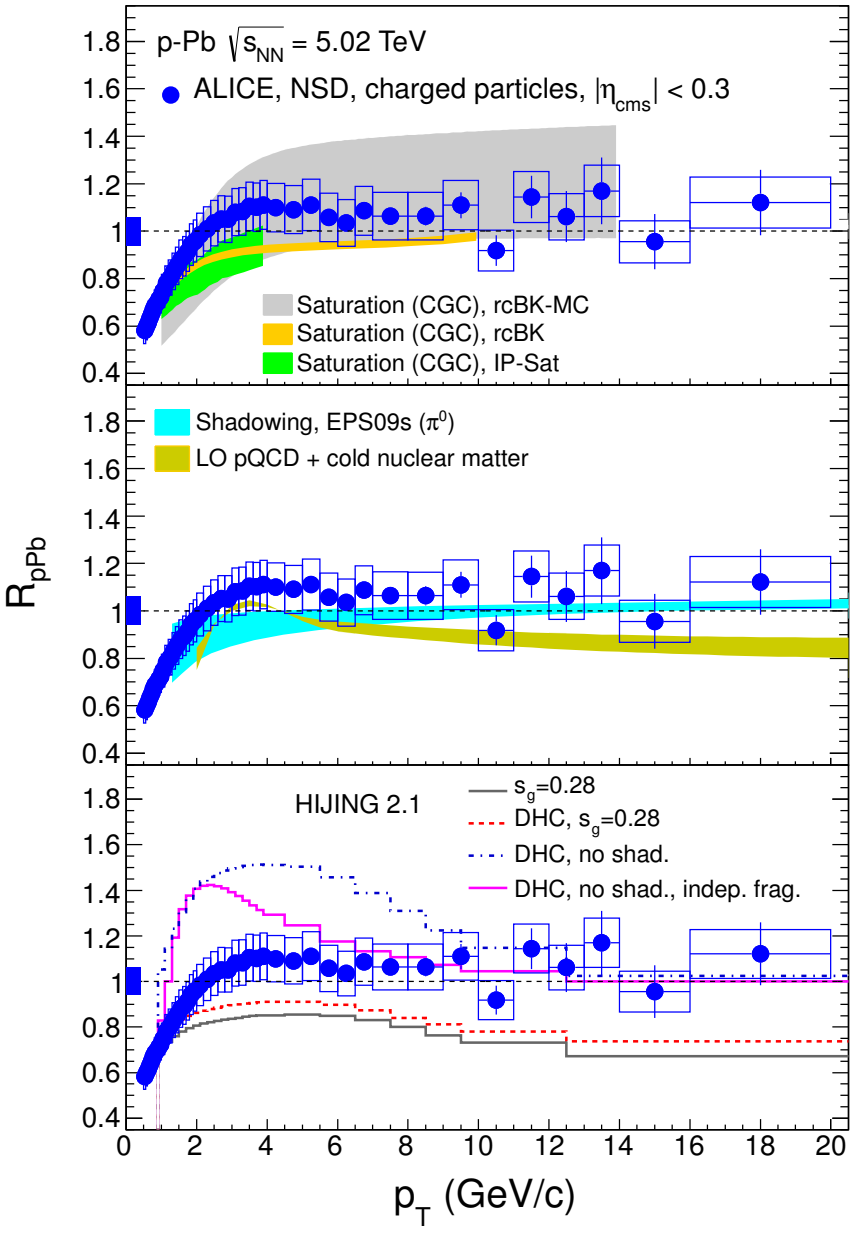


Figure 2.20: Nuclear modification factor for charged particles in minimum bias p+Pb collisions, from [98].

looking at reconstructed jets instead of single particles, the results presented in this work allow for a significantly extended measurement of the R_{dAu} in the different centrality bins.

There are also nPDF-inspired predictions for the R_{pPb} for pions at the LHC experiments[129]

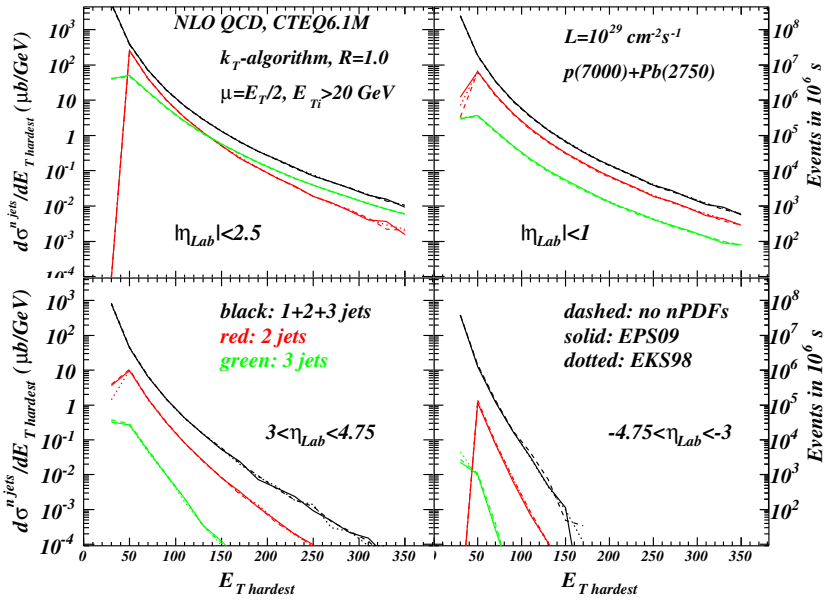


Figure 2.21: Example of prediction of the jet cross-section in p+Pb collisions at attainable LHC energies, from [98].

over a large p_{T} and rapidity range. In these calculations, overall effects are expected to be modest, at least at mid-rapidity. The nPDFs, along with NLO calculations of the hard processes, can be used to predict the overall jet cross-section in p+Pb collisions, shown for a putative upgraded LHC run in Figure 2.21.

The first R_{pPb} measurement at the LHC was performed by ALICE in a narrow pseudorapidity range ($|\eta_{\mathrm{LAB}}| < 0.3$) for minimum bias p+Pb collisions[98], shown in Figure 2.20. Above $p_{\mathrm{T}} > 3$ GeV, the minimum bias $R_{\mathrm{pPb}} \approx 1.1$, but is consistent with unity when the systematic uncertainties are taken into account. The results presented in this work will show the R_{pPb} within a much large η range, impact parameter dependence, and for fully reconstructed jets instead of inclusive hadrons.

Finally, a recent preliminary CMS measurement of the high-p_{T} (> 100 GeV) dijet systems in different p+Pb centrality selections[130] may probe nPDF modification. The η^* distribution of the dijet system, which is correlated with the nuclear x distribution, appears to narrow with centrality in a way that is qualitatively consistent with shadowing and the EMC effect (suppression at the edges of the η^* distribution) and antishadowing (enhancement in the center of the distribution).

However, a more detailed (or directly related to x) treatment is required to quantitatively extract the extent of nuclear effects.

2.2.4 Saturation

HERA results on the lepton-proton scattering[131] revealed a steep rise at fixed Q^2 in the proton structure function F_2 down to values as low as $x > 10^{-5}$. The derivative

$$-\left.\frac{\partial \log\left(F_2\right)}{\partial \log\left(x\right)}\right|_{Q^2} = \lambda\left(x, Q^2\right) \tag{2.90}$$

was observed to be independent of x for $x < 10^{-2}$ and $Q^2 \gtrsim 1.5$ GeV, with λ increasing with $\log\left(Q^2\right)$ but generally in the range $\approx$ 0.25-0.30. At these small values of x, the proton structure function is dominated by the gluonic contribution. Thus, an application of the BFKL[41; 42] evolution equations to determine the gluon density at ever lower x would imply that it rises without limit as a power law,

$$xg(x) \propto x^{-\lambda} \tag{2.91}$$

This increase in the gluon density has implications for the high-$\sqrt{s}$ behavior of the pp cross-section, resulting in a power-law growth at high-energy via

$$\sigma \propto s^{+\lambda} \tag{2.92}$$

This would unavoidably violate the Froissart bound[132], which limits the growth of hadronic cross-sections in QCD to no faster than $\sigma \propto \log^2\left(s\right)$. Thus, it is generally thought that other mechanisms must become important which slow the growth of the gluon density.

As the gluon density increases, the gluons begin to overlap in phase space. Whereas the BFKL evolution had previously been dominated by $1 \to 2$ $g \to gg$ gluon splitting, at sufficiently high density, the $2 \to 1$ $gg \to g$ gluon recombination, or gluon fusion, processes become important[133]. The resulting evolution in $\log(1/x)$ therefore becomes non-linear. In the large number of color (N_C) limit, it is described by the Balitsky-Kochegov (BK)[134; 135; 136] evolution equations. Schematically, the BK evolution is the solution to equations of the form

$$\frac{\partial \mathcal{N}}{\partial \log(1/x)} = \omega \left(\mathcal{N} - \mathcal{N}^2 \right) \tag{2.93}$$

where $\mathcal{N}$ is the gluon number density, $\omega\mathcal{N}$ is the linear (BFKL) term from gluon splitting and $\omega\mathcal{N}^2$ is the non-linear term from gluon recombination. It is easily observed that Equation 2.93 has fixed points at $\mathcal{N} = 0$ and, due to the non-linear term, $\mathcal{N} = 1$ (in the high-$\log(1/x)$, or low-x, limit). The generalization of the BK equations to a more realistic number of colors are known as the Jalilian-Marian-Iancu-McLerran-Weigert-Leonidov-Kovner (JIMWLK)[137; 138; 139; 140] equations, which have the structure of renormalization group equations in the variable $Y = \log(1/x)$. Furthermore, while the original BK equations were a leading order (LO) calculation with a fixed coupling constant α_s, the running coupling BK (rcBK)[141; 142] equations extend the calculation to NLO accuracy. Figure 2.22 summarizes the regimes in which the DGLAP, BFKL and BK/JIMWLK evolution equations are appropriate.

For a given value of x, the momentum scale at which non-linear QCD effects are expected to affect the available amount of gluons is called the *saturation scale*, Q_s[144] which grows with decreasing x,

$$Q_s^2 \propto (1/x)^{+\lambda} \ [\mathrm{GeV}^2] \tag{2.94}$$

For sufficiently high $\sqrt{s}$ (e.g. low-x), the saturation scale may even be in the non-perturbative regime $Q_s \gg \Lambda_{QCD}$. Now consider how the gluonic structure in affected inside colliding nuclei. By fairly general arguments, the gluon wavelength is longer than the longitudinal size of the Lorentz-contracted nucleus, causing all the gluons in the transverse plane to overlap and increasing the density by a factor $A^{1/3}$. With the approximation of $\lambda = 1/3$, the saturation scale in heavy nuclei can be written as

$$Q_s^2 \propto (A/x)^{1/3} \ [\mathrm{GeV}^2] \tag{2.95}$$

Equation 2.95 reveals the advantage of probing saturation phenomena with heavy nuclei. To increase the saturation scale by as much as the presence of a gold (Au) or lead (Pb) nucleus, the center of mass energy would have to increase by a factor of $A \approx 200$! Furthermore, it is possible that saturation effects could be probed differentially as a function of impact parameter, since the saturation scale could be related to local nucleon density, via

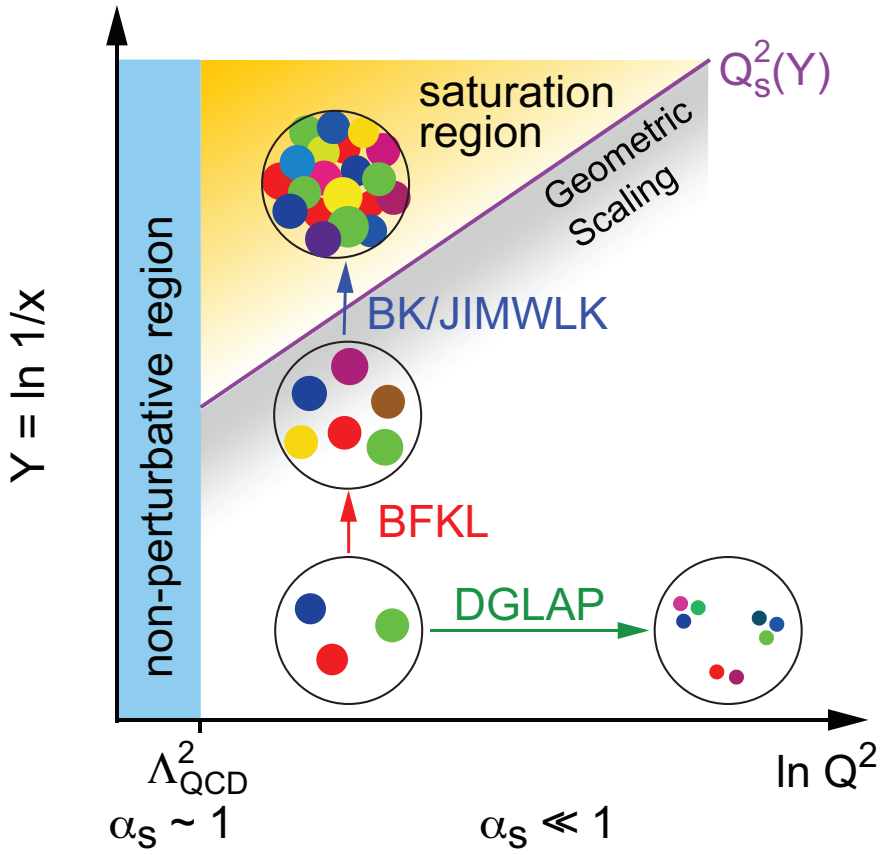


Figure 2.22: Diagram of (x, Q^2) evolution of PDFs, from [143].

$$Q_s^2 \propto (N_{\text{coll}})(1/x)^{+\lambda} \text{ [GeV}^2\text{]} \tag{2.96}$$

The *color glass condensate* (CGC) framework is a unified treatment of the non-linear gluon evolution as an effective classical field theory based on the McLerran-Venugopalan (MV) model[145]. In the CGC formalism[146; 147], the dynamics of the large-x and low-x ("wee") partons are separated. On the time scale of the wee partons, the large-x partons are essentially "frozen" by Lorentz dilation and are treated as static color charges. The wee partons behave as classical Yang-Mills gauge fields sourced by the (random) static color charges. In large nuclei and at sufficiently high energy where the color charge per unit area is $\gg \Lambda_{QCD}$, the effective theory is a weakly coupled one. *Color* refers to the SU(3) of the theory, *Glass* refers to the "frozen disorder" of the static color sources and *Condensate* to the high occupation numbers of the small-x gluons. The CGC

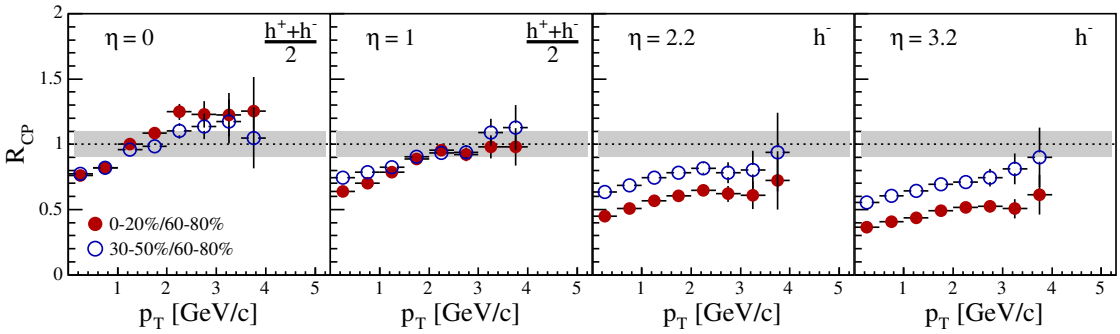


Figure 2.23: BRAHMS experiment measurement of the central-to-peripheral nuclear modification factor of charged hadrons in d+Au collisions, from [119].

framework aims for a comprehensive description of the initial state of low-x partons in p+A and A+A collisions. Many saturation calculations take place within this framework.

There are several proposed consequences of low-x saturation which could be measured experimentally in d+Au collisions at RHIC and p+Pb collisions at the LHC. The most straight-forward one is the decreased amount of gluons at a given (low) x in nuclei, which would result in a decreased rate of low-x gluons jets relative to that in protons. As shown in Section 2.1.5, the nuclear x is generally related to the transverse momentum and rapidity of the hard-scattered partons, via $x \sim p_{\mathrm{T}} e^{-y}$. Furthermore, in such a saturated regime, gluons cannot be regarded as independent for the purposes of collinear factorization, and collective interference effects may become important. Four historical putative signals of saturation, the suppression of forward hadrons and the appearance of forward monojets (at RHIC), and the suppression of the total multiplicity and the appearance of the *ridge* (at the LHC) are discussed here.

Fits to DIS data [148] predict the saturation scale in d+Au collisions at RHIC to be $Q_s^2 \approx 1.5$-2 GeV^2. Unfortunately, this scale is right at the interchange of hard and soft processes, which are difficult to separate. Thus, saturation effects are better probed with measurements in the very forward (d-going) direction. The first such measurement is by the BRAHMS experiment[119], shown in Figure 2.23, which observed an η-dependent suppression of charged hadrons qualitatively consistent with the behavior of Q_s as inferred by measurements in HERA. In fact, such a suppression had been predicted at large rapidities within the CGC[149; 150; 151] for some time. It was even argued that the shadowing from higher-twist effects could not explain the full magnitude of the

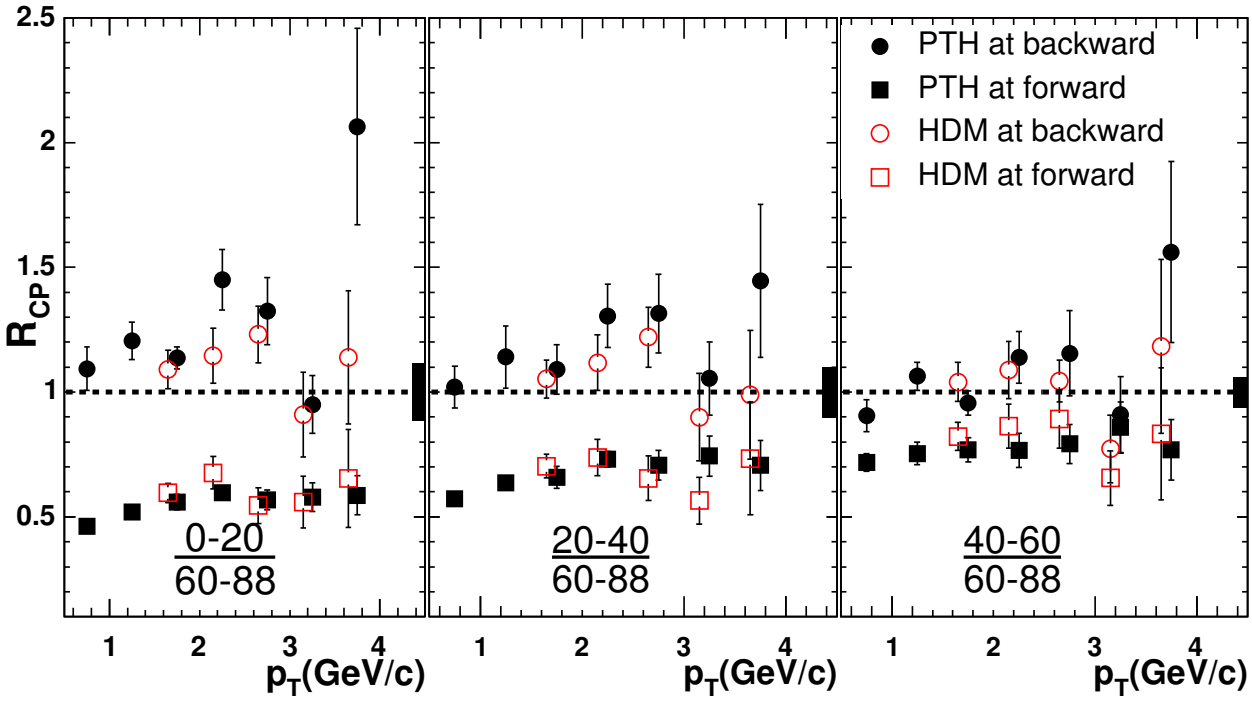


Figure 2.24: PHENIX experiment measurement of the central-to-peripheral nuclear modification factor of charged hadrons in d+Au collisions, from [120].

suppression[152]. The PHENIX experiment confirmed this suppression[120], shown in Figure 2.24, but also showed an surprising small enhancement at backward rapidity, qualitatively consistent with antishadowing in the large nuclear-x region. However, more detailed treatments revealted that the R_{dAu} could potentially be sensitive to other effects not necessarily related to the change in partonic structure. In fact, the data could also be sensibly explained as the result of energy loss[153], nuclear enhanced power corrections to the structure functions (shadowing)[154] and even as the breakdown of QCD factorization for high-x partons (in this case, the high x in the deuteron, since the nuclear x is very small for very forward production)[155].

In addition to the suppression of single particles in the very forward direction, di-hadron correlations are thought to be a potentially more sensitive probe of saturation effects, since measuring the p_{T} and y of both partons (or, in this case, high-p_{T} fragments from the original parton) better reconstructs the original nuclear x (see Equation 2.57). PHENIX has measured the back to back correlation of hadrons at mid-mid, mid-forward and forward-forward rapitidies[156]. As seen

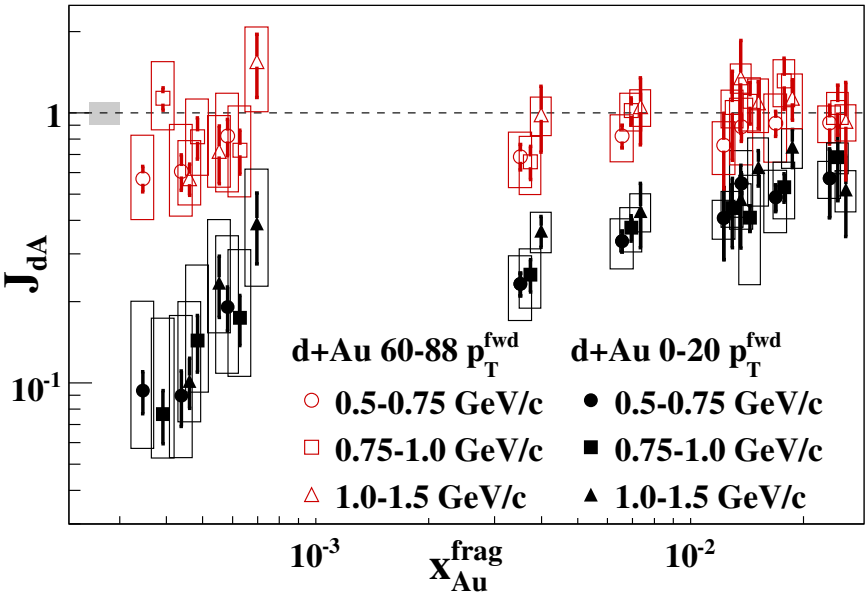


Figure 2.25: Nuclear modification factor if pair yields in d+Au collisions for two centrality selections, from [156].

in Figure 2.25, the results show a suppression in central d+Au collisions relative to the pp baseline that is systematic with decreasing x. On the contrary, the peripheral collisions show a much weaker effect. Similarly, the STAR experiment has shown that in central d+Au collisions, the amplitude of the away-side ($\Delta\phi \approx \pi$) yield in di-hadron pairs is suppressed and broadened[158; 157] relative to the same signal in peripheral d+Au collisions and pp collisions (see Figure 2.26). The disappearance of the away-side correlation is sometimes referred to as the *monojet* phenomenon. These effects are able to be reproduced within CGC calculations[159], but other arguments stressed the role of multiple parton interactions and energy loss[160] in influencing azimuthal correlations at large rapidities.

Additionally, the total particle multiplicity (extrapolated to $p_{\mathrm{T}} = 0$) could be very sensitive to saturation effects, since the number of gluons available for particle production would be modified. The ALICE Collaboration has recently measured the pseudorapidity density $(1/N_{evt})(dN/d\eta)$ for charged particles in $|\eta_{\mathrm{LAB}}| < 2$, for a sample of p+Pb events corresponding to 96% of the non-single diffractive p+Pb cross-section and limited contamination from other events[161], shown in Figure 2.27. The data are compared to several models of particle production. In particular, the

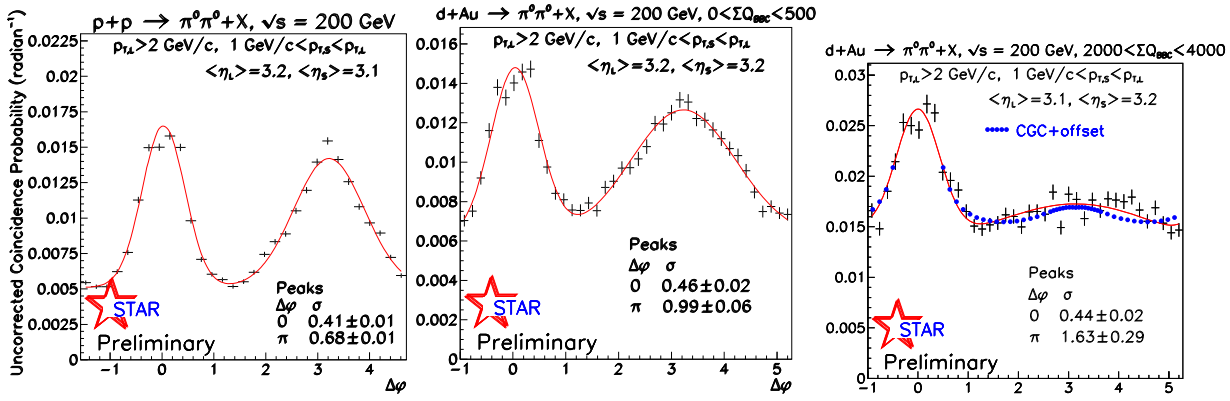


Figure 2.26: Coincidence probability between two π^0's with $\langle\eta\rangle \approx$ 3.1-3.2 in pp (left), peripheral d+Au (middle) and central d+Au (right) collisions, from [157].

strength of the forward-backward asymmetry is examined, since saturation effects would generally have a y dependence. Currently, saturation calculations seem to overpredict the slope of the $dN/d\eta$ distribution.

Another novel feature of p+A collisions with possible relevance to saturation is the observation of long-range, rapidity separated correlations in the production of soft particles in very high multiplicity p+A collisions, referred to as the *ridge*. The history of the ridge in A+A, high-multiplity pp and then finally p+A collisions is somewhat complicated, but the p+A results admit several interpretations. All three large LHC experiments observed a ridge in p+Pb collisions[163; 164; 162], with the ATLAS analysis showing that the excess near-side yield contributes equally to the away-side in a way that is monotonic with event activity (as measured by the sum of the energy in the Pb-going forward calorimeter), shown in Figure 2.28. In the CGC framework the nearside and awayside (rapidity separated) peaks come from a combination of "back to back" QCD graphs contributing to the away-side and "glasma graphs", strongly enhanced by high powers of α_s^{-1} in the saturation regime, which contribute to both the near- and away-sides[165]. On the other hand, the ATLAS analysis suggests that the ridge could have an origin as a collective final-state, instead of initial-state, effect.

While the RHIC data on forward single hadrons and monojets, and the early LHC ridge results, could be said to be tentatively consistent with saturation effects in the Au or Pb nuclei, they can also be individually explained by other models. Thus, more definitive tests are needed. For

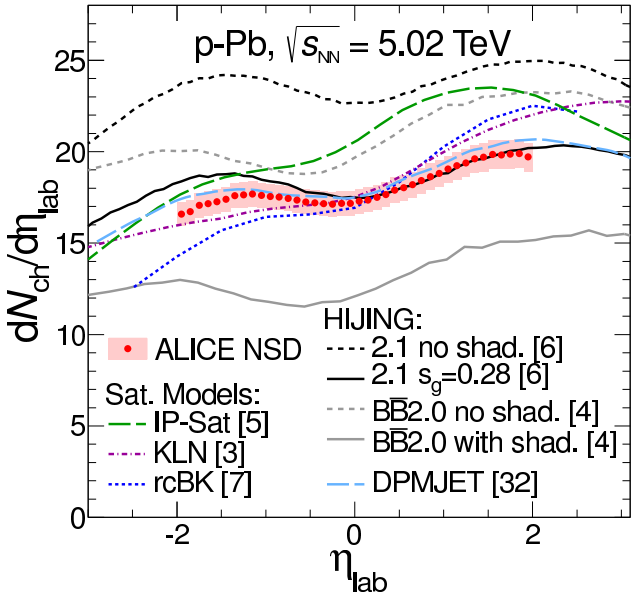


Figure 2.27: Measurement of charged particle multiplicity in minimum bias p+Pb collisions, from [161].

example, there are saturation models based on a high-energy factorization[166] treatment of p+Pb scattering with a non-linear gluon density at low-x that predict a substantial modification in the azimuthal correlation of low-p_{T} central-forward dijets at LHC energies[167]. Furthermore, if the forward hadron suppression at RHIC energies is a saturation effect, a naive application of the $Y = \log(1/x)$ dependence of the saturation scale would imply that the effects seen at $y \approx 3.5$ at RHIC energies would be visible at $y \approx 0$ at the LHC! Indeed, saturation calculations within the CGC framework[168] predict very strong effects in the R_{pPb} for single particles out to high-p_{T} and over a wide range of rapidity, in stark contrast to collinear factorization approaches[169] which predict more modest effects from the nuclear parton distribution functions. Figure 2.29 illustrates that (for the models and $\sqrt{s_{\mathrm{NN}}}$ pictured there) the difference between pQCD and saturation calculations can have a factor of 2 difference in the R_{pPb} even at mid-rapidity.

Thus, experimental control over jets for a wide range of forward rapidities and at as low a p_{T} as can be experimentally achieved are crucial for distinguishing between models. This work is the first measurement of forward jet production at LHC energies.

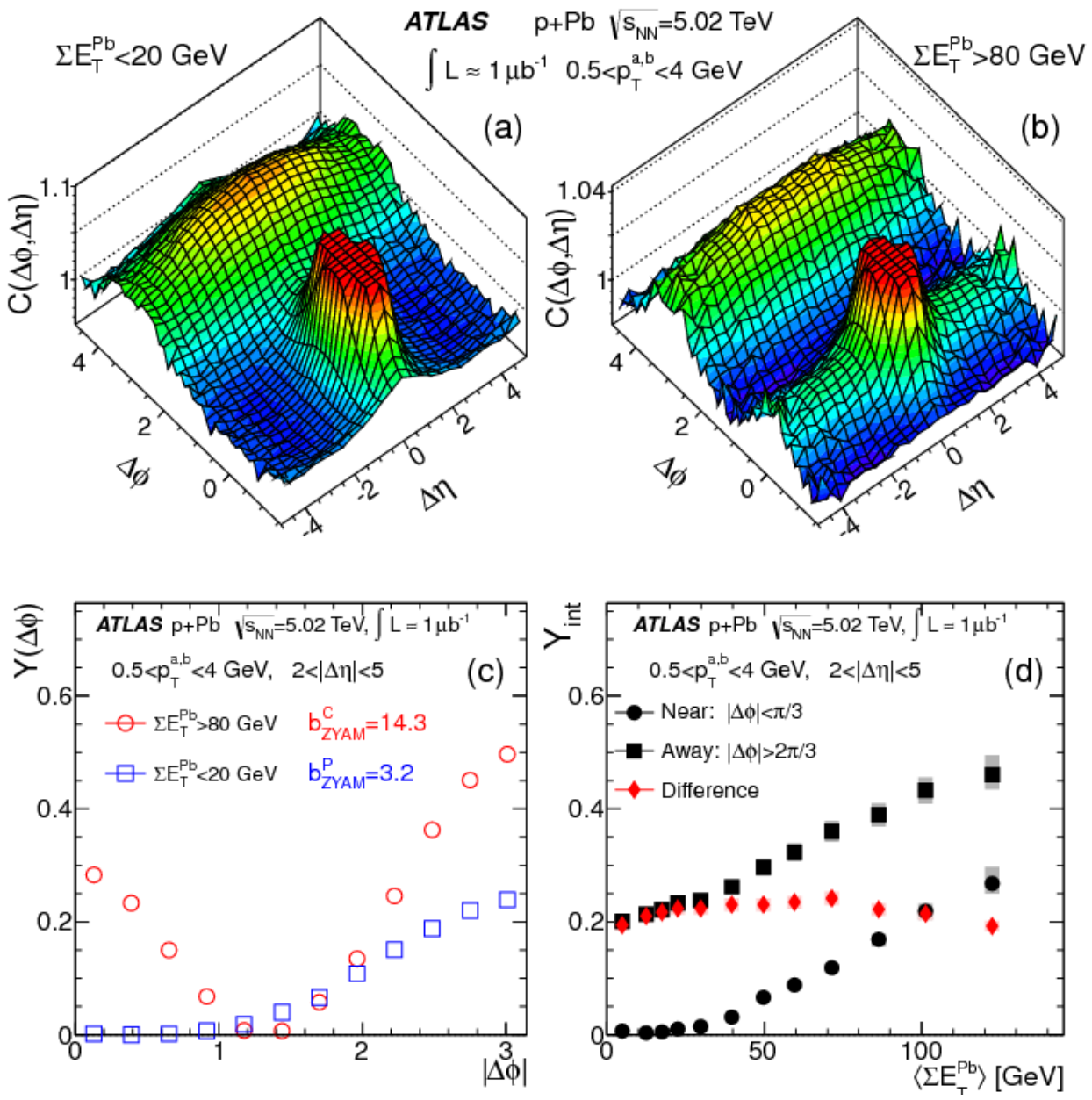


Figure 2.28: a) Two-particle $\Delta\eta, \Delta\phi$ correlation functions in low-activity p+Pb events, and b) the same in high-activity p+Pb events. c) Rapidity-separated yields vs. $\Delta\phi$ in low- and high-activity events. d) Integrated near- and away-side yields vs. event activity, from [162].

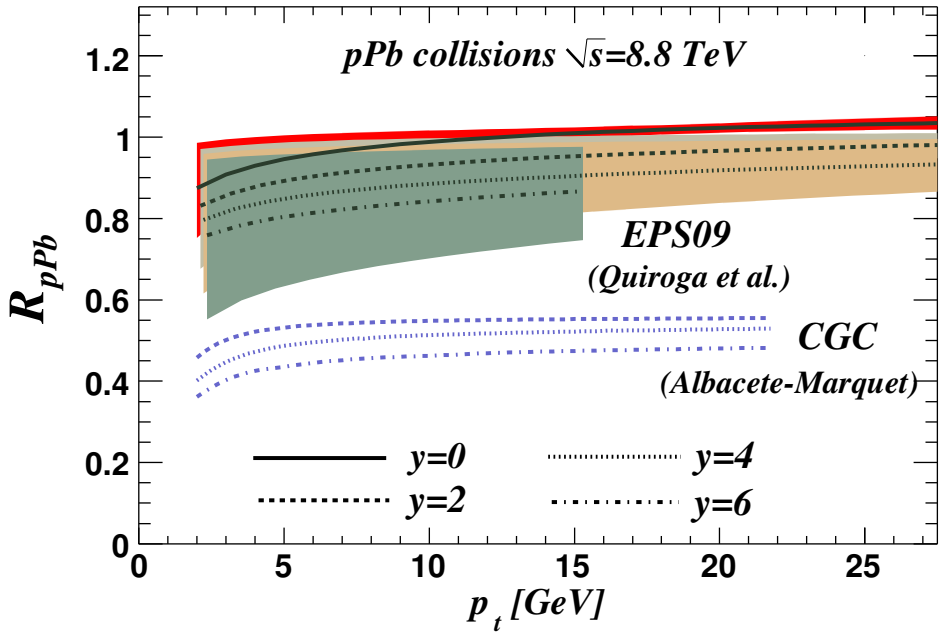


Figure 2.29: Comparison of predictions for the R_{pPb} at LHC energies between conventional pQCD and saturation calculations, from [114].

Chapter 3

Hadron and Ion Colliders

But, as we often see, against some storm,
A silence in the heavens, the rack stand still,
The bold winds speechless and the orb below
As hush as death, anon the dreadful thunder
Doth rend the region

Hamlet, Act II, Scene II

3.1 Principles of a hadron synchotron

In the present era, the study of high-energy physics relies crucially on the capability and performance of particle accelerators. High energy collisions probe distance scales far smaller than the size of the typical hadron and have enough center of mass energy to create heavy, unstable particles. The electroweak bosons ($W^{\pm}$, Z^0 and now the Higgs), the heavy quarks and leptons were discovered, and QCD and the Standard Model in general were ratified through the analysis of high-energy collisions of elementary particles. For practical reasons, the highest energy machines that exist today are hadron synchotron colliders. Complementary to these, lepton-lepton colliders (such as the active BEPC II or the decomissioned LEP) and lepton-hadron colliders (such as HERA) are critical to having a full experimental understanding of the field. However, these have somewhat different design concerns and considerations. Though many aspects of their operation are similar, we will focus on the details of hadron colliders here.

One of the major motivations behind colliding beams of particles with the same energy instead of one beam onto a stationary target is the efficient increase of the center of mass energy with the increasing energy of the beams. In a fixed target experiment where the beam has energy E (for simplicity, we consider situations were both particles in the collision have the same mass m), the kinematics can be described by

$$p_1^\mu = (E, 0, 0, p)\,,\ p_2^\mu = (m, 0, 0, 0) \tag{3.1}$$

$$s = (p_1 + p_2)^2 = (E + m)^2 - p^2 = 2m(m + E) \tag{3.2}$$

$$\Rightarrow \sqrt{s} = \sqrt{2}\sqrt{m^2 + Em} \tag{3.3}$$

whereas for two beams colliding head on with energy E, the center of mass energy is

$$p_1^\mu = (E, 0, 0, p)\,,\ p_2^\mu = (E, 0, 0, -p) \tag{3.4}$$

$$\Rightarrow \sqrt{s} = \sqrt{4E^2} = 2E \tag{3.5}$$

Thus, a $\sqrt{s} = 10$ TeV $= 10^4$ GeV pp collision would require an achievable 5 TeV beams in a collider, and an exceedingly large 50×10^3 TeV in a fixed target experiment of protons on a hydrogen target.

Another major reason is the fact that any detector arrangement in a fixed target experiment has an acceptance in the center of mass frame which depends on the incident energy and the target mass. This can lead to potentially burdensome systematics when comparing measurements with different targets or at different beam energies. In a collider experiment the laboratory and center of mass frames coincide, so the acceptance is exactly the same, regardless of the beam energies. (Unfortunately, what colliders gain in center of mass energy over fixed target linear accelerators, they trade in luminosity, as discussed below.)

The early precursor to modern synchrotrons is the *cyclotron*, first developed in 1934 by Ernest Lawrence[170]. The basic cyclotron design is two "D" shaped electrodes held at opposite oscillating potential (geometrically, this can be thought of a short cylinder with a strip along a diameter removed) that lie in a constant magnetic field perpendicular to the electrode plane. Charged particles injected near the center of the apparatus with be accelerated by the potential in the

plane of the cyclotron while also curving along a circular trajectory due to the magnetic field. The frequency of the electrode potential oscillation (called the radiofrequency or RF) must be synchronous with the revolution frequency which is given classically by

$$f = qB/2\pi m \tag{3.6}$$

so that as the particle passes between the electrodes it always sees an accelerating (not deccelerrating) field. In this way, the energy of the particle is successively increased until they reach the largest radius supported by the cyclotron, at which point the energetic particles may be directed as a beam outside the device.

In practice, this simple design cannot work for very high energy particles, for which the oscillation frequency becomes velocity-dependent,

$$f = qB/2\pi\gamma m \tag{3.7}$$

and can thus become out of sync with the oscillating potential. There are improvements to this design which change the frequency of the electric potential to match the changing revolution frequency (called a *synchocyclotron*) or which keep the frequency constant but change the B field in a radius-dependent way such that the revolution frequency is also constant (called an *isochronous cyclotron*).

The modern implementation is the *synchrotron*, in which the radius of the trajectory is fixed and the magnetic field is increased in step with the increase in momentum. In a synchrotron, the beam is controlled by a number of specialized sections, including radiofrequency cavities which accelerate the beam as it passes through them and also keep it together longitudinally, and magnet systems which keep the beam along the required trajectory and contain the slow transverse expansion of the beam. Thus, unlike the early cyclotrons where the magnetic field permeates the entire apparatus, a synchotron uses a very strong magnetic field at select positions.

The beam itself is contained in a small pipe held in very high vacuum and consists of radiofrequency *bunches*, small localized groups of very many (typically 10^8-10^{11}) hadrons. The bunches are enforced by the oscillating radiofrequency field. The center of the bunch sits in synch with the falling edge of an RF oscillation. Charged particles closer to the crest feel more of a force, pushing them to the center, and particles closer to the minimum feel less of a force, also pushing them to the

center of the bunch. Thus, the bunches and their spacing are naturally guided by the RF period. In linear accelerators, charged particles typically go through a series of RF cavities to accelerate, while in a synchotron the particles go through the same RF cavities with every revolution.

The dipole magnets are simple electromagnets with a single North/South pole pair that creates a uniform transverse field between them. As the beam passes through the magnet, it is bent in the other transverse direction according to the Lorentz force law. The dipole magnets in effect define the beam momentum, since particles with a slightly different momentum will not follow the reference trajectory and be lost. The strength of the curving can be characterized by the *magnetic rigidity* $B\rho = p/q$, where ρ is the gyroradius and p and q are the momentum and charge, respectively. Beginning with the Tevatron, the high beam energy has required magnetic fields so strong that they cannot be produced by conventional metals which saturate at $\approx 1 - 2$ Tesla. Thus, superconducting magnets and the cryogenic systems required to keep the cool are necessary in modern colliders.

The transverse sizes of the beams are kept from growing through the use of focusing *quadrupole* magnets. Such a magnetic configuration might have a field proportional to the displacement from the nominal center, $\vec{B} \propto (y, x, 0)$, which results in a force $\vec{F} \propto (-x, +y)$ that is restoring in the x direction but anti-restoring in y (particles exactly in the nominal center feel no force). By analogy with optics, they function as an optical lens which is convex along one transverse axis and concave along the other. When the beams are heavy ions, the intrabeam scattering increases by a factor Z^4/A^2 and more attention must be paid to the focusing capability.

Generally, a sequence of quadrupoles with alternating restoring planes can be built in such a way that the net effect is a focusing of the tranverse beam size one in both planes. This sequence is interspersed with dipole magnets (and possible other, more specialized sextupole or octopole magnets which provide finer corrections described below) and is called the *FODO* (*focusing-orbit-defocusing-orbit*) pattern or the accelerator's alternating gradient *lattice*. The basic unit of magnets in sequence which is repeated throughout the collider is called a FODO cell. The net focusing effect was revolutionary when first introduced in 1952[171] and is now the principle on which all modern synchotrons are built[172].

The beams are generally kept in separate but adjacent chambers and share the same vacuum chamber only in the interaction regions. To minimize the amount of beam-beam interactions that

can occur from two beams sharing the same chamber for a long stretch, the beams are often collided at a very small crossing angle. A counterexample to this is the Tevatron, in which protons and antiprotons circulate in the same chamber but are separated when not in the interaction region through additional electrostatic elements. After the bunches are injected, they are accelerated to the desired energy, in a procedure called a *ramping*, then brought into collision (called *cogging*).

For optimal operation, it is important to understand the transverse dynamics of the beam bunches. The oscillation of the bunch particles in the transverse plane is characterized by the betatron amplitude $\beta(s)$, and the beam emittance ϵ is defined as the region of $d^3\vec{x}d^3\vec{p}$ phase space which contains some nominal fraction (for example, 95% or a 1σ contour) of the bunch particles.

Formally, the equation of motion for a beam particle relative to the beam bunch center in one of the transverse dimensions is given by the differential equation,

$$x'' + K(s)x = 0 \tag{3.8}$$

where s is a longitudinal coordinate. K is periodic in s and is given by the details of the accelerator lattice. Equation 3.8 is called Hill's equation[173] and is known to be solvable using Fourier methods. The homogenous solution is

$$x(s) = \sqrt{\epsilon\beta_x(s)}\cos\left(\mu(s) + \mu_0\right) \tag{3.9}$$

which corresponds to betatron oscillation around the nominal center with a longitudinal position-dependent amplitude, and phase given by

$$\mu(s) = \int_0^s d\xi/\beta(\xi) \tag{3.10}$$

Strictly speaking, Equation 3.8 only holds for particles with no longitudinal displacement from the bunch. Any difference in the longitudinal momentum Δp from the nominal p results in additional dispersion terms in the differential equation of the form $D(s)(\Delta p/p)$, where $D(s)$ is set by the lattice, and thus results in a slightly different trajectory in phase space. This effect is called *chromaticity*. To mitigate the resulting chromatic abberations, synchotrons use the sextupole and octopole magnets described above to keep the longitudinal expanse of the bunches small.

Another crucial quantity is the tune Q which is the number of total oscillations of the phase incurred during a single synchrotron revolution,

$$Q = (1/2\pi) \oint d\xi/\beta(\xi) \tag{3.11}$$

A particle having precisely the same phase each time it sees a given focusing or steering magnet will exhibit resonant behavior. Thus it is necessary to avoid integral values of Q.

In the longitudinal center of the bunch, the transverse extent of the beam at position s can therefore be given by

$$\sigma_x(s) = \sqrt{\epsilon\beta_x(s)} \tag{3.12}$$

Because the instaneous luminosity from two crossed beams (see below) depends inversely on the transverse size $\mathcal{L} \propto 1/\sigma_x\sigma_y$, it is important to minimize the betatron function at the IP, denoted by $\beta^* = \beta(s = s_{IP})$. Thus, near any interaction region that requires a high instantaneous luminosity, a specialized set of magnets *squeeze* the beam to decrease the β^* by a factor of ten or more.

3.1.1 Luminosity

One of the primary measures of collider performance is the *luminosity* L. It has the units of inverse area and is quoted in a range of units from μb^{-1} (small luminosity) to fb^{-1} (large luminosity), where the conversion to natural units is 1 barn = 2568 GeV^2. Luminosity can be thought of as a measure of the amount of data recorded, corresponding to a number of events N_{evt} for a process with cross section σ via

$$N_{evt} = L\sigma \tag{3.13}$$

For example, when searching for rare events (such as the production of the Higgs boson or various conjectured beyond the Standard Model particles) with a known cross-section, the luminosity conveys information about how many events are expected to have occured. The instantaneous luminosity $\mathcal{L} = dL/dt$ is measured in $\text{cm}^{-2}\text{s}^{-1}$ and relates the cross-section of a given process to the rate (in events/second),

$$R_{evt} = \mathcal{L}\sigma \tag{3.14}$$

The instantaneous rate can also be rewritten in terms of the revolution frequency f_r of the machine and the mean number of interactions per crossing μ, $R = f_r\mu$. The instantaneous luminosity in a collider with revolution frequency f_r, and n_b bunches with intensities n_1 and n_2 in beams 1 and 2 is given by

$$\mathcal{L} = \frac{f_r \Sigma n_1 n_2}{2\pi \Sigma_x \Sigma_y} \tag{3.15}$$

where $\Sigma^2 \approx \sigma^2_{beam-1} + \sigma^2_{beam-2}$ are the convolved beam overlap profiles. In the case where all bunches have the same intensity, $\Sigma n_1 n_2 = n_b n_1 n_2$. Equation 3.15 can also be written in terms of the normalized emittance $\epsilon_N = \beta\gamma\epsilon$ but the exact form depends on the choice of what fraction of the phase space ϵ encloses (see above).

To monitor the instantaneous luminosity and determine the overall integrated luminosity for a set of data, a luminosity calibration is needed. This is done by determining the *visible cross-section* σ_{vis} for some detector (or coincidence of detectors) and then inverting Equation 3.14 in a way that instantaneous luminosity depends on the visible mean number of interactions per crossing μ^{vis}:

$$\mathcal{L} = R\frac{1}{\sigma_{vis}} = \mu^{vis}\frac{f_r}{\sigma_{vis}} \tag{3.16}$$

When the mean number of interactions is low μ can be approximated by the probability P of firing the luminosity counter, which is simply the ratio of the number of observed counts divided by the number of bunch crossings $P = n/n_{BC}$. When μ is high, a more sophisticated relationship must be used which relates the (Poisson) probability distribution of at least one event occuring to the observed rate.

In principle, σ_{vis} can be determined for different sets of luminosity detectors, each of which report a different luminosity rate and can give a cross-check on $\mathcal{L}$. Thus, the problem reduces to determining the visible cross-section for the luminosity detector(s) in question. This is done via the method of Simon van der Meer[174]. In a *van der Meer scan* (also called *Vernier scan*), the transverse separation of the beams is varied in one plane while held constant in the other, and μ^{vis} is recorded as a function of the separation (δ_X or δ_Y, depending on which is the scan plane). It can be shown that the beam overlap profiles are determined by

$$\Sigma_{X,Y} = \frac{1}{\sqrt{2\pi}} \frac{\int \mu^{vis}(\delta_{X,Y}) d\delta_{X,Y}}{\mu^{vis}(0)} \tag{3.17}$$

$\mu^{vis}(\delta_X)$ is called the luminosity curve and is proportional to the overlap of beam densities $\int d^3x \rho_1(x-\delta_X/2, y, z)\rho_2(x+\delta_X/2, y, z)$. For $\mu^{vis}(\delta_X)$ described by a gaussian function, Σ is simply the width of the distribution (although in general the transverse density of bunches produced by real colliders are not precisely gaussian). Furthermore, it should be remembered that for two beams each with gaussian width σ, $\Sigma = \sqrt{2}\sigma$.) The luminosity curve for a bunch during horizontal and vertical separation scans in ATLAS is shown in Figure 3.1. The data in the figure are fit to a double gaussian profile,

$$\mu_{vis}(\delta) = A \left(\frac{f e^{-(\delta-\delta_0)^2/2\sigma_1^2}}{\sigma_1} + \frac{(1-f) e^{-(\delta-\delta_0)^2/2\sigma_2^2}}{\sigma_2} \right) \tag{3.18}$$

where $1/\Sigma = (f/\sigma_1 + (1-f)/\sigma_2)$. For beam profiles that are similar to, but not quite, gaussian, the double gaussian function is a popular choice. Recent vdM analyses at ATLAS[175] have also investigated a "gaussian $\times$ polynomial" function which models the deviations from a purely gaussian envelope with a fourth-order polynomial (excluding the linear term):

$$\mu_{vis}(\delta) = A \left(1 + c_2(\delta-\delta_0)^2 + c_3(\delta-\delta_0)^3 + c_4(\delta-\delta_0)^4\right) e^{-(\delta-\delta_0)^2/2\sigma^2} \tag{3.19}$$

Once the product $\Sigma_X \Sigma_Y$ in the van der Meer scan has been determined, Equations 3.16 and 3.15 can be combined to extract the visible cross-section via

$$\sigma_{vis} = \mu^{vis}_{MAX} \frac{2\pi \Sigma_X \Sigma_Y}{\Sigma (n_1 n_2)} \tag{3.20}$$

In reality, the analysis of van der Meer scans is complicated by many additional factors, such as the contribution of beam-gas and afterglow (long-lived activation of materials in the IR) to the measured rate, the possible transverse coupling and non-factorization of the transverse profile into independent x and y profiles, the uncertainty in the measurement of the beam intensities, the slow growth of the emittance and decay of the beam during the scan itself, the time-dependent uncertainty in the assumed beam positions, and other effects.

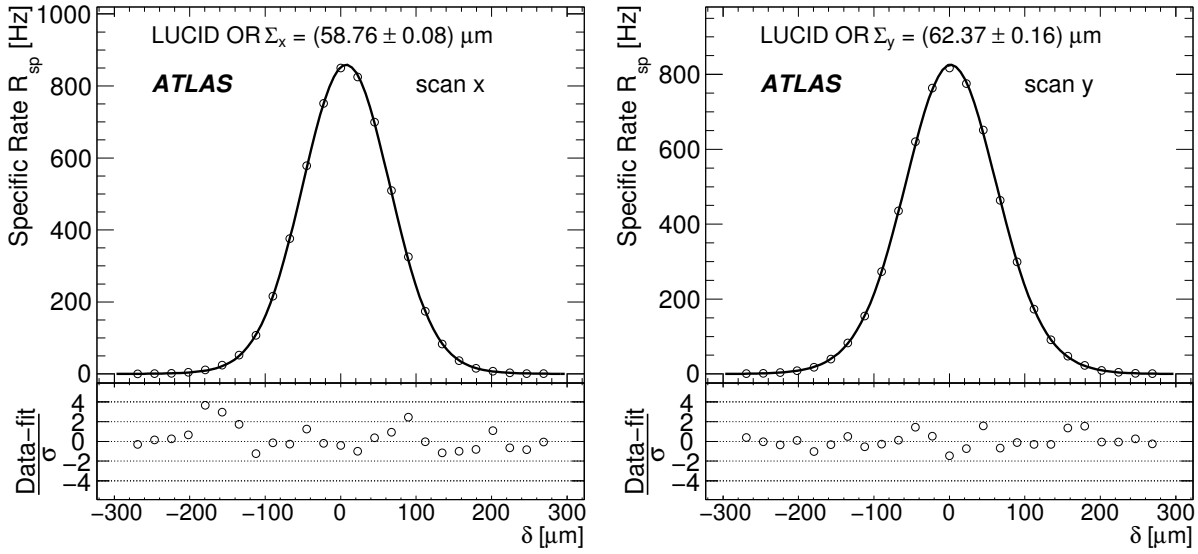


Figure 3.1: Example of beam profile overlap curves from horizontal (left) and vertical (right) separation scans performed during an ATLAS vdM scan, from [176].

3.2 The Relativistic Heavy Ion Collider

The Relativistic Heavy Ion Collider (RHIC)[177; 178; 179] is one of two operating heavy ion colliders in the world (the other being the LHC), the second highest energy proton-proton collider and, since the closing of the Tevatron, the only operating hadron collider in the United States. It is also the only spin-polarized proton collider in operation[180], an important capability for studying the spin structure of the proton. RHIC is located in Brookhaven National Laboratory (BNL) in Upton, New York, about 60 miles East of New York City.

Of the four original detectors at RHIC, PHOBOS and the Broad RAnge Hadron Magnetic Spectrometers Experiment (BRAHMS) have completed operations after finishing their scientific mission, while the Pioneering High Energy Nuclear Interaction eXperiment (PHENIX) and the Solenoidal Tracker at RHIC (STAR) continue to operate.

RHIC had its first collisions in 2000 and has been operating at design energy since 2001. The machine has provided collisions in a wide variety of systems (including unpolarized pp, $\uparrow p$+$\uparrow p$, Au+Au, d+Au, Cu+Cu and more recently Cu+Au and U+U) and collision energies (as high as $\sqrt{s} = 500$ GeV for pp running and as low as $\sqrt{s_{NN}} = 7.7$ GeV in Au+Au running for the beam

energy scan program). One of the charges of RHIC experiments was to verify the formation of the quark-gluon plasma and study its properties. In 2005, the four experiments published review articles summarizing their experimental perspective on the QGP[181; 182; 183; 184].

3.2.1 RHIC Accelerator Chain

The RHIC injection and accelerator chain is pictured in Figure 3.2.

A major difference between RHIC and proton-proton collider designs is that RHIC is able to provide collisions between beams of different species. Each ring must accomodate Z/A values from 1 (protons) to ~ 0.4 (gold ions) while maintaining the same frequency. Additionally, while proton-proton machines may only be interested in operation at the highest reachable energy, RHIC was designed to operate at a number of energies for experimental reasons.

RHIC itself is an intersecting storage ring with two independent rings, called yellow and blue, each 3834 m in circumference. The RHIC rings are hexagonally shaped with rounded corners and cross each other at six points so that the total distance seen by each ring is the same. The RHIC radiofrequency is 28 MHz, which defines 360 RF buckets of approximately 35.7 ns. In normal RHIC operation, at most every third RF bucket is used, allowing for 120 distinct bunches. The superconducting magnets are cooled by liquid helium to 4 K to operate.

RHIC utilizes 396 dipole (shown in Figure 3.3) and 492 quadrupole magnets in both rings, with a dipole field of ≈ 3.5 Tesla for 100 GeV Au ions and quadrupole gradient of 71 T/m, and over 900 smaller sextupole and corrected magnets. Inside the RHIC interaction regions, both beams share a single beam-pipe. Figure 3.5 shows the betatron function as the beam is squeezed in preparation for collisions in the IPs. β^* is the lowest for IP's 6 and 8 (the PHENIX and STAR interaction points) and somewhat higher in the other four. At injection, $\beta \approx 10$ m.

The injection sequence for Au and Cu ions is,

$$\text{Tandem} \xrightarrow{HITL} \text{Booster} \rightarrow \text{AGS} \xrightarrow{ATR} \text{RHIC}$$

in the following description, we focus on Au ions specifically. The sequence is pictured diagramatically in Figure 3.4.

The Tandem Van de Graaf generates negatively charged ions (Au^{-1}) from a sputter ion source, which are then accelerated and run through a stripping foil which removes many of the outer

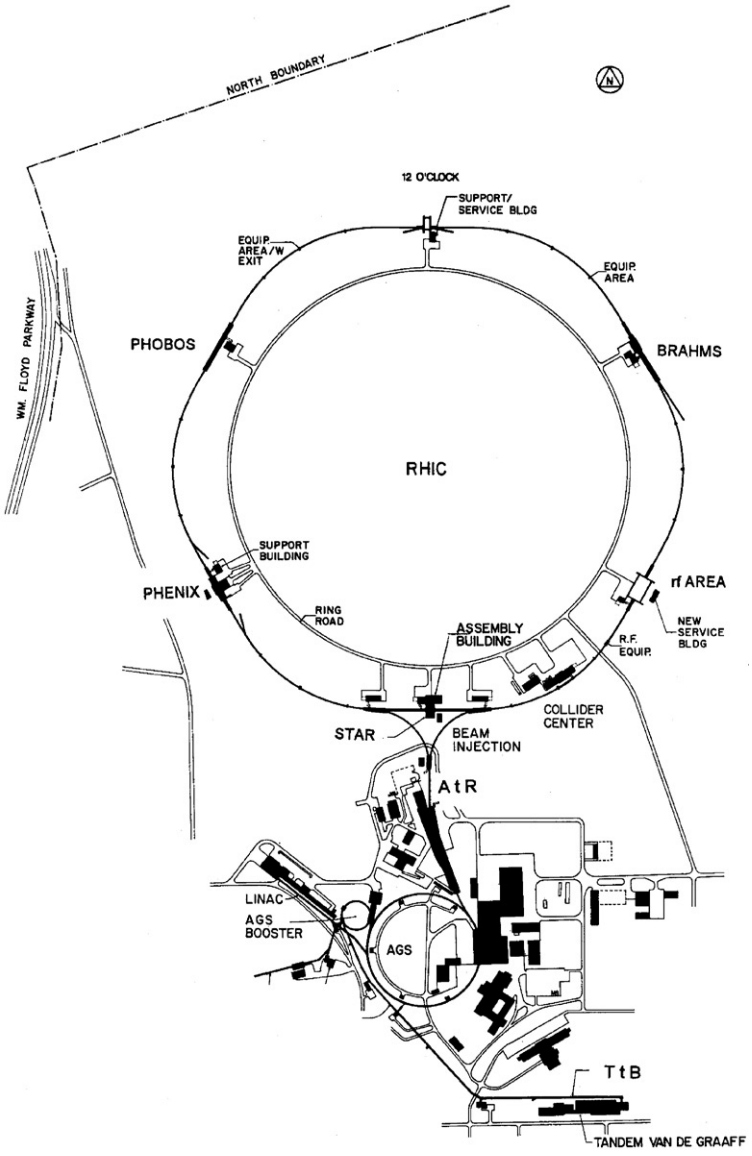


Figure 3.2: Diagram of the RHIC complex, from [178].

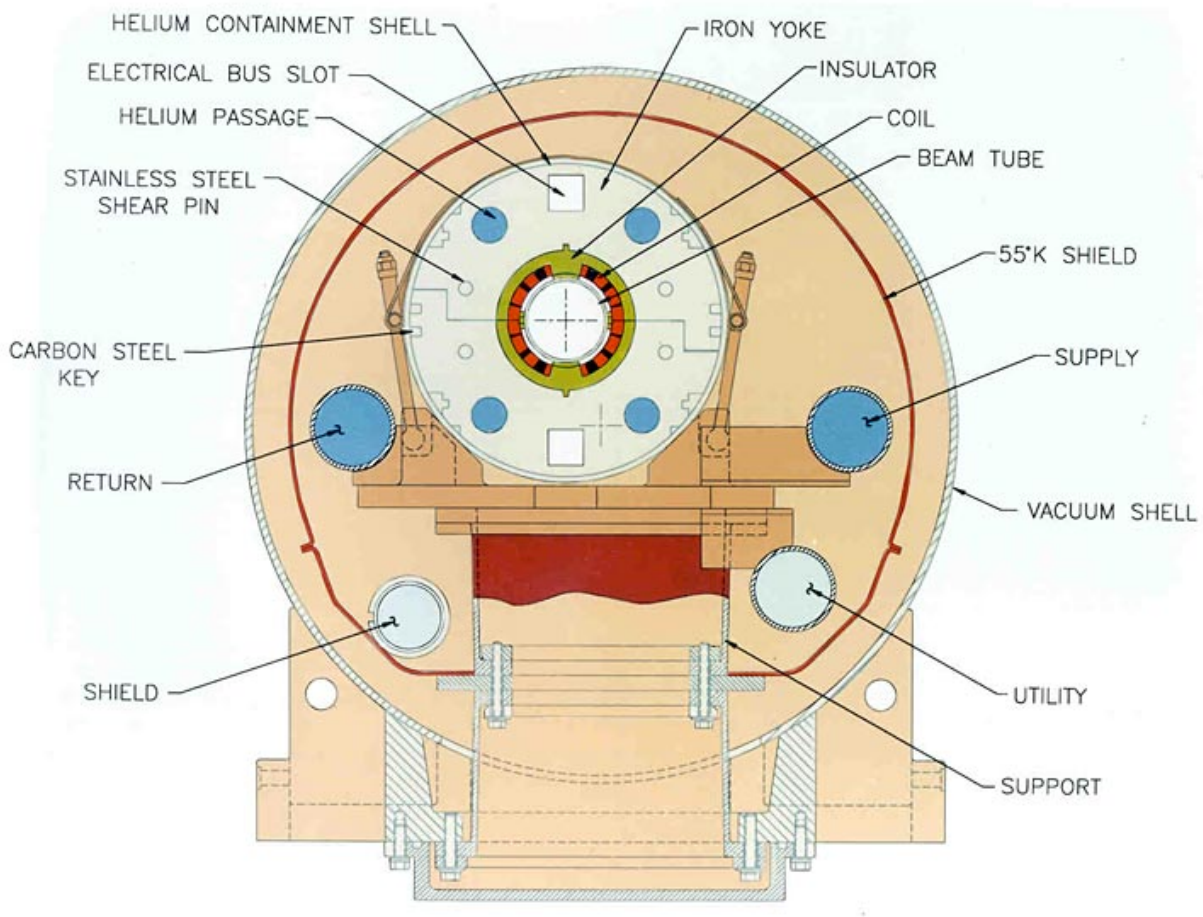


Figure 3.3: Cross-section drawing of a RHIC dipole.

electrons, leaving it in the state Au^{+12} ($Q_T = +12$). The ions are then accelerated back to ground, having gained an energy of 1 MeV/nucleon in the process.

They then move through an 850 m transfer line called the Heavy Ion Transfer Line to the Booster Synchotron, along the way becoming further ionized to the charge state Au^{+32}. There are two tandems that can work in parallel, with the second serving as a redundant backup (in the case of A+A collisions) or the source of the second specie (in the case of asymmetric A+B collisions. A new Electron Beam Ion Source (EBIS) was comissioned in 2011 and has taken over from the Tandem in providing ions to the Booster. The EBIS is able to provide new species of ions entirely, such as the uranium used in 2012 running.

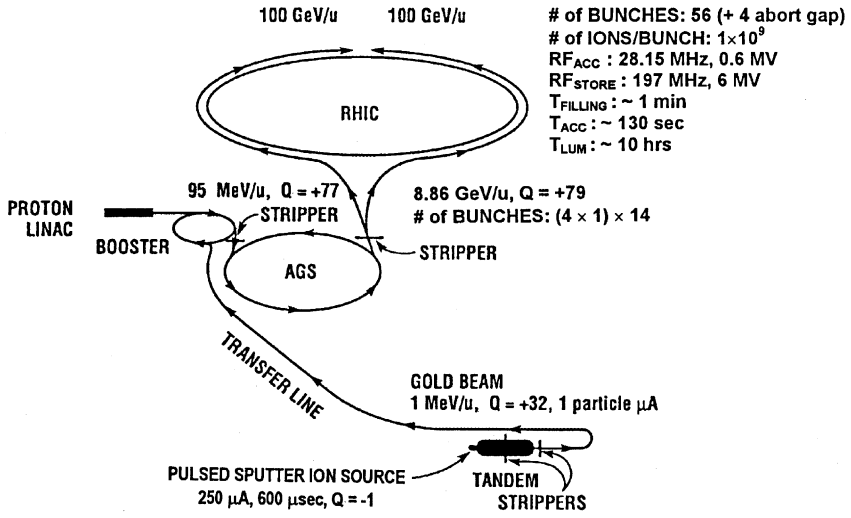


Figure 3.4: Diagram of the Au injection sequence in RHIC, from [178].

The Booster Synchotron is a compact hadron synchotron which accepts long ion pulses from the tandem and accumulates them into 6 bunches, which are accelerated to 95 MeV/nucleon and then injected into the AGS. A foil at the Booster strips almost all of the remaining electrons except the two most tightly bound $1s$ pair, leaving the ion in a Au^{+77} state.

The AGS is filled with 24 Au bunches from 4 Booster fills which are then rebunched into four Au bunches to be used in RHIC and accelerated to 8.86 GeV/nucleon. Ions exit into the AGS-to-RHIC (AtR) transfer line, becoming full stripped (Au^{+79}) on the way out. After injection into the Alternating Gradient Synchotron (AGS), the last two electrons are stripped, resulting in fully ionzed Au^{+79}, and gold nuclei are accelerated to 8.9 GeV/nucleon. At this point, they are injected into the yellow or blue rings via the AGS-to-RHIC transfer line. The highest achievable energy for ions at RHIC has been 100 GeV.

For protons the injection sequence is slightly different,

$$\text{Linac} \rightarrow \text{Booster} \rightarrow \text{AGS} \xrightarrow{ATR} \text{RHIC}$$

Protons are generated in the linear accelerator (Linac) from a hydrogen ion source and accelerated through nine RF cavities in a 140 m tunnel to an energy of 200 MeV, reach an energy of 1.5 GeV in the Booster and 25 GeV in the AGS. The highest achievable proton energy is 250 GeV.

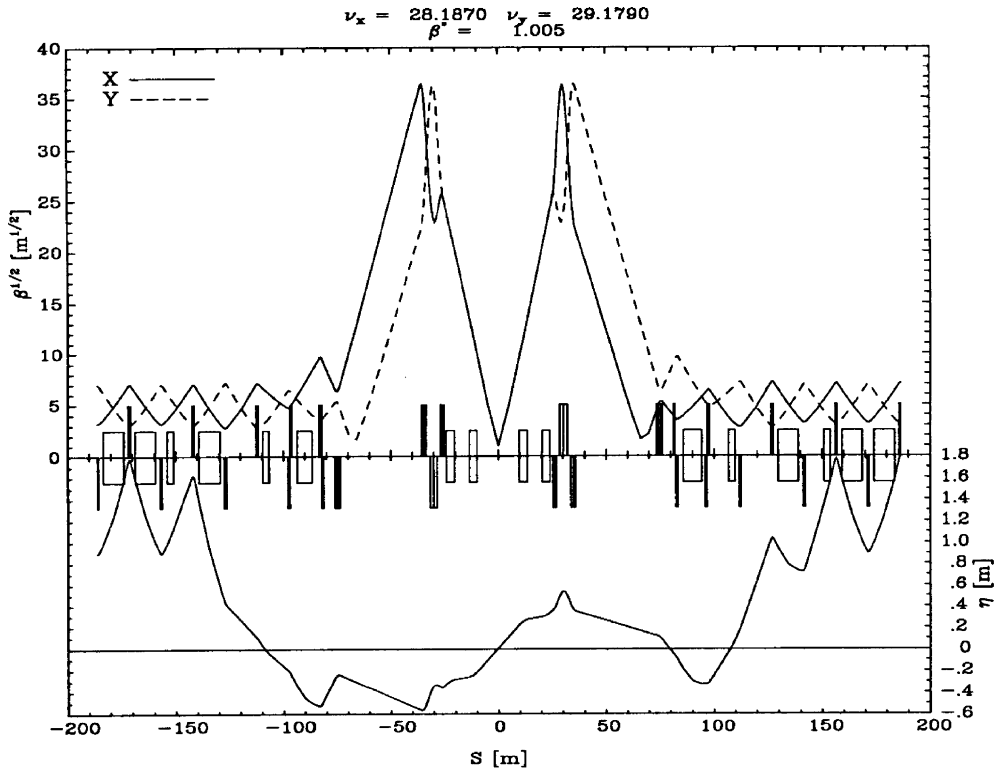


Figure 3.5: RHIC betatron $\beta(s)$ function in the vicinity of the IP, from [178].

In d+Au collisions, the magnetic rigidity at injection is different between the two beams (86 T·m for Au^{+79} and 69.48 T·m for d) such that the frequency is the same. Deuterons and gold ions were provided separately by the two Tandems.

3.2.2 Performance for 2007-2008 d+Au and pp runs

RHIC Run 8 ran between November 2007, when cryogenic cooling of the RHIC magnets began, to March 2008. The machine ran in three modes, the third of which was a low energy Au+Au mode which is not relevant for this work.

The first mode[185] was 100.7 GeV / nucleon d on 100.0 GeV / nucleon $^{197}Au^{+79}$ collisions, with deuterons and gold ions in the blue and yellow rings, respectively. The highest luminosity achieved fill scheme consisted of 95 bunches with 10^{11} deuterons and 10^9 ions per bunch and $\beta^* = 0.7$ m, corresponding to a peak instantaneous luminosity of $L_{peak} = 31 \times 10^{28}$ cm^{-2}s^{-1}. Over 63

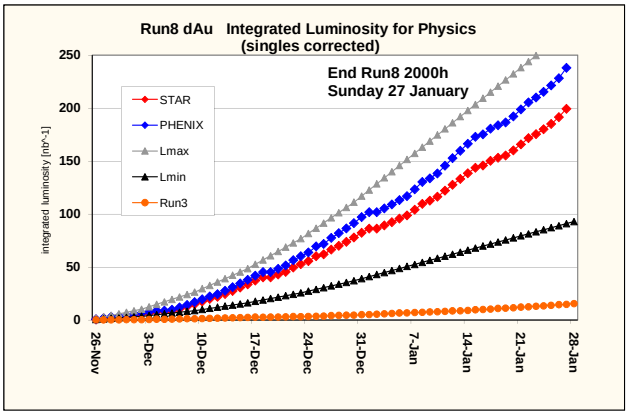


Figure 3.6: Integrated luminosity in the PHENIX Run 8 d+Au running as a function of time, from [185]. Luminosity delivered to PHENIX is shown in blue.

days of physics data-taking, PHENIX collected 238 nb^{-1} of data, about ten times more integrated luminosity than the Run 3 d+Au data in 2003[186], which occured with a $\beta^* \approx 2.5$ times as large, half as many bunches and half the physics days. The integrated luminosity as a function of time is shown in Figure 3.6.

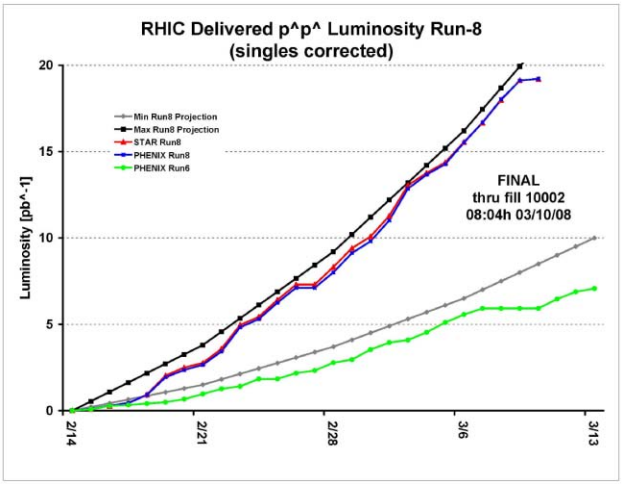


Figure 3.7: Integrated luminosity in the PHENIX Run 8 pp running as a function of time, from [187]. Luminosity delivered to PHENIX is shown in blue.

The second mode[187] was 100.2 GeV on 100.2 GeV polarized proton-proton collisions. The highest luminosity achieved fill scheme consisted of 109 bunches with 1.5×10^{11} protons/bunch and $\beta^* = 1$ m, corresponding to a peak luminosity of 35×10^{30} cm^{-2}s^{-1}. Over 24 days of physics data-taking, PHENIX collected 19.2 pb^{-1} of data, about half as much data as was taken in 100 GeV *pp* operation in Run 6. The average store polarization (predominantly vertical in PHENIX) was 50% (40%) in the Blue (Yellow) ring, although for the purposes of this work, the polarization is not relevant. The integrated luminosity as a function of time is shown in Figure 3.7.

3.3 The Large Hadron Collider

The Large Hadron Collider (LHC) is a proton-proton collider designed to answer the foremost open questions in high-energy physics, such as the existence of the Higgs boson and beyond the Standard Model particles predicted by supersymmetry that could serve as viable dark matter candidates. It is the largest and highest-energy collider ever built. The LHC is located at CERN (originally the *Conseil Européen pour la Recherche Nucléaire*, now stylized in English as the European Organization for Nuclear Research), on the outskirts of Geneva, Switzerland, although the underground LHC tunnel falls along the Swiss-French border.

There are seven detectors at the LHC. The two largest are A Toroidal LHC Apparatus (ATLAS) and the Compact Muon Solenoid (CMS) which are large, general-purpose detectors capable of exploring new physics that could be reached at LHC energies. A Large Ion Collider Experiment (ALICE) was originally designed as the detector dedicated to exploring proton-nucleus and nucleus-nucleus collisions at the LHC. However, as the theoretical understanding and experimental focus of ultrarelativistic heavy ion physics has evolved since the original designs of ALICE, ATLAS and CMS, it has turned out that the latter two are equally suited to explore these collisions. There are four other, more specialized detectors: the Large Hadron Collider beauty (LHCb) detector, which focuses on CP violation in the bottom quark sector; the TOTal Elastic and diffractive cross section Measurement (TOTEM) detector shares an intersection point with CMS and measures the different types of proton-proton cross-sections; the Large Hadron Collider forward (LCHf) detector shares an interaction point with ATLAS and measures physics processes in the very forward direction as a way to understand cosmic ray astroparticle physics; and the Monopole and Exotics Detector At

the LHC (MoEDAL) shares a cavern with LCHb and searches for rare, new physcs objects such as magnetic monopoles.

The first proton beams circulated in the LHC in September 2008 but later that month an electric fault caused a magnet quench and the explosive expansion of liquid helium, resulting in substantial damage to the LHC magnet and cryogenic systems. Operations were restored and the first collisions at 450 GeV per beam occured in November 2009. Since then, the LHC has collided protons at a number of energies including 1.38 TeV, 3.5 TeV and 4 TeV, as well as lead beams at 1.38 GeV/nucleon ($= \frac{Z}{A}$3.5 TeV) and lead ions on protons at 1.58 TeV/nucleon on 4 TeV.

On July 4th, 2012, coincidentally the author's 26th birthday, both ATLAS and CMS announced the observation of a new particle with a mass of 125 GeV[188; 189]. Further studies[190; 191] of the couplings and quantum numbers of the particle have so far showed that it is consistent with a Standard Model Higgs boson.

3.3.1 LHC Accelerator Chain

The LHC[192; 193; 194; 195] sits in the 26.7 km long former Large Electron Positron Collider (LEP) tunnel at a depth ranging from 70 to 140 m. Construction began in 2001 after the shutdown of LEP. The LHC rings are octagons with rounded edges, consisting of eight arc regions with strong dipole fields and eight straight regions, with an interaction region in the middle of each. IR1 is the Southern most and the location of ATLAS. The remaining IR's are numbered clockwise, with the injection into ring 1 at IR2, the injection into ring 2 at IR8, and beam dumps at IR6. The beams switch inner/outer position at IR1, 2, 5 and 8, so that both beams see the same total distance.

The LHC RF system is composed of 8 RF cavities per beam and operates at 400 MHz. This defines RF buckets of 2.5 ns. However, the fill structure of the SPS generally limits the bunches to multiples of 25 ns apart, resulting in a total of (26.7 km/3 $\times$ 10^8 m/s)/25 ns = 3560 buckets that can be filled with beam. These are enumerated with a bunch-crossing identifier (BCID). The Run I bunch spacings have been 50 to 200 ns long, with high luminosity 25 ns fill schemes planned after the long shutdown.

To steer the 3.5 GeV beams, the superconducting LHC dipole magnets must produce an 8.3 T field, which requires a temperature of 1.8 K. The two counter-rotating beams, which lie in separate vacuum chambers, are each rotated by the same magnets as shown in Figure 3.8. Each arc contains

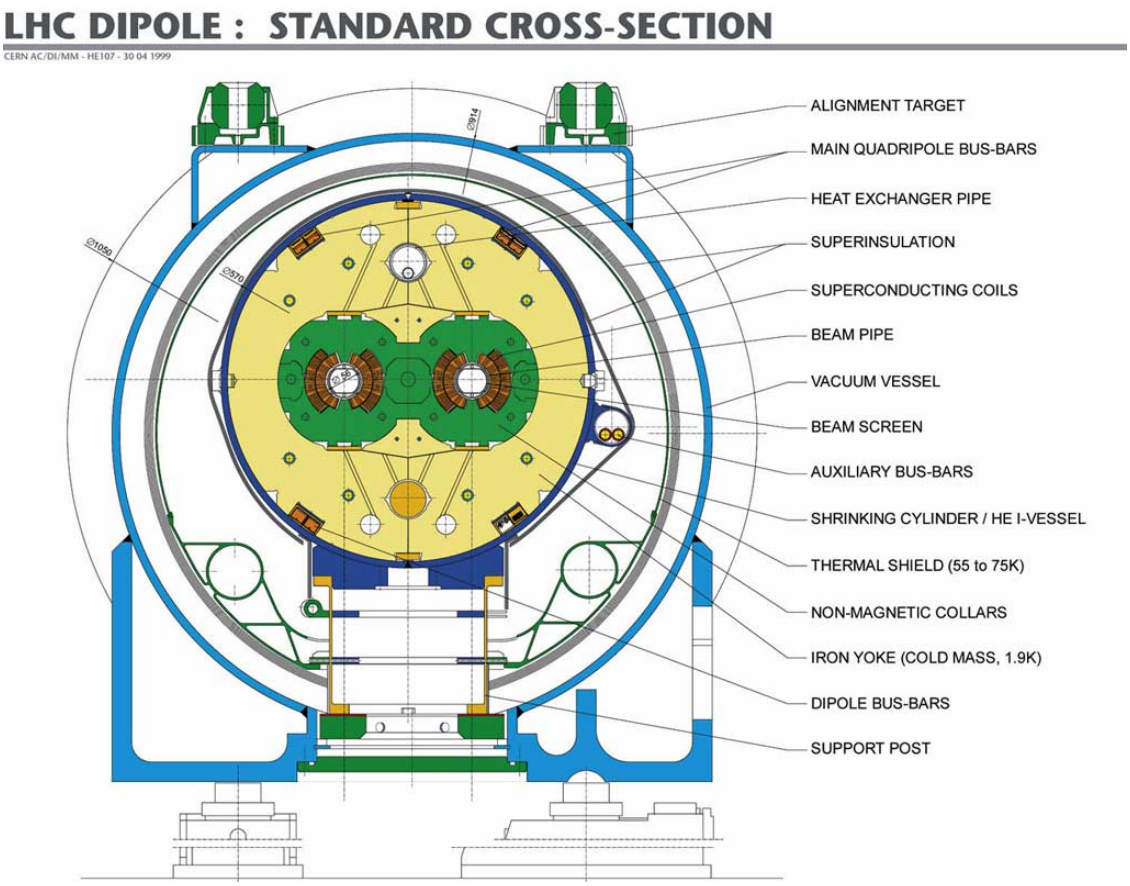


Figure 3.8: Drawing of LHC cryodipole.

392 main quadrupoles and 1232 main dipoles and over 6000 smaller multipole corrector magnet. In the ATLAS IR, the crossing plane is typically the vertical plane with a crossing angle of ≈ 280 μrad. β at injection is typically ≈ 10 m, but the LHC design calls for squeezing to β^* as low as 0.55 m at the high luminosity IPs (see Figure 3.9).

The beam is dumped with a septum magnet which redirects the beams vertically into two dump blocks 700 m away which consist of an 8 m-deep graphite absorber, a low-Z material which spreads the resulting hadronic shower over a large transverse area. The dump blocks are surrounded with concrete shielding.

The injection sequence for protons is

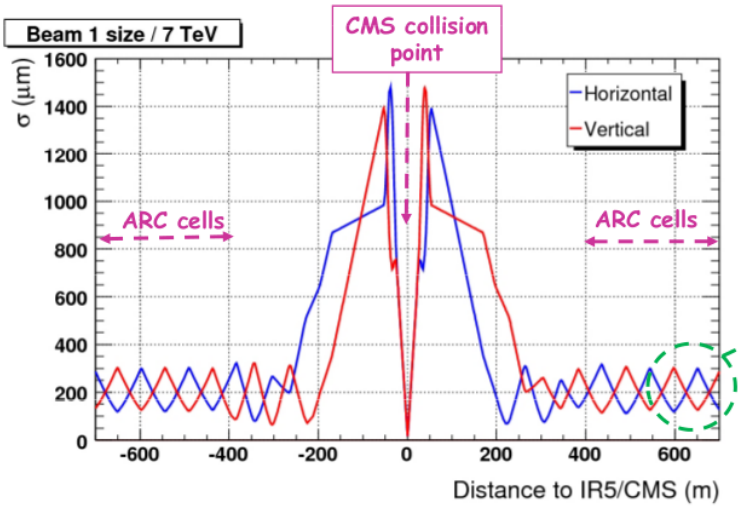


Figure 3.9: Cartoon of $\beta(s)$ near an LHC interaction point.

$$\text{Linac 2} \rightarrow \text{PS Booster} \rightarrow \text{Proton Synchrotron}$$
$$\rightarrow \text{Super Proton Synchrotron} \rightarrow \text{LHC}$$

Protons are produced from an ionized hydrogen source and accelerated to 50 MeV in the Linac 2, then transferred to the Proton Synchrotron Booster (PSB) where they reach 1.4 GeV. Then, proton bunches are accelerated to 26 GeV in the Proton Synchrotron (PS) and up to 450 GeV, the LHC injection energy, in the Super Proton Synchrotron (SPS). After injection into the LHC, the beams are brought to the desired energy, 4 TeV in the case of the p+Pb data considered here. The entire sequence is illustrated in Figure 3.10.

The injection sequence for lead ions begins in a different way,

$$\text{Linac 3} \rightarrow \text{Low Energy Ion Ring} \rightarrow \text{Proton Synchrotron}$$
$$\rightarrow \text{Super Proton Synchrotron} \rightarrow \text{LHC}$$

Lead ions with charge state Pb^{+27} are generated in the Linac 3, where they are stripped to Pb^{+42} via carbon foil and accelerated to 4.2 MeV/nucleon. They reach 72 MeV/nucleon in the Low Energy Ion Ring (LEIR), and 6 GeV in the PS. The lead beam is fully stripped with an

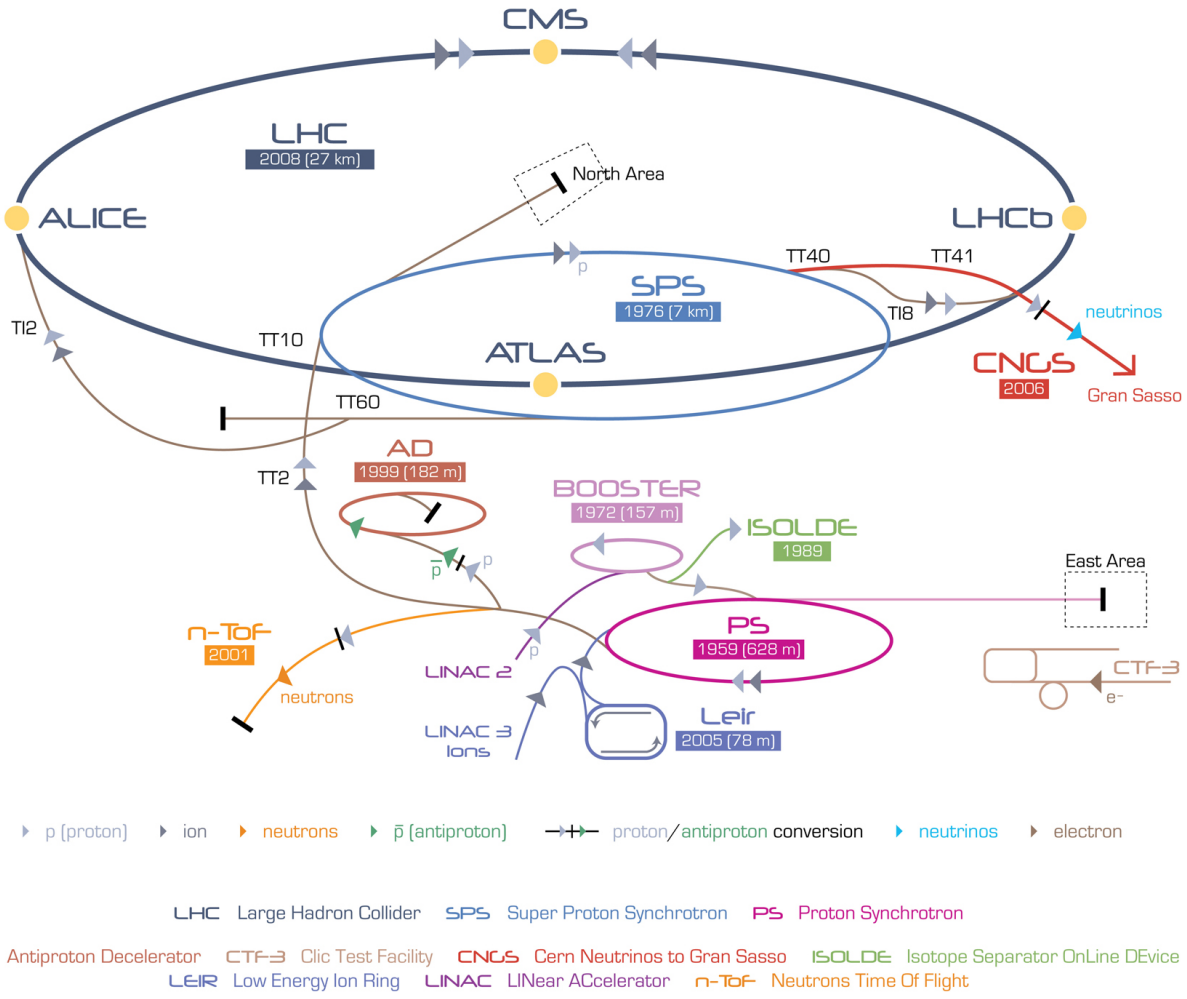


Figure 3.10: Diagram of LHC injection chains.

aluminum foil to Pb^{+82} as it is transferred from the PS to the SPS, and then accelerated to 177 GeV/nucleon, the LHC injection energy.

3.3.2 Performance for 2013 p+Pb and pp runs

The ATLAS p+Pb and pp running took place in January and February 2013.

The first run mode was 4 TeV p on 1.58 TeV/nucleon $^{208}Pb^{+82}$. The successful p+Pb running followed a failed attempt at a short p+Pb run in November 2011 and a short, few hour pilot run in September 2012 resulting in a small 1 μb^{-1} of data. The kinematics of the p+Pb running are dictated by the requirement of equal magnetic rigidity, corresponding to 4 TeV per unit charge.

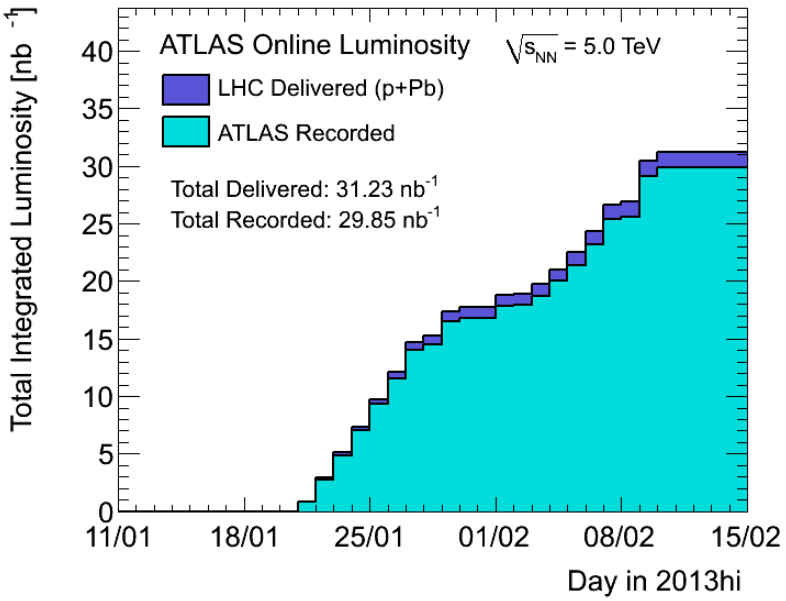


Figure 3.11: Integrated luminosity in the ATLAS 2013 p+Pb running as a function of time.

This results in a center of mass energy (neglecting the mass) of

$$p_p^\mu = (E, 0, 0, +E)\,,\ p_{Pb}^\mu = \left(\frac{Z}{A}E, 0, 0, -\frac{Z}{A}E\right) \tag{3.21}$$

$$s = (1 + Z/A)^2\, E^2 - (1 - Z/A)^2 E^2 = 4(Z/A)E^2 \tag{3.22}$$

$$\Rightarrow \sqrt{s} = 1.255E = 5.02 \text{ TeV} \tag{3.23}$$

The center of mass system $(p_p{+}p_{Pb})^\mu = (+5.58 \text{ TeV}, 0, 0, +2.42 \text{ TeV})$ is related to the laboratory frame by a boost $\beta = 0.434$ (or rapidity shift $y = \tanh^{-1}(\beta) = 0.465$) in the proton direction.

The highest luminosity fill scheme was 296 bunches colliding in ATLAS (338 bunches total), 1.6×10^{11} and 12×10^7 protons and ions per bunch, and $\beta^* = 0.8$ m. The peak instantaneous luminosity of 1.1×10^{29} cm^{-2}/s^{-1}. All of these were a dramatic improvement over the pilot run, which featured an $\mathcal{L}$ of 10^3 smaller, a β^* more than an order of magnitude larger and only 8 colliding bunches. After taking approximately 18 nb^{-1} the species in the rings were flipped, resulting in 13 nb^{-1} of collisions with the reversed Pb+p kinematics. Of the total 31 nb^{-1} delivered by the LHC, ≈ 30 nb^{-1} was recorded by ATLAS.

The integrated luminosity as a function of time is shown in Figure 3.11.

Chapter 4

The PHENIX Detector

CORIOLANUS

These eyes are not the same I wore in Rome.

VIRGILIA

The sorrow that delivers us thus changed

Makes you think so.

CORIOLANUS

Like a dull actor now,

I have forgot my part, and I am out,

Even to a full disgrace

Coriolanus, Act V, Scene III

The Pioneering High-Energy Nuclear Interaction eXperiment (PHENIX)[196] is a general-purpose detector designed to explore nuclear matter under extreme conditions and explore the spin structure of the proton, with particular strengths in the detection of rare high-p_{T} and electromagnetic probes. The PHENIX project originally began as the merger of four failed proposals for detectors at RHIC (called the TALES, SPARC, OASIS and DIMUON detectors). The initial detector was designed and built through 2001, when collisions at RHIC began, and has been undergoing detector upgrades and the installation of entirely new subsystems ever since. Most recently, the silicon vertex tracker (VTX)[197] was installed at mid-rapidity in 2011 and the forward vertex tracker (FVTX)[198] was installed in 2012.

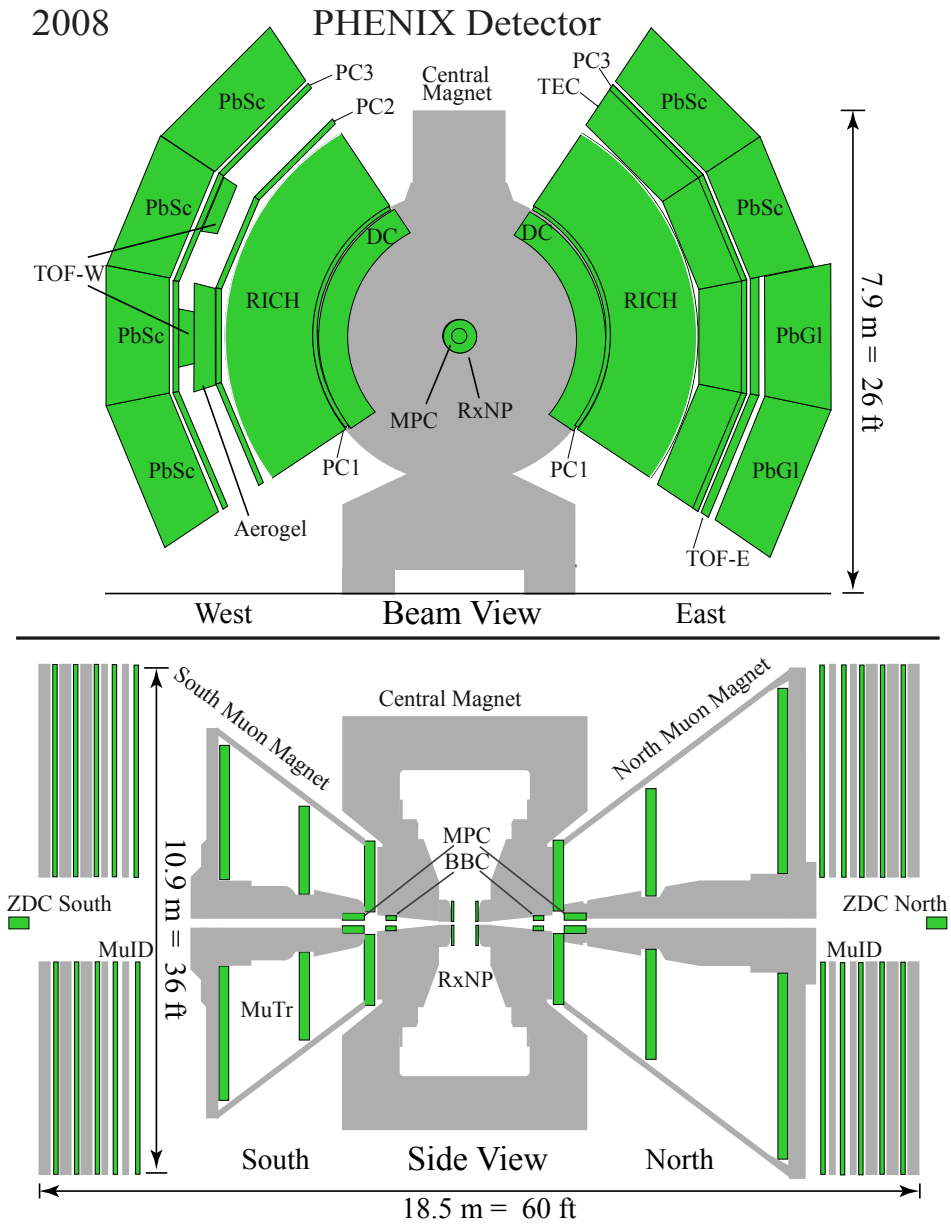


Figure 4.1: Diagram of the PHENIX detector systems during d+Au and pp data-taking in 2008. Cross-section looking down the beam-axis, in the $\hat{x}$-$\hat{y}$ plane (top) and in the $\hat{z}$-$\hat{y}$ plane (bottom).

PHENIX consists of four major spectrometer arms and a number of global or forward detectors. The detector subsystems are shown in Figure 4.1. Two of the four arms are the North and South muon arms, which sit at forward and backward rapidity on either side of the collision point. They are not used in this work.

In the immediate future, an extension to the muon-piston calorimeter (MPC-EX)[199] is due to be installed in time for a putative p+Au run in 2015. The MPC-EX is designed to separate photons from π^0's in the forward region and add significant experimental capability for probing the gluon nPDF, among other physics. On a longer timescale, PHENIX has proposed a redesign of the detector as sPHENIX[200], a dedicated jet detector featuring hermetic hadronic calorimetry and other improvements.

4.1 PHENIX Central Arm Detectors

Two of the major original design goal of the PHENIX central tracking system are the ability to reliably reconstruct the trajectory of a charged particle as it passes through the central arm spectrometer from detector to detector over a wide range of momenta in a high-multiplicity Au+Au collision and the rejection of hadron backgrounds from measurements of electron tracks by a factor of 10^4. These are achieved through a combination of drift and pad chambers which allow for the measurement of the particle trajectory, particle identification detectors and a calorimeter capable of shower shape analysis.

The PHENIX Central Magnet (CM)[202] consists of two sets of inner and outer circular coils which provide an axial field parallel to the beam axis, with total field integral ranging from 0.43 to 1.15 T·m depending on the trajectory of the charged particle. The final design of the CM was guided by several physics motivations, such as a minimal field integral past $R > 200$ cm where the Drift Chamber begins and dense material near the collision zone to serve as hadron absorbers for the muon arms. The field was mapped using a number of techniques (including a conventional measurement of the field components on a large grid filling the volume, an analytic method which numerically solved Laplace's equation under the boundary conditions of the surface enclosing the field, and a detailed 3D simulation program) with an accuracy of 1 in 10^3. During 2008 d+Au running, the inner and outer coils were switched from a "++" to a "−−" polarity, indicating a

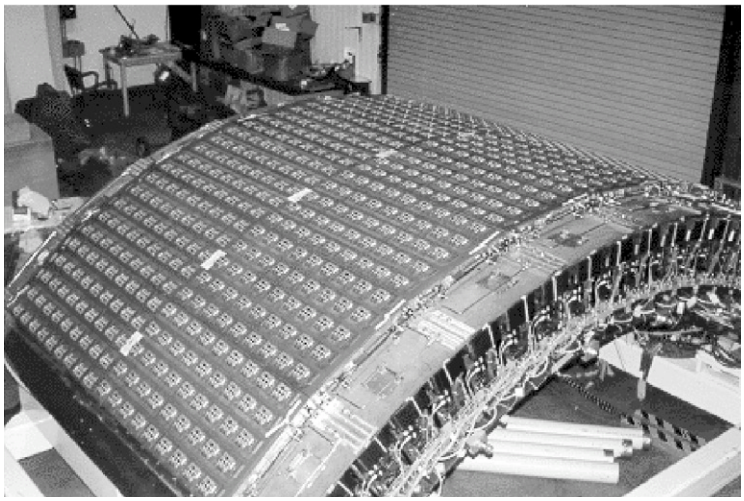


Figure 4.2: Photograph of a Drift Chamber / Pad Chamber 1 assembly in one PHENIX arm, from [201]

complete inversion of the field in the central region, and were kept in the "−−" configuration throughout the *pp* running.

In the PHENIX coordinate system, the West and East arms are in the $+\hat{x}$ and $-\hat{x}$ directions from the interaction region, respectively, the $+\hat{z}$ direction points North and the $+\hat{y}$ direction points away from the Earth. The detectors in the PHENIX central arms used in this work are described below in order of increasing radial distance from the beamline.

The central arms extend from $z = -90$ cm to $z = +90$ cm and begin at $\rho = +246$ cm in a cylindrical coordinate system. For a displaced vertex at z_{vtx}, the range of accessible angle θ is

$$\theta \in \left(\arctan\left(-\frac{246}{90 - z_{vtx}} \right), \arctan\left(\frac{246}{90 - z_{vtx}} \right) \right) \tag{4.1}$$

Or, equivalently, the pseudorapidity limits as a function of z-vertex are:

$$\eta^{-} = \log\tan\left(\frac{1}{2} \arctan\left(-\frac{246}{90 - z_{vtx}} \right) \right) \tag{4.2}$$

$$\eta^{+} = \log\tan\left(\frac{1}{2} \arctan\left(\frac{246}{90 - z_{vtx}} \right) \right) \tag{4.3}$$

At the nominal vertex position $z_{vtx} = 0$ cm, these equations give $(\eta^{-}, \eta^{+}) = (-0.358, +0.358)$,

which is the nominal PHENIX acceptance. At $z_{vtx} = \pm 20$ cm, the pseudorapidity ranges are $(\eta^-, \eta^+) = (-0.433, +0.281)$ and $(-0.281, +0.433)$, respectively. (The size of the pseudorapidity range is $\Delta\eta = 0.715$ in all cases.)

The azimuthal acceptance of the East arm is $\phi \in (\frac{41}{60}\pi, \frac{71}{60}\pi)$ (the West arm, which is not used in this analysis, is $\phi \in (-\frac{11}{60}\pi, \frac{19}{60}\pi)$) with a total range of $\Delta\phi = \frac{30}{60}\pi$.

4.1.1 Drift chamber and pad chambers

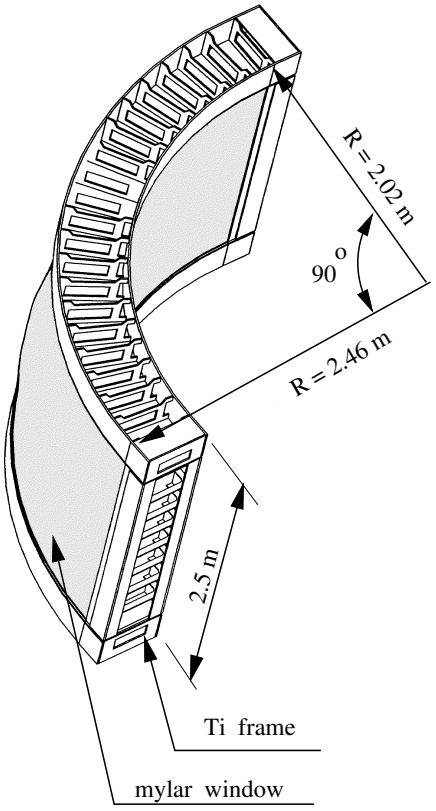


Figure 4.3: Schematic of the PHENIX Drift Chamber, from [203].

The PHENIX Drift Chamber (DC)[203] is the closest detector to the beampipe, with mirror images of it in each arm. Each DC is a cylindrical volume sitting between approximately 2 to 2.4 m from the beampipe, covering $\Delta\phi = \pi/2$ and ± 1.25 m in the $\hat{z}$ direction. The $\hat{r}$-$\hat{\phi}$ and $\hat{r}$-$\hat{z}$ planes are defined by a titanium frame, while the $\hat{\phi}$-$\hat{r}$ planes (the ones traversed by charged particles from

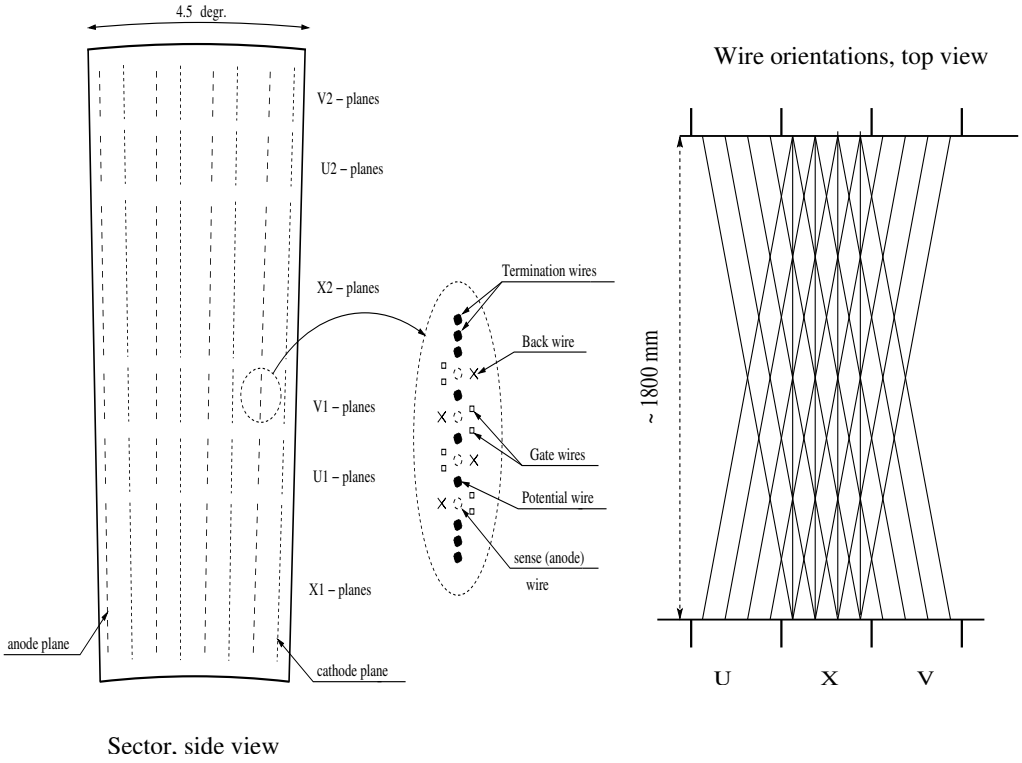


Figure 4.4: Wire arrangement in a PHENIX Drift Chamber module in the $\hat{\phi}$-$\hat{r}$ plane (left) and $\hat{\phi}$-$\hat{z}$ plane (right), from [203].

the interaction point) are 0.5 mm thick aluminum-mylar windows. Each arm consists of twenty sectors (also called keystones) stacked azimuthally (visible in Figure 4.3), with each sector covering $\pi/40 = 4.5°$. Each sector consists of six wire modules, stacked radially (shown in the $\hat{\phi}$-$\hat{r}$ plane on the left side of Figure 4.4). Each module consists of four cathode wire planes alternating with four anode wire planes, which lie in the $\hat{r}$-$\hat{z}$ plane. The X1 (first radially) and X2 (fourth radially) wire modules have twelve wires per plane and run along the beam axis. The U1 (second), V1 (third), U2 (fifth) and V2 (sixth) modules have four wires per plane and are slanted at a 6° angle with respect to the X1/X2 wires. Thus, while the X1 and X2 wires begin and end in the same sector, the U and V wires terminate in a sector adjacent to the one in which they begin (shown on the right side of Figure 4.4).

The DC is a multiwire proportional chamber (MWPC) filled with an ethane-argon mixture. Charged particles traversing a MWPC ionize the gas medium, resulting in the positive ions and

electrons drifting towards the anode and cathode wires, respectively, causing an avalance of other ionized particles as they get closer to the high electric field region of the wire. Since a constant drift velocity is necessary for resolving the position of the initial charge ionization, the electric field in the MWPC should be as constant as possible.

The PHENIX DC design includes some additional features beyond the basic MWPC concept. The inset in Figure 4.4 shows the wire configuration in the vicinity of the sense (anode or cathode) wires. Two gate wires and one back wire straddle each anode wire and the anode wires are separated by potential-shaping wires in the $\hat{r}$ direction. This "controlled geometry" design resolves the left-right ambiguity in each anode wire (since the back wire prevents charged particles drifting to it from that side), focuses drifting charge onto the sense wires (by shaping the field with the gate wires, thus decreasing the drift time) and ensures that the individual sense wires are isolated (by interspersing them with potential wires). In practice, a left-right ambiguity still exists for tracks that pass through the ± 2 region between the back wire and sense wire planes. In addition, to improve the performance of the pattern recognition in the expected high-multiplicty Au+Au collisions, the sense wires were cut in half, with each half able to read out separately.

The electric field in the drift chamber is modeled using the gas detector simulation code GARFIELD[204; 205]. The nominal single wire efficiency is $> 95\%$ with back wire efficiency (the efficiency for charge leaking past the back wire onto the sense wire) of $\approx 5\%$, and the single wire position resolutions were determined to be 165 μm and 2 mm in the $\hat{r}$-$p\hat{h}i$ plane and z direction, respectively.

The PHENIX central arms also contain three pad chambers (PC1, PC2 and PC3)[201]. They are multiwire proportional chambers filled with an argon-ethane gas mixture. PC1 and PC3 lie 2.5 m and 4.9 m from the interaction region, respectively, sandwiching the particle ID detectors (such as the RICH, see below) between them. The PC2, which due to financial considerations was only built in the PHENIX West arm, is not used in this analysis.

The PC1 in each arm is divided along the $\hat{\phi}$ direction into eight 49.5 cm-wide sectors, each 1.975 m long in the $\hat{z}$ direction and consisting of (in order of increasing radial distance), a cathode plane, an anode wire plane and a pixel pad plane, shown in Figure 4.5. The 58 anode wires are spaced at 8.4mm intervals along the $\hat{\phi}$ direction. To provide fine position resolution, save on readout channels and suppress hits from electronic noise, every 8.4×8.2 mm^2 cell in the pad plane is segmented into

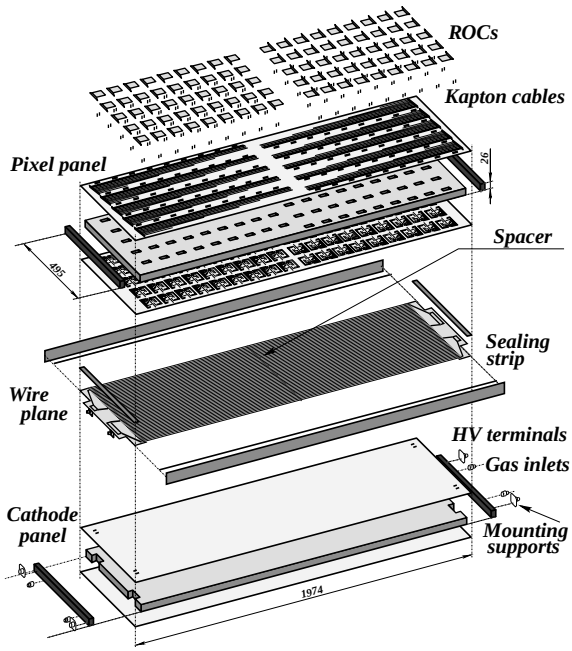


Figure 4.5: Schematic of the PHENIX Pad Chamber, from [201].

three pixels, all of which must sense an avalance from an anode wire to make a valid hit in the cell. Nine pixels from a 3×3 array of cells are ganged together in a common readout channel so that no pixel shares a readout channel with each of its neighbors, shown in Figure 4.6. The end result is a z position resolution of ± 1.7 mm. In total, each PC1 sector is 5.8cm thick in the radial direction, with the material inside corresponding to a total 1.2% of a radiation length. PC3 has a similar design but all dimensions in the $\hat{\phi}$ direction roughly twice as large since it is roughly twice as far away from the interaction region as PC1 (e.g. the anode wires are 16 mm apart and the cells are four times larger in area).

The pad chambers are the only non-projective elements of the central tracking system and thus critical for fully reconstructing charged particle trajectories. Although the DC provides a weak determination of the z position of a charged track due to the angle of the stereo wires, the measurement provided by PC1 is considered the definitive one. Similarly, although PC1 has some ability to resolve the ϕ position, this is more accurately determined by the DC. Together, the DC and PC1 provide direction vectors through the particle ID detectors and the PC3 is used to resolve ambiguities in the trajectories of particles landing in the outer calorimeter and can reject secondary

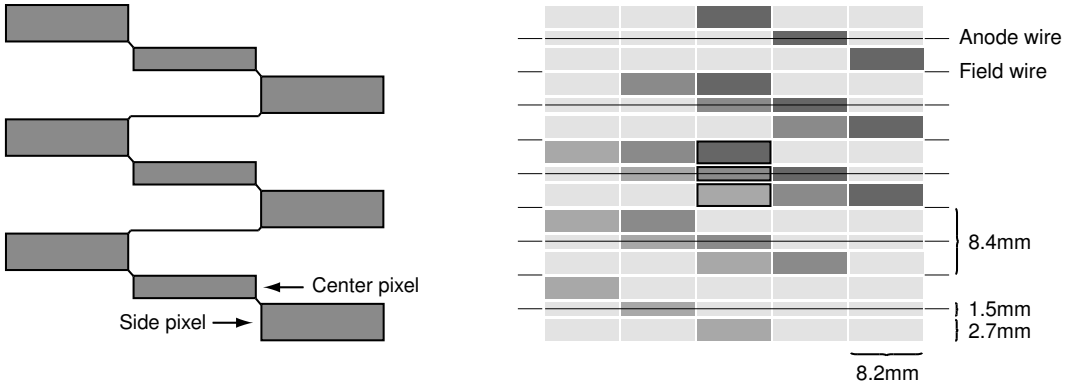


Figure 4.6: Diagram of the ganged pixel layout, from [203].

tracks produced by late decays and conversions.

The momentum resolution of the combined tracking systems is

$$\frac{\delta p}{p} = \frac{0.7\%}{p(GeV/c)} \oplus 1.1\% \tag{4.4}$$

and is multiplicity independent over the range of multiplicities observed in d+Au collisions.

In PHENIX, the quality of reconstructed tracks is related to which of the six DC modules the track has hits in, whether or not there is a matching PC1 hit, and if these hits are shared between other reconstructed tracks. The six quality bits are:

- 2^0 : hit in an X1 module
- 2^1 : hit in an X2 module
- 2^2 : hit in the UV modules, possibly shared
- 2^3 : unique hit in the UV modules
- 2^4 : PC1 hit, possibly shared
- 2^5 : unique PC1 hit

Typically, 4-6 of the quality bit are required for a good reconstructed track with possible additional requirement on a match to hits in the PC3. This results in a final track efficiency somewhat

lower than the single wire efficiency. Additionally, a small number of X1, X2, U and V wires are known to be broken during the d+Au and pp running. So as not to create localized regions of extremely poor efficiency, a detector position-aware set of quality cuts were ultimately used in this analysis (see Section 6.1.2).

4.1.2 Ring-imaging Ĉerenkov detector

Each of the PHENIX central arms includes a Ring-imaging Ĉerenkov (RICH) detector[206], which sits behind the pad chambers and in front of the electromagnetic calorimeter. The RICH detectors are crucial for providing $e^{\pm}/\pi^{\pm}$ discrimination capability, which is used to reject stiff drift chamber tracks from decay or conversion electrons masquerading as charged hadrons with a well-reconstructed momentum. The RICH is pictured in Figure 4.7.

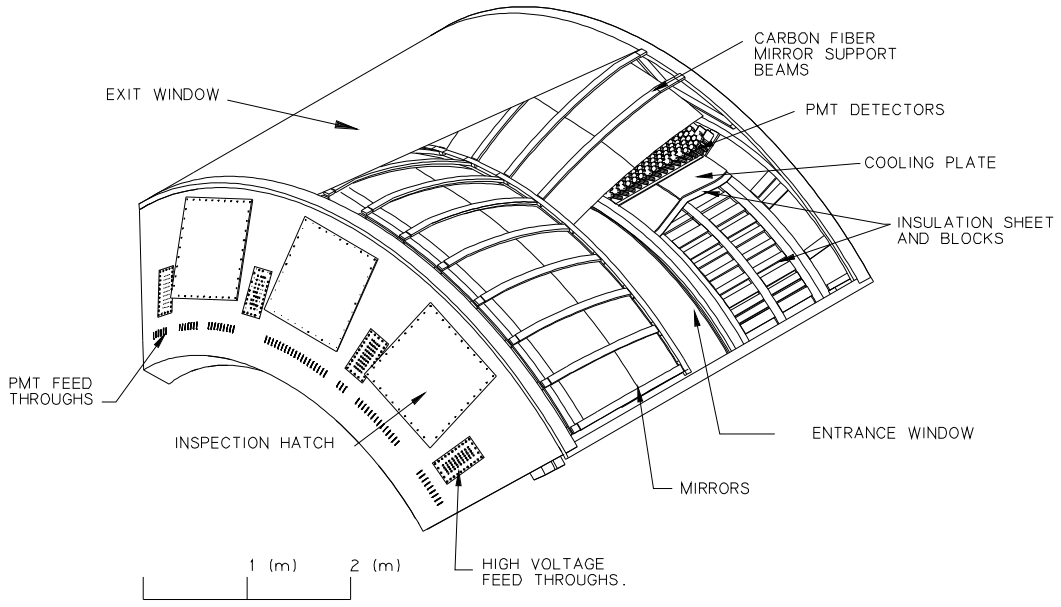


Figure 4.7: Schematic of the Ring Imaging Ĉerenkov detector, from [206].

Each detector has a volume of 40 m^3, and contains 48 composite mirror panels which form two intersecting spherical surfaces with a total reflecting area of 20 m^2. The mirrors focus Ĉerenkov light onto two arrays of 1280 PMTs, which are located on either side of the RICH entrance window. The RICH is filled with carbon dioxide (CO_2) gas, though it is also designed for operation with

ethane (C_2H_6) which has slightly different π/e separation capabilities but the disadvantage of being flammable. CO_2 has an index of refraction $n = 1 + 41 \times 10^{-5}$, and a charged particle moving faster than the speed of light in the CO_2 medium, c/n, will emit Ĉerenkov radiation which is reflected by the mirrors and collected in the PMTs. This correspond to $\gamma > 35$ particles, which is $p > 18$ MeV/c for $e^{\pm}$ but only $p > 4.8$ GeV/c for $\pi^{\pm}$. When filled with CO_2, the Ĉerenkov rings have a diameter of 11.8 cm and consist of 12 photons on average. Thus, searching for PMT's within this nominal radius of a charged track can distinguish electrons from other hadrons below ≈ 4.8 GeV/c.

In addition to the RICH, PHENIX also includes the Time of Flight (TOF) detector, which provides additional particle identification. The TOF sits 5.1m away from the collision vertex and consists of 10 panels of 96 segments, each with a scintillator slat and PMT. It follows the central arms in its pseudorapidity coverage ($\Delta\eta = 2 \times 0.35$) and subtends $\Delta\phi = \pi/4$ in azimuth, just in front of the lead-glass calorimeter sectors in the East arm. The TOF has a design timing resolution of $\sigma_{TOF} \approx 100$ ps and, in conjunction with information from the BBC and the tracking systems, can distinguish $\pi^{\pm}$ from $K^{\pm}$ up to 2.4 GeV/c and $K^{\pm}$ from $p/\bar{p}$ up to 4.0 GeV/c. Due to this capability, the TOF is used to calibrate the tracking momentum scale in the PHENIX.

4.1.3 Electromagnetic calorimeter

The PHENIX electromagnetic calorimeter (EMCal)[207; 208] complements the charged particle measurements performed by the DC and PCs with measurements of electromagnetic energy from photons and electrons. Two major design requirements of the EMCal were the ability to reconstruct the 2γ decays of highly Lorentz boosted π^0's (where the resulting photons can "merge" into an undistinguishable single energy deposit), and to discriminate hadronic from electromagnetic energy on the basis of the transverse shape of the shower.

The EMCal is comprised of a separate lead-scintillator (PbSc) and lead-glass (PbGl) calorimeters, two very different technologies with somewhat different capability. Historically, using both sections has served as an important systematic cross-check on the energy scale of key PHENIX results. The EMCal consists of eight sectors, six of which are PbSc and two (the bottom half of the East arm) are PbGl, situated just outside the PC3 at approximately 5 meters away from the interaction region. The performance of both technologies was evaluated with e^- test beams at the AGS at BNL (0.5-5 GeV) and the SPS at CERN (10-80 GeV). In both detectors, each individual

tower is optically isolated from the rest.

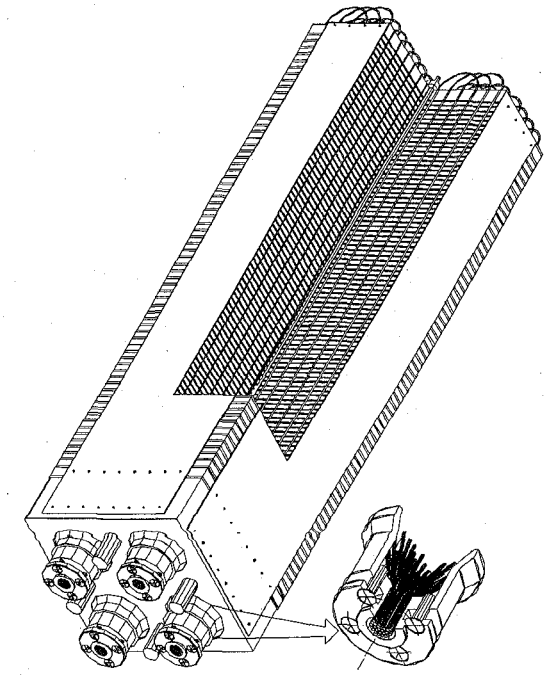


Figure 4.8: Cut-away drawing of PbSc tower, showing alternating tiles, wave-length shifting fibers and fiber used for laser monitoring system, from [208].

Each PbSc sector is an array of 36×72 (in the $\hat{\phi} \times \hat{z}$ directions, corresponding to a $\Delta\eta \times \Delta\phi = 0.01 \times 0.01$ segmentation) calorimeter towers, one of which is shown in Figure 4.8. The towers are "shashlik" style sampling calorimeter calorimeters composed of 66 alternating lead and scintillator tiles in which scintillation light is read out via wavelength shifting fibers[209] running perpendicularly to the plane of the tiles. Each tower is 18 radiation lengths long with nominal energy resolution,

$$\frac{\sigma_E^{PbSc}}{E} = \frac{8.1\%}{\sqrt{E(GeV)}} \oplus \frac{2.1\%}{E(GeV)} \tag{4.5}$$

where the first value is the stochastic term and the second value is from intrinsic nonuniformities in the detector such as tower boundaries and variations in the amount of light leaking out of either end of the tower. The likelihood of a given reconstructed cluster in the PbSc calorimeter coming from an electron or photon instead of a hadron is determined through a χ^2 parameter that compares

the topology of the deposited energy to reference data of real electromagnetic clusters. The timing resolution for > 500 MeV particles is ~ 120 ps for electrons and protons and ~ 270 ps for pions, and the position resolution for > 1 GeV particles is ~ 7 mm.

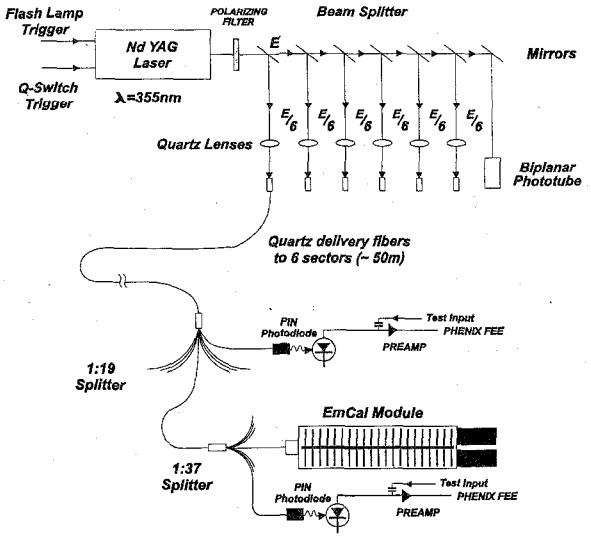


Figure 4.9: Daigram of PbSc laser monitoring system, from [207].

During the experimental run, the PbGl is monitored by a laser light distribution system (diagrammed in Figure 4.9) which supplies light at regular intervals to each calorimeter sector, which is then optically split onto a 2 mm diameter fiber which pentrates the center of each individual 2×2 tower module. The fiber is grated such that light exits in a way simulating the depth profile of a 1 GeV electromagnetic shower.

Each of the two PbGl sectors is an array of 48×96 (corresponding to a $\Delta\eta \times \Delta\phi = 0.008 \times 0.008$ segmentation) calorimeter towers which were originally used in the WA80 and WA98 experiment at CERN, shown in Figure 4.10. The PbGl towers are Ĉerenkov detectors homogenously composed of a lead, glass and lead-oxide (51%) medium which both serves as the Ĉerenkov radiator ($n = 1.648$) and produces the secondary particles in proportion to the incident particle energy, which are measured by the PMT at the end of the module. Each tower is ≈ 14 radiation lengths long with nominal energy resolution,

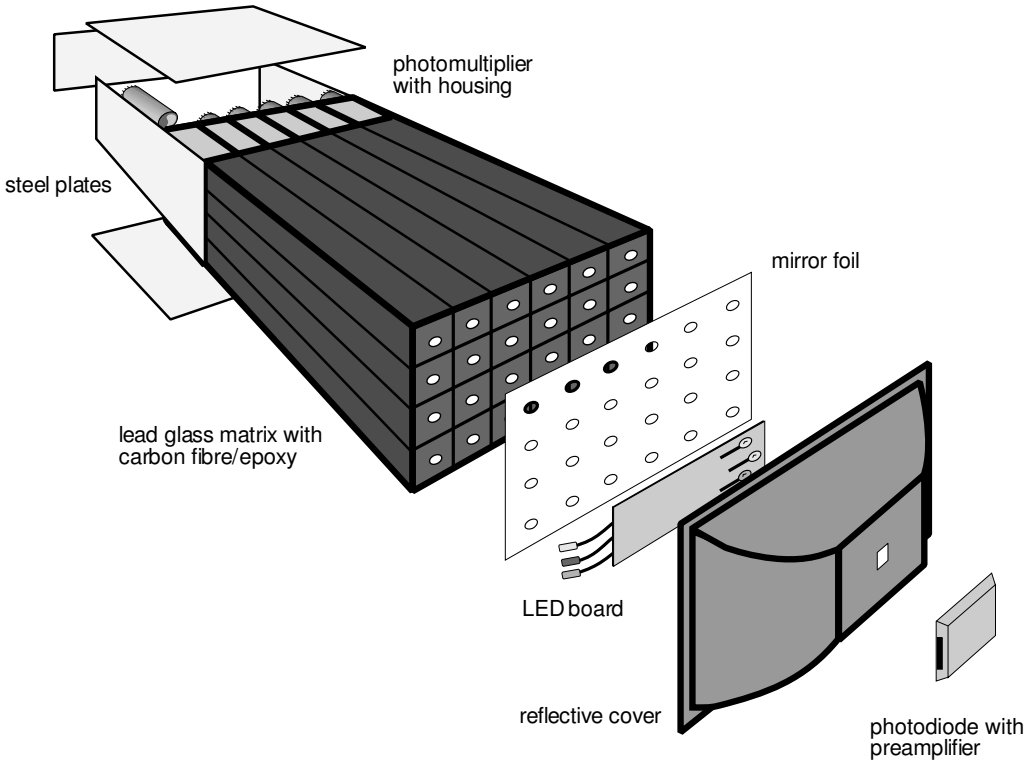


Figure 4.10: Schematic of PbGl supermodule assembly, from [207].

$$\frac{\sigma_E^{PbGl}}{E} = \frac{5.9\%}{\sqrt{E(GeV)}} \oplus \frac{0.8\%}{E(GeV)} \tag{4.6}$$

Like the PbSc, the PbGl can also discriminate between electromagnetic and hadronic showers on the basis of the shower shape (although hadrons deposit only a small fraction of their energy, due to the much larger interaction length). The timing resolution for > 500 MeV particles is ~ 200 ps, although hadronic showers have their apparent arrival time shifted by ≈ 800 ps since they reach the end of the tower before the Ĉerenkov radiation they produce at the front of the module. The position resolution is ≈ 10 mm for > 1 GeV particles.

The PbGl has a separate gain monitoring system from the original WA98 experiment which utilizes an LED system with pulse shapes resembling that of real electromagnetic showers[210].

For readout and fast trigger logic purposes (see Section 4.2.2), 12×12 groups of towers in each sector are collected in an EMCal SuperModule (SM), corresponding to 3×6 and 4×8 SM's per

PbSc and PbGl sector, respectively.

Although the shower shape and time of flight information is used to reject hadronic showers in most PHENIX analyses, in the present work it was found that the standard methods of rejecting these was causing a serious bias in the ability to properly reconstruct high-p_{T} jets which fragmented to a high-p_{T} π^0. In MC simulations, merged clusters originating from high-p_{T} π^0 decays had poor efficiency for the χ^2-based cuts, resulting in a bias on jet energy scale at high p_{T}. This, and the fact that the neutral hadron contamination fell with increasing cluster p_{T}, motivated the inclusion of all clusters, even if they are hadronic in origin. Even though this is not normally done, PHENIX has previously studied[211] how well the MC simulation describes the response of the EMCal to hadrons.

The offline energy scale for the PHENIX calorimeters has been traditionally investigated by the position of the invariant mass peak of $\pi^0 \to \gamma\gamma$ decays. However, following the method of [47], in this analysis the EMCal is calibrated sector-by-sector and run-by-run by considering the energy-to-monmentum ratio (E/p) of selected electrons (a PHENIX experimental strength) associated with energetic clusters in the calorimeter. This has the advantage of anchoring the EMCal energy scale to the overall tracking momentum scale, which is well-determined and well-modeled in MC.

4.2 PHENIX Data and Trigger systems

Since the highest luminosity pp fills to date have reached BBC-firing rates as high as $\sim$ 10 MHz and the maximum PHENIX data-acquisition (daq) rate is $\sim$ 6-7 kHz, the daq cannot record every crossing and prescaling of events is required. PHENIX uses a Level 1 (LVL1) online trigger to select which events to read out in real time. For the purposes of this analysis, there are two triggers of interest: the online Minimum Bias trigger and the EMCal/RHIC Trigger (ERT), a trigger for selecting rare probes and high-p_{T} physics by triggering on clusters in the EMCal. In PHENIX convention, a scaledown s means that s events are prescaled for every one that is taken, e.g. $1/(s+1)$ of triggered events are eventually recorded.

Signals from the PHENIX subdetector systems are digitized in Front End Modules (FEMs)[212; 213], which are typically attached to each detector in the IR. A Level-1 trigger system processes information from a subset of detectors (namely, the EMCal, the ZDC's, the BBC, the RICH and

parts of the muon system) and accepts or rejects the events depending on the Level-1 triggers fired and the prescale rotation. In some subsystems, digitization occurs as soon as the analog signals are receives by the FEMs, while in other the signals are stored in analog memory units and are only digitized upon the receipt of an accept signal from the trigger. The FEMs can pipeline up to 40 beam crossings while waiting for the trigger decision to arrive. The timing is synchronized by a Master Timing Module (MTM), which communicates with the RHIC ring clocks, and Granule Timing Modules (GTMs) under its control which interact with the FEMs.

After digitization by the FEMs, all detector-specific signals have been repackaged in a uniform manner. Data is sent along fiber to the Data Collection Modules (DCMs), which are grouped into Partition Modules (PMs) and are located in the server rack room of the PHENIX Counting House (CH). The DCMs perform zero suppression and basic error and data integrity checks before sending blocks of events to the Sub-Event Buffers (SEBs), each of which buffers many event fragments with the same subdetector information. The DCM to SEB data transfer is accomplished through the use of the JSEB ("Jack's Sub-Event Buffer" card, named after the original developer of the FPGA code)[214] PCI card which uses direct memory access (DMA) to copy the event data coming in over fiber to SEB memory. The SEBs are connected through a 10-gigabit switch to the Assembly and Trigger Processors (ATPs), which assemble the full events. The Event Builder (EvB) brokers the assignment of event fragments from all the SEBs to individual ATPs. After assembly, the data is stored locally in six buffer boxes at the CH before being transferred to the RHIC Computing Facility's High-Performance Storage System (HPSS) tape.

The entire setup is controlled by the PHENIX Online Computing System (ONCS), which uses an implementation of the Common Object Request Broker Architecture (CORBA)[215] standard to communicate with the various components of the daq. The original design of the overall system in shown in Figure 4.11.

The original PHENIX design, which operated with a constraint of 20MB/s on the tape writing rate, included an additional Level-2 trigger in the ATPs which would further select interesting physics events and prescale uninteresting ones. However, a huge increase in the available data rate (writing to the buffer boxes can reach $\approx$ 1400 MB/s in 2013) and the implementation of zero-suppression and data compression methods in the ATPs themselves made the Level-2 trigger ultimately unnecessary.

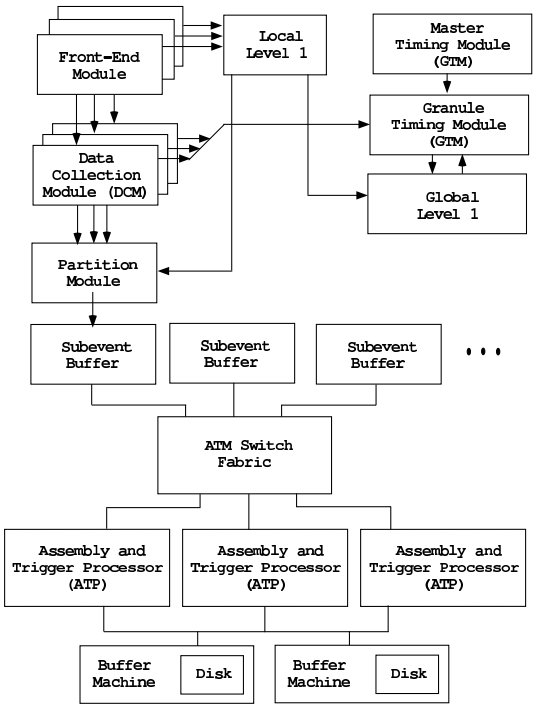


Figure 4.11: Diagram of front-end electronics, timing and event assembly systems, from [213].

4.2.1 PHENIX BBCs and the Minimum Bias Trigger

The beam-beam counter (BBC) detectors[216; 217] are located along the beamline at ±114 cm from the nominal collision vertex, subtending a full 2π in azimuth and covering $3.0 < |\eta| < 3.9$ ($2.4^\circ < \theta < 5.7^\circ$). The BBCs are used for the minimum bias event triggers in pp and d+Au collisions and to measure the position of the collision vertex. Additionally, they are one of the detectors used to characterize the event centrality in PHENIX (see Section 6.2.3). Each counter is composed of 64 quartz Ĉerenkov radiators each on top of a photomultiplier tube (PMTs), which have a dynamic range spanning 1 to ≈ 30 minimum ionizing particles (MIPs) per PMT. The tubes convert the Ĉerenkov radiation of fast ($\beta > 0.7$) charged particles to a total charge on the order of 40 pC per MIP depending on the gain. The BBC is shown in Figure 4.12.

The arrival time in each BBC is taken to be the average of the time in each fired PMT, and can be related to the collision vertex through $z_{vtx} = c(t_N - t_S)/2$, where t_N and t_S are the North and South BBC times, respectively. The timing resolution $\sigma_t \approx 50$ ps allows the determination of

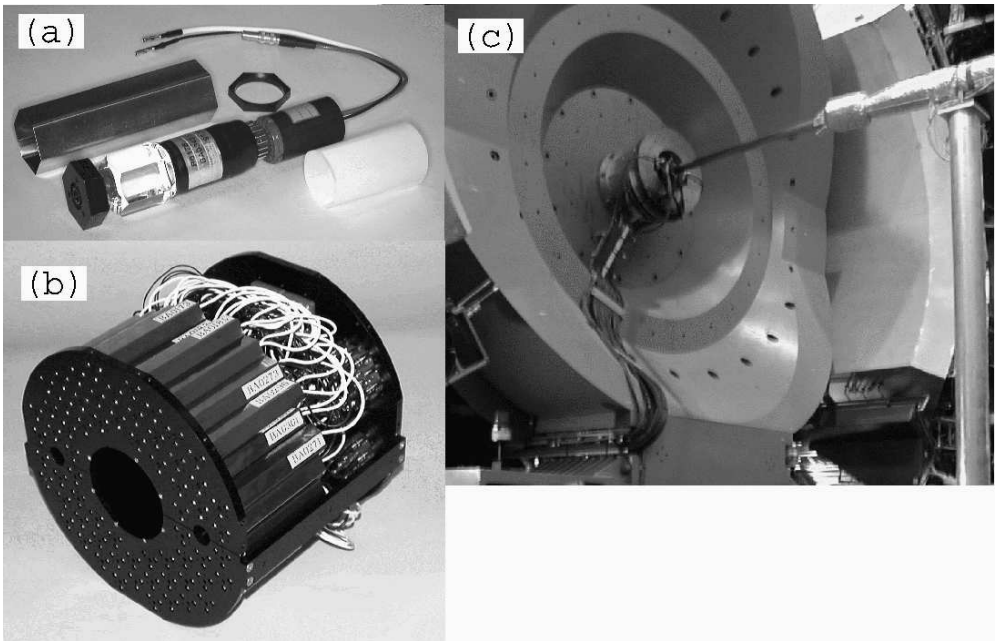


Figure 4.12: Photographs of the BBC: (a) a single BBC element, (b) fully assembled 64-tube BBC detector, (c) the BBC installed around the beamline in the PHENIX experimental hall, from [216].

the event vertex to within $\approx$ 1cm.

The online minimum bias event definition for d+Au collisions consists of the the coincident firing of each BBC ($\geq$ 1 one PMT fired in each) within a 20 ns window, corresponding to an (online) reconstructed vertex position within $|z_{vtx}| < 30$ cm of the nominal collision point. From the most recent van der Meer analysis of Run 4-6 data in PHENIX[218], the BBCLL1-firing cross-section is 24.5 mb $\pm$10% in pp collisions (which correspond to $\sim 55 - 60\%$ of the nominal though unofficial $\sigma_{inel} = 42$ mb). The BBCLL1 fires on $\approx 88\%$ of the d+Au cross-section. In both cases, an additional offline cut is imposed, requiring $|z_{vtx}| < 20$ cm.

4.2.2 EMCal/RHIC Triggers

Since the tracking reconstruction algorithms run on a timescale smaller than the trigger decision time, the main trigger for rare and high-p_{T} probes is implemented in the EMCal. The ERT is implemented by considering the energy (really the ADC sums, since it is before the final offline calibration) in 2×2 tiles of calorimeter towers, in sliding 4×4 windows of four adjacent 2×2 tiles,

and by the presence of RICH light from electrons. In line with the PHENIX design, these trigger naturally select events with high-p_{T} electromagnetic probes (in the case of the calorimeter energy sums) or the presence of heavy flavor decays (in the case of the RICH). Only the 4×4 triggers are used for the results presented here.

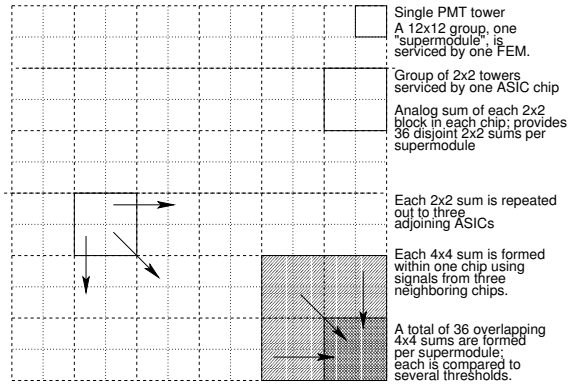


Figure 4.13: Diagram of sliding window tiles in Electromagnetic/RHIC trigger, from [207].

The analog PMT signal from each 2×2 tower group is summed by an ASIC chip and transmitted to its three nearest neighbors[207]. Logic circuitry in every sector then considers the values in any four adjacent 2×2 tiles in a 4×4 arrangement and compares them against three different thresholds, as diagrammed in Figure 4.13. In order of increasing threshold in the d+Au running these triggers are called the $4 \times 4c$ ($\gtrsim$ 2.1 GeV in the PbSc, $\gtrsim$ 1.6 GeV in the PbGl), $4 \times 4a$ ($\gtrsim$ 2.8 GeV in the PbSc, $\gtrsim$ 2.1 GeV in the PbGl) and $4 \times 4b$ ($\gtrsim$ 3.5 GeV in the PbSc, $\gtrsim$ 2.8 GeV in the PbGl). The ASIC logic connects them across SM, but not sector, boundaries. The other two triggers consider just the energy in 2×2 windows and a minimum number of photoelectrons in a sliding window of adjacent RICH PMTs, respectively, but they are not used in this work.

In the pp running, the $4 \times 4c$ was turned off in the PbGl and the thresholds in the other two were systematically decreased: thus, only the $4 \times 4a$ ($\gtrsim$ 2.1 GeV in the PbSc, $\gtrsim$ 1.6 GeV in the PbGl, corresponding to the $4 \times 4c$ in d+Au running) and the $4 \times 4b$ ($\gtrsim$ 2.8 GeV in the PbSc, $\gtrsim$ 2.1 GeV in the PbGl, corresponding to the $4 \times 4b$ in d+Au running) triggers were used. As a rule, an event that fired a trigger with a higher threshold always fires the lower-threshold triggers as well, although due to differences in prescales, an event may only have the live and not the scaled trigger

bits set for the other triggers.

The $4 \times 4c$ and $4 \times 4a$ triggers required the coincidence of a BBCLL1 trigger to fire, while the $4 \times 4b$ did not. This makes the $4 \times 4b$ the only trigger which can independently check the bias factor (the enhancement above minimum bias trigger efficiency that a pp event with a high-p_{T} jet will fire the BBCLL1 trigger).

Chapter 5

The ATLAS detector

Draw thy sword,
That, if my speech offend a noble heart,
Thy arm may do thee justice: here is mine.

King Lear, Act V, Scene III

The ATLAS detector[219] is one of two large, general purpose detectors at the LHC. Design requirements for the ATLAS detector are strongly driven by the desire to measure new TeV-scale physics signatures, such as the Higgs boson, dark matter, light supersymmetric particles, excitations from unification theories and other exotics. These signatures require the ability to reconstruct photons, jets, all leptons, displaced vertices and even large missing transverse energy. To this end, ATLAS has been built with several key principles in mind, including a large acceptance in pseudorapidity and full azimuthal coverage, good charged particle momentum resolution and secondary vertex identification, precise and segmented electromagnetic calorimetry (for photons and electrons), hermetic hadronic calorimetry (for jets), muon capabilities and efficient triggering, all in a high-luminosity, high-radiation environment. This chapter describes the major features of the ATLAS detector used in this analysis.

ATLAS uses a right-handed coordinate system centered at the nominal interaction point. $\hat{z}$ is coincident with the beam axis, with $+\hat{z}$ pointing towards the "A" side of the detector and $-\hat{z}$ towards the "C" side. The $+\hat{x}$ axis points towards the center of the LHC ring, which is roughly

North-East, and the $+\hat{y}$ axis points skyward. Thus the $+\hat{z}$ direction points counterclockwise when the ring is seen from above. The pseudorapidity η is defined $\eta = -\log\tan(\theta/2)$ where θ is the polar angle in the coordinate system. For a particle with mass m and transverse momentum p_{T}, this is related to the rapidity y through

$$y = \log \frac{\sqrt{m^2 + {p_{\mathrm{T}}}^2 \cosh^2 \eta} + p_{\mathrm{T}} \sinh \eta}{\sqrt{m^2 + {p_{\mathrm{T}}}^2}} \tag{5.1}$$

which is $\approx \eta$ in the limit $p_{\mathrm{T}} \gg m$.

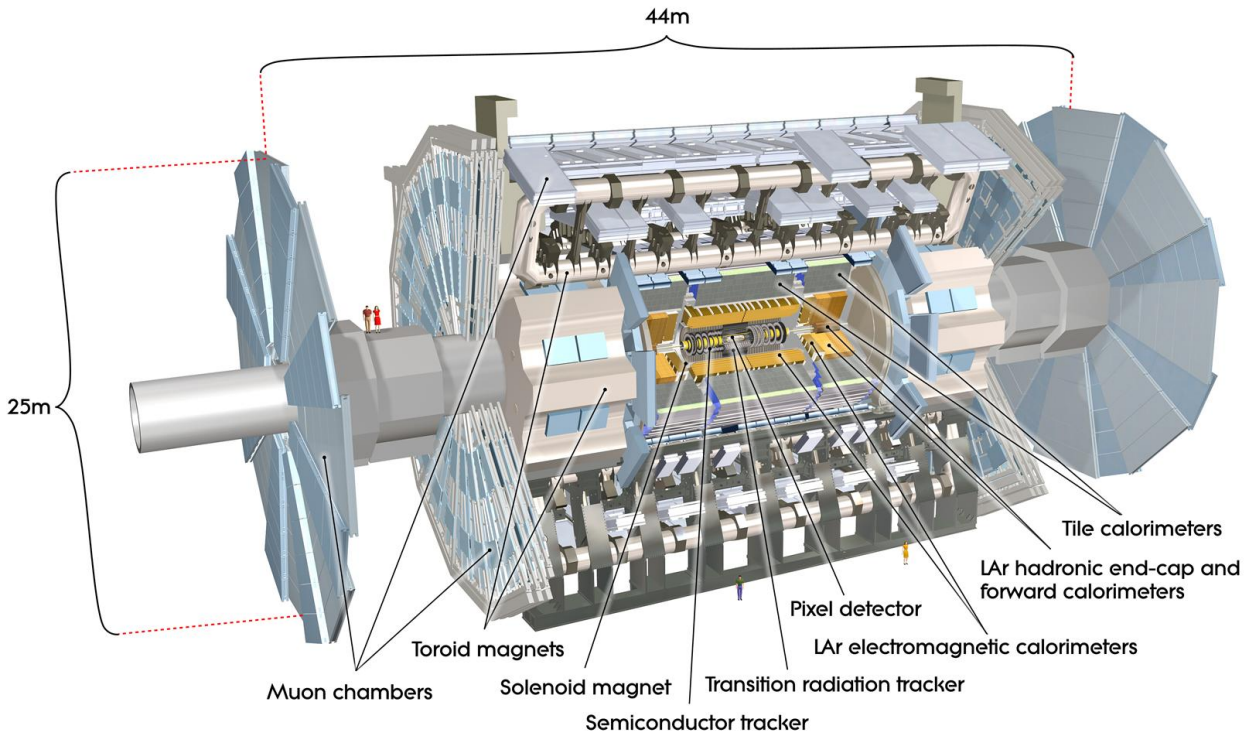


Figure 5.1: Cut-away view of the ATLAS detector indicating some of the major detector systems, from [219].

ATLAS is nominally forward-backward symmetric. Figure 5.1 shows the position of some major ATLAS detectors and magnets. The various muon-specific detectors, instrumentation and triggers, while critical for the overall physics goals of the experiment, are not used in this work.

5.1 ATLAS Central Solenoid

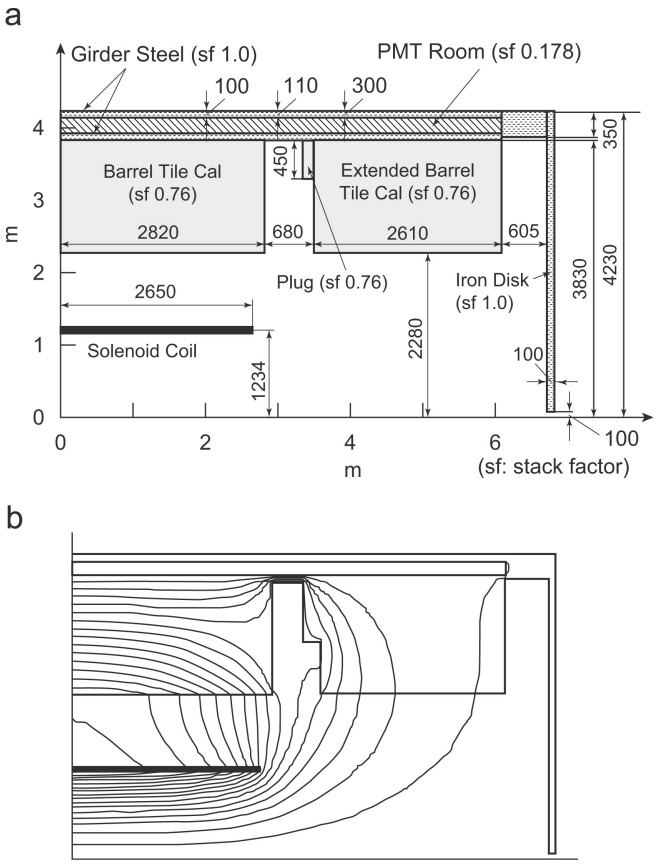


Figure 5.2: Diagram of ATLAS solenoid. a) Solenoid position in $\hat{r}$-$\hat{z}$ plane along with the iron of the hadron calorimeter. b) Magnetic field lines. Figure from [220].

The ATLAS magnet system consists of a central solenoid (CS)[220] which provides the magnetic field for determining the momentum of primary vertex particles, and three toirodal magnets for the muon systems. The CS produces an axial magnetic field of 2 T throughout the inner detector (ID) tracking volume, which is shown in Figure 5.2. Some of the major design criteria included an axially symmetric, reversible field, minimum total radiation thickness, integration into the detector between the ID and EM calorimeter with minimal additional support or vaccuum structures, and stable, maintenance-free operation for up to nine months out of the year, when the solenoid is generally inaccessible inside the ATLAS detector. The superconducting cable consists of 12 strands of niobium-titanium (NbTi) filaments in a copper (Cu) matrix, housed in a pure aluminum (Al)

sheath, and residing in a high-strength aluminum alloy cylindrical coil.

The solenoid was installed just within the EM calorimeter volume, directly in front of its first layer and with a 2 mm thin heat shield separating it from the cryostat wall. It extends between 1.23 m $< r <$ 1.28 m radially and $z \in \pm 2.9$ m along the axis. Having the two detector elements share a cryostat significantly reduces the extra material that would be needed to form two additional vacuum walls (one to close the solenoid cryostat, and one to enclose the EMCal). A further choice was made to limit the coil thickness (despite the consequently large stored energy-to-mass ratio E/m) to as small as is reasonably safe. Thus, the CS material thickness corresponds to a radiation lenth of $0.64X_0$ and nuclear interaction length of $0.13\lambda_I$. Behind the EMCal, the iron in the Tile calorimeter and the Tile support structure serve as the return yoke for the magnetic field. Figure 5.2 shows the position of the solenoid and iron yoke elements, as well as the magnetic field lines. The CS is near to an "ideal" solenoid, which is defined as one with constant double field integral

$$I_2 = \int_{r=0}^{r=L\sin\theta} \int_{l=0}^{l=r/\sin\theta} \left| \vec{B} \times \vec{dl} \right| dr \tag{5.2}$$

over the pseudorapidity range it covers. For the ATLAS CS, deviations of I_2 at the 5-10% level only appear near the edges at $\eta = \pm 1.6$ and are well understood. The solenoid is cooled from room temperature to 100 K using helium gas, and from 100 K down to the operating temperature of 4.5 K using supercritical helium from the main refridgerator (MR) of the cryo system.

In 2006, the solenoid field was mapped with a Hall probe array[221] with several different values of the current. The solenoid and toroidal magnets were installed in their final positions during the mapping. 20,000 data points were taken which allowed, along with a detailed simulation of the solenoid elements, a good fit of the field strength to a functional form that obey's Maxwell's equations. The fit residuals have an RMS of less than < 0.5 mT. The final systematic uncertainty on the tracking momentum scale from the uncertainty in the magnetic field integral is 0.02% to 0.12% depending on the track rapidity.

5.2 ATLAS Global and Forward Detectors

5.2.1 Minimum Bias Trigger Scintillator

The Minimum Bias Trigger Scintillator (MBTS)[222] detector is used to trigger on minimum bias pp and p+Pb collisions with high efficiency while keeping the ability to reject out of time beam background. The MBTS is a pair of 2 cm thick octagonal polystyrene scintillators positioned just in front of the electromagnetic endcap (see Section 5.3 below) at $z = \pm 3.56$ m. The detector on each side consists of an inner and outer layer of 8 trapezoidal wedges each covering $\Delta\phi = 2\pi/8$. The innermost layer is defined by 15.3 cm $< r <$ 42.6 cm ($3.84 < |\eta| < 2.82$) and the outermost by 42.6 cm $< r <$ 89.0 cm ($2.82 < |\eta| < 2.09$).

Light is collected from each wedge by wavelength-shifting (WLS) optical fibers and delivered to a photomultiplier tube (PMT). The PMT signal is shaped and amplified by repurposed TileCal electronics, which applies a leading edge discriminator and sends a pulse to the Central Trigger Processor (CTP, see Section 5.5). The MBTS multiplicity is calculated for each side independently (saturating at 16, when each of the 16 wedges in the given side have a hit above discriminator threshold) and used to create L1 triggers of the form `MBTS_N`, which requires a total of N hits split either way between the detector sides, or `MBTS_N_N`, which requires N hits in each. The MBTS operating voltage has changed over time as the understanding of the detector performance improves, with the recent PMT high voltage of 850 V chosen so that even a single minimum ionizing particle (MIP) signal is accepted with good efficiency. The efficiency of the various MBTS triggers has been studied with respect to the track multiplicity and other indicators of a real pp collision[223].

The individual wedges are timing capable, and each side of the MBTS reports a time t_A or t_C which is the average over the wedges that have a hit of the time relative to the LHC clock. The time difference $(t_A - t_C)$ is used as an offline cut to reject out of time background not connected to a collision at the nominal interaction point.

5.2.2 Luminosity detectors

Many methods have been proposed to monitor the online luminosity and return an integrated offline luminosity[176], including specialized detectors in the forward region, counting hits or reconstructed vertices with the inner detector, and even using the production of Z bosons as a standard luminosity

candle. In this section, we describe the two more common detectors used for this purposes in ATLAS and the ones for which a luminosity calibration exists for the 2.76 TeV *pp* reference data.

The LUminosity measurement using Ĉerenkov Integrating Detector (LUCID)[224] is an ATLAS detector designed for luminosity measurements. This design was chosen in part because the hard radiation environment in the forward region prohibits the use of scintillator counters that have been used for this purpose in previous generations of detectors. Other design requirements include the need for good acceptance for minimum bias *pp* collisions and sufficiently good time resolution to resolve one bunch crossing from another (25 ns).

Each LUCID detector is a cylinder of sixteen optically reflecting aluminum tubes filled with C_4F_{10} gas at overpressure (resulting in a Ĉerenkov threshold of 10 MeV for electrons and 2.8 GeV for pions). The detectors are situated a radial distance of 10 cm away from the beamline, at a distance of approximately $z = \pm 17$ m, with pseudorapidity coverage within $5.6 < |\eta| < 5.8$. Ĉerenkov light from fast charged particles reflects along the walls of the tubes and is collected by a PMT at the other end, which amplifies the signal and subjects it to a constant fraction discriminator.

The sixteen signals from each detector are processed by a custom-built electronics card (LUminosity Monitor of the ATLAS experiment, or LUMAT). For each crossing LUMAT records the number of phototubes on either side which fired. Currently, the luminosity is determined separately by examining two logical conditions on the number of hits: `LucidEvt_OR`, which requires ≥ 1 hit on either side, and `LucidEvt_AND`, which requires ≥ 1 hits on both. LUMAT then sends information to the CTP, if necessary, about the presence of L1 trigger items (see Section 5.5) and is used to monitor the instantaneous luminosity. Finally, the two LUCID event algorithms are used in the offline luminosity determination.

The primary purpose of the ATLAS Beam Condition Monitors (BCM)[225] is to detect beam anomalies with the potential to cause serious detector damage and provide a fast abort signal to mitigate this. However, they are also used as luminosity detectors. The BCM consists of four diamond sensors on each side of the interaction point, positioned at $r = 5.5$ cm and $z = \pm 1.84$ m (corresponding to $|\eta| = 4.2$) and arranged equidistantly in azimuth. Charged particles ionize the 1×1 cm^2 polycrystalline Chemical Vapour Deposited (CVD) diamond sensors when passing through them, and the resulting charge is collected by an RF amplifier and read out. As with other detector elements situated right by the beam pipe, the BCMs are designed to be very radiation

hard. Furthermore, the BCM detectors have a timing resolution of $\approx$ 0.7 ns, and can resolve coincident hits as coming from an event at the interaction point or from beam backgrounds. The BCM luminosity algorithms used are `BCMHEvt_Or` and `BCMVEvt_Or`, which are defined as one hit on either side in the horizontal or vertical pairs of monitors, respectively.

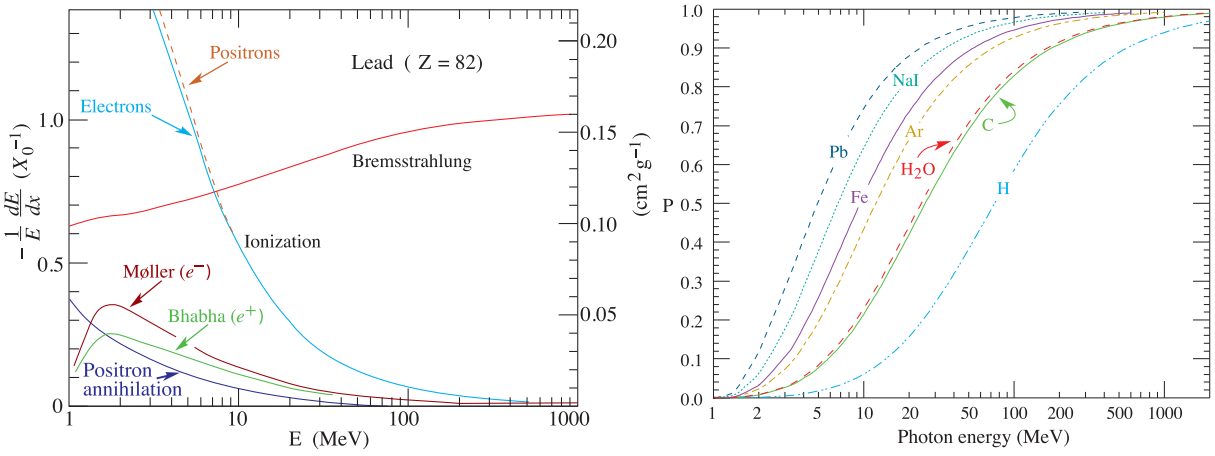
5.3 ATLAS Calorimeter

The ATLAS calorimeter system contains electromagnetic (in the form of the Electromagnetic Barrel and Electromagnetic End-cap detectors) and hadronic (consisting of the Hadronic End-Cap, the Tile Barrel and Extension and Forward Calorimeter detectors) over the range $|\eta| < 4.9$, covering nearly ten units of pseudorapidity. These can be seen in Figure 5.1.

The ATLAS calorimeter system measures the energy of all produced particles except muons (which leave minimum ionizing radiation in the detector until they are detected by the muon system) and neutrinos (which can only be detected indirectly through a transverse energy imbalance). A major design feature of the calorimeter system is that (with the exception of the forward calorimeter) it is *projective*, in that the detector elements fit within planes of constant ϕ and η from the interaction point. Furthermore, the calorimeter is designed to provide fine granularity at mid-rapidity for precision measurements of electrons and photons and sufficiently fine granularity over the entire pseudorapidity range for jet reconstruction. It must contain hadronic showers to limit the rate of punch-through hadrons being detected by the muon system and also to be able to properly reconstruct any missing energy in the event (an important signal in some beyond the Standard Model processes).

Two important characteristics of calorimeters are the *radiation length* X_0 and the *nuclear interaction length* λ_I. Both are material dependent.

Consider electrons and photons travelling through dense matter[1]. High-energy electrons lose energy primarily through Bremmstrahlung radiation (which rises linearly with the energy of the electron), while low-energy electrons lose energy primarily through ionization (which rises logarithmically with energy and dominates in the low-energy regime) and Møeller scattering. The point below which the energy loss is dominated by ionization is called the *critical energy* E_c, and is = 7.43 MeV in a lead (Pb) medium. This turning point is shown on the left side of Figure 5.3. X_0 is the



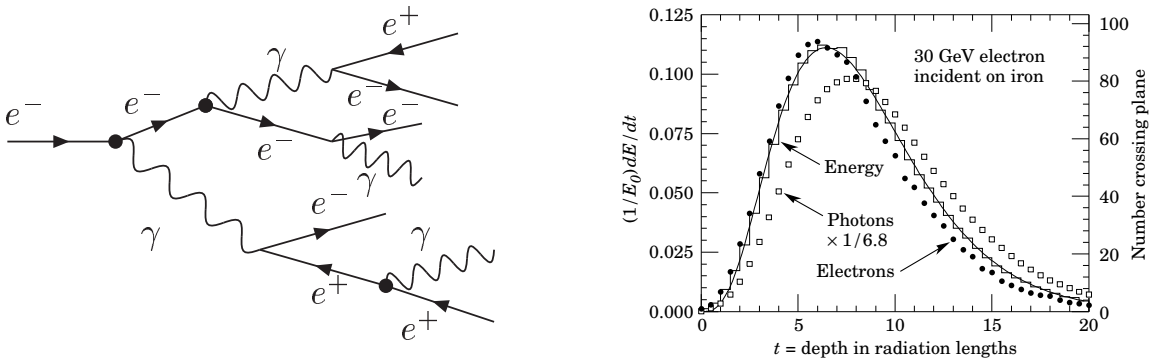
(a) (b)

Figure 5.3: Relative rates of energy loss for electrons and photons passing through matter, from [1]. a) Fractional energy loss per radiation length in lead as a function of electron energy. Curves are shown for different processes. b) Probability that a photon interaction will result in pair-production as a function of photon energy. Curves are shown for different materials.

characteristic length over which the energy loss, and is defined as the mean distance travelled by a high-energy electron before its energy has been reduced to a fraction $1/e$ of the original energy. X_0 increases with the A of the material and decreases (that is, more energy is lost per unit length) with increasing Z.

On the other hand, high-energy photons predominantly undergo e^+e^- pair production when they interact with material. The right side of Figure 5.3 shows that even at energies as low as $E \approx 5$ MeV pair-production in Pb is still as likely as any other processes (such as the photoelectric effect, Compton scattering and Rayleigh scattering). The characteristic length for pair-production is $^7/_9 X_0$.

A high-energy electron or photon incident on a thick target will initiate an electromagnetic *shower* or *cascade*, with the electrons producing bremmstrahlung radiation (photons) and photons undergoing pair production (electrons and positrons), and the products of those processes undergoing their own bremmstrahlung and pair production processes, etc. Thus, a single high-energy electron or photon eventually ends up as a large number of low-energy electrons and photons, called



(a) (b)

Figure 5.4: a) Developing electromagnetic shower. b) Longitudinal profile of electromagnetic shower from 30 GeV e^- and γ on an iron target, from [1].

a shower. This is represented symbolically on the left side of Figure 5.4.

The shower develops with a characteristic longitudinal and transverse scale. Longitudinally, the mean profile of energy deposition can be modeled as a Gamma distribution via,

$$\frac{1}{E_0}\frac{dE}{d(x/X_0)} = b\frac{(bx/X_0)^{a-1}e^{-bx/X_0}}{\Gamma(a)} \tag{5.3}$$

with $a \approx 0.5$ and $b \approx 1 \pm 1 + 2\log(E_0/E_C)$, where the sign is + for photons and − for electrons. Examples of the longitudinal profile of 30 GeV electron and photon showers are shown on the right side of Figure 5.4. In this example, an electromagnetic shower may need > 15 radiation lengths to contain most of its energy.

The transverse profile of an electromagnetic shower is characterized by its Molière radius R_M, which is defined as the value for which a cylinder with radius R_M contains 90% of the shower energy. The Molière radius is given by

$$R_M = X_0\frac{\sqrt{4\pi/\alpha_{QED}}m_e c^2}{E_C} = X_0\frac{21.2 \text{ MeV}}{E_C} \tag{5.4}$$

For example, $R_M \approx 3X_0$ in lead.

Unfortunately, the *hadronic showers* aused by the passage of high-energy pions, long-lived kaons, protons, neutrons, and so forth passing through dense material are somewhat more complicated.

The energy loss through bremmstrahlang radiation is strongly suppressed by the large (relative to the $e^{\pm}$) hadron mass. Generally, high-energy hadrons undergo inelastic collisions with nuclei in the material, producing many light hadrons (mostly pions) and/or knocking out nucleons. Of the pions, the neutral π^0's decay electromagnetically and initiate their own electromagnetic showers. Some of the energy may be invisible entirely if the products are captured by nuclei, or if there is significant recoil of the nuclear target. In general, there are large fluctuations both in the electromagnetic component fraction of hadronic showers and also in the measurable shower energy. Furthermore, hadronic showers have a somewhat broader transverse development than electromagnetic ones.

In any case, the characteristic longitudinal scale for a hadron losing energy is the nuclear interaction length λ_I, which is the distance at which the hadron energy has dropped to $1/e$ of the original. λ_I scales with the A of the material ($\propto A^{1/3}$) and is independent of Z. Typically, it is much larger than the radiation length in the same material – in Pb, $\lambda_I \approx 30X_0$. For this reason, the ATLAS electromagnetic calorimetry is placed before the hadronic calorimetry. The EM calorimeters precisely capture most of the electromagnetic shower, leaving the hadronic calorimeters to measure the rest of the hadronic energy.

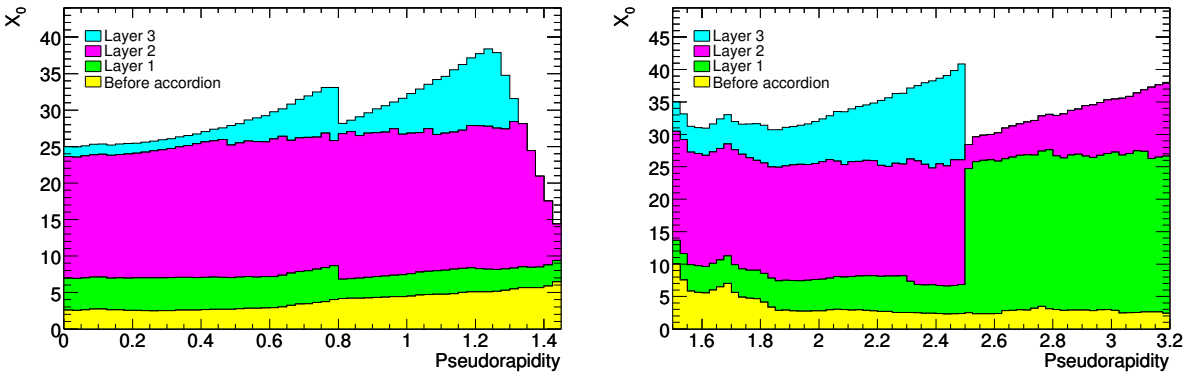


Figure 5.5: Cumulative number of radiation lengths as a function of pseudorapidity, in units of X_0, for the ATLAS calorimeter system, from [219].

The total integrated X_0 and λ_I for the ATLAS calorimeters (and material before and after them) are shown in Figures 5.5 and 5.6 as a function of η. The segmentation of the detector into projective, longitudinal (η, ϕ) cells is summarized in Table 5.1.

Layer	η Coverage	$\Delta\eta \times \Delta\phi$	Absorber/Medium
EM Barrel			
Presampler (EMB0)	$\|\eta\| < 1.52$	0.025×0.1	LAr / -
EMB1	$\|\eta\| < 1.475$	0.003125×0.025	LAr/Lead
EMB2		0.025×0.025	
EMB3		0.05×0.025	
EM End-Cap			
Presampler (EME0)	$1.5 < \|\eta\| < 1.8$	0.025×0.1	LAr / -
EME1	$1.375 < \|\eta\| < 3.2$	$0.03125\text{-}0.05 \times 0.025$ ($\|\eta\| < 2.5$) 0.1×0.1 ($\|\eta\| > 2.5$)	LAr/Lead
EME2		$0.025\text{-}0.05 \times 0.025$ ($\|\eta\| < 2.5$) 0.1×0.1 ($\|\eta\| > 2.5$)	
EME3	$1.5 < \|\eta\| < 2.5$	0.05×0.025	
Tile Barrel			
TileBar1, 2	$\|\eta\| < 1.0$	0.1×0.1	Polysterene/Steel
TileBar3		0.2×0.1	
Tile Extension			
TileExt1, 2	$0.8 < \|\eta\| < 1.7$	0.1×0.1	Polysterene/Steel
TileExt3		0.2×0.1	
Hadronic End-Cap			
HEC0, 1, 2, 3	$1.5 < \|\eta\| < 3.2$	0.1×0.1 ($\|\eta\| < 2.5$)	LAr/Copper
		0.2×0.2 ($\|\eta\| > 2.5$)	
FCal			
FCal1	$3.1 < \|\eta\| < 4.9$	$\approx 0.2 \times 0.2$ (non-projective)	LAr/Copper
FCal2, 3			LAr/Tungsten

Table 5.1: Summary of ATLAS electromagnetic and hadronic detector elements.

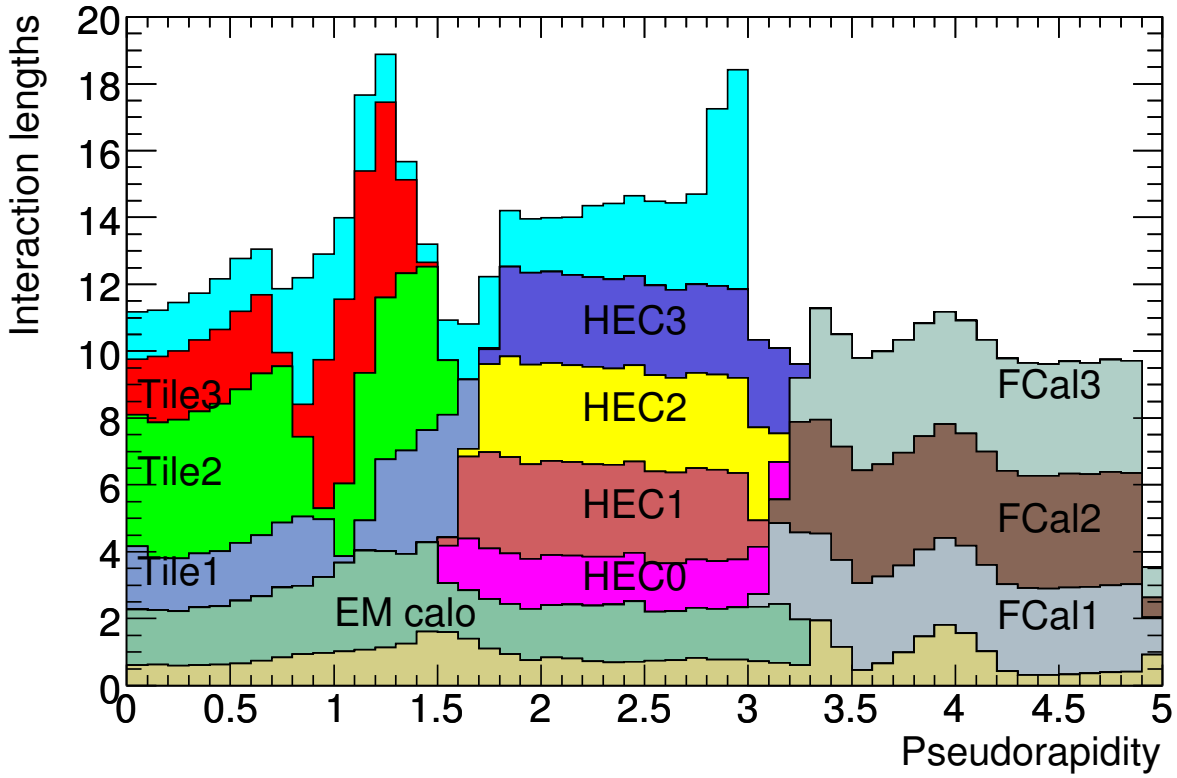


Figure 5.6: Cumulative number of nuclear interaction lengths as a function of pseudorapidity, in units of λ_I, for the ATLAS calorimeter system, from [219].

5.3.1 Electromagnetic calorimetry

All ATLAS calorimeters are *sampling* calorimeters, which alternate absorber and an active material. High-energy electromagnetic particles ($e^{\pm}$ and γ) and hadrons initiate showers through interactions in the absorber material which are then measured through ionized charge or scintillation light in the active material. Absorber and sampling regions are alternated for enough interaction lengths that a large fraction of the energy in most showers is sampled. With the exception of the hadronic Tile calorimeter discussed later, all ATLAS calorimeters use liquid Argon (LAr) technology as the active medium. The LAr detectors are contained in three separate cryostats, one for the barrel and two for each end-cap, kept at 80 K (the melting point of Argon is $\approx$ 84 K) with nitrogen refrigeration.

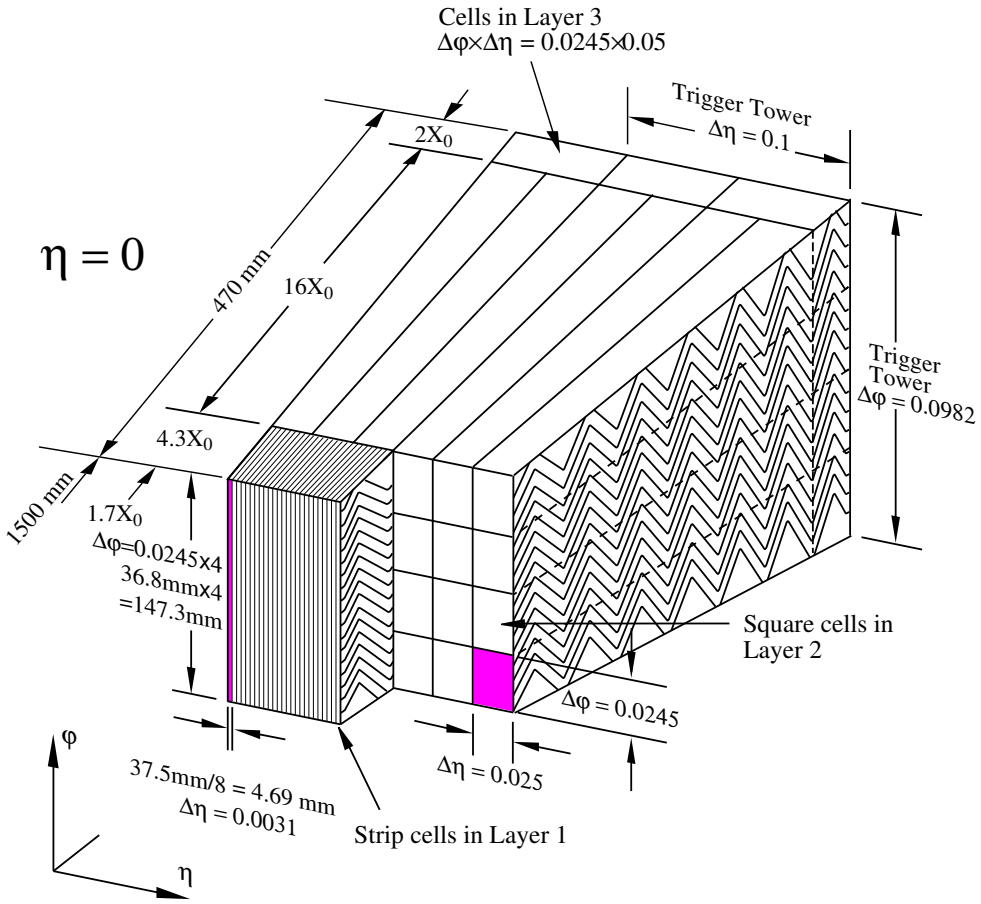


Figure 5.7: Segmentation of readout cells in an EM Barrel module, [219].

The EM calorimeter[226; 227] consists of a cylindrical barrel detector centered at $\eta = 0$ and two end-cap detectors on either side of it, which can be seen in Figure 5.1. The EM Barrel (EMB) extends radially between 1.25 m $< r <$ 2.05 m and along the beam axis according to $|z| <$ 3.1 m ($|\eta| < 1.475$). The EMB is situated in a cryostat which also contains the ATLAS central solenoid and is separated into two half-barrels symmetric around $\eta = 0$. The detector is divided into three radial layers. The first layer has very fine pseudorapidity segmentation ($\Delta\eta \times \Delta\phi = 0.003125 \times 0.1$) to measure the developing electromagnetic shower in detail over a radiation length of $4.3X_0$. The second layer is coarser in pseudorapidity but finer in azimuth ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) and captures a large fraction of the remainder of the shower over its $16X_0$ radiation lengths. The third layer captures the remainder of the energy (1-10X_0 depending on η) with coarser segmentation

($\Delta\eta \times \Delta\phi = 0.05 \times 0.025$). The layer-by-layer segmentation in the barrel is shown in Figure 5.7.

The EM End-caps (EME) are at the same radial position but are centered at $z = \pm 3.7$ m, covering a region $1.375 < |\eta| < 3.2$. The EME is divided into two coaxial endcaps called wheels. The EME is also split into three longitudinal layers, the third of which only exists from $1.5 < |\eta| < 2.5$. The segmentation follows the same general pattern (fine segmentation in η at the expense of coarse segmentation in ϕ in the first layer, $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the second layer, etc.) but the situation is more complicated than in the EMB. Table 5.1 lists the details of the η- and layer-dependent segmentation in the End-Cap.

The fundamental geometry of the EM calorimeters is an azimuthally stacked series of accordian-shaped lead absorber plates interleaved with readout electrodes. The accordion waves zig-zag in the $\hat{r}$-$\hat{\phi}$ plane and the folding angle decreases with radial depth. This choice of geometry makes the amount of material seen uniform in azimuth and provides a finely segmented readout. The absorber plate thickness ranges in the Barrel from 1.53 mm in the $|\eta| < 0.8$ part to 1.13 mm for $|\eta| > 0.8$ (to keep the sampling fraction from rising), and in the end-cap from 1.7 mm in the region $|\eta| < 2.5$ to 2.2 mm beyond this. Each half-barrel and each end-cap consist of 1024 absorbers each.

The electrodes[228] are three layer Cu-polyimide printer circuit boards, with the outer layers held at high potential and the inner layer serving as the readout. In the barrel, the LAr gap is a constant 2.1 mm (really, this is width of the half-gap between the inner layer and each outer layer) and the potential is +2000 V. In the End-Cap, the gap size and potential increase with radius to keep the response constant. The electrodes are actually quite large ($\sim$ 1-2 m^2 before accordion folding), with the barrel and end-cap each covered by two types of electrodes. They are etched into readout cells in a way projective with η, and result in the final granularity listed in Table 5.1. The projective segmentation of the four electrode types (before folding) can be seen in Figure 5.8. For example, in the top-left electrode which covers $|\eta| < 0.8$, the different $\Delta\eta$ segmentation of the three layers can be clearly seen.

Finally, the EM system is instrumented with an EM Presampler situated just in front of the EM detector within $|\eta| < 1.8$ (covering all of the Barrel and part of the End-Caps). The Presampler is a 1.1 cm (in the Barrel) and 0.5 cm (in the End-Cap) layer of LAr and helps correct for losses from interactions upstream of the EM calorimeter.

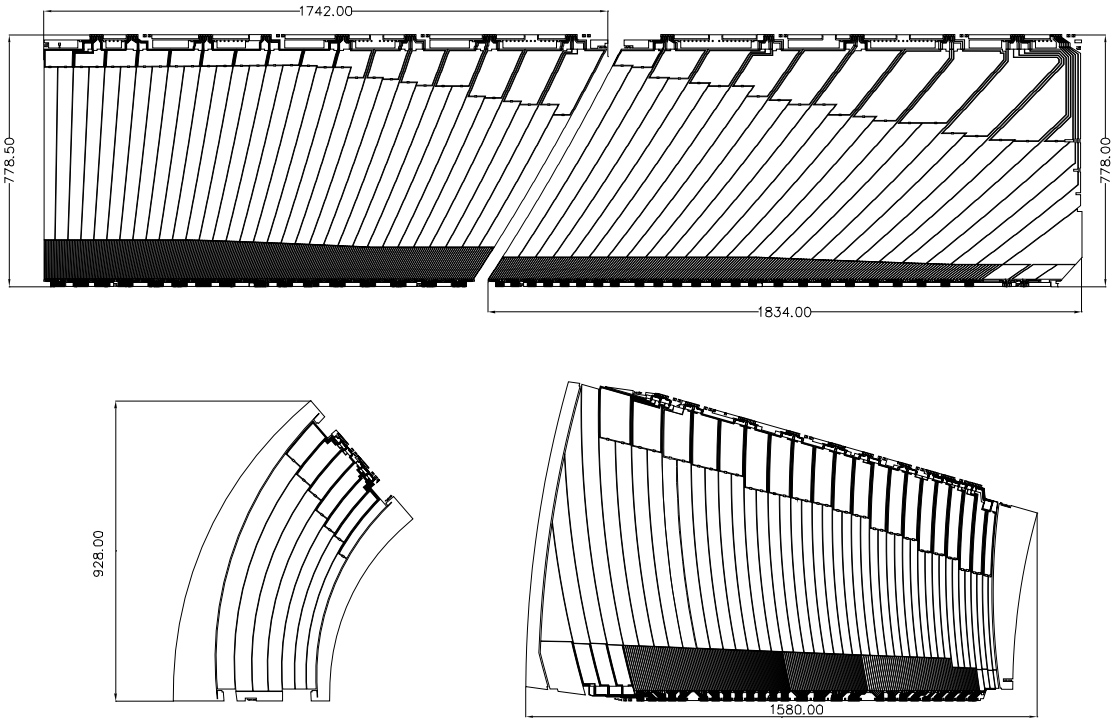


Figure 5.8: Segmentation of readout cells in an EM Barrel module, from [219]. Left to right, top to bottom, these are Type A (Barrel $|\eta| < 0.8$), Type B (Barrel, $0.8 < |\eta| < 1.475$), Type C (end-cap inner wheel, $2.5 < \eta < 3.2$) and Type D (end-cap outer wheel, $1.375 < \eta < 2.5$) electodes.

5.3.2 Hadronic calorimetry

ATLAS is also instrumented with hadronic calorimetry. Starting from $\eta = 0$ and going out to $\eta = 4.9$, the hadronic calorimeter system consists of the tile barrel and extention detectors, which complement the barrel electromagnetic calorimeter (EMB), the hadronic end-cap calorimeter (HEC), which complements the electromagnetic end-cap calorimeter (EME), and the forward calorimeter (FCal), of which the first and last two layers function as an electromagnetic calorimeter and hadronic calorimeter, respectively.

The Tile calorimeter[229], which complements the EM Barrel at mid-rapidity with hadronic calorimeter capability, and consists of Barrel and Extended elements. The Tile detectors are the only component of the ATLAS calorimeter system which is not LAr based. Both detectors are sampling calorimeters with alternating layers of steel absorber and polysterene plastic as the scin-

tillating medium. The Tile Barrel is situated directly outside the EM barrel cryostat envelope between $r = 2.28$ m to 4.25 m and extending symmetrically across $\eta = 0$ with $|z| < 2.9$ m ($|\eta| < 1.0$). The Extended Tile barrels sit at the same radial position, but are separated from the Tile Barrel by a 60 cm gap and is situated between $2.9 < z < 5.5$ m ($0.8 < |\eta| < 1.7$).

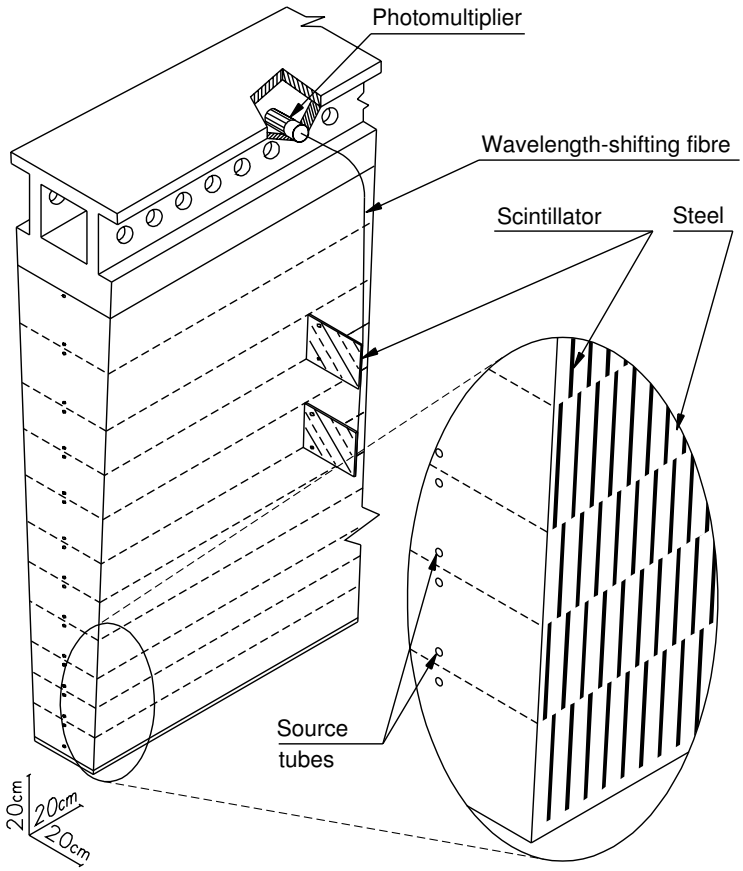


Figure 5.9: Tile calorimeter module showing the radial layers, unique orientation of scintillator and absorber tiles, and readout, from [219].

The Tile detectors consist of 64 azimuthally-segmented segments, each composed of 11 radially stacked layers of alternating tile and steel. One of the modules is shown in Figure 5.9. Unlike in other sampling calorimeters, the tiles are oriented so that their largest dimensions lie in the $\hat{r}$-$\hat{\phi}$ plane rather than the $\hat{z}$-$\hat{\phi}$ plane. This design allows for an almost seamless azimuthal coverage, with just 1.5 mm azimuthal gaps between the segments. The polysterene scintillator tiles are 3 mm thick (in the $\hat{z}$ direction) and vary in their radial and azimuthal extent from $\Delta r \times r\Delta\phi = 97 \times 200$

nm^2 to 187×400 nm^2 depending on which layer they are in. The absorber tiles are 5 mm thick steel plates with 4 mm spacer plates on each side. Overall, there are 460,000 individual tiles in the barrel and two extensions.

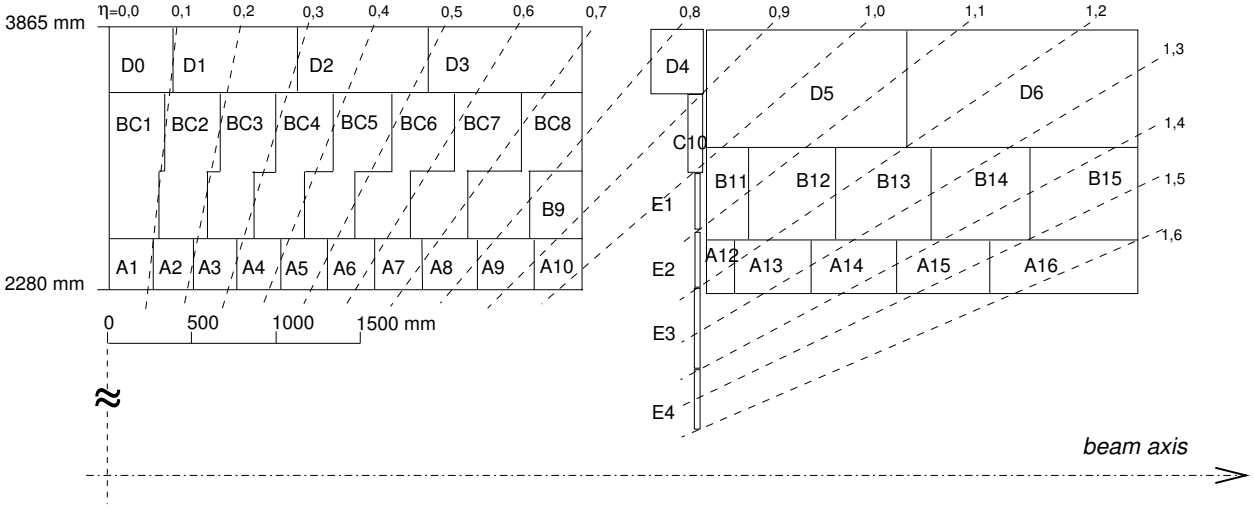


Figure 5.10: Schematic of the segmented tile-calorimeter modules, along with the trajectories of particles with varying η, from [219].

Readout fibers from the tiles are grouped in a way that forms projective detector elements called modules. Figure 5.10 illustrates the location of the Tile modules. The modules define longitudinal layers with nuclear interaction lengths $1.5\lambda_I$, $4.1\lambda_I$ and $1.8\lambda_I$ in the closest to farthest radial layers, respectively. The first two layers have modules with a granularity of $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$, with the pseudorapidity granularity rising to $\Delta\eta = 0.2$ in the last layer. The tile extension has slightly different longitudinal segmentation ($1.5\lambda_I$, $2.6\lambda_I$ and $3.3\lambda_I$ in the three layers). Readout PMTs and other electronics are contained in a steel girder at the far radial end of the detector, which is also used as the flux return for the ATLAS solenoid (see Section 5.1).

As can be seen in Figure 5.6, the total nuclear thickness in the Tile-instrumented region, including the EM barrel and endcaps, is $9.7\lambda_I$ at $\eta = 0$, falling to $\approx 8\lambda_I$ at $\eta = \pm 0.8$ where coverage is weakest as the hadrons pass from the barrel to the extended barrel and rising to $\approx 17\lambda_I$ at $\eta = \pm 1.3$, when hadrons traverse the diagonal of the extended barrel. These geometric considerations are easier to visualize with the aid of Figure 5.10, which shows the detector elements traversed by high-p_{T} hadrons as a function of η.

In addition, three sets of calorimeters at $z \approx \pm 2.9$ m provide instrumentation in the region between the EM/Tile barrel and the End-cap cryostat/Extended tile detectors. They are read out either independently or as part of the Tile electronics. In Figure 5.10, these are the modules denoted D4+C10 (the plug tile calorimeter, which is a reduced section of a standard Tile module), E1+E2 (the "gap scintillators", which cover $1.0 < |\eta| < 1.2$) and E3+E4 (the "cryostat scintillators", which cover $1.2 < |\eta| < 1.6$).

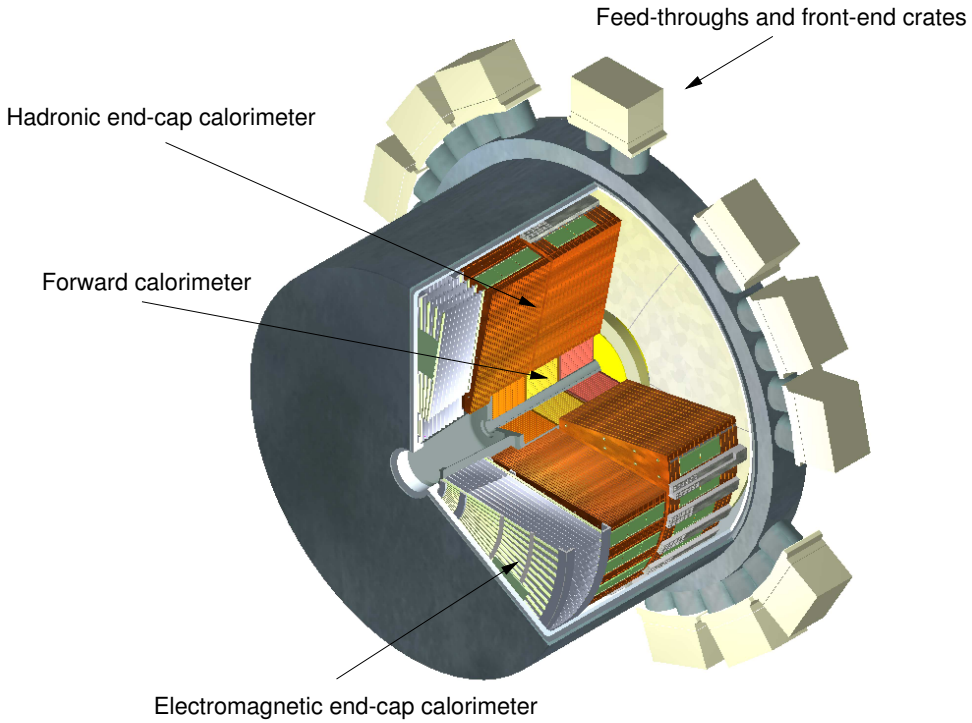


Figure 5.11: Cut-away view of the end-cap cryostat, showing the orientation of the Electromagnetic end-cap, Hadronic end-cap and Forward calorimeters, from [219].

The hadronic end-cap (HEC) calorimeters sit directly behind the EME calorimeters and within the same cryostats, which are also shared with the FCals, as seen in Figure 5.11. Each HEC detector is segmented into two wheels, which are further segmented into two longitudinal sections. Radially in to out, they are referred to as the HEC0, HEC1, HEC2 and HEC3 layers and their contribution to the total nuclear interaction length is shown in Figure 5.6. The HEC covers the region $1.5 < |\eta| < 3.2$ (or 4.3 m $< |z| <$ 6.1 m).

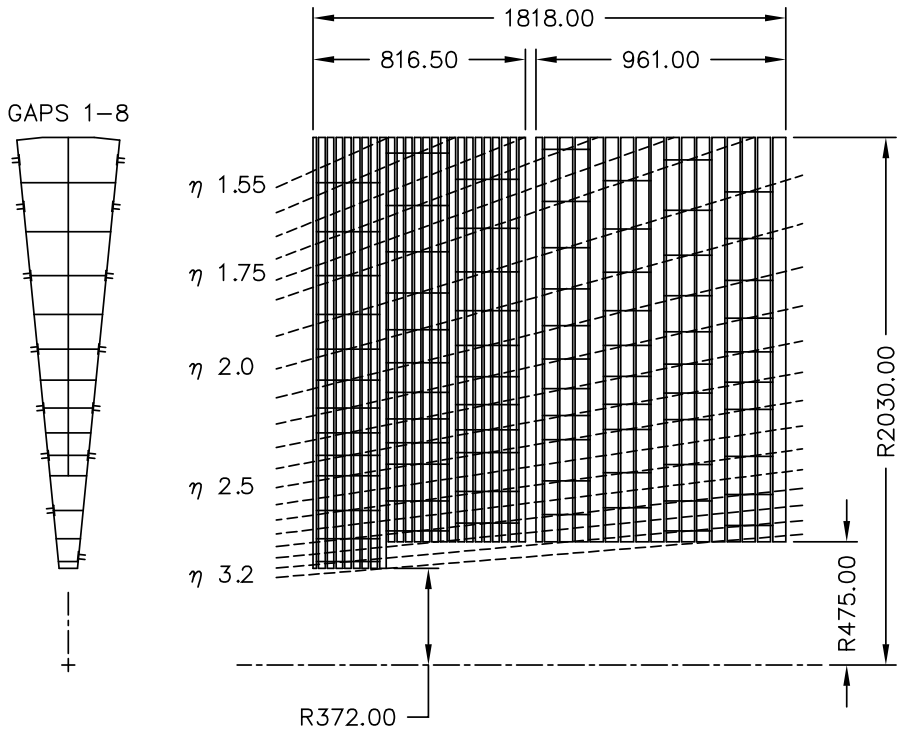


Figure 5.12: Schematic of the hadronid end-cap (HEC) absorber plates in the front and rear wheels, along with the trajectories of particles with varying η and segmentation of the pad readout, from [219].

Unlike the EM calorimeter accordian design, the HEC is a straightforward sampling calorimeter with alternating absorber and active material layers. The front wheel is made up of 24 $\Delta z = 25$ mm thick copper plates with a $\Delta z = 12.5$ mm thick front plate, while the rear wheel has 16 $\Delta z = 50$mm thick copper plates and a half-thickness front plate. Radially, the plates extend from $r = 0.372$ m for the first nine plates in the front wheel or $r = 0.475$ m for the remaining plates to $r = 2.03$ m. This slight difference in inner radii can be seen in Figure 5.11, where it is clear that the additional length of the nine plates sits longitudinally just before the FCal. Figure 5.12 shows the layout of the HEC wheels and plates in relation to the trajectories of high-p_{T} hadrons at varying values of η. The 8.5mm-sized regions between the plates are structurally supported with a honeycomb sheet kept in place with longitudinal support rods, and are filled with four 1.8 mm LAr drift zones alternated with three electrodes, shown in Figure 5.13. The electrodes are kept at a voltage of

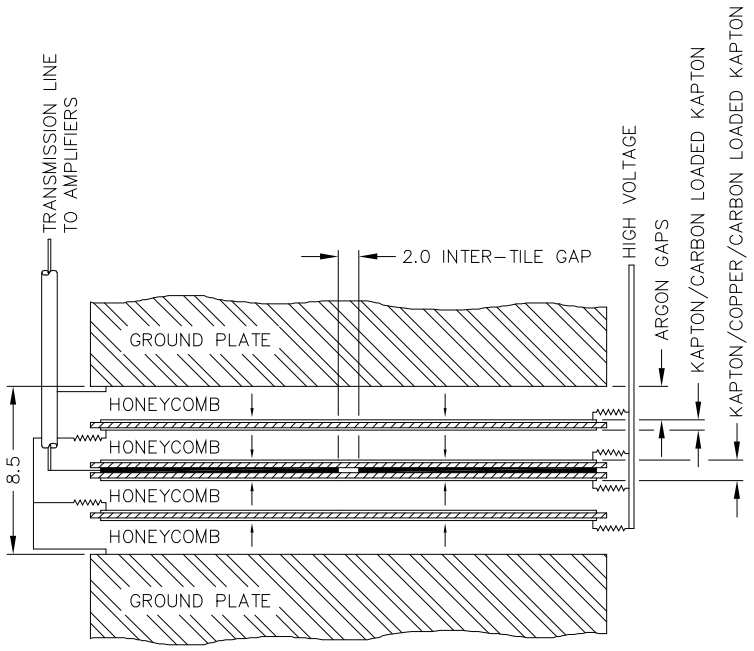


Figure 5.13: Absorber plate, LAr gap and electrode/readout structure in the hadronic end-cap (HEC) calorimeter, from [219].

+1800 V, with the middle electrode containing a segmented readout pads. Figure 5.12 shows how the readout segmentation varies from $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in for the parts of the HEC traversed by $|\eta| < 2.5$ to $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $|\eta| > 2.5$. The hadronic coverage is weakest in the transition region from the HEC to the FCal ($3.0 < |\eta| < 3.2$), as can also be seen in Figure 5.6.

The two Foward Calorimeters (FCal)[230] share cryostats with the end-cap calorimeters (see Figure 5.11) and are situated symmetrically at $3.1 < |\eta| < 4.9$, approximately $z = \pm 4.9$ m from the nominal beam interaction point. The FCal is segmented longitudinally (at this large η, $\hat{z}$ rather than $\hat{r}$ is the longitudinal direction) into three $\Delta z = 45$cm-long layers called FCal1, FCal2 and FCal3. Like most other elements of the calorimeter system, the FCal layers use LAr as the active material. Each FCal layer consists of a hexagonal array of electrodes embedded in an absorber matrix with an electrode nearest neighbor distance of $d = 7.5$ mm in the first layer and increasing in the second and third layers. In the first layer, the absorber matrix is a series of longitudinally

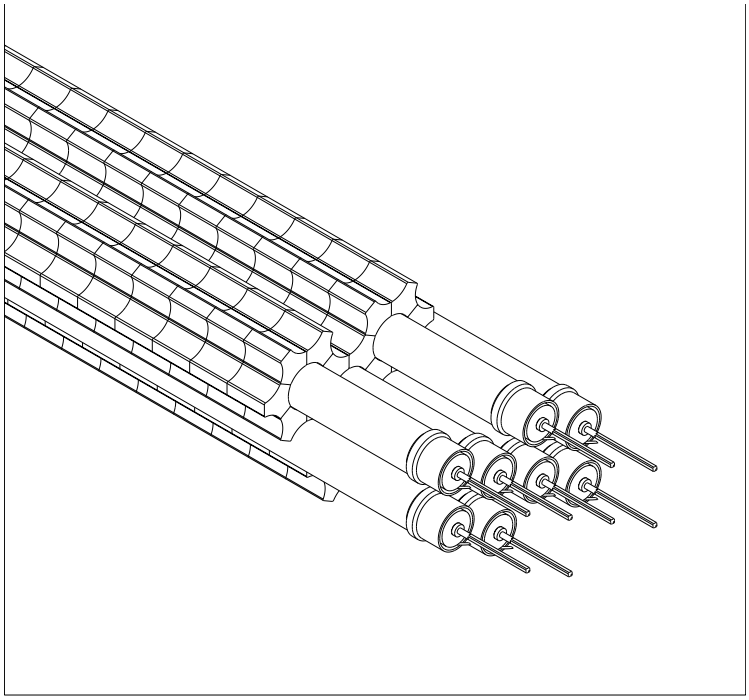


Figure 5.14: Tungsten slug absorber and electrode structure of the second and third hadronic FCal layers, from [219].

stacked copper plates, with holes drilled in them to hold the electrodes. In the second and third layers, the absorber matrix has copper front and end plates but is filled with tungsten slugs as shown in Figure 5.14. The electrodes consists of an outer copper cathode tube held at ground, an intermediary layer of LAr serving as the active material, and an inner anode rod made of the absorber material held at a voltage of +250 V (FCal1) to +500 V (FCal3). A radiation hard fiber helix is wound between the cathode and anode to keep them structurally and electrically separated. The structure of the electrode (cathode + LAr gap + fiber + anode) is shown in Figure 5.15. Note that the Molière radius (r = 17 mm) of the material, which defines the characteristic transverse area containing 90% of the electromagnetic shower, spans several electrodes and ensures that the response is reasonable uniform.

The LAr gap width (the width of the annular cylinder between the anode and cathode) ranges from 0.269 mm in the first layer to 0.508 mm in the third layer, where it can be bigger due to the

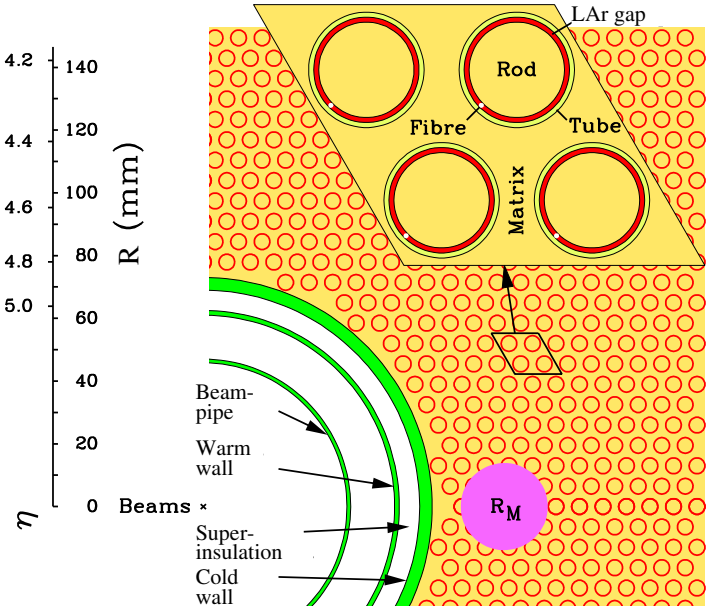


Figure 5.15: Electrode structure of the first FCal layer, with Molière radius for comparison, from [219].

decreasing ionization density. This small gap width helps prevent the build-up of LAr ions and keeps the drift time small. The first layer uses copper as the absorber material and functions as an electromagnetic calorimeter (with total radiation length $27.6X_0$ and nuclear interaction length $2.66\lambda_I$), while the second and third layers use tungsten as the absorber material and function as hadronic calorimeters (with nuclear interaction lengths $3.68\lambda_I$ and $3.60\lambda_I$, respectively).

Test beam measurements at CERN[231] have measured the energy resolution of electrons in the first FCal layer as

$$\frac{\sigma(\Delta E^{EM})}{E^{EM}} = 3.5(0.1)\% \pm \oplus \frac{28.5(1.0)\%}{\sqrt{E[\text{GeV}]}} \tag{5.5}$$

and the (hadronic) energy resolution of pions in the full FCal as

$$\frac{\sigma(\Delta E^{had})}{E^{had}} = 7.5(0.4)\% \oplus \frac{94.2(1.6)\%}{\sqrt{E[\text{GeV}]}} \tag{5.6}$$

5.4 ATLAS Inner Detector

The goal of the ATLAS Inner Detector (ID) design is to provide hermetic and accurate measurements of charged tracks to low p_{T} (> 0.5 GeV by original design but as low as > 0.1 GeV in ongoing performance studies) within five units of pseudorapidity ($|\eta| < 2.5$) and full 2π coverage in azimuth, as well as electron identification and secondary vertex reconstruction capabilities. In addition to these physics motivations, the high-radiation environment (the nearest ID element is 5 cm from the interation point) imposes significant requirements on the detector and electronic readout design.

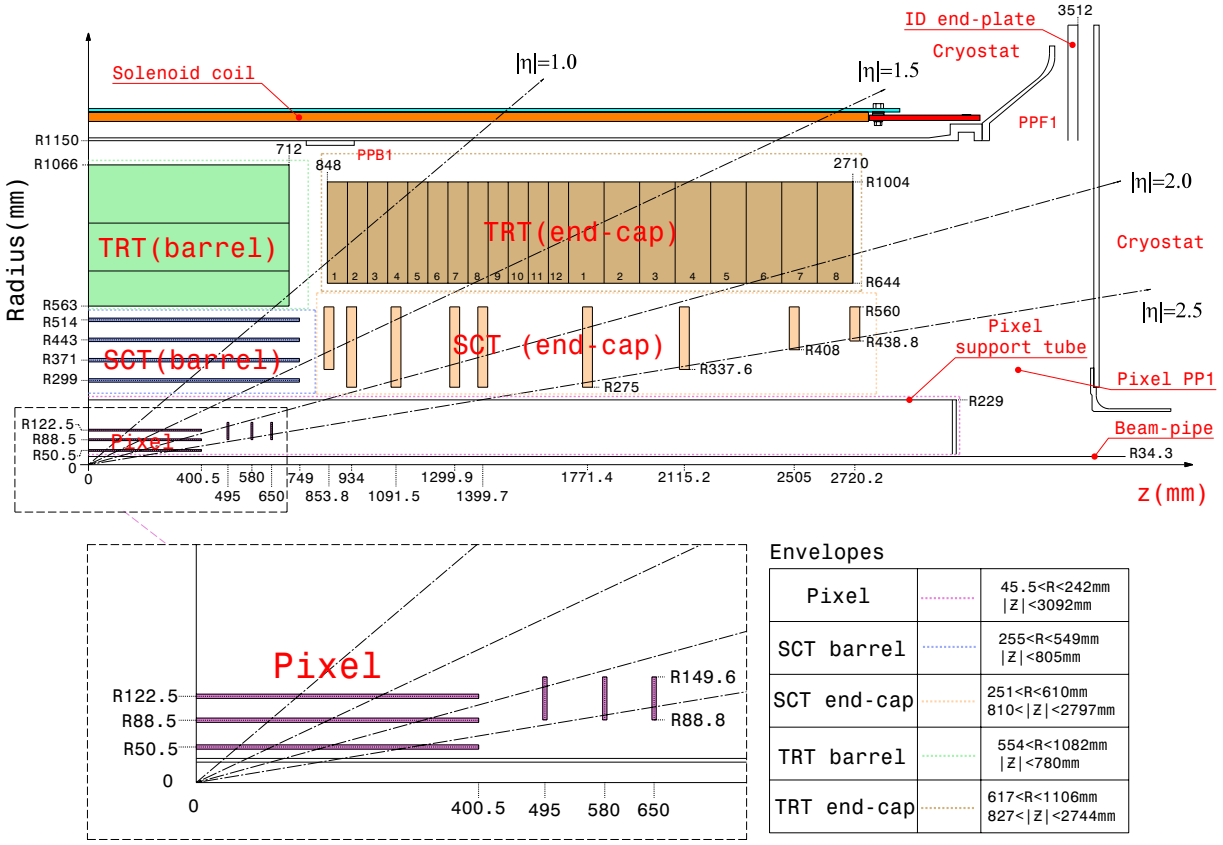


Figure 5.16: Schematic of the ATLAS inner detector showing the positions and dimensions of the Pixel, SCT and TRT components, from [219].

The ID is contained in a $z \in \pm 3.512$ m $\times\, r < 1.15$ m cylindrical envelope in an axial magnetic field of 2 T and consists of three distinct but complemenrary subsystems: the silicon pixel detector,

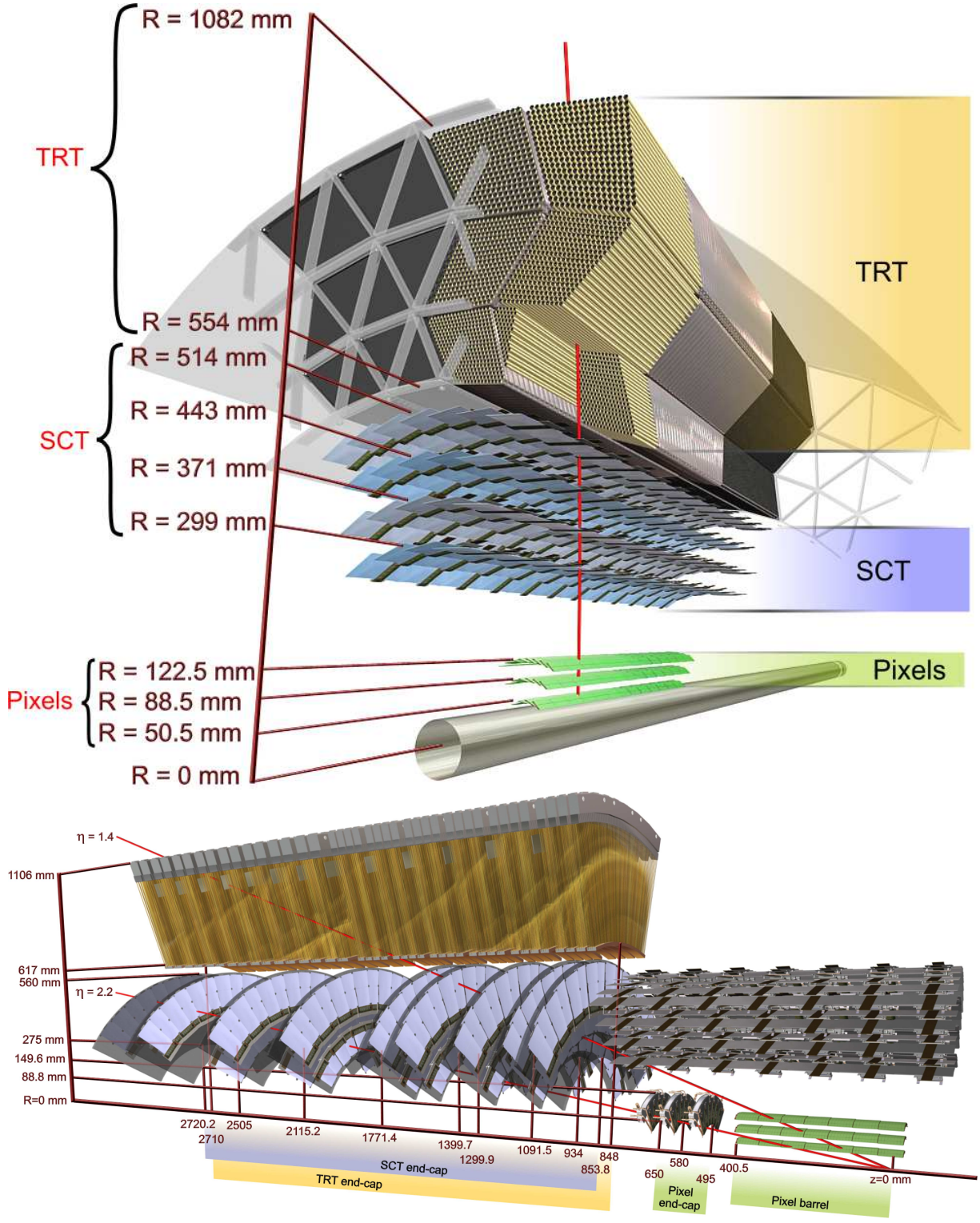


Figure 5.17: Drawing of ID detector elements traversed by 10 GeV charged particles that are close to mid-rapidity ($\eta = 0.3$, top) and in a more forward direction ($\eta = 1.4, 2.2$, bottom), from [219].

the silicon microstrip sensors (SCT) and the transition radiation tracker (TRT). The layout of the major detector components is shown in Figure 5.16. All three subsystems have "barrel" and "endcap" region, where the "transverse" plane sis considered the $\hat{z}$-$\hat{\phi}$ and $\hat{r}$-$\hat{z}$ plane, respectively. To this end, the components are situated differently in these two regions, as demonstrated in Figure 5.17. The semiconductor-based subdetectors (the Pixels and the SCT) are kept cooled to -10° C $< T < -5^\circ$ C by a C_3F_8 gas cooling system, while the TRT operates at room (really, cavern) temperature.

Taken together, the tracking system uses high-precision measurements close to the interaction point and lower-precision measurements farther out to best reconstruct the particle trajectories. The general performance goals of the tracking systems are to achieve a p_{T} resolution of

$$\frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} = 0.05\% \times p_{\mathrm{T}}[\mathrm{GeV}] \oplus 1\% \tag{5.7}$$

In recent measurements of the charged particle multiplicity in $\sqrt{s} = 0.9, 2.36$ and 7 TeV pp collisions[232], reconstruction efficiencies for charged particles of $\approx 85\%$ for > 10 GeV particles and 50-78% for > 100 MeV particles depending on the track η have been achieved.

5.4.1 Silicon pixel detector

The Pixel detector[233] consists of thee concentric cylindrical detectors in the barrel region (situated at $|z| < 40$ cm and with the layers at $r = 5.05, 8.85$, and 12.25 cm) and three endcap disks in each forward region (each is an annulus with radial extent 8.88 cm $< r <$ 14.96 cm and coaxial with the beam, with the six disks situated along the beamline at $z = \pm 49.5, 58.0, 65.0$ cm). In terms of pseudorapidity kinematics, the barrel detectors are positioned such that a high-momentum track with $\eta \approx \pm 1.9$ from the nominal interaction point just passes through all three barrel layers, and tracks with $|\eta| \gtrsim 1.7$ will traverse at least one disk.

The Pixels are semiconductor detectors with sensors made from n+-implants in an oxygenated n-bulk material (near the readout), with the p-n junction on the opposite end of the wafer surrounded by a guard ring structure which controls the potential drop. This design was chosen because it allows good charge collection efficiency and increased radiation tolerance. The layout of the active detector element is illustrated in Figure 5.18. The bias voltage for initial operation is 150 V, but expected to rise to as high as 600 V along with the accumulated radiation damage.

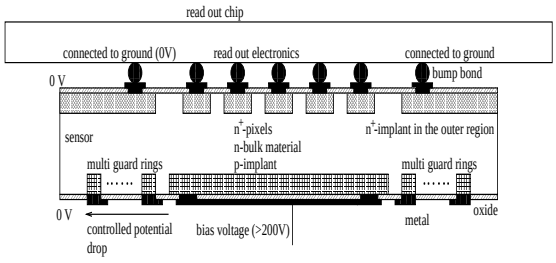


Figure 5.18: Cross-section of a pixel sensor, from [233].

The Pixels consist of 1744 24×63 mm^2 modules divided among the barrel layers (1456 modules) and endcap disks (288 modules). Each module is a stacked set of read-out electronics, bump bonds connecting the electronic channels to the pixel sensor output, the actual sensor tile, and a flex-hybrid circuit board. In the barrel layers, the modules are oriented so that the read-out electronics are the furthest radially, and the sensors see the smallest amount of material. During production and testing of the sensor-electronics assembly, the best performing modules were specially selected for inclusion in the layer-0 barrel region (the layer closest to the beampipe). Each sensor tile contains a 328×144 array of pixel cells, which are 50×400 μm^2 pixel cells (slightly larger in special areas such as near the front-end chips, where a special ganged-pixel arrangement is also required), resulting in a total of 80.4 million individual pixels (and slightly fewer distinct readout channels).

The pixel position resolution is 10 μm in the $\hat{r}$-$\hat{\phi}$ cylindrical plane and 115 μm in the $\hat{z}$ direction.

5.4.2 Silicon microstrip detector

The SCT detector[234; 235] consists of four cylindrical barrel layers (situated from $r = 29.9$ cm to 51.4 cm and within $z < \pm 80.5$ cm) and nine radial disks in the endcap region (of varying annular dimensions, but generally $r = 27.5$ cm to $r = 56.0$ cm, with positions along the beam axis from $z = \pm 81.0$ to ± 279.7 cm). The inner radius on the outermost four disks increases such that a particle with $\eta = 2.5$ just touches their inner radius. In general, particles with $|\eta| \lesssim 1.2$ traverse only the barrels, particles with $|\eta| \gtrsim 1.7$ traverse only the disks, and particles with an intermediate pseudorapidity traverse some layers in both.

Unlike the Pixels, the SCT is a traditional p-strip in high-resistivity n-bulk detector design. The SCT sensors operate at a bias voltage of 150 V but this is expected to slowly rise as high as

450 V over the lifetime of the detector, depending on the proximity to the interaction point and variations in the integrated radiation dose.

The SCT detector is composed of 4088 individual modules (2112 of which are in the barrel layers, ranging from 384 in the first to 672 in the fourth layers). Each barrel module consists of a pair of 6.4 cm long stereo strips daisy-chained together, with a pitch of 80 μrad and an angle of 40 mrad between them. The sensors are 285 μm thick in the transverse direction and consist of 770 individual readout strip (2 of which are non-functioning and serve to define the edges of the sensor). In the barrel, the strips are oriented near parallel to the beam axis, while in the endcap disks the are trapezoidal in shape and are arranged radially.

The SCT position resolution is 17 μm in the $\hat{r}$-$\hat{\phi}$ cylindrical plane and 580 μm in the $\hat{z}$ direction. The stability of the absolute position of the SCT elements is monitored through a frequency scanning interferometry (FSI) method[236]. This technique allows for possible shape changes to be detected in a high radiation, inaccessible environment.

5.4.3 Transition radiation tracker

The TRT is fundamentally a drift chamber detector. Like the Pixel and SCT detectors, the TRT detector consists of a cylindrical barrel component, situated at $z \in \pm 71.2$ cm and 56.3 cm $< r <$ 106.6 cm, and an endcap disk component situated at ± 84.8 cm $< z < \pm 271.0$ cm and 64.4 cm $< r <$ 100.4 cm. The basic element of the TRT is the straw tube, which are 4 mm diameter hollow cylinders. The straw exterior in a 0.2 μm Al shell which serves as the cathode, coated with ≈ 25 μm of graphite, polyimide and polyurethane layers for protection. A 31 μm diameter W wire coated with 0.5-0.7 μm of Au runs through the center of the straw and serves as the anode. The wires are kept at ground and connected directly to the front-end electronics while the cathode is kept at -1.53 kV. The straws are filled with a 70% Xe, 27% CO_2, 3% O_2 mixture. The interstitial region outside the straws but inside the modules is filled with CO_2 gas to prevent the accumulation of xenon, which would absorb transition radiation, and high-voltage discharges. Additionally, the CO_2 conducts heat from the straws to the module shells, which is further dissipated through a C_6F_14 gas cooling system.

Since the transition radiation from charged particles passing between different media is logarithmically proportional to the Lorentz factor of the particle and the frequency of produced photons

is more readily absorbed by the Xe in the gas mixture, electrons (which have a Lorentz factor $m_\pi/m_e \approx 270$ higher than $\pi^\pm$ of the same p_T) passing through the TRT leave a large transition radiation signal in the detector. In fact, two different gains are used in the front-end electronics for signals arising from transition radiation and that from energy deposited by charged particles. In this way, the TRT allows for electron identification within certain kinematic ranges.

The TRT barrel consists of three concentric collections of modules stacked radially, where the modules have a triangular cross-section in the $\hat{r}$-$\hat{\phi}$ plane and extend along the $\hat{z}$ direction. In the barrel region, the straws are oriented in the $\hat{z}$ direction inside the modules and are 144 cm long (and electrically split into two ≈ 71 cm active halves). Depending on the region of the barrel (the innermost 9 layers have only the 31.2 cm wire sections closest to the edges are active), there are 64 or 73 straw tube layers available to be crossed by charged particles. Each layer is rotated by 3/8ths of the straw diameter in ϕ to ensure that no ϕ position falls consistently on straw edges and thus has no acceptance.

In the TRT endcap, the straws are 37 cm long and oriented in the $\hat{r}$ direction. The detector is segmented into annular modules staggered along the $\hat{z}$ direction. There are twenty endcap modules, with the 12 nearest and 8 farthest from $z = 0$ being approximately $\Delta z = 7$ cm and 12 cm thick. Each endcap module holds 8 layers (which are spaced farther apart in the last 8 modules) ordered in $\hat{z}$, for a total of 160 straw tube layers in the entire endcap assembly.

Charged tracks with $p_\mathrm{T} > 0.5$ GeV and $|\eta| < 2$ typically traverse at least 36 straws (with decreased acceptance in the small barrel-to-end-cap transition). The intrinsic straw position resolution is 130 μm in the $\hat{r}$-$\hat{\phi}$ direction and provides essentially no z-discriminating information.

5.5 ATLAS Trigger and Data Acquisition

ATLAS selects which events to record based upon three levels of successively more detailed trigger decisions, narrowing the full 40 MHz crossing rate to a 100-200 Hz selected event rate. The guiding design principles of the trigger system include a fast trigger decision, good trigger efficiency for rare physics processes and an appropriate rejection factor for the final, limited bandwidth. The three levels are the Level-1 (L1), the Level-2 (L2) and Event Filter (EF), of which the first is implemented in hardware and the last two are grouped together into the software High-Level Trigger (HLT)

system. A trigger chain is an accepted sequence of $L1 \rightarrow L2 \rightarrow EF$ triggers. A trigger "menu" is a defined set of trigger chains to be used in data-taking, with instaneous luminosity-dependent sets of scaledowns which are adjusted during data-taking to keep the output event rate constant.

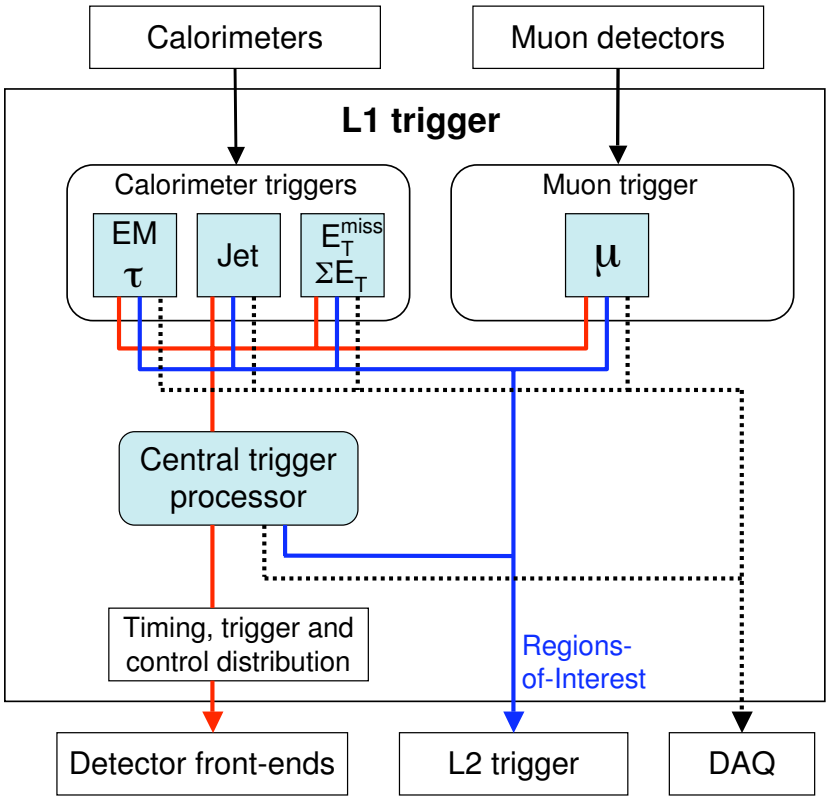


Figure 5.19: Diagram of the ATLAS Level-1 trigger decision system, from [219].

The L1 trigger system is managed by the Central Trigger Processor (CTP) and makes the fastest and therefore most basic decision from a combination of information from the calorimeter, ZDC, MBTS and muon subsystems. For example, the L1Calo system processes information from reduced-granularity $\Delta\eta \times \Delta\phi = 0.1\times = 0.1$ EM and Hadronic calorimeter towers, which are sent to the Cluster Processing (CP) and Jet/Energy-sum Processor (JEP) subsystems. The JEP identifies E_{T} sums in varying tower window sizes and compares these to pre-defined jet energy thresholds, resulting in L1 items of the form `L1Jx` (in the barrel) and `L1FJx` (in the FCal). While the event data fragments await the L1 trigger decision, they are pipelined in custom electronics near their detectors of origin. A schematic of the L1 operation is shown in Figure 5.19.

Overall, the L1 system can be programmed with up to 256 distinct items, each composed of specific requirements on the detector signals, and returns a trigger decision within 2.5 μs (or 100 bunch crossings), nearly 1 μs of which is cable delays rather than actual data processing. Due to the limited time allowed to make an L1 trigger decision, no tracking information from the inner detector (which requires a longer timescale for the full pattern recognition) can be used to form L1 trigger objects. The maximum L1 trigger accept rate is 75 kHz, a factor of > 500 smaller than the bunch crossing rate.

The CTP is also responsible for incrementing the *luminosity block* number (lumiblock, or lbn) every 120 seconds. Lumiblocks are the quanta used in offline data quality checks, and the subset of lumiblocks in each data-taking run which have stable detector performance is known as the Good Run List (GRL). Only events which fall in the GRL are used in a physics analysis. Lumiblocks are the smallest amount of time during which the luminosity detectors (see Section 5.2.2) can return a prescale- and deadtime-corrected integrated luminosity which is systematics rather than statistics limited. Finally, the CTP serves as the master timing system and is responsible for distributing the LHC clock information to all relevant subsytems. The timing information and trigger decisions are transferred from the CTP to the detector front-end and readout systems using the Trigger, Timing and Control (TTC) system over an optical-broadcast network.

Upon L1 acceptance by CTP, the event data is transferred to the DAQ/HLT in 1574 detector-specific fragments over Readout Links (ROL) and stored in the Readout System (ROS), to be used upon request in the two subsequent HLT stages. The L1 system also provides the Region of Interest (RoI), the specific region of the detector which actually fired the L1 trigger.

The L2 trigger stage examines the detector information in the RoI in more detail (for example, including the full calorimeter granularity and sampling information) and can apply more stringent and detailed trigger requirements than what is possible in the L1. Events selected by the L2 are tagged for full reconstruction before analysis by the EF, and are built in the one of the 48 event filter input nodes (SFI's). Somewhat analogous to the PHENIX Event Builder system (see Section 4.2), event fragments residing in the ROS are sent to the SFI for assembly in a process brokered by Data Flow Manager (DFM) nodes. For events not selected by the L2 trigger, the DFM informs the ROS that they can be safely expunged from memory. Overall, the maximum L2 event rate is 3.5kHz (a factor 20 smaller than the maximum L1 rate) with an average event-processing time of 40 ms.

The EF is the third and final stage of trigger processing. During this stage, the EF analyzes fully built events and can make sophisticated and detailed trigger decisions. The EF can consider the calorimeter information with full granularity, fully reconstructed jets (including multiple cone sizes and subtraction procedures), and even reconstructed tracks in the inner detector. The events selected by the EF are written out by the DAQ to the 6 event filter output nodes (SFO's), which buffer the data locally until they can be transferred to CERN's central data recording service. The SFOs can push a sustained output bandwidth of 300 MB/sec (with an instantaneous but not sustainable peak rate of 600 MB/sec). The typical event size in $\mathcal{L} = 10^{34}$ cm^2/s pp running is 1.3 MB, which corresponds to an event rate of 200 Hz for events of this size. Overall, the EF trigger stage has a maximum output rate of approximately 400 Hz (another factor of ≈ 10 reduction from the maximum L2 rate), with the typical time to process an event ranging from one to four seconds.

The HLT also associated each event with one or more data streams. In the present work, the two relevant streams are the `MinBias` stream, which contains basic variations on the minimum bias trigger (e.g. `MBTS_1_1`, `MBTS_2_2`, ZDC and space-point triggers, etc.) along with specialized triggers such as those selecting empty and unpaired bunches, and the `HardProbes` stream, which contains jet, photon, muon and other rare physics signature triggers.

Chapter 6

Direct jet reconstruction in d+Au collisions

I have no spur to prick the side of my intent

Macbeth, Act I, Scene VII

This chapter describes the measurement of inclusive jet production in 200 GeV d+Au and pp collisions. The single particle selection and jet reconstruction procedure are described in Section 6.1. The data selection and checks on the stability of jet yields are described in Section 6.2. The trigger selection and efficiency determination are described in Section 6.3. The Monte Carlo simulation, determination of the detector response and correction for underlying event effects are described in Section 6.4. The results are shown in Section 6.5. Systematic uncertainties are tabulated and a few additional cross-checks are considered in Section 6.6.

6.1 Jet reconstruction

In this work, electromagnetic clusters in the EMCal, arising from photons, electrons, decaying neutral pions and neutral hadrons, and charged tracks in the Drift Chamber, arising from charged hadrons and electrons, are reconstructed into full jets using the Gaussian filter algorithm with $\sigma = 0.2$ and $\sigma = 0.3$. Jets originating from clusters and tracks in the East arm only are used for the final result, whereas both East and West arms are used for several studies and checks. Reconstructed

$\sigma = 0.3$ jets are used for the final jet results, and $\sigma = 0.2$ are retained as a systematic control against the influence of the d+Au underlying event.

The gaussian filter is a seedless, infrared-safe algorithm which looks for local maxima in the energy density of an event to reconstruct a jet. Formally, it constructs a p_{T} density in η/ϕ space:

$$p_{\mathrm{T}}(\eta,\phi) = \Sigma_{\text{particle } i}\ {p_{\mathrm{T}}}^{i}\ \delta(\eta-\eta^{i})\ \delta(\phi-\phi^{i}) \tag{6.1}$$

To determine how consistent a given position (η,ϕ) is with that of a jet, it performs the convolution:

$$\tilde{p}_{\mathrm{T}}(\eta,\phi) = \int d\phi' \int d\eta'\ p_{\mathrm{T}}(\eta',\phi')h(\eta-\eta',\phi-\phi') \tag{6.2}$$

Where the filtering kernel h is a Gaussian distribution with parameter σ:

$$h(\Delta\eta,\Delta\phi) = e^{-(\Delta\eta^2+\Delta\phi^2)/2\sigma^2} \tag{6.3}$$

The final jet positions are the η^{jet}, ϕ^{jet} for which Equation 6.2 is a local maximum. The p_{T} of the jet is given by

$$p_{\mathrm{T}}^{jet} = \tilde{p}_{\mathrm{T}}(\eta^{jet},\phi^{jet}) \tag{6.4}$$

The gaussian filter is described in more detail in Section 2.1.5 and its performance is benchmarked against other algorithms in [46]. Previously, it has been successfully used in PHENIX in 200 GeV p+p and Cu+Cu collisions[47].

The p_{T} bins are chosen to grow slowly with p_{T} to compensate for the falling statistics at high-p_{T} and to aid in the interpretation of the data as possible nPDF effects (which are more easily visible on a logarithmic p_{T}-scale). There are 25 bins between 5 and 50 GeV/c, with widths that are constant in logarithmic p_{T}-space. Specifically, the 26 bin edges and 25 bin centers are defined by

$${p_{\mathrm{T}}}^{edge} = 5\times 10^{n/25}, n\in 0\ldots 25 \tag{6.5}$$

$${p_{\mathrm{T}}}^{center} = 5\times\frac{1}{2}\left(10^{n/25}+10^{(n+1)/25}\right), n\in 0\ldots 24 \tag{6.6}$$

Since $10^{1/25} \sim 1.1$, a handy rule of thumb is that the width of a bin is about $\sim 10\%$ the central bin p_{T} value.

6.1.1 Strategy in PHENIX

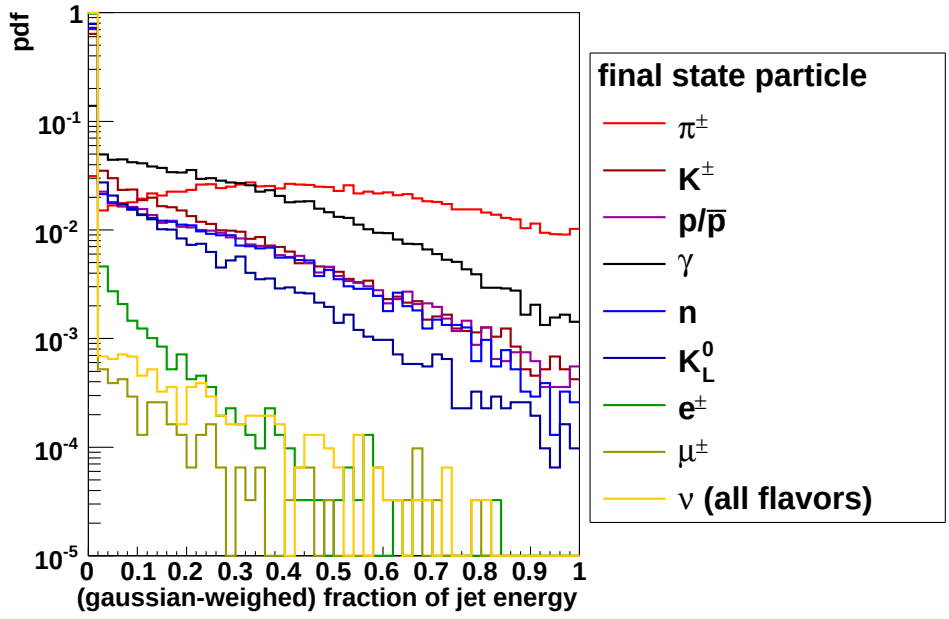


Figure 6.1: Mean contribution to the final truth jet energy, by particle species.

To calibrate the strategy for a successful jet reconstruction with the PHENIX detector, and to have a ballpark understanding of how to best capture the fragmenting jet energy, we perform a toy PYTHIA study with ${p_{\mathrm{T}}}^{truth} > 10 - 15$ GeV/c truth jets (although the results were found to hold broadly across the entire accessible ${p_{\mathrm{T}}}^{truth}$ range). The truth jets are statistically examined for the relative contributions of different final-state particles to the total energy. Figure 6.1 and Table 6.1 show the mean contribution to final truth jet energy from different sources, broken down by final state particle species (π^0's, ρ's and other short-lived neutral mesons are considered to have

$h^{\pm}$	62%	$\pi^{\pm}$	45±26%
		$K^{\pm}$	9±16%
		$p/\bar{p}$	8±16%
EM	26%	γ	25±21%
		$e^{\pm}$	1±2%
h^0	12%	$n, \bar{n}$	7±15%
		K_L^0	5±10%
misc.	$\sim 1\%$	$\mu^{\pm}$	$0.5 \pm 1.5\%$
		$\nu/\bar{\nu}_{\text{all}}$	$0.5 \pm 2\%$

Table 6.1: Summary of mean contribution (with RMS) to the final truth jet energy, by paticle species.

decayed into photons in this analysis). Although the fluctuations in particle composition on a jet to jet basis are very large due to different fragmentation scenarios, it can be seen that on average charged hadrons contribute an average of 62% of the jet energy, while electromagnetic energy sources ("EM", above) and neutral hadrons contribute 26% and 12% of the energy, respectively. There is also a very small muonic (and invisible neutrino) contribution from rare heavy flavor jets.

The strawman best jet performance using PHENIX instrumentation can be estimated by assuming a p_{T}-independent $\sim 80\%$ efficiency for $h^{\pm}$ reconstruction, $\sim 100\%$ efficiency for reconstructing γ's and a $\sim 0\%$ efficiency for reconstructing neutral h^0 (since PHENIX has no hadronic calorimetry). Furthermore, the energy resolution on the individual particles is substantially smaller than the resolution on the total jet energy due to fluctuations in the fragmentation, so the individual cuts on the particle level can be relatively looser. In this scenario, the PHENIX detector would typically capture $80\% \times 62\% + 100\% \times 26\% \sim 75\%$ of the jet energy. Thus, good reconstruction performance can be achieved by designing tracking and cluster reconstruction cuts that are as efficient as possible while still rejecting backgrounds. This is the philosophy with which the cuts in Sections 6.1.2 and 6.1.3 were developed, and their performance is quantified in Section 6.4.2.

The approximations above are rough. For example, there will always be losses of photons in the cracks between EMCal sectors, and the efficiency of the tracking reconstruction in PHENIX

is a geometry-dependent quantity due to differences in acceptance. Although the EMCal response to neutral hadrons is not good, a study[211] has investigated the ability of the EMCal to measure neutral hadronic energy and of PISA to properly model the interactions, and so special restriction on the contribution from neutral hadrons is attempted with the single particle cuts.

6.1.2 Electromagnetic cluster selection

Electromagnetic clusters are reconstructed as lightlike (massless) four-vectors. The direction vector (with arbitrary normalization) is $(x, y, z - z_{vtx})$, with x, y and z being the absolute reconstructed cluster coordinates. Note that the z-component of the direction is relative to the collision point at $(0, 0, z_{vtx})$. Since the electromagnetic calorimeter is used primarily to collect neutral electromagnetic energy (from direct and fragmentation photons), the "core" cluster energy (`emcClusterContent::ecore()`) is used. The kinematics are summarized by

- $\theta = \arccos\left(\frac{z-z_{vtx}}{\sqrt{x^2+y^2+(z-z_{vtx})^2}}\right)$
- $\phi = \arctan\left(\frac{y}{x}\right)$
- $\eta = -\log\left(\tan\left(\frac{\theta}{2}\right)\right)$
- $p_{\mathrm{T}} = \mathrm{ecore} \cdot \sin\theta$

Reconstructed clusters must then pass a series of cuts to ensure that they arise from the deposition of real energy.

First, only clusters with energy above 400 MeV are kept:

- $p_{\mathrm{T}} > 400$ MeV

Next, clusters from towers which have pathologically bad behavior are discarded. Each calorimeter tower is assigned one of four statuses:

- *Normal.* Clusters in these towers appear to be well-calibdation and have a time of flight distribution consistent with a Gaussian.
- *Dead.* These towers have no counts at all or only very few counts at very low E_{T}. These towers are masked off. (In principle, they can be allowed to contribute low-E_{T} clusters, but they are masked off in both data and MC to ensure the same acceptance.)

- *Hot.* These towers have many fake clusters from electronic noise that are impossible to distinguish from real (light-speed) particles. These towers are masked off.
- *Hot at low* E_{T}. These towers appear to be hot at low E_{T} but normal at high E_{T}. Treated as *Hot* for clusters with $E_{\mathrm{T}} < 1$ GeV, but *Normal* for clusters with $E_{\mathrm{T}} > 1$ GeV.

The status of a tower and its time of flight calibration are run-independent. The number of towers with each status is summarized in Table 6.2 and the tower maps are shown in Figure 6.2.

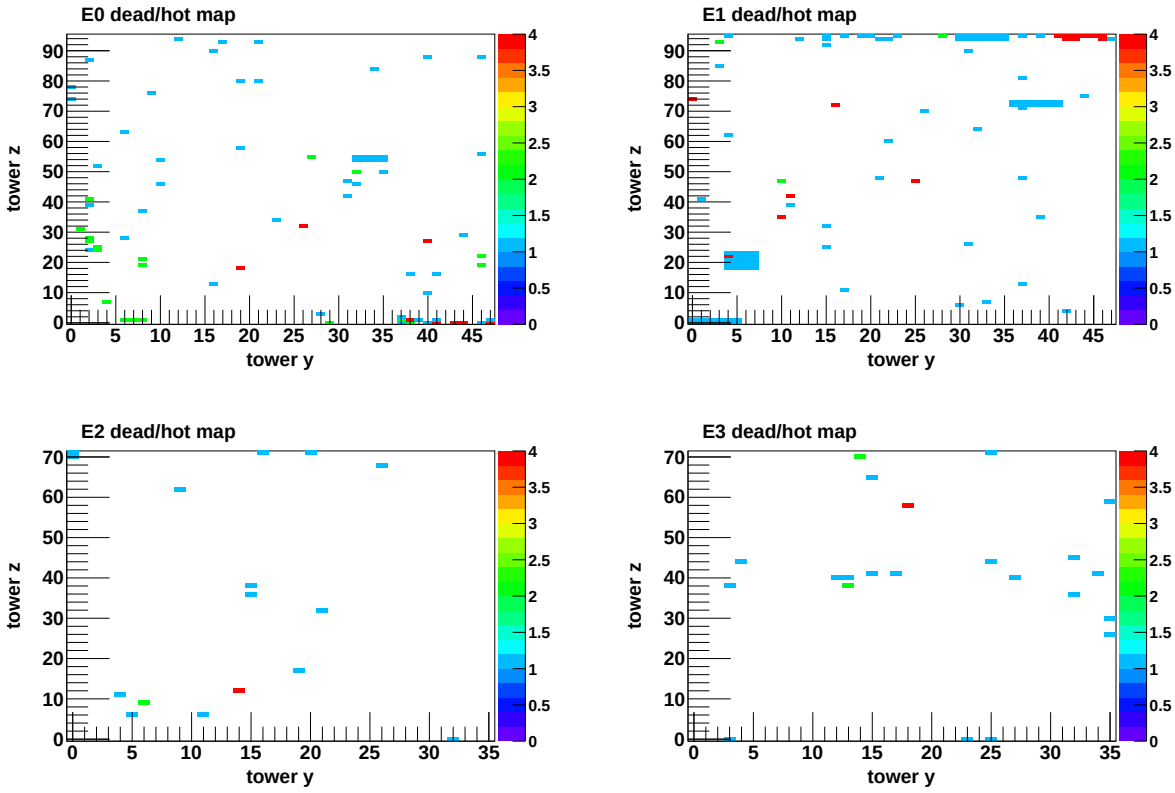


Figure 6.2: Tower statuses for the East arm. The histogram values 1, 2 and 4 correspond to *Dead*, *Hot at low* E_{T} and *Hot* towers, respectively. All unmarked towers are *Normal*.

Furthermore, clusters in *Normal* or *Hot at low-*E_{T} towers must have a time of flight consistent with that of a light-speed particle originating from the collision vertex. The time of flight is measured by the EMCal and calculated relative to the BBC-reported t_0 in the collision:

- ToF [ns] = `emcClusterContent::tof()` - t_0

Sector	Total towers	Normal towers	Dead towers	Hot towers	Hot at low E_{T}	total active area %
E0 (PbGl)	4608	4532	49	8	19	98.35%
E1 (PbGl)	4608	4495	95	15	3	97.55%
E2 (PbSc)	2592	2576	14	1	1	99.38%
E3 (PbSc)	2592	2570	19	1	2	99.15%
all East	14400	14173	177	25	25	98.61%

Table 6.2: Summary of hot/dead tower masks per sector.

The mean ToF μ and the resolution σ_{ToF} are determined by a gaussian fit to the ToF distribution for $E_{\mathrm{T}} > 400$ MeV clusters in each of the $\approx$ 21k towers. For most towers, a fit in the range -20 ns $< ToF < +20$ ns is sufficient, whereas a small number of towers have a displaced peak and thus require special treatment.

To be deemed consistent with a light-speed particle, a cluster must have a time of flight that is $\pm 5\sigma_{ToF}$ within the tower mean, where σ_{ToF} is the time of flight resolution in the tower. Some towers have a hot spot at a localized time of flight value within the $\pm 5\sigma_{ToF}$ window; these have their lower or higher ToF limit adjusted manually to exclude the noise. Thus, the cut is

- $-5\sigma_{ToF} + \mu <$ ToF $< \mu + 5\sigma_{ToF}$
- for select special towers, $ToF^{low} <$ ToF and/or ToF $< ToF^{high}$

Low and high time of flight limits are shown in Figures 6.3 and 6.4. It was found that the mean ToF slews at high-E_{T}, but that the $\pm 5\sigma_{ToF}$ window is large enough that no clusters are erroneously rejected.

Finally, the set of good clusters is compared with the set of good tracks to see if any of the two were associated during PHENIX event reconstruction, which arises from charged hadrons and electrons depositing energy in the EMCal in addition to being reconstructed as charged tracks. In these cases, the clusters are rejected and only the tracks are kept, since the DC has better momentum resolution and the EMcal only captures a small part of the full hadronic energy.

Therefore, we apply the cluster-track association cut,

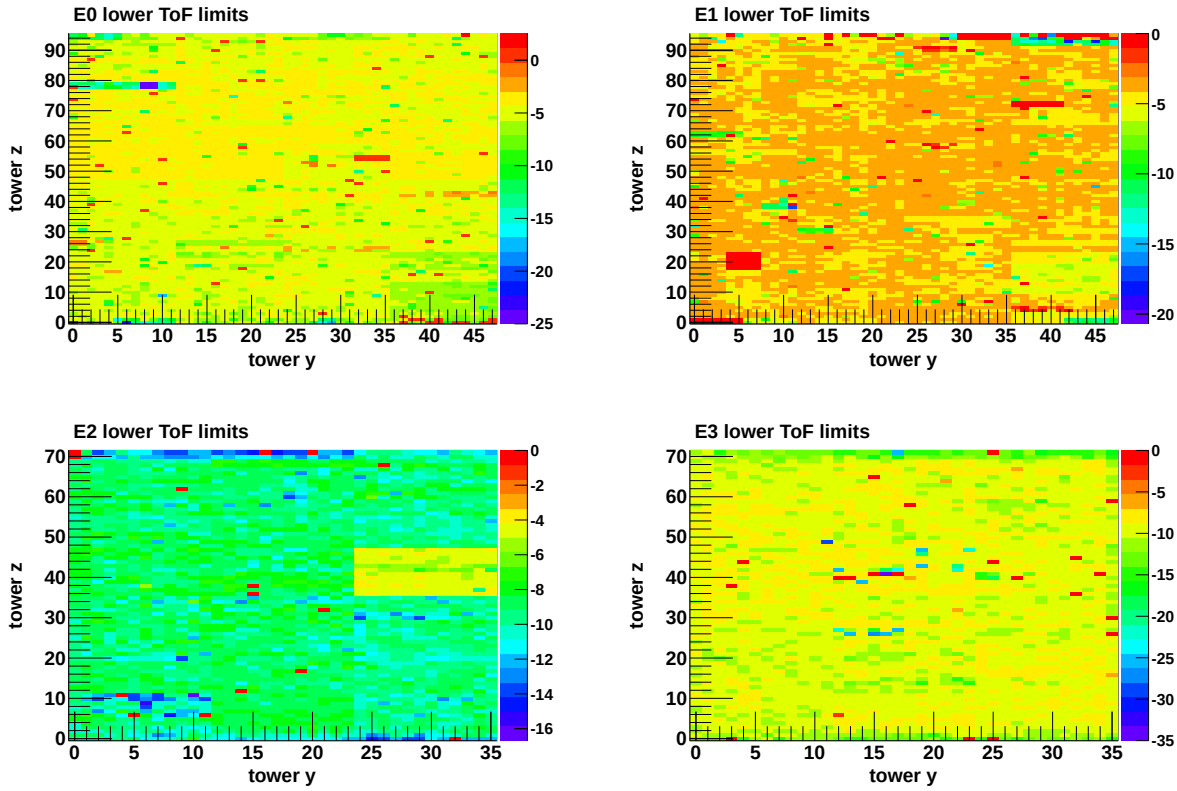


Figure 6.3: Lower time-of-flight limits in towers. Clusters with a time of flight below this value are excluded.

- after all cuts are performed, discard any good cluster which is associated with a good charged track

The resulting good clusters are packaged into a four-vector $(p_{\mathrm{T}}, \phi, \eta, m = 0)$ and used in jet reconstruction.

6.1.3 Charged track selection

The spatial direction and total momentum of the charged tracks is given by the θ_0, ϕ_0 and p from the Drift Chamber reconstruction. Tracks must pass several quality cuts, which are traditionally divided into "single track" cuts (cuts that are applied track by track) and "pair track" cuts (cuts which are checked against every combinatoric pair of tracks in the same event).

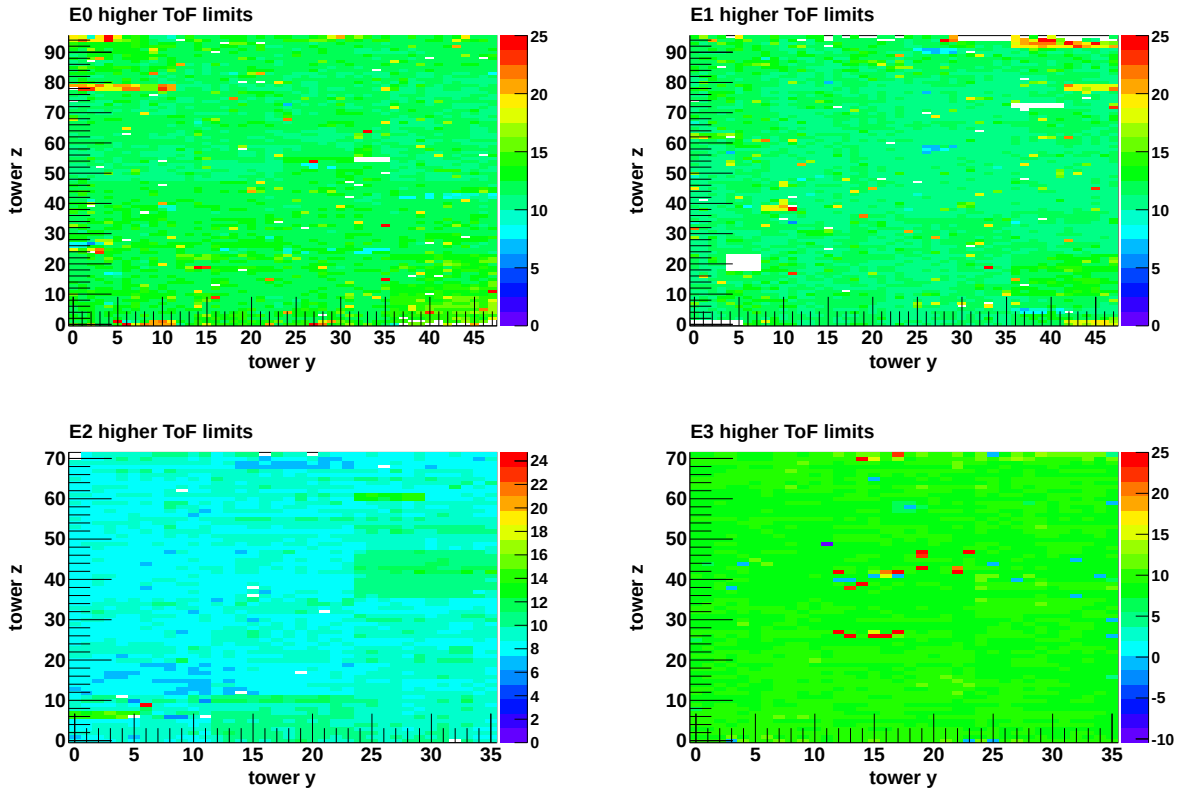


Figure 6.4: Higher time-of-flight limits in towers. Clusters with a time of flight above this value are excluded.

Tracks that are so ill-defined due to being composed of a small number of wire hits often do not have a definite z_{DC} or ϕ_{DC} assigned to them (instead these variables have the PHENIX "magic number" of -9999 instead). These are rejected,

- -100 cm $< z_{DC} < +100$ cm and $\phi_{DC} \in [0, 2\pi)$

The remaining tracks have enough longitudinal information to calculate the transverse momentum,

- $p_{\mathrm{T}} = p \cdot \sin\theta_0$

As with the clusters, we require a minimum p_{T} cut of 400 MeV. However, the track measured p_{T} spectrum above 25 GeV is completely dominated by mis-measured stiff tracks, so tracks also have a p_{T} *maximum*.

- $p_{\mathrm{T}} > 400$ MeV
- $p_{\mathrm{T}} < 25$ GeV

ϕ_{DC} range	z_{DC} range	comment
$2.40 < \phi_{DC} < 2.60$	$z_{DC} < 0$ cm	X1 mask
$3.00 < \phi_{DC} < 3.12$	$z_{DC} < 0$ cm	X1 mask
$2.40 < \phi_{DC} < 2.60$	$z_{DC} > 0$ cm	X1 mask
$3.00 < \phi_{DC} < 3.12$	$z_{DC} > 0$ cm	X1 mask
$3.20 < \phi_{DC} < 3.34$	$z_{DC} > 0$ cm	X1 mask
$2.61 < \phi_{DC} < 2.67$	$z_{DC} < 0$ cm	X2 mask
$2.91 < \phi_{DC} < 3.02$	$z_{DC} < 0$ cm	X2 mask
$3.34 < \phi_{DC} < 3.40$	$z_{DC} < 0$ cm	X2 mask
$2.91 < \phi_{DC} < 3.02$	$z_{DC} > 0$ cm	X2 mask
$2.10 < \phi_{DC} < 2.20$	$z_{DC} < 0$ cm	UV mask
$2.30 < \phi_{DC} < 2.45$	$z_{DC} < 0$ cm	UV mask
$2.70 < \phi_{DC} < 2.80$	$z_{DC} < 0$ cm	UV mask
$3.65 < \phi_{DC} < 3.80$	$z_{DC} < 0$ cm	UV mask
$2.10 < \phi_{DC} < 2.20$	$z_{DC} > 0$ cm	UV mask
$2.65 < \phi_{DC} < 2.75$	$z_{DC} > 0$ cm	UV mask
$3.65 < \phi_{DC} < 3.80$	$z_{DC} > 0$ cm	UV mask

Table 6.3: Broken wire masks, East arm

In addition to these minimum cuts, only relatively high-quality tracks according to the quality bit scheme in Section 4.1.1 are accepted. Due to the presence of broken X1, X2 and UV wires in the DC, the quality requirements are slightly loosened in areas near these broken wires, as summarized in Table 6.3. The same set of masks are used for the d+Au and pp runs. Of the four types of quality bits (X1, X2, UV, PC1), we require the quality bits q to be:

- X1 wire hit ($q \,\&\, 1 = 1$) unless the track is in a region of weak X1 acceptance (Table 6.3)

- X2 wire hit (q & $2 = 2$) unless the track is in a region of weak X2 acceptance (Table 6.3)
- Good PC1 and UV information. One of the following:
  - Ambiguous PC1 hit (q & $16 = 0$) and unique UV wire hit (q & $12 = 12$), or
  - Unique PC1 (q & $32 = 32$), and unique UV wire hit (q & $12 = 12$) unless the track is in a region of weak UV acceptance (Table 6.3).

As further evidence that a track corresponds to a real charged particle, and not noise in the Drift Chamber, we require that it is associated with a hit in the PC3 or (since the PC3 has numerous dead or inefficient region) a cluster in the EMCal. The expected distance between hits in the DC and hits in the PC/EMC is described by a normalized distribution in terms of the number of standard deviations σ in the z and ϕ directions. A so called "radial 3σ" cut is used, such that each track must be:

- associated with PC3 hit: $\sqrt{\sigma^2_{PC3,\Delta\phi} + \sigma^2_{PC3,\Delta z}} < 3.0$, or
- associated with EMCal hit: $\sqrt{\sigma^2_{EMC,\Delta\phi} + \sigma^2_{EMC,\Delta z}} < 3.0$

After the application of the previous single tracking cuts, it was observed that there are pairs of tracks that are extremely close together in $(\phi_{\rm DC}, z_{\rm DC})$ space. Furthermore, this phenomena tended to be associated preferentially with high-$p_{\rm T}$ tracks and also with tracks falling in or near a DC guard wire region. These tracks are dominated by "ghosts", which are multiple copies of the same hadron being reconstructed by the software, and conversions, the result of a photon converting on the face of the drift chamber and turning into two stiff tracks not affected by the magnetic field which the software erroneously determines must be extremely high-$p_{\rm T}$ to bend so little.

As it turns out, ghosting phenomena is easily observed in Monte Carlo (where two reconstructed tracks have the same `McSingleEvalList` ancestry, and matching to truth-level `PYTHIA` confirms that they come from the same single primary charged hadron). However, cuts on reconstructed tracks could confuse them for track pairs arising from a combinatorial overlap of two different charged particles. In fact, because of the segmentation of the PC1 in $z_{\rm DC}$ and the PC1 hit sharing that can happen during PHENIX charged tracking reconstruction, combinatoric pairs of tracks are often forced to be close together in $z_{\rm DC}$. The Monte Carlo allows us to distinguish these two

cases and derive a cut which identifies ghosts with only a small amount of false positives from combinatoric hadron pairs.

The ghost-pair selection cuts are:

- *close together in phase space:* $|\phi_{\rm DC}| < 0.05$ rad, $|z_{\rm DC}| < 0.02$ cm
- *close to an X1 or X2 guard wire region or in broken wire region:* one of the following:
  - both tracks have min. X1 distance < 0.3 cm, or
  - both tracks have min. X1 distance > 3 cm (e.g. no X1 hits), or
  - both tracks have min. X2 distance < 0.3 cm, or
  - both tracks have min. X2 distance > 3 cm (e.g. no X2 hits)

In jet-triggered Monte Carlo samples, these criteria select over 90% of identified ghost pairs, with a false positive rate corresponing to 0.1% of real jets having combinatoric track pairs selected.

Since the ghost track pair really does correspond to a real charged hadron, we should keep the track with more wire hits, since these tend to have the better approximation of the original hadron $p_{\rm T}$. The criteria for selecting the good track are phenomenologically driven from observations in MC and are based on the number of X1 and X2 hits (n_{X1}, n_{X2}):

- If one track has $n_{X1} < 3$ or $n_{X2} < 3$ while the other has $n_{X1} \geq 3$ and $n_{X2} \geq 3$, use the latter
- If both tracks have $n_{X1} < 3$ or $n_{X2} < 3$, keep the lower $p_{\rm T}$ one.
- If both tracks have $n_{X1} \geq 3$ and $n_{X2} \geq 3$, then consider the product $n_{X1}n_{X2}$, where each n is capped at 6:
  - If $n_{X1}n_{X2}$ for one track is > 1.25 as much as $n_{X1}n_{X2}$ of the other track, use the first track
  - If $n_{X1}n_{X2}$ for either track are within 25% of each other, keep the lower $p_{\rm T}$ track.

It has been demonstrated[237] that the effect of conversions in PHENIX can also be reduced with a set of cuts that calculate the distance of closest approach between oppositely-signed track pairs. Additionally, many PHENIX hadron measurements use light from the RICH as a veto

against conversion tracks. However, doing so in the present case would have a negative effect on the efficiency for high-p_{T} charged pions. Therefore, in this analysis, the remaining background from ghost and conversion track pairs is removed using cuts at the jet- rather than the single particle-level.

The final set of tracks are then assumed to be massless hadrons, and the four-vector ($p_{\mathrm{T}}, \phi, \eta, m = 0$) is used in jet reconstruction.

6.1.4 Stability of single particle yields

To evaluate any residual changes in the acceptance of the Drift Chamber or EMCal East arm, we examine the ϕ- and z- differential yields of charged tracks and clusters at beginning and ending of the Run 8 configurations:

- Run #246444, the first good d+Au run (the field is in a ++ configuration)
- Run #253701, the last good d+Au run (the field is in a -- configuration)
- Run #256450, the first good p+p run (the field is in a -- configuration)
- Run #259368, the last good p+p run (the field is in a -- configuration)

We examine 500,000 events in each run from the disk-resident Minimum Bias data, applying all event and particle selection cuts. For the d+Au runs, we focus on the 60-80% centrality, which is most like the p+p collision system in terms of N_{part} and N_{coll} and where any strong centrality-dependent effects are likely to be weak.

run	track ϕ_{DC}	track z_{DC}	cluster ϕ	cluster z
246444	1.000	1.000	1.000	1.000
253701	0.989(9)	0.988(9)	0.983(9)	0.975(10)
256450	0.565(5)	0.565(5)	0.570(5)	0.567(5)
259368	0.559(5)	0.559(5)	0.571(5)	0.570(5

Table 6.4: Mean ratio of differential yields in selected runs to the differential yield in Run #246444.

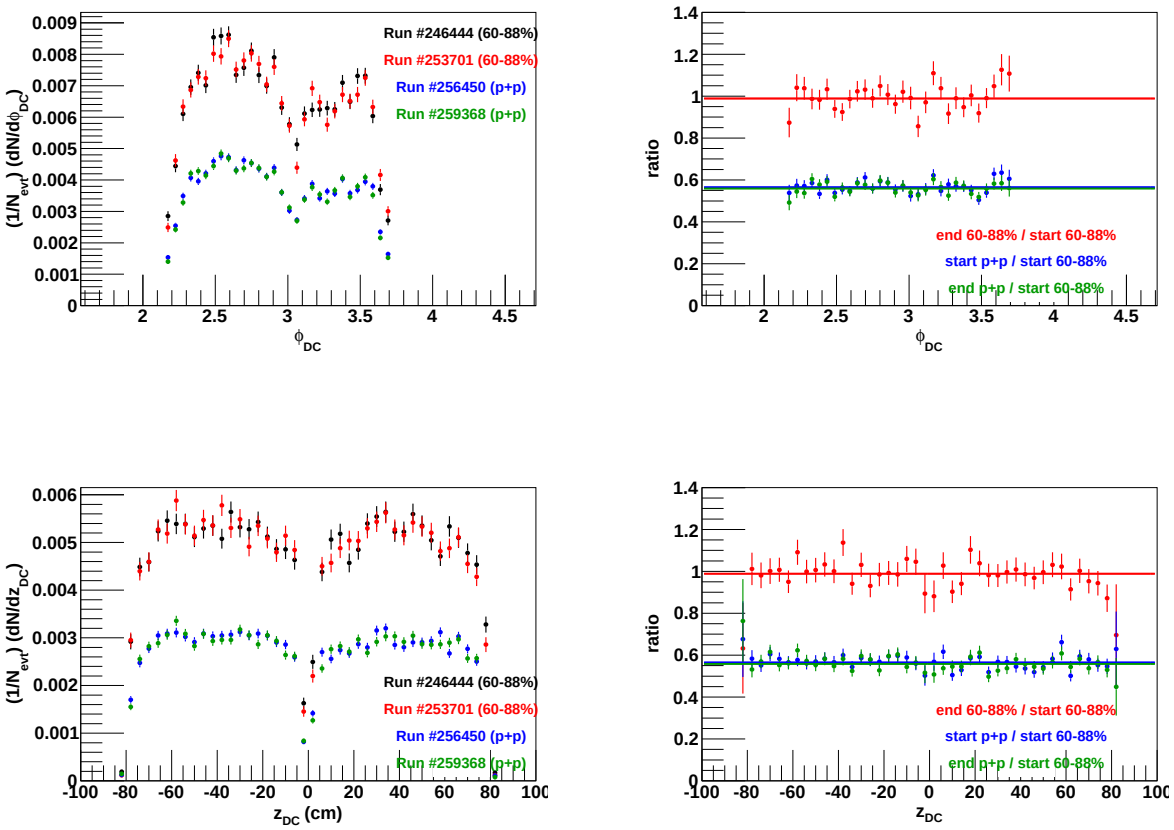


Figure 6.5: Differential charged track yields, $1 < p_{\mathrm{T}} < 2$ GeV/c, in minimum bias events per ϕ_{DC} (top plots) and z_{DC} (bottom plots). Differential yields are shown in the left plots, common ratio to the run #246444 reference are shown on the right.

Per-event $dN/d\phi_{\mathrm{DC}}$ dN/dz_{DC} yields for charged tracks passing all single track cuts and pair cuts with $1 < p_{\mathrm{T}} < 2$ GeV/c are plotted in Figure 6.5, and for clusters passing all tower masks and time of flight cuts with $1 < p_{\mathrm{T}} < 2$ GeV/c are plotted in Figure 6.6. A line is fit through the distributions. Table 6.4 summarizes the results. The ratios of ϕ- and z-differential yields before the beginning and end of the run appear to be consistent with statistical fluctuations. However, the ϕ- and z-integrated yields are consistent with a $-1 \pm 1\%$ systematic decrease in the charged tracks and a $-2 \pm 1\%$ decrease in the calorimetry over the course of the Run 8 d+Au. The pp run is stable within statistical uncertainties. However, the study above includes the ++ d+Au data, which are

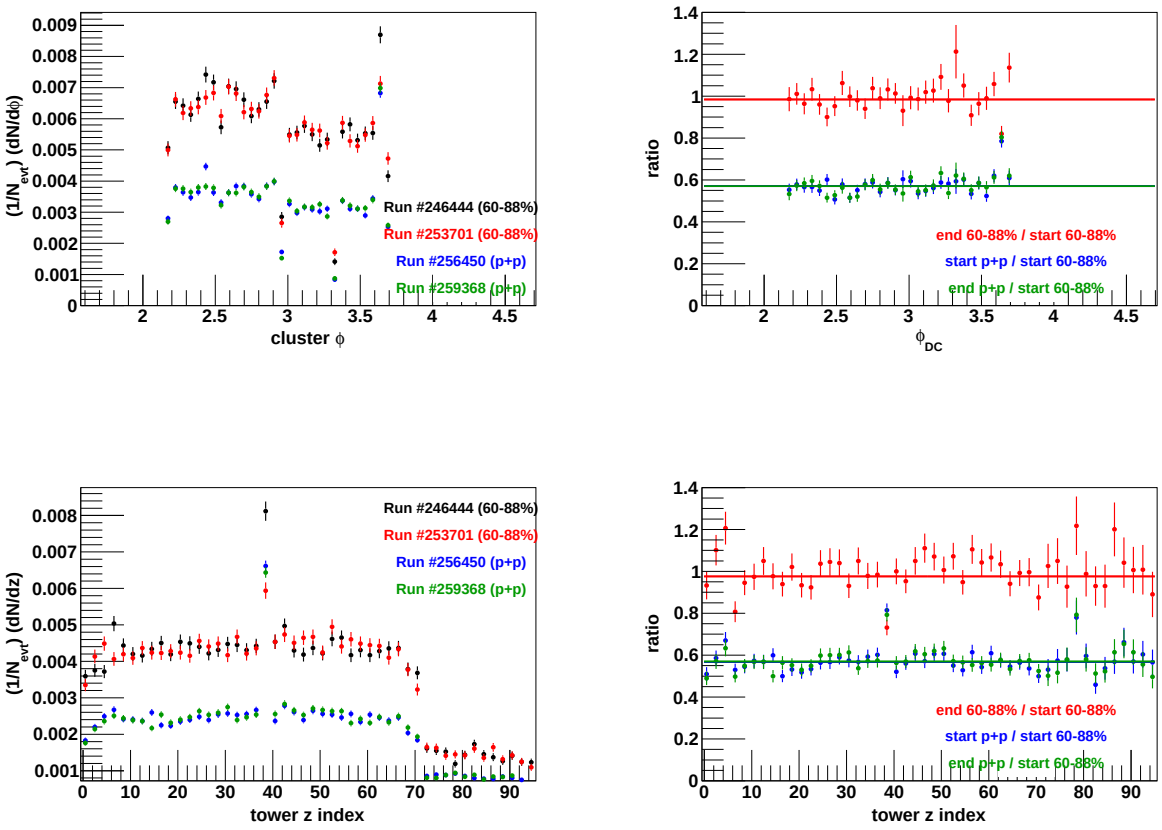


Figure 6.6: Differential cluster yields, $1 < p_{\mathrm{T}} < 2$ GeV/c, in minimum bias events per ϕ (top plots) and tower z position (bottom plots). Differential yields are shown in the left plots, common ratio to the run #246444 reference are shown on the right.

not used in the final data. The true measure of stability is to examine the run dependence of the per-event *jet* instead of particle yields, in Section 6.3.4.

6.1.5 Jet selection cuts

Reconstructed jets are also required to pass quality cuts that reject "fake" jets from combinatoric d+Au particles and those arising from high-p_{T} background in the charged tracking. Some of the the cuts are defined in terms of jet-level variables used to describe reconstructed jets.

The *charged fraction* of a gaussian filter jet is the fraction of the jet's (gaussian weighted) p_{T} which comes from charged particles.

$$c.f. = (1/{p_{\mathrm{T}}}^{\mathrm{jet}})\ \Sigma_{tracks}\ {p_{\mathrm{T}}}^{\mathrm{track}} \cdot \exp\left(-\Delta R^2_{\mathrm{jet,track}}/2\sigma^2\right) \tag{6.7}$$

where σ is the Gaussian filter algorithm parameter.

The *max z* of a gaussian filter jet is the largest fraction of its (gaussian weighted) p_{T} which comes from any single track or cluster.

$$\max\ z = (1/{p_{\mathrm{T}}}^{\mathrm{jet}})\ \max_{\mathrm{particles}}\left({p_{\mathrm{T}}}^{\mathrm{particle}} \cdot \exp\left(-\Delta R^2_{\mathrm{jet,particle}}/2\sigma^2\right)\right) \tag{6.8}$$

The *number of constituents* of a gaussian filter jet is the (gaussian weighted) number of > 400 MeV particles associated with the jet.

$$n.c. = \Sigma_{particles}\ 1 \cdot \exp\left(-\Delta R^2_{\mathrm{jet,particle}}/2\sigma^2\right) \tag{6.9}$$

Additionally, it is possible to define an integral number of constituents variable which is more analogous to that used in algorithms like anti-k_{T}, in which particles are exclusively associated with ("are owned by") one jet in particular.

$$n.c.^{\mathrm{int}} = \Sigma_{particles}\ 1 \cdot \theta\left(0.3 - \Delta R\right) \tag{6.10}$$

To cut down on jets reconstructed from combinatoric particles, jets must have a minimum number of constituents,

- $n.c. > 2.5$ (gaussian weighted)

Note that since the number of constituents are gaussian-weighted (e.g. each particle necessarily contributes < 1 to the sum), a minimum of three or more real particles are required to produce a weighted value of 2.5.

When using the integral number of constituents, this cut is instead,

- $n.c. \geq 3$ (integral)

Because of the presence of charged, high-p_{T} jet background coming from conversions and late decays that give rise to stiff tracks with an erroneously high measured p_{T}, we must use additional cuts at the jet-level.

This background is particularly pernicious in the Minimum Bias sample (the requirement of calorimeter energy in the ERT data supresses it somewhat), which is used to construct the ERT efficiency (Section 6.3.1), since any residual contamination will be treated as "inefficiency" and erroneously corrected for. The rate of high-p_{T} charged background is also higher in d+Au collisions than in pp collisions because of the additional amount of soft photon production that might cause a conversion.

Fortunately, the background is highly localized in the jet fragmentation space defined by the charged fraction and max z. Figure 6.7 shows the distribution of low- and high-p_{T} jets in $c.f.$ / max z space. Background jets are dominated by one very high-p_{T} track ($c.f. = \text{max } z \approx 1$) or by two high-$p_{\mathrm{T}}$ tracks with little other energy ($c.f. \approx 1$, $0.5 > z > 1$).

Three different jet-level cuts to reject background, shown in Figure 6.7, were defined and evaluated:

- most restrictive: cut out jets with $c.f. > 75\%$
- moderately restrictive: cut out jets with $c.f. > 85\%$ or ($c.f. > 75\%$ and max $z > 75\%$)
- least restrictive: cut out jets with $c.f. > 90\%$ or ($c.f. > 75\%$ and max $z > 75\%$)

Figure 6.8 shows the spectrum (in counts / GeV) of reconstructed jets in pp from both datastreams. Before any jet-level cuts (the green points), the spectrum does not obey a power-law behavior and has a kink around 25 GeV (where the maximum p_{T} cut on the charged tracks is causing a change in behavior). The spectra after each of the possible cuts is shown, and they exhibit much better behavior, with no artifical behavior through 25 GeV. Figures 6.9 and 6.10 show similar plots for the different centrality selections in d+Au MinBias and d+Au ERT streams, respectively.

After the implementation of the new ghost cuts in the tracking, there is no evidence of residual background using the least restrictive cut. Thus, jets are required to pass the following cut,

- cut out jets with $c.f. > 90\%$ or ($c.f. > 75\%$ and max $z > 75\%$)

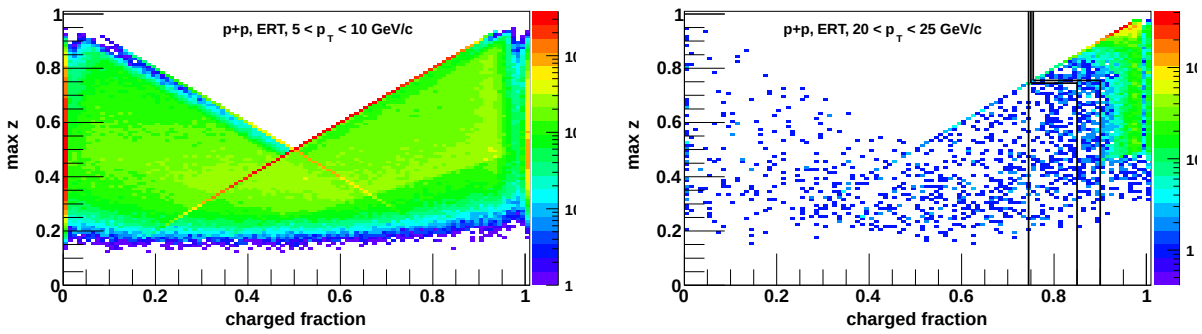


Figure 6.7: Demonstration of high-p_{T} charged background in triggered pp data. At low p_{T} (left), the fragmentation distribution shows no regions of high background. At high p_{T} (right), the effect of conversions and artificial high-p_{T} tracks can be seen in the upper right of the distribution. The three proposed jet-level cuts are shown as black lines on the right.

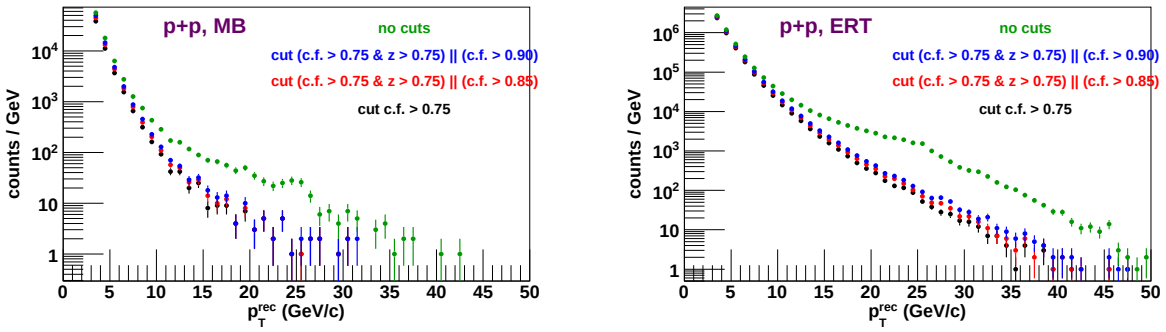


Figure 6.8: Reconstructed jets (in counts per GeV) from the pp datasets (left, MinBias and right, ERT). The effects of each of the three proposed jet-level cuts are shown.

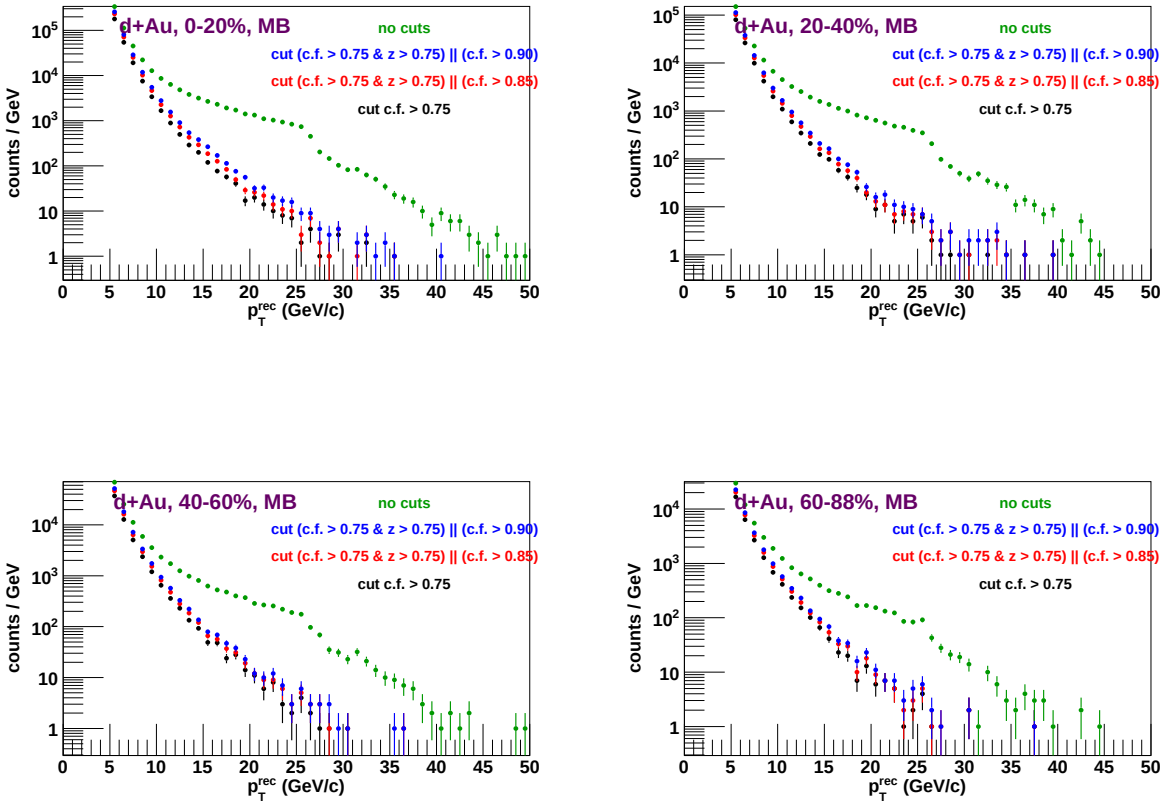


Figure 6.9: Reconstructed jets (in counts per GeV) from the d+Au MinBias stream. The effects of each of the three proposed jet-level cuts are shown.

The final jet-level cut is a fiducial cut which removes jets near the edge of the acceptance.

From Monte Carlo, we see that the jet reconstruction peformance for jets near the edge of the PHENIX acceptance is very bad, since many articles from the jet are outside the acceptance. Figure 6.11 shows the difference between truth and reconstructed jet p_{T} for jets which are $\Delta\eta, \Delta\phi > 0.05$ away from the edge of the detector compared with those that are not. Fiducial jets have a good correlation between the truth and reconstructed p_{T}, whereas the energy scale for edge jets is almost completely uncorrelated. Additionally, Figure 6.12 shows that that the reconstruction efficiency in the edge region drops 20% relative to that in the fiducial region.

Due to this poor performance, we require that reconstructed jets are at least 0.05 in pseudora-

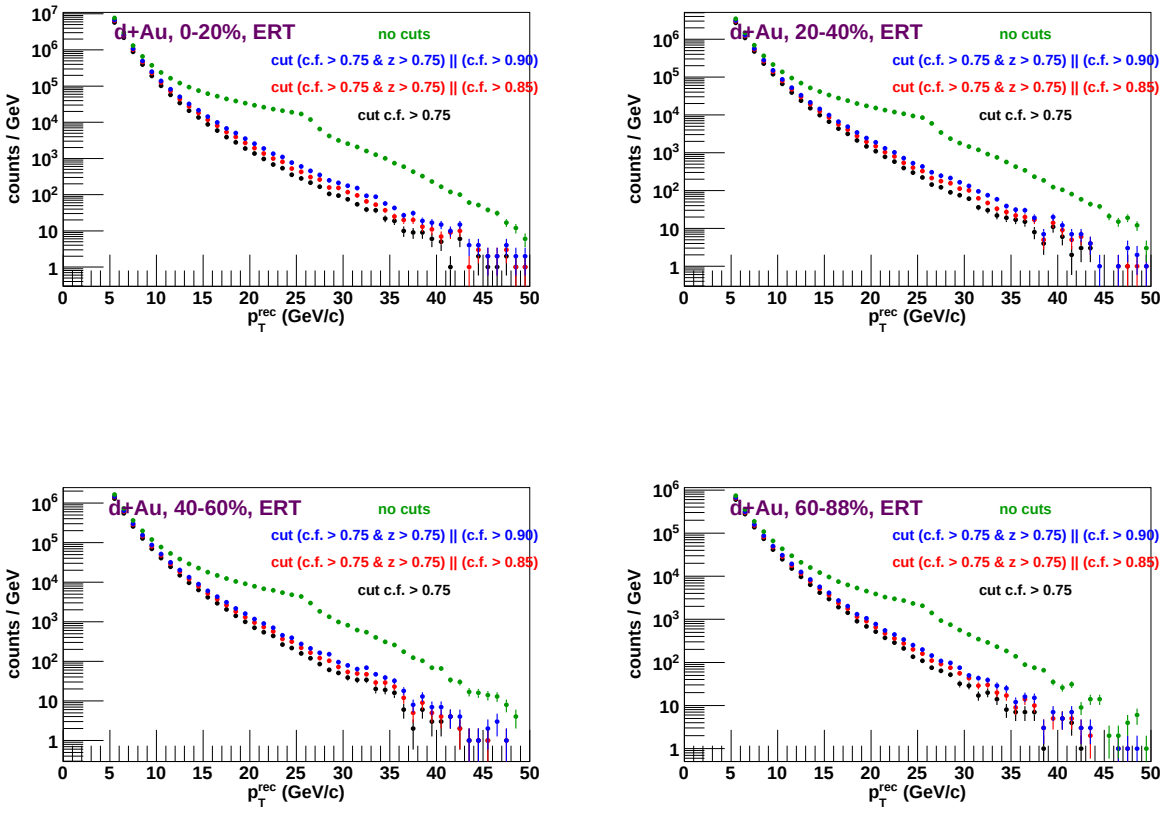


Figure 6.10: Reconstructed jets (in counts per GeV) from the d+Au MinBias stream. The effects of each of the three proposed jet-level cuts are shown.

pidity and $\frac{\pi}{60} \simeq 0.05$ in azimuth away from the edge of the PHENIX acceptance.

Thus, using Equations 4.2 and 4.3 for a given z_{vtx}, reconcstructed jets are required to lie within a restricted phase space given by:

- $\eta \in (\eta^{-}(z_{vtx}) + 0.05, \eta^{+}(z_{vtx}) - 0.05)$
- $\phi \in (\frac{42}{60}\pi, \frac{70}{60}\pi)$, in the East arm

The total phase space for reconstructed jets is therefore

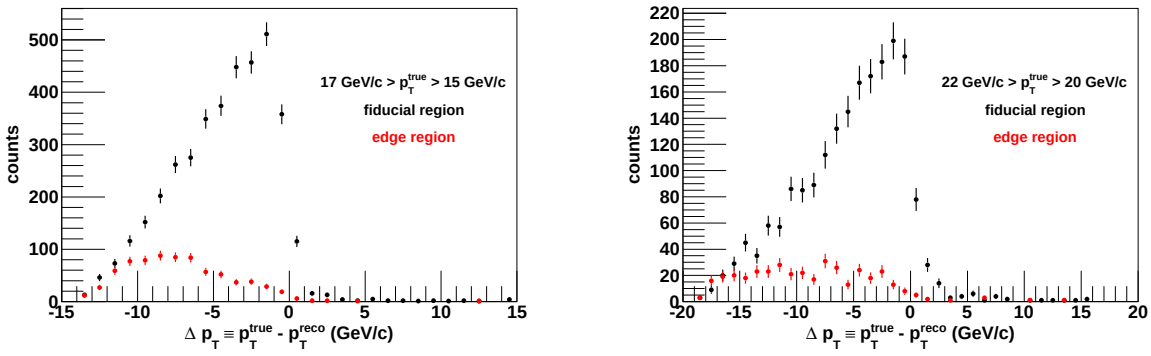


Figure 6.11: Δp_{T} between truth and reconstructed jets, comparing truth jets in the fiducial region vs. jets near the edge of the acceptance, two different ${p_{\mathrm{T}}}^{truth}$ selections.

$$\Delta\eta \times \Delta\phi = (0.715 - 2 \times 0.05) \times \left(\frac{30}{60}\pi - 2 \times \frac{\pi}{60}\right) = 0.615 \times \frac{28}{60}\pi \simeq 0.902 \tag{6.11}$$

6.1.6 Fake jet rejection capability

The aim of jet reconstruction algorithms is to reconstruct the kinematics of a fragmenting high-p_{T} parton that has undergone a hard scattering. However, combinatoric collections of soft particles from the heavy ion underlying event can be reconstructed as jets. These "fake jets" appear as a low-p_{T} excess in the reconstructed jet spectrum and are more prevalent in more central events. In 0-20% events, they can be seen as an excess of jets with $p_{\mathrm{T}} < 10$ GeV well above N_{coll} scaling. For a variety of reasons, it is not tractable to statistically subtract a yield of "fake jets" (partially because this method relies on Monte Carlo generators being able to characterize correlations in the underlying event that result in fake jets).

Instead, there is a way to extend the idea of the Gaussian filter in a way to preferentially select real jets, on a jet-by-jet basis. This method was first developed for use in Cu+Cu collisions at PHENIX, where the fake rate is substantially larger due to the larger underlying event[47]. It introduces an inefficiency to reconstructing real jets, but provides a high (quantifiable) rejection against fake jets.

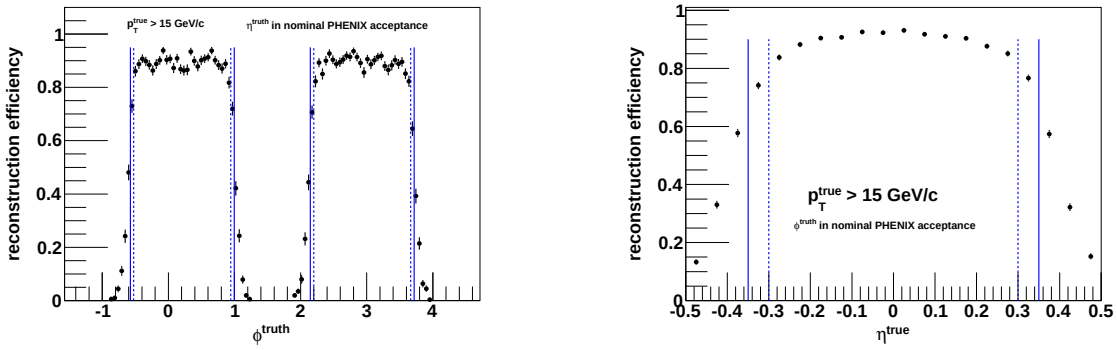


Figure 6.12: Reconstruction efficiency for $p_{\mathrm{T}} = 15$ GeV jets, as a function of ϕ (left) and η (right) of the truth jets. The solid lines are the nominal PHENIX acceptance and the dotted lines are the fiducial acceptance.

For each jets, we compute the discriminant g, defined by

$$g = \Sigma_{\mathrm{particle}\ i}\ \left({p_{\mathrm{T}}}^{i}\right)^2 \exp\left(-\Delta R^2_{\mathrm{jet,particle}}/2\sigma^2_{\mathrm{dis}}\right) \tag{6.12}$$

The parameter $\sigma_{\mathrm{dis}} = 0.1$ is chosen to be somewhat smaller than the Gaussian filter parameter σ. This discriminant gives a high value to jets with harder fragmentation kinematics (because of the ${p_{\mathrm{T}}}^2$ weighing) and to jets with an angularly tight "core" of particles (because $\sigma_{\mathrm{dis}} < \sigma$). Conversely, jets reconstructed from a combination of spread out, low-p_{T} particles, which are very likely to be fake jets, have a low value of g. In the PHENIX Cu+Cu reconstructed jet analysis[47], it was found that a lower power (${p_{\mathrm{T}}}^1$) does not have enough discriminating power to separate fake jets from real jets, and that a higher power (e.g. ${p_{\mathrm{T}}}^3$) does not significantly improve the rejection vs. efficiency tradeoff.

Originally, two possible fake jet cuts were tested in this analysis.

- $g > 5$ GeV^2
- $g > 10$ GeV^2

However, the fake rate in d+Au collisions was found to be negligible above $p_{\mathrm{T}} > 12$ GeV, with

the fake jet rejection cut only introducing additional inefficiency. Thus, an upper limit of 5% was set on the residual fake rate between 9 GeV $< p_{\mathrm{T}} <$ 12 GeV using another method, described in the next section.

6.1.7 Limit on residual fake rate

Although the West arm is not usable for the final R_{dAu} and R_{CP} results due to instability during the pp part of the run, we can use the start of the d+Au run (as described in Section 6.2.1 section) for a reconstructed di-jet analysis through which we can determine the fake rate before and after fake jet rejection. This is possible because the $\Delta\phi$ distribution of "fake jets" should look combinatoric while that for real jets should exhibit a back-to-back shape with a narrow width around $\Delta\phi = \pi$. For this study, it is assumed that the energy balance of real dijets is the same in each d+Au centrality – that is, there is no impact parameter dependent angular decorrelation within the sensitivity of this study.

This is done by dijet $\Delta\phi$ distributions in data into components that look like real dijets (called the "signal", obtained from Monte Carlo) and a combinatoric part (constructed from a mixed event method). Only East arm-West arm dijets are considered, and both jets are required to pass all cuts. Furthermore, East-arm jets are required to fire the ERT trigger, so that they have the same jet definition as the data. The $\Delta\phi$ distribution in constructed for dijets with jets in each arm binned in four p_{T} ranges of interest (5-6 GeV/c, 6-7 GeV/c, 7-9 GeV/c and 9-12 GeV/c), and for each d+Au centrality. Then, all $\Delta\phi$ data are normalized to 1, so the template fitting can be expressed in terms of fractions of the total per-pair $\Delta\phi$ yield.

The dijet signal distributions are constructed from $\sqrt{Q^2}$ = 6, 8 and 12 GeV/c Monte Carlo subsamples as described in Section 6.4.1 and combined according to their generator cross-section weighing. Separate $\Delta\phi$ are made for each possible dijet flavor (q+q dijet, q+g dijet, g+g dijet). The three separate QCD signal distributions are used to give the template fitting method more flexibility in describing different mixtures of quark and gluon jets in data than is generated in the MC. (We do not, at this stage of the analysis, propose that the template fitting method can actually separate subtle differences in quark vs gluon jet contributions).

The combinatoric $\Delta\phi$ distributions are generated through a mixed event method. One jet at random from each event in data is used, and is randomly combined with a jet in the opposite arm

from an event of the same centrality. Any individual reconstructed jet is used only once in the event mixing. This is done separately for each centrality.

The template fit function models the reconstructed jet distribution in data as the sum of combinatoric and signal jet components. Formally, the model is

$$\frac{dN}{d\phi}_{model} = f_0 \cdot \frac{dN}{d\phi}_{combinatoric} + f_1 \cdot \frac{dN}{d\phi}_{MC,q+q} + f_2 \cdot \frac{dN}{d\phi}_{MC,q+q} + f_3 \cdot \frac{dN}{d\phi}_{MC,g+g} \tag{6.13}$$

$$f_0 + f_1 + f_2 + f_3 = 1 \tag{6.14}$$

$$0 \leq f_0, f_1, f_2, f_3 \leq 1 \tag{6.15}$$

Where the second to last line is a constraint from the forced normalization of the individual $\Delta\phi$ distributions. To parameterize these constraints, we use the fit parameters $c_0, c_1, c_2 \in (0, 1)$ where

$$f_0 = c_0 \tag{6.16}$$

$$f_1 = (1 - c_0) \cdot c_1 \tag{6.17}$$

$$f_2 = (1 - c_0) \cdot (1 - c_1) \cdot c_2 \tag{6.18}$$

$$f_3 = 1 - f_0 - f_1 - f_2 \tag{6.19}$$

That is, c_0 is "the fraction of the yield from combinatoric", c_1 is "the fraction of the non-combinatoric yield that is q+q QCD jets" and so forth.

The template function is fit to the $\Delta\phi$ distributions in data in each centrality and p_{T} bin.

The template fitting results for $p_{\mathrm{T}}^{East} \in 9 - 12$ GeV/c $\otimes$ $p_{\mathrm{T}}^{West} \in 9 - 12$ GeV/c are shown in Figure 6.13. In general, the template fits do a good job of describing the data, with some imperfections in regions of low PHENIX dijet acceptance (at $\Delta\phi \sim 2.5, 3.7$). Additionally, we see that in general only one or two of the QCD dijet templates dominate the distribution. This implies that do not have sufficient control to separate the peak into qq, qg and gg components (although we can still separate it more generally into QCD and combinatoric components).

Then, the residual fake rate for 9-12 GeV jets is estimated in two ways.

In the first method, we assume a simple model where the jet yield from a symmetric $p_{\mathrm{T}}^{East} \otimes p_{\mathrm{T}}^{West}$ bin in either arm has a real jet purity r and fake jet contamination fraction f (such that r+$f = 1$). If these distributions were independent in each arm, the real component of the $\Delta\phi$ distribution should

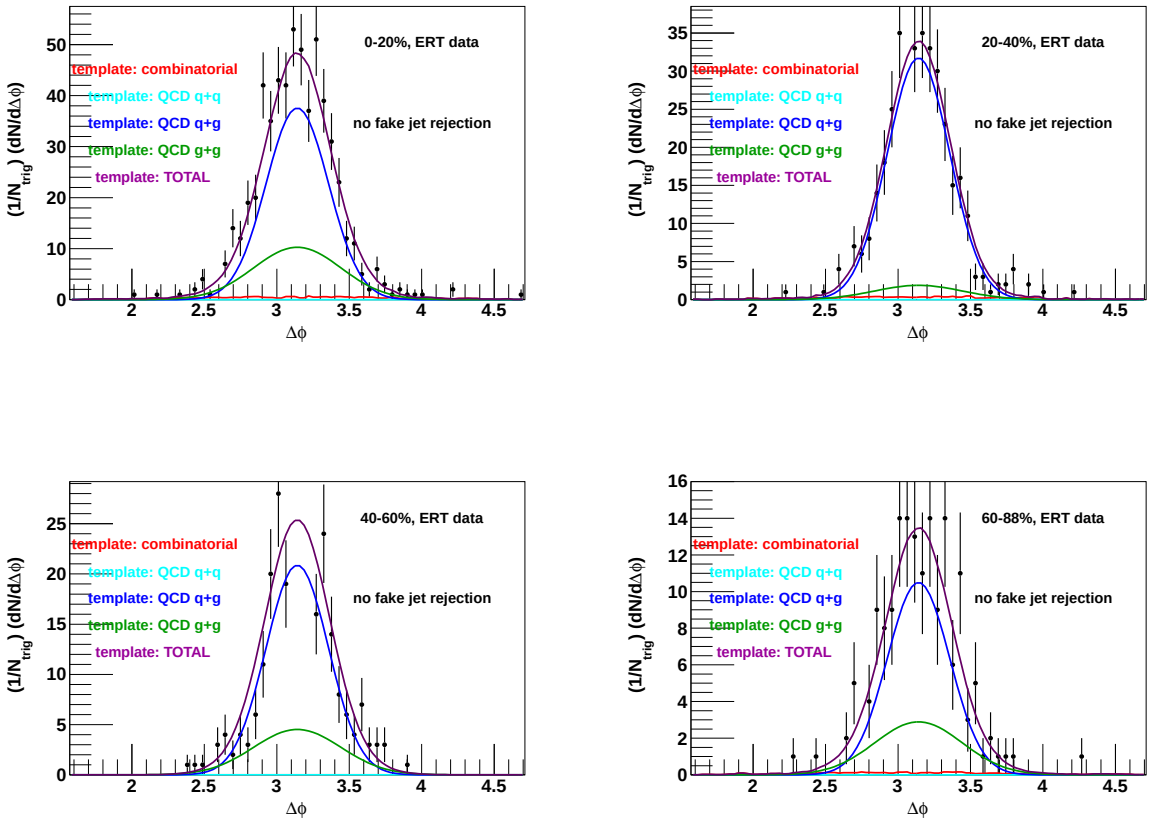


Figure 6.13: Template fits to dijet $\Delta\phi$ data. East and West arm jets both have $9 < pt < 12$ GeV/c. Each centrality is shown separately.

have a fraction r^2 (since only two real jets will give a back to back signal), and the combinatoric component a fraction $2rf + f^2$ (since a fake jet and a real jet, as well as two fake jets, should be uncorrelated). However, selecting a trigger jet in one arm biases the distribution in the other arm towards real jets. At worst, we can assume an away side jet is *always* reconstructed given a real near-side jet. In this formulation, the real signal should take up a fraction $r \cdot 1$ instead of r^2, and extracting it from the template fits gives a lower bound on the purity and an upper bound on the fake rate.

The template fitting results for the $9 - 12$ GeV/c bin in Figure 6.13 give the dijet signal (variously) as $95 \pm 2\%$ to $100 \pm 1\%$ of the distribution, depending on the centrality. Thus, an upper

	$p_{\mathrm{T}}{}^{asoc}$ (West arm) (GeV/c)				
centrality	$5-6$	$6-7$	$7-9$	$9-12$	total
0-20%	0.152(8)	0.124(6)	0.138(7)	0.107(6)	0.521
20-40%	0.160(10)	0.112(8)	0.141(9)	0.106(7)	0.519
40-60%	0.142(12)	0.122(10)	0.141(11)	0.128(9)	0.533
60-88%	0.141(13)	0.115(11)	0.152(14)	0.114(11)	0.522

Table 6.5: Per-trigger away-side yields as given by the real component of a template fit, for $p_{\mathrm{T}}{}^{East} \in 9-12$ GeV/c trigger jets.

limit of 5% is derived.

The second method is to use the peripheral bin to calibrate what the away-side per-trigger yield in central collisions should be, given 100% purity, and compare it to the real result. A smaller per-trigger yield in 0-20% events than in 60-88% implies an impure sample of trigger jets in the central bin.

Table 6.5 shows the away-side per-trigger jet yields for trigger jets in the East arm with 9-12 GeV in different centralities, as extracted using the template fitting. Although there are statistical (and systematic, from the fitting procedure) fluctuations in any given $p_{\mathrm{T}}{}^{West}$ bin or centrality, the associated-p_{T}-averaged yields are consistent between the different centralities. Thus, this method is not sensitive to the rsidual fake rate within statistics, which are of order 5-10%.

Between these two methods, we assign a systematic uncertainty of 5% on the fake rate for $9-12$ GeV jets.

6.2 Data selection

6.2.1 Run QA

Out of the 886 recorded runs in 2007-2008 $\sqrt{s_{NN}} = 200$ GeV d+Au running, data from the central arms was read out in 847. In the $\sqrt{s} = 200$ GeV pp running, 267 runs were included in the central arm production, but run 259465 has only an ERT stream without a corresponding MinBias stream. Therefore we only use the other 266.

An aggressive run quality procedure was applied to all runs, consisting of requirements of good centrality behavior (d+Au only), uniform acceptance in the drift chamber and good trigger behavior.

First, d+Au runs have to have a good centrality distribution. By plotting the fraction of Minimum Bias events in a given centrality category for each run, we determine that the following eleven runs have a non-flat centrality distribution:

- 249127, 249128, 249129, 249130, 249131, 249132 (adjacent runs)
- 249467
- 249867, 249868, 249869, 249870 (adjacent runs)

These same runs were noted to be bad in the first determination of d+Au centrality in Run 8[238], and are excluded.

Next, runs must have a uniform tracking acceptance in the East arm. Because good jet reconstruction capability requires a large and uniform acceptance (and because all of the data needs to have the same acceptance), we discard runs in which the there are temporary holes in the PC1 acceptance or broken wires in the DC.

In the d+Au runs, the 44 excluded runs (along with a description of the acceptance loss) are:

- 248165, 248166, 248167, 248170, 248171, 248171, 248174, 248175, 248177, 248178, 248182, 248183, 248184, 248287, 248288, 248289, 248291, 248293, 248295, 248298, 248302, 248304, 248305, 248308 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 2.3$ strip)
- 248353, 248355, 248359, 248363, 248364, 248365, 248367, 248368, 248380, 248382, 248383 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.5$ strip)
- 249858 ($z_{\rm DC} < 0, \phi_{\rm DC} \sim 3.1$ strip)
- 250837, 250838 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.5$ strip)
- 251857 (X2 wire at $\phi_{\rm DC} \sim 2.8$)
- 252123 (X2 wire at $\phi_{\rm DC} \sim 3.1$)

- 252127 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.5$ strip)
- 252140 ($z_{\rm DC} < 0, \phi_{\rm DC} \sim 3.1$ strip, X1 wire at $\phi_{\rm DC} \sim 3.3$))
- 252216 (X1 wire at $\phi_{\rm DC} \sim 3.3$)
- 253428 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.1$ strip)

Additionally, the X1 wire at $\phi_{\rm DC} \sim 2.5$ is broken for an extended period of time from runs 251969 to 252262. Due to the long duration, we integrate this region into the X1 broken-wire masks in Section 6.1.3 instead of excluding the runs.

In the *pp* runs, the 47 excluded runs (along with a description of the loss of acceptance) are:

- 257080, 257083, 257084, 257085 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.1$)
- 257473, 257477, 257478, 257479, 257480 ($z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.5$)
- 258260, 258262, 258263, 258264, 258265, 258266, 258267 (thin strip at $z_{\rm DC} < 0, \phi_{\rm DC} \sim 3.4$)
- 259369, 259370, 259371, 259372, 259474 (strips at $\phi_{\rm DC} \sim 2.3, 2.5, 2.7, 2.9$)
- 259466, 259467, 259470, 259472, 259476, 259488, 259489, 259491, 259492, 259493, 259499 ($\phi_{\rm DC} \sim 2.5$, additional wide strip at $z_{\rm DC} > 0, \phi_{\rm DC} \sim 3.5$ from runs 259488-259499; additionally, run #259465 would be excluded at this stage, were it not already excluded due to a different problem above)
- 259556, 259557, 259560, 259561, 259562, 259563, 259564, 259565, 259567, 259568, 259569, 259570, 259571, 259572, 259575 ($\phi_{\rm DC} \sim 2.5$)

Because the West arm drift chamber acceptance changes strongly and in unpredictable ways from run to run in the *d*+Au part of the run and because of a significant loss of uniformity of acceptance before the *pp* part of the run, the West arm is not used for the final results. However, the West arm can be used for the di-jet $\Delta\phi$ analysis in Section 6.1.7, which provides an independent data-driven estimate of the fake rate as a function of $p_{\rm T}$ in different *d*+Au centralities. Runs used for this study must pass the good run list based on the criteria above, and are only used until run #249437, after which PC1 acceptance problems in the West arm begin to rapidly develop (and unfortunately, the West arm is not usable at all in the *pp* runs).

Furthermore, runs must have a consistent and well understood trigger performance. The ++ part of Run 8 d+Au is characterized by changing ERT super module masks[239] and rejection factors[240] in the East arm[239]. In addition, several runs experienced intermittent problems with the ERT East low voltage, as reported in the PHENIX electronic logbook (ELOG)[241]. Thus, although the full trigger performance in this part of the run is still being evaluated, the d+Au -- runs, where the trigger performance is known to be stable, are used for the final results. This has the additional benefit of having a consistent field configuration between the d+Au and pp runs. Thus, only runs starting with run #250593 are used.

Finally, the Run 8 pp data contain several runs in which additional converting material was inserted around the beampipe. Because of the increased jet background from conversions, the following eighteen converter runs are excluded,

- 257841, 257842, 257844, 257848, 257850, 257851, 257920, 257927, 257928, 257929, 257930, 257931, 257936, 257937, 257939, 257940, 257947, 257948

The final data selection corresponds to 46×10^9 minimum bias d+Au events (approximately $L_{int} = 23$ nb^{-1}) and 4.7×10^9 pp events (approximately $L_{int} = 0.2$ pb^{-1}).

6.2.2 MB event requirements

The online BBC trigger is described in Section 4.2.1. The following tighter offline selection is applied to BBC L1-firing events:

- good BBC timing: $t_0 > -10$ ns
- narrow BBC-measured vertex position: -20 cm $< z_{vtx} < 20$ cm
- (for d+Au) good centrality value: $0 < cent < 88$

6.2.3 Centrality determination in d+Au

In d+Au collisions, the total charge in the Au-going BBC (BBC South in the 2008 d+Au run) is used to define the event centrality[242; 243].

The Glauber model (see Section 2.2.1) with nucleon-nucleon cross-section $\sigma_{NN} = 42$ mb and the default parameters for modeling the nucleon probability density in the deuteron and Au nucleus

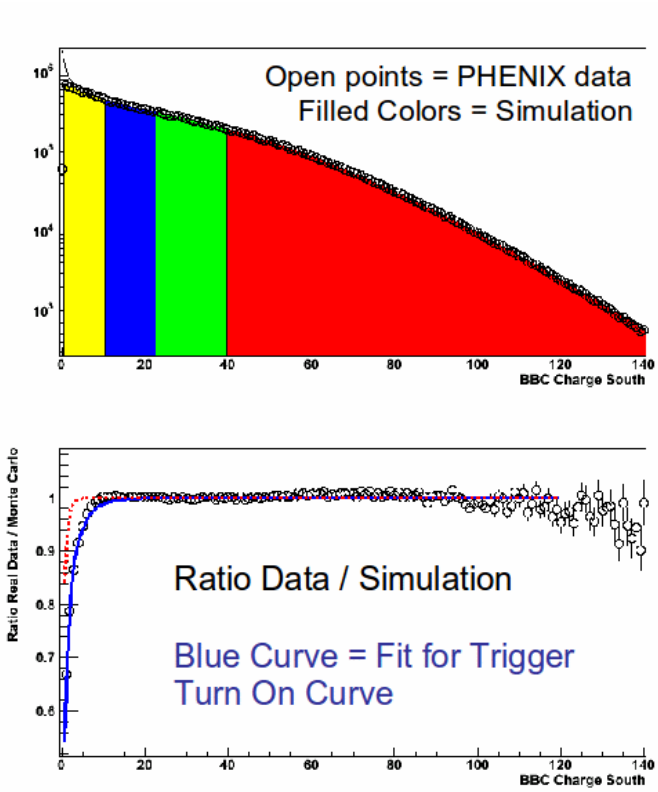


Figure 6.14: BBC South Charge distribution in data with the 0-20%, 20-40%, 40-60% and 60-88% centrality selections in red, green, blue and yellow (top), and ratio of the BBC South Charge distribution in data to the total Glauber distribution, showing the effect of the trigger eficiency, from [242].

is used to generate a set of Glauber d+Au events, classified by their $N_{\rm part}$ and $N_{\rm coll}$. The signal in the BBC from an event with given $N_{\rm coll}$ is assumed to be a negative binomial distribution (Equation 2.74) with parameters $k = k_0 N_{\rm coll}$ and μ. The distribution in data is fit with the Glauber $N_{\rm coll}$ distribution $dN/dN_{\rm coll}$ convolved with the BBC signal $P(\text{BBC Charge};\ N_{\rm coll})$, with the fit restricted to the > 20 region where the trigger efficiency is 100%.

Figure 6.14 shows the distribution in data and the best fit Glauber distribution. It is observed that for low values of the BBC Charge, the model overpredicts the data. This is due to the fact that the PHENIX BBC trigger is inefficient for the peripheral part of the d+Au cross-section. Comparing the model and data at low charge yields a trigger efficiency of $88 \pm 4\%$ (that is, the

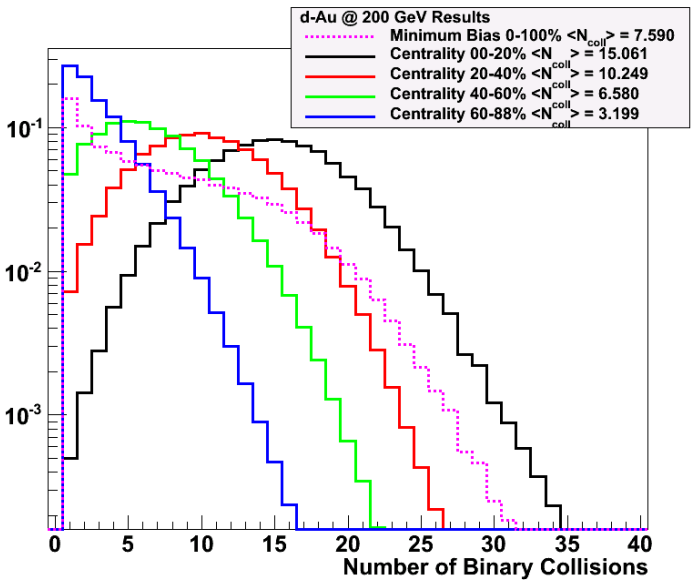


Figure 6.15: Distribution of $N_{\rm coll}$ values in the different centrality selections according to the Glauber model, from [242].

BBC fires on this fraction of the inelastic d+Au cross-section). Thus, the data is partioned into four centrality categories consisting of the highest-$^{20}/_{88}$ths of the data, called "0-20%", and all the way down to the lowest-$^{28}/_{88}$ths of the data, called "60-88%". Then, the mean $N_{\rm coll}$ and $N_{\rm part}$ are taken from the corresponding selection in the Glauber distribution. The distribution of $N_{\rm coll}$ values in each centrality bin is shown in Figure 6.15.

The final systematic uncertainties on the mean $N_{\rm coll}$ (and other geometric quantities) are derived from a number of sources, including variations in the Glauber MC parameters as well as repeating the fit procedure with the mean BBC Charge parameterized in terms of $N_{\rm part}$ instead of $N_{\rm coll}$.

The mean number of binary collisions in each centrality category are given by

- $N_{\rm coll} = 15.061 \pm 1.013$ for 0-20% collisions
- $N_{\rm coll} = 10.248 \pm 0.704$ for 20-40% collisions
- $N_{\rm coll} = 6.579 \pm 0.444$ for 40-60% collisions
- $N_{\rm coll} = 3.198 \pm 0.193$ for 60-88% collisions

- $N_{\rm coll} = 7.590 \pm 0.427$ for 0-88% (inclusive d+Au) collisions

Finally, when measuring the centrality-dependent yields of hard processes, there is one more factor that must be taken into account. It is known that high-Q^2 processes such as the production of a jet, heavy quark or heavy boson, occur preferentally in events with a larger multiplicity. Thus, a *centrality bias* can be introduced in which jet events can be assigned an erroneously more central centrality due to the increased multiplicity. Thus, selecting on central events according to the BBC charge will preferentially select a larger-than-$N_{\rm coll}$ sample of jets, and selecting on peripheral events according to the BBC charge will result in the opposite effect.

This effect has been investigated in PHENIX pp 200 GeV data, and the additional mean BBC charge in jet events used to calculate derive a correction to the yield which takes this into account. These correctios are called bias factors (BF) in PHENIX terminology and are equivalent to using a modified $\langle N_{\rm coll} \rangle$ which takes into account the difference in $P(\text{BBC Charge}; N_{\rm coll})$ between events with and without a hard-scattering. The bias factors are a multiplicative factor applied to the jet yield and are given by,

- $BF = 0.941 \pm 0.010$ for 0-20% collisions
- $BF = 1.000 \pm 0.006$ for 20-40% collisions
- $BF = 1.034 \pm 0.017$ for 40-60% collisions
- $BF = 1.031 \pm 0.055$ for 60-88% collisions
- $BF = 1$ for 0-88% collisions (by construction)

In fact, there has been further work investigating the full $p_{\rm T}$ depedence of the bias factors in the Glauber fit and in HIJING events[244], but the results here include the original d+Au centrality recommendation, which are listed above.

In Au+Au and other A+A collisions, the combined information from the BBC and ZDC can provide a better experimental handle on the geometry of the event, since the spectator neutrons measured in the ZDC can provide additional discriminating information. In d+Au collisions, the ZDC North (e.g. the deuteron-going direction) can be examined for a spectator neutron signal, and the resulting events tagged as p+Au collisions. This technique is used as a cross-check on the Glauber results, as well as explored as a cross-check on the d+Au results in Section 6.6.2.

6.2.4 Energy scale stability

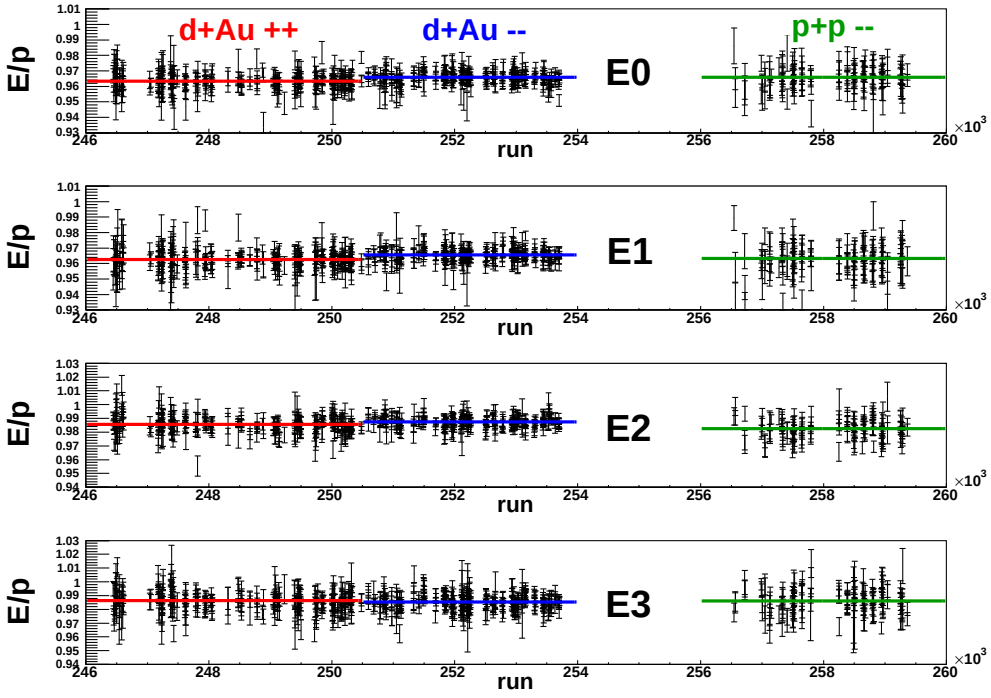


Figure 6.16: Run dependence of E/p peak position for selected electrons. The four East arm sectors are shown. The mean position for the d+Au ++, d+Au −− and pp runs are shown in red, blue and green, respectively.

The selected d+Au and pp runs were examined for the stability of the relative energy scale between the tracking and calorimeter subsystems. The energy scale is calibrates in data by selecting electrons with the DC and RICH which deposit energy in the calorimeter and comparing the ratio of the reconstructed electron energy and momentum E/p, which in PHENIX ≈ 0.98.

For each of the good runs above, electrons in the ERT dataset are selected according to:

- quality > 7
- $n_0 \geq 3$ (at least three PMT towers fired in the RICH)

- $p_{\mathrm{T}} > 1.5$ GeV/c

These cuts are required to cut down on the background in the E/p distribution (although we have redone this analysis with stricter and looser cuts, and the results appear to be consistent). The E/p distributions are calculated for each run and each sector in the East arm. If the $\frac{E}{p} \in (0.7, 1.2)$ region has more than 200 counts, we fit it with a gaussian signal plus exponential background:

$$\frac{dN}{d(E/p)} = f\left(\frac{E}{p}\right) = c_0 e^{-(\frac{E}{p}-c_1)^2/2c_2^2} + c_3 e^{-c_4 \frac{E}{p}} \tag{6.20}$$

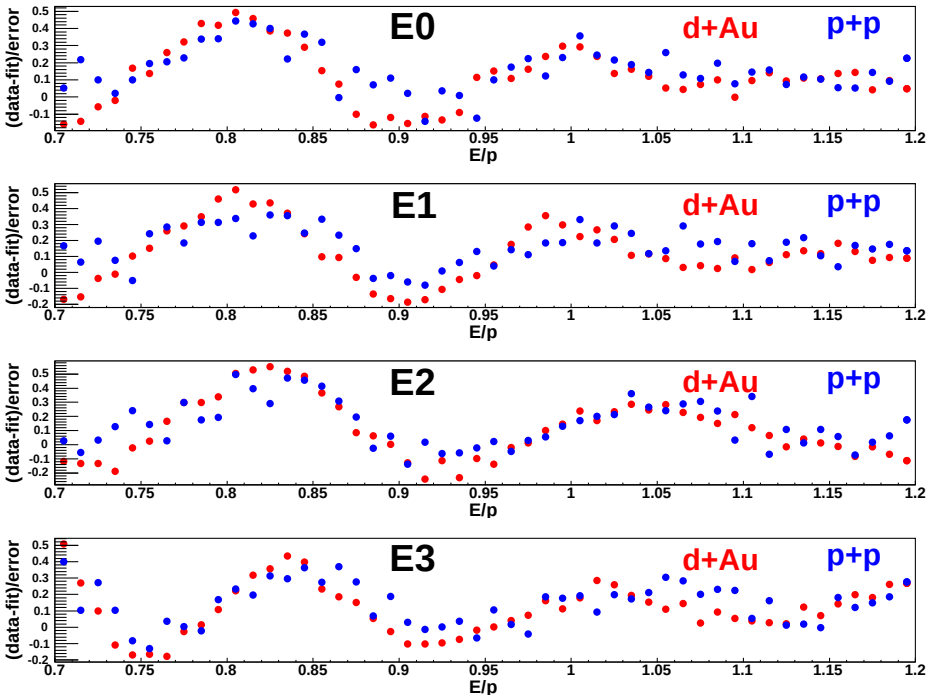


Figure 6.17: Mean pull (= (data − fit)/error) of E/p fits in the d+Au (red) and pp (blue) run for each sector.

If there are fewer than 200 counts, we do not include the run in this study. The full run dependence of the E/p peak position are shown in Figure 6.16. To ensure that the fit does an

adequate job of modeling the distribution, we keep track of the residual pull (= (data − fit)/error) of the fit and plot the mean pulls for each sector in Figure 6.17. The pattern of residuals is common to d+Au and pp and all sectors. The typical pull at $E/p \sim 0.98$ is 0.1, indicating that we are able to model the distribution very well in that region.

sector	d+Au ++	d+Au −−	pp
E0 (PbGl)	0.9634(2)	0.9659(1)	0.9657(4)
E1 (PbGl)	0.9626(2)	0.9656(1)	0.9633(5)
E2 (PbSc)	0.9858(2)	0.9874(1)	0.9825(4)
E3 (PbSc)	0.9864(2)	0.9853(2)	0.9862(4)

Table 6.6: Summary of mean E/p peak position per sector in different run groups.

The mean E/p peaks for the East arm sectors in each major run group are listed in Table 6.6. We can see that after run selection, the energy scale is stable with only a 0.5%-level residual change in the E/p position throughout the run.

6.3 Trigger selection

The results in this work are constructed from the ERT-selected data (see Section 4.2.2) with jets in d+Au measured with the 4x4c trigger, which has the lowest ADC energy threshold[239]. In d+Au running, the next lowest thresholds are the 4x4a and 4x4b triggers, but since their efficiency curves saturate more slowly, they are only used for cross-checks.

Between the d+Au and pp parts of Run 8, the ERT thresholds were lowered for all triggers. Furthermore, all 18 supermodules in each PbGl sector had their 4x4c trigger bits masked off throughout Run 8 pp, so the next lowest trigger, the 4x4a, is used for the pp results. Fortunately, because of the change in thresholds, the 4x4a trigger in the pp runs actually has the same threshold (approximately 2.1 GeV in 4x4 PbSc towers and 1.4 GeV in 4x4 PbGl towers) as the 4x4c trigger in the d+Au runs.

6.3.1 ERT efficiency

The ERT efficiency for jets is determined in minimum bias data.

Unlike the trigger efficiency for a single particle, which has an obvious definition, we must define what it means for a reconstructed jet to fire the ERT. The same definition must be used in a consistent manner to construct the efficiency (the probability that a reconstructed jet in a minimum bias event fires the trigger) and select jets in the triggered dataset (which are then corrected by the trigger efficiency).

To test whether a minimum bias reconstructed jet fires the trigger, we look for a nearby particle associated with the jet that is consistent with a live trigger bit. Since the trigger bits from all 4x4 towers in a Super Module are *OR*'d together online, the best we can ask is if a cluster/track l*ies in* a SuperModule which recorded a live trigger bit. (Reconstructed clusters have an obvious Super Module association. For charged tracks, we use the associated cluster.)

A reconstructed jet fires the trigger under the condition,

- a cluster within $\Delta R < 0.3$ of the jet axis lies in a Super Module with a live trigger bit, or
- a track within $\Delta R < 0.3$ of the jet axis is associated with a cluster which lies in a Super Module with a live trigger bit

Because the energy scale at the d+Au-reconstructed level has weak centrality-dependent differences, we must construct the efficiency as a function of $p_{\mathrm{T}}{}^{rec}$ for each centrality separately.

6.3.2 Efficiency fits

The trigger efficiency ϵ is constructed by considering the fraction of jets in minimum bias events which are consistent with firing the trigger in a nearby ERT SM as decribed above. The set of jets tested for firing the trigger are required to pass all analysis cuts.

The errors on the efficiency are the standard binomial errors, with the exception that bins with N counts and 100% efficiency have their error constructed as if they have $N-1$ efficient counts (the standard binomial error formula would assign these bins an efficiency of 1 ± 0 exactly). To summarize, if a given bin has A efficient counts over N total counts (an efficiency of $\epsilon = A/N$),

- if $A \neq N$, the error on ϵ is $\sqrt{\frac{\epsilon(1-\epsilon)}{N}}$

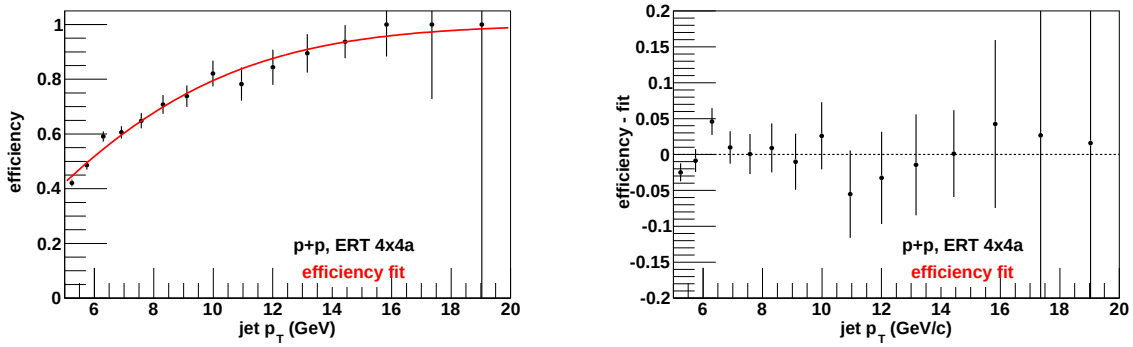


Figure 6.18: *pp* 4x4a trigger efficiency and fit (left) and residuals (right).

- if $A = N \neq 1$, the error on ϵ is $\sqrt{\frac{N-1}{N}\left(1-\frac{N-1}{N}\right)\frac{1}{N}} = \sqrt{(N-1)/N^3}$
- if $A = N = 1$, we set the error on ϵ equal to ± 1

To properly model the saturation at high-p_{T}, we use the error function $erf()$ which has an asymptote when the argument is large. To allow the function to describe the rise with p_{T}, the argument to erf is parameterized as a second-order polynomial (formally, for reasons of orthogonality, the fit is performed with the three lowest-order Chebyshev polynomials). The full function is

$$eff({p_{\mathrm{T}}}^{rec}) = c_0 \cdot \mathrm{erf}\left(c_1 + c_2 \cdot {p_{\mathrm{T}}}^{rec} + c_3 \cdot (2({p_{\mathrm{T}}}^{rec})^2 - 1)\right) \tag{6.21}$$

Although the high-p_{T} efficiency points have few counts (due to the much poorer rates for high-p_{T} jets in the minimum bias sample), they are important for determining the asymptotic behavior at high-p_{T}. Thus, we use logarithmic likelihood fits, which better deal with the small numbers of counts.

The trigger efficiency, fit result and fit residuals are shown for the 4x4a trigger in *pp* collisions in Figure 6.18. For the *d*+Au centralities, we use a slightly different procedure, as follows. Since the high-p_{T} asymptote of the efficiency should be driven by geometric factors alone (e.g. loss of trigger "acceptance" from masked off Super Modules, etc.), the fit functions in the different centrality

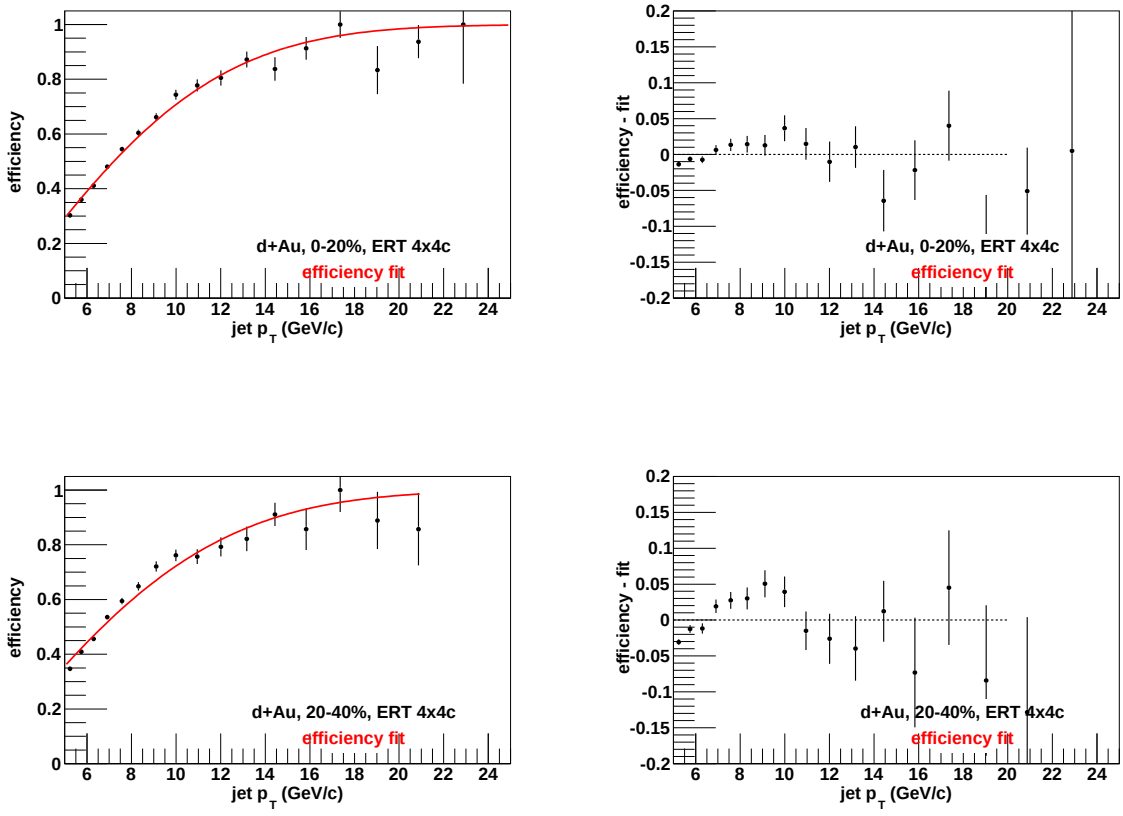


Figure 6.19: d+Au 4x4c trigger efficiency and fit (left) and residuals (right) for 0-20% (top) and 20-40% (bottom).

bins should asymptote at the same efficiency. (This is also important for limiting systematics in the R_{CP}, for which the trigger efficiency at high-p_{T} should cancel in the ratio.) Thus, we fit the minimum bias d+Au efficiency curve and constrain the individual centrality fits to use the same asymptote. Figures 6.19 and 6.20 show the efficiencies, fit results, and fit residuals for the 4x4c trigger in the d+Au centralities. In general, the fits were found to describe teh data well, and the magnitude of the fit residuals are used to assign a systematic uncertainty to the trigger efficiency.

For both the pp and d+Au cases, the efficiency asymptotes at 100% within uncertainty, indicating full saturation at high-p_{T}. This makes sense, since the total dead area from masked Super Modules corresponds to 3.5% of the towers in d+Au (and 2% in pp), and and a high-p_{T} jet may

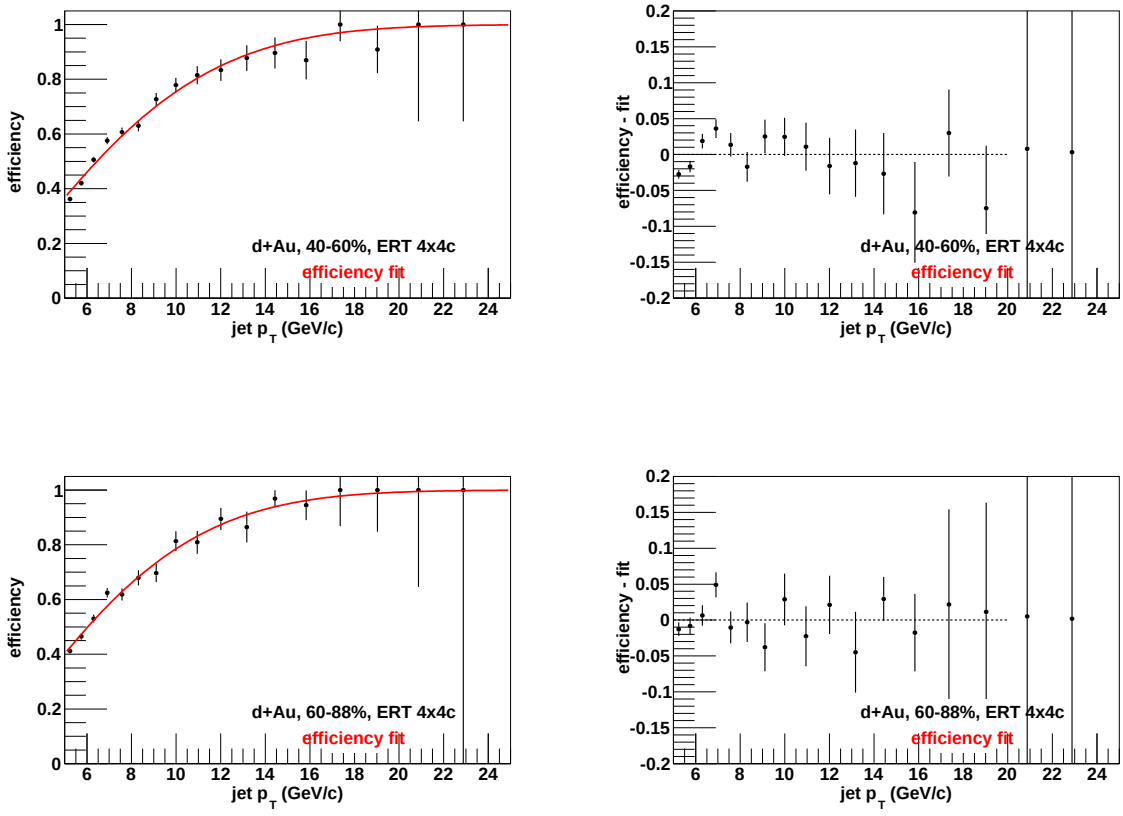


Figure 6.20: d+Au 4x4c trigger efficiency and fit (left) and residuals (right) for 40-60% (top) and 60-88% (bottom).

fragment in such a way that it fires the trigger bit in multiple SuperModules.

As a consistency check of the trigger efficiency determination and efficiency correction procedure, we examine the ratio of the trigger-corrected yields to the Minimum Bias yields, for the pp and each d+Au centrality. These are shown for the pp in Figure 6.21 and for the d+Au in Figure 6.22. The ratios are 1 within statistics, which are unfortunately large.

To determine systematic errors from the efficiency fit function modeling, we examine the fit residuals in Figures 6.18, 6.19 and 6.20 and the corrected/MinBias cross-checks in Figures 6.21 and 6.22, and quantify the maximum possible deviation suggested in data from the trigger efficiency fitting and correction procedure. The systematic uncertainties on the trigger efficiency correction

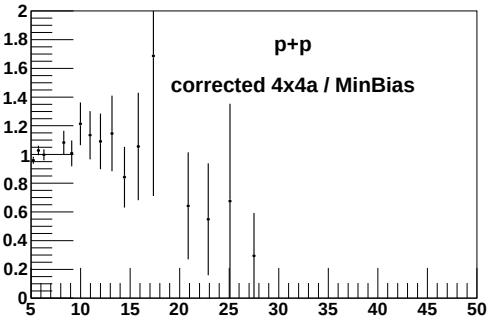


Figure 6.21: *pp* efficiency-corrected ERT 4x4a / MinBias yield ratio.

are:

- 5% for $p_{\mathrm{T}} < 15$ GeV/c, 5% for $p_{\mathrm{T}} > 15$ GeV/c in 0-20%
- 8% for $p_{\mathrm{T}} < 15$ GeV/c, 5% for $p_{\mathrm{T}} > 15$ GeV/c in 20-40%
- 5% for $p_{\mathrm{T}} < 15$ GeV/c, 5% for $p_{\mathrm{T}} > 15$ GeV/c in 40-60%
- 8% for $p_{\mathrm{T}} < 15$ GeV/c, 5% for $p_{\mathrm{T}} > 15$ GeV/c in 60-88%
- 8% for $p_{\mathrm{T}} < 15$ GeV/c, 5% for $p_{\mathrm{T}} > 15$ GeV/c in *pp*

6.3.3 Prescale-corrected yields

The total per-event jet yields in *d*+Au and *pp* are constructed after all jet-level selection cuts are applied, and it is these yields which are further corrected wtih trigger efficiency and unfolding factors.

For a given run, the per-event jet yield in a given p_{T} bin from the minimum bias stream is simply

$$\frac{1}{N^{(\mathrm{run})}_{\mathrm{MB,\ evt}}} \frac{dN^{\mathrm{MB,\ jet}}_{(\mathrm{run})}}{dp_{\mathrm{T}}} \tag{6.22}$$

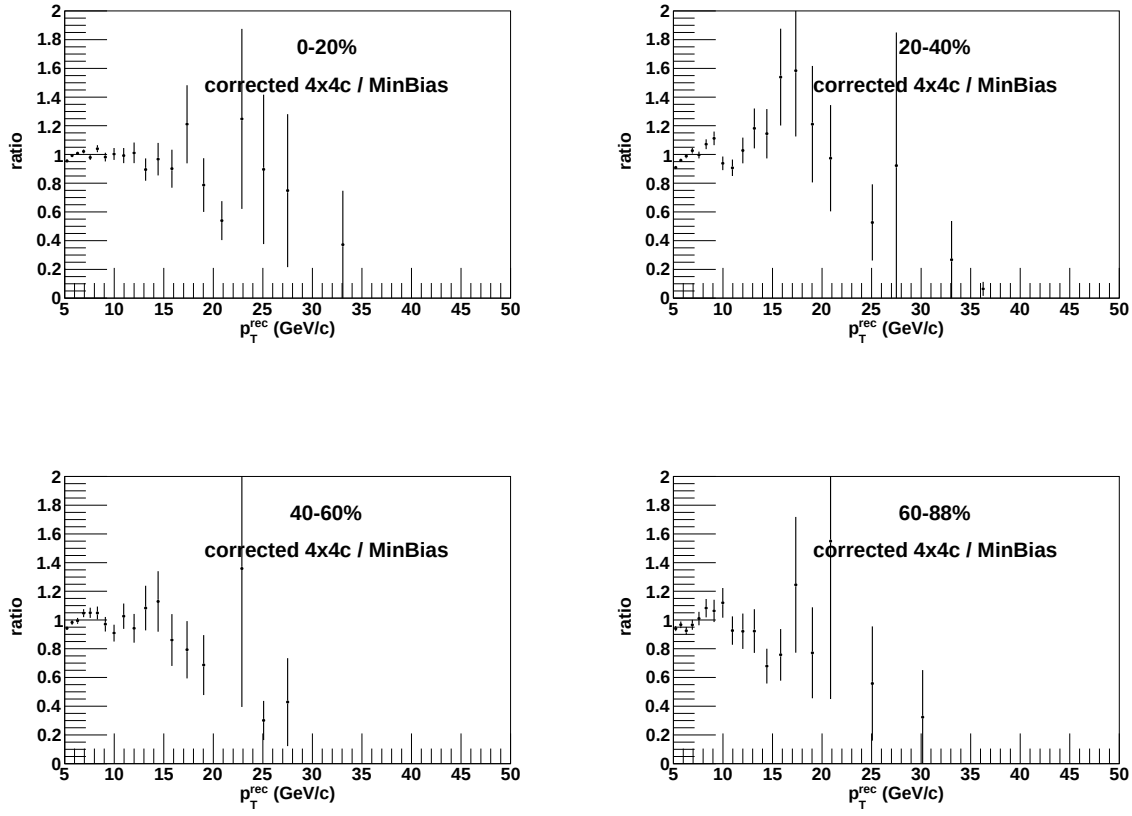


Figure 6.22: d+Au efficiency-corrected ERT 4x4c / MinBias yield ratio for each centrality.

where $N^{(\mathrm{run})}_{\mathrm{MB,\ evt}}$ is the number of Minimum Bias events analyzed in that run and $N^{\mathrm{MB,\ jet}}$ is the number of jets in those events (within the p_{T} bin in question).

These can combined from all runs in a straightforward manner to get the full Minimum Bias yield,

$$\frac{1}{N_{\mathrm{MB,\ evt}}}\frac{dN^{\mathrm{MB,\ jet}}}{dp_{\mathrm{T}}} = \frac{\Sigma_{\mathrm{run}} dN^{\mathrm{MB,\ jet}}_{(\mathrm{run})}/dp_{\mathrm{T}}}{\Sigma_{\mathrm{run}} N^{(\mathrm{run})}_{\mathrm{MB,\ evt}}} \tag{6.23}$$

In the triggered dataset, the situation is slightly different. Each run has a Minimum Bias and ERT scaledown s_{MB} and s_{ERT}. The total number of Minimum Bias events recorded by the DAQ (and thus, seen by the analysis train) is $N^{(\mathrm{run})}_{\mathrm{MB,\ evt}}$. However, the total number of events that passed

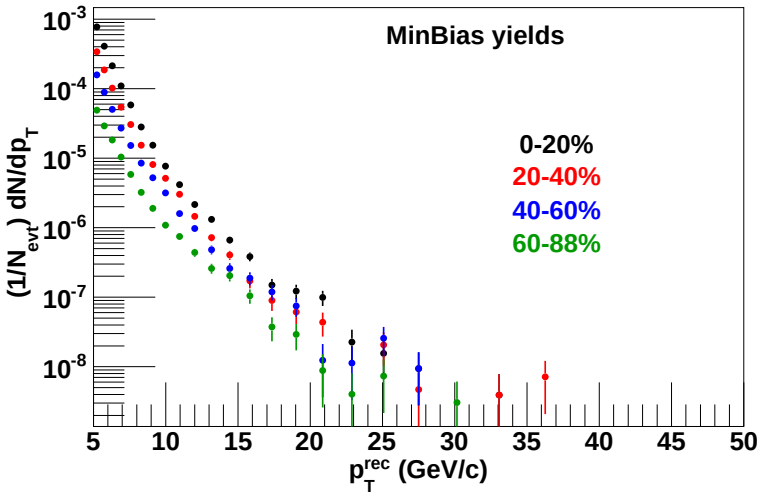


Figure 6.23: Per-event jet yields in d+Au collisions from the Minimum Bias datastream.

the MB criterion is $N^{(\mathrm{run})}_{\mathrm{MB,\ evt}} \times (1 + s_{\mathrm{MB}})$. And the total number of events seen by the ERT trigger must be corrected for the ERT scaledown:

$$N^{(\mathrm{run})}_{\mathrm{ERT,\ evt}} = N^{(\mathrm{run})}_{\mathrm{MB,\ evt}} \times (1 + s_{\mathrm{MB}})/(1 + s_{\mathrm{ERT}}) \tag{6.24}$$

Thus, the yield from the triggered datastream in any given run is,

$$\frac{1}{N^{(\mathrm{run})}_{\mathrm{ERT,\ evt}}} \frac{dN^{\mathrm{ERT,\ jet}}_{(\mathrm{run})}}{dp_{\mathrm{T}}} \tag{6.25}$$

And the combined yields from the triggered datastreams are,

$$\frac{1}{N_{\mathrm{ERT,\ evt}}} \frac{dN^{\mathrm{ERT,\ jet}}}{dp_{\mathrm{T}}} = \frac{\Sigma_{\mathrm{run}} dN^{\mathrm{ERT,\ jet}}_{(\mathrm{run})}/dp_{\mathrm{T}}}{\Sigma_{\mathrm{run}} N^{(\mathrm{run})}_{\mathrm{ERT,\ evt}}} = \frac{\Sigma_{\mathrm{run}} dN^{\mathrm{ERT,\ jet}}_{(\mathrm{run})}/dp_{\mathrm{T}}}{\Sigma_{\mathrm{run}} N^{(\mathrm{run})}_{\mathrm{MB,\ evt}} \times (1 + s_{\mathrm{MB}})/(1 + s_{\mathrm{ERT}})} \tag{6.26}$$

For the centrality-selected yields, $N^{\mathrm{ERT,\ jet}}$ are $N_{\mathrm{MB,\ evt}}$ are only measured in events with the relevant centrality selection.

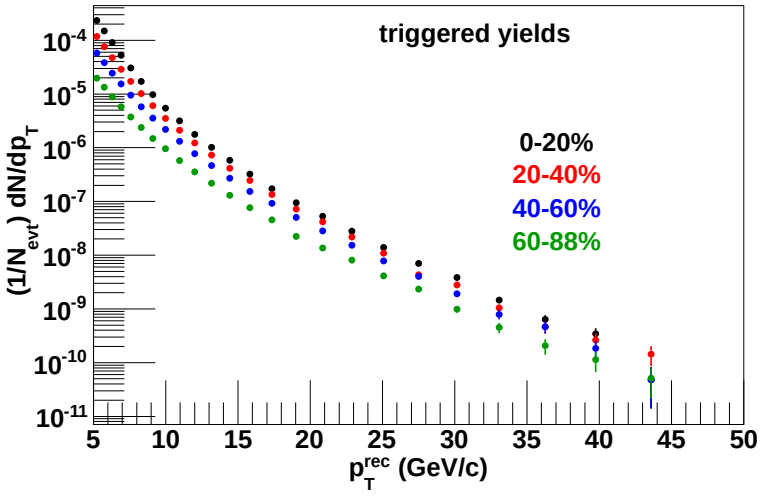


Figure 6.24: Per-event jet yields in d+Au collisions from the ERT datastream.

The minimum bias and triggered pet-event jet yields for the d+Au centralities are shown in Figures 6.23 and 6.24. The per-event jet yield in pp collisions is shown in Figure 6.25.

6.3.4 Stability of jet yields

To quantify an upper limit on any residual changes in the jet yields from changes in the detector acceptance, we plot the p_{T}-integrated jet yields in each run on the good run list. We use the minimum bias data streams for this cross-check, since the yields in the triggered datastream will have the effects of the changing trigger efficiency folded in (which is a separate issues from the acceptance).

The run dependence for the d+Au centralities and pp jets are shown in Figures 6.26 and 6.27. To quantify any possible change in the jet yields, a line was fit to the data and the change in the yield from the beginning to the end of the run was evaluated. The residual change in yields is:

- -0.54 ± 0.48 for 0-20% yields
- -1.1 ± 0.5 for 20-40% yields

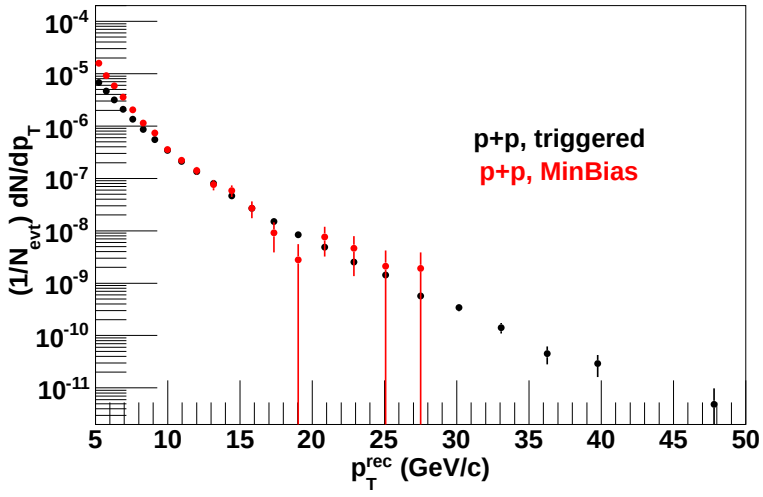


Figure 6.25: Per-event jet yields in *pp* collisions from Minimum Bias and ERT datastream.

- -1.2 ± 0.7 for 40-60% yields
- -1.7 ± 0.7 for 60-88% yields
- -0.37 ± 1.2 for *pp* jets

The worst of these are consistent with $< 2\%$. (Note that these are changes in the jet *yields* which translate directly into an uncertainty on the $R_{\rm CP}$ or $R_{\rm dAu}$, unlike possible changes in the energy scale examined in Section 6.2.4 which would enter as a fifth power due to the falling spectrum.) We quote a conservative systematic of 3% for possible residual run dependent effects, which include this observation here.

6.3.5 BBC cross-section & trigger bias in *pp*

The per-event yield in *pp* collisions must be converted into a cross-section. Since a luminosity calibration for Run 8 *pp* is unavailable to provide an integrated luminosity (in which case the cross-section could be constructed according to $d\sigma/dp_{\rm T} = (1/\mathcal{L})dN/dp_{\rm T}$), it is sufficient to know the

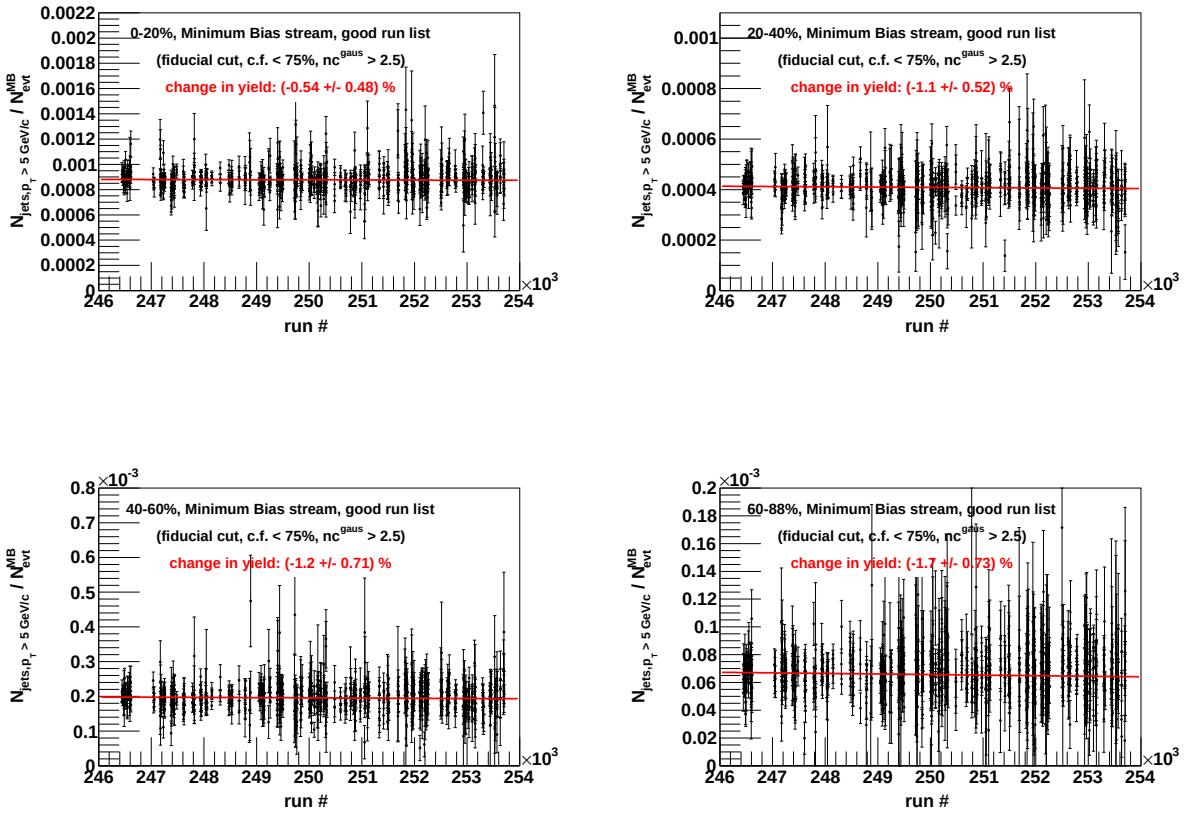


Figure 6.26: Run dependence of jet yields in minimum bias stream, for different centralities in d+Au runs.

BBC cross-section for jet events. At very high luminosities, when the mean number of interactions per crossing is large, "per-event" pp yields cannot be mapped to a total cross-section. However, for the peak luminosity in the 2008 pp running (see Section 3.2.2) it is safe to convert a per-event yield to a cross-section via $d\sigma/dp_{\mathrm{T}} = \left(\sigma_{MB}/N_{evt}^{MB}\right) dN/dp_{\mathrm{T}}$.

The BBC cross-section has not been measured directly in Run 8 pp, but the best estimate, which includes corrections derived in data from Run 5 and 6 luminosity scans[218], has a large systematic uncertainty and is given by

$$\sigma_{BBC} = 24.5 \text{ mb} \pm 10\% \tag{6.27}$$

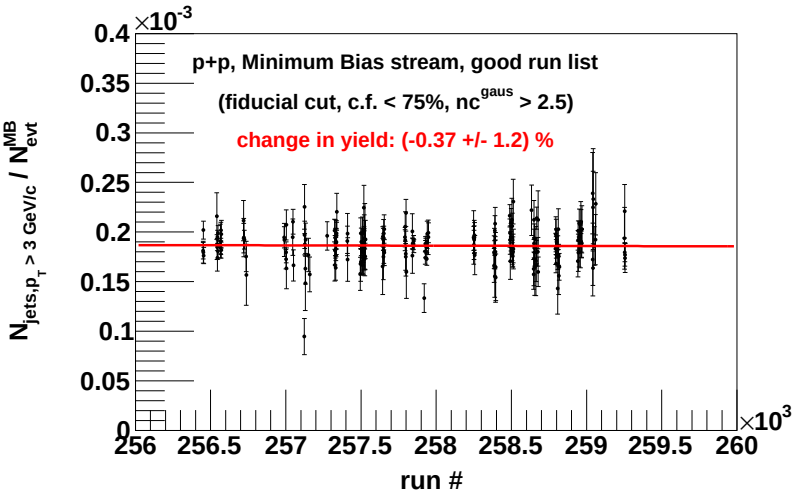


Figure 6.27: Run dependence of jet yields in minimum bias stream in *pp* runs.

Thus, the BBC LL1 trigger normally fires on $\sim 60\%$ of the *pp* inelastic cross-section. However, in the presence a jet in the Central arms, the BBC is susceptible to a *trigger bias* and fires with an efficiency ϵ_{BBC} which is higher than the nominal $\sim 60\%$ figure. The increased BBC multiplicity in the presence of a high-p_{T} jet is identically the same effect which gives rise to the bias factors in Section 6.2.3. To determine this efficiency, we select *pp* events in the ERT stream which fire the standalone 4x4b trigger and, if they contain a jet, ask with what efficiency they fire the BBC LL1 trigger as well.

Although some of the BBC LL1 inefficient *pp* events have a good reconstructed vertex (which is either outside the online ± 30 cm cut, or one that is reconstructed with the ZDC), most of the events have no reconstructed vertex at all. The latter class of no-vertex events complicates the determination of the BBC trigger efficiency and must be treated in a specialized manner, since the standard jet reconstruction procedure depends on the existence of a vertex:

- Since there is no z_{vtx} to construct the full cluster kinematics as given in Section 6.1.2, we assume $z_{vtx} = 0$. This distorts the jet energy reconstruction since the η of clusters is poten-

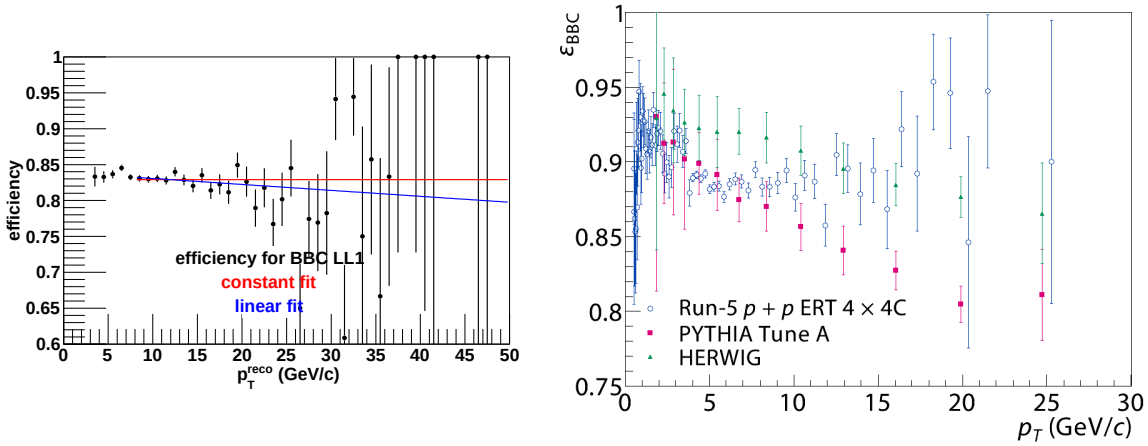


Figure 6.28: BBC LL1 trigger efficiency determination in this analysis (left) and in Run 5 pp (right).

tially misreconstructed, but since the PHENIX acceptance is relatively narrow, it is not a large effect in this study.

- Since there is no t_0 to provide time of flight information to the EMCal clusters, time of flight cuts are impossible and are disabled. Unfortunately, this invites out of time calorimetric background that is impossible to eliminate on a cluster by cluster basis.

To mitigate the latter problem, for the determination of the BBC LL1 efficiency only, we use two additional jet-level cuts:

- $c.f. > 5\%$, which eliminates purely-neutral jets that are dominated by EMCal noise
- $n.c. > 5.0$ (gaussian weighted), which eliminates low-constituent jets potentially influenced by EMCal noise

The resulting ϵ_{BBC} is shown in Figure 6.28 as a function of p_{T}. The high-p_{T} behavior is determined by fitting the efficiency above $p_{\mathrm{T}} > 8$ GeV/c (below this value, the sample of jets is potentially biased becase of the high-threshold requirement of the 4x4b trigger). The observed behavior is consistent with either a constant or linear behavior:

- constant fit: $\epsilon_{BBC} = 82.9 \pm 0.2$ %

- linear fit: $\epsilon_{BBC} = (83.9 \pm 0.6) - (0.084 \pm 0.054) \times p_{\mathrm{T}}$ [GeV] %

From a physics perspective, it is possible that the efficiency might drop at high-p_{T}, since requiring a high x in the hard-scattering can constrain the kinematics of the remaining partons in such a way as to make the BBC firing less likely. In fact, a moderate effect consistent with this possibility has been observed in PHENIX data.

Figure 6.28 also shows a comparison to the Run 5 pp efficiency, which is p_{T}-independent within statistics and has an asymptote of $\sim 88\%$. It also shows a comparison to the efficiency as determined by PYTHIA and HERWIG, both of which show a slow decrease with p_{T}. Because of the changes in the BBC calibration and lookup tables between Run 5 and Run 8, the difference in the asymptotes is not surprising.

For the results presented here, the p_{T}-independent efficiency is used. However, a systematic uncertainty is added to the data by considering the differences between the fits and the residuals in the constant fit. The assigned uncertainty on ϵ_{BBC} is

- 2% for the pp cross-section below 15 GeV
- 3% between $15 - 20$ GeV
- 5% for > 20 GeV results

6.4 Monte Carlo simulation

6.4.1 MC datasets

To fully understand the jet reconstruction performance across the kinematic range available to PHENIX, multiple Monte Carlo datasets with different minimum kinematic requirements on the hard scattering are used.

The underlying event generation is performed using PYTHIA Tune "A"[245] with all $2 \to 2$ QCD processes turned on. A total of nine different thresholds on $\sqrt{Q^2}$, the minimum momentum transfer in PYTHIA between the scattering partons, are used, from 4 to 40 GeV. The values of $\sqrt{Q^2}$, as well as the total number of generator events, are summarized in Table 6.7. Every stage of the MC simulation, from the event generation to the jet reconstruction of detector-level objects is handled within the PHPythia/Fun4All framework.

$\sqrt{Q^2}$ (GeV)	generator events	$\sigma(Q^2)$ (mb)	truth jet trigger	triggered events
4	4,050,000	4.81×10^{-1}	${p_{\mathrm{T}}}^{truth} > 3$ GeV/c	823k
6	2,560,000	6.88×10^{-2}	${p_{\mathrm{T}}}^{truth} > 3$ GeV/c	1,036k
8	400,000	1.55×10^{-2}	${p_{\mathrm{T}}}^{truth} > 5$ GeV/c	136k
12	300,000	1.58×10^{-3}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	73k
16	240,000	2.65×10^{-4}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	122k
20	200,000	5.79×10^{-5}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	135k
24	200,000	1.49×10^{-5}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	151k
28	160,000	4.18×10^{-6}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	130k
32	160,000	1.25×10^{-6}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	138k
40	50,000	1.28×10^{-7}	${p_{\mathrm{T}}}^{truth} > 10$ GeV/c	36k
total	2,210,000			0

Table 6.7: Monte Carlo dataset summary.

After the PYTHIA event is generated (with most short-lived unstable particles decayed by the generator), Gaussian filter jet reconstruction is run on all $pT > 400$ MeV final state, non-neutrino particles in the event record. The resulting set of jets are called *truth jets* with a transverse momentum ${p_{\mathrm{T}}}^{truth}$. These are the objects of interest.

Not every generator level event with a given $\sqrt{Q^2}$ will produce a fragmenting parton in a region of phase space accessible to PHENIX. To save CPU time processing these empty (to PHENIX) events, a trigger on the truth jet was implemented that requires at least one truth jet within $|\eta| < 0.5$ above a certain threshold ${p_{\mathrm{T}}}^{truth}$ which varies with $\sqrt{Q^2}$. Table 6.7) lists the p_{T} threshold and the number of events which fire the truth trigger. These events are then processed through the full detector simulation.

The detector response is simulated using a `GEANT3` code adapted to the PHENIX geometry called PISA. To properly model the full z range of the interaction point, the z-vertex is artificially smeared in a Gaussian way around 0 with a standard deviation of 20 cm, which is close to the width of the reconstructed z_{vtx}-distribution observed in d+Au data.

After this, the PISA hits are digitized and converted into a simulated datastream (simDST) for analysis, with run # 258795 (a late, stable *pp* run) serving as the reference for detector calibrations and dead maps. The run selection and broken wire masks described in Sections 6.1 and 6.2 have been devised so that the detector setup in data matches that in the Monte Carlo as closely and in as uniform a manner as possible. The full analysis module is then run on the simDST, implementing the full set of track quality, track-pair, cluster time of flight, kinematic, etc., cuts and performing jet reconstruction in the same way as it is done on data.

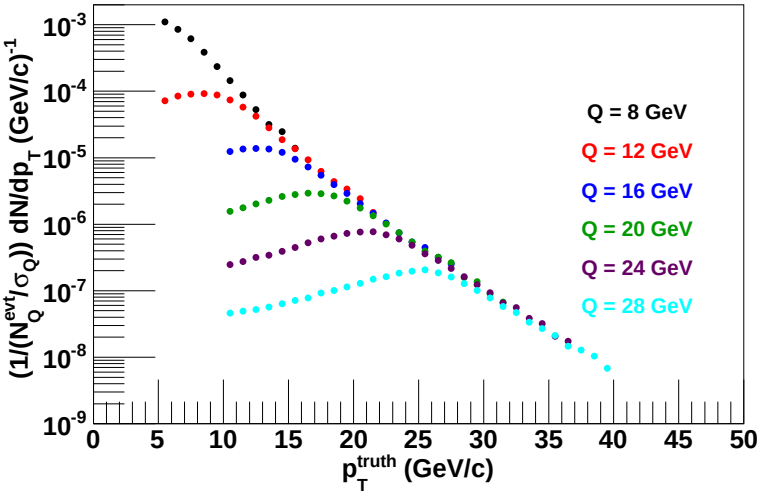


Figure 6.29: Per-event truth jet yield vs. p_{T} in select MC subsamples, from $\sqrt{Q^2} = 8$ to $\sqrt{Q^2} = 28$ GeV.

Because of the steep power law behavior of jet production from a hard scattering, no single MC subsample can describe the full kinematic range used in this analysis. Thus, the variables of interest (such as the jet spectrum or the reconstruction efficiency) must be combined fom multiple MC subsamples (denoted with the index Q).

A generic variable $\mathcal{O}$ is constructed from a weighted average of the $\mathcal{O}_Q$ from each MC sample, where the weighting is generally taken to be the PYTHIA generator cross-section of the sample scaled by the number of thrown events,

$$\mathcal{O} = \Sigma_Q \mathcal{O}_Q \cdot w_Q / \Sigma_Q w_Q \tag{6.28}$$

$$w_Q = \sigma_Q / N^Q_{evt} \tag{6.29}$$

To ward against fluctuations from low-statistics events at the high-p_{T} end of an MC subsample with low $\sqrt{Q}^2$ being given a high weighing and included in the determination of $\mathcal{O}$, the individual $\mathcal{O}_Q$ in Equation 6.28 are often pruned such that only bins with > 10 or > 20 in each $\sqrt{Q^2}$ contribute to the total weighted average.

The reconstruction efficiency must be constructed with a slightly different weight,

$$\epsilon(p_{\mathrm{T}}) = \Sigma_Q \epsilon_Q(p_{\mathrm{T}}) \cdot w_Q / \Sigma_Q w_Q \tag{6.30}$$

$$w_Q = \sigma_Q / N^Q_{evt} \times N^Q_{jet}(p_{\mathrm{T}}) \tag{6.31}$$

where $N^Q_{jet}(p_{\mathrm{T}})$ is the total number of truth jets in that p_{T} bin (e.g. the denominator in the ratio $\epsilon_Q(p_{\mathrm{T}})$).

The per-event jet yield also has a modified weight,

$$(1/N_{evt})(dN/dp_{\mathrm{T}}) = \Sigma_Q (1/N^Q_{evt})(dN/dp_{\mathrm{T}}{}^Q) \cdot w_Q / \Sigma_Q w_Q \tag{6.32}$$

$$w_Q = \sigma_Q \tag{6.33}$$

Figure 6.29 shows the per-event truth jet yield in the $\sqrt{Q^2} = 8$ to $\sqrt{Q^2} = 28$ GeV samples, after all pruning has been performed. At high-p_{T}, each $\sqrt{Q^2}$ subsample gives a yield consistent with that of the next-highest $\sqrt{Q^2}$ sample, as expected.

6.4.2 Single Particle Performance

Before the performance for full jets is examined, the full set of tracking and cluster cuts presented in Sections 6.1.2 and 6.1.3 were examined in MC for their efficiency with regard to charged hadrons and electrons and photons. Each charged hadron in the PYTHIA truth was associated with the nearest reconstructed track within $\Delta R < 0.02$ that passed tracking cuts. If there were no such

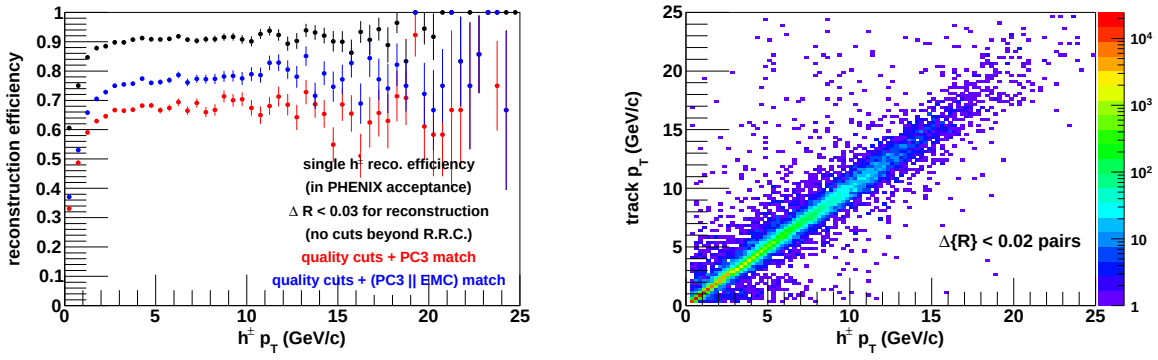


Figure 6.30: Efficiency for reconstructing charged hadrons (left) and transfer matrix between charged hadrons and the associated reconstructed track (right).

tracks within $\Delta R < 0.02$, that charged hadron was determined to have not been reconstructed. This was repeated for different sets of tracking cuts.

Figure 6.30 shows the reconstruction efficiency for charged hadrons as a function of p_{T} under a number of scenarios for the tracking cuts. Originally, reconstructed tracks were required to have a matching hit in the PC3. However, this was found to have poor efficiency in regions of weak PC3 acceptance. Thus, the cut was adjusted to alternately require a matching hit in the EMCal, both improving the overall efficiency to $\approx 80\%$ at high-p_{T} and making the efficiency much more uniform.

The same study was repeated for photons, electrons and π^0's (which quickly decay to γs) in the PYTHIA truth, but for matching to reconstructed electromagnetic clusters. Originally, the clusters were required to pass a set of signal shape cuts that preferentially selected clusters from photons and electrons over those from neutral hadrons. However, as is shown in Figure 6.31, it was found that merged clusters from a nearby pair of photons could not be distinguished from the clusters due to neutral hadrons, and rejecting them was causing a bias on the efficiency and energy scale at high-p_{T}. Thus, the signal shape cuts were dropped altogether. It can be seen from Figure 6.32 that at all cluster p_{T}, the contribution from single photons and π^0's dominated over the neutral hadrons anyway. With this set of cuts, it can be seen from Figure 6.33 that the reconstruction efficiency is $> 96\%$ (> 98 for $p_{\mathrm{T}} > 8$ GeV) for photons.

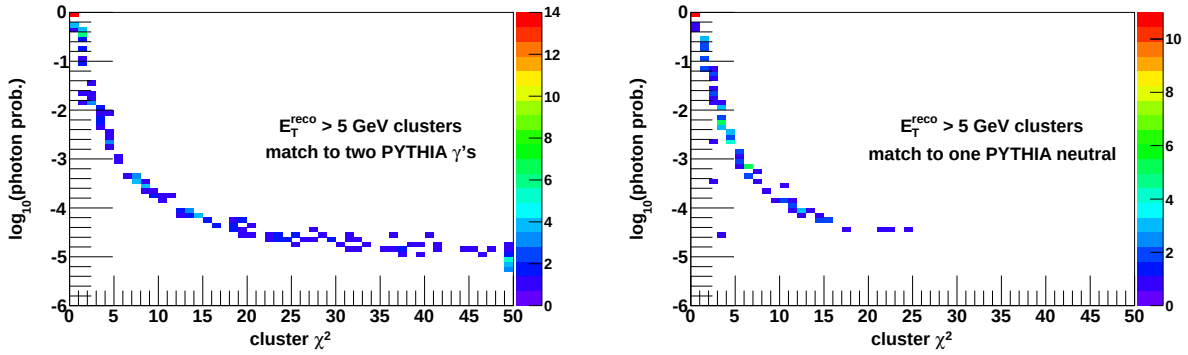


Figure 6.31: Scatterplot in χ^2 / "prob" space for $p_{\rm T} > 5$ clusters that originate from merged γ's (left) and neutral hadrons (right).

6.4.3 Jet Performance

In this section, we quantify the jet reconstruction efficiency and performance in terms of the distribution of ${p_{\rm T}}^{rec}$ values for a given ${p_{\rm T}}^{truth}$ range.

All truth jets within $|\eta| < 0.35$ (and any $\phi \in [0, 2\pi)$) are taken as the denominator in the efficiency. With this definition, the maximum theoretical reconstruction efficiency from the geometric acceptance of the PHENIX Central Arms is $\sim 50\%$. A truth jet is considered to be successfully reconstructed if the truth jet axis is within $\Delta R < 0.3$ of a reconstructed jet with ${p_{\rm T}}^{rec} > 3$ GeV/c jet. In the case of multiple matches, the closest reconstruction jet is considered to be the associated jet.

Figure 6.34 shows that the efficiency to see any jet at all (the black poins in the plot on the left) rises quickly with ${p_{\rm T}}^{truth}$ and reaches the "theoretical" $\sim 50\%$ by ${p_{\rm T}}^{truth} > 20$ GeV/c. The fiducial cut (red points) changes the value at which the efficiency saturates, since it effectively removes a part of the acceptance. The effect of a $g > 10$ GeV^2 fake jet rejection cut was also tested (see Section 6.1.6), and results in a substantial inefficiency below ${p_{\rm T}}^{truth} < 15$ GeV/c (blue points). The effect of the jet-level cut on the charged fraction and max z (green points) is a constant drop in efficiency, since the jet fragmentation space does not change significantly as a function of truth jet ${p_{\rm T}}^{truth}$.

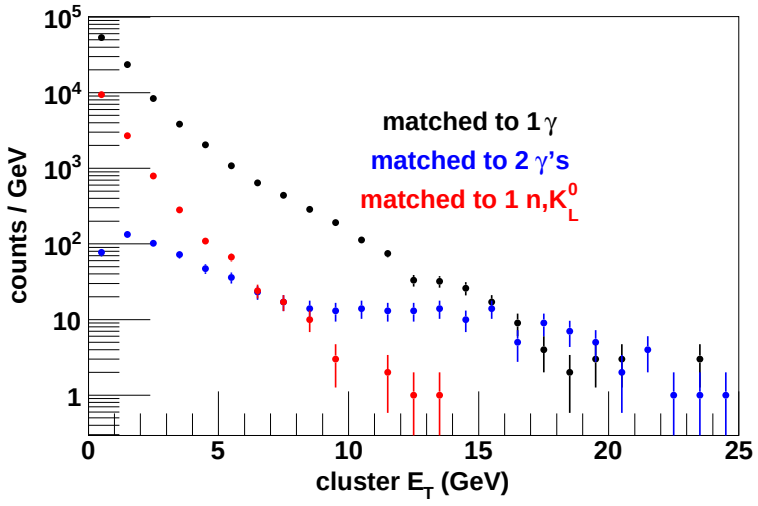


Figure 6.32: Relative contribution to the cluster spectrum in MC from single photons, decayed π^0's and neutral hadrons.

The right plot in Figure 6.34 compares the "raw" and final reconstruction efficiencies. The best comparison within PHENIX is the Run 5 Gaussian filter pp analysis[47], which parameterizes the efficiency with a Chebyshev polynomial expansion of an error function. Good agreement between the Run 5 and Run 8 efficiency is observed.

For truth jets which were successfully reconstructed, the transfer matrix $d^2N/dp_{\mathrm{T}}{}^{rec}dp_{\mathrm{T}}{}^{truth}$ is formed from matched pairs of $(p_{\mathrm{T}}{}^{truth}, p_{\mathrm{T}}{}^{rec})$ jets. The transfer matrix describes the range of reconstructed jet values for a given truth jet. In this way, the transfer matrix connects the true energy scale to the detector energy scale. Figure 6.35 shows the transfer matrix in East arm with and without jet-level cuts. The addition of the fiducial and jet-level cuts substantially improve the transfer matrix, making it more diagonal with and giving it a higher $\langle p_{\mathrm{T}}{}^{rec}\rangle$ for each $p_{\mathrm{T}}{}^{truth}$ slice.

Figure 6.36 shows slices of the transfer matrix for a wide range of $p_{\mathrm{T}}{}^{truth}$ bins. In each case, we see that the distribution of reconstructed $p_{\mathrm{T}}{}^{rec}$ has a long tail to low-p_{T}, reflecting that the largest contributor to the jet energy resolution in PHENIX is the fragmentation-dependent energy "loss" in the PHENIX detector from inefficiencies in the tracking and poor reconstruction of neutral energy,

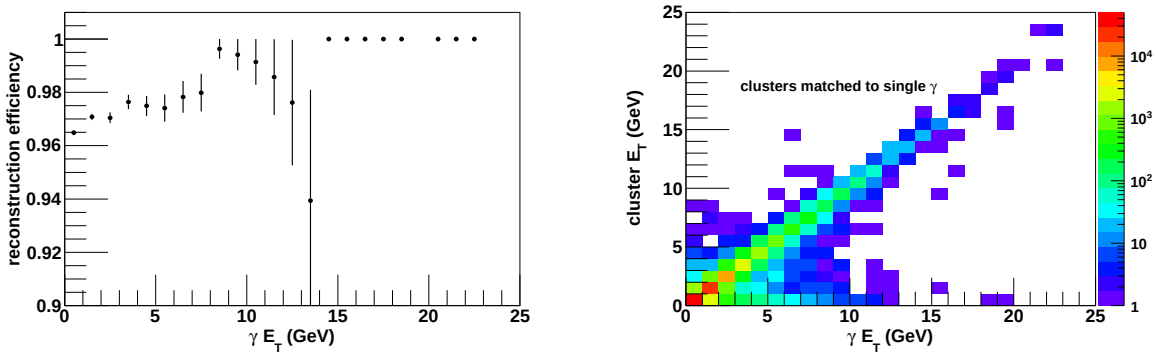


Figure 6.33: Reconstruction efficiency (left) and transfer matrix (right) for single photons matched to reconstructed clusters.

rather than the resolution on the individual tracks and clusters. The ${p_{\rm T}}^{rec}$ distributions are peaked close to the value ${p_{\rm T}}^{rec} \sim 0.75{p_{\rm T}}^{truth}$, which is the "back of the envelope" theoretical maximum of 75% for jets in the PHENIX detector, as derived in Section 6.1.1. (Perversely, an unfolding of the data from the *pp*-equivalent energy scale to the truth scale is made somewhat easier by the fact that tail is to *lower* $p_{\rm T}$ values. A tail to *higher* $p_{\rm T}$ values on a steeply falling spectrum makes the result of the unfolding more sensitive to small changes in the resolution, requiring additional care.)

The $p_{\rm T}^{\rm rec}$ distributions for this analysis are also shown compared to those in the Run 5 *pp* analysis. The Run 8 *pp* analysis appears to better reconstruct the jet energy. This is presumably from two major differences in particle-level cuts (both are re-posted here from Sections 6.1.2 and 6.1.3):

- In the present analysis we have taken care to not bias the cluster cuts against high-$p_{\rm T}$ merged clusters from π^0's, which often have `prob` and χ^2 characteristics that make them appear as if they are neutral clusters.

- Furthermore, reconstructed tracks appear to become less efficient for PC3 matching at high-$p_{\rm T}$, and thus in this analysis tracks may match to an EMCal hit in lieu of a PC3 hit.

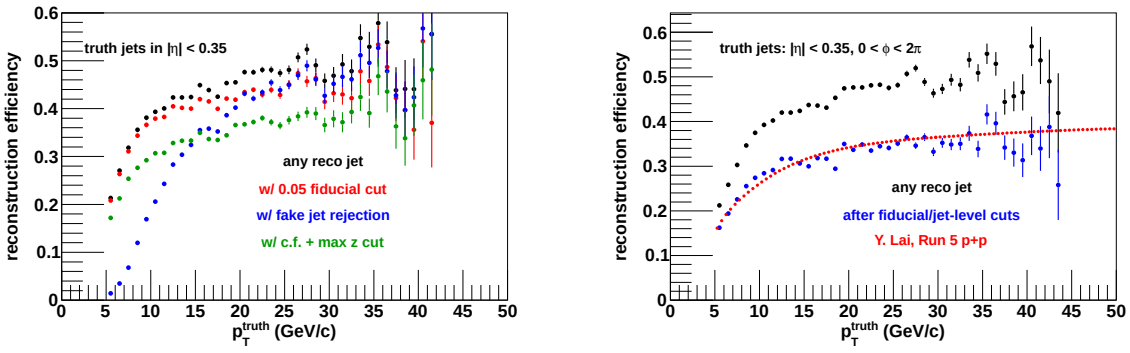


Figure 6.34: Left: Reconstruction efficiency as a function of $p_{\mathrm{T}}{}^{truth}$ with different jet-level reconstruction cuts. Right: Comparison of the reconstruction efficiency for seeing any jet at all (black), after all cuts (blue) and comparison to the Run 5 *pp* jet analysis (red line).

6.4.4 *d*+Au Embedding

To understand the effects of the soft *d*+Au underlying event on the jet reconstruction efficiency and energy scale, Monte Carlo *pp* jet events are embedded into real *d*+Au data. This is performed by reconstructing the MC *pp* event, merging its tracks and clusters with that of a real *d*+Au event, and running jet reconstruction on the result.

The *d*+Au events used for embedding are drawn from disk-resident Minimum Bias data. Each MC *pp* event is embedded four times total, into a separate background event from each of the four centrality categories. To better represent the real data (and model any residual run dependence), *d*+Au events from any given run om data are only used for up to 5% of a $\sqrt{Q^2}$-triggered MC sample. In total, *d*+Au events are currently drawn from more than 180 runs which pass run selection. The transfer matrices between the truth jet and the reconstructed jet in the *d*+Au environment are shown for the four centralities in Figure 6.37.

MC events are not required to already have a reconstructed jet in the *pp* event before they are embedded into the *d*+Au background. In this way, centrality-dependent differences in the reconstruction efficiency can be observed. For example, it is possible that a reconstructed jet will fail a jet-level cut (most likely the constituent requirement) in the *pp* system, but because of

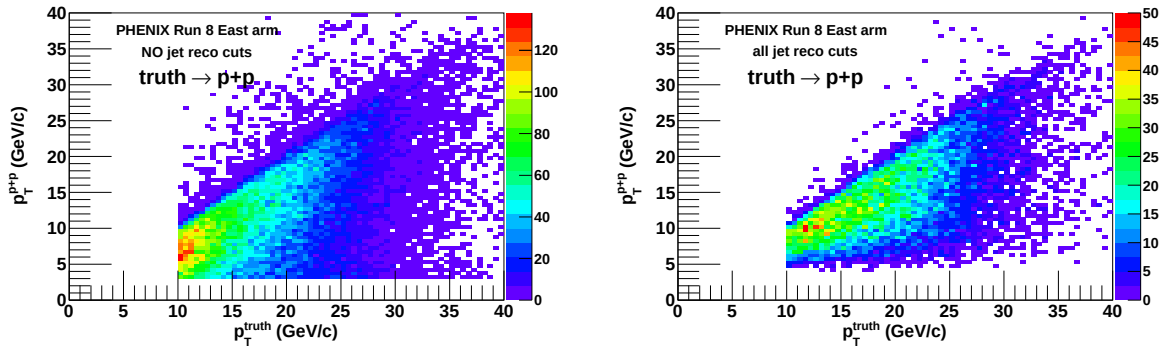


Figure 6.35: Transfer matrix for jets in PYTHIA *pp* events, before and after the full set of jet reconstruction cuts.

the effect of the *d*+Au underlying event, the reconstructed jet in the embedded event will pass. However, an examination of the centrality dependence of the reconstruction efficiency shows a $< 1\%$ effect in the reconstruction efficiency and in fact, the correction factors in the next section handle any changes in the efficiency in a natural way.

In actuality, the more important effect of the *d*+Au underlying event is the shift of the energy scale to higher ${p_{\rm T}}^{rec}$ from the addition of extra soft energy under the jet. This is called "$p_{\rm T}$-feeding" and is the dominant centrality-dependent effect in the *d*+Au jet yields.

6.4.5 Correction factors

To correct the centrality-dependent $p_{\rm T}$-feeding effect on the *d*+Au jet yields, a set of first order (also called "bin-by-bin") correction factors are constructed as follows. The end result is a set of *d*+Au jet yields which are all at the *pp*-equivalent reconstructed jet energy scale, and thus appropriate for comparisons between the different centralities and to the *pp* cross-section.

The $\sqrt{Q^2}$-combined total yield of truth-matched jets in different *d*+Au background events is compared to the yield of jets in *pp* collisions (e.g. just the PYTHIA simulation). As described in the embedding procedure, the set of truth events is the same for all yields. Then the correction factors are simply

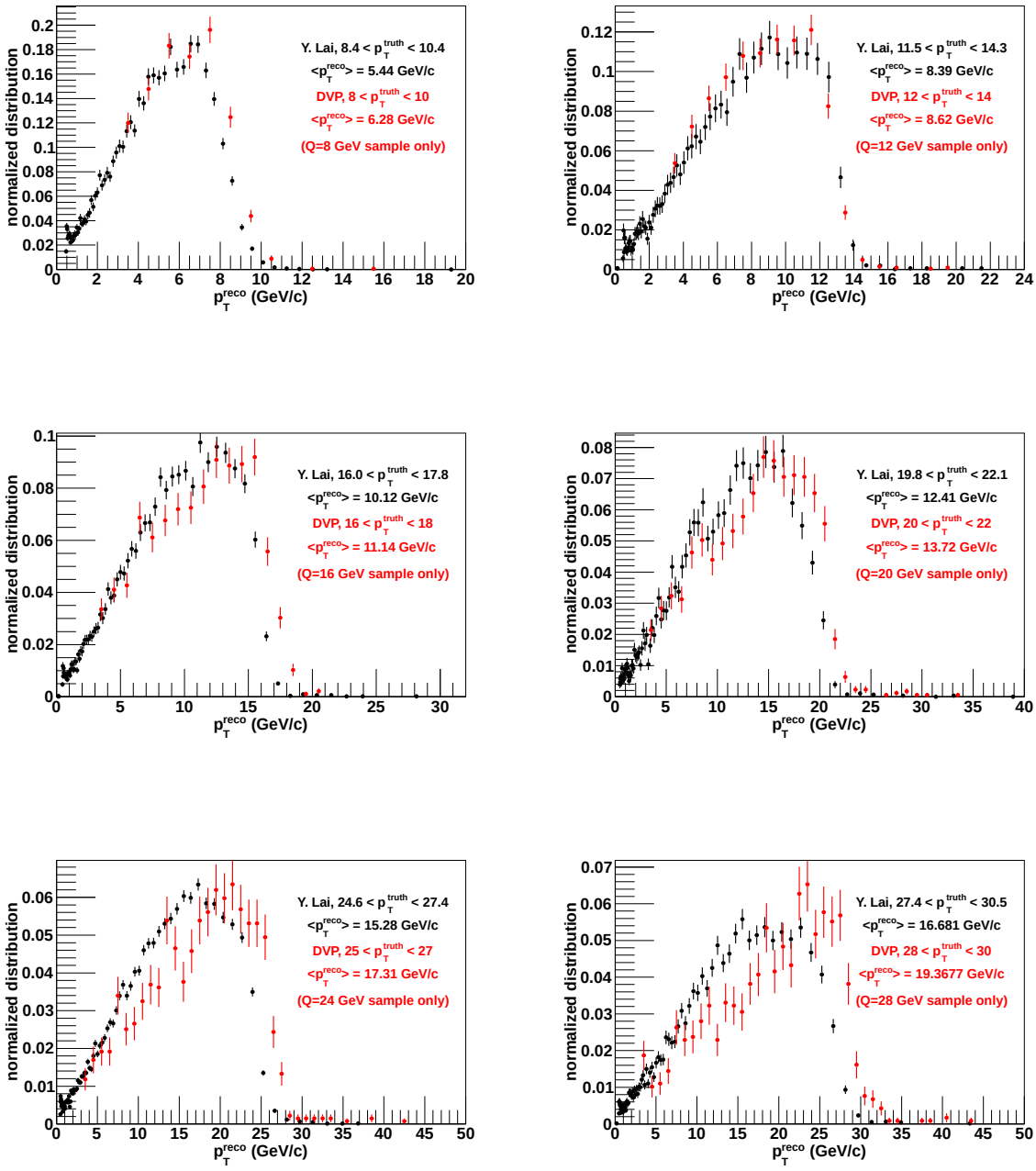


Figure 6.36: Distributions of ${p_{\mathrm{T}}}^{reco}$ given a narrow selection of ${p_{\mathrm{T}}}^{truth}$ in pp events. Each plot is drawn from one MC subsample. The mean of each distribution is shown and compared with a similar slice of the transfer matrix from [47].

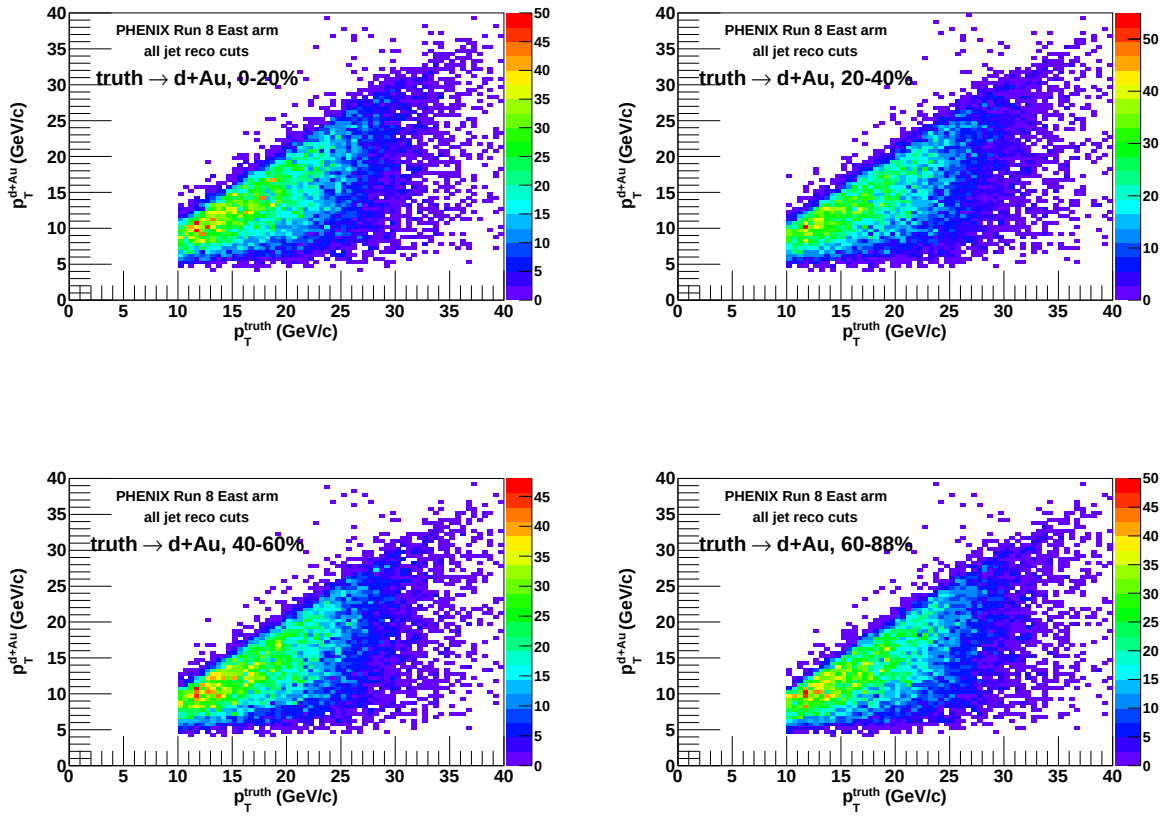


Figure 6.37: Transfer matrices for jets between the PYTHIA truth jet and the reconstructed jet in embedded d+Au background events, with each centrality shown separately.

$$C^{cent}(p_{\mathrm{T}}) = \frac{dN^{reco,cent}}{dp_{\mathrm{T}}} \Big/ \frac{dN^{reco,pp}}{dp_{\mathrm{T}}} \tag{6.34}$$

where the yields are evaluated at the same p_{T}. These could also be rewritten as the double ratio of the correction factors between the truth and reconstructed energy spectra, e.g.

$$C^{cent}(p_{\mathrm{T}}) = \left(\frac{dN^{reco,cent}}{dp_{\mathrm{T}}} / \frac{dN^{truth}}{dp_{\mathrm{T}}} \right) \Big/ \left(\frac{dN^{reco,pp}}{dp_{\mathrm{T}}} / \frac{dN^{truth}}{dp_{\mathrm{T}}} \right) \tag{6.35}$$

which reduces to Equation 6.34 since the underlying truth distribution is the same in both cases. Then, the d+Au yields measured in data are corrected via

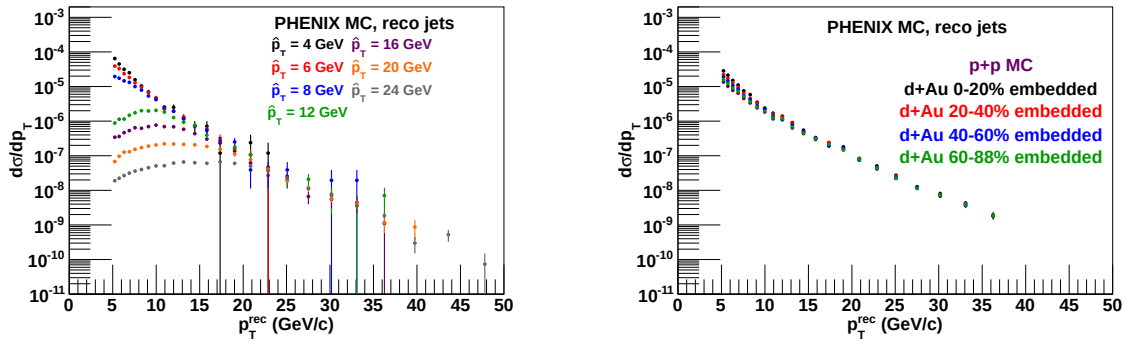


Figure 6.38: Left: truth-matched reconstructed jet spectra from each MC subset, *pp* Monte Carlo setup. Right: Truth-matched reconstructed jet spectra from *pp* MC events and from MC events embedded into different-centrality data heavy ion background.

$$\frac{1}{N_{evt}^{cent}}\frac{dN^{cent}}{dp_{\mathrm{T}}}\bigg|_{corrected} = \frac{1}{C^{cent}(p_{\mathrm{T}})}\,\frac{1}{N_{evt}^{cent}}\frac{dN^{cent}}{dp_{\mathrm{T}}}\bigg|_{raw} \tag{6.36}$$

Note that by construction the correction is $C = 1$ for *pp* events, since the *pp* energy scale is reference for the correction.

The truth-matched spectra are shown in Figure 6.38, although the centrality-dependent differences are difficult to see on a logarithmic scale. The correction factors $C^{cent}(p_{\mathrm{T}})$ are shown in Figure 6.39. Because of statistical fluctuations in the numerator and denominator, the correction factors were smoothed with a fit function.

Phenomenologically, the p_{T}-feeding can be modeled as the addition of small (centrality-dependent, but independent of the jet p_{T}) constant Δ to a jet's energy. On a falling $dN/dp_{\mathrm{T}} \sim {p_{\mathrm{T}}}^{-5}$ spectrum, this corresponds to a first-order multiplicative correction in the measured yield that has the functional form

$$(p_{\mathrm{T}}+\Delta)^5/(p_{\mathrm{T}})^5 = \left(1+\frac{\Delta}{p_{\mathrm{T}}}\right)^5 = 1+5\frac{\Delta}{p_{\mathrm{T}}}+O\left(\frac{\Delta}{p_{\mathrm{T}}}\right)^2 \tag{6.37}$$

Thus, we model the first-order correction to the d+Au yields with the form $c_0+c_1/p_{\mathrm{T}}+c_2/{p_{\mathrm{T}}}^2+c_3/{p_{\mathrm{T}}}^3$, which has the advantage of a natural high-p_{T} asymptote. The result of the fits to the C are

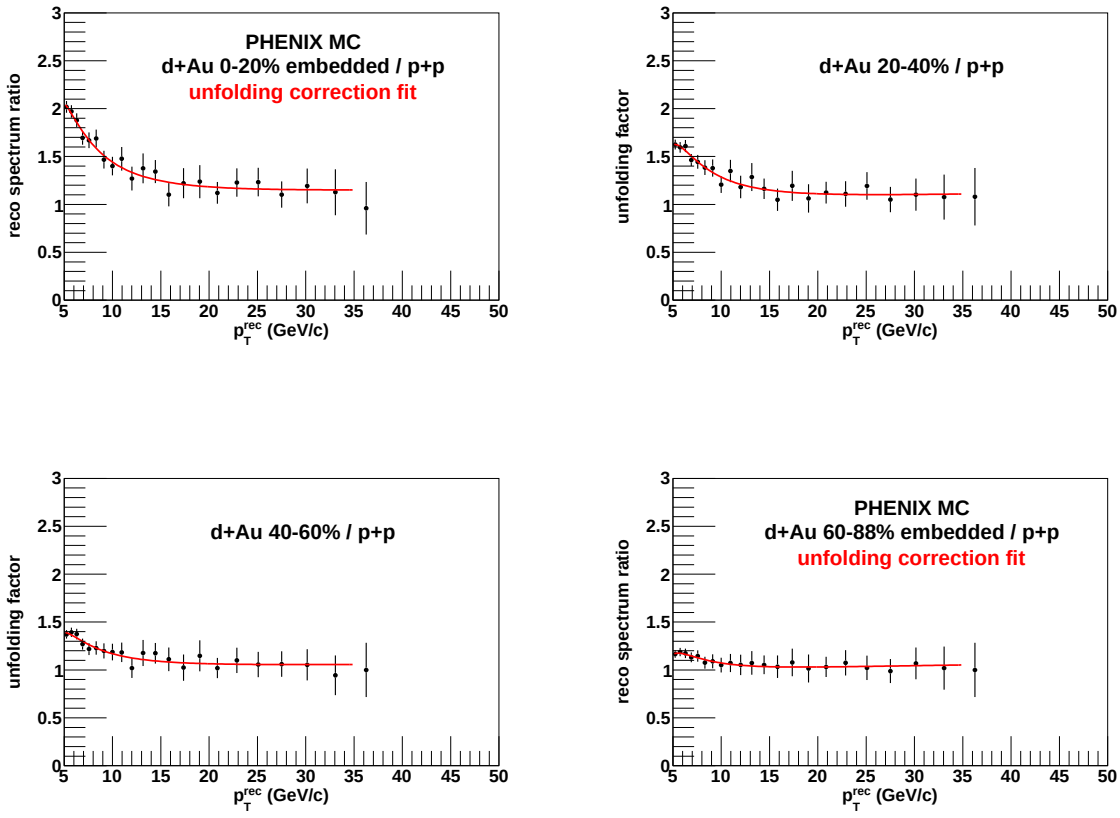


Figure 6.39: Ratio of truth-matched reconstructed jet spectra in embedded events to that in pp MC events, with first-order unfolding fits.

shown in Figure 6.39, and summarized together in Figure 6.40. The unfolding corrections follow the expected behavior in centrality (the jet yield in more central events have larger unfolding corrections because the increasing effect of the p_{T}-feeding of the d+Au underlying event), and they also follow the expected behavior in p_{T} (larger correction at lower p_{T}).

To determine the uncertainty on this correction, we examine the residuals in the fits to gauge the largest possible deviation from the fit model. The systematic uncertainties are

- 10% for $p_{\mathrm{T}} < 15$ GeV/c, 6% for $p_{\mathrm{T}} > 15$ GeV/c for 0-20%
- 8% for $p_{\mathrm{T}} < 15$ GeV/c, 6% for $p_{\mathrm{T}} > 15$ GeV/c for 20-40%

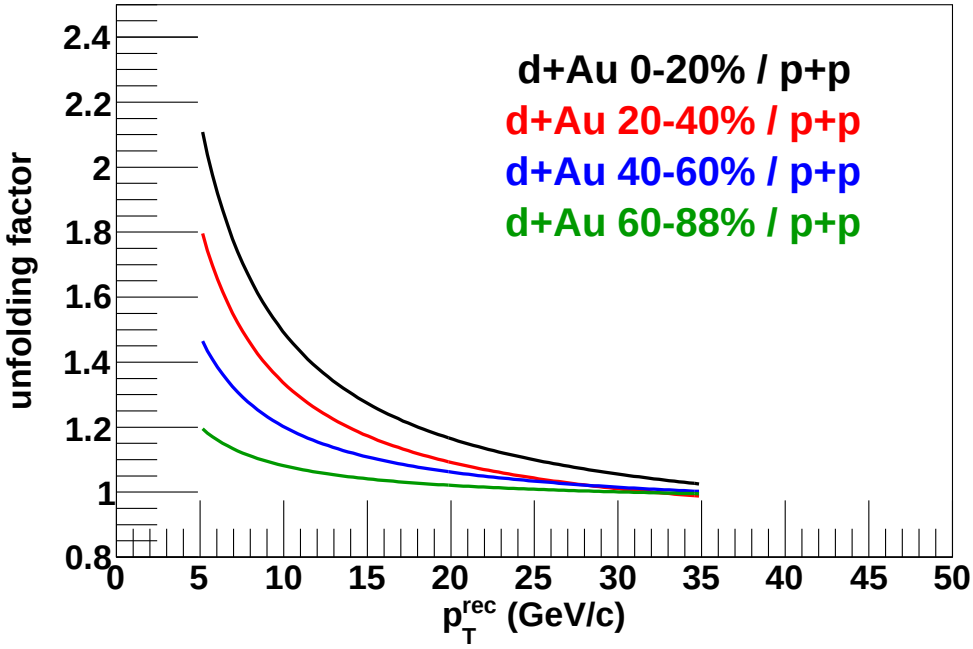


Figure 6.40: First-order unfolding fits for each centrality.

- 8% for $p_{\mathrm{T}} < 15$ GeV/c, 6% for $p_{\mathrm{T}} > 15$ GeV/c for 40-60%
- 4% for $p_{\mathrm{T}} < 15$ GeV/c, 4% for $p_{\mathrm{T}} > 15$ GeV/c for 60-88%

6.5 Results

6.5.1 Per-event jet yields in d+Au

The final d+Au yields are constructed from the ERT data and corrected for the p_{T}-dependent trigger efficiency in Section 6.3.2 and converted to the pp-equivalent energy scale using the correction factors in Section 6.4.5. These are normalized by $d^2/dp_{\mathrm{T}}d\eta$ and plotted in 6.41.

The $\langle T_{\mathrm{AB}}\rangle$-scaled yields are constructed via

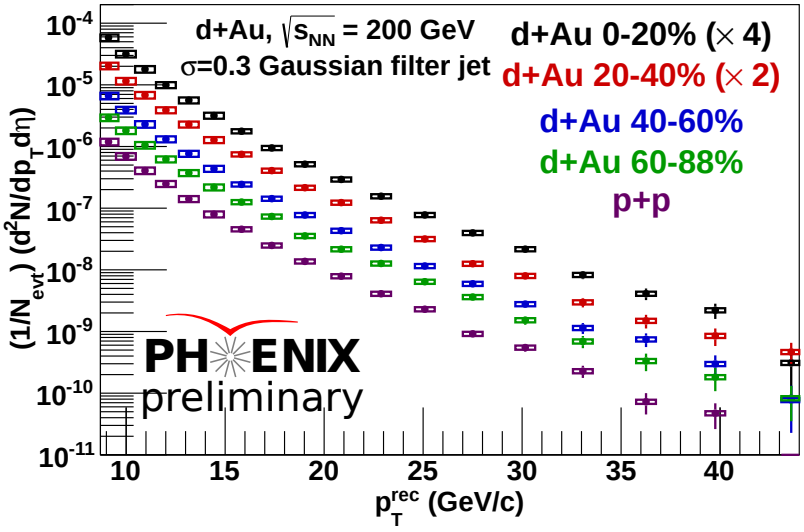


Figure 6.41: Reconstructed jet yields in d+Au at the $p_{\mathrm{T}}{}^{\mathrm{pp-rec}}$ energy scale.

$$\frac{1}{T_{\mathrm{AB}}^{\mathrm{cent}}}\frac{1}{N_{\mathrm{evt}}^{\mathrm{cent}}}\frac{dN^{\mathrm{cent}}}{dp_{\mathrm{T}}} \tag{6.38}$$

where the nuclear overlap function is $T_{\mathrm{AB}}^{\mathrm{cent}} = (N_{\mathrm{coll}}^{cent}/BF)/\sigma_{\mathrm{pp}}^{\mathrm{inelastic}}$, with the mean number of binary collisions N_{coll}, inelastic nucleon-nucleon cross-section σ_{NN} and bias factor BF determined according to the Glauber centrality analysis in Section 6.2.3.

6.5.2 Jet cross-section in *pp*

The efficiency-corrected p+p yields (which are at the pp-equivalent energy scale by definition) are converted into a cross-section according to

$$\frac{d\sigma}{dp_{\mathrm{T}}} = \frac{1}{\epsilon_{\mathrm{BBC}}}\sigma_{\mathrm{BBC}}\frac{1}{N_{\mathrm{evt}}}\frac{dN}{dp_{\mathrm{T}}} \tag{6.39}$$

Where σ_{BBC} is the BBC-firing part of the pp inelastic cross-section and ϵ_{BBC} is the BBC LL1 trigger efficiency for jet events, both of which are described in Section 6.3.5. The result is shown in Figure 6.42

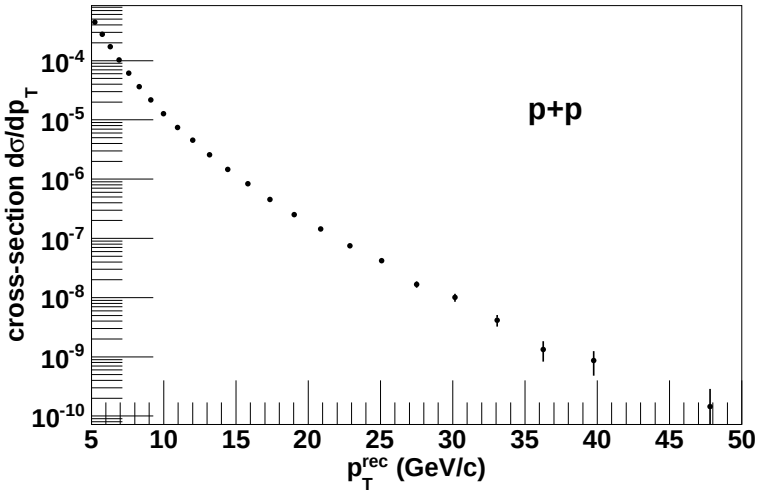


Figure 6.42: *pp* jet production cross-section at the $p_{\mathrm{T}}^{\mathrm{rec}}$-scale.

6.5.3 Jet R_{dAu}

The R_{dA} for a given centrality bin is formed as follows:

$$R_{\mathrm{dA}}^{\mathrm{cent}} = \frac{\left(1/N_{\mathrm{evt}}^{\mathrm{cent}}\right)\left(dN^{\mathrm{cent}}/dp_{\mathrm{T}}\right)}{T_{\mathrm{AB}}^{\mathrm{cent}} \cdot d\sigma/dp_{\mathrm{T}}} \tag{6.40}$$

As a reminder, the R_{dAu} is made from the ratio of a yield and a jet spectrum at the same *pp*-equivalent energy scale. The R_{dAu} in all four centralities is shown in Figures 6.43 and 6.45.

The R_{dAu} for inclusive (0-88%, or minimum bias) d+Au events can be constructed from the centrality-dependent R_{dAu}, since they are now all at the pp-equivalent detector scale. The inclusive R_{dAu} is given by the $N_{\mathrm{coll}} \times N_{\mathrm{evt}}$-weighted yields of the component R_{dAu}'s via

$$R_{\mathrm{dAu}}{}^{0-88\%} = \sum_{\mathrm{cent}} N_{\mathrm{evt}}^{\mathrm{cent}} N_{\mathrm{coll}}^{\mathrm{cent}} R_{\mathrm{dAu}}^{\mathrm{cent}} / \sum_{\mathrm{cent}} N_{\mathrm{evt}}^{\mathrm{cent}} N_{\mathrm{coll}}^{\mathrm{cent}} \tag{6.41}$$

where $N_{\mathrm{evt}}^{\mathrm{cent}}$ is the number of events ($N_{\mathrm{evt}}^{0-20\%} = N_{\mathrm{evt}}^{20-40\%} = N_{\mathrm{evt}}^{40-60\%} = \frac{28}{20} N_{\mathrm{evt}}^{60-88\%}$) and $N_{\mathrm{coll}}^{\mathrm{cent}}$ the number of binary collisions. Note that by construction the quantity in the denominator is just

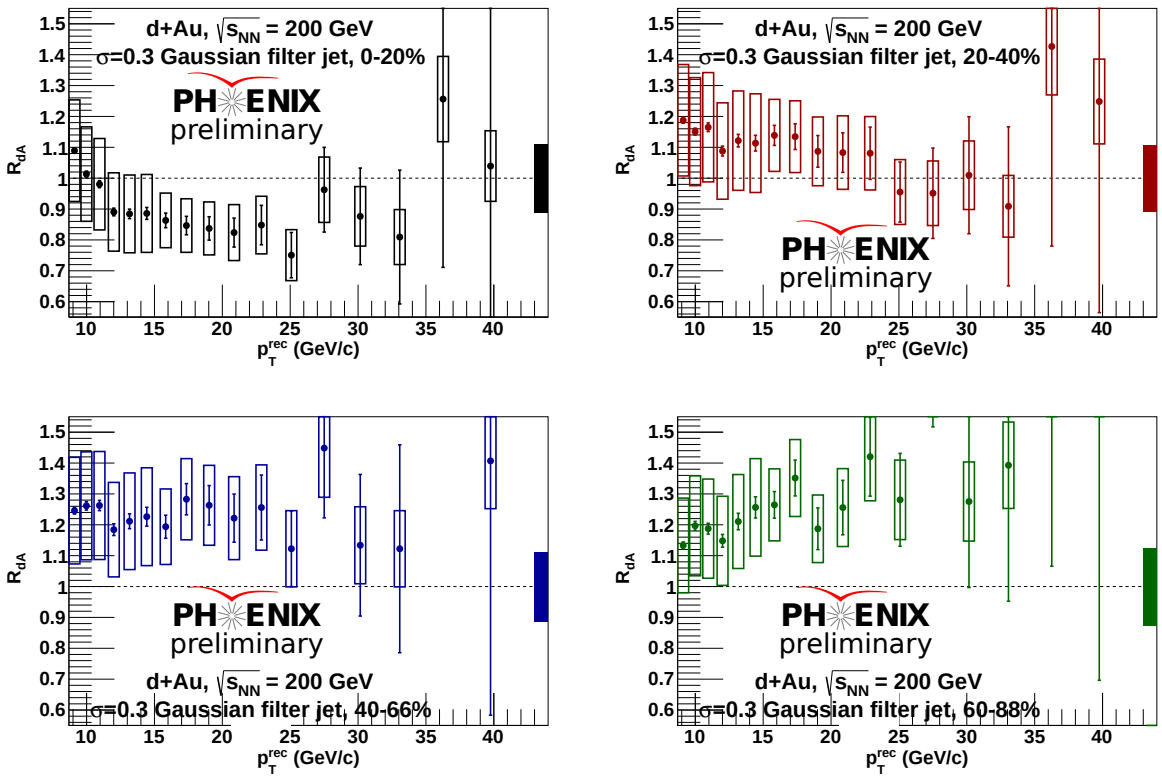


Figure 6.43: R_{dAu} for reconstructed jets at the pp energy scale, different centrality selections.

$$N_{\mathrm{evt}}^{\mathrm{cent}} N_{\mathrm{coll}}^{\mathrm{cent}} = N_{\mathrm{evt}}^{0-88\%} N_{\mathrm{coll}}^{0-88\%} \tag{6.42}$$

The statistical uncertainty is propagated in standard way through Equation 6.41. Since the centrality-selected R_{dAu} results all have a slightly different relative systematic uncertainty from differences in how well the efficiency fit functions and unfolding corrections describe the data, the inclusive R_{dAu} is chosen to conservatively inherit the worst systematic uncertainty from each of the four centrality-dependent R_{dAu}'s at a given p_{T} bin. Figure 6.44 shows the 0-88% R_{dAu}.

6.5.4 Jet R_{CP}

The R_{CP} relative to the most peripheral (60-88%) bin is:

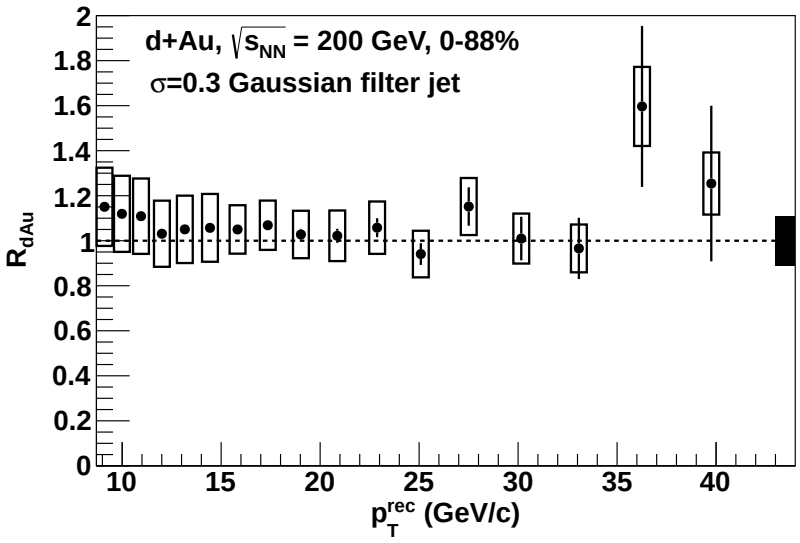


Figure 6.44: R_{dAu} for reconstructed jets at the pp energy scale, 0-88% d+Au events.

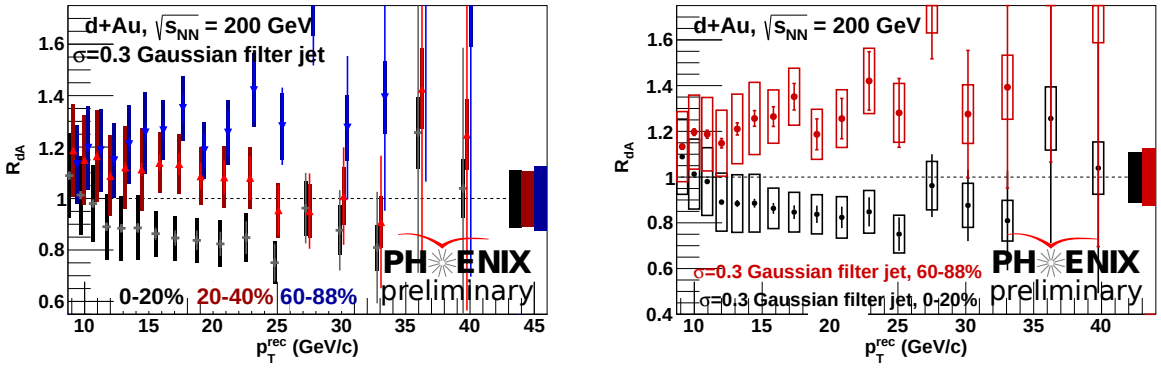


Figure 6.45: R_{dAu} for reconstructed jets at the pp energy scale, super-imposed centrality results.

$$R_{\mathrm{CP}}^{\mathrm{cent}} = \frac{\left(1/N_{\mathrm{coll}}^{\mathrm{cent}}\right)\left(1/N_{\mathrm{evt}}^{\mathrm{cent}}\right)\left(dN^{\mathrm{cent}}/dp_{\mathrm{T}}\right)}{\left(1/N_{\mathrm{coll}}^{60-88\%}\right)\left(1/N_{\mathrm{evt}}^{60-88\%}\right)\left(dN^{60-88\%}/dp_{\mathrm{T}}\right)} \tag{6.43}$$

The R_{CP} is constructed from the ratio of two yields, each of which has been corrected to the pp-equivalent energy scale. The R_{CP} is shown for each of the three non-peripheral centralities in Figures 6.46 and 6.47.

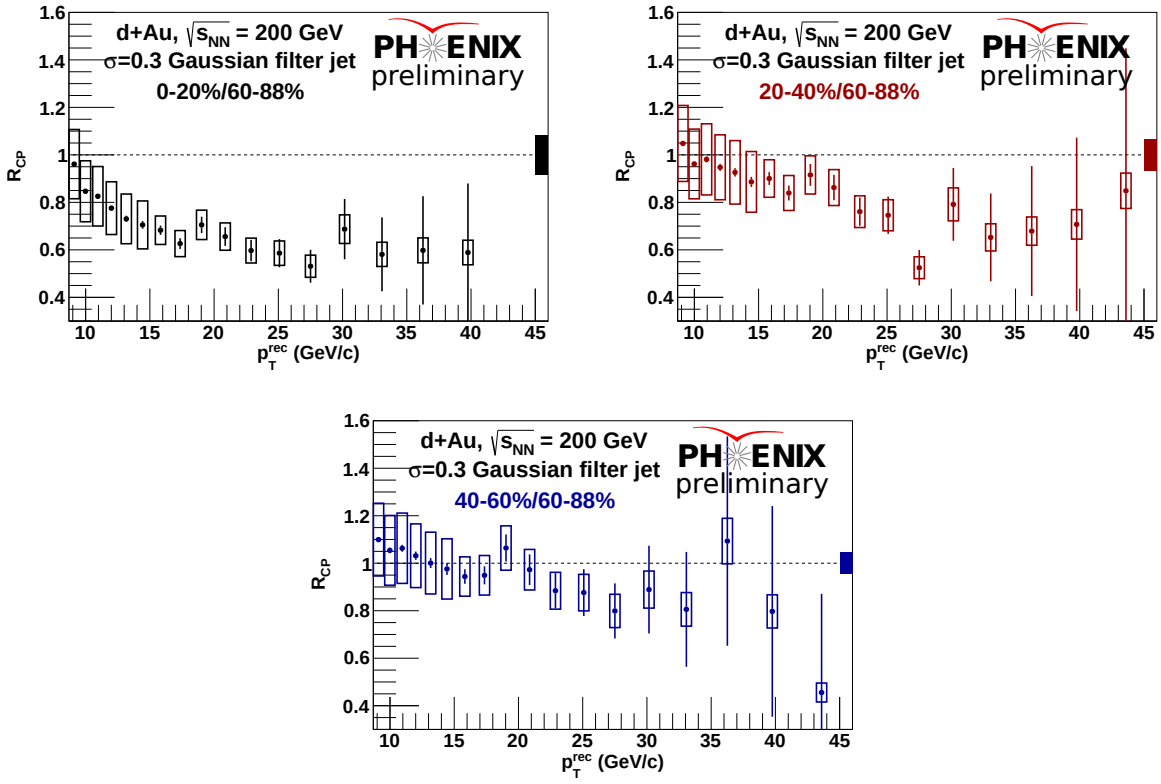


Figure 6.46: R_{CP} for reconstructed jets at the pp energy scale, different centrality selections.

6.6 Systematic Uncertainties & Cross-checks

6.6.1 Fiducial cut cross-check

Since the pp collisions at the PHENIX interaction point have a significantly wider z_{vtx} distribution and poorer z_{vtx} resolution in the BBC than d+Au collisions, it is important to investigate the kinds of jets reconstructed at a large η (relative to the apparent vertex) across both collisions. As a systematic check, consider the ratio of jets inside $|\eta| < 0.20$ (according to Equations 4.2-4.3, at $z_{\mathrm{vtx}} = 0$ cm this corresponds to jets within ±50cm when projected to the Central Arms) relative to the entire jet yield. Any change in this ratio as a function of p_{T} could be taken as a systematic resulting from the z-vertex differences in d+Au vs. pp.

The ratio of fiducially-restricted yields to the total are shown in Figure 6.48 for pp jets and 0-88% d+Au jets. The fraction of jet yields for both types of collisions is independent of p_{T} within

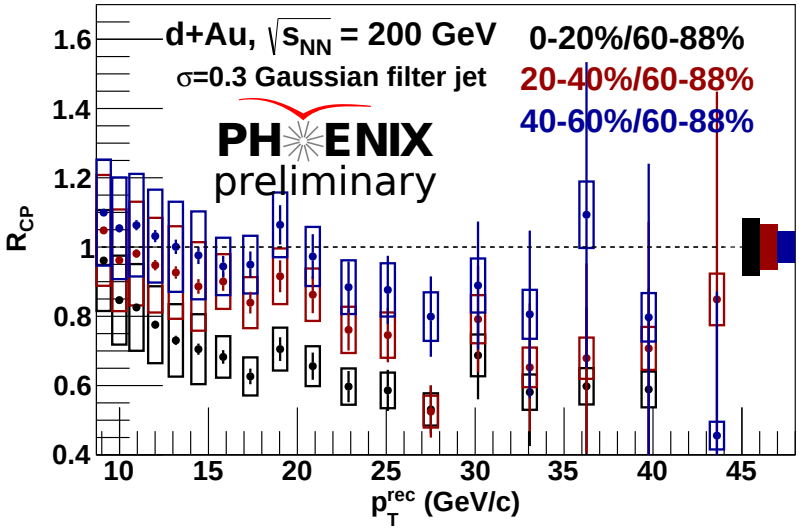


Figure 6.47: R_{CP} for reconstructed jets at the pp energy scale, superimposed centrality selections.

statistics: $72 \pm 1\%$ in pp collisions and $73 \pm 1\%$ in d+Au. This slight difference could arise from a number of sources, including residual differences in the acceptance. Plausibly, it could also arise from a slightly enhanced reconstruction efficiency in d+Au events for jets otherwise near the edge of the acceptance. If this is the case, such effects are included in the definition of the unfolding correction factors. In any case, the possible 1% difference is considered as part of the total 3% systematic uncertainty from residual acceptance differences.

6.6.2 Neutron-tagged yields

As a check against potentially exotic effects in the deuteron [247; 248], we can attempt to reduce the set of d+Au collisions to "p+Au" collisions in which only the proton interacts with the nucleus, by observing the intact spectrator neutron. This can be done using with the PHENIX zero-degree calorimeter in the d-going direction, as was done in a previous work[246].

Neutron-tagged d+Au events are selected by requiring a ZDC North Energy between 60 and 180 GeV. In principle, if the neutron-tagged events are "p+Au" events, the Glauber model can be used to estimate the N_{coll} in events that remain in each centrality bin after the neutron tag. These are

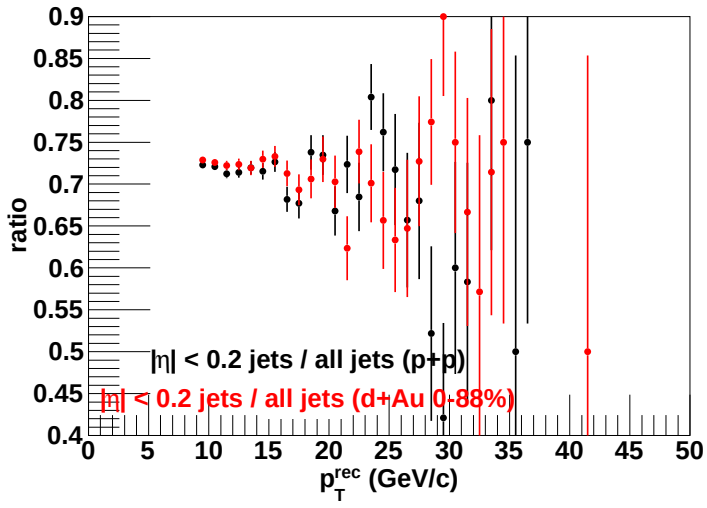


Figure 6.48: Ratio of jet yield inside $|\eta| < 0.2$ fiducial cut in pp and d+Au.

- $N_{\text{coll}}^{p+\text{Au}} = 7.20$ in 0-20%
- $N_{\text{coll}}^{p+\text{Au}} = 5.40$ in 20-40%
- $N_{\text{coll}}^{p+\text{Au}} = 4.04$ in 40-60%
- $N_{\text{coll}}^{p+\text{Au}} = 2.52$ in 60-88%

Figure 6.50 shows the ZDC North energy distribution in d+Au events, with the events selected by the neutron tag bracketed by dotted lines. A clear peak at ≈ 100 GeV, the nominal neutron beam energy, is observed in all centrality bins. As can be seen, the proportion of events with a neutron peak in the ZDC decreases in the more central events. This makes sense, since p+Au events do not have two nucleons in the deuteron to interact with the Au nucleus and thus have an N_{coll} distribution with a much smaller range than that in d+Au collisions.

On the assumption that these events are p+Au events with the given N_{coll}, we can perform a consistency check on the results by measuring the ratio of N_{coll}-scaled yields in p+Au events relative to d+Au events. Formally,

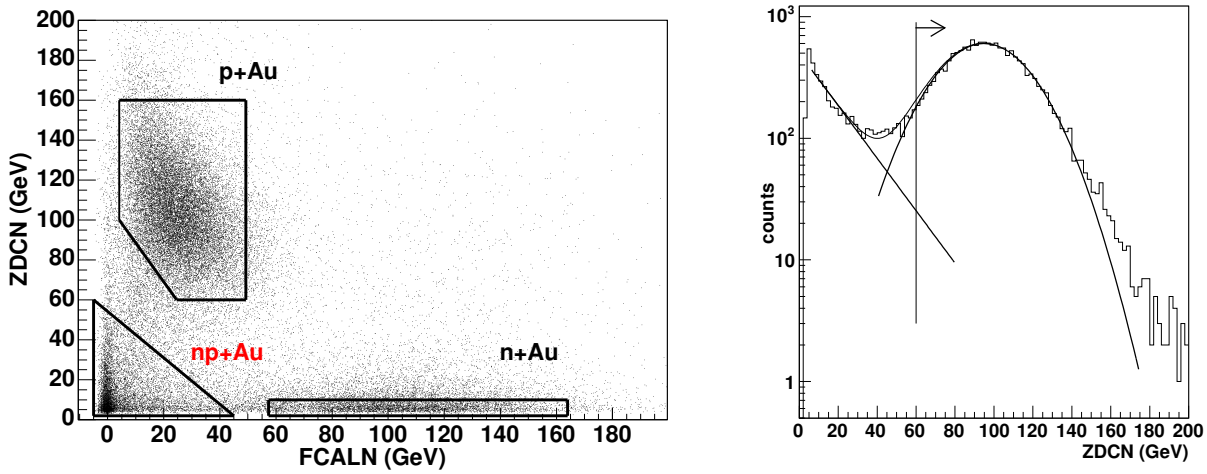


Figure 6.49: Scatterplot of ZDC North Energy vs. FCAL North Energy in 2003 d+Au data, illustrating the separation between d+Au, p+Au and n+Au collisions (left), and the ZDC North energy in d+Au events, demonstrating the single neutron peak (right), from [246].

$$R = \frac{1}{N_{\text{coll}}p\text{+Au}} \frac{1}{N_{evt}^{n-tag}} \frac{dN^{n-tag}}{dp_{\text{T}}} \Bigg/ \frac{1}{N_{\text{coll}}d\text{+Au}} \frac{1}{N_{evt}^{MB}} \frac{dN^{MB}}{dp_{\text{T}}} \tag{6.44}$$

where $N_{\text{coll}}p$+Au and $N_{\text{coll}}d$+Au are the number of binary collisions in p+Au and d+Au collisions, respectively, N_{evt}^{n-tag} and N_{evt}^{MB} are the number of neutron-tagged events and MB events, respectively, and N^{n-tag} and N^{MB} are the yield of jets in the neutron-tagged events and MB events, respectively. The ratio of N_{coll}-scaled yields is shown in Figure 6.51 as a function of reconstructed jet p_{T}. A constant fit to the values of R finds,

- $R = 1.48 \pm 0.02$ in 0-20%
- $R = 1.23 \pm 0.02$ in 20-40%
- $R = 1.07 \pm 0.02$ in 40-60%
- $R = 0.924 \pm 0.01$ in 60-88%

The ratio R increases quickly in more central events. Taken at face value, the 0-20% results indicate a 50% enhancement above the N_{coll}-scaled yields in p+Au collisions, relative to d+Au

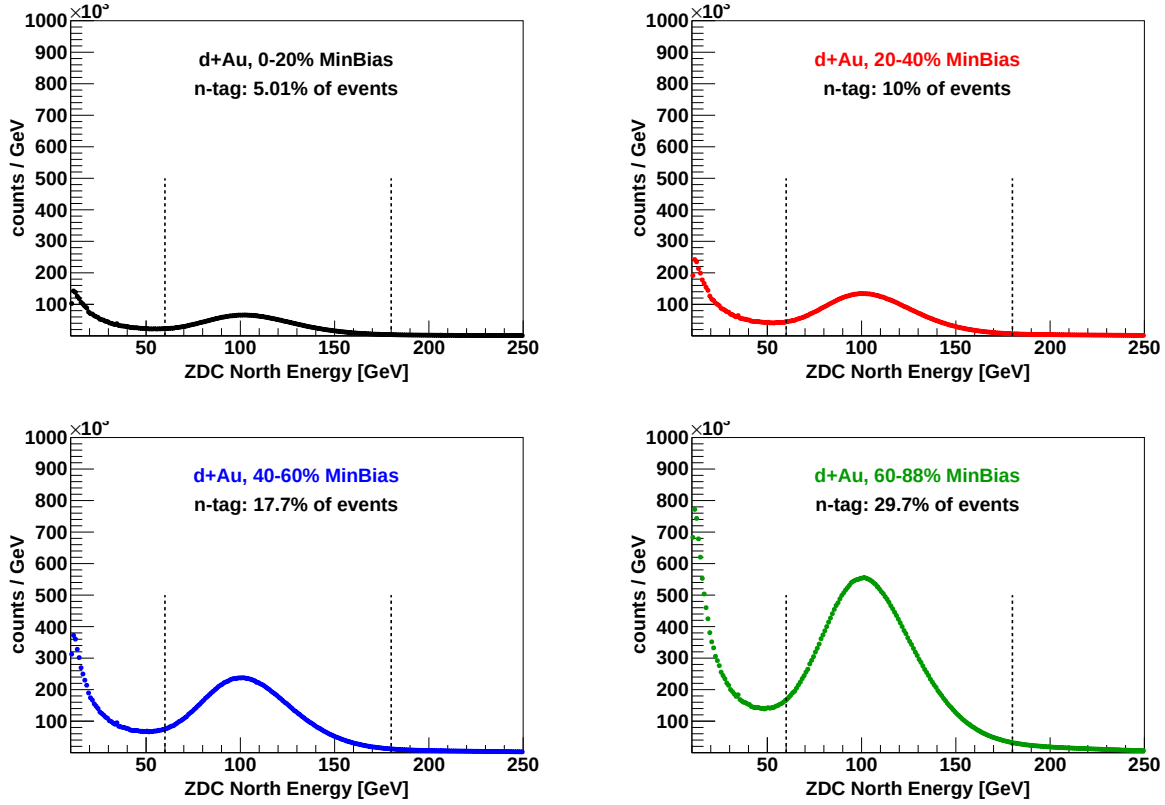


Figure 6.50: ZDC North energy distribution in d+Au data, with each centrality shown separately. The dotted line indicate the neutron-tagged events.

collisions! However, this is not the case. A more detailed analysis of the ZDC North energy distribution and double interaction rate in PHENIX[243] found that the neutron-tagged events in the 0-20% d+Au centrality selection are strongly contaminated, with 58% of them arising from in-time pileup (two or more real d+Au interactions in the same crossing). Thus, the rise in central events is artificial.

However, the contamination is calculated to be only 9% in the neutron-tagged 60-88% events. In this bin, $R = 92\%$, indicating a 8% change in the $N_{\rm coll}$-scaled yield in p+Au collisions relative to d+Au collisions in this centrality bin. Of course, the full extraction of the $R_{\rm pAu}$ is complicated by several additional factors, including a proper treatment of the contamination by d+Au pileup and the falling exponential distribution at small values of the ZDC North Energy. However, if the scale of the effect is 8%, this cannot by itself explain the large enhancement ($\approx 1.3 - 1.4$) seen in

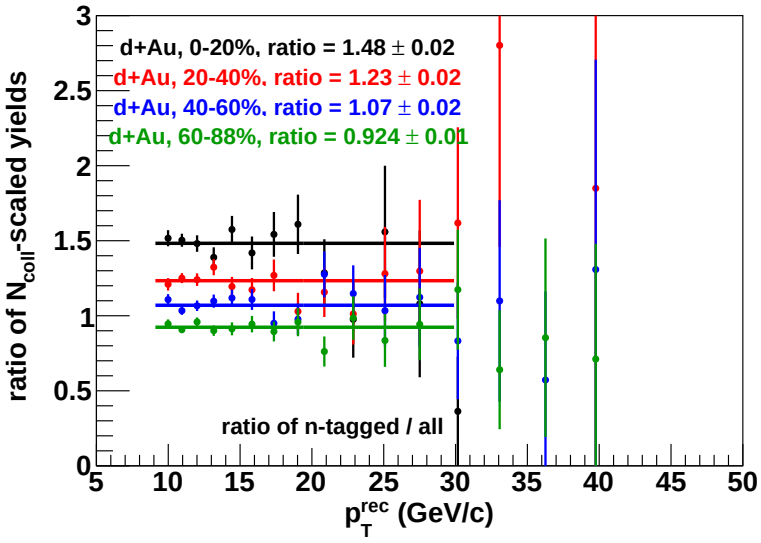


Figure 6.51: Ratio of $N_{\rm coll}$-scaled yields in neutron-tagged events relative to d+Au events, with the neutron-tagged $N_{\rm coll}$ made on the assumption that the neutron-tagged events are p+Au events.

the peripheral $R_{\rm dAu}$.

6.6.3 Overview of systematic uncertainties

Tables 6.8, 6.9 and 6.10 give an overview of the statistical and systematic uncertainties on the jet yields, jet $R_{\rm dAu}$ and jet $R_{\rm CP}$, respectively, as a function of jet $p_{\rm T}$. The errors are categorized following the PHENIX convention of error type "A" (statistical), "B" (uncorrelated point-to-point systematic) and "C" (correlated point-to-point systematic). Statistical uncertainties are indicated by vertical bars that go through the points and most systematic uncertainties are plotted as open boxes with the data point at the center. Although the so-called "normalization" uncertainties from sources such as the σ_{BBC} and $N_{\rm coll}$ are often are properly classified as type "C" errors, they are typically shown at the left or right side of the x-axis instead of being added in quadrature to the systematic error bars on the data points.

The full set of statistical and systematic uncertainties considered in these results are:

- Statistical uncertainties, which are unavoidable in any measurement and are present in the yield, $R_{\rm dAu}$ and $R_{\rm CP}$.

- Uncertainty on the BBC-firing pp cross-section σ_{BBC}, discussed in Section 6.5.2, is p_{T} independent and applies to the pp cross-section and thus the R_{dAu}.

- Uncertainty on N_{coll}, discussed in Section 6.5.1, are p_{T}-independent, but different in the different centrality selections. These uncertainties show up directly in the R_{dAu}. In the R_{CP}, care must be taken since there the uncertainty on N_{coll} is partially correlated between the centrality selection and partially cancels in the ratio.

- Uncertainty on the Bias Factors ($B.F.$), also discussed in Section 6.5.1. Like the uncertainty on N_{coll}, they appear directly in the R_{dAu} but partially cancel in the ratio in the R_{CP}.

- Uncertainty on the BBC trigger efficiency for jets ϵ_{BBC}, discussed in Section 6.3.5, which is separate from the uncertainty on σ_{BBC}. This uncertainty grows with p_{T}, and appears in the pp cross-section and thus the R_{dAu}.

- Uncertainty on the trigger efficiency correction, described in Section 6.3.2. This uncertainty affects all yields, and is centrality- and p_{T}-dependent, becoming smaller at high-p_{T} where the efficiency correction is smaller. This systematic uncertainty partially cancels in the ratio at high-p_{T}, where the efficiency in all d+Au centralities and in pp asymptote to 1.

- Uncertainty on the first-order correction factors C, described in Section 6.4.5. This uncertainty decreases in magnitude at high-p_{T} where the corrections are small, and applies to all d+Au yields, thus affecting the R_{dAu} and R_{CP} as well.

- The upper limit on the residual fake rate, described in Section 6.1.7. This uncertainty applies only to p_{T} bins between 9-12 GeV, and is applied to the d+Au yields, and also the relative ratios of any yields (e.g. it is not double counted in the R_{CP}).

- Systematic uncertainty arising from differences in the z-vertex distribution between d+Au and pp, described in Section 6.6.1. This uncertainty is p_{T} independent and only applies to the R_{dAu}.

- Systematic uncertainty arising from possible changes in the acceptance and/or energy scale between the d+Au and pp parts of the run, described in Sections 6.2.4 and 6.3.4. This uncertainty is p_{T} independent and only applies to the R_{dAu}.

source	type	9-12 GeV	12-15 GeV	12-20 GeV	20+ GeV
statistical	A	1%	1-3%	3-10%	6%-
trigger efficiency	B	5-8%	5-8%	5%	5%
unfolding	B	4-10%	4-10%	4-6%	4-6%
res. fake rate	B	5%	0%	0	0

Table 6.8: Error summary for d+Au and pp yields.

source	type	9-12 GeV	12-15 GeV	12-20 GeV	20+ GeV
statistical	A	1%	1-3%	3-10%	6%-
σ_{BBC}	C	$\leftarrow$ 10% $\rightarrow$			
N_{coll}	C	$\leftarrow$ 6-7% $\rightarrow$			
bias factors	C	$\leftarrow$ 1-6% $\rightarrow$			
ϵ_{BBC}	B	2%	2%	3%	5%
trigger efficiency	B	9-12%	9-12%	7%	7%
unfolding	B	4-10%	4-10%	4-6%	4-6%
res. fake rate	B	5%	0%	0	0
z-vertex dep.	C	$\leftarrow$ 1% $\rightarrow$			
acceptance/E-scale	C	$\leftarrow$ 3% $\rightarrow$			

Table 6.9: Error summary for R_{dAu} results.

source	type	9-12 GeV	12-15 GeV	12-20 GeV	20+ GeV
statistical	A	1%	1-3%	3-10%	6%-
N_{coll}	C	← 6-7% →			
bias factors	C	← 1-6% →			
trigger efficiency	B	9%	9%	5%	5%
unfolding	B	4-10%	4-10%	4-6%	4-6%
res. fake rate	B	5%	0%	0	0

Table 6.10: Error summary for $R_{\rm CP}$ results.

Chapter 7

Centrality determination in p+Pb collisions

I will have such revenges on you both,
That all the world shall–I will do such things,–
What they are, yet I know not: but they shall be
The terrors of the earth. You think I'll weep
No, I'll not weep

King Lear, Act II, Scene IV

7.1 Introduction

In this chapter, we aim to to understand the composition of events in the minimum bias p+Pb event sample, and to categorize the collision geometry of these events through a measure of soft underlying event activity.

In Pb+Pb collisions in ATLAS, the sum of the transverse energy in both FCal is used in the centrality determination[249], and similar measurements of the energy or multiplicity at large rapidities has also been used at other LHC and RHIC experiments. In p+Pb collisions, the Pb-going FCal transverse energy, $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, is used to the measurement of soft underlying event activity. p+Pb events are sorted according to their $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ and each p+Pb event is categorized by their centile in the full $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution. This centile is called the *centrality* of the event, and it ranges from 0%

to 100%. By convention, *central* events have a small centrality % (and are correlated with small b) and larger event activity, while *peripheral* events have a large centrality % and softer event activity (and are correlated with larger b).

On the other hand, the possible event geometries and their relative probability are generated by the Glauber model. Events are classified in terms of the number of participating nucleons in the p+Pb collision (N_{part}, of which $N_{\mathrm{part}}-1$ are in the nucleus, resulting in $N_{\mathrm{coll}} = N_{\mathrm{part}}-1$ individual nucleon-nucleon collisions). Increasing N_{part} is correlated with decreasing impact parameter b. The *raison d'être* of the centrality analysis is to quantitatively determine the correlation between $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ and N_{part}. A more detailed discussion of the the geometry of nuclear collisions and the Glauber model can be found in Section 2.2.1.

Once the number of participating nucleons is known, the yields of jets in different centrality selections can be compared with the superposition of N_{coll} incoherent p+p collisions. Deviations from this binary collision scaled expectation indicate the presence of nuclear effects in p+Pb collisions, such as initial state energy loss or the nuclear modification of parton distribution functions, including from the onset of partonic saturation.

A centrality analysis using the default Glauber model has been successfully used for 2010 and 2011 Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV in ATLAS. In each centrality bin, the N_{coll}-scaled pp reference was used as the baseline against which to test the rate of hard probe production. ATLAS has shown that the rate of high-p_{T} charged hadrons[250] and jets[60] fall significantly below the binary scaled expectation. On the other hand, the yields of colorless probes such as photons[251] and Z bosons[252] were shown to be consistent with the N_{coll}-scaling expectation. Taken together, these results indicate a strong final state effect in which highly energetic colored partons are significantly suppressed in the QGP.

In p+Pb collisions in ATLAS, the application of the "traditional" centrality analysis used in heavy ion collisions is complicated by several factors:

- The ATLAS trigger is sufficiently efficient that the peripheral part of the p+Pb cross-section is sampled in detail. This means that inelastic events which are not described by the Glauber model could be included in the minimum bias event sample. For example, in the Glauber model, single diffractive excitation is treated as the incoherent sum of p-nucleon diffraction, whereas the full set of single diffractive events may include contributions from coherent diffrac-

tion and photonuclear dissociation of the proton. More generally, the peripheral end of the observed p+Pb events is likely to be a mixture of these event classes, each of which will have a different trigger efficiency. Thus, these events must be identified before a proper mapping of the Glauber cross-section onto the data can be performed.

- It has been suggested[253] that a nucleon undergoing many inelastic scatterings with other nuclei will result in a slow shift of the COM frame of the produced soft particles, increasing with the number of collisions. This effectively appears as an acceptance effect, where the centrality-measuring detector has an $N_{\rm part}$-dependent acceptance. In Pb+Pb collisions, this effect may cancel in the average, since a roughly equal number of nucleons on either side are participating. However, in p+Pb collisions, where the proton can undergo many collisions at small impact parameter, it is possible that this is a non-trivial effect.

- At sufficiently high-energies, where the nucleon-nucleon cross-section is dominated by the low-x gluons, fluctuations in the configuration of the proton event to event could become important[82; 83; 84]. The standard Glauber model does not account for this possibility. Thus, Glauber-Gribov models with event-by-event fluctuations of the effective transverse size of the proton (and thus $\sigma_{\rm NN}$) may be necessary to properly describe the distribution of collision geometries. In this case, changes in the configuration of the proton wavefunction would introduce additional fluctuations into the number of participating nucleons at a given impact parameter.

Each of these issues will be addressed here.

The remainder of this chapter is outlined as follows: the remainder of this chapter (Section 7.1.1) introduces the basics of diffractive processes and their identification and relevance to the present analysis. The Glauber and Glauber-Gribov models and the MC parameters used are detailed in Section 7.2. The selection of the data and correction of the $\Sigma E_{\rm T}^{\rm Pb}$ distributions used in the global centrality analysis are discussed in Section 7.3. The Monte Carlo, which is used to understand the efficiency of p+Pb collisions and constrain the $\Sigma E_{\rm T}^{\rm Pb}$ from pp-like collisions, is described in Section 7.4. The models of particle production and fits to data are described in Section 7.5. An overview of the systematic uncertainties is given in Section 7.6. The geometric quantities are extracted in Section 7.7.

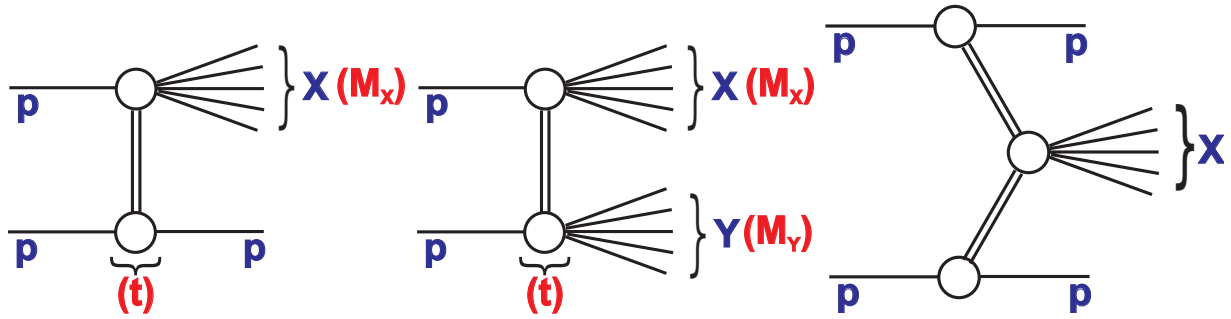


Figure 7.1: Schematic diagrams of single-diffractive (left), double-diffractive (center) and central diffractive (right) pp events, from [254].

7.1.1 Diffractive events

In addition to the class of inelastic proton-nucleus collisions where an exchange of color takes place, there are contributions to the total cross-section mediated by the exchange of vacuum quantum numbers[255]. These "diffractive" events are indicated by the presence of a large rapidity gap $\Delta\eta$ between the produced particles, and the mass of the two diffractive systems M_X and M_Y. For example, in a single diffractive process, $M_Y = m_p$ by convention and $\Delta\eta = -\log\left(M_X^2/s\right)$. Figure 7.1 illustrates the diagrams corresponding to single-, double- and central-diffraction, including the diffractive masses of the dissociating systems. The vertical axis can be thought of as rapidity, revealing the topology of the events in terms of the rapidity dependence of the event activity.

Thus, to identify diffractive events, pPb events are caterogized according to their topology. Within the ATLAS acceptance, pseudorapidity gaps are constructed by examining the 49 pseudorapidity intervals of size $\delta\eta = 0.2$ between $\eta = \pm 4.9$. At the MC or truth level, a pseudorapidity interval is considered empty if it has no particles above some threshold $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$. In data, a pseudorapidity interval is empty if it has no calorimeter clusters or reconstructed tracks with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$, where ${p_{\mathrm{T}}}^{\mathrm{cut}}$ is chosen to be the same in MC and data.

In each p+Pb event, the three types of pseudorapidity gaps considered are defined as follows (with the pilot run / period A kinematic convention),

- $\Delta\eta^{\mathrm{Pb}}$, the nuclear pseudorapidity gap, is 0.2 times number of sequential empty $\delta\eta$ intervals starting from $\eta = +4.9$ (the Pb-going pseudorapidity detector edge). Alternately, this gap is the distance from $\eta = +4.9$ to the first particle with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$, rounded down to the nearest

multiple of 0.2.

- $\Delta\eta^{\mathrm{p}}$, the proton pseudorapidity gap, is 0.2 times the number of sequential empty $\delta\eta$ intervals starting from $\eta = -4.9$ (the proton-going pseudorapidity detector edge). Alternately, this gap is the distance from $\eta = -4.9$ to the first particle with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$, rounded down to the nearest multiple of 0.2.

- $\Delta\eta^{\mathrm{F}}$, the *forward pseudorapidity gap*, is defined as the largest of the pseudorapidity gaps on either side of the detector, $\Delta\eta^{\mathrm{F}} = \max(\Delta\eta^{\mathrm{p}}, \Delta\eta^{\mathrm{Pb}})$. This gap variable is most useful in symmetric systems such as *pp* collisions, since in p+Pb collisions the $\Delta\eta^{\mathrm{p}}$ and $\Delta\eta^{\mathrm{Pb}}$ distributions are not a priori expected to be the same. However, they are still useful as an overall indicator of diffractive activity.

In the language of pseudorapidity gaps, we can model the various diffractive components of the p+Pb inelastic cross-section as follows:

- p+Pb $\to p+X$, single diffractive excitation of the nucleus. This process is consistent with a large pseudorapidity gap in the proton-going direction ($\Delta\eta^{\mathrm{p}} \gg 0$).

- p+Pb $\to X$+Pb, single diffractive excitation or electromagnetic dissociation of the proton. This process is consistent with a large pseudorapidity gap in the nucleus-going direction ($\Delta\eta^{\mathrm{Pb}} \gg 0$). The single diffractive excitation includes contributes from incoherent (where the proton interacts with one of the nuclei) and coherent (where the proton interacts diffractively with the nucleus as a whole) diffraction.

- p+Pb $\to X+Y$, double diffraction, where both systems dissociate but there is no central particle production. Because the beam rapidity ($y_B = 8.5$ for 5.02 TeV nucleon-nucleon collisions) extends for several units beyond the detector acceptance, it is possible and even likely that the "central gap" is actually reconstructed as an edge gap. Thus, double diffractive events are often characterized by $\Delta\eta^{\mathrm{p}} \gg 0$ and/or $\Delta\eta^{\mathrm{Pb}} \gg 0$ in ATLAS.

Rarer processes such as central diffraction (p+Pb $\to p$+Pb+X) are not considered here. The diffractive portion of hadronic cross-sections at LHC energies are an important part of the total

inelastic cross-section. For example, p+p MC generators model the single and double diffractive contributions to be as large as $\sim 20\%$ and $\sim 10\%$ of the total pp inelastic cross-section, respectively.

ATLAS has previously measured the cross-section for forward pseudorapidity gaps $d\sigma/d\Delta\eta^F$ over eight units of pseudorapidity[254] in $\sqrt{s} = 7$ TeV pp collisions, reaching a plateau of ~ 1 mb per unit of pseudorapidity at large $\Delta\eta^F$. In that work, values of $p_{\mathrm{T}}^{\mathrm{cut}} = 200$-$800$ MeV were considered. In this analysis, we use $p_{\mathrm{T}}^{\mathrm{cut}} = 200$ MeV is chosen to give the greatest discrimination between diffractive and non-diffractive events.

7.2 Glauber models

In the non-relativistic Glauber model[67], inelastic A+B collisions are treated as an incoherent set of individual nucleon-nucleon collisions, each of which are assumed to be independent of one another. Each A+B event is classified by the number of binary nucleon-nucleon collisions N_{coll} and the total number of participating nucleons N_{part}. In p+Pb collisions, these two are related simply through $N_{\mathrm{part}} = N_{\mathrm{coll}} + 1$.

The probability distribution of N_{part}, $P(N_{\mathrm{part}}) = \frac{1}{N_{evt}} \frac{dN}{dN_{\mathrm{part}}}$, is generated through a Glauber MC[75; 76], which samples the available range of impact parameters b, populates nucleons in the nucleus according to a description of the nuclear wavefunction, and uses a semiclassical description of inelastic nucleon-nucleon collisions to determine how many nucleons in the Pb nucleus collide with the proton. An example of a *central* collision with many participating nucleons and a relatively small impact parameter is shown in Figure 7.3 and an example of a *peripheral* collision with few participating nucleons and a relatively large impact parameter is shown in Figure 7.4.

One of the main geometric parameters in the Glauber model is the nucleon-nucleon (proton-proton) cross-section, which is not measured at $\sqrt{s} = 5.02$ TeV, must be interpolated from fits to world data[1; 74]. It is taken to be

$$\sigma_{NN}(\sqrt{s} = 5.02 \text{ TeV}) = \sigma_{inel}^{p+p} = 70 \text{ mb } (\pm 5 \text{ mb}), \tag{7.1}$$

where the uncertainty comes from the poor knowledge and model dependence of σ_{NN}. σ_{NN} is varied by this uncertainty, and the changes in the geometric quantities are used as a contribution to the systematic uncertainty on the results.

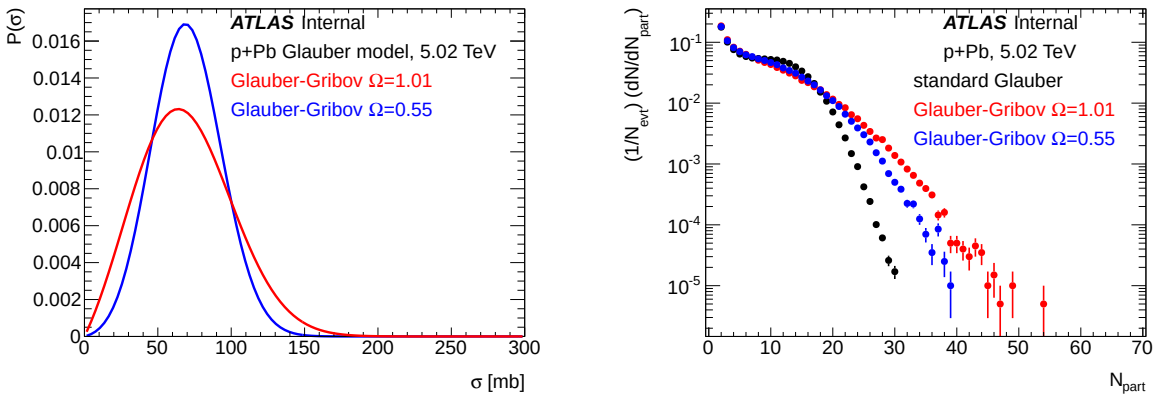


Figure 7.2: $P(\sigma)$ in the Glauber-Gribov models (left) and N_{part} distributions in the default Glauber and two Glauber-Gribov models (right).

7.2.1 Glauber-Gribov extension

As described above, a more recent extension of the Glauber model[84; 82] known as the Glauber-Gribov model attempts to incorporate fluctuations in the strength of the nucleon-nucleon interaction event by event. In these models, the fluctuations in N_{part} at fixed b arise not only from the geometric fluctuations of nucleon positions, but from fluctuations in the configuration of the proton (which is frozen for all of the individual proton-nucleon collisions).

This effect is described using a per-event probability distribution of the total cross-section σ_{tot} according to

$$P(\sigma_{\text{tot}}) = N \frac{\sigma_{\text{tot}}}{\sigma_{\text{tot}} + \sigma_0} \exp\left(-\frac{(\sigma_{\text{tot}}/\sigma_0 - 1)^2}{\Omega^2}\right) \tag{7.2}$$

where σ_0 controls the nominal mean nucleon-nucleon cross-section $\langle\sigma_{\text{tot}}\rangle$, Ω is a dimensionless parameter that describes the magnitude of the cross-section fluctuations, and N is a normalization factor to ensure $\int P(\sigma_{\text{tot}})d\sigma_{\text{tot}} = 1$. The inelastic nucleon-nucleon cross-section σ_{NN} is taken to be a fixed fraction of the total cross-section σ_{tot} according to $\sigma_{\text{NN}} = \lambda\sigma_{\text{tot}}$, so that $P(\sigma_{\text{NN}}) = \frac{1}{\lambda}P(\sigma_{\text{tot}}/\lambda)$. For each choice of the parameters σ_0 and Ω, we choose λ such that $\langle\sigma_{\text{NN}}\rangle = 70$ mb, as in the default Glauber.

Two different choices of σ_0 and Ω are considered. In the first choice, $\Omega = 0.55$ is taken from interpolating the results for $\sqrt{s} = 1.8$, 9 and 14 TeV in [82] to 5.02 TeV, and $\sigma_0 = 78.6$ mb is chosen

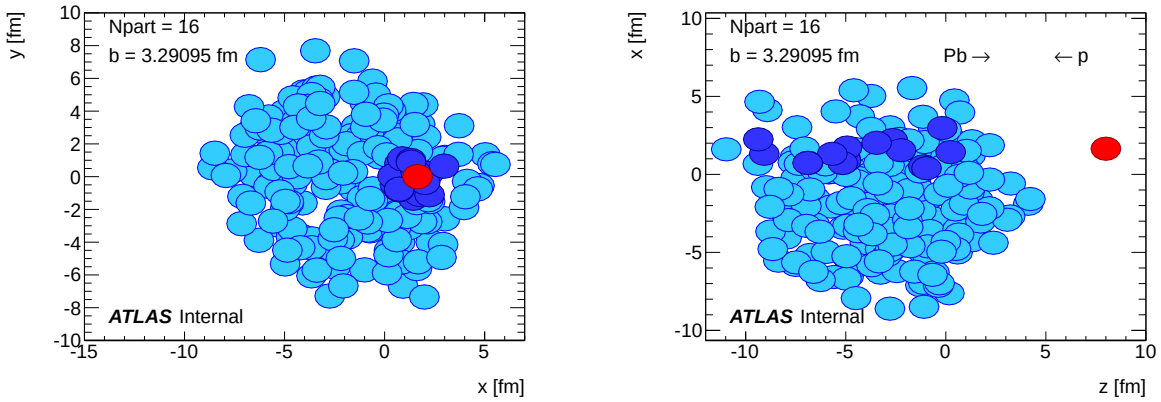


Figure 7.3: Example of a central collision with $N_{\text{part}} = 16$.

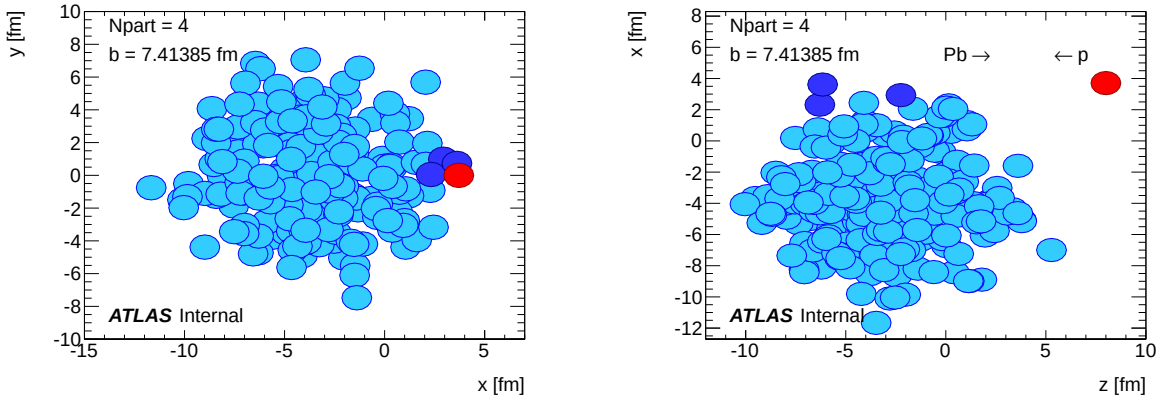


Figure 7.4: Example of a peripheral collision with $N_{\text{part}} = 4$.

to reproduce $\langle\sigma_{\text{tot}}\rangle = 86$ mb, which is consistent with [74]. In the second choice, Ω is taken to be $= 1.01$ from [84] and $\sigma_0 = 72.5$ is chosen to reproduce $\langle\sigma_{\text{tot}}\rangle = 94.8$ from the recent measurement in [73]. To give the desired $\langle\sigma_{\text{NN}}\rangle$, the λ are chosen to be 0.82 for $\Omega = 0.55$ and 0.74 for $\Omega = 1.01$.

Thus, each of the three Glauber/Glauber-Gribov models gives a different $P(N_{\text{part}})$ distribution for inelastic p+Pb collisions.

Figure 7.2 shows the $P(\sigma_{\text{NN}})$ distributions for the two choices of Ω, and compares the N_{part} distributions from the Glauber-Gribov models to that in the default Glauber model.

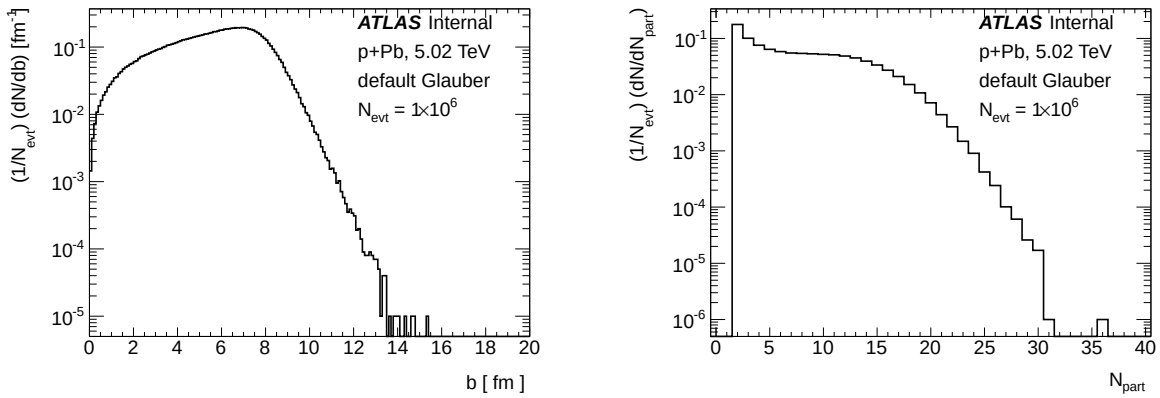


Figure 7.5: Distribution of impact parameter b (left) and number of participants N_{part} (right) from the default Glauber simulation.

7.2.2 Glauber MC parameters

A modified PHOBOS Glauber Monte Carlo code[75; 76] is used to simulate the geometry of p+Pb collisions. The default parameters used to populate the nucleus with nucleons drawn from the nucleon wavefunction are:

- Woods-Saxon parameters radius $R = 6.62$ fm and skin depth $a = 0.546$ fm to describe the Pb nuclear wavefunction.
- Hard core radius $d_{min} = 0.4$ fm

As described above, there are three different descriptions of the nucleon-nucleon cross-section:

1. in the default Glauber model, $\sigma_{NN} = 70$ mb is fixed in each event (i.e. $P(\sigma) = \delta(\sigma - \sigma_{NN})$)
2. in the Glauber-Gribov model with $\Omega = 1.01$, $P(\sigma_{NN})$ in each event is given by Equation 7.2 with $\sigma_0 = 72.5$ mb and $\lambda = 0.82$
3. in the Glauber-Gribov model with $\Omega = 0.55$, $P(\sigma_{NN})$ in each event is given by Equation 7.2 with $\sigma_0 = 78.6$ mb and $\lambda = 0.74$

10^6 events were simulated for each model. Figure 7.5 shows the impact parameter and N_{part} distribution for the standard Glauber model, and Figure 7.2 compares the N_{part} distributions among the three models.

If the p+Pb events recorded in data had exactly the same mixture of non-, single- and double diffractive events as are implied by the Glauber model, these simulations would immediately give the mean number of participants, $\langle N_{\rm part}\rangle$ and nuclear overlap function $\langle T_{\rm AB}\rangle$ ($=\langle N_{\rm part}\rangle/\sigma_{NN}$) (or $=\langle N_{\rm part}\rangle/\langle\sigma_{\rm NN}\rangle$ in the Glauber-Gribov case) for all inelastic p+Pb collisions. For this theoretical "0-100%" event sample, which has no inefficiency and no contamination from p+Pb events not described by the Glauber model,

- $\langle N_{\rm part}\rangle = 7.88$, 8.12 and 7.98 for 0-100% events in detauls Glauber, Glauber-Gribov $\Omega = 1.01$ and $\Omega = 0.55$, respectively
- $\langle T_{AB}\rangle = 0.0983$ mb^{-1}, 0.102 mb^{-1} and 0.0997 mb^{-1} for 0-100% events in default Glauber, Glauber-Gribov $\Omega = 1.01$ and $\Omega = 0.55$, respectively

It can be seen on the right plot in Figure 7.2 that although the $N_{\rm part}$ distributions between the three models have a very different shape, for minimum bias p+Pb collisions, the mean $N_{\rm part}$ is consistent to 3% between the models. Of course, since ATLAS is slightly inefficient for the most peripheral p+Pb collisions, the mean $N_{\rm part}$ (T_{AB}) in the observed data will be slightly higher (lower) than these.

7.3 p+Pb data selection

This section describes the data reconstruction, trigger and minimum bias event selection, and corrected $\Sigma E_{\rm T}^{\rm Pb}$ distribution used in the global fits. In the 2013 data, the measured $\Sigma E_{\rm T}^{\rm Pb}$ in each event is corrected for a vertex-dependent acceptance factor and run-dependent differences in the energy pedestal and scale, so that the centrality selection cuts behave in a uniform manner throughout the p+Pb run period. Since the 2012 pilot data consisted of a single short run under different conditions, the more detailed corrections are not applied to the $\Sigma E_{\rm T}^{\rm Pb}$ distribution in the pilot p+Pb data. Thus, the global fits are applied to both 2012 and 2013 datasets separately.

7.3.1 Datasets and statistics

Table 7.1 summarizes the datasets and statistics used in this analysis. In all three p+Pb run periods, $E = 4$ TeV protons collided with Pb nuclei with $E = 4\times\frac{Z}{A} \approx 4\times\frac{82}{208} = 1.58$ TeV per

Description	**Dataset**		
	Runs	N_{evt} (p+Pb MB)	N_{evt} (diffractive analysis)
pilot run, 2012	data12_hip.00210184.physics_MinBias.merge.NTUP_HI.f510_m1277/		
	1	2.14×10^6	-
period A, 2013	data13_hip.0021*.physics_MinBias.merge.NTUP_HI.f*_m1312/		
	16	1.50×10^7	2.65×10^6
period B, 2013	data13_hip.0021*.physics_MinBias.merge.NTUP_HI.f*_m1312/		
	16	2.80×10^7	2.26×10^6

Table 7.1: 5.02 TeV p+Pb datasets used in this analysis

nucleon, resulting in a collision system with $\sqrt{s_{NN}} = 5.02$ TeV in which the center of mass frame is shifted by $\Delta y = 0.465$ in the direction of the proton beam. The beam orientations in the three run periods are:

- pilot run (September 2012): Beam 1 is Pb ions, Beam 2 is protons. Protons go from $+\eta$ to $-\eta$ / in the "A→C" direction, according to the ATLAS convention.
- period A (first half of 2013 p+Pb running): Same kinematics as the pilot run.
- period B (second half of 2013 p+Pb running): Reversed kinematics. Beam 1 is protons, Beam 2 is Pb ions. Protons go from $-\eta$ to $+\eta$ / in the "C→A" direction, according to the ATLAS convention.

The three run periods correspond to approximately 1 μb^{-1}, 17nb^{-1} and 14nb^{-1} of data delivered by the machine, but an offline luminosity determination is not yet available. For more details of the LHC performance during the 2013 p+Pb run, see Section 3.3.2.

Events were only analyzed in the luminosity blocks (*lumiblocks*, the quantum of time in ATLAS data quality assurance) in which the beams were stable and all major components of the the ATLAS detector were determined to be in good operational condition. Specifically, lumiblocks were selected from version 60 of the official ATLAS Good Run List (GRL) for p+Pb running.

7.3.2 Reconstruction and tracking

The p+Pb data were reconstructed with low-μ version of the Minimum Bias reconstruction, and *LowPt* settings for the track and vertex reconstruction[256; 232].

However, for the final track selection, the following offline tracking cuts were used:

- Track author requirement that corresponds to inside-out tracking algorithms only (implemented by requiring *trk_patternReco1 & 1* or *trk_patternReco2 & 16*).
- $p_{\mathrm{T}} > 100$ MeV, $|\eta| < 2.5$
- $\geq 2, 4, 6$ SCT hits for $p_{\mathrm{T}} \in$ 100-200, 200-300 and > 300 MeV, respectively
- ≥ 1 Pixel hits
- At least one hit in the B-layer if one is expected.
- $|d_0|\,, |z_0 \sin\theta| < 1.5$ mm, where d_0 and z_0 are constructed with respect to the primary vertex.
- $|\sigma_{d_0}/d_0|\,, |\sigma_{z_0 \sin\theta}/z_0 \sin\theta| < 3$. Note that in constructing the uncertainty $\sigma_{z_0 \sin\theta}$, the full covariance matrix between z_0 and θ is used.

7.3.3 Topological cluster selection

Cells in the ATLAS calorimeter are combined via a topological clustering algorithm[257; 258] and corrected to the hadronic-scale energy. Only topological clusters (topoClusters) with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}} =$ 200 MeV are considered. It is well known that topoClusters with low p_{T} can be reconstructed from nearby cells with an upward fluctuation due to electronic noise. In the diffractive analysis, noisy clusters can disrupt long pseudorapidity gaps and cause incorrectly reconstructed $\Delta\eta^{\mathrm{p}}$ and $\Delta\eta^{\mathrm{Pb}}$ variables.

Thus, an additional selection is required to reject noisy topoClusters. This is done by considering the *cell significance* $S = E^{\mathrm{cell,max}}/E^{\mathrm{noise}}$ of the topoCluster, which is defined as the ratio of the energy of the highest-energy cell in the cluster to the RMS of the expected noise in cells in that region of the detector. Thus, the proper reconstruction of topoClusters depends on the correct modeling of the electronic noise. $S > 4$ is required in the default reconstruction. However, this cut

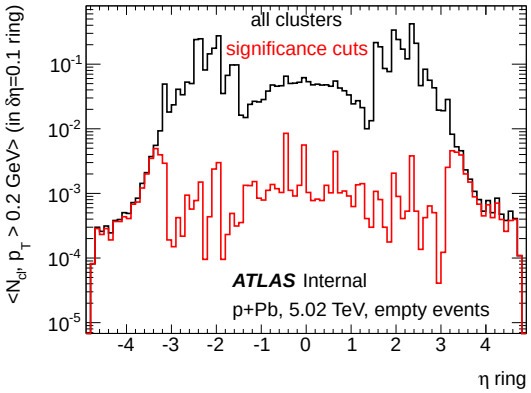


Figure 7.6: Mean number of clusters N_{cl} with $p_{\mathrm{T}} > 0.2$ GeV in empty events for all clusters (black) and after the subdetector-dependent significance cuts (red), in 2012 data.

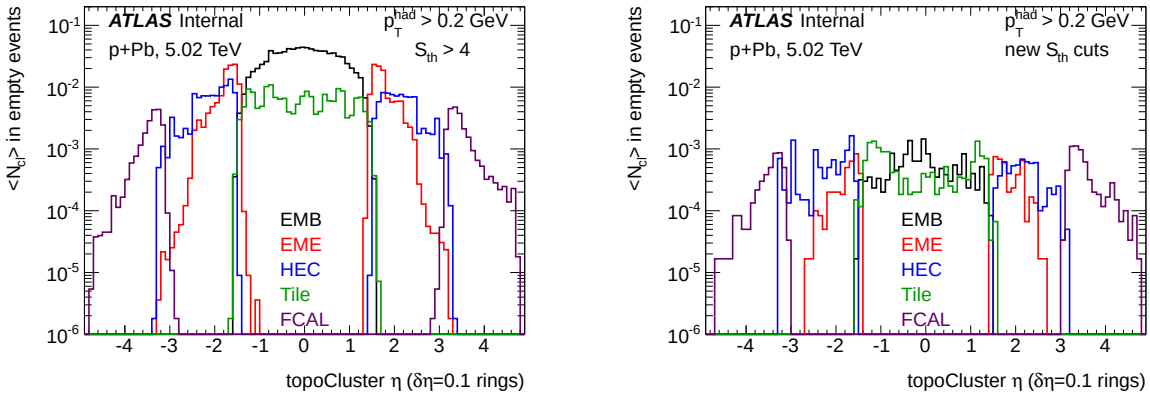


Figure 7.7: Mean number of clusters N_{cl} with $p_{\mathrm{T}} > 0.2$ GeV in empty events for all clusters (left) and after significance cuts (right), in 2013 data. Contributions from each sub-detector are shown in different colors.

in insufficient in reducing the rate of noisy clusters to an acceptable level. Thus, a new set of cell significance threshold cuts are determined.

These cuts are calibrated from examining empty events in data, selected by the `EF_rd0_empty_NoAlg` trigger, in which there is no physics event activity and only topoClusters from electronic noise are present. The mean number of clusters in each $\delta\eta = 0.1$ interval with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$ per event are plotted in Figures 7.6 and 7.7. For the default reconstruction requirement

for $S > 4$, the mean number of cluster per interval in empty events is observed to be large – these are shown as the black histogram in Figure 7.6 and the left plot on Figure 7.7, where they are broken down by subdetector – reaching $\langle N_{\mathrm{cl}} \rangle \sim 0.2$ in the Hadronic End Cap (HEC) in 2012 data, and $\langle N_{\mathrm{cl}} \rangle \sim$ 0.3-0.4 at mid-rapidity in the 2013 data. Then, the minimum requirement on S is adjusted in each subdetector region until the rate of noisy clusters per interval is $\langle N_{\mathrm{cl}} \rangle \approx 10^{-3}$ and is roughly constant as a function of η.

In the 2012 data, the description of the electronic noise in the HEC was found to be substantially smaller than what was observed in data, resulting in much larger tails in the cell significance distribution for noisy clusters. In this data, a simplified set of cell significance cuts was used,

- $S \geq 5.0$ for clusters originating in the EM Barrel, EM Endcap and Tile Calorimeter
- $S \geq 8.0$ for clusters originating in the Hadronic Endcap
- $S \geq 4.0$ for clusters originating in the FCal

The red histogram in Figure 7.6 shows the frequency of noisy clusters with these cuts in 2012 data.

In the 2013 data, the cuts used are

- $S \geq 5.0$ for clusters originating in the EM Barrel
- $S \geq 4.8$ for clusters originating in the EM Endcap
- $S \geq 4.6$ for clusters originating in the Hadronic Endcap
- $S \geq 5.4$ for clusters originating in the Tile Calorimeter
- $S \geq 4.4$ for clusters originating in the FCal

The right plot in Figure 7.7 shows the frequency of noisy clusters with these cuts in 2013 data.

7.3.4 Trigger and event selection

Three level-1 (L1) and event filter (EF) triggers were used in the data. In the low-μ pilot p+Pb run, the trigger menu was simple and all data were taken that fired the L1 `MBTS_2` trigger, which required ≥ 2 hits on either side of the Minimum Bias Trigger Scintillator (MBTS) detectors,

- `L1_MBTS_2`, in the pilot run

In the period A and period B 2013 running, there were two dedicated triggers used for minimum bias studies. Both were prescaled. These are:

- `EF_mbMbts_1_1`, which required ≥ 1 hit on *each* side of the MBTS. This trigger was used to select events for the final p+Pb minimum bias event definition and centrality determination. The difference in the minimum bias triggers between the pilot and period A/B runs selects a slightly different class of peripheral p+Pb events, which is discussed later.

- `EF_mbMbts_2`, which is less restrictive and requires only ≥ 2 hits in the MBTS total. This trigger was used to select an event sample within which to investigate the relative contribution of diffractive events in p+Pb.

As a reminder, all events were required to lie within the ATLAS GRL. For the `L1_MBTS_2`-selected pilot data and `EF_mbMbts_1_1`-selected 2013 data, a further set of event selection cuts were applied to define the class of minimum bias p+Pb events. These are,

1. Good MBTS timing, corresponding to less than 10 nanosecond difference in the time reported by MBTS A and C; e.g. $|\Delta t| = |t_A - t_C| < 10$ ns, where t_A, t_C are non-zero values.

2. At least one reconstructed vertex.

3. At least two tracks with $p_{\mathrm{T}} > 100$ MeV, where tracks are required to pass the cuts listed in Section 7.3.2.

4. To remove events consistent with pileup (two or more p+Pb collisions), no more than one reconstructed vertex with $\Sigma p_T^{assoc,track} > 5$ GeV is allowed. Events with two or more vertices, each of which has 5 GeV or more of tracks associated with it, are rejected.

One additional cut is needed to be able to perform the centrality analysis that has an important effect on the diffractive contributions to the p+Pb cross-section. Events consistent with the diffractive excitation of the proton (recognized by a long pseudorapidity gap on the nucleus-going side) do not deposit any energy in Pb-going FCal, and thus cannot be properly mapped to a particular value of N_{part}. For this reason, we reject events consistent with diffractive excitation of the proton, as given by:

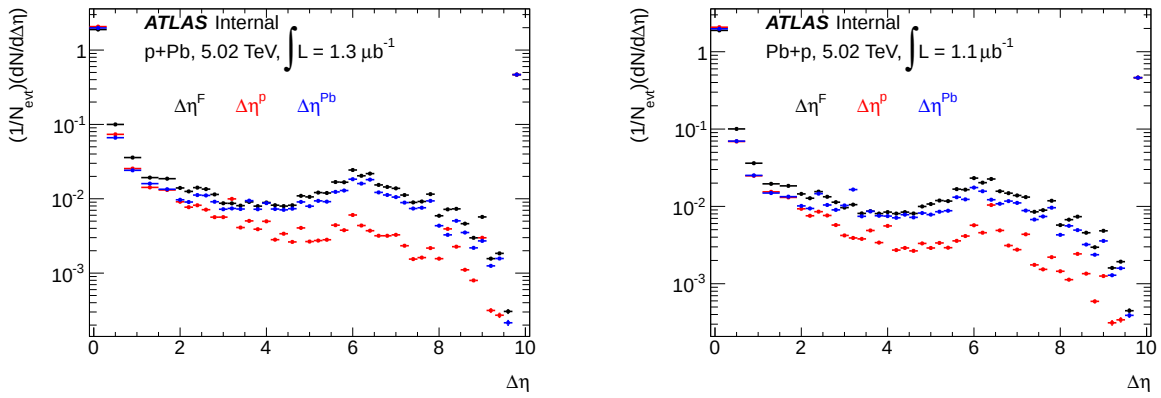


Figure 7.8: Forward, proton-side and Pb-side pseudorapidity gaps in period A (left) and period B (right) data.

7. a pseudorapidity gap on the nuclear side of $\Delta\eta^{Pb} \leq 2.0$, which corresponds to at least one particle with $p_{\mathrm{T}} > p_{\mathrm{T}}{}^{\mathrm{cut}} = 200$ MeV with $\eta > +2.9$ (in the pilot/period A convention)

In the MBTS_2-triggered period A and period B data, no additional cuts are applied after the GRL. Any further cuts, such as the requirement of a reconstructed vertex, or in-time hits on either side of the MBTS detectors, substantially bias the gap distribution towards non-diffractive events and have a low efficiency for events with large gaps. Since the goal of the diffractive/gap analysis is to learn about the full range of inelastic p+Pb collisions, the event selection is designed to fire on a minimally biasing indication of real event activity (e.g. the MBTS_2 trigger) without any further requirements.

Table 7.1 shows the total number of events selected for the minimum bias / centrality analysis and for the diffractive analysis.

7.3.5 Pseudorapidity gaps in data

Pseudorapidity gaps in the MBTS_2-triggered period A and period B data are reconstructed in with tracks and clusters that pass selection cuts and have $p_{\mathrm{T}} > p_{\mathrm{T}}{}^{cut}$. Figure 7.8 shows the gap distributions for each of the three variables and for both periods.

The distributions have some intriguing features. First, all the gap variables are dominated by events with $\Delta\eta = \Delta\eta^{\mathrm{p}} = \Delta\eta^{\mathrm{Pb}} = 0$, in which the majority of the detector is filled with particles

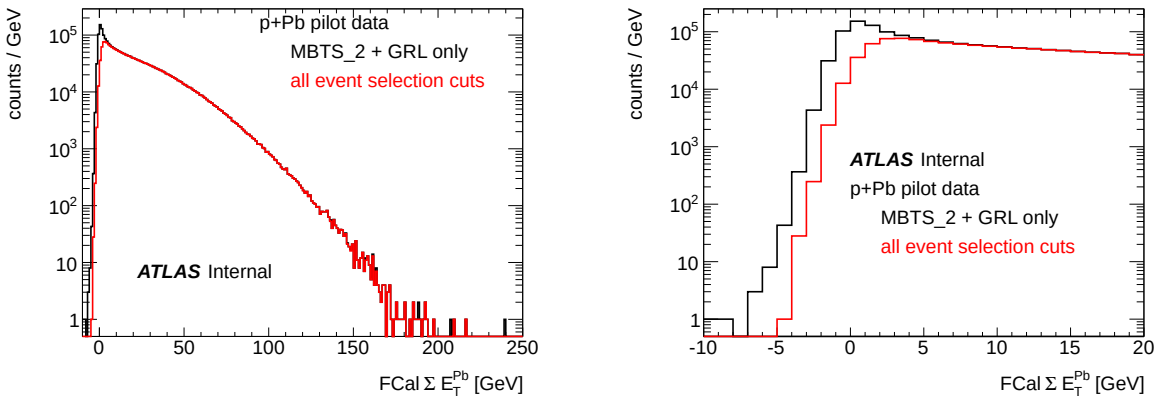


Figure 7.9: FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ for events passing the MBTS_2 trigger and GRL requirements (black) compared to events passing all event selection cuts (red), full distribution (left) and zoom to $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} <$ 20 GeV region (right).

from a high-multiplicity inelastic p+Pb collisions. However, all gap distributions also have a long tail towards $\Delta\eta \approx 8$. These are indicative of diffractive events. Intriguingly, the magnitude of the $dN/d\Delta\eta$ distribution only changes by a factor of ≈ 2 over the range $\Delta\eta = 2$ to 8. It can be shown from MC that events with $\Delta\eta > 8$ have poor efficiency for firing even the `MBTS_2` trigger. For this reason, previous analyses[254] only measure and correct reconstructed gaps out to $\Delta\eta = 8$.

Most interestingly, for gaps of $\Delta\eta > 2$, the distribution of $\Delta\eta^{\mathrm{Pb}}$ dominates over $\Delta\eta^{\mathrm{p}}$ by as much as a factor of ≈ 4. This means that events with a gap on the Pb-going side (e.g. consistent with the diffractive dissociation of the proton) are substantially more common than their counterpart in which the Pb nucleus undergoes the diffractive dissociation. This effect can be seen consistently in period A and period B data (where the orientation of the detector is flipped), and is therefore not an artifact due to differences in the detector response. This result underlines the increased importance of understanding and rejecting diffractive events in the centrality analysis. In fact, the events with large $\Delta\eta^{\mathrm{Pb}} \in$ 2-8 comprise approximately $(1/N_{\mathrm{evt}})(dN/d\Delta\eta^{\mathrm{Pb}}) \times 6 = 6\%$ of the p+Pb cross-section before additional event selection cuts.

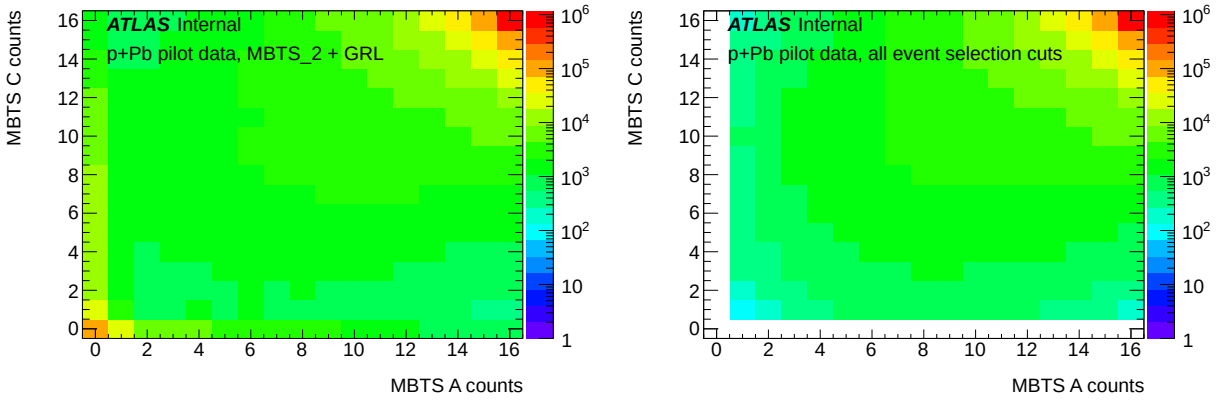


Figure 7.10: MBTS A vs. C count distributions for events passing the MBTS_2 trigger and GRL requirements (left) compared to events passing all event selection cuts (right).

7.3.6 p+Pb $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ analysis

Figure 7.9 shows the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in `L1_MBTS_2`-selected pilot data, before and after the event selection requirements. The reconstructed vertex, timing, two-track and cut against proton diffraction have a significant effect on the low-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ part of the distribution.

Similarly, Figure 7.10 shows the MBTS A and C count distributions for triggered events before and after the event selection cuts. Before event selection cuts, the MBTS distribution has contributions from events with zero or one hits on one side of the MBTS and a wide distribution of counts on the other side. These are consistent with a single-diffractive topology. Additionally, there is a large contribution with ≤ 1 hits on both sides, arising from low-activity events. After event selection cuts, these are significantly suppressed and the dominant remaining feature in the distribution is the peak at 16 hits in each MBTS detector, corresponding to large-multiplicity events typical of inelastic p+Pb collisions.

Figure 7.11 compares the per-event $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in the pilot and period A data. At high $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} > 10$, the ratio is statistically consistent with 1, indicating that the same types of p+Pb events are being selected during the two run periods. At low-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ ($0 < \Sigma E_{\mathrm{T}}^{\mathrm{Pb}} < 5$), the pilot data has an additional contribution that is not in the period A data sample. As mentioned above, this is due to the slightly different trigger selection, since the Level-1 `MBTS_2` trigger is somewhat more permissive for peripheral p+Pb events than the EF `MBTS_1_1` triggered events.

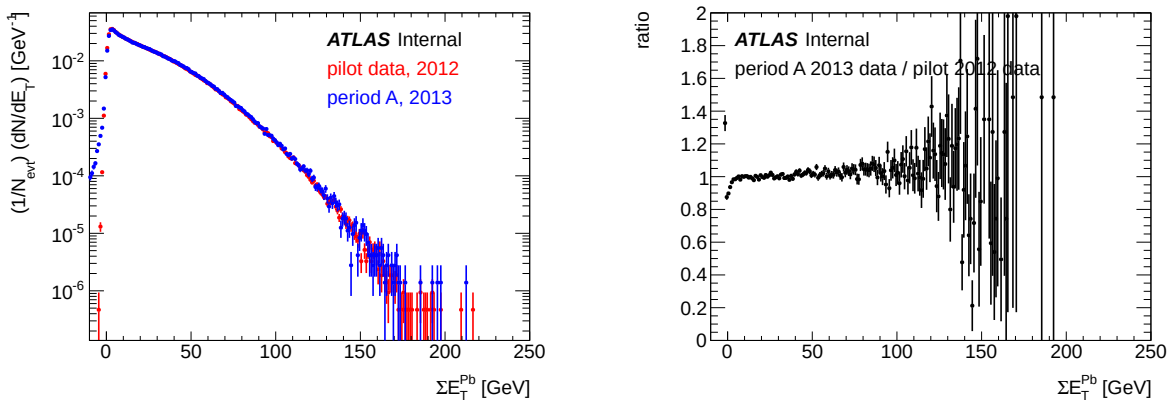


Figure 7.11: FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distributions in the 2012 pilot data and 2013 period A data, comparison (left) and ratio (right).

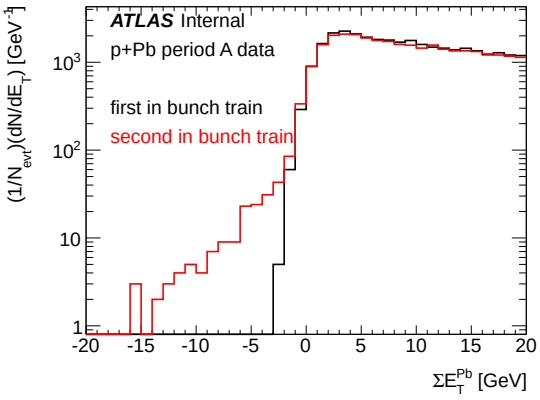


Figure 7.12: FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in the first and second bunches in bunch trains, in a selection of 2013 period A data, demonstrating the presence of out of time pileup in the 2013 data.

An additional feature in the 2013 data which is not present in the 2012 pilot run is the presence of *out of time pileup* from the significantly higher μ in 2013 p+Pb running. Figure 7.12 shows the FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution for the first and second bunches in each bunch train. The bunches at the very beginning of each bunch train have a steeply falling tail at $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} < 0$, from the falling edge of the electronic noise. On the other hand, every subsequent bunch (the second bunch in the bunch train is shown here) has a significant tail to large negative values of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. The BCID dependence indicates that this is caused by the presence of out of time pileup. The tail from out of time pileup

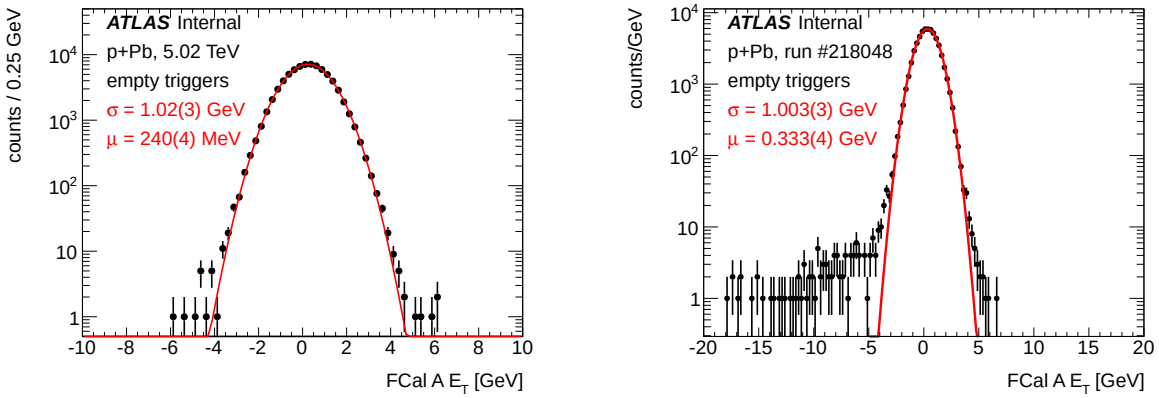


Figure 7.13: FCal A E_{T} distribution in empty events in 5.02 p+Pb data, fit to a Gaussian distribution in pilot data (left) and a period A run (right). (The x-axes have a different range in the two plots.)

is actually somewhat worse in the higher mean-μ period B data. Because of this fact, although the full $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution is used to determine the centrality cuts, the global fits are performed to the subset of the 2013 data which is obtained by selecting on the first BCID in each train. In this way, the tail from out of time pileup (which is not present in the MC and difficult to model) does not unduly influence the global fits in the low-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ part of the minimum bias distribution in data.

In the 2013 data, the raw $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ is corrected in each event for a slowly changing electronic pedestal, the vertex dependence of the acceptance, and (in period B) an overall energy scale difference between the Pb-going FCal (FCal A in period A but FCal C in period B). In addition, the electronic noise term σ_{noise} in data is also determined.

First, the pedestal in FCal A in each period A run and in FCal C in each period B run is determined by examining empty crossings, which are selected with the `EF_rd0_empty_NoAlg` trigger.

The left side of Figure 7.13 shows the FCal A E_{T} distributions in empty events in the pilot data, along with a Gaussian fit to the data. The electronic noise is statistically consistent with $\sigma_{noise} = 1.00$ GeV, and the pedestal is $\mu = 240$ MeV. Both of these are incorporated into the global fit to the data in a way that is described in Section 7.5.3. Using the same procedure, the electronic noise and pedestal is extracted in each run from the period A and B 2013 running. The right side of Figure 7.13 shows an example of this for a run in period A. In this particular run, the electronic

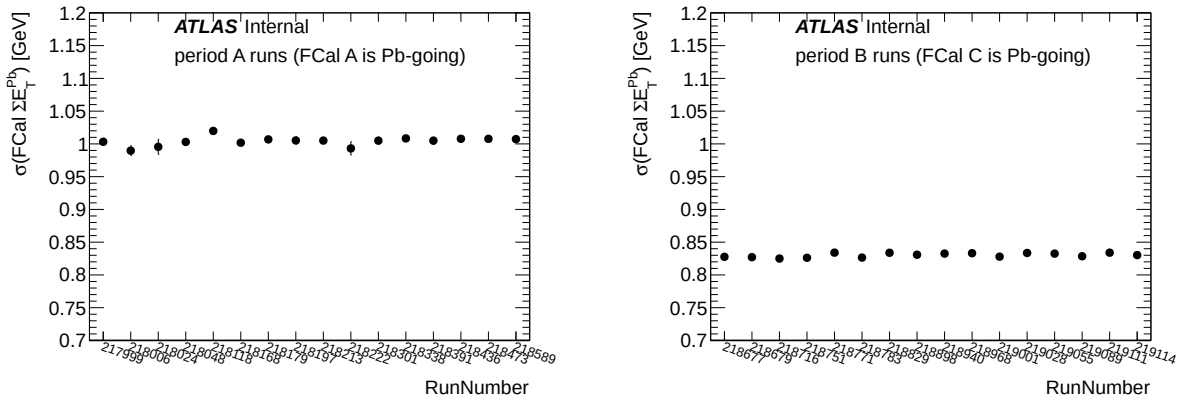


Figure 7.14: Electronic noise-induced width of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in empty events, as a function of run number for period A runs (left) and period B runs (right).

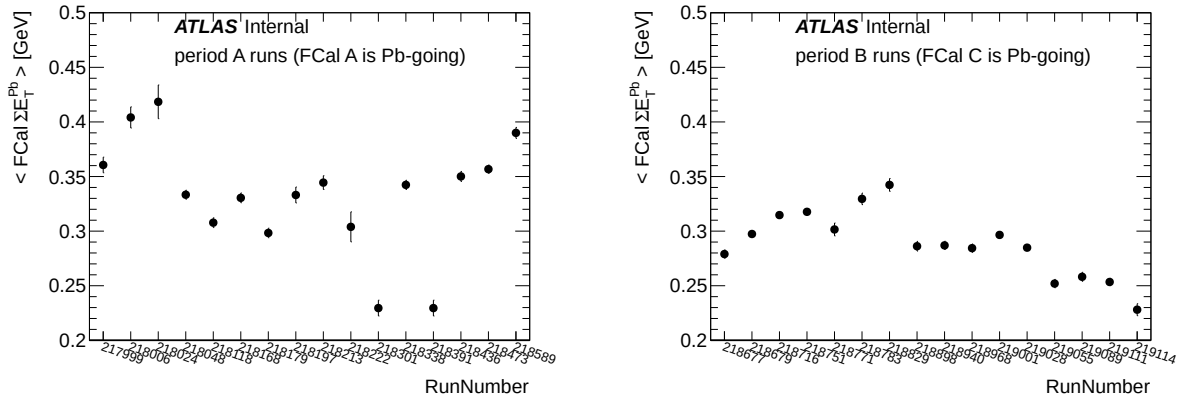


Figure 7.15: Pedestal in the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in empty events, as a function of run number for period A runs (left) and period B runs (right).

noise term is similar to that in the p+Pb pilot run, but the pedestal is different. Additionally, one can see the presence of a long tail to very negative $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ values, caused by out of time pileup (in filled bunches just before the empty crossings selected by this trigger).

Repeating this procedure for every period A and period B run, Figures 7.14 and 7.15 show the run-dependent electronic noise and pedestal. The electronic noise is close to constant in period A (FCal A, $\sigma_{\mathrm{noise}} = 1$ GeV) and period B (FCal C, $\sigma_{\mathrm{noise}} = 0.85$ GeV), but the pedestals show a significant run dependence on the order of 100-200 MeV. These values of the electronic noise

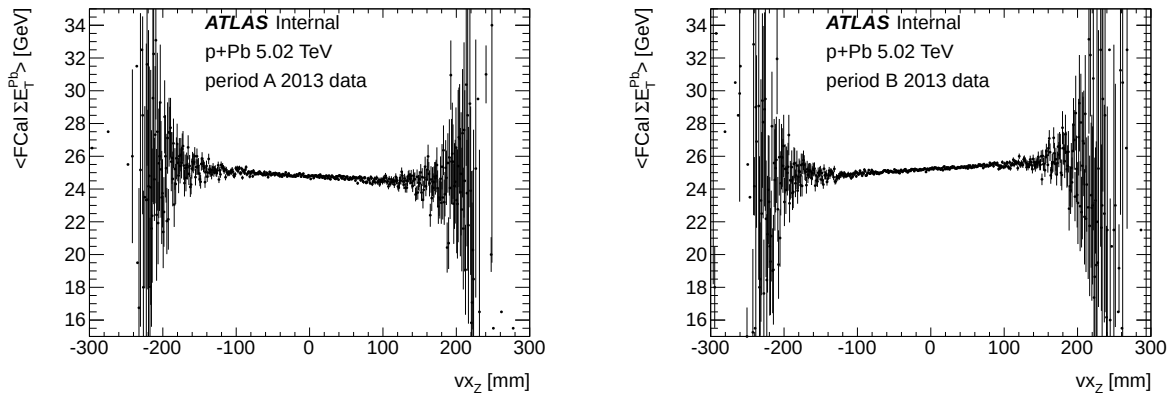


Figure 7.16: Mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ as a function of vertex position in minimum bias events, period A (left) and period B (right).

are used in the global model that describes the data. The first step in correcting the raw $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ measured in each event is to subtract out the contribution from the electronic pedestal in that run.

Second, the data were found to depend on the z-vertex of the event. That is, while the nominal acceptance of the Pb-going FCal (in period A convention) is $+3.1 < \eta < +4.9$ for events at the nominal interaction points, the part of the $dE_{\mathrm{T}}/d\eta$ distribution which makes it into the FCal surely changes with the position of the vertex.

Figure 7.16 shows the mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ in period A and period B data as a function of the vertex position. It can be seen by eye that the mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ is statistically correlated with the vertex position. The mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ decreases as the vertex position moves closer to the Pb-going FCal (FCal A in period A, located in the $+\eta$ direction), and increases when it moves farther away. In period B (where the FCal C is the Pb-going one and is located in the $-\eta$ direction), the relationship is reversed. This relationship is the result of two competing effects: as the vertex moves closer to the detector, the average $dE_{\mathrm{T}}/d\eta$ rises as η moves closer to zero. However, the interval $\Delta\eta$ corresponding to the FCal acceptance becomes smaller. The net effect is apparently a decrease in the total energy as the vertex moves closer to the FCal.

Fits to the vertex dependence reveal that this effect can me modeled as on the overall shift of the energy, with a magnitude of -0.010533% per mm for the FCal A and $+0.011299\%$ per mm for the FCal C. Thus, the pedestal-subtracted data is corrected by a scale factor $1 + 0.00010533 v_z$

in period A and $1 - 0.00011299 v_z$ in period B events (where v_z is measured in mm). After this correction, no statistically significant vertex dependence is observed in the mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

Third, the FCal A and the FCal C were found to have a small but significant energy scale difference. The pedestal-subtracted, acceptance-corrected period A and period B distributions were compared with 0.1%-sized changes in the energy scale. From this, it was determined that a 1.7% energy scale difference existed between the two FCals, in which the FCal C had a systematically higher $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. Thus, an energy scale correction of 0.983(1) was applied to the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ measured in period B to bring the total distribution in line with that in period A.

The full set of corrections can be summarized with,

$$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} = \beta^{\mathrm{scale}} \left(1 + \alpha^{\mathrm{acceptance}} \cdot z_{\mathrm{vertex}}\right) \left((\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^{\mathrm{raw}} - (\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^{\mathrm{pedestal}} \right) \tag{7.3}$$

where

- $(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^{\mathrm{raw}}$ is the raw FCal E_{T}
- $(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^{\mathrm{pedestal}}$ is the run-dependent pedestal in empty events
- $\alpha^{\mathrm{acceptance}}$ is the vertex-dependent correction factor, which is $+1.0533 \times 10^{-4}$ mm^{-1} for period A runs and -1.1299×10^{-4} mm^{-1} for period B runs
- β^{scale} is the FCal A/C energy scale factor, which is 1 for period A runs and 0.983 for period B runs

Figure 7.17 compares the fully corrected $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution for minimum bias events in a selection of period A and period B data. The ratio for the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} > 0$ part of the distribution (where there are no long tails introduced by the out of time pileup) is consistent with 1. The close agreement between the two distributions allows us to treat the corrected period A and period B data in the same manner in the centrality analysis.

Finally, the reconstructed pseudorapidity gaps in data are used to select real p+Pb data events to be used as a model of the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution for additional diffractive excitations of the nucleus not described by the Glauber model. Events that pass the minimum bias selection cuts are selected by requiring a large pseudorapidity gap on the proton side, $\Delta\eta^{\mathrm{p}} \geq 3.0$. Figure 7.18 shows the

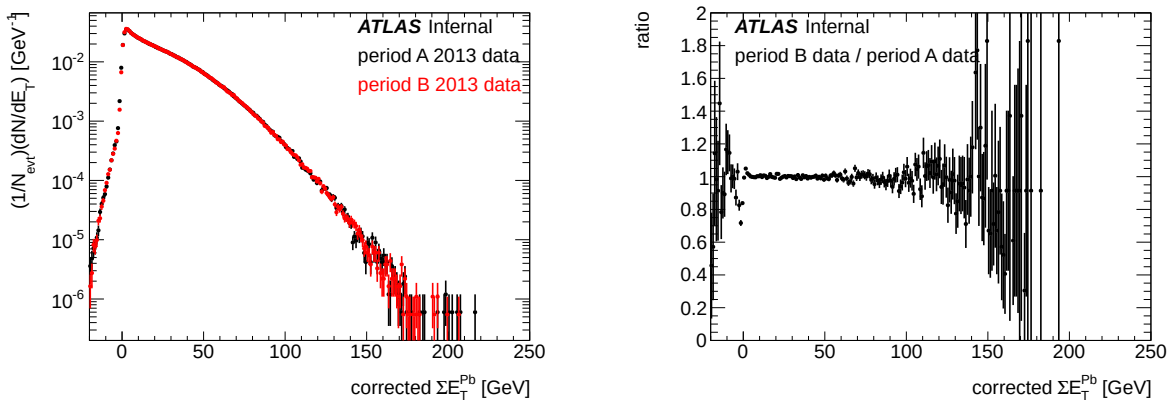


Figure 7.17: Corrected $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distributions in period A and period B data, comparison (left) and ratio (right).

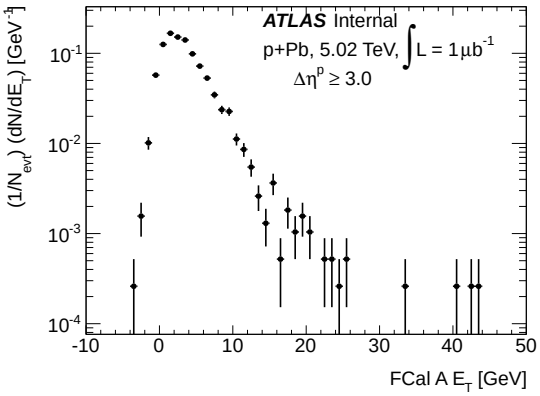


Figure 7.18: FCal A E_{T} distribution in p+Pb data consistent with diffraction of the Pb nucleus, selected with a large pseudorapidity gap on the proton side.

resulting distribution in the p+Pb pilot data. This distribution is later used directly in the global fit as a floating term which allows for the inclusion of residual difractive/photonuclear events.

7.4 Monte Carlo

Monte Carlo simulations of pp events with the same kinematics as the p+Pb data are used to understand the event selection efficiency and understand the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ signal from single nucleon-

Description	N_{evt}	σ_{PYTHIA}
Dataset		
PYTHIA 8, 5.02 TeV inelastic	10^5	67.9 mb
mc12_5TeV.147713.Pythia8_A2MSTW2008LO_minbias_inelastic_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		
PYTHIA 8, 5.02 TeV single diffractive	10^5	12.1 mb
mc12_5TeV.147714.Pythia8_A2MSTW2008LO_minbias_SD_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		
PYTHIA 8, 5.02 TeV double diffractive	10^5	7.73 mb
mc12_5TeV.147715.Pythia8_A2MSTW2008LO_minbias_DD_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		
PYTHIA 6, 5.02 TeV inelastic	10^5	67.8 mb
mc12_5TeV.147710.Pythia_AMBT2BCTEQ6L1_minbias_inelastic_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		
PYTHIA 6, 5.02 TeV single diffractive	10^5	13.4 mb
mc12_5TeV.147711.Pythia_AMBT2BCTEQ6L1_minbias_SD_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		
PYTHIA 6, 5.02 TeV double diffractive	10^5	8.71 mb
mc12_5TeV.147712.Pythia_AMBT2BCTEQ6L1_minbias_DD_pPb5TeV .recon.NTUP_HI.e1722_s1614_s1506_r4251/		

Table 7.2: Monte Carlo sets used in this analysis

nucleon collisions so that it can be extrapolated to p+Pb collisions with many binary collisions.

7.4.1 Datasets

The Monte Carlo sets used throughout this analysis are listed in Table 7.2. They consist of six data sets, broken into PYTHIA 6[259] and PYTHIA 8 p+p events at the 5.02 TeV p+Pb period A kinematics, and further split into equal amounts of inelastic, single- and double-diffractive event samples. (The inelastic sample include appropriately weighed fractions of non-, single- and double-diffractive, and provide the capability to identify the non-diffractive events at the generator level.) The PYTHIA 6 sample was produced using version 6.426 of the generator with the AMBT2[260] tune with the CTEQ6L1[261] pdf set. The PYTHIA 8 sample was produced using version 8.170 of

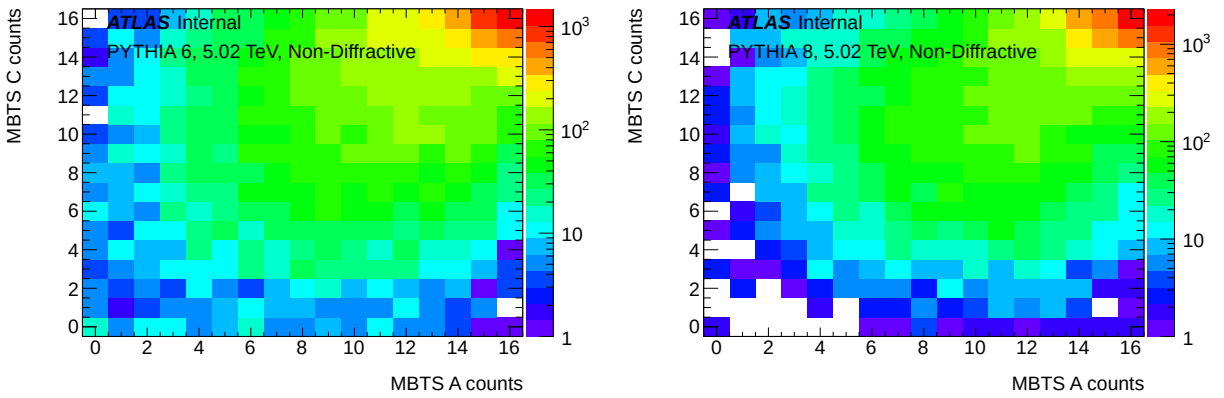


Figure 7.19: MBTS A count vs. C count correlation for PYTHIA 6 (left) and 8 (right), Non-Diffractive events only.

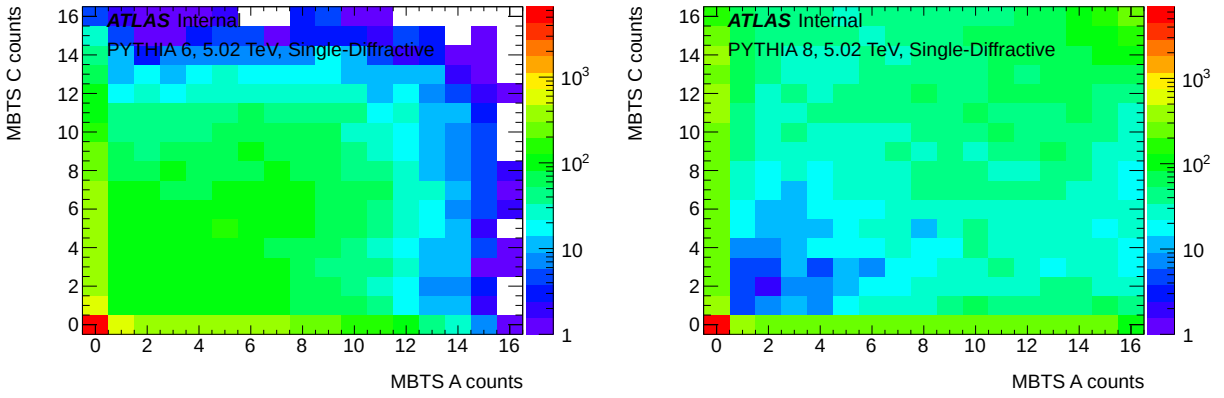


Figure 7.20: MBTS A count vs. C count correlation for PYTHIA 6 (left) and 8 (right), Single-Diffractive events only.

the generator with the A2 tune[262] and the MSTW2008LO pdf set.

At the generator level, the PYTHIA 8 inelastic cross-section is composed of 71% non-diffractive, 18% single diffractive and 11% double diffractive events. The PYTHIA 6 inelastic cross-section is composed of 67% non-diffractive, 20% single diffractive and 13% double diffractive events. PYTHIA 8 has been found to better describe the minimum bias physics observed in ATLAS (see, for example, [254]) and is used as the default MC reference. PYTHIA 6, which has a softer description of minimum bias collisions, especially in forward directions, is used as an extreme alternate description

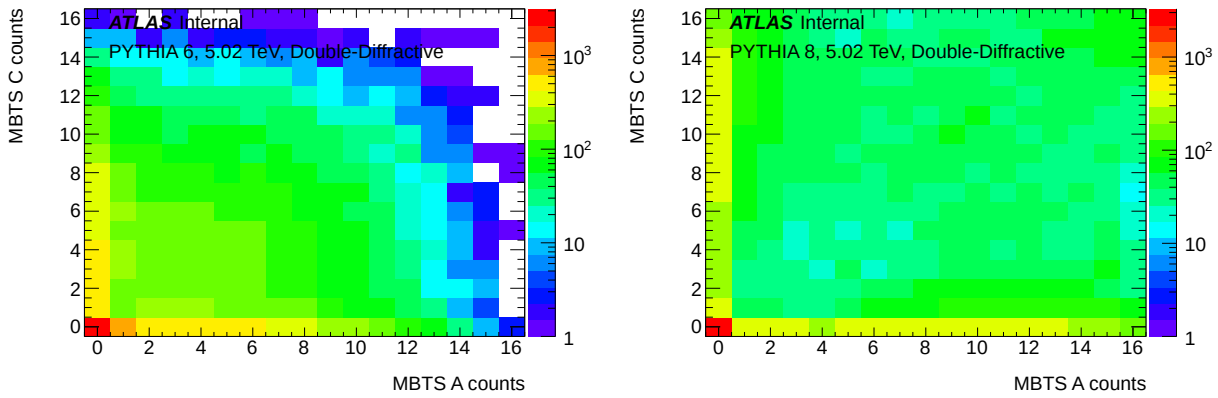


Figure 7.21: MBTS A count vs. C count correlation for PYTHIA 6 (left) and 8 (right), Double-Diffractive events only.

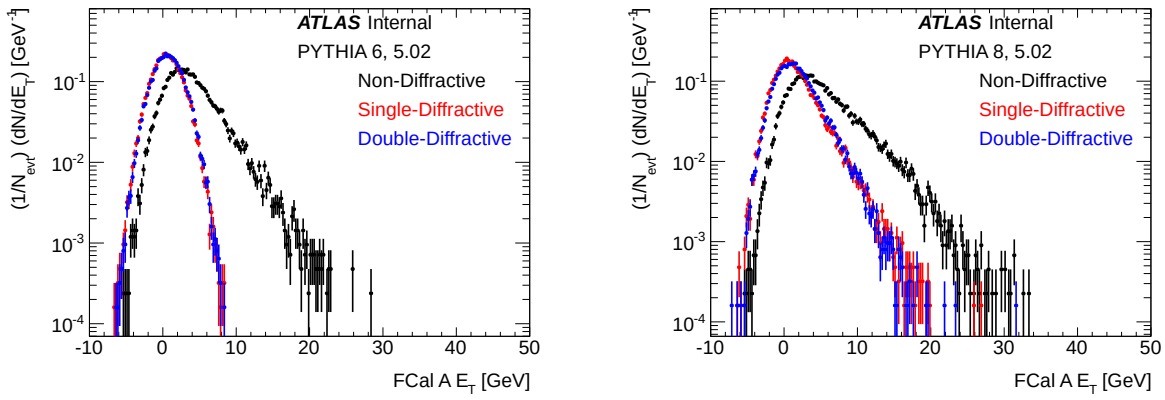


Figure 7.22: FCal A E_{T} distributions, normalized per-event, for PYTHIA 6 (left) and PYTHIA 8 (right), with Non-Diffractive (black), Single-Diffractive (red) and Double-Diffractive (blue) contributions shown separately.

of the minimum bias physics to determine the systematic uncertainty from the choice of model.

7.4.2 Pseudorapidity gaps in MC

The pseudorapidity gap distributions have been examined in MC at the truth and reconstructed levels. For the sake of consistent notation with data, the gaps on the C-side / $\eta < 0$ and A-side / $\eta > 0$ edges of the detector are still referred to as $\Delta\eta^{\mathrm{p}}$ and $\Delta\eta^{\mathrm{Pb}}$, respectively, even though the

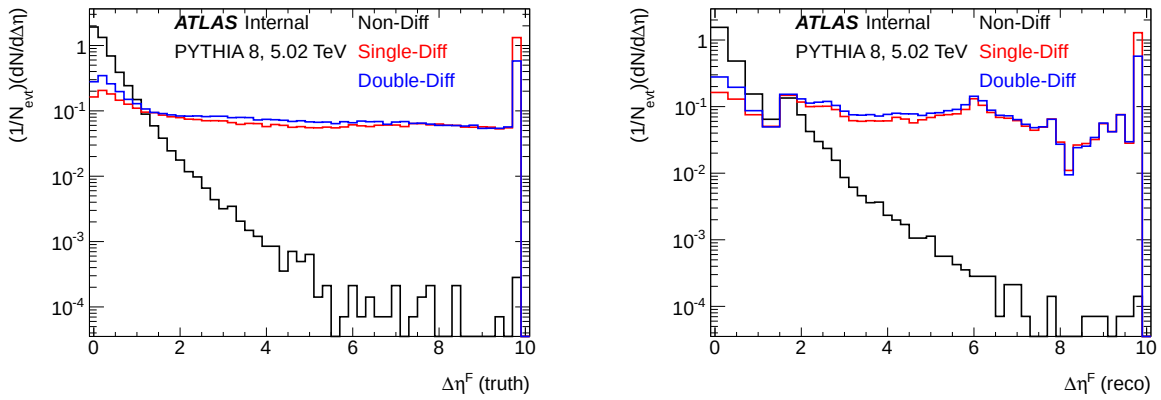


Figure 7.23: Forward pseudorapidity gaps $\Delta\eta^{\mathrm{F}}$ at the generator level (left) and detector level (right) in pp MC, with contributions from non-, single- and double-diffractive events shown in different colors.

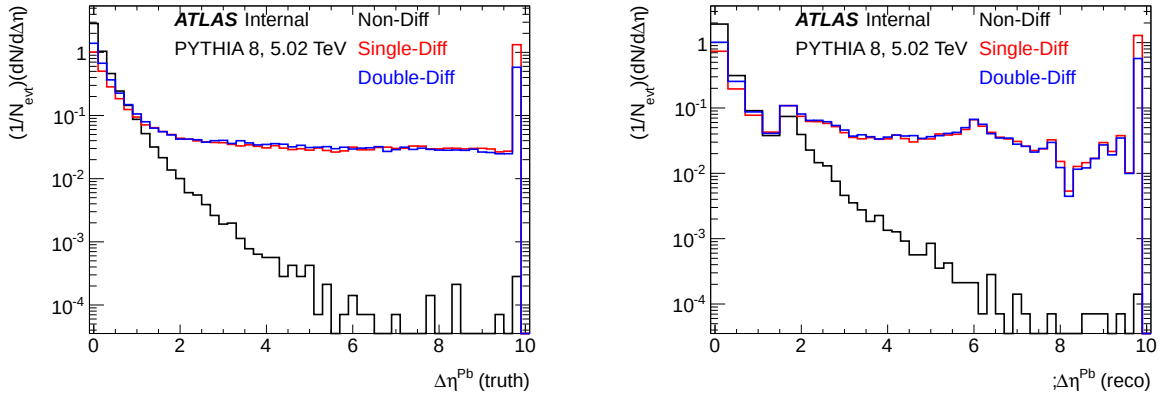


Figure 7.24: Pb-going pseudorapidity gaps $\Delta\eta^{\mathrm{Pb}}$ at the generator level (left) and detector level (right) in pp MC, with contributions from non-, single- and double-diffractive events shown in different colors.

MC consists of pp collisions. The truth gaps are constructed with final-state PYTHIA particles with $p_{\mathrm{T}} > 200$ MeV, while the reconstructed gaps are made with the same track and topoCluster selections as in data.

Figure 7.23 shows the forward gap distributions $dN/d\Delta\eta^{\mathrm{F}}$ at the truth (generator) and reconstructed levels in the PYTHIA 8 MC. In the generator-level distributions, the contribution from

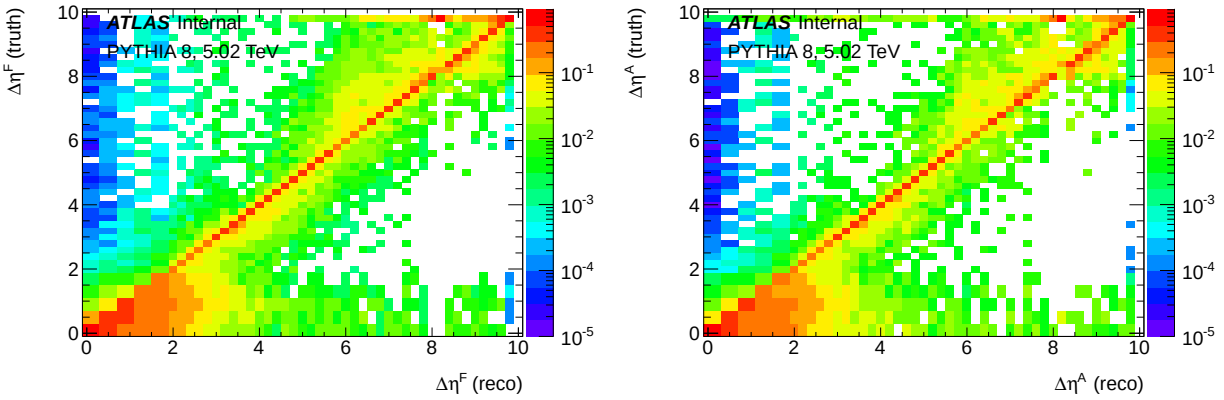


Figure 7.25: Transfer matrices for forward pseudorapidity gaps (left) and Pb-going pseudorapidity gaps (right) at the generator level (y-axis) to detector level (x-axis). The matrices are normalized so that columns sum to 1.

inelastic events has its most likely value at $gapF = 0$ and falls off exponentially, reflecting fluctuations in the hadronization process of particles produced in a real non-inelastic collision. On the other hand, the single- and double-diffractive contributions are nearly flat with $\Delta\eta^{\mathrm{F}}$, overtaking the non-diffractive part of the distribution past $\Delta\eta^{\mathrm{F}} > 2$. The detector-level distributions are not as smooth, reflecting differences in the detector response as a function of pseudorapidity, but the general separation of non-diffractive from diffractive events around $\Delta\eta^{\mathrm{F}} \approx 2$ remains. Figure 7.24 shows the distributions of $\Delta\eta^{\mathrm{Pb}}$ at the truth and reconstructed level, with similar conclusions.

The fact that the non-diffractive and diffractive events distributions intersect near $\Delta\eta approx 2$ is what motivates the cut against diffractive excitation of the proton in the event selection cuts (see Section 7.3.4), where $\Delta\eta^{\mathrm{Pb}} \geq 2$ is required. Furthermore, it is also what motivates the selection of nuclear diffractive events via $\Delta\eta^{\mathrm{p}} \geq 3$ in the data.

Figure 7.25 shows the response matrices $P(\Delta\eta^{\mathrm{F,reco}}|\Delta\eta^{\mathrm{F,truth}})$ and $P(\Delta\eta^{\mathrm{Pb,reco}}|\Delta\eta^{\mathrm{Pb,truth}})$. It can be seen that the increase in the reconstructed gap distributions at $\Delta\eta^{\mathrm{F,reco}} \approx 1.8$ comes from small truth gaps reconstructed at systematically higher values of $\Delta\eta^{\mathrm{F}}$, presumably from the imperfect efficiency for single particles in the FCal. Fortunately, this bin migration is mostly restricted to small gaps with $\Delta\eta < 2$, and thus do not significantly effect the efficiencies of the gap cuts on the reconstructed level.

7.4.3 Event selection efficiency

In this section we derive the efficiency for the p+Pb event selection cuts in MC.

Some of the MC events generated cannot normally be detected within ATLAS. For example, single diffractive events with an extremely small system mass, or double diffractive events with a pseudorapidity gap $\Delta\eta > 9.8$ larger than the entire ATLAS acceptance. Since they leave no signal of any kind and have no efficiency for passing event selection cuts, it is meaningless to include these in the model description of the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution (to do so would require the centrality analysis to map zero signal to a range of N_{part} values). Thus, in order to have an MC reference for inelastic p+p events which is meaningful in a centrality analysis, we require at least one PYTHIA truth particle with $p_{\mathrm{T}} > 200$ MeV within the nominal Pb-going FCal acceptance ($+3.1 < \eta < +4.9$). This has the added benefit of mirroring the last event selection cut in Section 7.3.4. In this way, the event sample in data and the MC reference have equivalent requirements at the truth and reconstructed level.

Figures 7.19, 7.20, and 7.21 show the MBTS count distributions for PYTHIA 6 and PYTHIA 8 Non-, Single- and Double Diffractive events, respectively. These are important for evaluating the efficiency of the different events for passing the event selection cuts in data. Figure 7.22 shows the FCal A E_{T} distributions for the different event classes in PYTHIA 6 and 8, before any event selection cuts are applied.

The efficiency ϵ for passing the basic event selection cuts (which require a reconstructed vertex and good MBTS timing) as well as the event selection cuts plus the additional cut against pseudorapidity gaps on the nuclear-going side ($\Delta\eta^{Pb} \leq 2.0$) are determined and shown in Table 7.3. The efficiency is shown for all inelastic MC events, as well as for individual components of the inelastic pp cross-section (non-, single- and double-diffractive).

The results have several interesting features. As expected, it can be seen that non-diffractive events, where the multiplicity is higher, are much more efficient than single- and double-diffractive events. Thus, most of the inefficiency for inelastic events comes from the diffractive part of the inelastic cross-section. Additionally, all PYTHIA 8 MC subprocesses are systematically more efficient than their PYTHIA 6 counterparts. It is not too surprising that single and double diffractive events have similar efficiency, since many double diffractive events have a gap so large that one of the diffractive systems is outside the ATLAS acceptance, causing the event to have an observed

MC events	ϵ for event selection	ϵ for event selection + $\Delta\eta^{Pb}$ cut
PYTHIA 6 inelastic	76.7%	72.4%
PYTHIA 6 non-diffractive	97.5%	95.5%
PYTHIA 6 single diffractive	33.3%	25.1%
PYTHIA 6 double diffractive	34.9%	25.8%
PYTHIA 8 inelastic	81.5%	78.3%
PYTHIA 8 non-diffractive	98.7%	97.4%
PYTHIA 8 single diffractive	38.0%	30.8%
PYTHIA 8 double diffractive	39.3%	31.7%

Table 7.3: Event selection efficiency with and without the pseudorapidity gap cut, in PYTHIA 6 and 8 MC events.

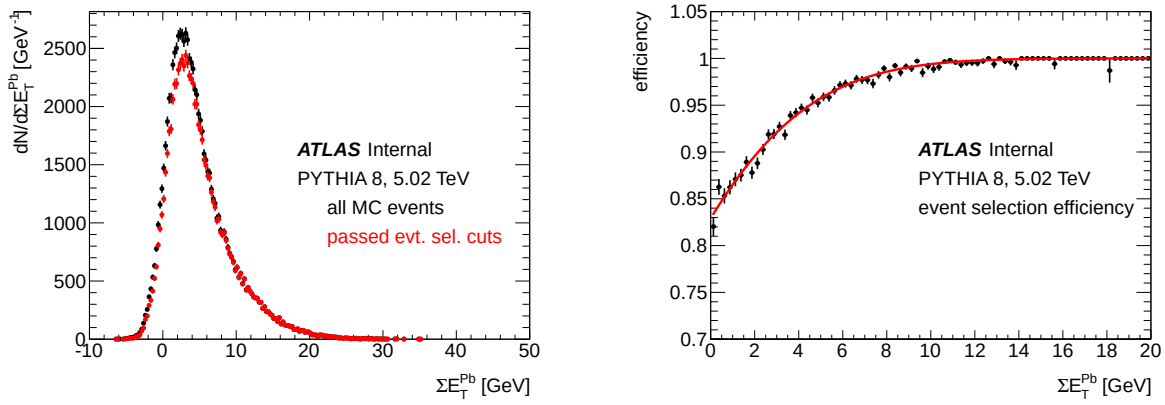


Figure 7.26: FCal A E_{T} distribution in 5.02 TeV PYTHIA 8 (left) inelastic events that pass event selection cuts analogous to those used in the p+Pb data, and efficiency (right).

topology similar to single diffractive events. Furthermore, the gap cut removes an additional 20-30% of the diffractive events remaining after the event selection cuts (presumably from the single particle reconstruction inefficiency in the Pb-going FCal).

The efficiency for the event selection cuts is assumed to depend only on $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, and the MC reference is used to determine this efficiency $\epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$. Figure 7.26 shows the FCal E_{T} distribution for pp MC events which pass the full event selection cuts and a fit to the data as a function of

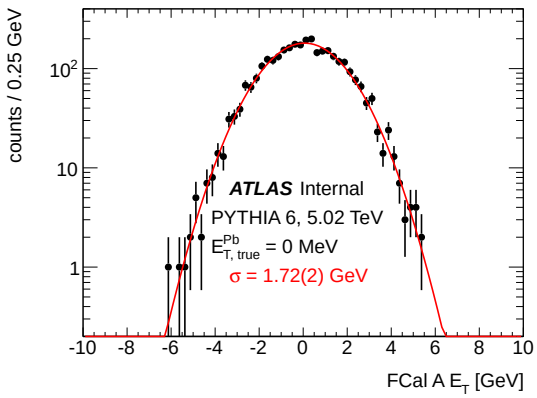


Figure 7.27: FCal A E_{T} distribution in empty events in 5.02 TeV PYTHIA 6 inelastic events, fit to a Gaussian distribution.

$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. The efficiency is modeled as

$$\epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}) = 0.557 + 0.443 \cdot \mathrm{erf}\left(0.113 \cdot \Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\ [\mathrm{GeV}] + 0.614\right) \tag{7.4}$$

where the error function `erf` is chosen phenomenologically and for the feature that it saturates at high $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. The efficiency is $\approx 80\%$ at $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} \approx 0$ and rises monotonically, reaching $> 99\%$ at $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} = 10$ GeV and eventually asymptoting at $= 1$ in the MC.

Furthermore, an important component of the FCal response to inelastic p+p-like collisions is the electronic noise. To quantitatively understand the contribution to the FCal resolution from electronic noise, the FCal distribution in events with no generator-level activity in the FCal A pseudorapidity range is examined, as shown in Figure 7.27. The electronic noise is evaluated as $\sigma_{\mathrm{noise}} = 1.72(2)$ GeV. In MC, no electronic pedestal is found within sensitivity.

7.4.4 $N_{\mathrm{part}} = 2$ Reference

Since a direct p+p 5.02 TeV reference at the shifted center of mass is unavailable, the MC is used to describe the $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution for $N_{\mathrm{part}} = 2$ p+Pb events, which are hypothesized to be similar to pp events with the same $\sqrt{s} = 5.02$ TeV kinematics.

Figure 7.28 shows the distribution $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ for inelastic pp MC events. We model this as a Gamma distribution with k_0 and θ_0. Generally, the signal in the Pb-going FCal for a p+Pb

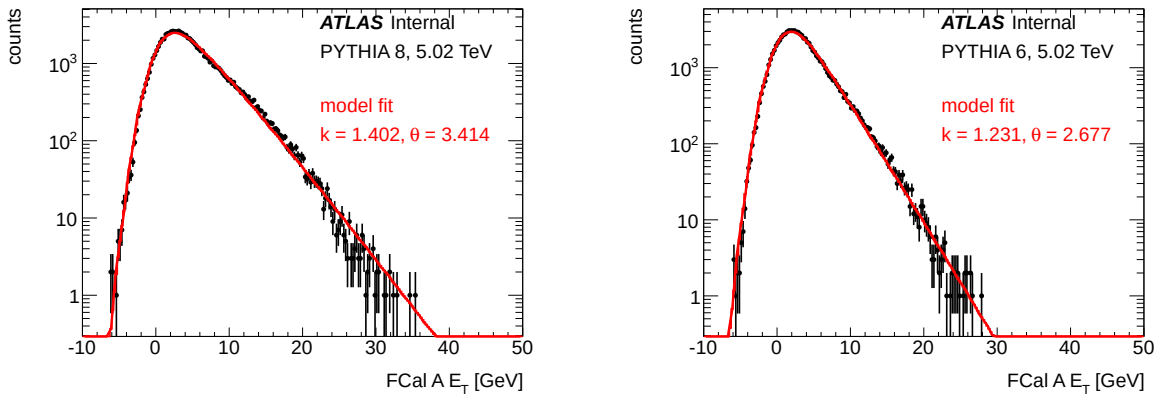


Figure 7.28: Left: FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution for pp inelastic events in MC (black). The best-fit Gamma distribution (with electronic noise contribution) in $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ is shown in red. PYTHIA 8 (6) is shown on the left (right).

collision with N_{part} participants is modeled as a Gamma distribution with parameters that depend on N_{part}, $k(N_{\mathrm{part}})$ and $\theta(N_{\mathrm{part}})$. The MC is used to constrain this behavior for $N_{\mathrm{part}} = 2$, such that $k(N_{\mathrm{part}} = 2) = k$) and $\theta(N_{\mathrm{part}} = 2) = \theta_0$.

These parameters are determined by a fit to a Gamma function, which describes the underlying particle production as a function of the true transverse energy in the FCal $\Sigma E_{\mathrm{T}}^{\mathrm{Pb\,truth}}$, and then convoluted with the electronic noise in MC to give the measured distribution $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. The model $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in the MC is

$$dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\big|_{N_{\mathrm{part}}=2} = A \cdot Gamma(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}}; k_0, \theta_0) * \exp\left(-\Delta\Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^2/2\sigma^2\right) \quad (7.5)$$

$$= A \int_{-\infty}^{+\infty} d\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}} Gamma(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}}; k_0, \theta_0) \exp\left(-(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}} - \Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^2/2\sigma_{\mathrm{noise}}^2\right) \quad (7.6)$$

Where A sets the overall normalization and $\sigma_{\mathrm{noise}} = 1.7$ GeV is the electronic noise in MC. The results of the fit to the PYTHIA 8 and PYTHIA 6 reference are shown in Figure 7.28, and the fit parameters are

- $k_0 = 1.40(1)$, $\theta_0 = 3.41(2)$ for the PYTHIA 8 reference
- $k_0 = 1.23(2)$, $\theta_0 = 2.67(5)$ for the PYTHIA 6 reference

The fits do a good but not perfect job of describing the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in MC, with some residual tension in the shape of the high-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ tail, but with a remarkably good description of the electronic noise in the negative-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ tail.

7.5 Global fits to data

This section describes the global fits that map the N_{part} distribution generated in the Glauber model of inelastic p+Pb events onto the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution observed in data.

7.5.1 Global models

The model to which the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data is fit consists of the sum of the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distributions from Glauber p+Pb events described by a number of participants N_{part}, weighed by the relative probability of having N_{part} participants in the full set of Glauber p+Pb events. Let $P(N_{\mathrm{part}}) = (1/N_{\mathrm{MC,evt}})(dN/dN_{\mathrm{part}})$, where $N_{\mathrm{MC,evt}}$ is the total number of Glauber events. Furthemore, let $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ be the probability distribution of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ values from events with N_{part} participants. Naively, the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution from all inelastic p+Pb events is then given by

$$\left.\frac{dN}{d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}}\right|_{\mathrm{model}} = C\Sigma_{N_{\mathrm{part}}} P(N_{\mathrm{part}}) P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}}) \tag{7.7}$$

where C is a constant chosen to match the total number of counts in the data. In fact, fitting Equation 7.7 to the data is complicated by several factors. First, peripheral events may have some inefficiency for the minimum bias event selection criteria, which will decrease the observed $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ at low $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. Second, if $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ describes the distribution of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ values at the truth level, the measured $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution will surely be smeared by the electronic noise in the FCal σ_{noise}. Third, we must account for any residual events in the minimum bias sample not fully described by the Glauber or Glauber-Gribov picture.

Thus, the full model is

$$\left.\frac{dN}{d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}}\right|_{model} =$$
$$C\left((1-f_{\mathrm{diff}})\Sigma_{N_{\mathrm{part}}}\left\{P(N_{\mathrm{part}})P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}}|N_{\mathrm{part}})\epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}})\right\} * (\mathrm{noise}) + f_{\mathrm{diff}}P_{\mathrm{diff}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})\right) \quad (7.8)$$

where $\epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ is the event selection efficiency as a function of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, f_{diff} is the relative contribution from extra diffractive events not normally modeled by Glauber, $P_{\mathrm{diff}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ is the distribution of FCal energy for these processes, C is an overall normalization factor to the total counts in data, and the electronic noise term is

$$\exp\left(-(\Sigma E_{\mathrm{T}}^{\mathrm{Pb,truth}} - \Sigma E_{\mathrm{T}}^{\mathrm{Pb}})^2/2\sigma_{\mathrm{noise}}^2\right) \quad (7.9)$$

where σ_{noise} was determined above from empty events in data to be ≈ 1 GeV. $\epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ is taken from MC, while $P_{\mathrm{diff}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ and σ_{noise} are taken from data, and f_{diff} is treated as a free parameter in the fit.

The signal from $N_{\mathrm{part}} = 2$ ("*pp*-like") p+Pb events, $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}} = 2)$ is determined using the 5.02 TeV inelastic *pp* MC. However, an additional model is needed to describe how $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ evolves with increasing N_{part}. In this analysis, $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ is modeled as a *Gamma Distribution*. A Gamma distribution is defined by parameters k (the "shape parameter") and θ (the "scale parameter"), and for non-zero numbers $x > 0$ has the probability density function

$$Gamma(x;k,\theta) = \frac{1}{\Gamma(k)\theta^k}x^{k-1}e^{-x/\theta} \quad (7.10)$$

Above, $\Gamma(k) = \int_0^{+\infty} t^{z-1}e^{-t}dt (= (k-1)!$ for integer values of $k)$ is the Gamma function. Note that the Gamma distribution is normalized so that $\int_0^{+\infty} GammaDist(x;\ k,\ \theta)\ dx = 1$. A useful fact is that the mean and variance of the distribution are $k\theta$ and $k\theta^2$, respectively. Since the Gamma distribution is completely described by its k and θ parameters, the different fit models therefore describe the N_{part} dependence of $k(N_{\mathrm{part}})$ and $\theta(N_{\mathrm{part}})$.

7.5.2 Models of N_{part} dependence

Two main models of the N_{part} dependence in the Gamma distribution parameters were used, plus an alternate third model as a check on the results.

- Fit model #1. *Linear-k with θ Acceptance Term*

  - $k(N_{\text{part}}) = k_0 + k_1(N_{\text{part}} - 2)$
  - $\theta(N_{\text{part}}) = \theta_0 + \theta_1 \log(N_{\text{part}} - 1)$

  This model includes a term in k that is linear in N_{part}. This is the "wounded nucleon" term, resulting in the behavior that the distribution corresponding to N_{part} is simply the $N_{\text{part}} - 1$ distribution convoluted with a Gamma distribution with the same parameter θ but $k = k_1$.

  Additionally, this model includes an acceptance term in θ which varies with the logarithm of $N_{\text{coll}}(= N_{\text{part}} - 1)$. This term is motivated by the observation that additional proton-nucleon collisions after the first one may systematically shift the center of mass of the system in the direction of the nucleus[253]. Since the dN/dy distribution is downward sloping in the region of the FCal, boosts in the nuclear direction actually increase the mean energy falling within its acceptance. The logarithmic term has the proper behavior at $N_{\text{coll}} = 1$ (there is no correction since $\log 1 = 0$). The logarithm is motivated by MC studies which show that the energy of truth particles within $3.1 < \eta^{\text{lab}} < 4.9$ increase, with good approximation, as a logarithm of the number of successive boosts.

- Fit model #2. *Quadratic-k with Constant θ*

  - $k(N_{\text{part}}) = k_0 + k_1(N_{\text{part}} - 2) + k_2(N_{\text{part}} - 2)^2$
  - $\theta(N_{\text{part}}) = \theta_0$

  This model is closer to "traditional" Glauber approaches, in which the N_{part} dependence is contained entirely in the k parameter, with a quadratic term k_2 which allows for non-linear corrections at higher N_{part}.

In both cases, k_0 and θ_0 are held fixed so that $k(N_{\text{part}} = 2) = k_0$ and $\theta(N_{\text{part}} = 2) = \theta_0$ reduce to the distribution $P(\Sigma E_{\text{T}}^{\text{Pb}})$ in MC pp events.

A key difference between these two models is the size of fluctuations in $P(\Sigma E_{\text{T}}^{\text{Pb}}|N_{\text{part}})$ at high-N_{part}. The mean of a Gamma Distribution is $\mu = k\theta$, and the variance is $\sigma^2 = k\theta^2$. Thus, the size of the relative standard deviation / mean is $\sigma/\mu = 1/\sqrt{k}$. In fit model #2, $k(N_{\text{part}})$ must

account for all of the shift in the mean of $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ (since θ is constant with N_{part}) and grows substantially with N_{part}. Thus, the relative fluctuations decrease at high N_{part}. However, in fit model #1, the N_{part} dependence in θ can account for some of the growth in the mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ and k can remain much smaller as a function of N_{part}. Thus, in fit model #1 the relative fluctuations can remain relatively large, even at high N_{part}.

Alternate models were attemped, but ultimately these two models were found to give a good description of the extremes of the possible parameter space. However, as an additional systematic check on the results, we consider one additional fit model:

- Fit model #3. *Wounded Nucleon with Non-Linear Terms*
  - $k(N_{\mathrm{part}}) = \frac{k_0}{2} N_{\mathrm{part}} + k_2(N_{\mathrm{part}} - 2)^2$
  - $\theta(N_{\mathrm{part}}) = \theta_0 + \theta_1 \log(N_{\mathrm{part}} - 2)$

  This model explicitly enforces as a wounded nucleon scaling by setting $k_1 = k_0/2$, but then allows for the possibility of non-linear terms both in k (the k_2 term seen in fit #1) and θ (the θ_1 term seen in fit #2). This fit model was used see if the data could be reproduced by enforcing the wounded nucleon behavior with either the non-linear behavior in k or the acceptance term in θ (or a combination of both).

7.5.3 Application to data

This section describes the fit of the model described by Equation 7.8 to the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data. The fit is performed for each of the three versions of the Glauber model (each of which provide a different $P(N_{\mathrm{part}})$) and for each fit model (which have different descriptions of $k(N_{\mathrm{part}})$ and $\theta(N_{\mathrm{part}})$ in the Gamma distribution $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$). Additionally, the fits are performed to 2012 and 2013 data separately as a cross-check.

In the 2012 data, where the μ was low, all good minimum bias events are used in the global fit, providing a total of 2.1×10^6 events. In the 2013 data, since the model cannot reproduce the effects of out of time pileup, the global fits are performed only to data from the first bunch in each bunch train, corresponding to $\approx 1 \times 10^6$ minimum bias triggered events. The 2012 results are shown and discussed in detail here. The 2013 results are extremely consistent with the 2012

model	k_1	k_2	θ_1	f_{diff}	total eff.
fits to 2012 data					
standard Glauber, fit #1	0.425(2)	-	+1.32(1)	$0 \pm 0.01\%$	98.1%
standard Glauber, fit #2	0.70(12)	+0.025(6)	-	$0 \pm 1.7\%$	98.2%
standard Glauber, fit #3	0.701 (fixed)	+0.020(26)	+0.08(46)	$0.4 \pm 1.0\%$	98.2%
Glauber-Gribov $\Omega = 1.01$, fit #1	1.139(3)	-	-0.209(2)	$1.83 \pm 0.04\%$	98.3%
Glauber-Gribov $\Omega = 1.01$, fit #2	0.994(1)	-0.00457(7)	-	$1.08 \pm 0.04\%$	98.2%
Glauber-Gribov $\Omega = 1.01$, fit #3	0.701 (fixed)	-0.0101(1)	+0.689(3)	$0 \pm 0.01\%$	98.1%
Glauber-Gribov $\Omega = 0.55$, fit #1	0.901(3)	-	+0.074(4)	$1.31 \pm 0.05\%$	98.2%
Glauber-Gribov $\Omega = 0.55$, fit #2	0.929(1)	+0.00227(9)	-	$1.40 \pm 0.04\%$	98.2%
Glauber-Gribov $\Omega = 0.55$, fit #3	0.701 (fixed)	-0.0057(1)	+0.564(4)	$0.40 \pm 0.04\%$	98.2%
fits to 2013 data					
standard Glauber, fit #1	0.435(5)	-	+1.30(2)	$0 \pm 0.01\%$	98.0%
standard Glauber, fit #2	0.712(3)	+0.0250(2)	-	$0 \pm 0.01\%$	98.0%
standard Glauber, fit #3	0.704 (fixed)	+0.0196(6)	+0.094(10)	$0 \pm 0.01\%$	98.0%
Glauber-Gribov $\Omega = 1.01$, fit #1	1.146(6)	-	-0.200(4)	$1.22 \pm 0.09\%$	98.1%
Glauber-Gribov $\Omega = 1.01$, fit #2	1.003(3)	-0.0043(1)	-	$0.43 \pm 0.09\%$	98.1%
Glauber-Gribov $\Omega = 1.01$, fit #3	0.704 (fixed)	-0.0103(1)	+0.713(7)	$0 \pm 0.01\%$	98.0%
Glauber-Gribov $\Omega = 0.55$, fit #1	0.90(16)	-	+0.092(20)	$0.6 \pm 1.7\%$	98.1%
Glauber-Gribov $\Omega = 0.55$, fit #2	0.937(3)	+0.0027(2)	-	$0.76 \pm 0.09\%$	98.1%
Glauber-Gribov $\Omega = 0.55$, fit #3	0.704 (fixed)	-0.0056(2)	+0.575(7)	$0 \pm 0.01\%$	98.0%

Table 7.4: Summary of fit results.

results, with typical changes in the geometric observables on the order of $< 1\%$. Thus, the 2013 fits and results are relegated to Appendix A, and are mentioned only when they are different enough from the 2012 results to mention.

Table 7.4 shows the fit parameters for global fits using each of the three fit models, each of the three Glauber models, and to 2012 and 2013 data. Generally, the 2012 and 2013 data give very similar results in terms of the fit parameters, with slight differences reflecting the different treatment of the uncorrected $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. However, a consistent feature is that fits to the 2012 data find a higher residual diffractive contribution (an additional $\approx 1\%$ of the total observed event sample) than do the corresponding fits in the 2013 data. This is consistent with what is described in Section 7.3.6 when comparing the low-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ region of data from the two run periods. Table 7.4

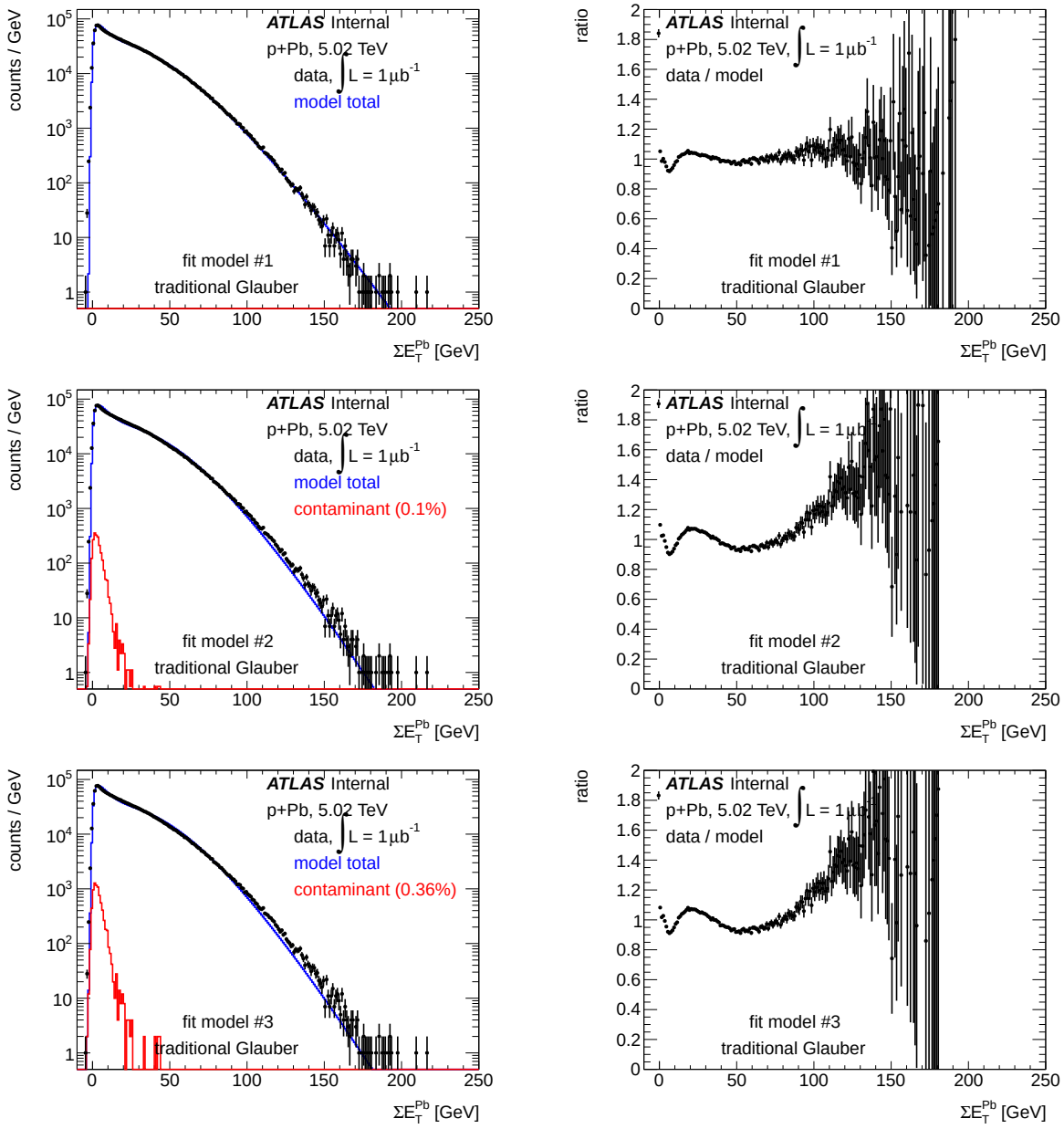


Figure 7.29: Fits to 2012 p+Pb data using the default Glauber model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

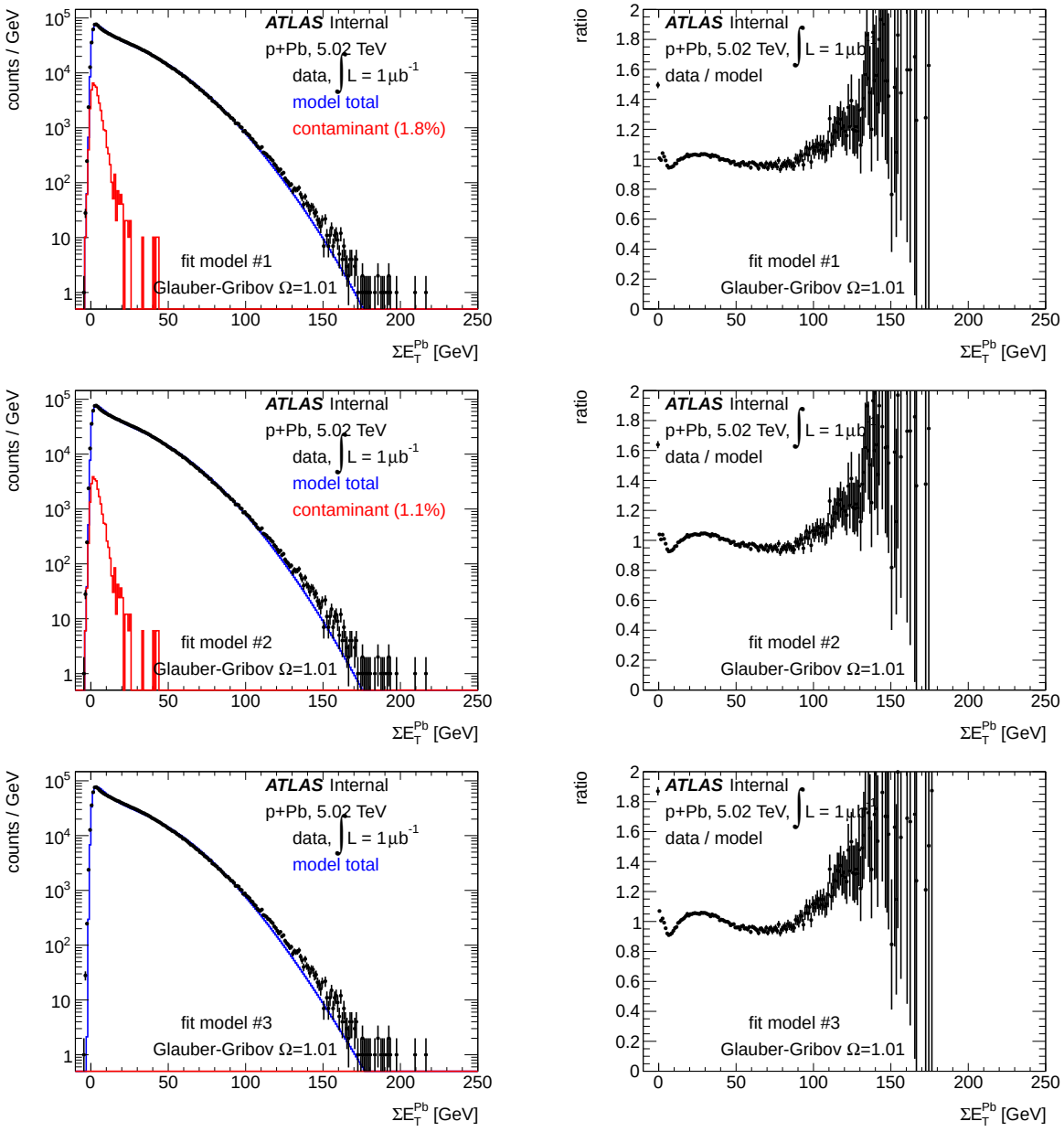


Figure 7.30: Fits to 2012 p+Pb data using the Glauber-Gribov $\Omega = 1.01$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

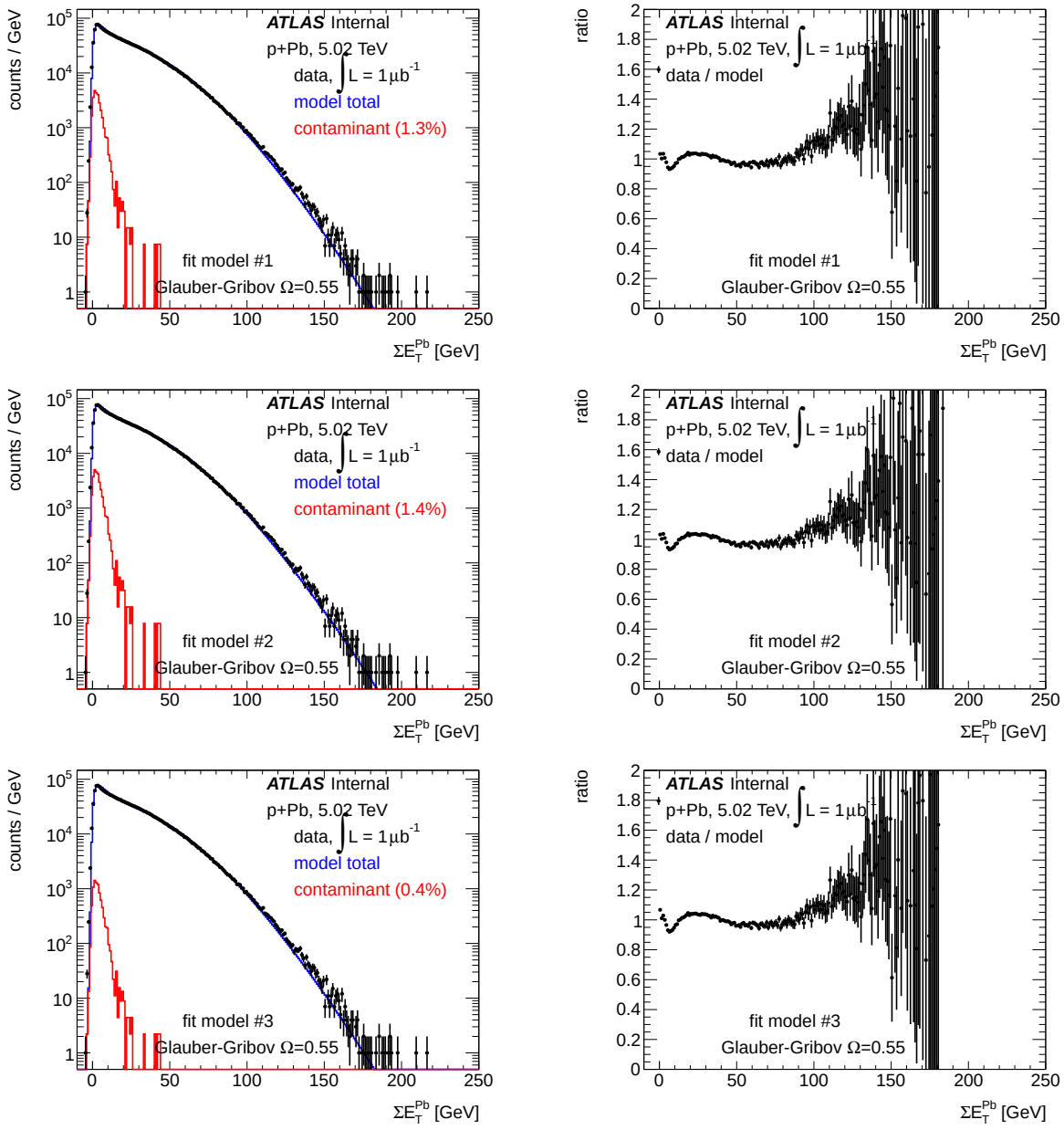


Figure 7.31: Fits to 2012 p+Pb data using the Glauber-Gribov $\Omega = 0.55$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

also shows the event selection efficiency for the full inelastic Glauber cross-section, which is defined below in Equation 7.11.

Figures 7.29, 7.30 and 7.31 show the best fit models to 2012 data. Figures A.1, A.2 and A.3 in the Appendix show the fits to 2013 data. Fits to the 2012 data using PYTHIA 6 to model the efficiency and set the $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ for $N_{\mathrm{part}} = 2$ are shown in Figures A.4, A.5 and A.6 in the Appendix. The fit models with different PYTHIA references and Glauber models vary in how well they describe the data. Generally, the best fits are able to describe the data to within $< 5\%$ over three orders of magnitude, typically with slowly divergent behavior in the very high-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ tail in data. No fit is able to describe to describe the data perfectly, and fit functions #1 and #2 with the PYTHIA 8 reference are seen being equally plausible descriptions of the data.

Since the efficiency is a function of model $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, the total efficiency for the p+Pb model $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution can be estimated only after performing the global fit. The total efficiency $\epsilon^{\mathrm{total}}$ for the model distribution $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\big|_{\mathrm{model}}$ is given by

$$\epsilon^{\mathrm{total}} = \frac{\int d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} \epsilon(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})\; dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\big|_{\mathrm{model}}}{\int d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\; dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\big|_{\mathrm{model}}} \tag{7.11}$$

The fits find a 98% efficiency for the inelastic p+Pb Glauber cross-section and a 0-2% residual contaminant, depending on the details of the fit. (As a reminder, since the contaminant contribution is modeled using p+Pb data events with a large proton-side pseudorapidity gap, it is consistent with residual diffractive events passing event selection cuts or photonuclear dissociation of the Pb nucleus. Although this component is often referred to as "residual diffraction", it is also consistent with the presence of residual photonuclear events.) On the other hand, the PYTHIA6 reference shows a systematically lower total efficiency (97%) and also less diffractive contribution. Thus, conservatively considering the full variation of fit models, the total efficiency plus residual inefficiency is taken to be $98 \pm 2\%$ (comprised of 97-98% efficiency for the total Glauber cross-section and a 0-2% residual contaminant). The $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ cuts in data will be drawn according to this 98% figure, but 96% and 100% total efficiency+contaminant will be used to determine the systematic uncertainties on the results.

As can be seen in the first two rows of Table 7.4, in the standard Glauber model, the N_{part} dependence is very sensitive to the choice of fit model. Both models have a significant non-linear term (θ_1 in the case of fit model #1, k_2 in fit model #2) in the best fit to data. In fit model #1,

a significant portion of the $N_{\rm part}$ dependence of $P(\Sigma E_{\rm T}^{\rm Pb}|N_{\rm part})$ is contained in the θ_1 parameter. On the other hand, fit model #2 gives a k_1 close to the wounded nucleon expectation of $k_0/2$ along with a statistically significant k_2. In the default Glauber model (Figure 7.29), fit model #1 gives a substantially better fit to the data. On the other hand, in the two Glauber-Gribov models (Figures 7.30 and 7.31), the fit models do an approximately equal job. In fact, the fit results for both Glauber-Gribov models have smaller non-linear terms and are thus less sensitive to the details of which fit model is used. In particular, the $\Omega = 0.55$ model finds very similar values for k_1 in fit models #1 and #2 and very small but statistically significant non-linear terms. Additionally, both fit models agree on a 1.5% residual diffractive contaminant. This implies that the Glauber-Gribov variants (especially the $\Omega = 0.55$ model) can describe the $\Sigma E_{\rm T}^{\rm Pb}$ distribution in data with a much simpler assumption about how the $\Sigma E_{\rm T}^{\rm Pb}$ distribution grows in events with increasing $N_{\rm part}$. Furthermore, in the default Glauber model the best fits do not find a statistically significant residual contamination from diffraction f_{diff}, while the Glauber-Gribov models see a possible role for a residual diffractive component.

Another interesting way to characterize the fits is to examine the the mean and standard deviation of the $\Sigma E_{\rm T}^{\rm Pb}$ as a function of $N_{\rm part}$,

$$\left\langle \Sigma E_{\rm T}^{\rm Pb} \right\rangle_{N_{\rm part}} = \int \Sigma E_{\rm T}^{\rm Pb} P(\Sigma E_{\rm T}^{\rm Pb}|N_{\rm part}) d\Sigma E_{\rm T}^{\rm Pb} \tag{7.12}$$

$$\sigma\left(\Sigma E_{\rm T}^{\rm Pb}\right)_{N_{\rm part}} = \sqrt{\int \Sigma E_{\rm T}^{\rm Pb\,2} P(\Sigma E_{\rm T}^{\rm Pb}|N_{\rm part}) d\Sigma E_{\rm T}^{\rm Pb} - \left(\int \Sigma E_{\rm T}^{\rm Pb} P(\Sigma E_{\rm T}^{\rm Pb}|N_{\rm part}) d\Sigma E_{\rm T}^{\rm Pb}\right)^2} \tag{7.13}$$

In a wounded nucleon model ($k_1 = k_0/2$, no other non-linear terms in k or θ), $\left\langle \Sigma E_{\rm T}^{\rm Pb} \right\rangle \propto N_{\rm part}$. Figure 7.32 plots the mean energy per participant $\left\langle \Sigma E_{\rm T}^{\rm Pb} \right\rangle N_{\rm part}$ as a function of $N_{\rm part}$ in the two fit models and three Glauber models. In the default Glauber model, it grows substantially faster than linear with $N_{\rm part}$ $\left(\text{i.e. } d\left\langle \Sigma E_{\rm T}^{\rm Pb} \right\rangle / dN_{\rm part} > 0\right)$. Furthermore, it increases much more quickly in fit model #2 than fit model #1. In the Glauber-Gribov model with $\Omega = 1.01$, the mean $\Sigma E_{\rm T}^{\rm Pb}$ per $N_{\rm part}$ has a different behavior, with the mean energy growing faster than linear at low-$N_{\rm part}$ and then turning over and growing more slowly than linear at high-$N_{\rm part}$. This implies that in this Glauber-Gribov variant, the best models which describe the data are those in which additional collisions of the proton contribute increasingly less and less energy. On the other hand, the mean

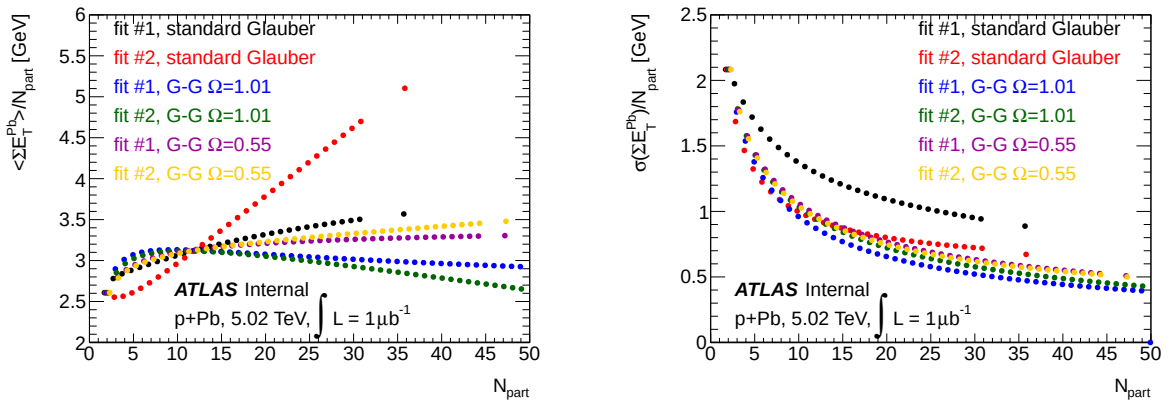


Figure 7.32: Mean $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ per N_{part} as a function of N_{part} (left) and standard deviation of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ per N_{part} as a function of N_{part} (right), for all three Glauber models and two fit models.

$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ per N_{part} in the Glauber-Gribov $\Omega = 0.55$ fits rise slowly at low-N_{part} and then are close to constant at high-N_{part}, and with only a small difference between the fit models.

Figure 7.32 also plots the standard deviation per participant $\sigma(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})_{N_{\mathrm{part}}}/N_{\mathrm{part}}$. In the default Glauber model, the two fit models give substantially different values for the widths of the $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ distributions as a function of N_{part}. In fit model #1, the fluctuations are substantially larger at all N_{part} than in fit model #2, as predicted at the end of Section 7.5.2. In the Glauber-Gribov $\Omega = 1.01$ model, the relative width of the $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ distribution shrinks strongly with increasing N_{part}. The Glauber-Gribov $\Omega = 0.55$ model is somewhere in between the two other models, and is also the model for which the two fit models give the closest results in terms of $\sigma(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})_{N_{\mathrm{part}}}/N_{\mathrm{part}}$.

Finally, the correlation between $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ and N_{part} can be further investigated by projecting slices of $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ for various N_{part} (e.g. the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution from events with a given N_{part}) and slices of dN/dN_{part} for various $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ (e.g. the relative contribution to this $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ from events with a different N_{part}). This is shown in Figure 7.33 for fit model #1 and Figure 7.34 for fit model #2. It can be seen that for a given N_{part}, the default Glauber model gives very fit-dependent results on the shape of $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$. Fit model #1 looks more like a Gamma distribution (since $k(N_{\mathrm{part}})$ remains small even at high-N_{part}) while fit model #2 looks more like a Gaussian distribution (which is the large-k limit of a Gamma distribution). Similarly, fit model #2 gives $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$

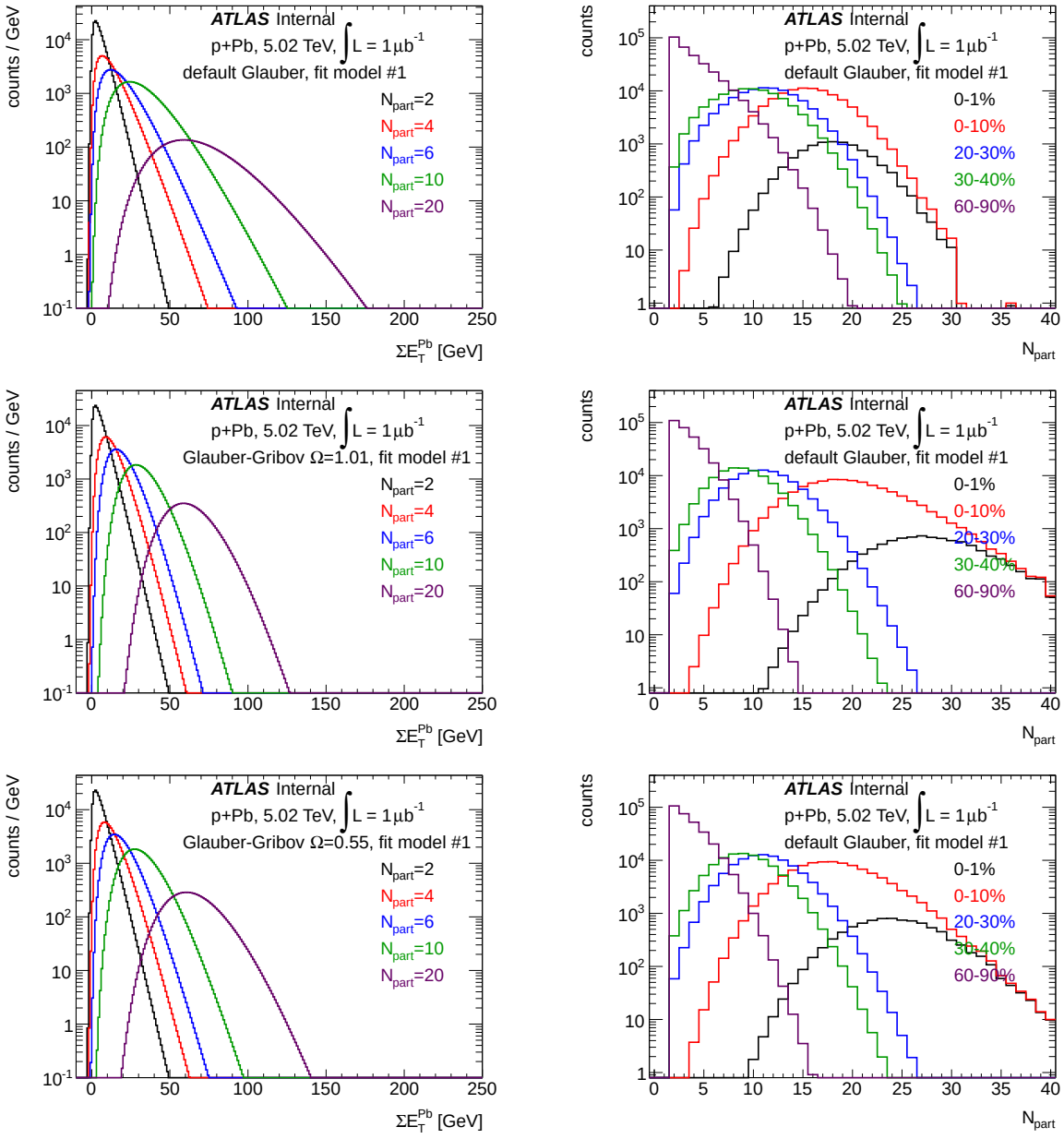


Figure 7.33: Correlation between $N_{\rm part}$ and $\Sigma E_{\rm T}^{\rm Pb}$ from global fits. Left column: contribution to the total $\Sigma E_{\rm T}^{\rm Pb}$ distribution from $N_{\rm part} = 2, 4, 6, 10, 20$ events. Right columns: $N_{\rm part}$ values contributing to the given $\Sigma E_{\rm T}^{\rm Pb}$ selections. Results are shown for fit model #1 with Glauber (top row), Glauber-Gribov $\Omega = 1.01$ (middle row) and Glauber-Gribov $\Omega = 0.55$ (bottom row).

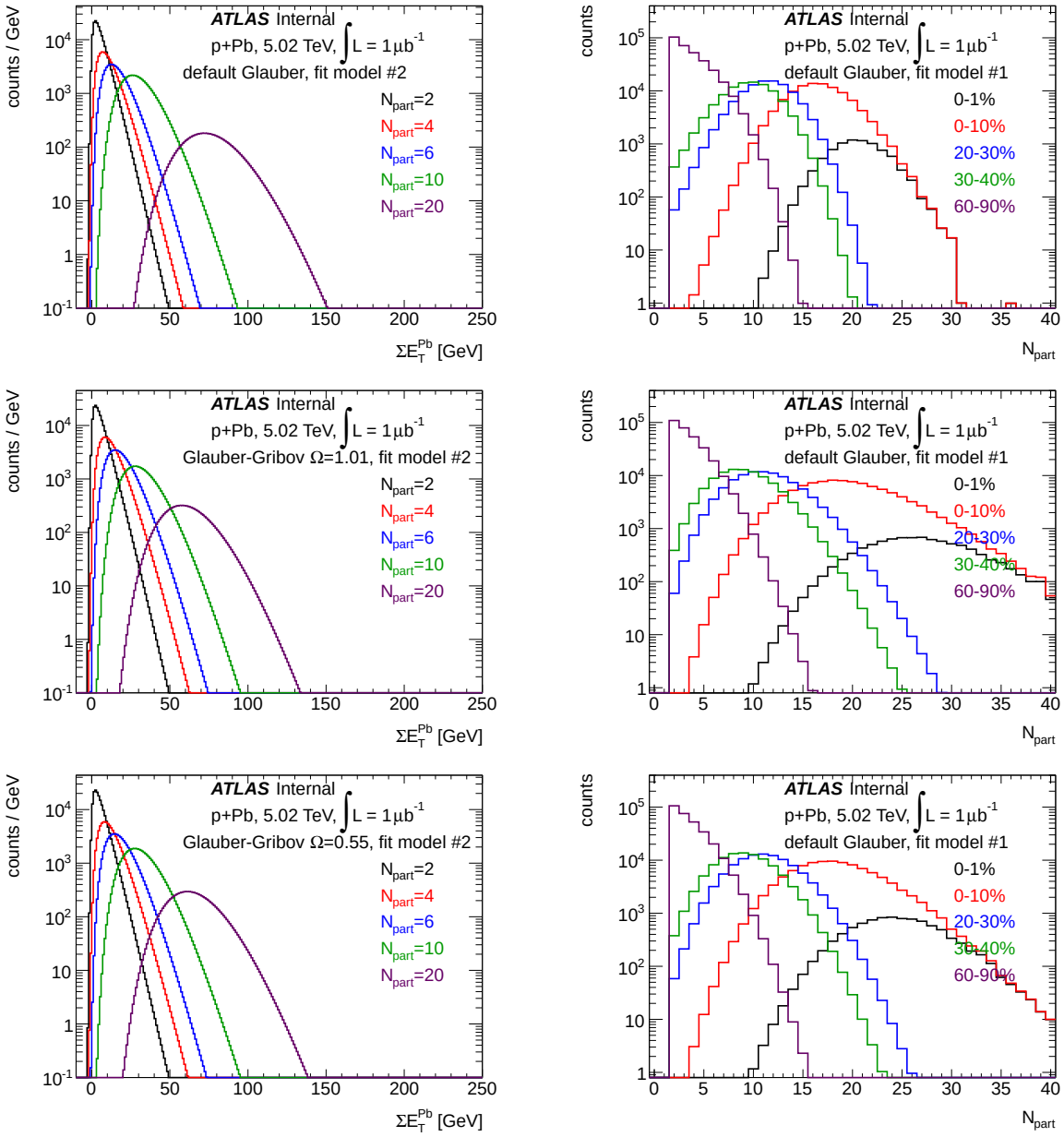


Figure 7.34: Correlation between N_{part} and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ from global fits. Left column: contribution to the total $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution from $N_{\mathrm{part}} = 2, 4, 6, 10, 20$ events. Right columns: N_{part} values contributing to the given $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ selections. Results are shown for fit model #2 with Glauber (top row), Glauber-Gribov $\Omega = 1.01$ (middle row) and Glauber-Gribov $\Omega = 0.55$ (bottom row).

distributions that are more localized in $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. In fit model #1, the figures show that a very wide range of N_{part} values contribute to the top 0-1% of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ events. In fit model #2, the range is somewhat narrower. That is, the correlation between $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ and N_{part} are tighter in fit model #2 than in fit model #1.

7.6 Systematic Uncertainties

The systematic uncertainty on the final extracted geometric quantities are discussed here. There are five major sources of systematic uncertainty, which are each detailed below. Each systematic uncertainty is associated with a systematic variation of the global fit procedure, which modifies the final results. The systematic uncertainty from any variation (or vategory of variatoins) quantifies the sensitive of the results to that detail, choice or assumption in the analysis.

Of the systematic uncertainties presented here, it should be noted that the differences in the results between the Glauber models are *not* considered a systematic uncertainty on the "final" results. Rather, the different Glauber models are considered to be equally likely descriptions of the underlying physics of p+Pb collisions and results for each are determined separately.

The effect of the systematic variations on the results are summarized in Figures 7.35 and 7.36 for the default Glauber model. The relative change in the mean N_{part} to that in the default is shown for each systematic variation. Similar results for the systematic change in the $\langle N_{\mathrm{part}} \rangle$ results within the Glauber-Gribov models are shown in the Appendix in Figures A.7 and A.8 (for $\Omega = 1.01$) and in Figures A.9 and A.10 (for $\Omega = 0.55$).

7.6.1 Uncertainty from choice of fit model and MC reference

This systematic accounts for different models of the N_{part} dependence of $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ as described in Section 7.5.2, and for using PYTHIA 6 instead of PYTHIA 8 as the MC reference as described in Section 7.4.4. The $\langle N_{\mathrm{part}} \rangle$ results using PYTHIA 6 are systematically higher than that their PYTHIA 8 counterparts, because PYTHIA 6 has a softer description of inelastic collisions and thus the fit necessitates a smaller residual contamination at low-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$. Thus, more of the low-$N_{\mathrm{part}}$ part of the total distribution is contained below the 90% centile, enhancing the N_{part} in the other bins.

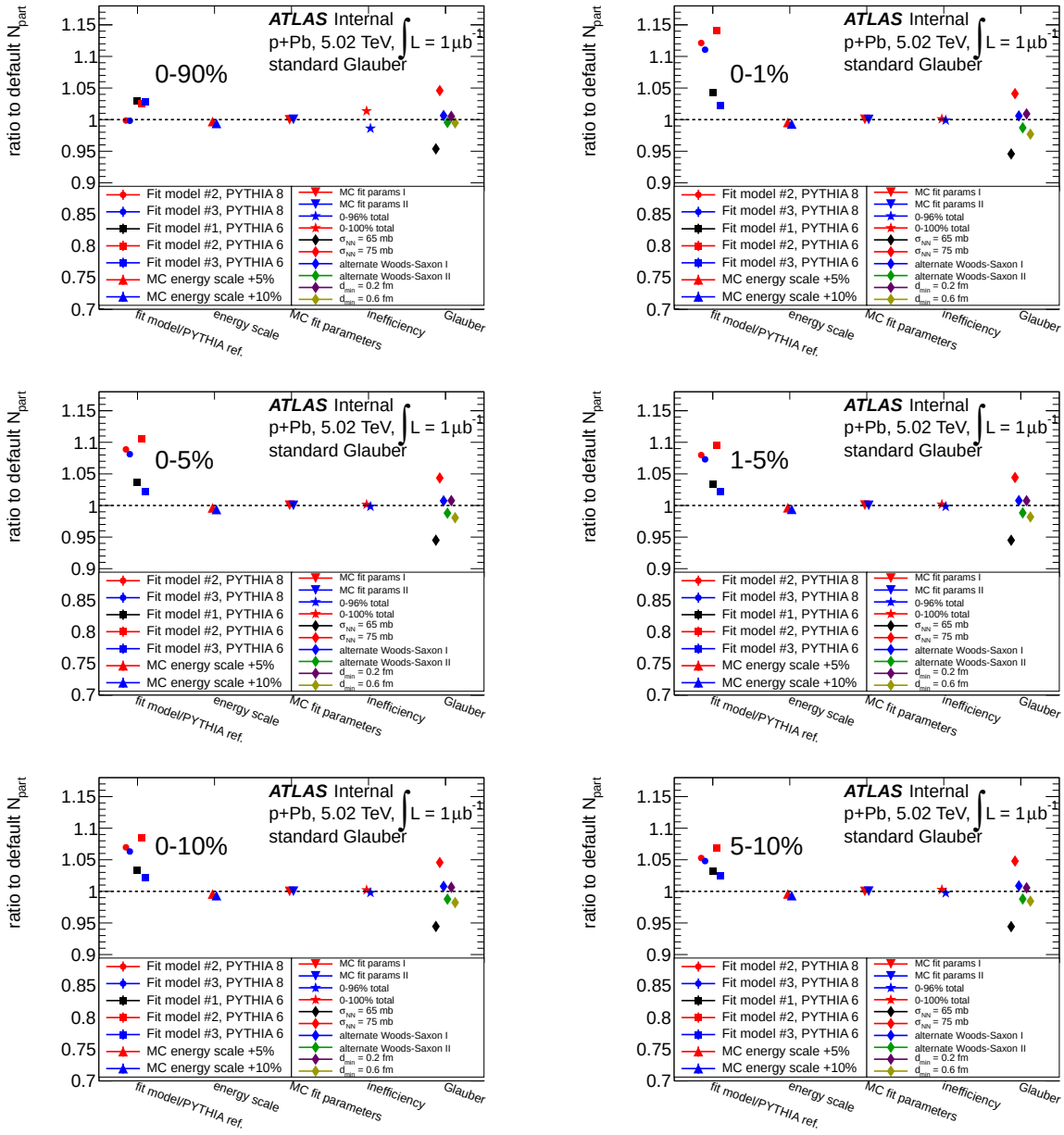


Figure 7.35: Overview of systematic uncertainty for the default Glauber model. The change in $\langle N_{\mathrm{part}} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on N_{part} in the 0-90% (top left), 0-1% (top right), 0-5% (middle left), 1-5% (middle right), 0-10% (bottom left) and 5-10% (bottom right) centrality bins.

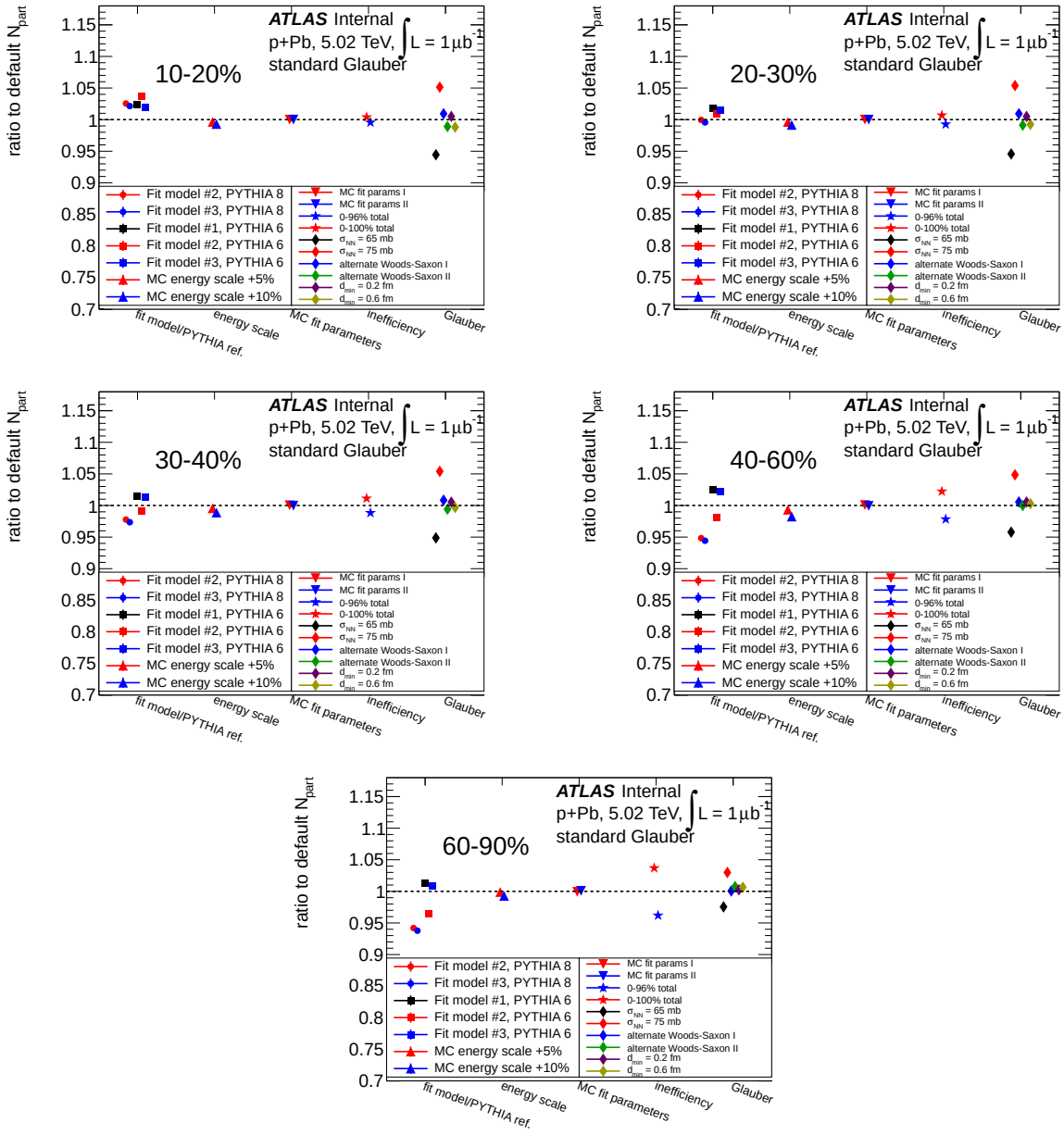


Figure 7.36: Overview of systematic uncertainty for the default Glauber model. The change in $\langle N_{\mathrm{part}} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on N_{part} in the 10-20% (top left), 20-30% (top right), 30-40% (middle left), 40-60% (middle right) and 60-90% (bottom) centrality bins.

The choice of fit model is one of the dominant systematic uncertainties. As described in Section 7.5.3, fit model #2 has a quantitatively different description of the fluctuations in $P(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}|N_{\mathrm{part}})$ at high-N_{part}, especially in the default Glauber model. Because of this, results with fit model #2 tend to produce somewhat smaller $\langle N_{\mathrm{part}}\rangle$ in peripheral selections (e.g. see last plot in Figure 7.36) and larger $\langle N_{\mathrm{part}}\rangle$ in central ones (including as much as 15% higher in 0-1%, as shown in the second plot of Figure 7.35). This reflects the fact that in the default Glauber fits, fit model #2 has a tighter correlation between N_{part} and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

7.6.2 Uncertainty from MC energy scale

To account for a possible mismatch between the FCal energy scale in the Monte Carlo and data, the θ_0 parameter in the fits to MC which is used to set the $N_{\mathrm{part}} = 2$ behavior is scaled up by 5% and 10%. This is appropriate, because θ in a Gamma distribution controls the overall scale of the distribution (it is the "scale parameter", whereas k is the "shape parameter"). 10% is seen is as an energy scale difference between the FCal in data and MC that is plausible from previous studied of the performance in Pb+Pb events[249], but on the upper limit of possibilities.

As it turns out, the results are relatively insensitive to such changes in the energy scale, since other free parameters in the model can absorb changes in the energy scale in the fit. Thus, most of the (small) differences are at low $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

7.6.3 Uncertainty from modeling of the MC reference

The fit parameters in Section 7.4.4 are varied by their uncertainty in a way that respects the correlations them. The (k, θ) of the Gamma distribution is alternately modeled as $(k + \sigma_k, \theta - \sigma_\theta)$ and $(k - \sigma_k, \theta + \sigma_\theta)$ where σ_k and σ_θ are the errors on k and θ from the fit, respectively. For similar reasons as given in 7.6.2, as well as the small level of uncertainty on the $N_{\mathrm{part}} = 2$ fits, the final results are relatively insensitive to these variations.

7.6.4 Uncertainty from total inefficiency

The nominal total efficiency (efficiency + residual contamination) for the Glauber cross-section is estimated to be 98%, with a ± 2 uncertainty from differences in the total efficiency and the amount of residual diffraction. Thus, the final $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ centrality bins can be redefined assuming a 96% and

100% efficiency, respectively. (For example, under the default analysis the 0-10% bin encloses the highest-$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ $\frac{10}{98}^{th}$'s of the data, while an efficiency of 96% implies that the 0-10% bin should enclose the highest-E_{T} $\frac{10}{96}^{th}$'s.) The N_{part} values are then determined again using these alternate definitions.

The most peripheral bins are the most sensitive to systematic variations in the total inefficiency, with the systematic uncertainty reaching just under 5% in the 60-90% bin (last plot in Figure 7.36). The effect of this variation decreases with increasing centrality, becoming $< 1\%$ by 20% centrality.

7.6.5 Uncertainty from Glauber MC variations

Six independent variations of the parameters in the Glauber MC model are considered, as follows:

- The nucleon-nucleon inelastic cross-section σ_{NN} at $\sqrt{s} = 5.02$ TeV is varied by $\pm$5mb, from 65 mb to 75 mb.
- The Woods-Saxon radius $R = 6.62 \pm 0.06$ fm and skin depth $a = 0.546 \pm 0.010$ fm parameters are varied by their uncertainty in a way that respects the anticorrelation between them. That is, the default parameters are $(R, a)_{default} = 0.546$ fm) and the two alternate are $(R, a)_{syst,I} = (6.68 \text{ fm}, 0.536 \text{ fm})$ and $(R, a)_{syst,I} = (6.56 \text{ fm}, 0.556 \text{ fm})$.
- The hard core radius d_{min} is varied by $\pm$0.2 fm, from 0.2 fm to 0.6 fm.

For each variation, an alternate N_{part} distribution entirely is generated, and the global fits are repeated to this different description of the Glauber p+Pb events.

Of these six variations, the σ_{NN} changes have the largest effect (as expected, a higher σ_{NN} increases the range N_{part} values, and vice versa) and are roughly constant as a function of centrality.

7.6.6 Total systematic uncertainties

The final systematic uncertainty on $\langle N_{\mathrm{part}} \rangle$ is derived from adding the largest deviation in each of the five categories above in quadrature. Positive and negative uncertainties are tabulated separately, and the final results for all centrality bins are shown in Table 7.6.

To determine the systematic uncertainty on $\langle T_{\mathrm{AB}} \rangle$, the variations in $\langle T_{\mathrm{AB}} \rangle = \langle N_{\mathrm{part}} - 1 \rangle / \sigma_{NN}$ are tabulated and the largest deviations in each of the fit categories are added in quadrature.

Centrality selection	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (default)**	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (0-96%)**	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (0-100%)**
2012 data			
0-1%	$>$ 90.87 GeV	$>$ 90.59 GeV	$>$ 91.17 GeV
0-5%	$>$ 65.84 GeV	$>$ 65.50 GeV	$>$ 66.17 GeV
0-10%	$>$ 53.65 GeV	$>$ 53.27 GeV	$>$ 54.02 GeV
10-20%	39.96-53.65 GeV	39.52-53.27 GeV	40.39-54.02 GeV
20-30%	31.06-39.96 GeV	30.59-39.52 GeV	31.53-40.39 GeV
30-40%	24.15-31.06 GeV	23.63-30.59 GeV	24.66-31.53 GeV
40-60%	13.55-24.15 GeV	13.00-23.63 GeV	14.10-24.66 GeV
60-90%	2.817-13.55 GeV	2.284-13.00 GeV	3.319-14.10 GeV

Table 7.5: $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ ranges in 2012 p+Pb data for given centrality selections.

Thus, much of the uncertainty from varying σ_{NN} in the Glauber MC is reduced from the partial cancellation of uncertainties in the ratio. The final uncertainty is shown in Table 7.7.

Similarly, the systematic uncertainty in R_{coll} is determined by considering the worst change in $R_{\mathrm{coll}} = \langle N_{\mathrm{part}} - 1 \rangle^{\mathrm{cent}} \,/\, \langle N_{\mathrm{part}} - 1 \rangle^{60-90\%}$ in each category and adding these in quadrature. Again, the uncertainty from varying σ_{NN} cancels to a large extent in the ratio. However, the uncertainty from the choice of fit model is correlated in a way to give a large one-sided uncertainty in R_{coll}. This is understood to be because of the differences in how tightly $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ is correlated with N_{part} in fit models #1 and #2. In fit model #2, the central events have a higher mean N_{part} and (to compensate), the peripheral events have a lower mean N_{part} than in fit model #1. Thus, the effect is amplified in the central-to-peripheral ratio as expressed in the R_{coll}. The final uncertainties are shown in Table 7.8.

7.7 Results

7.7.1 Centrality divisions

As found in Section 7.5.3, the best fit of the Glauber N_{part} distribution to data is consistent with a nominal 98% total efficiency, comprising 96-98% of the Glauber inelastic events plus an additional 0-2% of the data being consistent with non-Glauber diffractive events. We define ranges of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ which correspond to given fractions of the Glauber cross section. That is, a centrality selection

X^{low}%-X^{high}% is given by a range of $\Sigma E_{\text{T}}^{\text{Pb}}$ values $\in (\Sigma E_{\text{T}}^{\text{Pb,low}}, \Sigma E_{\text{T}}^{\text{Pb,high}})$ such that X^{low}% / 98% of minimum bias p+Pb events have a higher $\Sigma E_{\text{T}}^{\text{Pb}}$

$$\frac{\int_{\Sigma E_{\text{T}}^{\text{Pb,high}}}^{+\infty} d\Sigma E_{\text{T}}^{\text{Pb}} \frac{dN}{d\Sigma E_{\text{T}}^{\text{Pb}}}}{\int_{0}^{+\infty} d\Sigma E_{\text{T}}^{\text{Pb}} \frac{dN}{d\Sigma E_{\text{T}}^{\text{Pb}}}} = \frac{X^{\text{,low}}}{98} \tag{7.14}$$

and that the $98 - X^{\text{,high}}$% / 98% of minimum bias p+Pb events have a lower $\Sigma E_{\text{T}}^{\text{Pb}}$,

$$\frac{\int_{0}^{\Sigma E_{\text{T}}^{\text{Pb,low}}} d\Sigma E_{\text{T}}^{\text{Pb}} \frac{dN}{d\Sigma E_{\text{T}}^{\text{Pb}}}}{\int_{0}^{+\infty} d\Sigma E_{\text{T}}^{\text{Pb}} \frac{dN}{d\Sigma E_{\text{T}}^{\text{Pb}}}} = \frac{98 - X^{\text{,high}}}{98} \tag{7.15}$$

The divisions of the data into centrality bins are shown in the second column of Table 7.5 in the 2012 data. Because of the slightly different treatment of the uncorrected $\Sigma E_{\text{T}}^{\text{Pb}}$ values in 2012 and 2013 data, the $\Sigma E_{\text{T}}^{\text{Pb}}$ cuts corresponding to the centrality divisions are slightly different between the two (though the geometric quantities within those bins are extremely similar). The 2013 centrality divisions are shown in Table A.1 in the Appendix. The Tables also show how the $\Sigma E_{\text{T}}^{\text{Pb}}$ cuts change if they are to be redrawn with different assumptions on the total efficiency.

The most peripheral centile for which geometric quantities are extracted is 90% centile. Events with an $\Sigma E_{\text{T}}^{\text{Pb}}$ lower than this threshold are not used for two reasons. First, the determination of their N_{part} values suffers from a large systematic uncertainty due to the uncertainty in the total efficiency. Second, in the Glauber model, $N_{\text{part}} = 2$ p+Pb collisions are ≈ 18% of the total cross-section. Thus, selecting events with such low-$\Sigma E_{\text{T}}^{\text{Pb}}$ actually preferentially selects pp like events but with a suppressed soft underlying event and multiplicity of produced particles, which biases the events. Because of these effects, the 60-90% bin is the most peripheral one which should be used in centrality-dependent p+Pb analyses.

7.7.2 Determination of N_{part}, T_{AB} and R_{coll}

The geometric quantities N_{part}, R_{coll} and T_{AB} are extracted as follows. To determine the mean N_{part} within a given centrality selection, the global fit is used to determine the composition of N_{part} events which fall into the given $\Sigma E_{\text{T}}^{\text{Pb}}$ range. Formally, for a centrality bin defined by $\Sigma E_{\text{T}}^{\text{Pb}} \in (\Sigma E_{\text{T}}^{\text{Pb,low}}, \Sigma E_{\text{T}}^{\text{Pb,high}})$, the mean number of participants N_{part} for those events is

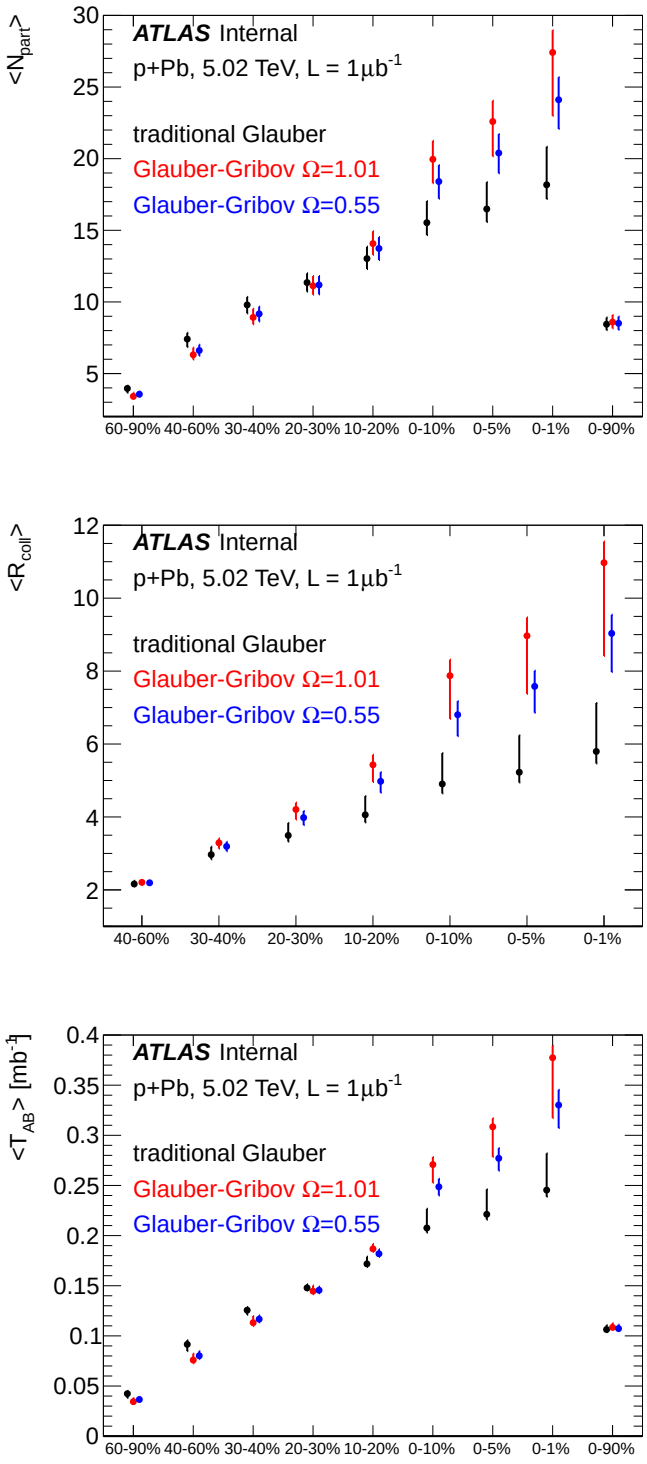


Figure 7.37: Final $\langle N_{\rm part}\rangle$ (top left), $R_{\rm coll}$ (top right) and $\langle T_{\rm AB}\rangle$ (bottom) results in selected centrality bins, with final systematic errors. Results are shown from fits to 2012 data, and from all three Glauber models.

$$\langle N_{\rm part}\rangle = \frac{\int_{\Sigma E_{\rm T}^{\rm Pb,low}}^{\Sigma E_{\rm T}^{\rm Pb,high}} d\Sigma E_{\rm T}^{\rm Pb} \sum_{N_{\rm part}} N_{\rm part} P(N_{\rm part}) P(\Sigma E_{\rm T}^{\rm Pb}|N_{\rm part}) \frac{dN}{d\Sigma E_{\rm T}^{\rm Pb}}}{\int_{\Sigma E_{\rm T}^{\rm Pb,low}}^{\Sigma E_{\rm T}^{\rm Pb,high}} d\Sigma E_{\rm T}^{\rm Pb} \frac{dN}{d\Sigma E_{\rm T}^{\rm Pb}}} \tag{7.16}$$

The ratio of number of nucleon-nucleon collisions in a given centrality bin *cent* and the 60-90% bin is simple determine via the relationship $N_{\rm part} = N_{\rm coll} + 1$ in p+A collisions. Simply,

$$\langle R_{\rm coll}\rangle = \frac{\left\langle {N_{\rm part}}^{\rm cent}\right\rangle - 1}{\left\langle {N_{\rm part}}^{60-90\%}\right\rangle - 1} \tag{7.17}$$

where the numerator and denominator are each determined separately in their $\Sigma E_{\rm T}^{\rm Pb}$ bins. The nuclear thickness overlap function is

$$\left\langle {T_{\rm AB}}^{\rm cent}\right\rangle = \left(\left\langle {N_{\rm part}}^{\rm cent}\right\rangle - 1\right)/\sigma_{\rm NN} \tag{7.18}$$

For most of the centrality divisions, the $\Sigma E_{\rm T}^{\rm Pb}$ ranges used are taken directly from data as described above, even though the fraction of events in the model and in data within the given $\Sigma E_{\rm T}^{\rm Pb}$ range may differ slightly. However, when cutting on small centrality selections at the high tail of the $\Sigma E_{\rm T}^{\rm Pb}$ distribution, this difference is more problematic. Because no fit describes the large-$\Sigma E_{\rm T}^{\rm Pb}$ tail well, it is possible that the $\Sigma E_{\rm T}^{\rm Pb}$ range in data which contains the top 1% of events may contain up to 50% more or fewer events in the model (e.g. the model is known to over- or undershoot). Thus, the 1% and 5% centile divisions are redrawn to match what they should be in the model only for the purposes of extracting $\langle N_{\rm part}\rangle$.

Tables 7.6, 7.7 and 7.8 shows the final results for $\langle N_{\rm part}\rangle$, $\langle T_{\rm AB}\rangle$ and $\langle R_{\rm coll}\rangle$ for each of the three Glauber models along with their final systematic uncertainties. Additionally, these are summarized graphically in Figure 7.37.

Furthermore, Tables A.2, A.3 and A.4 in the Appendix show the same results in 2013 data. The agreement between the 2012 and 2013 is extremely good, with most results in agreement to within $< 1\%$ or less.

The results are discussed in Chapter 9.

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2012 data			
60-90%	$3.964^{+0.195}_{-0.306}\left(^{+4.92\%}_{-7.73\%}\right)$	$3.408^{+0.261}_{-0.155}\left(^{+7.67\%}_{-4.55\%}\right)$	$3.558^{+0.191}_{-0.171}\left(^{+5.37\%}_{-4.81\%}\right)$
40-60%	$7.408^{+0.437}_{-0.558}\left(^{+5.89\%}_{-7.53\%}\right)$	$6.313^{+0.501}_{-0.335}\left(^{+7.94\%}_{-5.3\%}\right)$	$6.613^{+0.403}_{-0.371}\left(^{+6.09\%}_{-5.6\%}\right)$
30-40%	$9.791^{+0.56}_{-0.587}\left(^{+5.72\%}_{-5.99\%}\right)$	$8.92^{+0.606}_{-0.486}\left(^{+6.8\%}_{-5.44\%}\right)$	$9.168^{+0.506}_{-0.529}\left(^{+5.51\%}_{-5.78\%}\right)$
20-30%	$11.35^{+0.648}_{-0.632}\left(^{+5.71\%}_{-5.57\%}\right)$	$11.12^{+0.686}_{-0.621}\left(^{+6.16\%}_{-5.58\%}\right)$	$11.19^{+0.624}_{-0.659}\left(^{+5.58\%}_{-5.89\%}\right)$
10-20%	$13.03^{+0.826}_{-0.732}\left(^{+6.34\%}_{-5.62\%}\right)$	$14.07^{+0.865}_{-0.79}\left(^{+6.14\%}_{-5.61\%}\right)$	$13.73^{+0.794}_{-0.811}\left(^{+5.78\%}_{-5.9\%}\right)$
0-10%	$15.53^{+1.5}_{-0.869}\left(^{+9.69\%}_{-5.6\%}\right)$	$19.96^{+1.29}_{-1.67}\left(^{+6.47\%}_{-8.37\%}\right)$	$18.4^{+1.14}_{-1.2}\left(^{+6.21\%}_{-6.52\%}\right)$
5-10%	$14.58^{+1.22}_{-0.82}\left(^{+8.4\%}_{-5.62\%}\right)$	$17.43^{+1.14}_{-1.07}\left(^{+6.51\%}_{-6.13\%}\right)$	$16.47^{+1}_{-1}\left(^{+6.1\%}_{-6.08\%}\right)$
0-5%	$16.49^{+1.87}_{-0.912}\left(^{+11.4\%}_{-5.53\%}\right)$	$22.59^{+1.45}_{-2.43}\left(^{+6.44\%}_{-10.7\%}\right)$	$20.4^{+1.32}_{-1.42}\left(^{+6.48\%}_{-6.97\%}\right)$
1-5%	$16.06^{+1.68}_{-0.889}\left(^{+10.5\%}_{-5.54\%}\right)$	$21.4^{+1.43}_{-1.98}\left(^{+6.7\%}_{-9.26\%}\right)$	$19.46^{+1.28}_{-1.28}\left(^{+6.56\%}_{-6.59\%}\right)$
0-1%	$18.18^{+2.66}_{-0.995}\left(^{+14.6\%}_{-5.47\%}\right)$	$27.41^{+1.55}_{-4.43}\left(^{+5.67\%}_{-16.1\%}\right)$	$24.11^{+1.57}_{-2.03}\left(^{+6.53\%}_{-8.44\%}\right)$
0-90%	$8.444^{+0.473}_{-0.411}\left(^{+5.6\%}_{-4.87\%}\right)$	$8.594^{+0.5}_{-0.444}\left(^{+5.82\%}_{-5.16\%}\right)$	$8.51^{+0.469}_{-0.458}\left(^{+5.52\%}_{-5.38\%}\right)$

Table 7.6: Mean N_{part} values for all centrality bins in 2012 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2012 data			
60-90%	$0.04234^{+0.00283}_{-0.00433}\left(^{+6.68\%}_{-10.2\%}\right)$	$0.03439^{+0.00373}_{-0.0021}\left(^{+10.8\%}_{-6.1\%}\right)$	$0.03655^{+0.00266}_{-0.00224}\left(^{+7.27\%}_{-6.13\%}\right)$
40-60%	$0.09154^{+0.00418}_{-0.00673}\left(^{+4.56\%}_{-7.35\%}\right)$	$0.0759^{+0.0065}_{-0.00325}\left(^{+8.57\%}_{-4.29\%}\right)$	$0.08019^{+0.00459}_{-0.00334}\left(^{+5.72\%}_{-4.17\%}\right)$
30-40%	$0.1256^{+0.00327}_{-0.00454}\left(^{+2.61\%}_{-3.61\%}\right)$	$0.1131^{+0.00664}_{-0.00333}\left(^{+5.87\%}_{-2.94\%}\right)$	$0.1167^{+0.00377}_{-0.00316}\left(^{+3.23\%}_{-2.71\%}\right)$
20-30%	$0.1479^{+0.00357}_{-0.00259}\left(^{+2.41\%}_{-1.75\%}\right)$	$0.1446^{+0.00561}_{-0.00311}\left(^{+3.88\%}_{-2.15\%}\right)$	$0.1455^{+0.00364}_{-0.00303}\left(^{+2.5\%}_{-2.08\%}\right)$
10-20%	$0.1718^{+0.00722}_{-0.00296}\left(^{+4.2\%}_{-1.72\%}\right)$	$0.1868^{+0.00501}_{-0.00292}\left(^{+2.68\%}_{-1.56\%}\right)$	$0.1819^{+0.0044}_{-0.00313}\left(^{+2.42\%}_{-1.72\%}\right)$
0-10%	$0.2076^{+0.0192}_{-0.00465}\left(^{+9.23\%}_{-2.24\%}\right)$	$0.2708^{+0.00754}_{-0.0182}\left(^{+2.78\%}_{-6.72\%}\right)$	$0.2486^{+0.00803}_{-0.00865}\left(^{+3.23\%}_{-3.48\%}\right)$
5-10%	$0.194^{+0.0146}_{-0.00391}\left(^{+7.5\%}_{-2.01\%}\right)$	$0.2348^{+0.00658}_{-0.00687}\left(^{+2.8\%}_{-2.93\%}\right)$	$0.221^{+0.00617}_{-0.00498}\left(^{+2.79\%}_{-2.25\%}\right)$
0-5%	$0.2212^{+0.0249}_{-0.00538}\left(^{+11.3\%}_{-2.43\%}\right)$	$0.3084^{+0.00858}_{-0.03}\left(^{+2.78\%}_{-9.73\%}\right)$	$0.2771^{+0.0102}_{-0.0124}\left(^{+3.67\%}_{-4.47\%}\right)$
1-5%	$0.2152^{+0.022}_{-0.00506}\left(^{+10.2\%}_{-2.35\%}\right)$	$0.2915^{+0.00789}_{-0.0227}\left(^{+2.71\%}_{-7.79\%}\right)$	$0.2637^{+0.00914}_{-0.00968}\left(^{+3.47\%}_{-3.67\%}\right)$
0-1%	$0.2454^{+0.0366}_{-0.00664}\left(^{+14.9\%}_{-2.71\%}\right)$	$0.3773^{+0.0119}_{-0.0601}\left(^{+3.16\%}_{-15.9\%}\right)$	$0.3302^{+0.0153}_{-0.023}\left(^{+4.62\%}_{-6.96\%}\right)$
0-90%	$0.1063^{+0.00443}_{-0.00265}\left(^{+4.16\%}_{-2.49\%}\right)$	$0.1085^{+0.00432}_{-0.00241}\left(^{+3.99\%}_{-2.22\%}\right)$	$0.1073^{+0.00393}_{-0.00259}\left(^{+3.67\%}_{-2.41\%}\right)$

Table 7.7: Mean T_{AB} values for all centrality bins in 2012 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2012 data			
40-60% / 60-90%	$2.162^{+0.089}_{-0.0648}\left(^{+4.12\%}_{-3\%}\right)$	$2.207^{+0.0529}_{-0.0611}\left(^{+2.4\%}_{-2.77\%}\right)$	$2.194^{+0.0571}_{-0.0589}\left(^{+2.6\%}_{-2.69\%}\right)$
30-40% / 60-90%	$2.967^{+0.218}_{-0.129}\left(^{+7.36\%}_{-4.34\%}\right)$	$3.29^{+0.122}_{-0.16}\left(^{+3.72\%}_{-4.86\%}\right)$	$3.192^{+0.127}_{-0.133}\left(^{+3.99\%}_{-4.16\%}\right)$
20-30% / 60-90%	$3.493^{+0.344}_{-0.17}\left(^{+9.84\%}_{-4.87\%}\right)$	$4.205^{+0.188}_{-0.275}\left(^{+4.48\%}_{-6.55\%}\right)$	$3.982^{+0.183}_{-0.206}\left(^{+4.6\%}_{-5.17\%}\right)$
10-20% / 60-90%	$4.058^{+0.514}_{-0.209}\left(^{+12.7\%}_{-5.15\%}\right)$	$5.43^{+0.276}_{-0.474}\left(^{+5.09\%}_{-8.72\%}\right)$	$4.977^{+0.251}_{-0.313}\left(^{+5.05\%}_{-6.29\%}\right)$
0-10% / 60-90%	$4.904^{+0.848}_{-0.259}\left(^{+17.3\%}_{-5.28\%}\right)$	$7.875^{+0.439}_{-1.19}\left(^{+5.58\%}_{-15.1\%}\right)$	$6.803^{+0.373}_{-0.588}\left(^{+5.48\%}_{-8.64\%}\right)$
5-10% / 60-90%	$4.582^{+0.715}_{-0.241}\left(^{+15.6\%}_{-5.25\%}\right)$	$6.826^{+0.376}_{-0.793}\left(^{+5.51\%}_{-11.6\%}\right)$	$6.046^{+0.326}_{-0.453}\left(^{+5.39\%}_{-7.49\%}\right)$
0-5% / 60-90%	$5.226^{+1.01}_{-0.283}\left(^{+19.4\%}_{-5.41\%}\right)$	$8.968^{+0.504}_{-1.59}\left(^{+5.62\%}_{-17.8\%}\right)$	$7.581^{+0.429}_{-0.724}\left(^{+5.66\%}_{-9.55\%}\right)$
1-5% / 60-90%	$5.082^{+0.935}_{-0.271}\left(^{+18.4\%}_{-5.33\%}\right)$	$8.475^{+0.488}_{-1.36}\left(^{+5.76\%}_{-16.1\%}\right)$	$7.215^{+0.413}_{-0.64}\left(^{+5.72\%}_{-8.86\%}\right)$
0-1% / 60-90%	$5.796^{+1.33}_{-0.329}\left(^{+22.9\%}_{-5.67\%}\right)$	$10.97^{+0.582}_{-2.56}\left(^{+5.31\%}_{-23.3\%}\right)$	$9.034^{+0.512}_{-1.06}\left(^{+5.66\%}_{-11.7\%}\right)$

Table 7.8: Mean R_{coll} values for all centrality bins in 2012 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

Chapter 8

Inclusive jet production in p+Pb collisions

Thou'st spoken right. 'Tis true.
The wheel is come full circle. I am here.

King Lear, Act V, Scene III

This chapter describes the measurement of inclusive jet production in 5.02 TeV p+Pb and 2.76 TeV pp collisions. The measured quantities and choice of binning are defined in Section 8.1. The data, reconstruction and event selection are described in Section 8.2. The jet reconstruction and calibration procedure, as well as the correction for selection efficiency is described in Section 8.3. The event filter trigger selection and determination of the trigger efficiency is described in Section 8.4. The jet performance in ATLAS and the corrections from detector effects are described in Section 8.5. Additional cross-checks on the jet energy resolution and scale are described in Section 8.6. The main results are shown in Section 8.7, with a description of the systematic uncertainties given in Section 8.8.

8.1 Definition of Measured Quantities

The basic quantity of interest is the double differential per-event jet yield $(1/N_{\text{evt}})(d^2N/dp_{\text{T}}dy)$ for jets with given transverse momentum p_{T} and rapidity y. Experimentally, the yield is measured in

p_{T} and y bins with center $p_{\mathrm{T}}{}'$ and y' and of size Δp_{T} and Δy, respectively, such that

$$\frac{d^2N}{dp_{\mathrm{T}}dy}(p_{\mathrm{T}}{}', y') \approx \int_{p_{\mathrm{T}}\in\Delta p_{\mathrm{T}}} \frac{dp_{\mathrm{T}}}{\Delta p_{\mathrm{T}}} \int_{y\in\Delta y} \frac{dy}{\Delta y} \frac{d^2N}{dp_{\mathrm{T}}dy} = \frac{1}{\Delta p_{\mathrm{T}}\Delta y} N_{\mathrm{jet}}(p_{\mathrm{T}} \in \Delta p_{\mathrm{T}}, y \in \Delta y) \tag{8.1}$$

In this work, the rapidity of jet $y\left(=\frac{1}{2}\log\frac{E+p_z}{E-p_z}\right)$ is always referred to in the center of mass frame of the p+Pb collision system and is denoted y^* whenever possible. Rapidity in the laboratory frame is always denoted y^{lab}. The per-event yield is given by

$$\frac{1}{N_{\mathrm{evt}}} \frac{1}{\Delta p_{\mathrm{T}}\Delta y} N_{\mathrm{jet}}(p_{\mathrm{T}} \in \Delta p_{\mathrm{T}}, y \in \Delta y) \tag{8.2}$$

The number of events N_{evt} is not the number of jet events, but rather the total effective number of p+Pb events sampled in which the jet yield was measured.

8.1.1 Centrality-dependent jet yields

Per-event jet yields are defined in the different centrality selections by restricting the jet yield and the N_{evt} only to events with the given centrality. They are denoted

$$\frac{1}{{N_{\mathrm{evt}}}^{\mathrm{cent}}} \frac{d^2N^{\mathrm{cent}}}{dp_{\mathrm{T}}dy}(p_{\mathrm{T}}{}', y') \tag{8.3}$$

In this work, the centrality selections considered are 0-10%, 10-20%, 20-30%, 30-40%, 40-60% and 60-90%. The per-event yield without any requirement on event centrality in Equation 8.2 is sometimes referred to as the 0-98% or the minimum bias per-event yield.

The two centrality-dependent physics observables are the R_{pPb} and the R_{CP}.

The p+Pb-to-pp nuclear modification factor ${R_{\mathrm{pPb}}}^{\mathrm{cent}}(p_{\mathrm{T}}, y)$ for a given centrality selection *cent* is the ratio of the per-event jet yield in p+Pb to the jet production cross-section $d^2\sigma^{pp}(p_{\mathrm{T}}, y)/dp_{\mathrm{T}}dy$ in pp collisions scaled by the nucleonic luminosity $\left\langle T_{AB}^{\mathrm{cent}}\right\rangle$ (also known as the *nuclear overlap function*) for p+Pb collisions in that centrality class. The R_{pPb} quantifies the deviation of jet production in different geometric configurations of p+Pb collisions relative to the behavior expected from a naive geometric scaling of nucleon-nucleon collisions.

$${R_{\mathrm{pPb}}}^{\mathrm{cent}}(p_{\mathrm{T}}, y) = \frac{1}{{N_{\mathrm{evt}}}^{\mathrm{cent}}} \frac{d^2N^{\mathrm{cent}}}{dp_{\mathrm{T}}dy}(p_{\mathrm{T}}, y) \Big/ \left\langle T_{AB}^{\mathrm{cent}}\right\rangle \frac{d^2\sigma^{pp}}{dp_{\mathrm{T}}dy}(p_{\mathrm{T}}, y) \tag{8.4}$$

The central-to-peripheral nuclear modification factor ${R_{\rm CP}}^{\rm cent/60-90\%}(p_{\rm T}, y)$ is the ratio of the $\langle T_{AB}\rangle$-scaled yields between the jet yield for a given centrality selection *cent* to that in the peripheral centrality selection 60-90%. The $R_{\rm CP}$ quantifies the deviation of jet production between different geometric configurations of p+Pb collisions relative to the ratio assumed from a geometric scaling argument.

$$ {R_{\rm CP}}^{\rm cent/60-90\%}(p_{\rm T}, y) = \frac{1}{\left\langle T_{AB}^{\rm cent}\right\rangle}\frac{1}{{N_{\rm evt}}^{\rm cent}}\frac{d^2N^{\rm cent}}{dp_{\rm T}dy}(p_{\rm T}, y) \Bigg/ \frac{1}{\left\langle {T_{\rm AB}}^{60-90\%}\right\rangle}\frac{1}{{N_{\rm evt}}^{60-90\%}}\frac{d^2N^{60-90\%}}{dp_{\rm T}dy}(p_{\rm T}, y) \tag{8.5} $$

Although the $R_{\rm CP}$ has historically been used to quantify physics effects when pp reference data are not available, it actually offers several advantages over the traditional quantity $R_{\rm pPb}$ such as the partial or full cancellation of large systematic uncertainties from the model-dependent determination of $\langle T_{AB}\rangle$, and other uncertainties such as the trigger efficiency correction and unfolding of detector effects.

It can be written in a way that accounts for the partial cancellation of systematic uncertainties on $\langle T_{AB}\rangle$ as

$$ {R_{\rm CP}}^{\rm cent/60-90\%}(p_{\rm T}, y) = \left(\frac{1}{R_{\rm coll}^{\rm cent}}\right)\frac{1}{{N_{\rm evt}}^{\rm cent}}\frac{d^2N^{\rm cent}}{dp_{\rm T}dy}(p_{\rm T}, y) \Bigg/ \frac{1}{{N_{\rm evt}}^{60-90\%}}\frac{d^2N^{60-90\%}}{dp_{\rm T}dy}(p_{\rm T}, y) \tag{8.6} $$

where $R_{\rm coll}^{\rm cent}$ is the ratio of the number of binary collisions in events with centrality *cent* relative to that in 60-90%. $R_{\rm coll}$ ranges from ≈ 2 as many binary collisions in 40-60% collisions as in 60-90%, to ≈ 5 as many collisions in 0-10% events relative to 60-90% (see Section 8.2.4).

8.1.2 Choice of binning

Several different choices of binning are used in this analysis, with separate binning options for the p+Pb jet yields and pp jet spectra.

To better measure the $p_{\rm T}$ dependence with good statistics over more than an order of magnitude, the jet spectra are binned with exponentially growing bin sizes.

Before corrections for detector effects and jet selection cuts, the following binning is used to examine and analyze the p+Pb and pp jet yields at the detector scale:

$$p_{\mathrm{T}} \in \left(10 \cdot 10^{n/20}, 10 \cdot 10^{(n+1)/20}\right) \text{ GeV} \tag{8.7}$$

for bins $n = 0$ (lower limit $p_{\mathrm{T}} = 10$ GeV) to $n = 39$ (high limit $p_{\mathrm{T}} = 1$ TeV).

The binning used for the physics results (the R_{CP} and the R_{pPb}) is

$$p_{\mathrm{T}} \in \left(10 \cdot 10^{n/10}, 10 \cdot 10^{(n+1)/10}\right) \text{ GeV} \tag{8.8}$$

for bins $n = 0$ (lower limit $p_{\mathrm{T}} = 10$ GeV) to $n = 19$ (high limit $p_{\mathrm{T}} = 1$ TeV).

Additionally, since the 5.02 TeV reference pp spectrum is constructed by x_{T}-scaling the 2.76 TeV pp spectrum, there needs to be a binning in the 2.76 TeV spectrum which is the $\sqrt{s}$-scaled version of the binning used in 5.02 TeV. This binning is

$$p_{\mathrm{T}} \in \frac{2.76}{5.02}\left(10 \cdot 10^{n/10}, 10 \cdot 10^{(n+1)/10}\right) \text{ GeV} \tag{8.9}$$

for bins $n = 0$ (lower limit $p_{\mathrm{T}} = 5.5$ GeV) to $n = 19$ (high limit $p_{\mathrm{T}} = 550$ GeV).

Finally, the 2013 pp spectrum is compared to the 2011 pp spectrum, which uses a standard set of bins common to all ATLAS jet results, as a cross-check. In this case, the binning is given by the following bin edges

$$20, 30, 45, 60, 80, 110, 160, 210, 260, 310, 400, 500, 600, 800, 1000 \text{ GeV} \tag{8.10}$$

The y^* bins are chosen to be identical to that used in the 2.76, 7 and 8 TeV jet cross-section measurements, which comprise 13 bins,

$$y^* \in (-0.3, +0.3),\ (\pm 0.3, \pm 0.8),\ (\pm 0.8, \pm 1.2),\ (\pm 1.2, \pm 2.1), \tag{8.11}$$

$$(\pm 2.1, \pm 2.8),\ (\pm 2.8, \pm 3.6),\ (\pm 3.6, \pm 4.4) \tag{8.12}$$

In the pp jet analyses (where $y^{\mathrm{lab}} = y^*$), these bins have the benefit that for most bins the bin boundaries coincide with the boundaries of distinct subdetectors, such that the instrumentation is the same within any bin. In the p+Pb system, which has a center of mass shift relative to the laboratory frame, this is no longer the case. However, these bins are kept to be consistent with the ATLAS pp data, which will serve as the reference spectra for the R_{pPb} measurement.

8.2 Data Selection

8.2.1 Run conditions

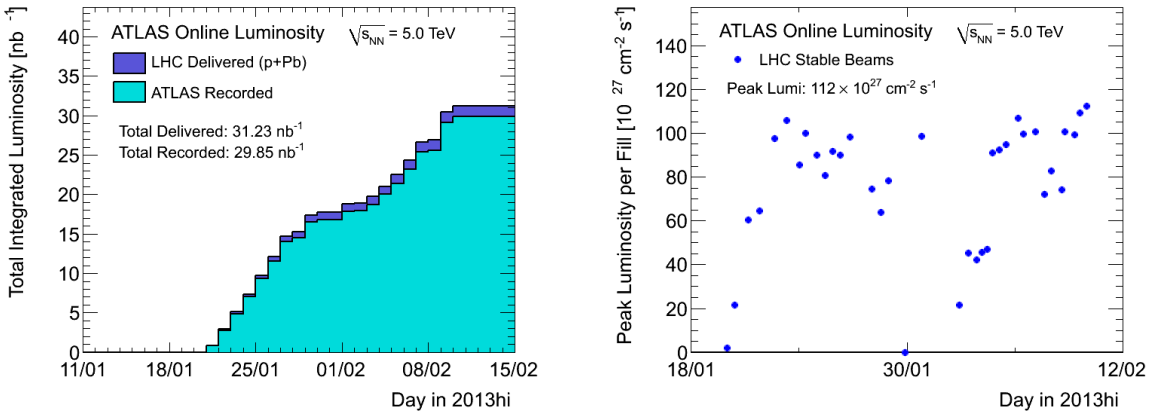


Figure 8.1: Cumulative integrated luminosity in 2013 p+Pb and Pb+p running (left) and daily peak instaneous luminosity (right) vs. day.

This analysis analyzes the LHC proton-nucleus (p+Pb) and proton-proton (pp) run in January and February 2013. The proton-nucleus running consisted of two beam orientations. The first is denoted "period A" or "p+Pb" running, with $E = 4$ TeV proton bunches in beam 1 (going from the A to C side of ATLAS) colliding with bunches of $E = Z \times 4$ TeV Pb nuclei in beam 2. The per-nucleon beam energy for the Pb nucleus is $(Z/A) \times 4$ TeV $= 1.58$ TeV. The resulting proton-nucleon center of mass frame has energy $\sqrt{s} = 5.02$ TeV and is boosted with respect to the laboratory frame with a rapidity shift $\Delta y = -0.465$. That is, $y^* = 0$ coincides with $y^{\text{lab}} = -0.465$. The LHC operated in this mode for most of January, delivering approximately 17 nb^{-1}, a luminosity estimated using MC estimates of the luminosity detector response and a p+Pb inelastic cross-section of $\sigma = 2.12$ barns.

The second mode is denoted "period B" or "Pb+p" running, with the same proton and nucleus kinematics but now switched between the two beams (protons in beam 2, nuclei in beam 1). In this work, y^* is always referred to by the convention of the period A kinematics. The LHC operated in this mode until mid-February, delivering approximately 14 nb^{-1}. The cumulative integrated luminosity and peak luminosity per day for the proton-nucleus running are shown in Figure 8.1.

Finally, the LHC delivered pp collisions with $\sqrt{s} = 2.76$ TeV, an energy chosen to coincide with the nucleon-nucleon center of mass energy in $E = 3.5$ TeV Pb+Pb running. The total delivered luminosty was 4.7 pb^{-1}, an increase of 20 in statistics from the March 2011 2.76 TeV run.

8.2.2 Reconstruction

The data were reconstructed with low-μ settings, using the *LowPt* option for the charged track reconstruction selection cuts[256; 232].

A further set of more stringent tracking cuts were applied offline:

- Track author requirement that corresponds to inside-out tracking algorithms only (implemented by requiring *trk_patternReco1 & 1* or *trk_patternReco2 & 16*).
- $p_{\mathrm{T}} > 100$ MeV, $|\eta| < 2.5$
- $\geq 2, 4, 6$ SCT hits for $p_{\mathrm{T}} \in$ 100-200, 200-300 and > 300 MeV, respectively
- ≥ 1 Pixel hits
- At least one hit in the B-layer if one is expected.
- $|d_0|\,, |z_0 \sin\theta| < 1.5$ mm, where d_0 and z_0 are constructed with respect to the primary vertex
- $|\sigma_{d_0}/d_0|\,, |\sigma_{z_0 \sin\theta}/z_0 \sin\theta| < 3$. Note that in constructing the uncertainty $\sigma_{z_0 \sin\theta}$, the full covariance matrix between z_0 and θ is used.

description	data set names	# runs
p+Pb, minimum bias	`data13_hip.0021*.physics_MinBias.merge.NTUP_HI.f*_m1312`	32
p+Pb, hard probes	`data13_hip.0021*.physics_HardProbes.merge.NTUP_HI.f*_m1312`	32
pp, minimum bias	`data13_2p76TeV.00219*.physics_MinBias.merge.NTUP_HI.f519_m1313`	6
pp, hard probes	`data13_2p76TeV.00219*.physics_HardProbes.merge.NTUP_HI.f519_m1313`	6

Table 8.1: Summary of data D3PD's used in the p+Pb and pp jet analysis.

8.2.3 p+Pb data and event selection

The data used in this analysis utilize the full 2013 p+Pb and Pb+p run statistics. p+Pb events in filled crossings are analyzed in lumiblocks when all detector components had good operational status. Specifically, v.60 of the Good Run List for p+Pb is used,

- `data13_hip.periodAllYear_DetStatus-v60-pro15_DQDefects-00-01-00_PHYS_HeavyIonP_All_Good.xml`

which comprises 32 runs (16 each in the period A and B orientations).

A handful of additional lumiblocks (approximately ten, in four different runs) are excluded because they lay on the edge of a series of GRL lumiblocks and had a large change in EF prescales relative to the other lumiblocks.

The minimum bias event definition in p+Pb collisions is the same as described in [263], and consists of:

- At least one hit in each minimum bias trigger scintillator (MBTS), with the two sides reporting an event time no more than 10 ns apart
- At least one reconstructed vertex, but no more than one reconstructed vertex with $\sum p_{\mathrm{T}} > 6$ GeV of tracks associated with it (to reject in-time p+Pb pileup events)
- At least two tracks in the event with $p_{\mathrm{T}} > 100$ MeV which pass minimum bias tracking cuts (see Section 8.2.2)
- A pseudorapidity gap smaller than $\Delta\eta^{\mathrm{Pb}} < 2.0$ on the Pb-going edge of the detector acceptance (to reject events consistent with diffractive excitations of the proton)

The minimum bias trigger consists of at least one hit in each MBTS (`MBTS_1_1`). The number of total p+Pb events sampled, used in Equation 8.2 is determined by the prescale-corrected set of recorded events which pass the above minimum bias event selection cuts. Formally,

$$N_{\mathrm{evt}} = \sum_{i \,\in\, \mathrm{MBTS\_1\_1}} C_i \tag{8.13}$$

where the sum is over recorded events meeting the conditions above and C_i is the full event filter prescale for the event. The prescale for the minimum bias trigger ranged from 2 in low-luminosity running to up to $\approx 10^4$ in the highest luminosity running. After event selection, the

prescale-corrected number of minimum-bias p+Pb crossings is $N_{\text{evt}} = 5.34 \times 10^{+10}$, corresponding to an integrated luminosity of approximately 25 nb^{-1}. Table 8.1 summarizes the dataset names.

The η conventions for the first period of p+Pb running are used, with $\eta < 0$ corresponding to the p-going, or "forward", and $\eta > 0$ corresponding to the Pb-going, or "backward", direction.

8.2.4 Centrality determination

The centrality in p+Pb collisions is measured using the sum of the transverse energy (at the EM scale) in the Pb-going forward hadronic calorimeter (FCal), $\Sigma E_{\text{T}}^{\text{Pb}}$, situated at $3.1 < \eta < 4.9$, and is described in detail in [263; 264]. The raw $\Sigma E_{\text{T}}^{\text{Pb}}$ is corrected for a small run-dependent electronic pedestal and an acceptance factor that changes with the vertex of the p+Pb event. The geometry of p+Pb collisions is determined using a Glauber model[68] with a nucleon-nucleon cross-section at $\sqrt{s} = 5.02$ TeV of 70 mb. Additionally, an alternative to the Glauber model (the "Glauber-Gribov" model) is considered in which the nucleon-nucleon cross-section fluctuates event to event[84; 82], resulting in a different underlying N_{part} distribution.

The minimum bias event selection efficiency for inelastic p+Pb collisions is determined to be 96-98% with a 0-2% residual contaminant from p+Pb events consistent with photonuclear or diffractive excitation of the Pb nucleus. The signal in the measured quantity $\Sigma E_{\text{T}}^{\text{Pb}}$ is assumed to vary with N_{part} according to several different models, the simplest one of which uses a PYTHIA 8 reference for the $\Sigma E_{\text{T}}^{\text{Pb}}$ distribution in inelastic pp collisions at $\sqrt{s} = 5.02$ and posits that the contribution to $\Sigma E_{\text{T}}^{\text{Pb}}$ from p+Pb events with a given number of participants N_{part} is equal to the pp signal convoluted $N_{\text{part}}/2$ times. Thus, selecting events with a given range of $\Sigma E_{\text{T}}^{\text{Pb}}$ is equivalent to selecting p+Pb events with a given mean N_{part}.

A similar determination of the mean number of participants using the E_{T} in both FCals has been performed in Pb+Pb collisions in ATLAS[249; 265].

In this analysis, the mean nuclear thickness $\langle T_{AB} \rangle$ and the ratio of the number of binary collisions between centrality bins R_{coll} is estimated using the default Glauber model (and not either of the Glauber-Gribov variants). Table 8.2 summarizes the centrality selections, $\langle T_{AB} \rangle$ and R_{coll} $\left(= {N_{\text{coll}}}^{\text{cent}} / {N_{\text{coll}}}^{60-90\%} \right)$ values, and the total number of prescale corrected minimum bias events N_{evt}, for the seven centrality selections used in this analysis.

There is one additional correction made to the centrality determination. The mapping of events

centrality	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ range (GeV)	$\langle T_{AB} \rangle$ (mb^{-1})	R_{coll}	N_{evt}
0-10%	> 53.7	$0.208^{+0.019}_{-0.005}$	$4.89^{+0.83}_{-0.27}$	$5.45 \times 10^{+9}$
10-20%	40.0-53.7	$0.172^{+0.007}_{-0.003}$	$4.05^{+0.49}_{-0.21}$	$5.45 \times 10^{+9}$
20-30%	31.0-40.0	$0.148^{+0.004}_{-0.002}$	$3.48^{+0.33}_{-0.18}$	$5.45 \times 10^{+9}$
30-40%	24.1-31.0	$0.126^{+0.003}_{-0.004}$	$2.96^{+0.21}_{-0.14}$	$5.45 \times 10^{+9}$
40-60%	13.4-24.1	$0.092^{+0.004}_{-0.006}$	$2.16^{+0.08}_{-0.07}$	$1.09 \times 10^{+10}$
60-90%	2.56-13.4	$0.043^{+0.003}_{-0.004}$	1	$1.64 \times 10^{+10}$
0-90%	> 2.56	$0.107^{+0.007}_{-0.006}$	n/a	$4.91 \times 10^{+10}$

Table 8.2: Summary of centrality bins used in the p+Pb jet analysis.

with a given centrality to a class of p+Pb geometries with a mean number of nucleon-nucleon collisions rests on the assumption that energy in the centrality-determining detector comes from soft particles in the p+Pb underlying event. Thus, the presence of a high-energy jet can bias the determination of an event's centrality. In particular, measuring centrality-selected jet yields in or near the region of the Pb-going FCal can result in strong autocorrelations between the jet yield and the centrality of the event.

In this analysis, the contribution to $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ from jets is subtracted out by examining R=0.4 jets. Reconstructed jets with an E_{T} above 8 GeV at the electromagnetic (EM) scale which contain towers within the Pb-going FCal, have the energy in those towers subtracted from the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ in the event according to

$$\left(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\right)^{\mathrm{corr}} = \left(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}\right)^{\mathrm{uncorr}} - \sum_{\mathrm{jets}\ \eta>2.8} {E_{\mathrm{T}}}^{\mathrm{EM,FCal}} \cdot \theta(E_{\mathrm{T}}^{\mathrm{EM}} - 8\ \mathrm{GeV}) \tag{8.14}$$

where $E_{\mathrm{T}}^{\mathrm{EM}}$ is the electromagnetic-scale energy of the jet, ${E_{\mathrm{T}}}^{\mathrm{EM,FCal}}$ is the subset of that energy in FCal cells and $\eta > 2.8$ selects (R=0.4) jets which overlap with the Pb-going FCal. Since the HI reconstructed jets have underlying event subtraction, the subtraction of the jet energy from $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ preserves the underlying event pedestal under the jet. The choice of $E_{\mathrm{T}}^{\mathrm{EM}} > 8$ GeV was made by comparing the jet yield in the FCal in p+Pb data, which contains jets from hard scatterings as well as underlying event fluctuations, to the truth-matched yield in MC, which only contains the

former. > 8 GeV was chosen so that jets from UE fluctuations would not be subtracted from the FCal, which would incorrectly bias the minimum bias FCal distribution. Additionally, the cut is varied from > 6 GeV to > 10 GeV to determine the systematic uncertainty.

8.2.5 *pp* data and event selection

The data analyzed in the 2013 *pp* 2.76 TeV running is selected from v.60 of the GRL for the period,

- `data13_2p76TeV.periodAllYear_DetStatus-v60-pro15_DQDefects-00-01-00_PHYS_HeavyIonP_All_Good.xml`

Since in *pp* collisions the goal is to measure the jet-cross section (instead of the per-event yield, as in p+Pb collisions), the event selection cuts are kept to a minimum, requiring only a reconstructed vertex. Table 8.1 summarizes the dataset names.

8.3 Jet Selection and Reconstruction

8.3.1 HI Jet Reconstruction

The jet reconstruction procedure in p+Pb collisions is similar to that used in ATLAS Pb+Pb jet results[59; 60; 266; 267], and is designed to function in a heavy ion environment with a large, fluctuating underlying event. The main details of the underlying event subtraction are summarized here. Further detail on the reconstruction procedure is available at [268; 269; 270].

The main jet reconstruction procedure consists of seven steps: (1) a preliminary determination of *seeds*, areas consistent with jets which should be excluded from the determination of the underlying event background, (2) a preliminary determination and subtraction of the underlying event background, (3) a more sophisticated determination of the seeds, (4) a new background determination based on the updated seed positions, (5) jet finding, (6) correction of resulting jet energies from residual misalignment with seeds, and (7) correction of the electromagnetic (EM) scale jet energy to the final jet energy scale. These steps are described below.

1. First, regions of the calorimeter with a high E_{T} consistent with jet production must be excluded from the background determination. Failure to properly exclude jets from the background determination results will lead to an oversubtraction of energy under the jet. This effect is called the *self-energy bias* (SEB).

Calorimeter cells at the electromagnetic scale are treated as four vectors with E_{T}, η and ϕ given by the cell energy and position (and $m = 0$). The cells are grouped through the addition of their four-vectors into $\Delta\eta = 0.1 \times \Delta\phi = 0.1$ sized projective calorimeter *towers*. Cells with a large granularity which would span multiple towers have their energy divided equally between the towers in question. anti-k_{T} jet reconstruction with $R = 0.2$ is run inclusively on the 64×98 grid of towers. The seeds are taken to be jets in which the highest-$E_{\mathrm{T}}^{\mathrm{EM}}$ tower has $E_{\mathrm{T}}^{\mathrm{EM}} > 3$ GeV and for which the discriminant D, which is the ratio of the $E_{\mathrm{T}}^{\mathrm{EM}}$ in the highest-$E_{\mathrm{T}}^{\mathrm{EM}}$ tower to the mean tower $E_{\mathrm{T}}^{\mathrm{EM}}$, is

$$D = \frac{\max\left(E_{\mathrm{T}}^{\mathrm{EM}}\right)}{\left\langle E_{\mathrm{T}}^{\mathrm{EM}} \right\rangle} > 4 \tag{8.15}$$

This preliminary set of seed jets is used in the next stage.

2. The EM scale energy measured in each calorimeter cell is modeled as the sum of contributions from jet energy and the underlying p+Pb event via

$$\frac{dE_{\mathrm{T}}}{d\eta d\phi}^{\mathrm{total}} = \frac{dE_{\mathrm{T}}}{d\eta d\phi}^{\mathrm{UE}} + \frac{dE_{\mathrm{T}}}{d\eta d\phi}^{\mathrm{jet}} \tag{8.16}$$

In p+Pb collisions, where no significant modulation of the underlying event in ϕ is expected, the contribution from the UE is modeled as a rapidity-dependent energy density,

$$\frac{dE_{\mathrm{T}}}{d\eta d\phi}^{\mathrm{UE}} = \rho(\eta) \tag{8.17}$$

$\rho(\eta)$ is determined separately in each calorimeter layer and in $\Delta\eta = 0.1$ strips from the average E_{T} in the cells which are not part of a seed jet,

$$\rho(\eta) = \frac{1}{N} \sum_{i \in N} E_{\mathrm{T}i} / \Delta\phi\Delta\eta \tag{8.18}$$

where the sum is over the N cells in that layer and $\Delta\eta = 0.1$ strip which were had not been clustered into a seed jet in step #1.

The cell energies are adjusted by subtracting this estimate of the underlying event via,

$$E_{\mathrm{T}}{}^{\mathrm{subtracted}} = E_{\mathrm{T}}{}^{\mathrm{total}} - \rho(\eta)\Delta\phi\Delta\eta \tag{8.19}$$

where $\rho(\eta)$ also depends on the calorimeter layer.

3. The background subtracted cells are again grouped into towers, and jet reconstruction is rerun with $R = 0.2$. The resulting jets with $E_{\mathrm{T}}^{\mathrm{EM}} > 4$ GeV are taken to be the new seed jets, replacing the preliminary set of seeds determined in step #1.

4. The layer-by-layer background in each η strip is determined again as in step #2, but this time using the updated seeds to exclude cells from the background determination.

5. Jet reconstruction is run for the third time on the second iteration of the background subtracted towers, this time with the resolution parameter $R = 0.2, 0.3, 0.4, 0.5$ and 0.6. With the exception of one additional correction applied in step #6, the resulting jets are the final set of jets which are used in the analysis.

6. In this step, any residual misalignment between the seeds in step #3 and the final set of jets in step #5 is corrected, in a procedure called the *self-energy bias correction*. The procedure considers the full set of reconstructed jet with $E_{\mathrm{T}}^{\mathrm{EM}} > 8$ GeV and the original seed jets. In each η strip, the effect on $\rho(\eta)$ of excluding jet towers not originally excluded by the seeds is computed, and any over-subtracted energy is added back to the jet. For jets which are well correlated with a seed jet, the effect of the SEB correction is negligible. For jets which have no overlap with the seeds, the correction is on the order of 10% for $R = 0.4$ jets (a geometric quantity which is derived from $\pi R^2/(2R)(2\pi) \approx 0.1$). Due to the low threshold on the E_{T} of the seed jets (> 4 GeV, see step #3), the SEB correction has only a minor effect on jets used in this analysis, since all reconstructed jets with $E_{\mathrm{T}} > 15$ or so are very well correlated with a seed jet.

7. Finally, the self-energy bias corrected jets with EM-scale energy $p_{\mathrm{T}}^{\mathrm{EM}}$ are converted to a full jet energy p_{T} through a multiplicative correction for the detector response derived from the *numerical inversion* procedure, which is described separately in Section 8.3.2.

In the Pb+Pb jet analysis, there is normally another step known as *fake jet rejection* which is designed to reject jets from upward fluctuations of underlying event energy ("fake jets"). In

this step, jets are required to be angularly associated with a charged track jet or a high-energy calorimeter cluster. In the p+Pb data, where the fake rate is substantially lower, this step is not applied.

The present results use the $R = 0.4$ and $R = 0.6$ jet collections. The jet reconstruction performance in ATLAS is detailed in Section 8.5.

8.3.2 Energy Scale Calibration

Jets at the EM are scale are corrected to the full hadronic scale through a procedure called numerical inversion (NI), which closely follows the procedure used to calibrate jets in 7 TeV pp [271].

The NI procedure uses the 5.02 TeV Signal MC (see Section 8.5) to derive a correction factor which converts the jet EM-scale transverse momentum $p_{\mathrm{T}}^{\mathrm{EM}}$ to a full jet momentum p_{T}. The correction is η dependent such that $p_{\mathrm{T}} = p_{\mathrm{T}}\left(p_{\mathrm{T}}^{\mathrm{EM}}; \eta\right)$. Jets at a given $p_{\mathrm{T}}^{\mathrm{EM}}$ are corrected to an energy p_{T} which is equal to the mean $p_{\mathrm{T}}^{\mathrm{truth}}$ of truth jets which are reconstructed at the given EM scale momentum.

Specifically, for each $p_{\mathrm{T}}^{\mathrm{truth}}$ bin, the mean EM-scale reconstructed $\left\langle p_{\mathrm{T}}^{\mathrm{EM}} \right\rangle (p_{\mathrm{T}}^{\mathrm{truth}})$ is determined. Then, then mean response $\left(= \dfrac{p_{\mathrm{T}}^{\mathrm{EM}}}{p_{\mathrm{T}}} \right)$ is determined as a function of the truth p_{T}, $\left\langle \dfrac{p_{\mathrm{T}}^{\mathrm{EM}}}{p_{\mathrm{T}}} \right\rangle (p_{\mathrm{T}}^{\mathrm{truth}})$. The mean response and uncalibrated momentum at each truth value are associated with each other, and the response as a function of $p_{\mathrm{T}}^{\mathrm{EM}}$ is parameterized by fitting polynomial up to fourth order in log-$p_{\mathrm{T}}^{\mathrm{EM}}$ space.

Thus, the fit interpolates between the $(p_{\mathrm{T}}^{\mathrm{EM}}, \dfrac{p_{\mathrm{T}}^{\mathrm{EM}}}{p_{\mathrm{T}}})$ points and results in the response as a function of the uncalibrated (not truth) transverse momentum,

$$Response(p_{\mathrm{T}}^{\mathrm{EM}}; \eta) = \left\langle \frac{p_{\mathrm{T}}^{\mathrm{EM}}}{p_{\mathrm{T}}} \right\rangle \left(\left\langle p_{\mathrm{T}}^{\mathrm{EM}} \right\rangle \right) \tag{8.20}$$

Then, the reconstructed jet p_{T} is given by

$$p_{\mathrm{T}}^{\mathrm{reco}} = p_{\mathrm{T}}^{\mathrm{EM}} / Response(p_{\mathrm{T}}^{\mathrm{EM}}; \eta) \tag{8.21}$$

The residual closure of the numerical inversion is determined in MC by measuring the mean fractional difference between the reconstructed and truth p_{T} as a function of the truth p_{T}, $\left\langle (p_{\mathrm{T}}^{\mathrm{truth}} - p_{\mathrm{T}}^{\mathrm{reco}})/p_{\mathrm{T}}^{\mathrm{truth}} \right\rangle \left(p_{\mathrm{T}}^{\mathrm{truth}}\right)$. The energy scale is evaluated fully in Section 8.5.3.

8.3.3 Independence of subtracted E_{T}

Since the underlying event subtraction procedure is designed to remove the soft pedestal which is otherwise not associated with the production of the jet, an important test of the procedure is the independence of the subtracted energy from the jet energy. Here, the subtracted energy is defined as the sum of the EM-scale E_{T} subtracted from each of the constituent towers in the reconstructed jet.

Figure 8.2 shows the mean subtracted energy as a function of jet EM-scale E_{T} and centrality selection, in selected rapidity bins, for R=0.4 and R=0.6 jets. As expected, the mean subtracted energy increases systematically from peripheral to central events. The mean subtracted energy is observed to have a small increase with jet energy. This effect is strongest near mid-rapidity, where it is observed to increase ≈ 200 MeV from $E_{\mathrm{T}}^{\mathrm{EM}} = 20$ to $E_{\mathrm{T}}^{\mathrm{EM}} = 100$ GeV. Furthermore, this effect is present in all centrality bins. Thus, this effect is consistent with a residual self-energy bias effect, in which a small fraction of jet energy falls outside the nominal anti-k_{T} cone, and is thus included in the background determination. As the jet energy increases, the fraction of the jet's energy that this effect corresponds to decreases. Conservatively, a residual over-subtracted of 200 MeV for $E_{\mathrm{T}}^{\mathrm{EM}} = 100$ GeV jets is consistent with a $< 0.2\%$ effect on the final energy scale (since the final jet energy $E_{\mathrm{T}} > E_{\mathrm{T}}^{\mathrm{EM}}$).

8.3.4 Cleaning cuts

To reject reconstructed jets consistent with background, jets are required to pass "BadMedium" (which rejects jets consistent with noise spikes in the HEC, coherent noise in the EM calorimeter and out-of-time energy deposits from cosmic rays and beam backgrounds) and "Ugly" (which rejects jets depositing a majority of their energy in the Tile Barrel to Tile Extension transition region or in dead cells, since the total energy is not well measured in these cases) quality cuts. The efficiency of this selection is estimated in Section 8.3.5 and the measured jet yields are corrected for it.

Figure 8.3 shows the fraction of R=0.4 reconstructed jets in data rejected by each of the two cuts as a function of p_{T} and η. As expected, the BadMedium cut mostly affects jets in the EM Barrel and HEC regions, but also several jets in the FCal that are consistent with beam background and cosmic rays. Also as expected, the Ugly cut rejects jets in or near the Tile transition region. Although Figure 8.3 shows several high-p_{T} bins regions where the rejected fraction is 100%, these

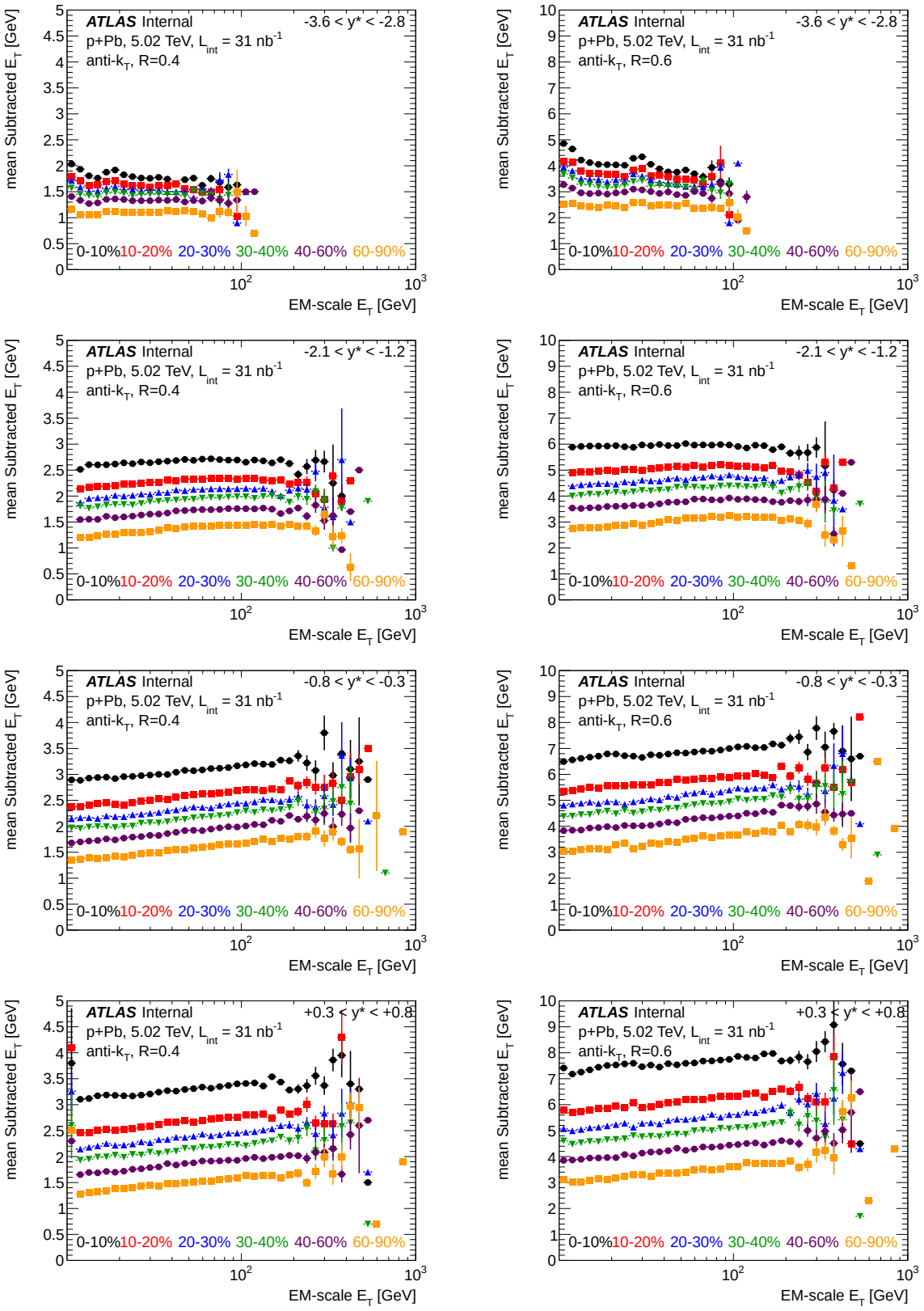


Figure 8.2: Mean subtracted E_{T} vs. EM-scale reconstructed jet E_{T} in selected rapidity bins for R=0.4 jets (left) and R=0.6 jets (right), with jets in different centrality selections shown in different markers.

are typically bins with only one jet in them far past the statistical reach for real jets.

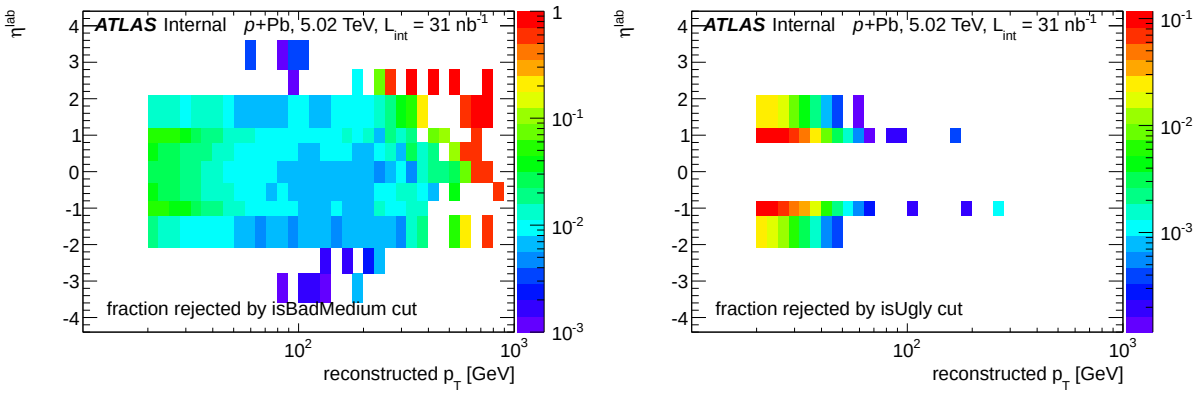


Figure 8.3: Fraction of jets rejected by BadMedium (left) and Ugly (right) cuts, for R=0.4 jets as a function of p_{T} and η.

Figures 8.4 and 8.5 show the cleaning cut survival fraction in data as a function of p_{T} in separate η bins, for R=0.4 and R=0.6 jets. The cleaning cuts have minimal effect on jets within $\left|\eta^{\mathrm{lab}}\right| > 2.1$, where the survival fraction is $> 99.5\%$ in all p_{T} bins. For jets within $\left|\eta^{\mathrm{lab}}\right| < 2.1$, the survival probability is typically smallest at low-p_{T}, but then saturates at high-p_{T}. For R=0.4 (R=0.6) jets the smallest survival probability is 93% (86%) for $p_{\mathrm{T}} = 20$ Gev jets in the $0.8 < \left|\eta^{\mathrm{lab}}\right| < 1.2$ bin, and $> 98\%$ for $p_{\mathrm{T}} > 40$ GeV jets in all bins.

8.3.5 Data-driven efficiency for real jets

The efficiency for real jets passing the cleaning cuts is determined in a data-driven way using a tag & probe method similar to what is done in [272]. The leading (highest-p_{T}) and subleading (next-highest-p_{T}) jet in p+Pb events are analyzed without cleaning cuts. We consider events where the leading and subleading jet are well-balanced according to their angular correlation and asymmetry,

$$\Delta\phi = \phi^{\mathrm{subleading}} - \phi^{\mathrm{subleading}} > 2.6 \tag{8.22}$$

$$A = \frac{{p_{\mathrm{T}}}^{\mathrm{subleading}} - {p_{\mathrm{T}}}^{\mathrm{subleading}}}{\frac{1}{2}({p_{\mathrm{T}}}^{\mathrm{subleading}} + {p_{\mathrm{T}}}^{\mathrm{subleading}})} < 0.4 \tag{8.23}$$

If the leading jet is within $\left|\eta^{\mathrm{lab}}\right| < 2$ and passes a stricter set of cleaning cuts, "BadTight" and "Ugly", then this jet is "tagged" and the event included in the *good jet sample*. Using the

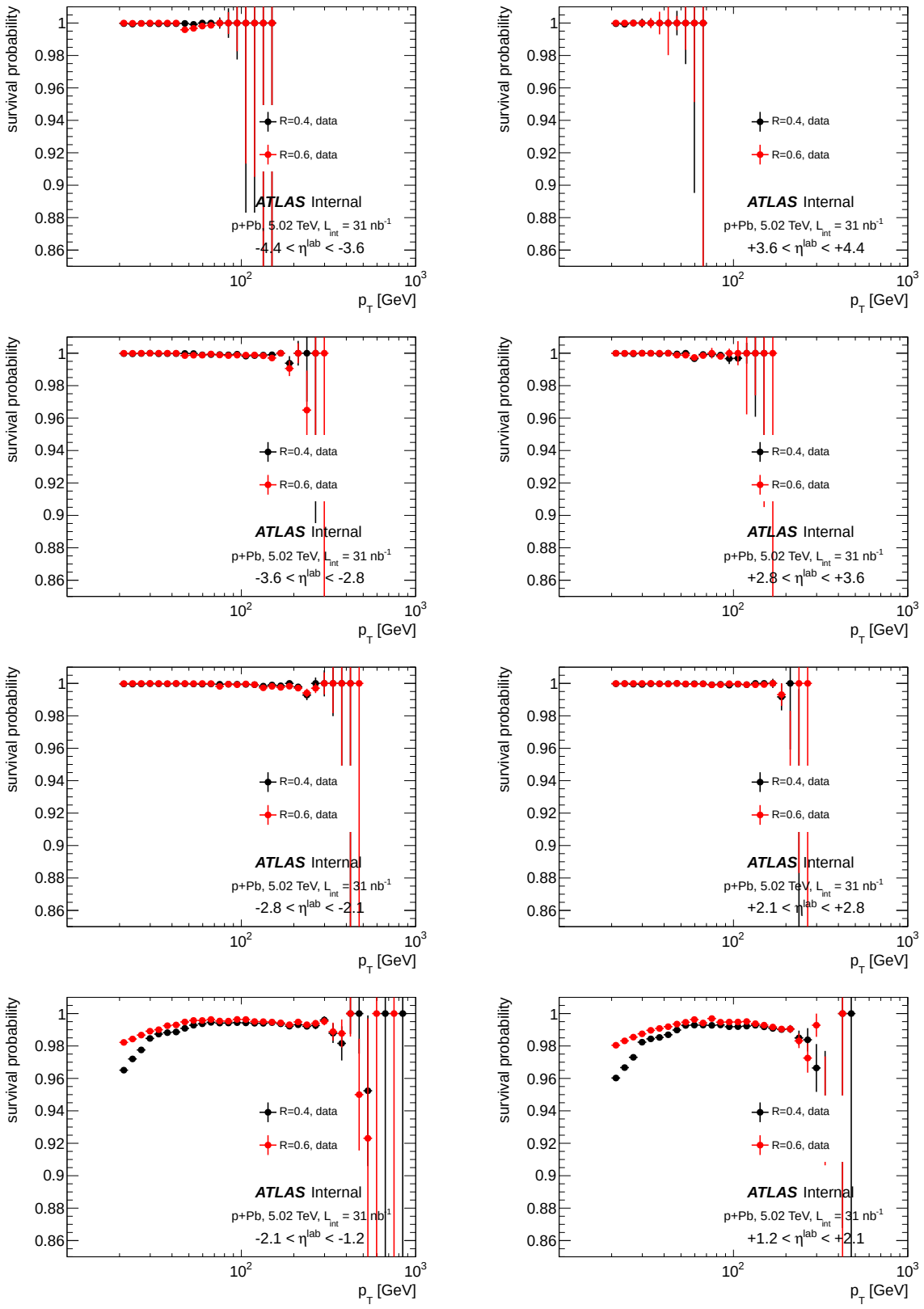


Figure 8.4: Fraction of jets surviving cleaning cuts in data, as a function of p_{T}. Each plot shows a different pseudorapidity range, with R=0.4 and R=0.6 jets shown separately. This figure shows the range $\left|\eta^{\mathrm{lab}}\right| > 1.2$.

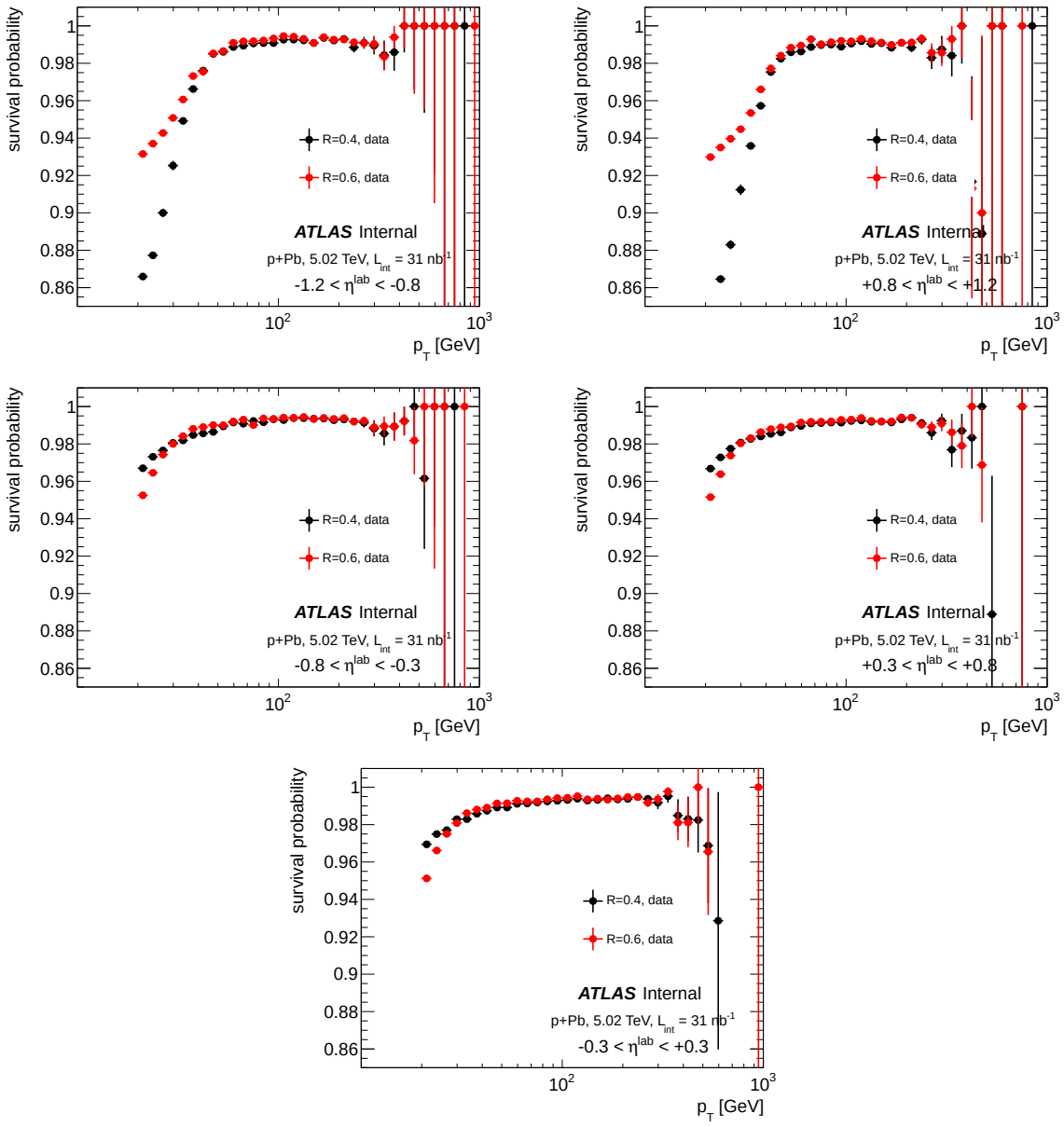


Figure 8.5: Fraction of jets surviving cleaning cuts in data, as a function of p_{T}. Each plot shows a different pseudorapidity range, with R=0.4 and R=0.6 jets shown separately. This figure shows the range $\left|\eta^{\mathrm{lab}}\right| < 1.2$

good jet sample, the cleaning cut efficiency is determined by examining the subleading jet as a function of η^{lab} and p_{T}. To smooth the statistical fluctuations, the data are fit to a function, which is phenomenologically chosen to be

$$\epsilon_{\text{cleaning}}(p_{\text{T}};\eta) = c_0 + c_1 \cdot \text{erf}\left(c_2 \cdot p_{\text{T}} + c_3\right) \tag{8.24}$$

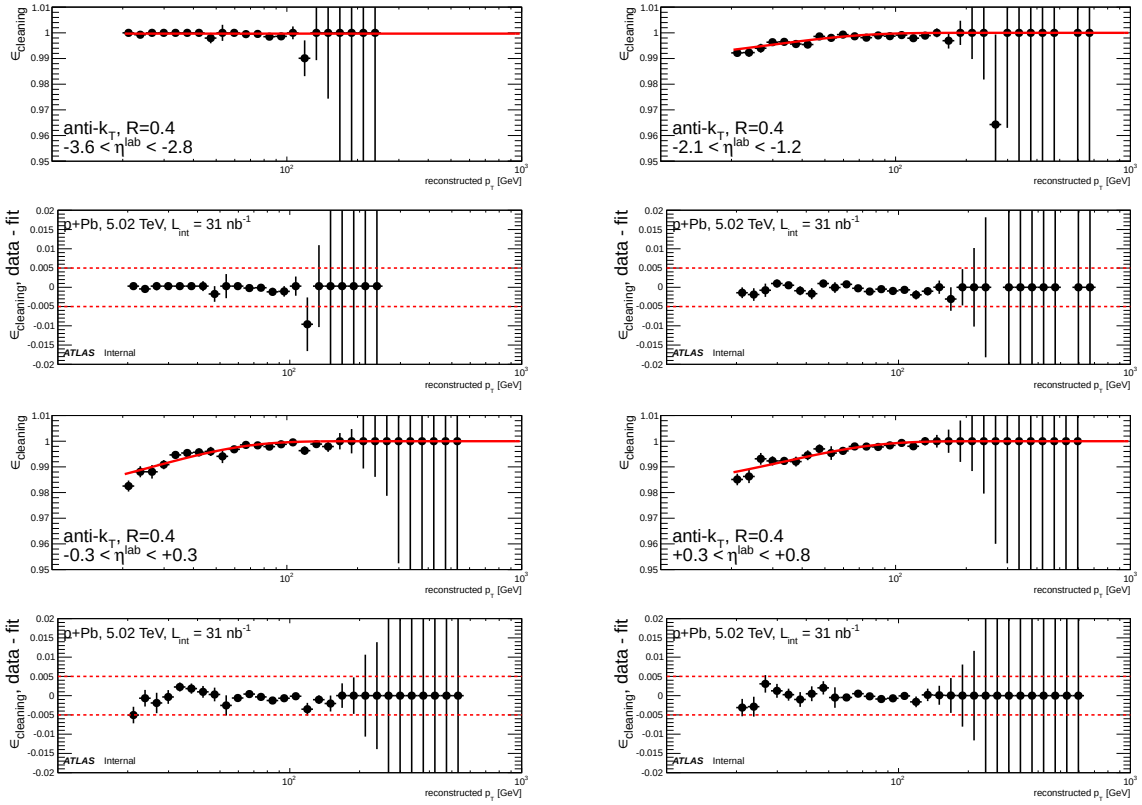


Figure 8.6: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}. Select pseudorapidity bins are shown for R=0.4 jets.

The fit function in Equation 8.24 was chosen because of its natural asymptotic behavior (at high p_{T}, $\epsilon \to c_0 + c_1$). Figures 8.6 shows the selection efficiency in the good jet sample in a selected sample of η^{lab} bins, along with the fits to the data. The bottom panel on each figure shows the data minus fit difference, with dotted red line indicating the $\pm 0.5\%$ difference. The fits to a good job of describing the data. In all cases, the data minus fit residuals are statistically consistent with lying inside the $\pm 0.5\%$ range. The full set of data-driven cleaning efficiencies in all η^{lab} bins, both

cone sizes and for the p+Pb and pp data are shown Figures B.1, B.2, B.3, B.4, B.5 and B.6 in Appendix B.1.

Furthermore, the procedure was repeated using looser and stricter selection criteria on the dijet pair,

- stricter dijet requirement, $\Delta\phi > 2.8, A < 0.3$
- looser dijet requirement, $\Delta\phi > 2.4, A < 0.5$

and the observed changes on the fit function are $< 0.1\%$ in all bins and at all p_{T}.

The data are corrected for the cleaning selection inefficiency by evaluating the fit function at the p_{T} value of each jet. Note that the cleaning efficiency correction is applied as a function of jet η^{lab}, whereas the centrality-dependent yields are constructed in bins of y^*. 0.5% is taken to be the systematic uncertainty on the cleaning efficiency correction.

8.4 Trigger Selection

Due to the high luminosity in the p+Pb data-taking, a trigger menu for jets with many EF trigger jet thresholds was used, with triggers for higher p_{T} jets having sequentially smaller prescales. To maximize the available statistics at every p_{T}, the per-event jet yields are constructed from eight trigger streams, with each p_{T} bin populated by one only trigger. All EF trigger jets used the R=0.4 anti-k_{T} algorithm, but used either a hadronic topocluster ("a4tchad") reconstruction without underlying event subtraction or heavy-ion subtraction and reconstruction ("a4hi"). The performance of the HI trigger jets has been previously evaluated in [273].

Towards the latter half of period A running, the trigger jet type was switched from a4tchad to a4hi, but in diffrent runs for the intermediate (j20-j40) triggers than the high-p_{T} (j50-j75) triggers. However, since the ultimate trigger scheme chosen only utilizes triggers where they are $> 99\%$ efficient, the a4tchad and a4hi versions of a given triggers can be considered interchangeable. To construct the per-event yield of low-p_{T} jets for which the j20 trigger had not yet saturated, MB events were used.

Table 8.3 gives an overview of the triggers used in p+Pb, the runs in which they applied, and the total number of triggered events sampled. R=0.4 and R=0.6 offline jets were selected by requiring

EF jet threshold	trigger	used in runs 217999-8222	used in runs 218301-38	used in runs 218391-9114	approx. $N_{\text{evt}}{}^{\text{trigger}}$
MB	EF_mbMbts_1_1	✓	✓	✓	4.3×10^7
j20	EF_j20_a4tchad_EFFS_L1MBTS	✓			2.2×10^6
	EF_j20_a4hi_EFFS_L1MBTS		✓		4.4×10^5
	EF_j20_a4hi_EFFS_L2mbMbts			✓	2.0×10^6
j30	EF_j30_a4tchad_EFFS_L1MBTS	✓			1.1×10^6
	EF_j30_a4hi_EFFS_L1MBTS		✓		8.8×10^4
	EF_j30_a4hi_EFFS_L2mbMbts			✓	1.5×10^6
j40	EF_j40_a4tchad_EFFS_L1MBTS	✓			4.1×10^5
	EF_j40_a4hi_EFFS_L1MBTS		✓		2.8×10^4
	EF_j40_a4hi_EFFS_L2mbMbts			✓	5.3×10^5
j50	EF_j50_a4tchad_EFFS_L1J10	✓	✓		9.9×10^5
	EF_j50_a4hi_EFFS_L1J10			✓	2.1×10^6
j60	EF_j60_a4tchad_EFFS_L1J15	✓	✓		1.2×10^6
	EF_j60_a4hi_EFFS_L1J15			✓	8.8×10^5
j75	EF_j75_a4tchad_EFFS_L1J15	✓	✓		4.7×10^6
	EF_j75_a4hi_EFFS_L1J15			✓	2.1×10^6
fj15	EF_fj15_a4tchad_EFFS_L1FJ0	✓			8.3×10^5
	EF_fj15_a4hi_EFFS_L1FJ0		✓	✓	1.4×10^6

Table 8.3: Overview of triggers used in the p+Pb data, showing the groups of runs in which they are used and the total number of sampled events.

a trigger jet within $\Delta R < 0.4$ of the R=0.4 offline jet. The highest trigger threshold used is the j75 trigger. Although j80-j100 triggers are sometimes included in the trigger menu, j75 is the first fully unprescaled trigger so they are not needed.

Section 8.4.1 describes the determination of the trigger efficiency, Section 8.4.2 describes the trigger scheme chosen including the jet p_{T} and η^{lab} ranges for which different triggers were used, and Section 8.4.4 investigates the stability of the trigger-constructed yields.

8.4.1 Trigger efficiency

The trigger efficiency $\epsilon^{\text{trigger}}$ is determined in data using MB jets in $\left|\eta^{\text{lab}}\right| < 3.2$ for the j20, j30, j40 triggers, MB jets with $\left|\eta^{\text{lab}}\right| > 3.2$ for the fj15 trigger and jets matched to j20 triggers in $\left|\eta^{\text{lab}}\right| < 3.2$

EF jet threshold	η	p_{T} (R=0.4 jets)	p_{T} (R=0.6 jets)
MB	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	< 31.6 GeV	< 44.7 GeV
	$\lvert\eta^{\mathrm{lab}}\rvert > 3.2$	< 31.6 GeV	< 39.8 GeV
j20	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	31.6-39.8 GeV	44.7-56.2 GeV
j30	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	39.8-50.1 GeV	56.2-70.8 GeV
j40	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	50.1-63.1 GeV	70.8-89.1 GeV
j50	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	63.1-79.4 GeV	89.1-100 GeV
j60	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	79.4-100 GeV	100-125 GeV
j75	$\lvert\eta^{\mathrm{lab}}\rvert < 3.2$	> 100 GeV	> 125 GeV
fj15	$\lvert\eta^{\mathrm{lab}}\rvert > 3.2$	> 31.6 GeV	> 39.8 GeV

Table 8.4: Trigger scheme in the p+Pb data, listing which triggers are used in the p_{T}, y bins.

trigger	p_{T}, y range R=0.4	p_{T}, y range R=0.6	$\mathcal{L}_{int}$ $[nb^{-1}]$
EF_mbSpTrk	$p_{\mathrm{T}} < 20$, $\lvert y\rvert < 3.2$ $p_{\mathrm{T}} < 22$, $\lvert y\rvert > 3.2$	$p_{\mathrm{T}} < 20$, $\lvert y\rvert < 3.2$ $p_{\mathrm{T}} < 32$, $\lvert y\rvert > 3.2$	1.05858 / 8
EF_j10_a4tchad_EFFS_L2mbMbts_L1RD0	$p_{\mathrm{T}} \in$ 20-32, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} \in$ 20-40, $\lvert y\rvert < 3.2$	11.7184 / 8
EF_j20_a4tchad_EFFS_L2mbMbts_L1RD0	$p_{\mathrm{T}} \in$ 32-50, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} \in$ 40-70, $\lvert y\rvert < 3.2$	12.9786 / 8
EF_j40_a4tchad_EFFS_L1J5	$p_{\mathrm{T}} \in$ 50-65, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} \in$ 70-79, $\lvert y\rvert < 3.2$	288.586
EF_j50_a4tchad_EFFS_L1J10	$p_{\mathrm{T}} \in$ 65-79, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} \in$ 79-100, $\lvert y\rvert < 3.2$	832.942
EF_j60_a4tchad_EFFS_L1J15	$p_{\mathrm{T}} \in$ 79-89, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} \in$ 100-125, $\lvert y\rvert < 3.2$	1659.44
EF_j75_a4tchad_EFFS_L1J15	$p_{\mathrm{T}} > 89$, $\lvert y\rvert < 3.2$	$p_{\mathrm{T}} > 125$, $\lvert y\rvert < 3.2$	4095.26
EF_fj10_a4tchad_EFFS_L2mbMbts_L1RD0	$p_{\mathrm{T}} > 22$, $\lvert y\rvert > 3.2$	$p_{\mathrm{T}} > 32$, $\lvert y\rvert > 3.2$	12.9786 / 8

Table 8.5: Trigger scheme in the pp data, listing which triggers were used in the p_{T}, y bins, and the total luminosity sampled.

for j50, j60 and j75 triggers. Thus, while for the j20-j40 and fj15 triggers, $\epsilon^{\text{trigger}}$ is the efficiency with respect to `EF_mbMbts_1_1` triggers, for j50-j75 triggers the $\epsilon^{\text{trigger}}$ is the efficiency with respect to `EF_j20` triggers. However, the efficiency for j50-j75 triggers turns on in a region where the j20 trigger has completely saturated ($\epsilon = 100\%$) with respect to the MB trigger.

Each offline jet is tested for firing the trigger by examining if it is within $\Delta R < 0.4$ of a trigger jet above the threshold (${p_{\text{T}}}^{\text{trigger}} > 20$ for j20 triggers, etc.) in question. In this way, the trigger condition used in determining the efficiency is the same as is that used to select jets in the offline analysis. Because the behavior of the turnon curve is observed to be underlying-event dependent, the efficiency is investigated as a function of η^{lab} and centrality for each cone size (R=0.4 and R=0.6). The efficiency is determined separately for online `a4hi` and `a4tchad` jets, but the reference offline jets are always HI jets, since these are what are used in the final results.

Figure 8.7 shows the η-integrated and centrality-blind, trigger efficiency for each of the 7 triggers. Efficiency curves are shown for both trigger jet variants (`a4tchad` and `a4hi`) and for R=0.4 and R=0.6 offline jets. Some general features of the efficiency are as follows: for low-p_{T} triggers (the j20-j40 and fj15 triggers), the efficiency for `a4tchad` trigger jets turns on earlier than the `a4hi` trigger jets. This is because, while the p_{T}-feeding from the underlying event pushes up the p_{T} of the `a4tchad` trigger jet, this effect is mitigated by the underlying event subtraction in the `a4hi` jets. Furthermore, the turn on curve for R=0.6 offline jets is substantially wider than it is for R=0.4 offline jets, due to additional resolution and relative energy scale difference between R=0.6 offline jets and R=0.4 trigger jets, which is not there when offline and trigger jets have the same cone size. Finally, in the forward region, the fj15 trigger efficiency does not appear to saturate until $p_{\text{T}} \approx 30$ GeV. This is due to the significant change in the jet energy calibration in the forward region between the trigger and offline jets.

Finally, it is important to determine any centrality- or η dependence in the trigger efficiencies. Figure 8.8 plots $1 - \epsilon$ in η bins for j20, j50 and fj15 triggers and both types of trigger jets in the different η bins. It can be seen that the η dependence of the trigger turn-on curve is stronger in `a4tchad` jets (where the underlying event is playing a role) than in `a4hi` jets and stronger in low-p_{T} than in high-p_{T} trigger thresholds. Figures B.7, B.8, B.9, B.10 and B.11 in the Appendix plot the full centrality dependence of all triggers, showing that there are modest centrality-dependent differences in the turn-on curve, especially in the low-p_{T} trigger thresholds.

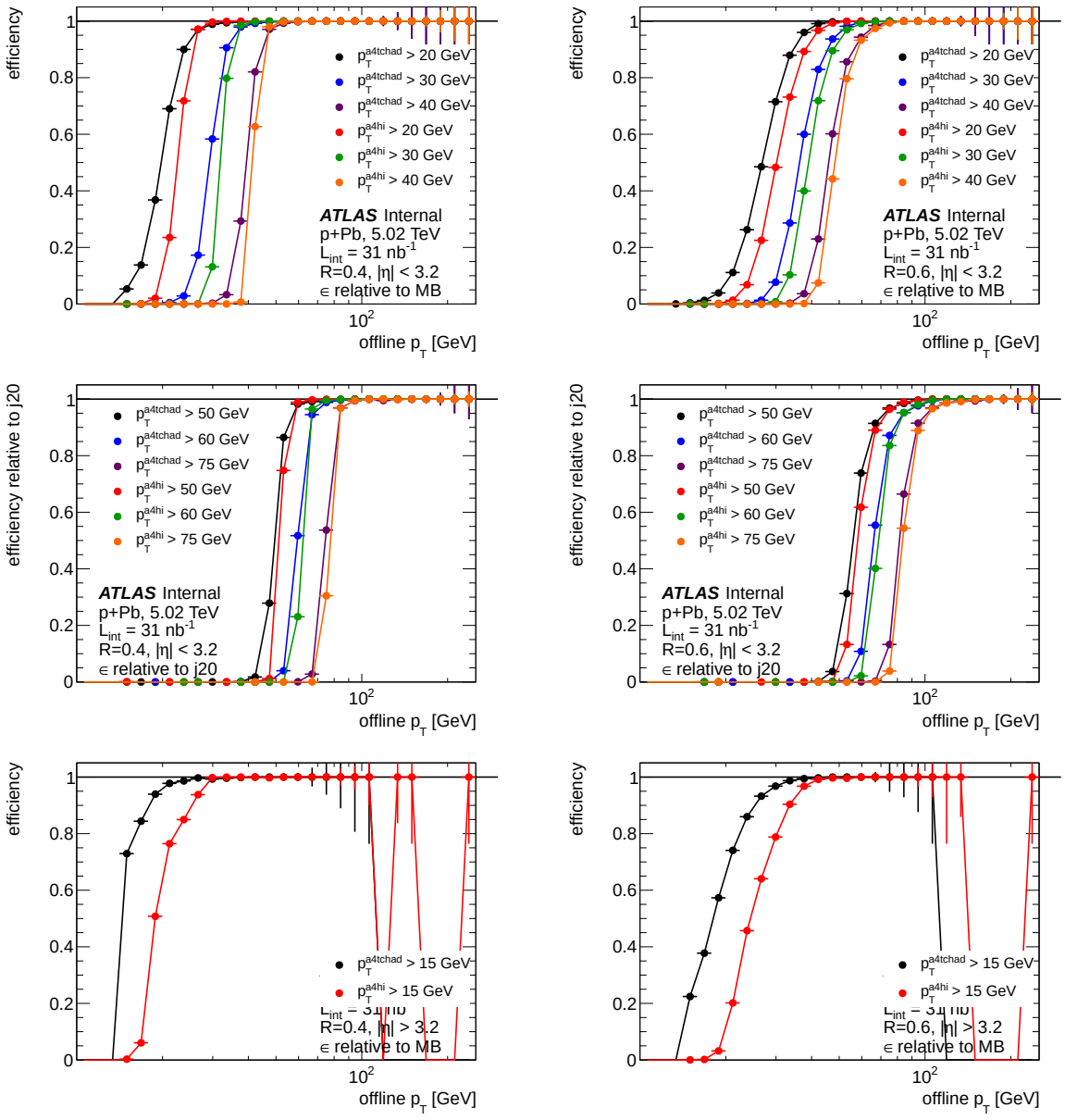


Figure 8.7: Trigger efficiencies of offline jets for the seven EF jet triggers used in this work, showing j20, j30 and j40 triggers (top row), j50, j60 and j75 triggers (middle row) and fj15 triggers (bottom row). `a4tchad` triggers are shown as open squares and `a4hi` triggers are shown as filled circles. Results for R=0.4 (left column) and R=0.6 (right column) jets are shown. The results shown here are centrality-blind and η-integrated within the η-range of the EF trigger.

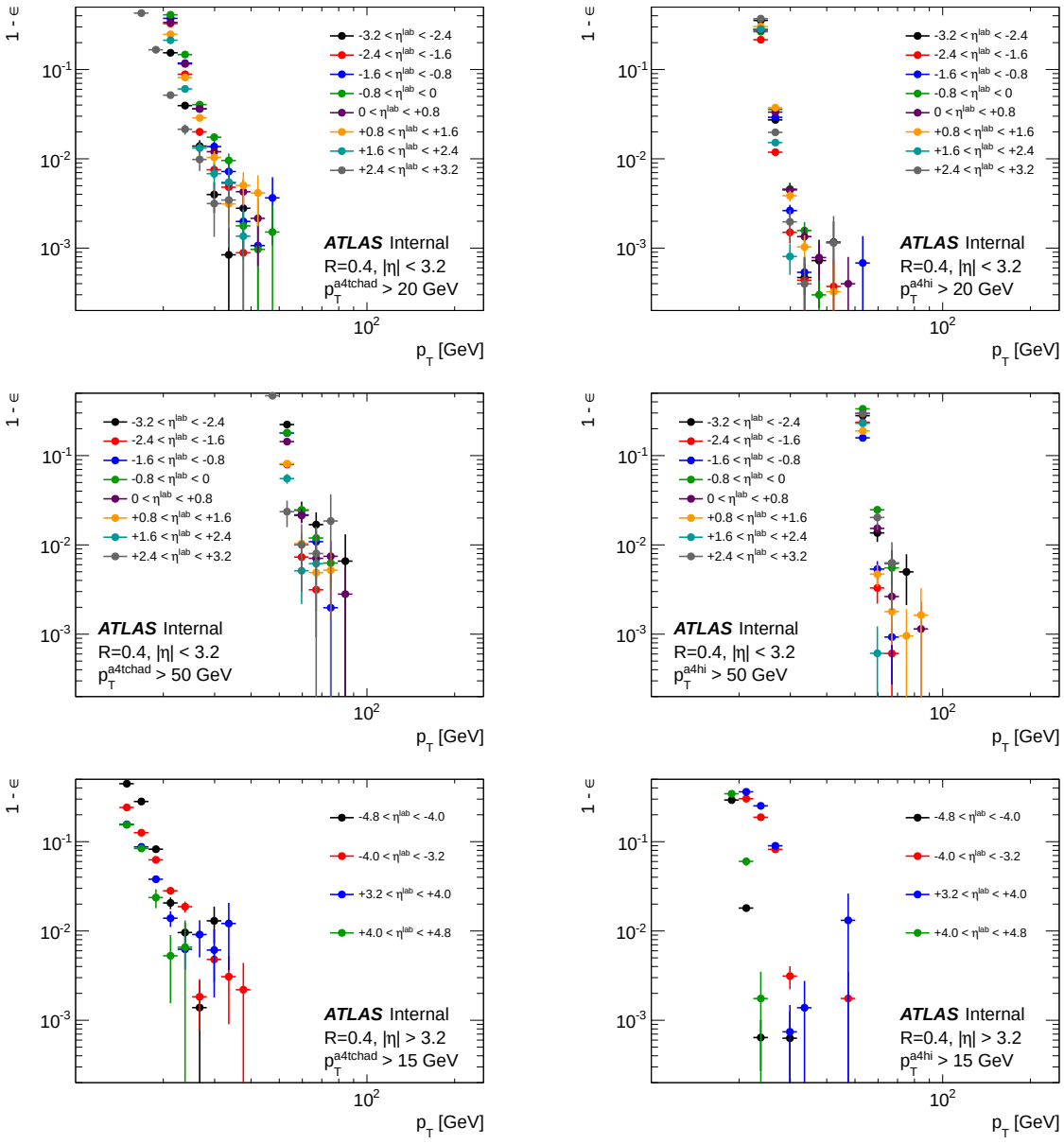


Figure 8.8: η dependence of the jet trigger efficiency, shown for j20 triggers (top row), j50 triggers (middle row) and fj15 triggers (bottom row). `a4tchad` triggers (left column) and `a4hi` triggers (right column) are shown separately. Results are plotted in the form $1 - \epsilon$, with p_{T} bins with perfect efficiency not shown on the logarithmic vertical axis. Within each plot, the efficiency is shown for different regions of η^{lab}.

As discussed in the next section, jets matched to the various triggers are only used when the trigger efficiency in question is $> 99\%$ (e.g. $1 - \epsilon < 1\%$) at all η and in all centralities.

8.4.2 p+Pb jet trigger scheme

The p+Pb trigger scheme for the construction of the per-event yields was chosen with several goals in mind. For rapidity bins that lie entirely within the barrel or the FCal, each p_{T} bin is only populated from one trigger, which is chosen to be the highest-luminosity trigger which has reached $> 99\%$ efficiency in all centrality bins for jets in that p_{T} bin. For the center-of-mass rapidity bins which overlap with both the $|\eta| > 3.2$ and $|\eta| < 3.2$ regions (in this case, the $-2.8 < y^* < -2.1$ and $3.6 < y^* < 4.4$ bins), the bin is filled by an OR of the relevant triggers in the scheme (e.g. both the fj15 and j20-75 triggers contribute to these bins).

The resulting trigger scheme which satisfied these requirements is shown in Table 8.4 for the p+Pb analysis. The systematic uncertainty on the possible residual inefficiency or residual centrality- or η dependence is taken to be 1% within the range 31.6-112 GeV.

8.4.3 pp jet trigger scheme

The determination of the trigger efficiency and final trigger scheme in pp data is substantially simpler than that used in p+Pb, for two reasons. First, only `a4tchad` trigger jet reconstruction was run. Second, there is no "centrality dependence" to check in the trigger efficiency.

Figure 8.9 shows the overall (e.g. η-integrated) trigger efficiency for all EF triggers used. The j10 trigger efficiency is constructed with respect to jets in the minimum bias (`EF_mbSpTrk`) stream, the j20 and j40 efficiency with respect to j10-triggered jets and the j50, j60 and j75 efficiency with respect to j40-triggered jets. As expected, the turn-on curve is sharper for R=0.4 offline jets (since they have the same R parameter as the trigger jets) than for R=0.6 offline jets (where the turn-on curve is smeared out due to the energy scale difference between trigger and offline jets). The η dependence of the trigger efficiency has also been investigated, to ensure that the chosen trigger scheme has $> 99\%$ efficiency separately in each region of the detector. Figures B.12 and B.13 in Appendix B.2 show the full $|\eta|$-dependent trigger efficiency in all triggers.

Table 8.5 summarizes the triggers used in the pp analysis. The p_{T} ranges chosen for a given trigger are those for which the efficiency is $> 99\%$. There are additional EF jet triggers in the

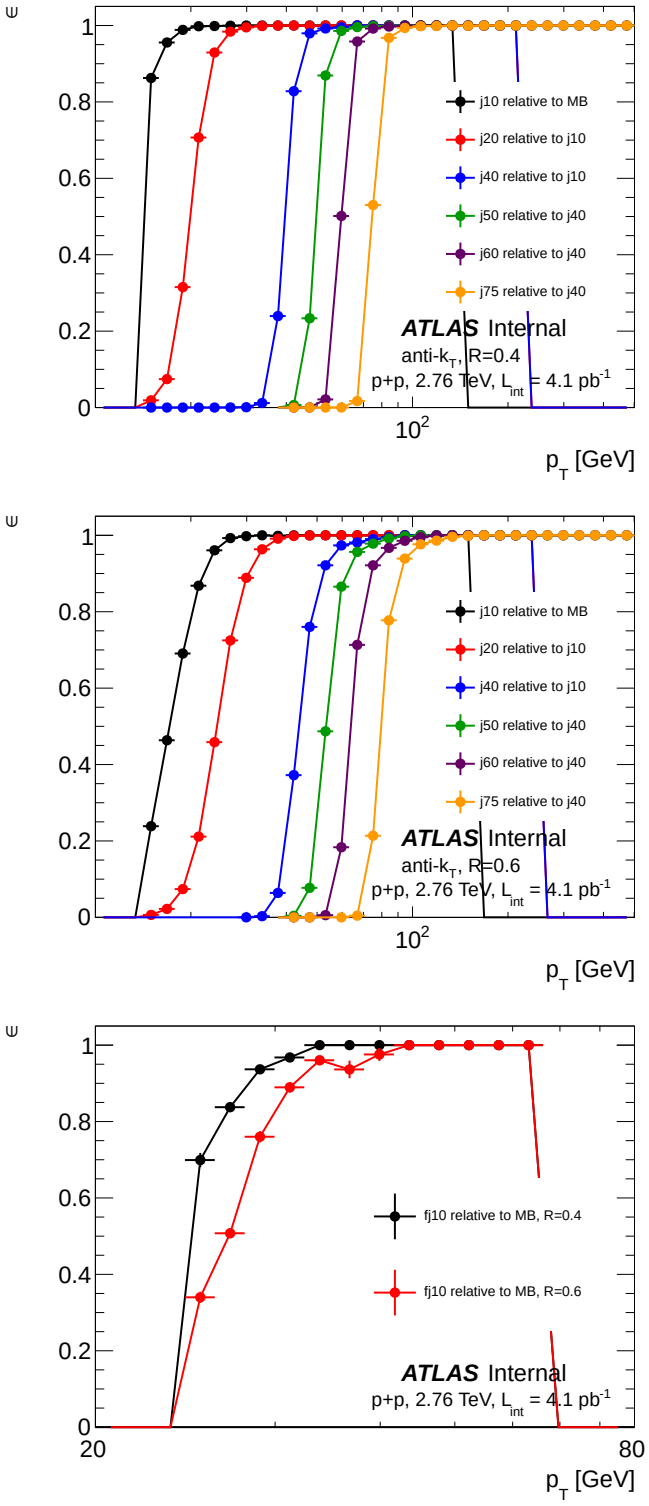


Figure 8.9: Summary of jet trigger efficiency in *pp* data. The efficiency of mid-rapidity triggers for R=0.4 (upper plot) and R=0.6 (middle plot) offline jets are shown. Each curve is the efficiency for a different trigger threshold, with respect to a minimum bias or lower threshold trigger. The efficiency of the fj10 trigger is shown for both R=0.4 and R=0.6 offline jets (bottom plot).

trigger menu at the time of running, but were not used because they had the same luminosity as a lower-threshold trigger (this includes the j30 and fj15 triggers as well as any trigger higher than the j75).

Although there exists a luminosity calibration for the March 2011 2.76 TeV pp data-taking[175], the final luminosity for the February 2013 running is still being prepared. Therefore, the integrated luminosity L_{int} for the triggers was determined using the ATLAS luminosity calculator (lumicalc) with `L1_TE10` used for the livefraction, and using the online calibration as the luminosity tag. For the `L1RD0`-seeded triggers, an extra factor of 8 is applied as an effective prescale (since the Level-1 `RD0` trigger fires with precisely $\frac{1}{8}^{\text{th}}$ the LHC clock rate).

8.4.4 Stability and distribution of triggered yields

The stability of the per-event jet yields was checked across the 16 period A and 16 period B runs for each of the triggers used in the trigger scheme. For each of the eight triggers used in the trigger scheme above (Section 8.4.2) and each p+Pb and Pb+p run, the yield of jets entering the final spectrum from that trigger is calculated, divided by the effective number of minimum-bias events in that run, $N_{\text{jet}}/N_{\text{evt}}$. For example, for the j20 trigger this corresponds to the number of j20 trigger-matched jets with $|\eta| < 3.2$, $p_{\text{T}} > {p_{\text{T}}}^{\text{threshold,j20}}$ and $p_{\text{T}} < {p_{\text{T}}}^{\text{threshold,j30}}$. For each trigger, the run-dependent per-event yield is compared to the total yield of jets from that trigger selection in the full p+Pb and Pb+p data, denoted $\langle N_{\text{jet}}/N_{\text{evt}}\rangle$. The results are shown in Figure 8.10. For all trigger-selected yields, the deviations of the run-dependent per-event yield from the mean are consistent with run-to-run statistical fluctuations.

The η^{lab} and ϕ positions of $p_{\text{T}} > 20$ GeV jets in MB p+Pb events, as well as that for jets matched to online j20, j50 and fj15 triggers, are plotted in Figure 8.11. Jets from Period A and Period B p+Pb data are shown separately.

8.5 Monte Carlo Simulation and Performance

8.5.1 MC jet samples

Several types of MC sets are used to understand the performance of jet reconstruction and quantify the effects of the p+Pb underlying event. In all cases, the underlying event generator is `PYTHIA`

Dataset name				
MC sample	orientation	N_{evt}	σ [nb] $\times \epsilon$	R=0.4 truth jet p_{T}, η^* range
mc12_5TeV.147760.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04.recon.e2166_s1675_s1586_r4893				
JZ0	period A	150×10^3	$(45.7 \times 10^6) \times (0.996)$	10 GeV $< p_{\text{T}} <$ 20 GeV
mc12_5TeV.147761.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04.recon.e2166_s1675_s1586_r4893				
JZ1	period A	625×10^3	$(45.7 \times 10^6) \times (4.09 \times 10^{-3})$	20 GeV $< p_{\text{T}} <$ 80 GeV
mc12_5TeV.147762.Pythia_AUET2BCTEQ6L1_jetjet_JZ2R04.recon.e2166_s1675_s1586_r4893				
JZ2	period A	625×10^3	$(11.0 \times 10^3) \times (4.82 \times 10^{-2})$	80 GeV $< p_{\text{T}} <$ 200 GeV
mc12_5TeV.147763.Pythia_AUET2BCTEQ6L1_jetjet_JZ3R04.recon.e2166_s1675_s1586_r4893				
JZ3	period A	625×10^3	$(165) \times (3.18 \times 10^{-2})$	200 GeV $< p_{\text{T}} <$ 500 GeV
mc12_5TeV.147764.Pythia_AUET2BCTEQ6L1_jetjet_JZ4R04.recon.e2166_s1675_s1586_r4893				
JZ4	period A	250×10^3	$(1.29) \times (1.36 \times 10^{-2})$	500 GeV $< p_{\text{T}} <$ 1 TeV
mc12_5TeV.147752.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04_MaxEta_m2p7.recon.e2166_s1675_s1586_r4893				
JZ0	period A	500×10^3	$(45.7 \times 10^6) \times (0.996)$	10 GeV $< p_{\text{T}} <$ 20 GeV, $\eta^* < -2.7$
mc12_5TeV.147753.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04_MaxEta_m2p7.recon.e2166_s1675_s1586_r4893				
JZ1	period A	500×10^3	$(45.7 \times 10^6) \times (4.00 \times 10^{-4})$	20 GeV $< p_{\text{T}} <$ 80 GeV, $\eta^* < -2.7$
mc12_5TeV.147760.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04.recon.e2167_s1675_s1586_r4893				
JZ0	period B	250×10^3	$(45.7 \times 10^6) \times (0.996)$	10 GeV $< p_{\text{T}} <$ 20 GeV
mc12_5TeV.147761.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04.recon.e2167_s1675_s1586_r4893				
JZ1	period B	625×10^3	$(45.7 \times 10^6) \times (4.09 \times 10^{-3})$	20 GeV $< p_{\text{T}} <$ 80 GeV
mc12_5TeV.147762.Pythia_AUET2BCTEQ6L1_jetjet_JZ2R04.recon.e2167_s1675_s1586_r4893				
JZ2	period B	625×10^3	$(11.0 \times 10^3) \times (4.82 \times 10^{-2})$	80 GeV $< p_{\text{T}} <$ 200 GeV
mc12_5TeV.147763.Pythia_AUET2BCTEQ6L1_jetjet_JZ3R04.recon.e2167_s1675_s1586_r4893				
JZ3	period B	625×10^3	$(165) \times (3.18 \times 10^{-2})$	200 GeV $< p_{\text{T}} <$ 500 GeV
mc12_5TeV.147764.Pythia_AUET2BCTEQ6L1_jetjet_JZ4R04.recon.e2167_s1675_s1586_r4893				
JZ4	period B	250×10^3	$(1.29) \times (1.36 \times 10^{-2})$	500 GeV $< p_{\text{T}} <$ 1 TeV
mc12_5TeV.147767.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04_MinEta_2p7.recon.e2267_s1675_s1586_r4893				
JZ0	period B	500×10^3	$(45.7 \times 10^6) \times (0.996)$	10 GeV $< p_{\text{T}} <$ 20 GeV, $\eta^* > 2.7$
mc12_5TeV.147768.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04_MinEta_2p7.recon.e2267_s1675_s1586_r4893				
JZ1	period B	500×10^3	$(45.7 \times 10^6) \times (4.00 \times 10^{-4})$	20 GeV $< p_{\text{T}} <$ 80 GeV, $\eta^* > 2.7$

Table 8.6: MC Signal datasets (boosted 5.02 TeV PYTHIA)

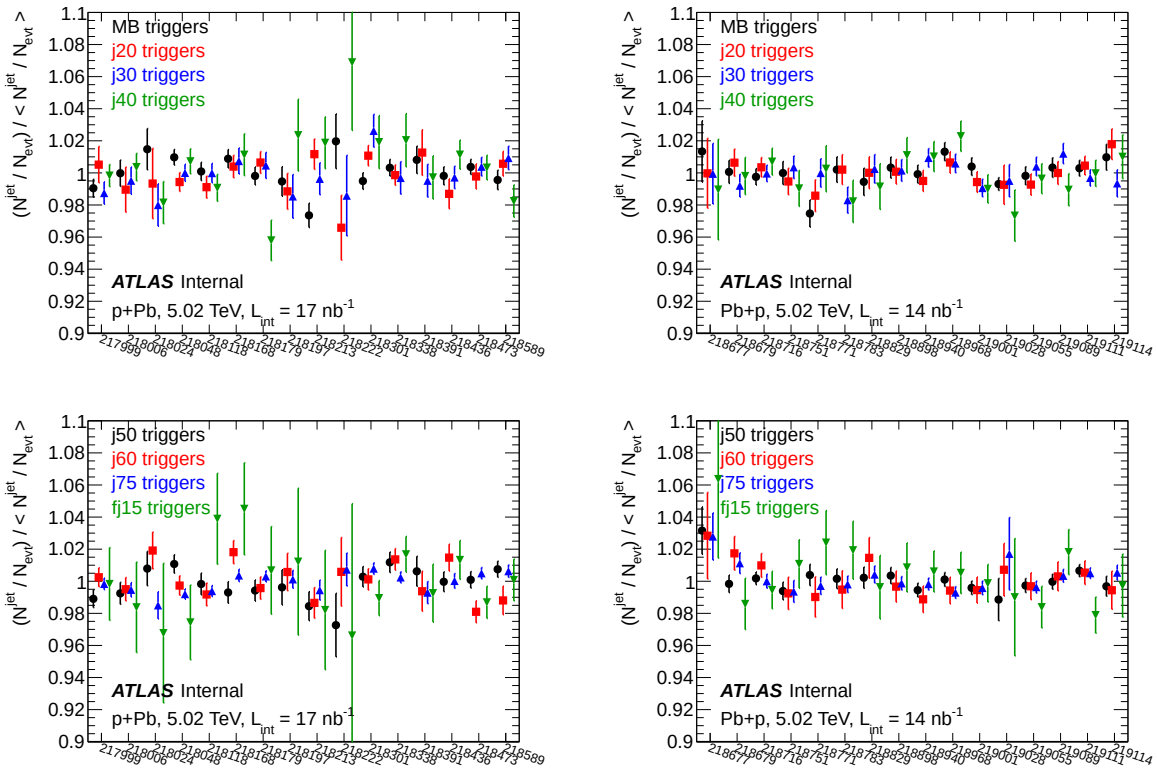


Figure 8.10: Run dependence of the per-event jet yield $N_{\text{jet}}/N_{\text{evt}}$ normalized by the mean per-event yield in all runs, shown for jets in the p_{T} ranges of the triggers from which the final yields are constructed. Period A p+Pb (Period B Pb+p) runs are shown in the left (right) column. The top row monitors the yield constructed from MB, j20, j30 and j40 triggers, while teh bottom row monitors the yields constructed from the j50, j60, j75 and fj15 triggers. The errors bars are statistical.

6.4[259], with parameters taken from the "AUTE2B" tune[260] to ATLAS 7 TeV pp and 1.96 TeV CDF and D0 $p\bar{p}$ data and using the CTEQ6L1[261] pdf sets. The three MC sets used are the 5.02 TeV Signal MC (pure PYTHIA pp events, with the p+Pb COM energy and kinematics), 5.02 TeV Overlay MC (PYTHIA pp events embedded into real p+Pb data events) and the 2.76 TeV MC (pure PYTHIA pp events, with $\sqrt{s} = 2.76$ TeV), and are described below.

Table 8.6 lists the 5.02 TeV pp Signal MC samples, which consist of a full detector simulation of PYTHIA dijets at $\sqrt{s} = 5.02$ TeV. Two versions of the Signal MC are generated, corresponding to

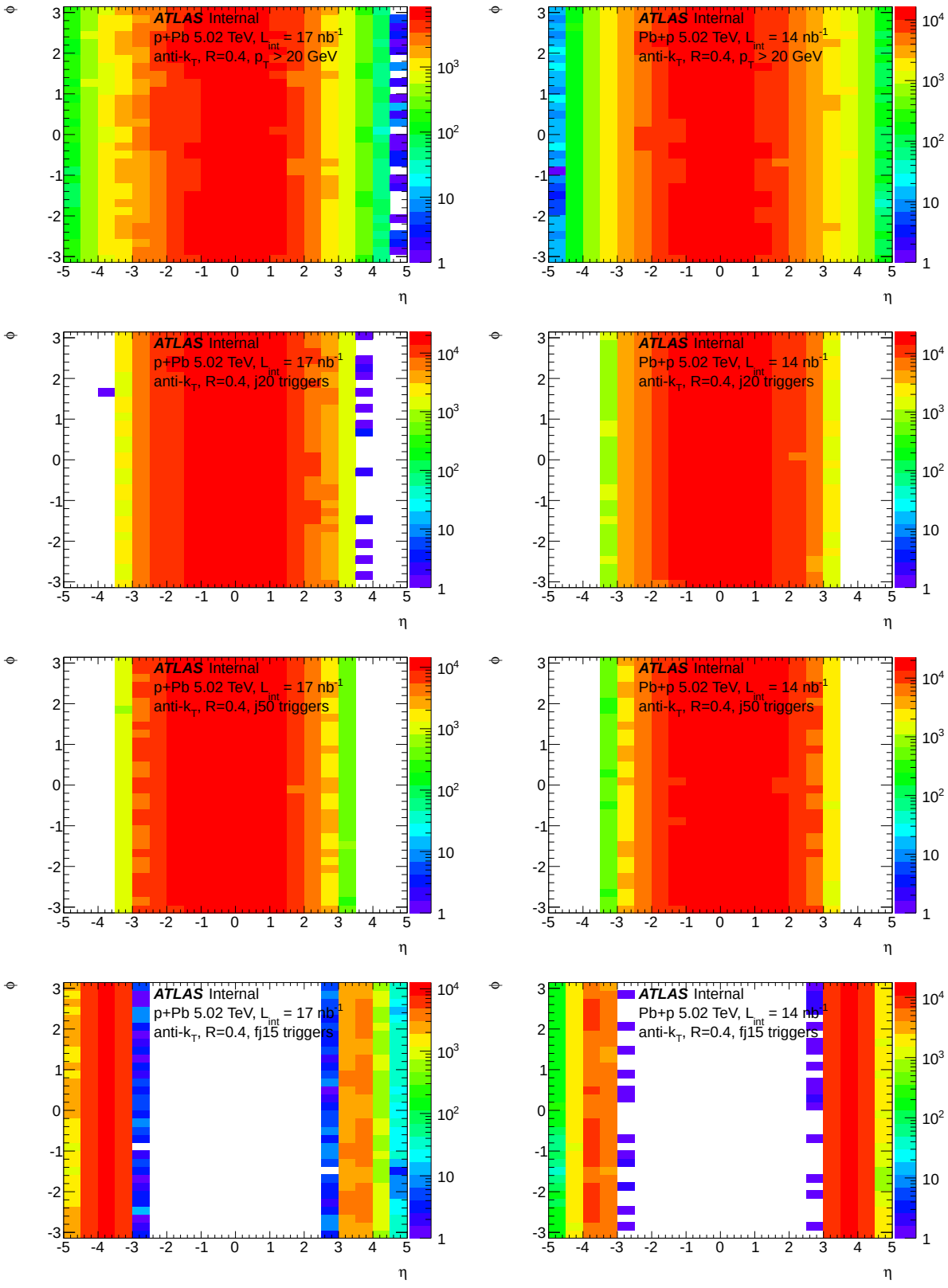


Figure 8.11: η/ϕ positions of jets in MB events (with $p_{\mathrm{T}} > 20$ GeV, top row), jets matched to j20 triggers (second row), j50 triggers (third row) and fj15 triggers (bottom row). Jets are shown from period A (left column) and period B (right column) separately.

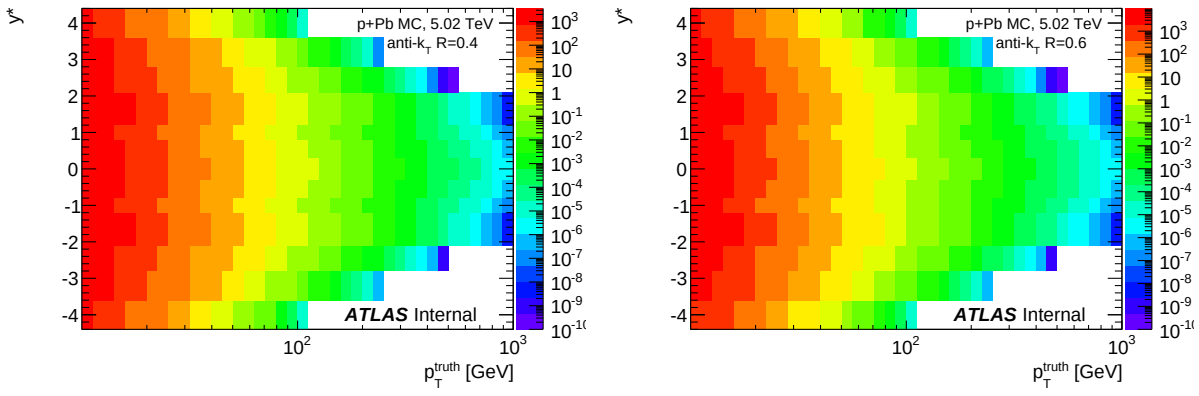


Figure 8.12: Total JZ-combined spectrum of jets in the p+Pb Overlay MC for R=0.4 (left) and R=0.6 (right) truth jets.

the period A and period B kinematics (e.g. the resulting generator event is boosted by $\Delta y = \pm 0.465$ depending on the kinematics). The generator-level events are filtered on the requirement of a R=0.4 truth jet within a certain kinematic range. Five different kinematic ranges are used to cover the range 20 GeV $< p_{\mathrm{T}}^{\mathrm{truth}} <$ 1 TeV, referred to as $JZ0$ (10-20 GeV), $JZ1$ (20-80 GeV), $JZ2$ (80-200 GeV), $JZ3$ (200-500 GeV) and $JZ4$ (500 GeV - 1 TeV). Furthermore, there are two additional MC sets with $JZ0$ and $JZ1$ truth jets, but filtered on the truth jet being at forward rapidity, corresponding to $\eta^* < -2.7$ in the period A kinematics sample and $\eta^* > 2.7$ in the period B kinematics sample.

Table 8.7 lists the p+Pb Overlay MC samples, which consist of $\sqrt{s} = 5.02$ TeV PYTHIA dijets overlayed onto real minimum bias p+Pb data, along with the number of generated events, the PYTHIA cross-section and efficiency of trigger truth reconstruction. In the MC Overlay, PYTHIA truth jets with p+Pb period A kinematics (e.g. with an overall center of mass shift $\Delta y = -0.465$) are embedded into p+Pb data from period A. Just like the MC Signal, the MC Overlay sets are also generated in five different JZ samples, with an additional two samples selecting forward jets with $\eta^* < -2.7$. The centrality of the original p+Pb events used in the MC Overlay was recorded, meaning that the performance can be evaluated in a centrality-dependent way. The Overlay MC is known to have a defect in which the energy from jets captured in Tile Barrel and Tile Extension cells is degraded. This occurs because of a software bug in the reconstruction of the overlaid data, in which Tile cells switching from the high gain to low gain setting due to the presence of additional

Dataset name			
MC sample	N_{evt}	σ [nb] $\times \epsilon$	R=0.4 truth jet p_{T}, η^* range
mc12_5TeV.147760.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04.recon.NTUP_HI.e2166_d773_r4642_tid01320078_00			
JZ0	1.3×10^6	$(45.7 \times 10^6) \times (0.996)$	10 GeV $< p_{\text{T}} <$ 20 GeV
mc12_5TeV.147761.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04.recon.NTUP_HI.e2166_d773_r4642_tid01320079_00			
JZ1	1.3×10^6	$(45.7 \times 10^6) \times (4.09 \times 10^{-3})$	20 GeV $< p_{\text{T}} <$ 80 GeV
mc12_5TeV.147762.Pythia_AUET2BCTEQ6L1_jetjet_JZ2R04.recon.NTUP_HI.e2166_d773_r4642_tid01320082_00			
JZ2	1.3×10^6	$(11.0 \times 10^3) \times (4.82 \times 10^{-2})$	80 GeV $< p_{\text{T}} <$ 200 GeV
mc12_5TeV.147763.Pythia_AUET2BCTEQ6L1_jetjet_JZ3R04.recon.NTUP_HI.e2166_d773_r4642_tid01320085_00			
JZ3	1.3×10^6	$(165) \times (3.18\times^{-2})$	200 GeV $< p_{\text{T}} <$ 500 GeV
mc12_5TeV.147764.Pythia_AUET2BCTEQ6L1_jetjet_JZ4R04.recon.NTUP_HI.e2166_d773_r4642_tid01320088_00			
JZ4	1.3×10^6	$(1.29) \times (1.26 \times 10^{-2})$	500 GeV $< p_{\text{T}} <$ 1 TeV
mc12_5TeV.147752.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04_MaxEta_m2p7.recon.NTUP_HI.e2166_d773_r4642_tid01320102_00 mc12_5TeV.147752.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04_MaxEta_m2p7.recon.NTUP_HI.e2166_d773_r4642_tid01320103_00			
JZ0, MaxEta2p7	2.6×10^6	$(45.7 \times 10^6) \times (0.9996)$	10 GeV $< p_{\text{T}} <$ 20 GeV, $\eta^* < -2.7$
mc12_5TeV.147753.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04_MaxEta_m2p7.recon.NTUP_HI.e2166_d773_r4642_tid01320106_00 mc12_5TeV.147753.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04_MaxEta_m2p7.recon.NTUP_HI.e2166_d773_r4642_tid01320107_00			
JZ1, MaxEta2p7	2.6×10^6	$(45.7 \times 10^6) \times (4.00 \times 10^{-4})$	20 GeV $< p_{\text{T}}$, $\eta^* < -2.7$

Table 8.7: p+Pb MC Overlay samples (boosted 5.02 TeV PYTHIA embedded into minimum bias p+Pb data)

energy from the PYTHIA event are inadvertently masked off. Thus, the EM-scale energy of truth-associated jets is systematically low for high-p_{T} jets. This effect turns on between 100 GeV and 200 GeV. Therefore, with these caveats, the Overlay MC is only used for jets outside the Tile-instrumented region ($|\eta| > 1.7$) and for $p_{\text{T}} \lesssim 200$ GeV jets within $|\eta| < 1.7$, and *only* to understand the centrality dependence in the response and not the absolute response.

Table 8.8 details the 2.76 TeV pp MC samples, which consist of a full detector simulation of PYTHIA dijets at $\sqrt{s} = 2.76$ TeV. As with the other MC samples, the 2.76 TeV MC is divided into five separate subsets, each selecting on a different range of R=0.4 truth jet p_{T}. Unlike the 5.02 TeV Signal and Overlay MC, there are no forward jet enhanced subsamples.

In the inclusive p+Pb jet analysis, the Signal MC is used to perform the jet calibration (see

Dataset name			
MC sample	N_{evt}	σ [nb] $\times \epsilon$	R=0.4 truth jet p_{T}, η range
mc12_2TeV.147760.Pythia_AUET2BCTEQ6L1_jetjet_JZ0R04.recon.e2201_s1647_s1586_r4895			
JZ0	1.9×10^6	$(41.3 \times 10^6) \times (0.998)$	10 GeV $< p_{\text{T}} <$ 20 GeV
mc12_2TeV.147761.Pythia_AUET2BCTEQ6L1_jetjet_JZ1R04.recon.e2201_s1647_s1586_r4895			
JZ1	5.0×10^6	$(41.3 \times 10^6) \times (1.56 \times 10^{-3})$	20 GeV $< p_{\text{T}} <$ 80 GeV
mc12_2TeV.147762.Pythia_AUET2BCTEQ6L1_jetjet_JZ2R04.recon.e2201_s1647_s1586_r4895			
JZ2	5.0×10^6	$(3.23 \times 10^3) \times (3.65 \times 10^{-2})$	80 GeV $< p_{\text{T}} <$ 200 GeV
mc12_2TeV.147763.Pythia_AUET2BCTEQ6L1_jetjet_JZ3R04.recon.e2201_s1647_s1586_r4895			
JZ3	5.0×10^6	$(33.3) \times (1.96 \times 10^{-2})$	200 GeV $< p_{\text{T}} <$ 500 GeV
mc12_2TeV.147764.Pythia_AUET2BCTEQ6L1_jetjet_JZ4R04.recon.e2201_s1647_s1586_r4895			
JZ4	2.0×10^6	$(0.129) \times (3.41 \times 10^{-3})$	500 GeV $< p_{\text{T}} <$ 1 TeV

Table 8.8: 2.76 TeV PYTHIA MC datasets

Section 8.3.2) and evaluate the jet reconstruction efficiency, energy resolution, any residual non-closure in the jet energy scale and other indicators of the performance. On the other hand, the Overlay MC is used to gauge the centrality dependence in the jet performance introduced by the p+Pb underlying event in different centrality bins. The Overlay MC is used to determine the p_{T} range in which the detector effects on the p+Pb yields are centrality-independent. Then, the Signal MC is used to correct the measured p+Pb yields to the truth level through a bin-by-bin unfolding method.

In the inclusive pp jet analysis, the 2.76 TeV MC is used to check the energy scale and resolution, and used to unfold the p+Pb jet yields though a bin-by-bin unfolding.

In any of the MC samples, variables related to the jet performance are typically constructed by combining the results from the individual JZ samples in accordance with the generator cross-section for each subsample. The total spectrum of truth jets in a given y^* bin can be constructed by weighing the spectra in the individual JZ samples according to

$$\frac{dN^{\text{truth}}}{dp_{\text{T}}}(p_{\text{T}}) = \sum_{JZ} w_{JZ} \frac{dN^{\text{truth}}{}_{JZ}}{dp_{\text{T}}}(p_{\text{T}}) \tag{8.25}$$

where $\frac{dN^{\text{truth}}{}_{JZ}}{dp_{\text{T}}}(p_{\text{T}})$ is the spectrum in a given JZ sample, and

$$w_{JZ} \equiv \frac{\sigma_{JZ} \times \epsilon_{JZ}}{\sum_{JZ'} \sigma_{JZ'} \times \epsilon_{JZ'}} \tag{8.26}$$

is the cross-section times generator efficiency weight, normalized to 1. Figure 8.12 shows the total JZ-combined spectrum as a function of p_{T} and y^*.

8.5.2 Reconstruction efficiency

For each truth jet with $p_{\mathrm{T}}^{\mathrm{truth}} > 10$ GeV, an attempt is made to find an associated reconstructed jet within $\Delta R = \sqrt{(\eta^{\mathrm{truth}} - \eta^{\mathrm{reco}})^2 + (\phi^{\mathrm{truth}} - \phi^{\mathrm{reco}})^2} < 0.4$. Only reconstructed jets with $p_{\mathrm{T}}^{\mathrm{reco}} > 10$ GeV are considered. If there are multiple reconstructed jets within $\Delta R < 0.4$, by default the nearest jet is taken to be the associated jet (but note that when determining the jet energy scale closure and resolution, an isolation cut is applied to ensure a $1 \leftrightarrow 1$ truth to reconstructed corresponce). If there are no reconstructed jets above threshold within $\Delta R < 0.4$, the truth jet is taken to not be successfully reconstructed.

Note that since the data is corrected for the efficiency of the cleaning cuts using a data-driven method, and that the MC is not known to model the effect of the cleaning cuts well, the cleaning cuts are not included in the definition of the reconstruction efficiency. However, as an investigation of how well the cleaning cut relevant variables are modeled in MC, the efficiency in MC is plotted alongside the survival fraction in data in Figures 8.4 and 8.5.

Unlike the cross-section weighted spectrum in Equation 8.25, the total reconstruction efficiency must be determined in another way. In each JZ sample, let ${N^{\mathrm{truth}}}_{(JZ)}$ be the total number of truth jets at some p_{T} and $N_{(JZ)}^{\mathrm{matched}}$ be the subset of those successfully reconstructed. Then the efficiency in a given JZ sample is

$$\epsilon(p_{\mathrm{T}}; JZ) = N_{(JZ)}^{\mathrm{matched}}(p_{\mathrm{T}}) / {N^{\mathrm{truth}}}_{(JZ)}(p_{\mathrm{T}}) \tag{8.27}$$

with the variance on ϵ estimated assuming a binomial distribution with $P = \epsilon$ and $N = {N^{\mathrm{truth}}}_{(JZ)}$. The total weighted efficiency is

$$\epsilon(p_{\mathrm{T}}) = \frac{\sum_{JZ} w_{JZ} \cdot {N^{\mathrm{truth}}}_{(JZ)}(p_{\mathrm{T}}) \cdot \epsilon(p_{\mathrm{T}}; JZ)}{\sum_{JZ} w_{JZ} \cdot {N^{\mathrm{truth}}}_{(JZ)}(p_{\mathrm{T}})} \tag{8.28}$$

where the uncertainty on each $\epsilon(p_{\mathrm{T}}; JZ)$ is propagated to the total efficiency $\epsilon(p_{\mathrm{T}})$.

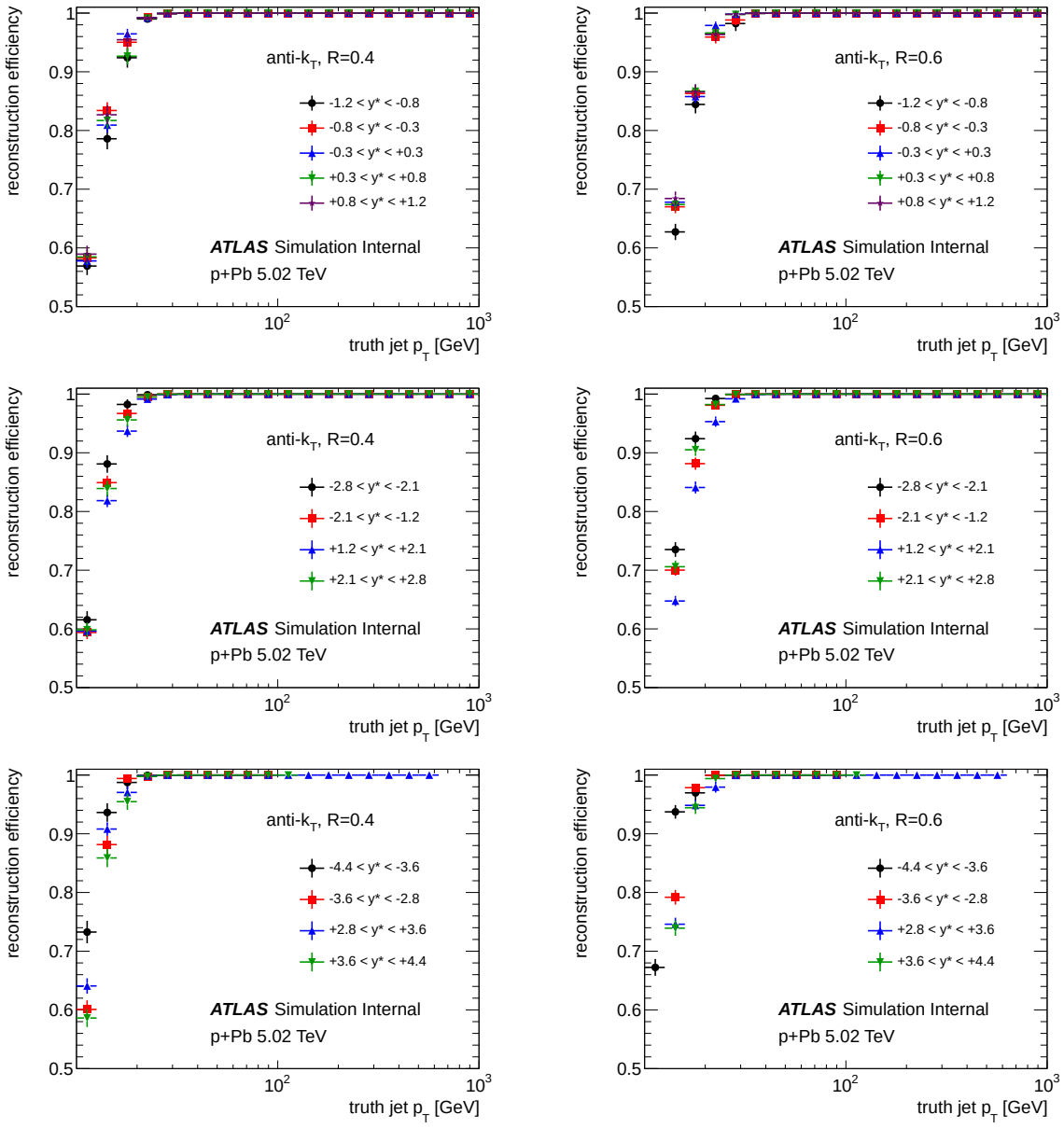


Figure 8.13: Reconstruction efficiency vs. truth jet p_{T} in the 5.02 TeV Signal MC, for R=0.4 (left column) and R=0.6 (right column) jets, showing jets in mid-rapidity bins (top row), at intermediate rapidities (middle row) and forward and backward rapidities (bottom row).

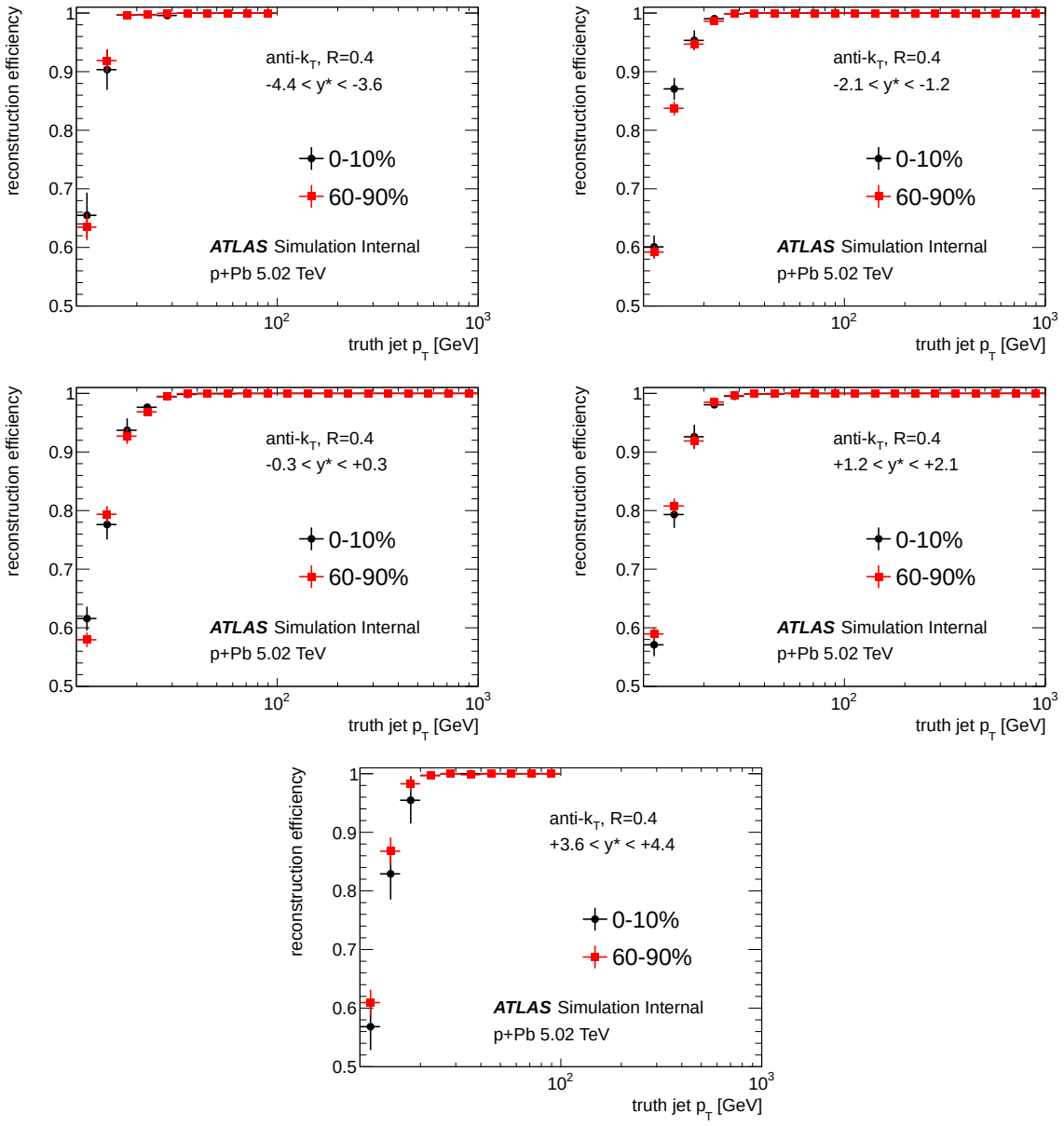


Figure 8.14: Reconstruction efficiency vs. truth jet p_{T} in the 5.02 TeV Overlay MC, for R=0.4 jets, for selected rapidity bins. The efficiency in central (peripheral) events is plotted with black (red) markers.

The reconstruction efficiency in the 5.02 TeV Signal MC is shown as a function of truth jet p_{T} in different y^* bins in 8.13. In all rapidity bins, the reconstruction efficiency is $> 97\%$ ($> 92\%$) above $p_{\mathrm{T}} > 20$ for R=0.4 (R=0.6) truth jets, and is $> 99.5\%$ at $p_{\mathrm{T}} > 30$ GeV for both R values. Additionally, the centrality dependence of the efficiency has been investigated in all rapidity bins in the 5.02 TeV Overlay MC (where it is possible to tag the p+Pb event centrality). Figure 8.14 shows the efficiency in the most central 0-10% and the most peripheral 60-90% centrality selections in selected rapidity bins for R=0.4 bins. The centrality dependence of the reconstruction efficiency is $< 1\%$ for 20 GeV truth jets in all rapidity bins.

8.5.3 Jet energy scale closure and resolution

Matched pairs of truth and reconstructed jets are used to evaluate the response of the detector to jets. In order to measure the response of the detector to a single jet precisely as possible, several isolation criteria are required to eliminate contributions to $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{truth}}$ from splitting at the truth or reconstructed level, as well as (in the case of the Overlay MC) the presence of jets in the original p+Pb data event. In this way, only "$1 \leftrightarrow 1$" correspondences between truth and reconstructed jet are used to determine the jet energy scale (JES) closure and jet energy resolution (JER).

The isolation criteria are:

- No other truth jet (above $p_{\mathrm{T}}^{\mathrm{truth}} > 10$ GeV) within $\Delta R < 1.5 \times R$, where R is the jet resolution parameter
- No other reconstructed jet (above $E_{\mathrm{T}}^{\mathrm{EM}} > 7$ GeV) within $\Delta R < 1.5 \times R$, where R is the jet resolution parameter

Furthermore, when evaluating the JES closure or JER in the MC Overlay, one additional isolation cut is used:

- No jet from the original p+Pb data event (above $E_{\mathrm{T}}^{\mathrm{EM}} > 7$ GeV) within $\Delta R < 1.5 \times R$, where R is the jet resolution parameter

Matched pairs of truth and reconstructed jets are contained in a transfer matrix, A_{ij}, where i (j) are indices into the $p_{\mathrm{T}}^{\mathrm{truth}}$ ($p_{\mathrm{T}}^{\mathrm{reco}}$) bins, and each entry in the transfer matrix is the (JZ cross-section

weighted) number of truth-reco jet pairs falling into the given p_{T} bin. The transfer matrices from each JZ sample are combined according to

$$A_{ij} = \sum_{JZ} A_{ij}^{(JZ)} w_{JZ} \tag{8.29}$$

Transfer matrices in all y^* bins and both R parameter values are shown in Figures 8.15, 8.16, 8.17 and 8.18. Note that the quantity plotted is A_{ij} and not the p_{T} bin-size normalized $d^2N/dp_{\mathrm{T}}^{\mathrm{truth}}p_{\mathrm{T}}^{\mathrm{reco}}$.

The JES closure and JER are determined by first calculating the distribution of $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{truth}} = (p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{truth}})/p_{\mathrm{T}}^{\mathrm{truth}}$ for matched pairs of truth and reconstructed jets, in bins of $p_{\mathrm{T}}^{\mathrm{truth}}$. The $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{truth}}$ distributions are combined from each JZ sample by weighing according to the generator cross-section in the standard way. Then, the JES closure and JER are determined by Gaussian fits to the resulting $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{truth}}$ distributions within a restricted range. At low $p_{\mathrm{T}}^{\mathrm{truth}}$, the lower end of the fitted range is restricted to account for the minimum p_{T} of the reconstructed jets.

Figures 8.19 and 8.20 shows the JES closure and JER in the p+Pb Signal MC, while Figures 8.21 and 8.22 show the JES closure and JER in the pp Standalone MC.

8.5.4 Constraint on centrality dependence of correction factors in p+Pb

The measured jet yields in p+Pb and pp collisions are corrected for the jet energy resolution and residual non-closure through the use of bin-by-bin correction factors. Three separate sets of bin-by-bin correction factors are derived from each of the three MC samples for different purposes.

The Overlay MC is used to investigate how the correction factors change in the different p+Pb centrality bins. (The actual correction factors applied to the p+Pb and pp yields are described below in Sections 8.5.5 and 8.5.6) Although the Overlay MC is known to have an energy scale problem for high-p_{T} jets ($p_{\mathrm{T}} > 200$ GeV or so), it is still used at low-p_{T} to understand centrality-dependent differences in the bin migration. Ultimately, the p+Pb yields are only used in a p_{T} region where the bin migration effects have become centrality-independent. This is determined by comparing the correction factors in central to peripheral overlay events.

The bin-by-bin machinery and correction factors are defined as follows. Working with truth and reconstructed bins of the same size, the bin-by-bin quantities can be defined as follows. Let T_i be the spectrum of truth jets, where i denotes the i^{th} truth p_{T} bin. T_i can be separated into two

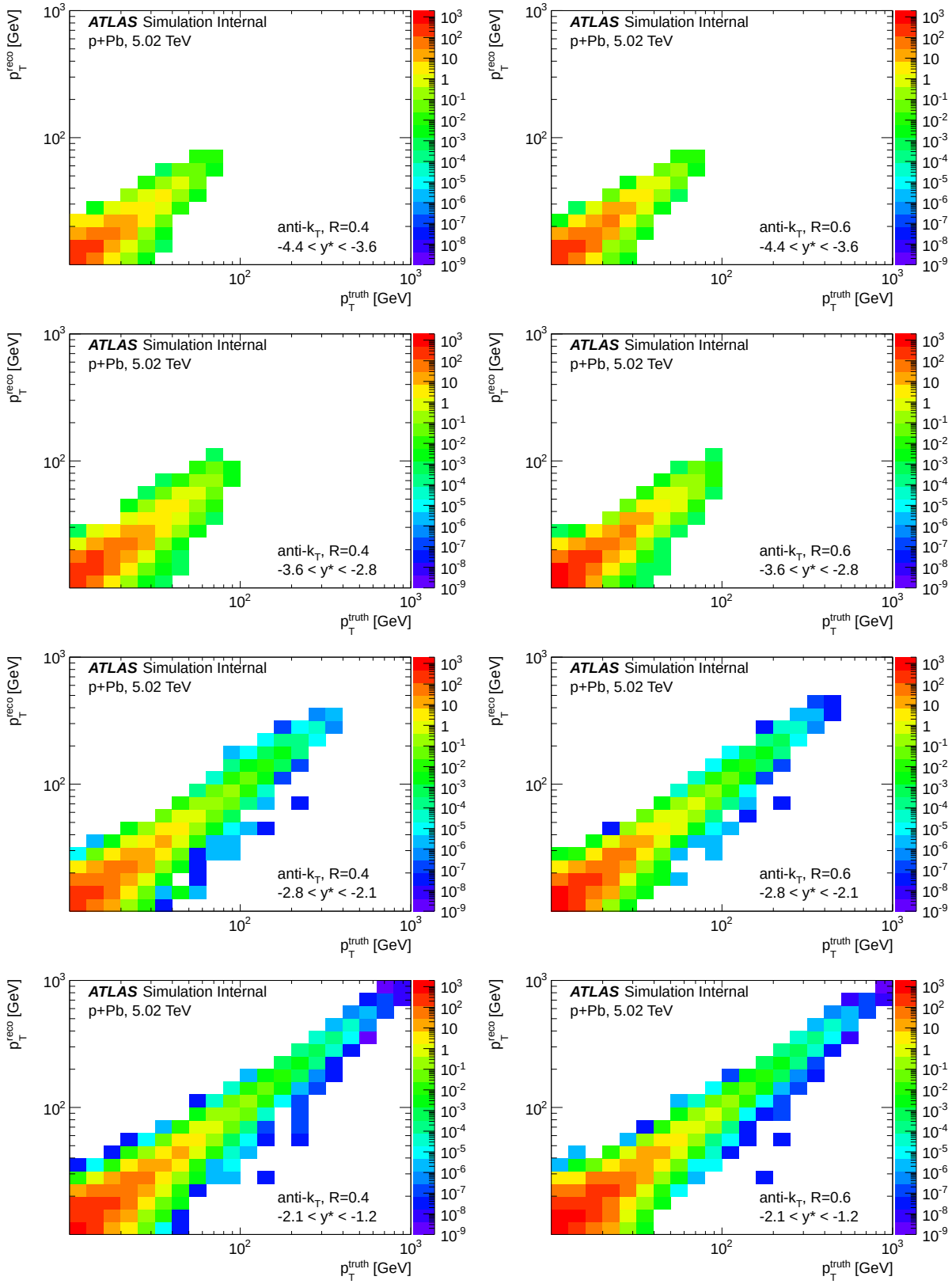


Figure 8.15: Transfer matrices A_{ij} in the 5.02 TeV Signal MC. R=0.4 jets (left column) and R=0.6 jets (right column) are shown separately. The matrices here show the rapidity bins $-4.4 < y^* < -1.2$.

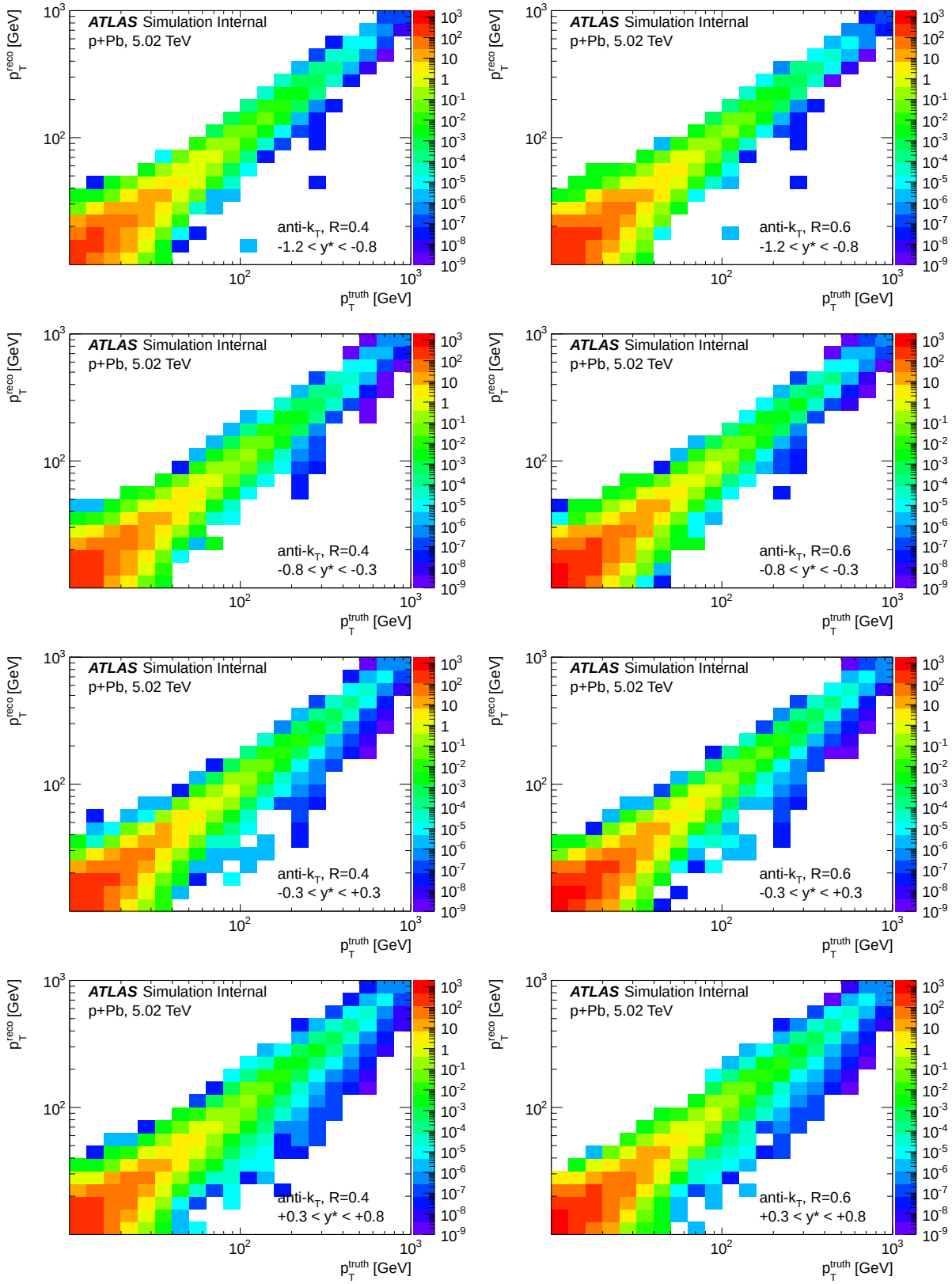


Figure 8.16: Transfer matrices A_{ij} in the 5.02 TeV Signal MC. R=0.4 jets (left column) and R=0.6 jets (right column) are shown separately. The matrices here show the rapidity bins $-1.2 < y^* < +0.8$.

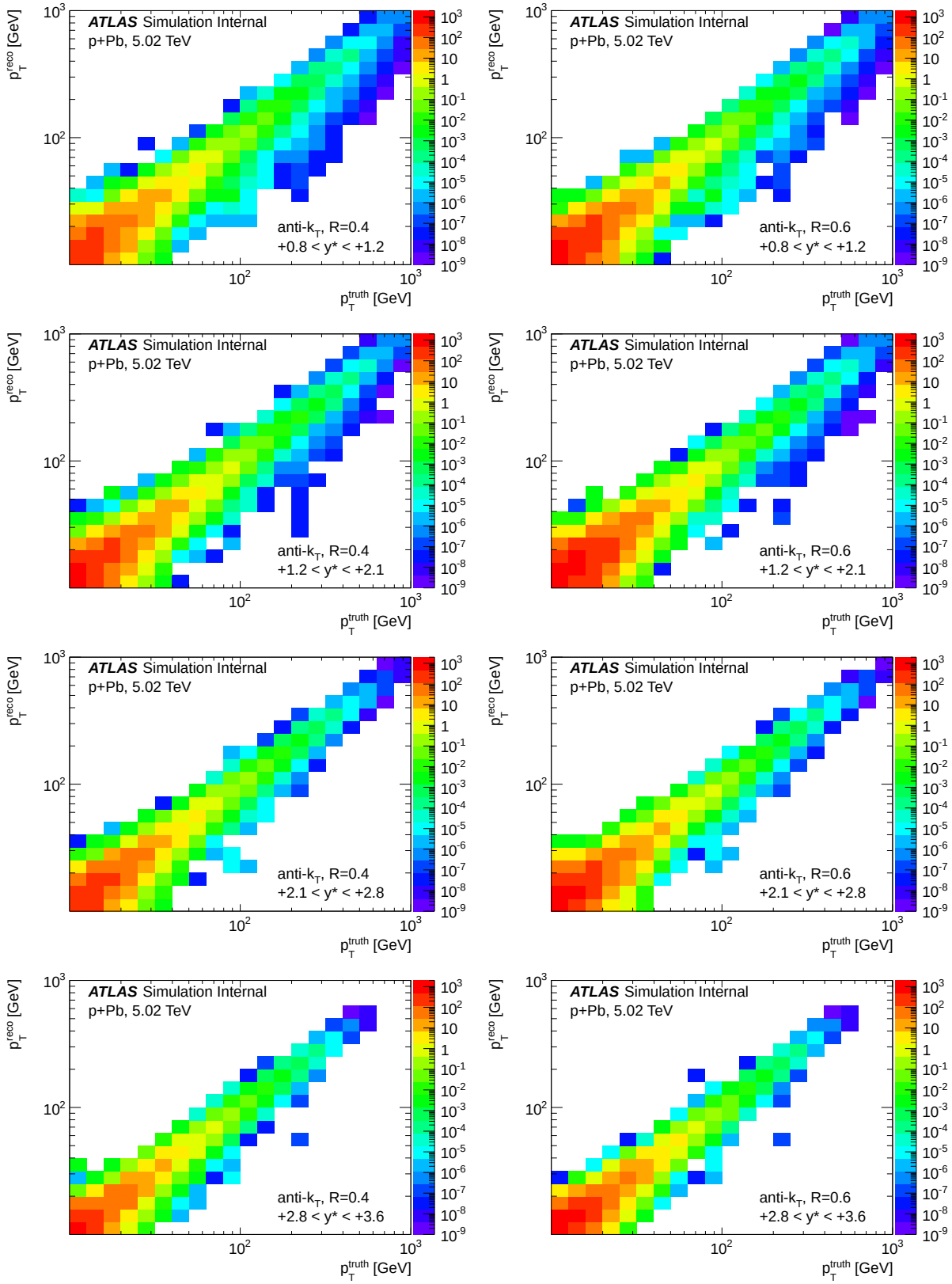


Figure 8.17: Transfer matrices A_{ij} in the 5.02 TeV Signal MC. R=0.4 jets (left column) and R=0.6 jets (right column) are shown separately. The matrices here show the rapidity bins $+0.8 < y^* < +3.6$.

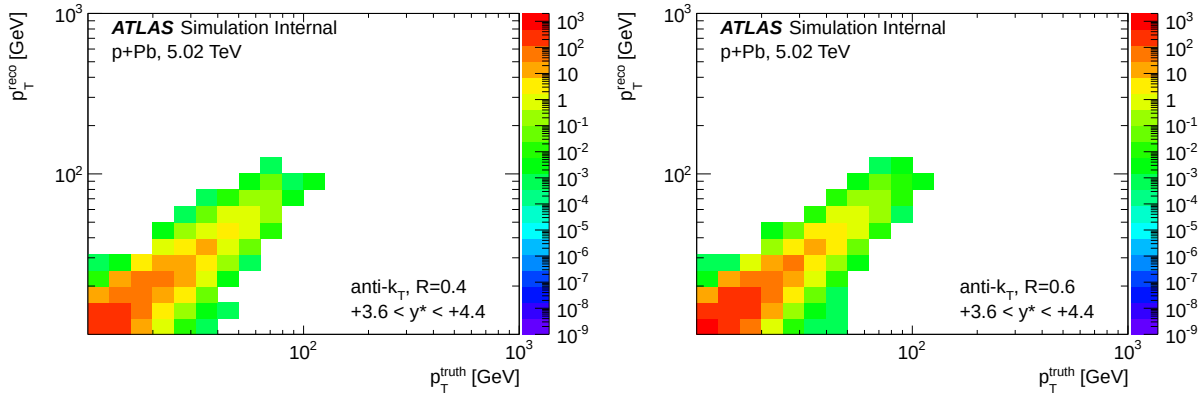


Figure 8.18: Transfer matrices A_{ij} in the 5.02 TeV Signal MC. R=0.4 jets (left column) and R=0.6 jets (right column) are shown separately. The matrices here show the rapidity bins $+3.6 < y^* < +4.4$.

components, the set of truth jets which are reconstructed (e.g. matched with a reconstructed jet) T_i^{matched} and those which are not $T_i^{\text{unmatched}}$,

$$T_i = T_i^{\text{matched}} + T_i^{\text{unmatched}} \tag{8.30}$$

Let R_i be the spectrum of the truth-matched reconstructed jets, where i denotes the i^{th} reco p_{T} bin. Let A_{ij} be the (i,j) element of the $(p_{\text{T}}^{\text{truth}}, p_{\text{T}}^{\text{reco}})$ transfer matrix. The elements of the vectors and matrices are the number of counts per bin, and *not* the bin size normalized dN/dp_{T} or $d^2N/dp_{\text{T}}^{\text{truth}}dp_{\text{T}}^{\text{reco}}$. A_{ij} is related to R_i and T_i^{matched} via

$$R_i = \sum_i A_{ij} \tag{8.31}$$

$$T_k^{\text{matched}} = \sum_l A_{kl} \tag{8.32}$$

where to recover the truth-matched reco spectrum, we sum over the truth axis of the transfer matrix and to recover the truth spectrum with a match, we sum over the reco axis of the transfer matrix.

Two quantities of merit which quantify the extent of the bin migration are the bin-by-bin *efficiency* E_i, which is the fraction of jets in a given truth bin which remain in that bin, and the

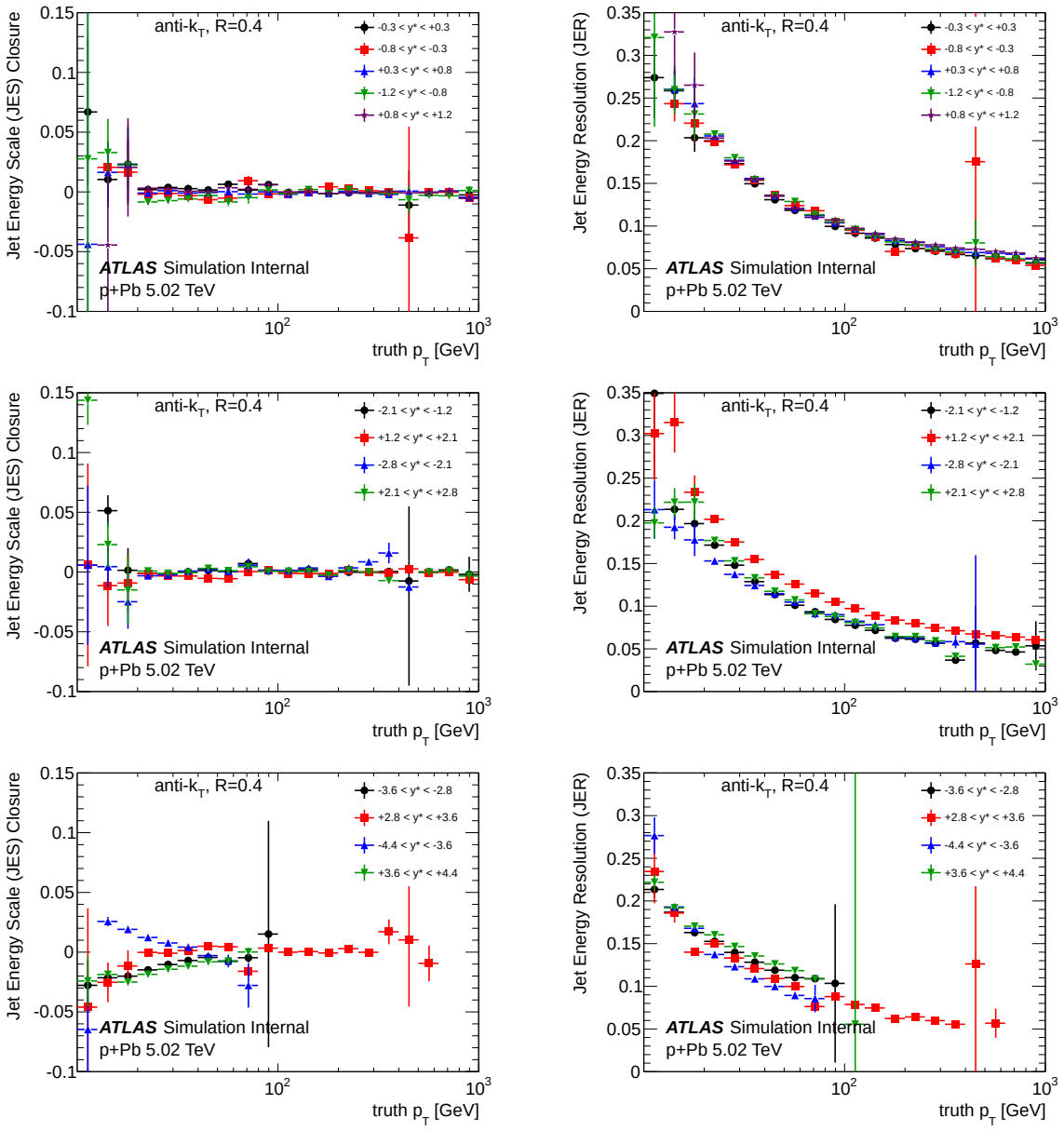


Figure 8.19: Jet energy scale closure (left column) and resolution (right column) for R=0.4 jets in the 5.02 TeV Signal MC, in bins of y^*. Each plot shows a selection of rapidities grouped into mid-rapidity (top row, $|y^*| < 1.2$), intermediate rapidity (middle row, $1.2 < |y^*| < 2.8$) and forward/backward rapidity ($2.8 < |y^*| < 4.4$).

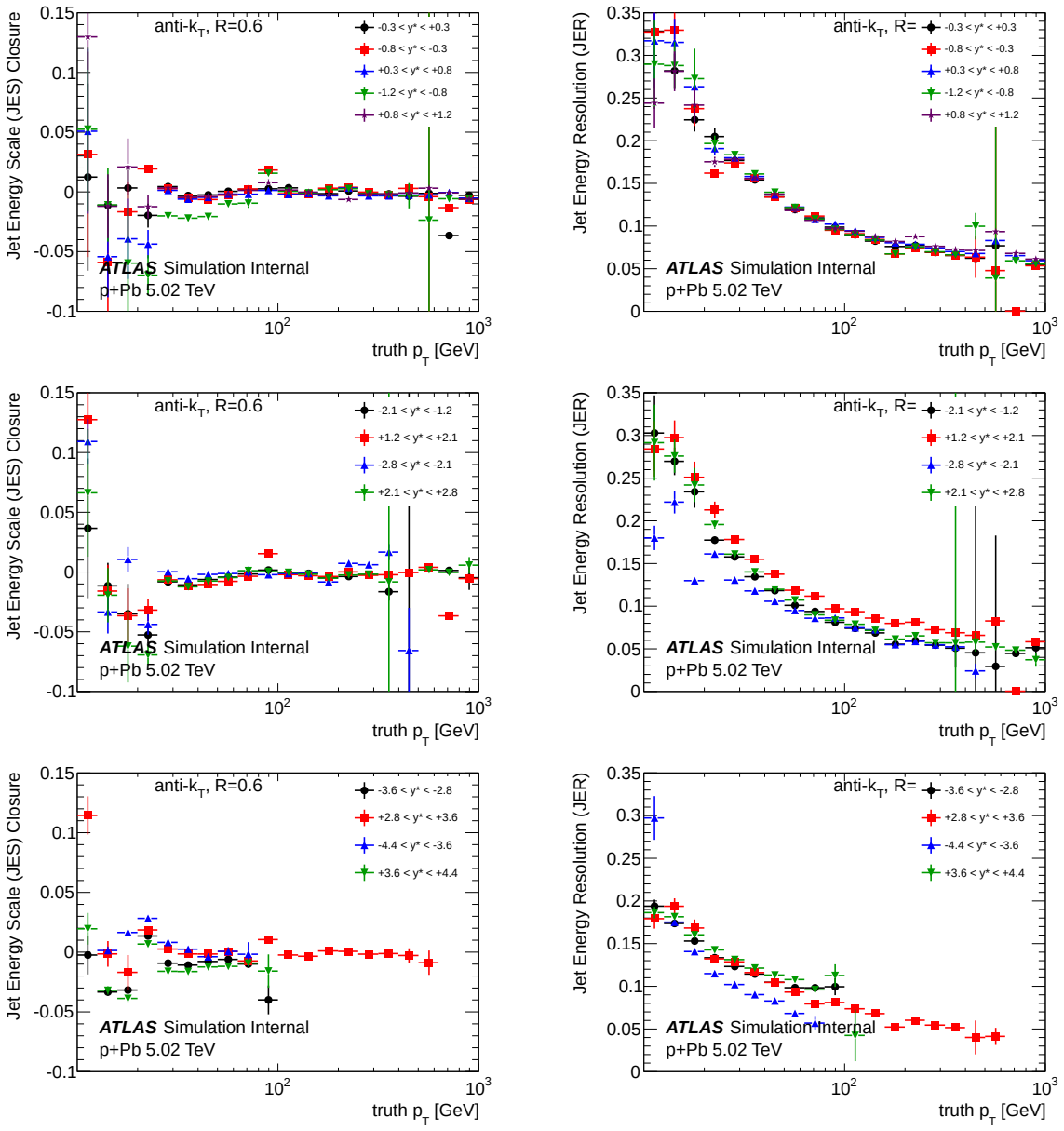


Figure 8.20: Jet energy scale closure (left column) and resolution (right column) for R=0.6 jets in the 5.02 TeV Signal MC, in bins of y^*, in the 5.02 TeV MC. Each plot shows a selection of rapidities grouped into mid-rapidity (top row, $|y^*| < 1.2$), intermediate rapidity (middle row, $1.2 < |y^*| < 2.8$) and forward/backward rapidity ($2.8 < |y^*| < 4.4$).

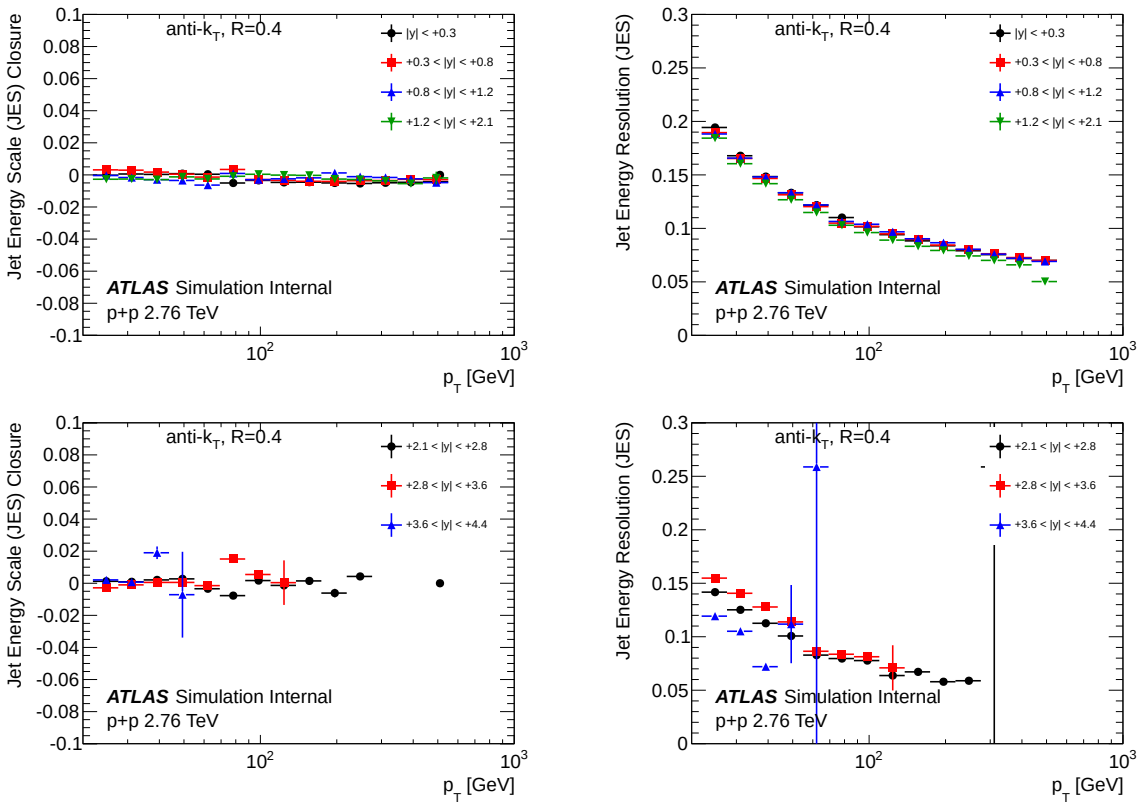


Figure 8.21: Jet energy scale closure (left column) and resolution (right column) for R=0.4 jets in the 2.76 TeV Standalone MC, in bins of y^*, in the pp 2.76 TeV MC. Each plot shows a selection of rapidities grouped into mid-rapidity (top row, $|y^*| < 2.1$) and intermediate/forward rapidity (middle row, $2.1 < |y^*| < 4.4$).

bin-by-bin *purity* P_i, which is the fraction of jets in a given reco bin which came from the same truth bin.

$$E_i = A_{ii}/T_i^{\text{matched}} \tag{8.33}$$

$$P_i = A_{ii}/R_i \tag{8.34}$$

note that A_{ii} denotes the (i,i) element and not an Einstein summation over the diagonal elements. Note that the in this definition, the bin-by-bin efficiency is defined from the subset of truth jets which were reconstruction– e.g. it does not contain the effect of the truth reconstruction

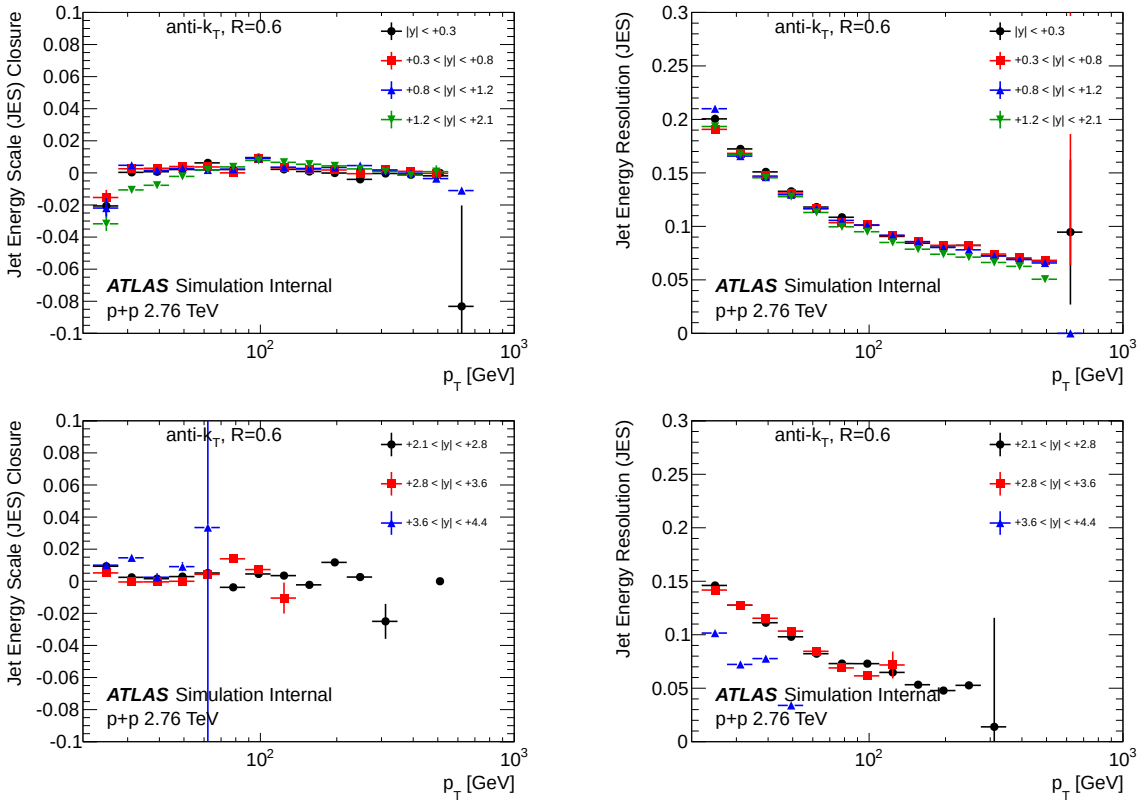


Figure 8.22: Jet energy scale closure (left column) and resolution (right column) for R=0.6 jets in the 2.76 TeV Standalone MC, in bins of y^*, in the pp 2.76 TeV MC. Each plot shows a selection of rapidities grouped into mid-rapidity (top row, $|y^*| < 2.1$) and intermediate/forward rapidity (middle row, $2.1 < |y^*| < 4.4$).

efficiency. This reconstruction efficiency (see Section 8.5.2) is corrected for separately, after the effects of the bin migration. The errors on the E_i (P_i) are determined according to a binomial distribution with sampled probability $p = E_i$ ($p = P_i$) and a total number of counts $N = T_i^{\text{matched}}$ ($N = R_i$).

E_i and P_i quantify the bin migration in the chosen bins. In the limit where they are both high ($\gtrsim 0.5$) and the input spectrum is well constrained, a set of correction factors C_i can be determined via

$$C_i = T_i/R_i \, (= P_i/E_i) \tag{8.35}$$

where the uncertainty is actually constructed for $1/C_i$, according to

$$\frac{1}{C_i} = \frac{R_i}{T_i} = \frac{A_{ii}}{T_i} + \frac{\sum_{j\neq i} A_{ij}}{T_i} \tag{8.36}$$

where the uncertainty on the A_{ii} (the jets that remained in the same bin) is binomial and the uncertainty on the quantity $\sum_{j\neq i} A_{ij}$ (the jets that migrated from outside the bin) is taken to be Poisson.

The correction factors are used as a multiplicative factor on the data to correct the measured yields for the effects of the bin migration via

$$\frac{dN_i}{dp_{\mathrm{T}}}^{\mathrm{corr}} = C_i \frac{dN_i}{dp_{\mathrm{T}}}^{\mathrm{uncorr}} \tag{8.37}$$

We would like to determine how different the correction factors in the different centralities. This is equivalent to deriving the correction factors on the R_{CP}, which is the binary collision scaled ratio of jet yields in central events relative to peripheral events. A centrality-dependent correction factor C_i^{cent} should be applied to the numerator and the denominator in Equation 8.6, via

$$R_{\mathrm{CP}}{}^{\mathrm{corr}} = \frac{C_i^{0-10\%}}{C_i^{60-90\%}} R_{\mathrm{CP}}{}^{\mathrm{uncorr}} = \frac{T_i^{0-10\%}/R_i^{0-10\%}}{T_i^{60-90\%}/R_i^{60-90\%}} R_{\mathrm{CP}}{}^{\mathrm{uncorr}} \tag{8.38}$$

where the 0-10% and 60-90% bins are used just for the sake of example, and the equation holds for other centrality bins.

These centrality dependent correction factors are derived from the Overlay MC, which contains the real effects of the p+Pb underlying event on the bin migration. That is, the transfer matrix A_{ij}^{cent} is determined separately for MC events in which the original p+Pb data event was of a given centrality class.

Normally, the ratio $C_i^{0-10\%}/C_i^{60-90\%}$ is susceptible to statistical fluctuations in the underlying truth distribution (as well as the isolation criteria removing slightly different amounts of truth, reco jet pairs in the different centralities). Therefore, the transfer matrices $A_{ii}^{\mathrm{cent,uncorr}}$ are re-weighted along the truth axis such that the underlying truth distribution T_i^{full} is the full centrality-independent truth jet spectrum before isolation cuts,

$$A_{ij}^{\mathrm{cent}} = \frac{T_i^{\mathrm{full}}}{T_i^{\mathrm{cent}}} A_{ij}^{\mathrm{cent,uncorr}} \tag{8.39}$$

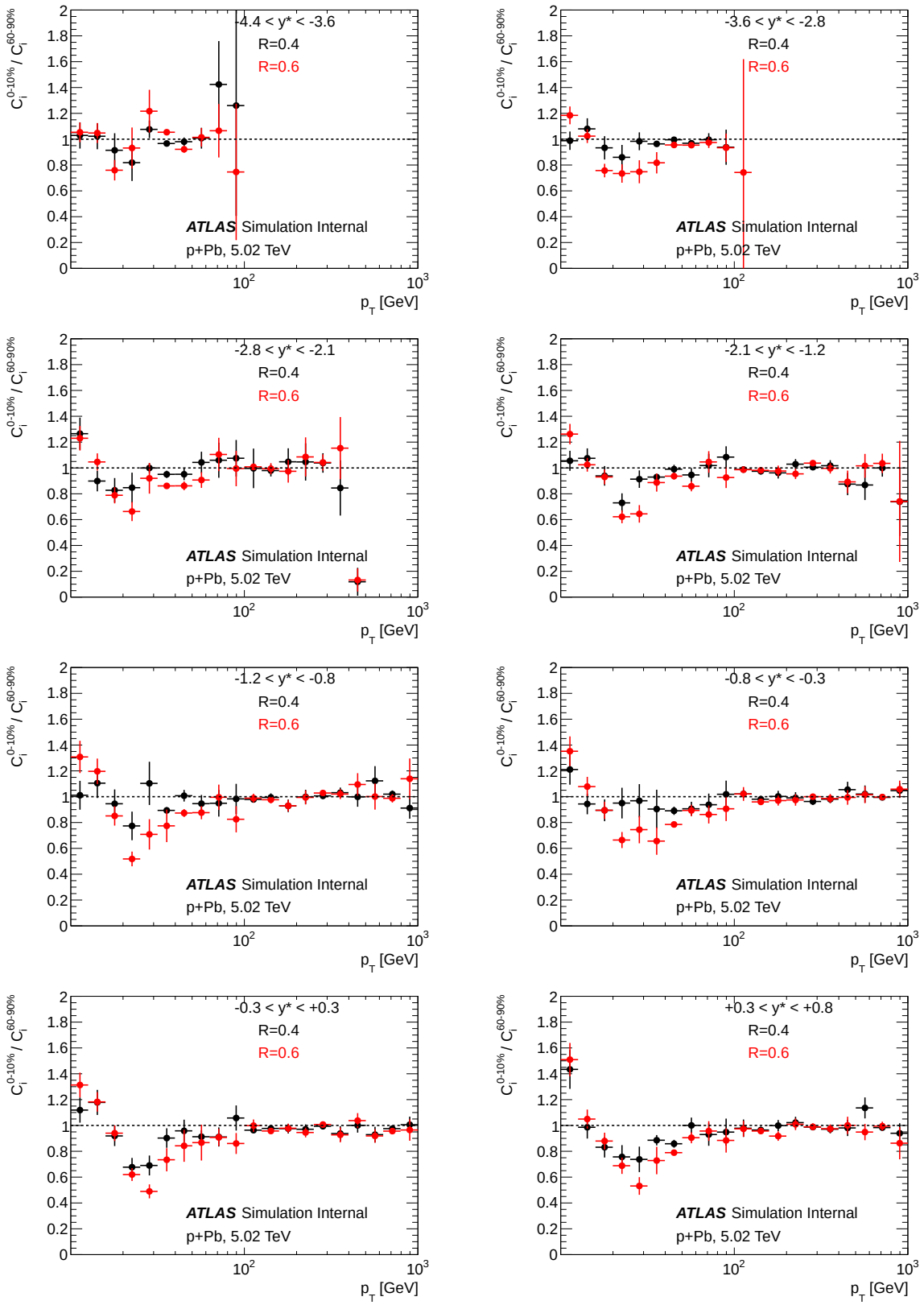


Figure 8.23: Bin-by-bin correction factors on the 0-10% / 60-90% R_{CP}, shown for R=0.4 jets (black points) and R=0.6 (red points), derived from the 5.02 TeV Overlay MC. Results are shown for the eight rapidity bins $y^* < 0.8$.

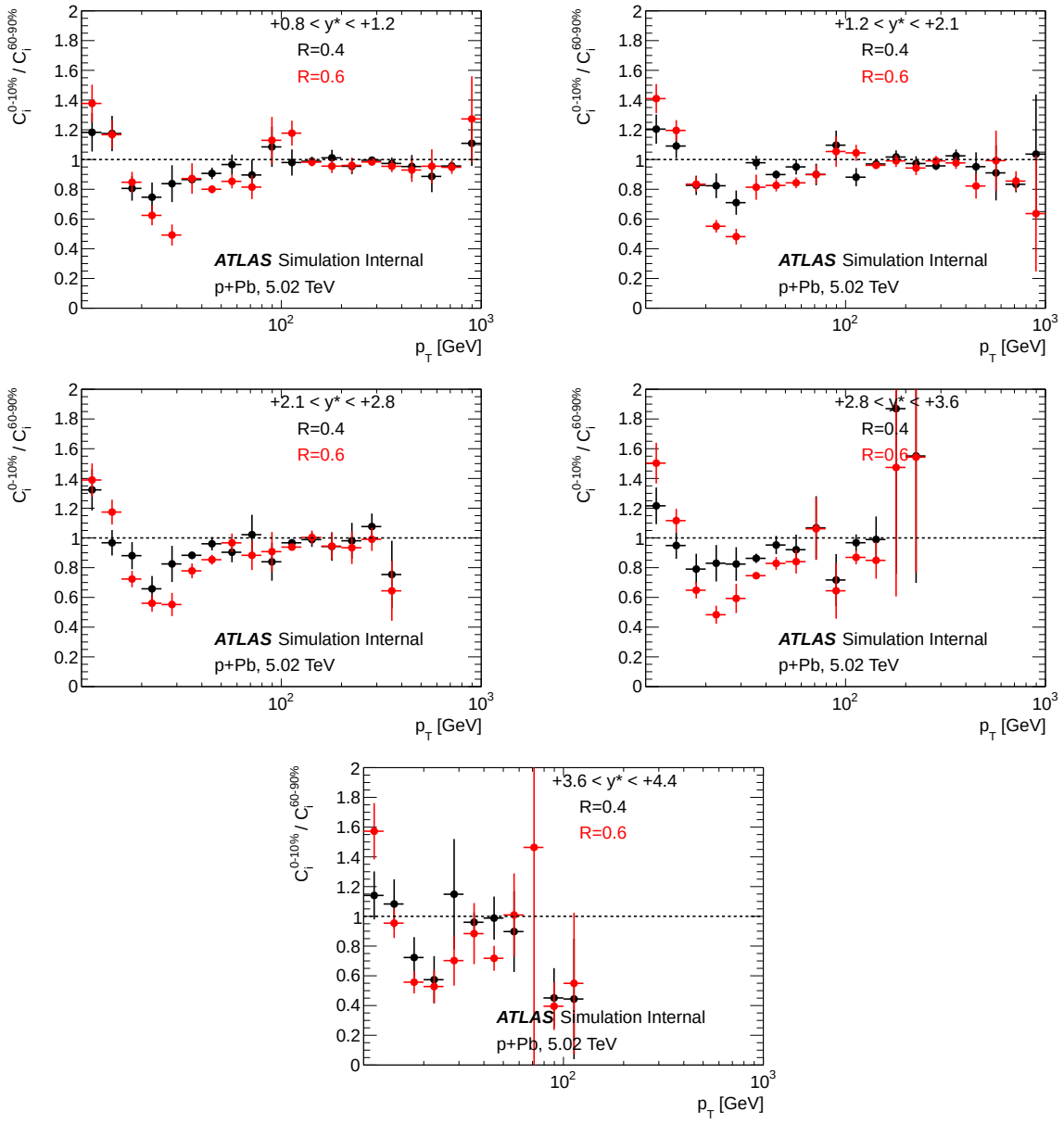


Figure 8.24: Bin-by-bin correction factors on the 0-10% / 60-90% R_{CP}, shown for R=0.4 jets (black points) and R=0.6 (red points), derived from the 5.02 TeV Overlay MC. Results are shown for the five rapidity bins $y^* > 0.8$.

By construction, the T_i^{cent} terms in the ratio in Equation 8.38 now cancel, and the correction factors on the R_{CP} reduce to

$$R_{\text{CP}}{}^{\text{corr}} = \frac{C_i^{0-10\%}}{C_i^{60-90\%}} R_{\text{CP}}{}^{\text{uncorr}} = \frac{R_i^{60-90\%}}{R_i^{0-10\%}} R_{\text{CP}}{}^{\text{uncorr}} \tag{8.40}$$

where the errors on the R_i^{cent} are statistically independent, since they come from disjoint MC events. In the p_{T} region where the correction factors on the R_{CP} are 1, the response for jets in p+Pb is centrality-independent.

The correction factors on the R_{CP} are shown in Figures 8.23 and 8.24. At high-p_{T} (e.g. $p_{\text{T}} > 100$ GeV at mid-rapidity), the corrections on the R_{CP} are ≈ 1, since at high-p_{T} the bin migration effects are the same for all centralities. At lower p_{T} (e.g. 40-80 GeV at mid-rapidity), the correction factors are < 1 and decrease with decreasing p_{T}, as the centrality-dependent differences in the JER have more of an effect. That is, this region of the R_{CP} is artificially high at the raw scale and should be corrected down. The size of this correction is larger at mid-rapidity and backward rapidities ($y^* > 1$) than it is at forward rapidities ($y^* < -1$) due to the differences in the underlying event. Furthermore, they are larger for R=0.6 jets, which are more susceptible to the underlying event. At very low p_{T} (e.g. $p_{\text{T}} < 30$ GeV), the correction factors come back up, but this is an artifact of the minimum p_{T} in the transfer matrix, since the full effects of the migration into bins near the > 10 GeV threshold are not fully captured.

The p_{T} range for which the effect of the correction factors cancels (and thus demonstrated that they are centrality independent) is determined in a systematic way. For each rapidity selection $y* < 2.8$, the lowest p_{T} value is determined above which all the correction factors are consistent with 1 within the statistical uncertainty on the MC. This value is called $p_{\text{T}}{}^{\text{cut}}$, and varies with rapidity y^* (it is smaller in the more forward bins, where the underlying event is smaller) and cone size (it is larger for R=0.6 than R=0.4 jets, since they are more suscpetible to the underlying event).

8.5.5 Correction factors on the p+Pb yields

Having determined that p+Pb jets above a certain p_{T} threshold have a detector response and bin migration effects independent of centrality, the jets above this p_{T} are corrected for detector effects from the jet energy resolution and any residual non-closure.

y^*	${p_{\mathrm{T}}}^{\mathrm{cut}}$, R=0.4	${p_{\mathrm{T}}}^{\mathrm{cut}}$, R=0.6
$-4.4 < y^* < -3.6$	25 GeV	25 GeV
$-3.6 < y^* < -2.8$	30 GeV	40 GeV
$-2.8 < y^* < -2.1$	30 GeV	60 GeV
$-2.1 < y^* < -1.2$	40 GeV	60 GeV
$-1.2 < y^* < -0.8$	80 GeV	100 GeV
$-0.8 < y^* < -0.3$	80 GeV	100 GeV
$-0.3 < y^* < +0.3$	100 GeV	100 GeV
$+0.3 < y^* < +0.8$	100 GeV	100 GeV
$+0.8 < y^* < +1.2$	100 GeV	100 GeV
$+1.2 < y^* < +2.1$	100 GeV	100 GeV

Table 8.9: Value of ${p_{\mathrm{T}}}^{\mathrm{cut}}$ in each rapidity bin and cone size for which the bin-by-bin correction factors on the p+Pb jet yields above $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$ are centrality-independent.

The raw p+Pb yields are corrected for detector effects using a bin-by-bin unfolding method. The Signal MC is used to understand the bin-by-bin purity and efficiency, and derive the correction factors. As a reminder, only p+Pb yields with $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$ as determined in the previous section are used.

The bin-by-bin purity, efficiency and correction factors are shown in Figures 8.25, 8.26, 8.27 and 8.28 for both jet R parameters and in all y^* bins. At very low p_{T} ($p_{\mathrm{T}} < 20$), the purity and efficiency are artificially high due to the low p_{T} cutoff in the transfer matrix – the migration of jets near the threshold to lower p_{T} values is not properly accounted for. For $p_{\mathrm{T}} > 20$, the purity and efficiency slowly increase, reflecting both the growing size of the p_{T} bins and also the decreasing jet energy resolution. Thus, the correction factors asymptote to a value near 1 as the growing bin size and increasing absolute resolution roughly cancel out.

Unfortunately, due to the choice of bins in which the ${p_{\mathrm{T}}}^{\mathrm{truth}}$ truth jet filtering is performed at the generator level and choice of analysis bins that lie very close to these, the raw correction factors suffer from statistical fluctuations near the regions of transition from one MC JZ subsample to another. This is especially pronounced in the JZ1 sample, which covers the range $20 < {p_{\mathrm{T}}}^{\mathrm{truth}} < 80$

GeV, which has to cover a factor of 4 change in the p_{T}. (Comparatively, the JZ2 sample covers the range $80 < {p_{\mathrm{T}}}^{\mathrm{truth}} < 200$ GeV, which is a $\times 2.5$ range.) The fluctuations are visible by eye in the purity and efficiency, usually in the bin just below or above 80 GeV. In these cases, the purity and efficiency are smoothed in the region $70 < p_{\mathrm{T}} < 100$ GeV through a fit to data points in the region. These fits are shown as dotted red and black lines in the plots on Figures 8.25-8.28.

Before the correction factors are applied to the uncorrected per-event yields, the data/MC ratio at the reconstructed p_{T} scale is examined for any p_{T} dependence. In fact, the MC underpredicts the slope of the 0-90% p+Pb yields by a small amount (a spectral index difference of about ≈ 0.15). Thus, to derive a set of correction factors appropriate for the observed p+Pb spectral shape, the transfer matrices are reweighted in each rapidity bin along the truth axis by the data/MC ratio. This way, the MC used for the unfolding has the same shape as the data and thus properly describes the effects of bin migration on a falling spectrum.

The reweighing procedure is repeated separately for the finer centrality selections (0-10% to 60-90%), since the p+Pb yields in different centrality bins have slightly different spectral shapes (in fact, as it turns out, this is a physics effect).

8.5.6 Correction factors on the *pp* spectrum

Bin-by-bin correction factors are also computed for the *pp* cross-section using the 2.76 TeV MC following the same methodology. The correction factors are determined for two sets of bins, the bins used in the default ATLAS jet analyses (so a comparison of the 2013 results can be made to the 2011 results) and the $\sqrt{s}$-scaled version of the binning used for the p+Pb yields (so that the cross-section at $\sqrt{s} = 2.76$ TeV can be used to construct the *pp* reference for the R_{pPb}). The purity, efficiency and correction factors for the former are shown in Figures 8.29 and 8.30. For the p_{T} range of interest, the bin by bin purity and efficiency are $> 50\%$, with the discontinuous behavior at high-p_{T} arising from changes in bin sizes. Similarly, Figures 8.31 and 8.32 show the same quantities but for the second set of binning.

Before the correction factors are applied to the data, the data/MC ratio at the reconstructed p_{T} scale is examined for any p_{T} dependence. As can be seen in Figures B.16 and B.17 in the Appendix, the MC underpredicts the slope of reconstructed data. Just as with the p+Pb correction factors, the transfer matrices for the *pp* unfolding are reweighted in each rapidity bin along the truth axis

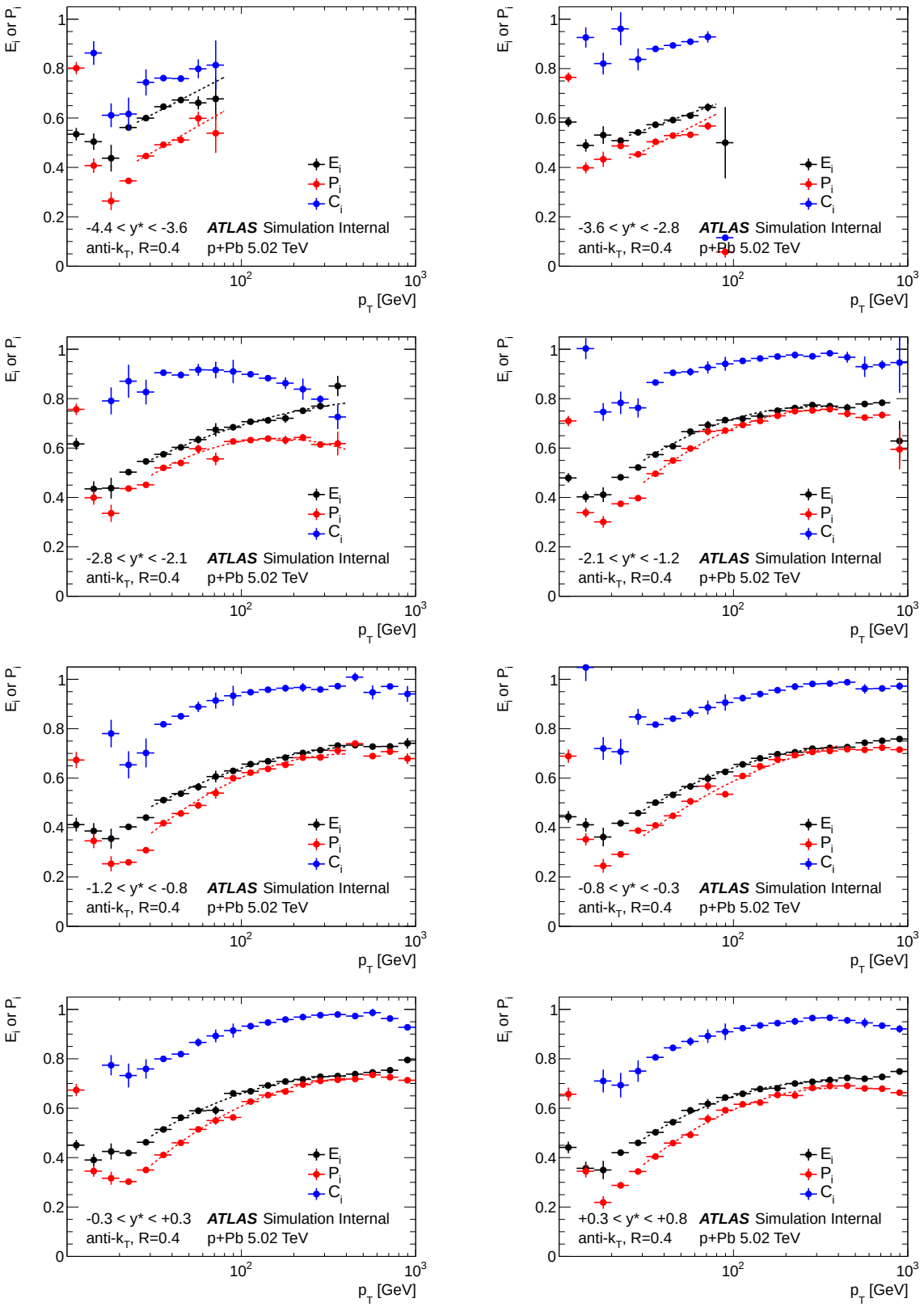


Figure 8.25: Bin-by-bin efficiency E_i (black points) and purity P_i (red points) for R=0.4 jets. Results are shown for the eight rapidity bins $y^* < 0.8$.

by the data/MC ratio. This way, the MC used for the unfolding has the same shape as the data and thus properly describes the effects of bin migration on a falling spectrum.

8.6 Jet Energy Scale and Resolution

8.6.1 Fluctuation analysis

As a check on the underlying event energy subtracted from the jet (Section 8.3.3) and the contribution to jet energy resolution from event-to-event fluctuations of the underlying event E_{T} (Section 8.5.3), an independent study of the mean and fluctuations of the underlying event E_{T} was performed. This study follows what was done in a similar study in Pb+Pb[274].

First, the calorimeter is segmented into 14 $\Delta\eta = 0.7$ slices, from $\eta = -4.9$ to $+4.9$. Second, each η-slice is segmented into non-overlapping $\Delta\phi = 7 \times \frac{\pi}{32} \approx 0.687$ regions, called *windows*, which are each a 7×7 square of 0.1×0.1 calorimeter towers. The result is 9 non-overlapping windows per η slice. The windows were chosen so that the size of the window area corresponds to that of a circular R=0.4 anti-k_{T} jet ($\pi(0.4)^2 \approx 0.50$).

Then, the EM-scale transverse energy in each window (${E_{\mathrm{T}}}^{7\times7}$) is recorded in many p+Pb events, as a function of the E_{T} in the Pb-going FCal ($\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$, which is used to determine event centrality). $L_{\mathrm{int}} = 0.5\mu\mathrm{m}^{-1}$ of period A p+Pb data was used for this study. Figure 8.33 shows the scatterplot of the ${E_{\mathrm{T}}}^{7\times7}$ plotted against $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ for windows in select η slices. In every η slice, the mean ${E_{\mathrm{T}}}^{7\times7}$ increases monotonically with increasing $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

To determine the mean $\left(\left\langle {E_{\mathrm{T}}}^{7\times7}\right\rangle\right)$ and standard deviation $\left(\sigma_{{E_{\mathrm{T}}}^{7\times7}}\right)$ of the window E_{T} for events within the p+Pb centrality bins, the following procedure is used. First, the total $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in the event sample, $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}})$ is determined. Second, the mean and standard deviation of the ${E_{\mathrm{T}}}^{7\times7}$ are determined in fine bins of $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ (in this case, $\Delta\Sigma E_{\mathrm{T}}^{\mathrm{Pb}} = 1$ GeV), and are denoted $\left\langle {E_{\mathrm{T}}}^{7\times7}\right\rangle(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}};\eta)$ and $\sigma_{{E_{\mathrm{T}}}^{7\times7}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}};\eta)$, where the η dependence is a reminder that this procedure is performed separately for windows in each η slice. Third, the centrality-dependent mean and standard deviation are assembled by taking the $dN/d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$-weighted average of the $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$-dependent mean and standard deviation for $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ values which fall into that centrality bin. In this way, the standard deviation in each centrality bin is not artificially large due to the change in the mean window E_{T} in events within the centrality bin.

Formally, the mean and standard deviation of the ${E_{\mathrm{T}}}^{7\times7}$ in p+Pb events with centrality *cent*, where the centrality bin is defined by $\Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{low}} < \Sigma E_{\mathrm{T}}^{\mathrm{Pb}} < \Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{high}}$, is

$$\left\langle {E_{\mathrm{T}}}^{7\times7} \right\rangle (cent;\eta) = \sum_{\Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{low}} < \Sigma E_{\mathrm{T}}^{\mathrm{Pb}} < \Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{high}}} \frac{dN}{d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}} \left\langle {E_{\mathrm{T}}}^{7\times7} \right\rangle (\Sigma E_{\mathrm{T}}^{\mathrm{Pb}};\eta) \tag{8.41}$$

$$\sigma_{{E_{\mathrm{T}}}^{7\times7}}(cent;\eta) = \sum_{\Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{low}} < \Sigma E_{\mathrm{T}}^{\mathrm{Pb}} < \Sigma {E_{\mathrm{T}}^{\mathrm{Pb}}}^{\mathrm{high}}} \frac{dN}{d\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}} \sigma_{{E_{\mathrm{T}}}^{7\times7}}(\Sigma E_{\mathrm{T}}^{\mathrm{Pb}};\eta) \tag{8.42}$$

Figure 8.35 shows the final $\left\langle {E_{\mathrm{T}}}^{7\times7} \right\rangle$ and $\sigma_{{E_{\mathrm{T}}}^{7\times7}}$ as a function of window η and event centrality. Within each η slice, the mean window E_{T} increases from peripheral to central events, as expected. The difference between peripheral and central events becomes larger for η slices closer to the centrality-determining FCal ($3.1 < \eta < 4.9$). Similarly, the standard deviation is ordered in centrality, being larger in more central events and smaller in peripheral events. The standard deviation also increases at mid-rapidity and decreases at large values of η. This serves as a cross-check on the subtracted E_{T} in jets as a function of η and centrality (see Section 8.3.3), which display the same pattern.

This cross-check has also been repeated for 11x11 windows, chosen to approximately correspond to the area of circular R=0.6 jets ($\pi(0.6)^2 \approx 1.13$). This is done in eight η slices of width $\Delta\eta = 1.1$ from $\eta = -4.4$ to $\eta = +4.4$, with only five windows fitting in each $\Delta\eta$ slice.

Figure 8.34 shows the scatterplot of the ${E_{\mathrm{T}}}^{7\times7}$ plotted against $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ for windows in select η slices. Figure 8.36 shows the final $\left\langle {E_{\mathrm{T}}}^{11\times11} \right\rangle$ and $\sigma_{{E_{\mathrm{T}}}^{11\times11}}$ as a function of window η and event centrality. The results look very similar to the 7×7 results in Figure 8.35, with the mean E_{T} larger by a factor $\approx (11/7)^2 = 2.5$ and the standard deviation larger by a factor $(11/7) = 1.6$.

8.6.2 Associated track jets

The HI jet energy scale has been cross-checked against that of the track jet collections. The track jets are anti-k_{T} R=0.4 jets run on $p_{\mathrm{T}} > 0.5$ GeV tracks with a nominal set of reconstruction cuts.

Each HI jet is associated with the closest track jet which is within $\Delta R < 0.4$ of the HI jet. For every such HI + track jet pair, the relative energy scale between HI and track jets is evaluated as the ratio

$$p_{\mathrm{T}}{}^{\text{track jet}}/p_{\mathrm{T}}{}^{\mathrm{HI}} \tag{8.43}$$

and is plotted as a function of the track jet p_{T}. Figure 8.37 shows the mean p_{T} ratio as a function of track jet p_{T}, in different rapidity bins, in p+Pb data. For track jets with $p_{\mathrm{T}} > 20$ GeV up to $p_{\mathrm{T}} \approx 200$ GeV, the ratio is very nearly constant at $\approx 75\%$, setting an upper limit on changes in the relative energy scale.

8.6.3 Associated topo cluster jets

The HI jet energy scale has also been cross-checked against that of the `LCTopo` jet collections, which are used in pp jet analyses. Each HI jet is associated with the closest LCTopo jet which is within $\Delta R < 0.4$ of the HI jet. For every such HI + LCTopo jet pair, the relative energy scale difference between HI and LCTopo jets is defined as

$$\frac{\Delta p_{\mathrm{T}}}{p_{\mathrm{T}}}(p_{\mathrm{T}}{}^{\mathrm{LCTopo}}) = \left(p_{\mathrm{T}}{}^{\mathrm{HI}} - p_{\mathrm{T}}{}^{\mathrm{LCTopo}}\right)/p_{\mathrm{T}}{}^{\mathrm{LCTopo}} \tag{8.44}$$

and is plotted as a function of the LCTopo jet p_{T}. There is a caveat with this comparison, which is that no official calibration for the LCTopo jets yet exists in the low-μ 2013 data. That is, residual energy scale differences between the collections are possible. Thus, the purposes of this cross-check are to see the scale of the difference and make sure that the comparison looks sensible.

Figure 8.38 summarizes the mean ($\langle \Delta p_{\mathrm{T}}/p_{\mathrm{T}} \rangle$) and standard deviation ($\sigma_{\Delta p_{\mathrm{T}}/p_{\mathrm{T}}}$) of the relative energy scale as a function of $p_{\mathrm{T}}{}^{\mathrm{LCTopo}}$ in the η^{lab} bins in p+Pb data. At small values of LCTopo jet p_{T} ($<$ 20-30 GeV), the mean energy scale is affected by the $>$ 15 GeV minimum cut on the HI jet. At intermediate values, the energy scale is *negative* due to the underlying event subtraction in HI jets which is not present in the LCTopo jets. That is, with respect to a not-subtracted LCTopo jet, the HI jet has less p_{T}. As the jet p_{T} increases, the differences due to the subtraction becomes a smaller effect, and the relative energy scale difference converges to 0.

Using the current calibration, the energy scale for R=0.4 jets at mid- and intermediate- rapidities is $<$ 1-2% between the two collections. It is somewhat worse in the forward regions, asymptoting to $< 5\%$ at high-p_{T}. The lack of an UE subtraction in LCTopo jets has a larger effect for R=0.6 jets than for R=0.4 jets, since the effect of the p_{T}-feeding in the LCTopo jets scales with the area of the jet.

8.7 Results

8.7.1 Centrality dependent yields in 5.02 TeV p+Pb collisions

As noted above, measuring the centrality dependence of jet yields requires the selection of crossings with one and only one inelastic p+Pb event. Thus, the construction of the per-event jet yields must be done in a different manner than the measurement of the cross-section in pp analyses (where the cross-section for each trigger is simply $(1/L_{int})(dN_{\rm jet}/dp_{\rm T})$ where L_{int} is the total integrated luminosity for the trigger in the selected runs). For $p_{\rm T}$ bins populated from MB events, the per-event yield is calculated in a straightforward manner by

$$N_{\rm jet}/N_{\rm evt} = \frac{\sum_{lbn} N_{{\rm jet}_{(lbn)}}}{\sum_{lbn} N^{evt,MB}_{(lbn)}} \tag{8.45}$$

where $N_{{\rm jet}_{(lbn)}}$ is the number of jets recorded in MB events in the given lumiblock, and $N^{evt,MB}_{(lbn)}$ is the number of MB events in the lumiblock. The situation is more complicated for $p_{\rm T}$ bins which are filled from a jet trigger. In a given lumiblock, the per-event yield is instead,

$$(N_{\rm jet}/N_{\rm evt})_{(lbn)} = N_{{\rm jet}_{(lbn)}} \times P^{\rm trigger}_{(lbn)} \Big/ N^{\rm evt,MB}_{(lbn)} \times P^{\rm MB}_{(lbn)} \tag{8.46}$$

where $P^{\rm trigger}_{(lbn)}$ and $P^{\rm MB}_{(lbn)}$ are the prescales of the jet trigger and the MB trigger in the lumiblock, respectively. That is, $N^{\rm evt,MB}_{(lbn)} \times P^{\rm MB}_{(lbn)}$ is the total number of (unprescaled) p+Pb events sampled by the DAQ in the lumiblock and $N_{{\rm jet}_{(lbn)}} \times P^{\rm trigger}_{(lbn)}$ are the effective number of (prescale-corrected) jets in the lumiblock. The sum over lumiblocks can be done in two ways, either via

$$N_{\rm jet}/N_{\rm evt} = \frac{\sum_{(lbn)} N_{{\rm jet}_{(lbn)}} \times P^{\rm trigger}_{(lbn)}}{\sum_{(lbn)} N^{evt,MB}_{(lbn)} \times P^{\rm MB}_{(lbn)}} \tag{8.47}$$

which can be thought of as constructing the total pre-scale corrected jet yield and dividing by the total number of MB events. Alternately, the sum can be accomplished via

$$N_{\rm jet}/N_{\rm evt} = \frac{\sum_{(lbn)} N_{{\rm jet}_{(lbn)}}}{\sum_{(lbn)} N^{evt,MB}_{(lbn)} \times P^{\rm MB}_{(lbn)}/P^{\rm trigger}_{(lbn)}} \tag{8.48}$$

which can be thought of as measuring the yield of jets in a given trigger and then normalizing by the effective fraction of MB events which are theoretically available to be sampled by the trigger, given the prescale.

In the limit of slowly-changing trigger prescales and large statistics, the two converge to the same result. Furthermore, for the very high-p_{T} triggers (such as the j75 for which $P^{\mathrm{j75}} = 1$) and Equations 8.47 and 8.48 are identical. In this work, the method in Equation 8.48 is used.

The centrality-dependent yields are constructed in a similar manner, except that the N_{evt} and N_{jet} are made only from events within the given centrality selection.

Furthermore, proper care is taken to construct the statistical covariance in data.

The uncorrected per-event jet yields in all (0-98%, or minimum bias) p+Pb collisions are shown in Figure 8.39 in the rapidity bins. The yields are shown at the calibrated jet energy scale and are corrected for the cleaning selection efficiency, but are not corrected for other detector effects such as the jet energy resolution and residual non-closure of the jet energy scale. Figures 8.40 and 8.41 show the per-event yields in the six centrality bins, from the most central (0-10%) to most peripheral (60-90%).

The corrected yields, in the rapidity bins used in the final results and truncated above $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$ (the p_{T} range determined in MC in which the jet response is centrality-independent), are shown in Figures 8.42.

8.7.2 Jet R_{CP}

The geometric-scaled central-to-peripheral ratio R_{CP} is constructed as a function of y^* and with different selections on the central bin.

In each rapidity bin, the R_{CP} is constructed with the p_{T} region of the jet yields in which the bin migration effects and detector response are independent of the p+Pb event centrality ($p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$, see above in Section 8.5.4). As described above, the numerator and denominator of the R_{CP} are corrected for detector effects through the use of bin-by-bin correction factors. Although the same transfer matrix is used, it is reweighed for use in the numerator and the denominator to have the same shape as that observed in data. Thus, the corrected R_{CP} differs from the uncorrected R_{CP} by the ratio of reweighted correction factors. The effect of this reweighing (also discussed in the systematic uncertainty section) is small, typically 2-3% and only as large as 5-10% at very low p_{T}.

The R_{CP} is shown in Figures 8.43, 8.44 and 8.45 for R=0.4 and R=0.6 jets. The R_{CP} is constructed as in Equation 8.5, using the N_{coll} estimated from the default Glauber model[263] (and also Chapter 7).

Figures 8.46 and 8.47 shows the same data, but organized such that each plot shows the R_{CP} for a different centrality selection in the six most forward rapidity bins (from $-0.8 < y^* < -0.3$ to $-4.4 < y^* < -3.6$). Figures 8.48 and 8.49 show the same data but scale the x-axis in a y^*-dependent way. Specifically, the results are plotted against the quantity $p = p_{\mathrm{T}} \cosh \langle y^* \rangle$, where $\langle y^* \rangle$ are the midpoints of the y^* bin (e.g. $\langle y^* \rangle = -4$ for $-4.4 < y^* < -3.6$).

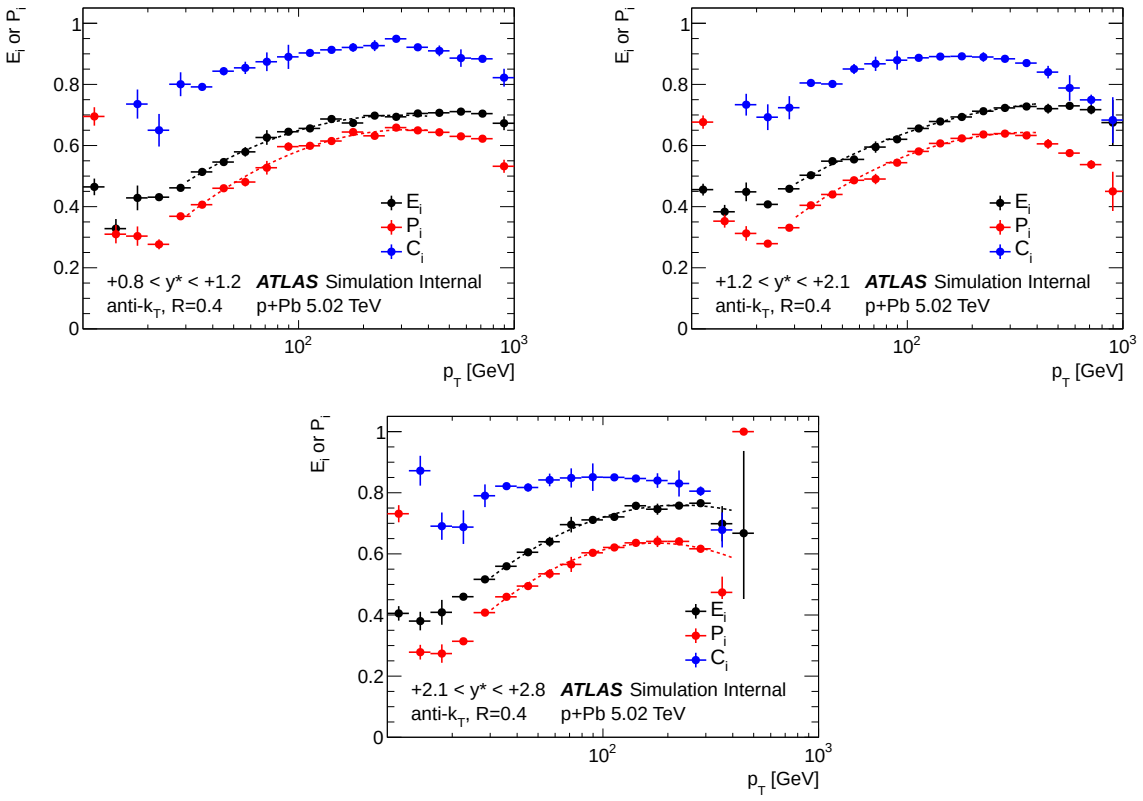


Figure 8.26: Bin-by-bin efficiency E_i (black points) and purity P_i (red points) for R=0.4 jets. Results are shown for the five rapidity bins $y^* > 0.8$.

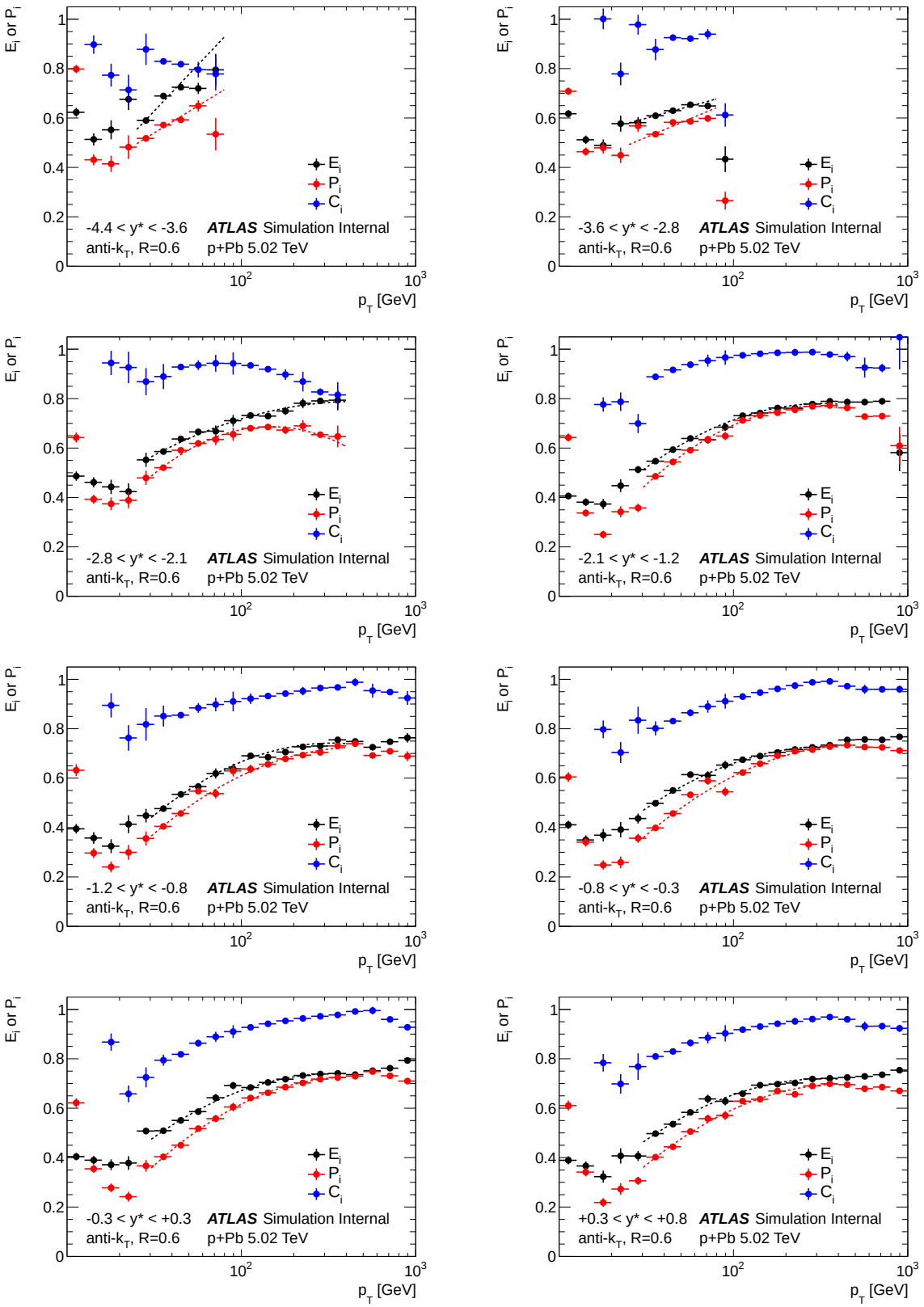


Figure 8.27: Bin-by-bin efficiency E_i (black points) and purity P_i (red points) for R=0.6 jets. Results are shown for the eight rapidity bins $y^* < 0.8$.

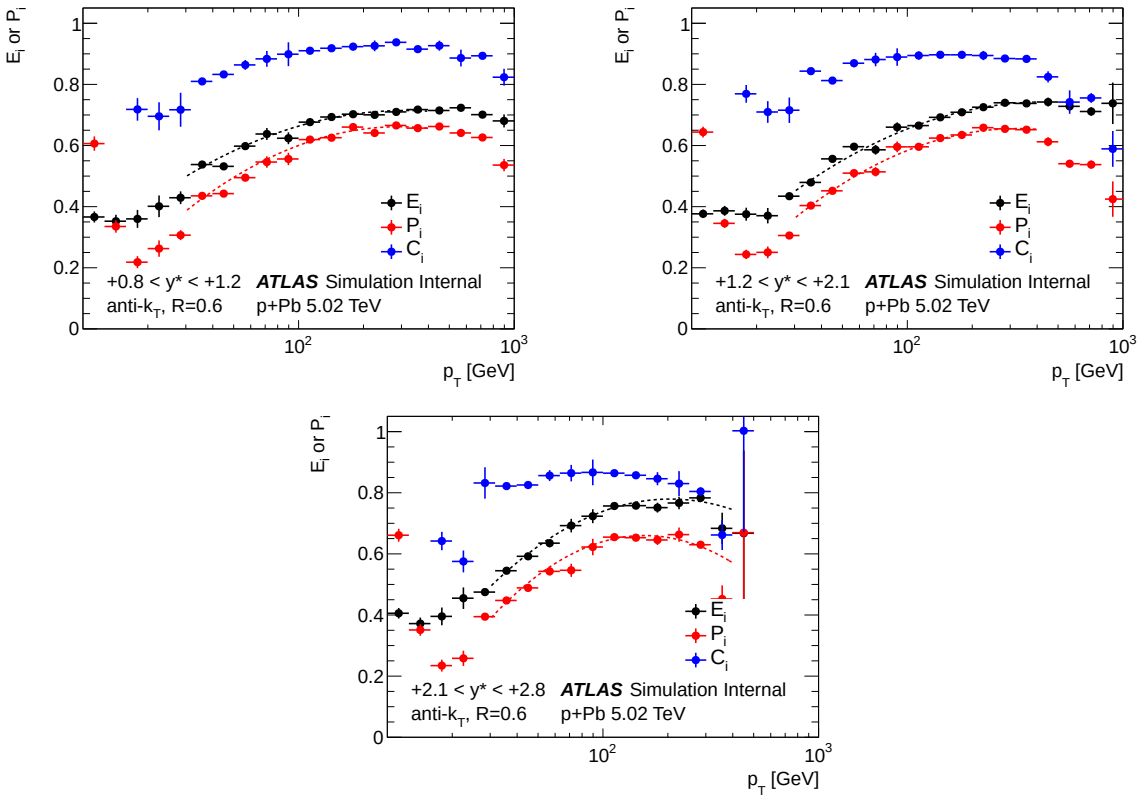


Figure 8.28: Bin-by-bin efficiency E_i (black points) and purity P_i (red points) for R=0.6 jets. Results are shown for the five rapidity bins $y^* > 0.8$.

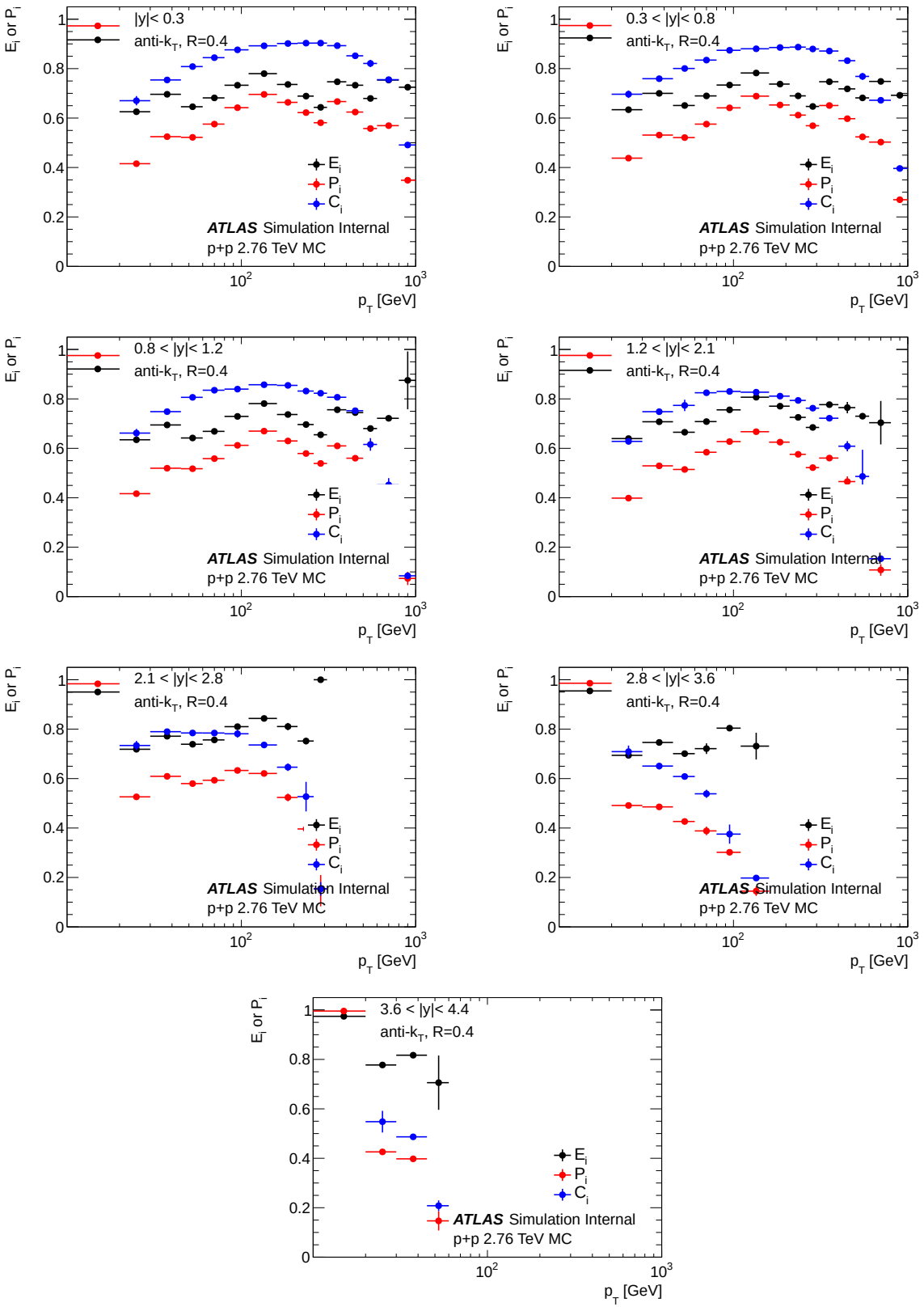


Figure 8.29: Bin-by-bin efficiency E_i (black points), purity P_i (red points) and correction factors C_i (blue points) for R=0.4 jets in the 2.76 TeV MC, in the ATLAS *pp* jet analysis bins.

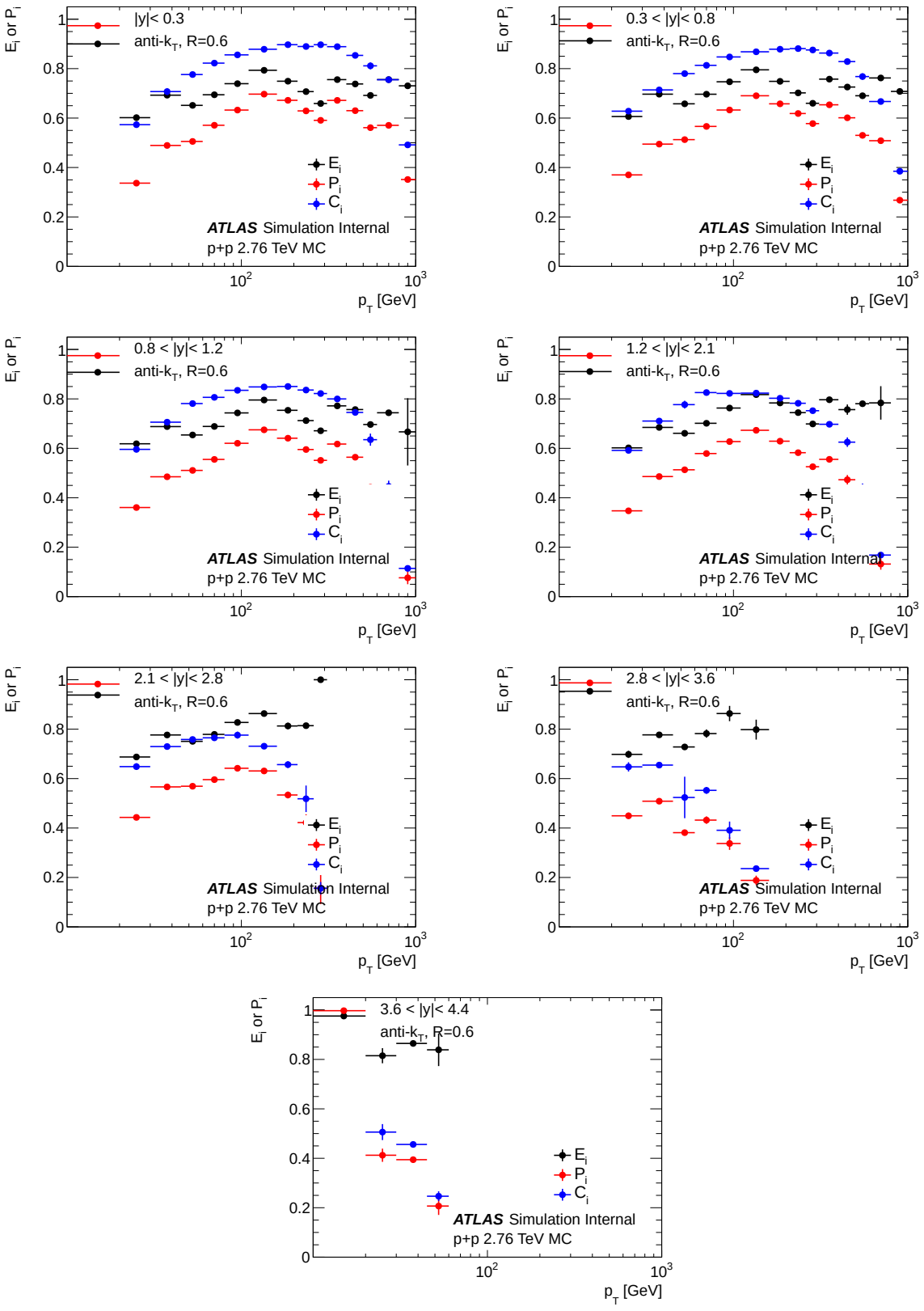


Figure 8.30: Bin-by-bin efficiency E_i (black points), purity P_i (red points) and correction factors C_i (blue points) for R=0.6 jets in the 2.76 TeV MC, in the ATLAS *pp* jet analysis bins.

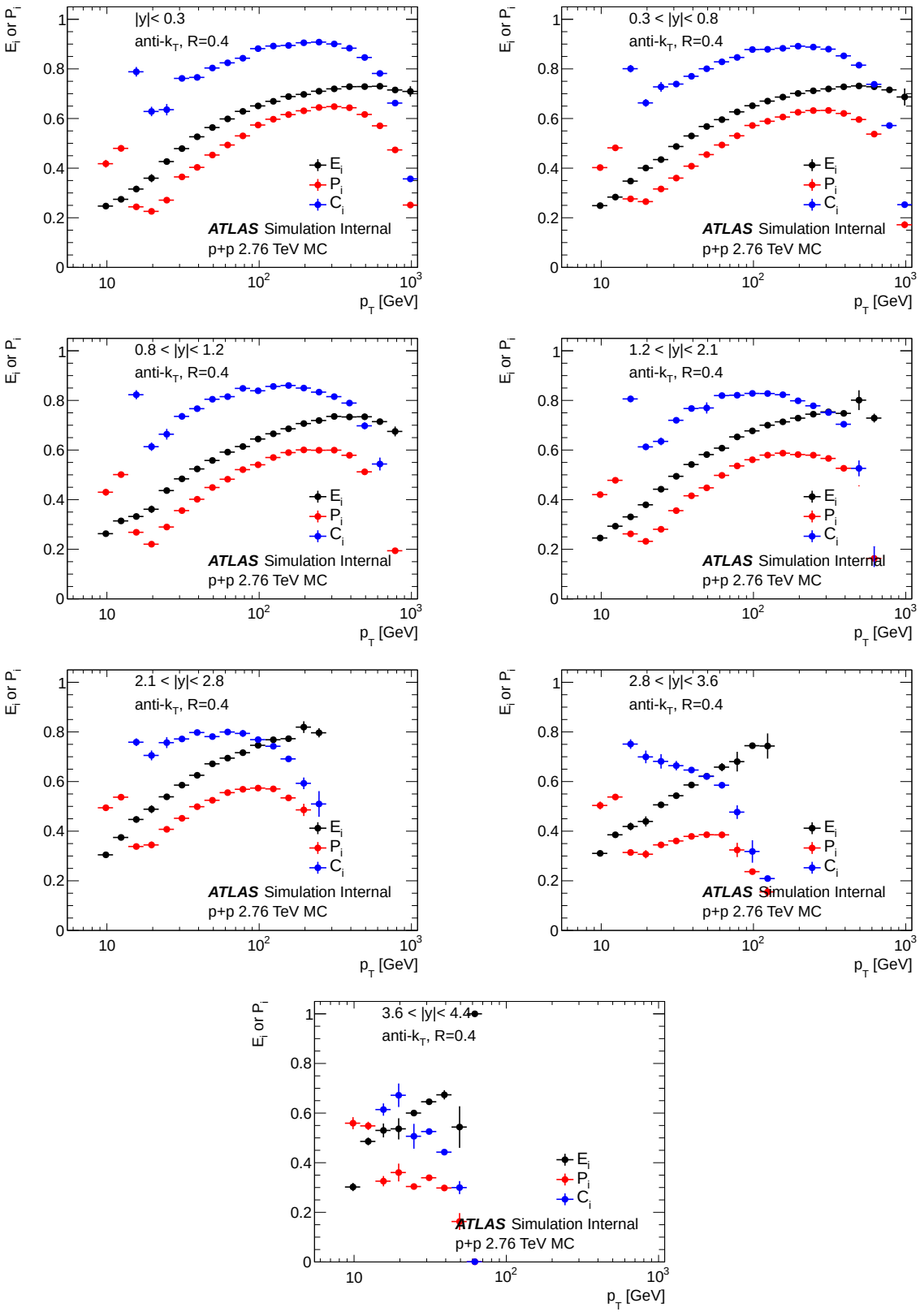


Figure 8.31: Bin-by-bin efficiency E_i (black points), purity P_i (red points) and correction factors C_i (blue points) for R=0.4 jets in the 2.76 TeV MC, in the $\sqrt{s}$-scaled binning of the p+Pb bins.

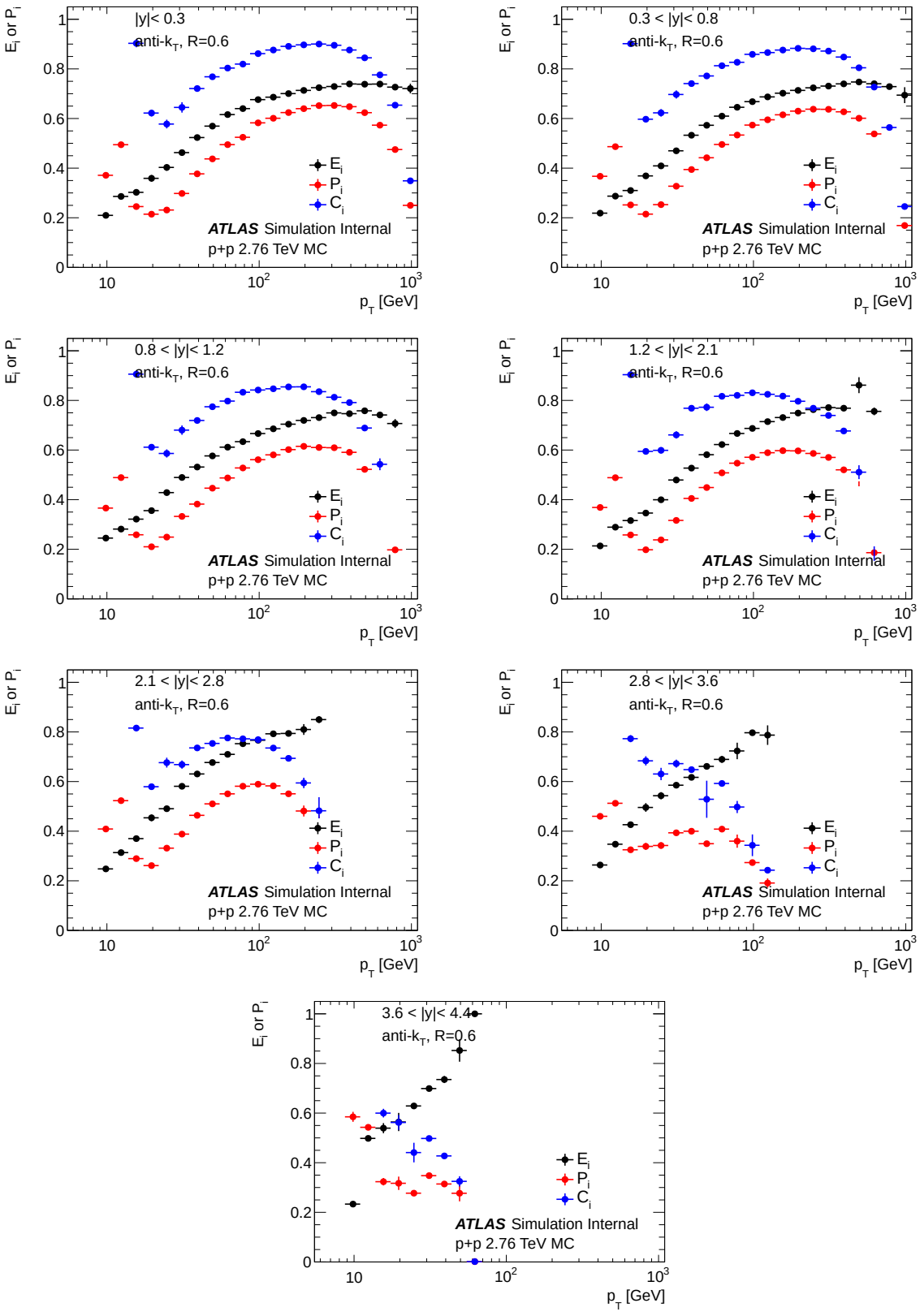


Figure 8.32: Bin-by-bin efficiency E_i (black points), purity P_i (red points) and correction factors C_i (blue points) for R=0.6 jets in the 2.76 TeV MC, in the $\sqrt{s}$-scaled binning of the p+Pb bins.

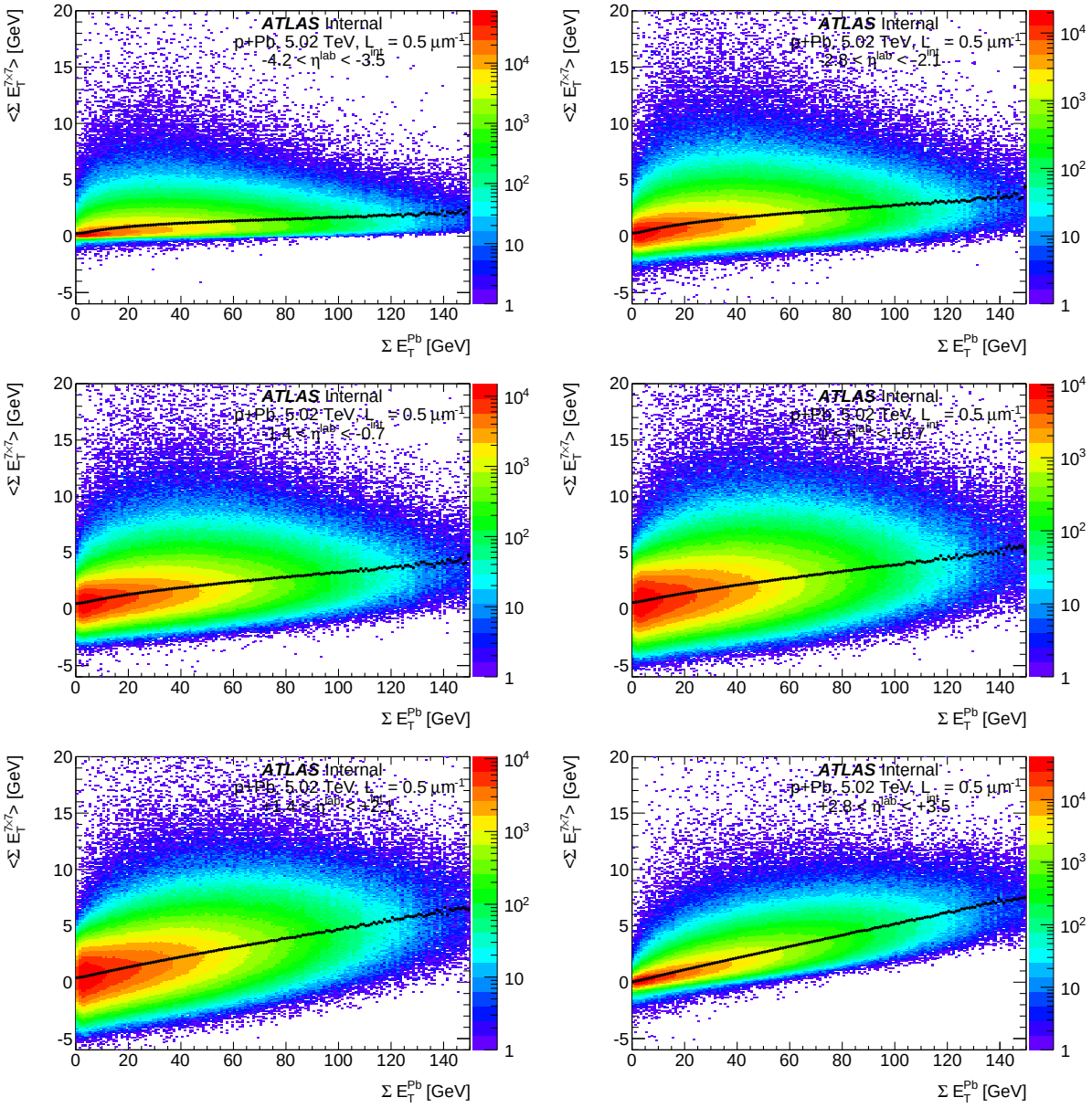


Figure 8.33: Distribution of the E_{T} in 7×7 tower windows as a function of Pb-going FCal E_{T}, for windows in select η locations shown in different plots. The trend lines show the mean window E_{T} at the given $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

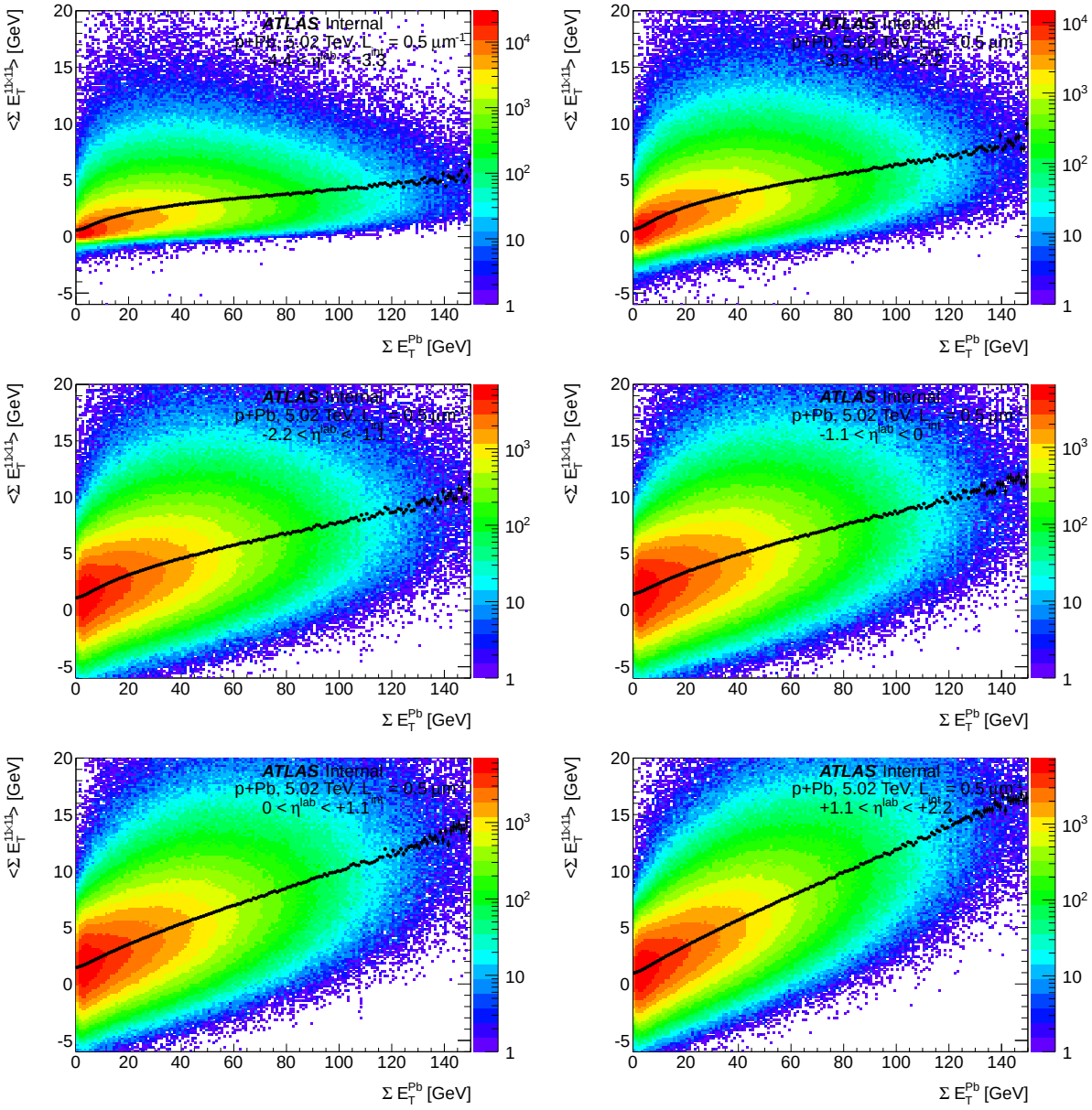


Figure 8.34: Distribution of the E_{T} in 11×11 tower windows as a function of Pb-going FCal E_{T}, for windows in select η locations shown in different plots. The trend lines show the mean window E_{T} at the given $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$.

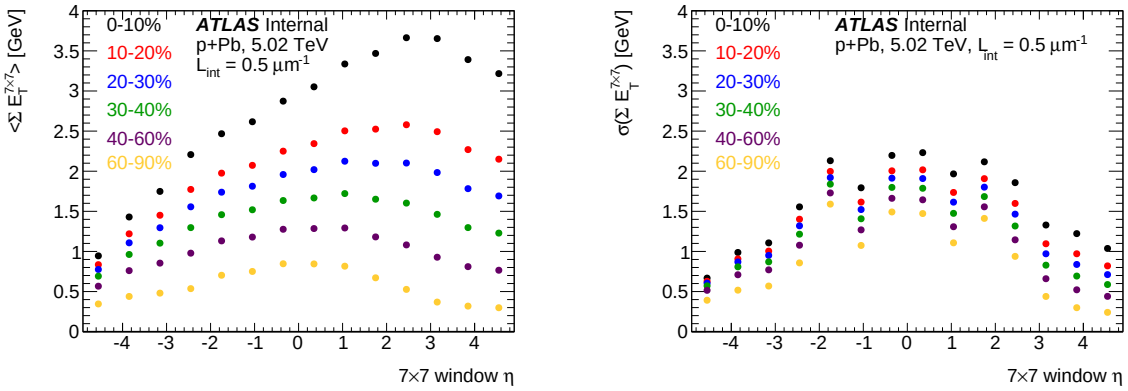


Figure 8.35: Mean (left) and standard deviation (right) of the E_{T} distribution in 7×7 tower windows as a function of the window η position, shown for different selections on the p+Pb event centrality.

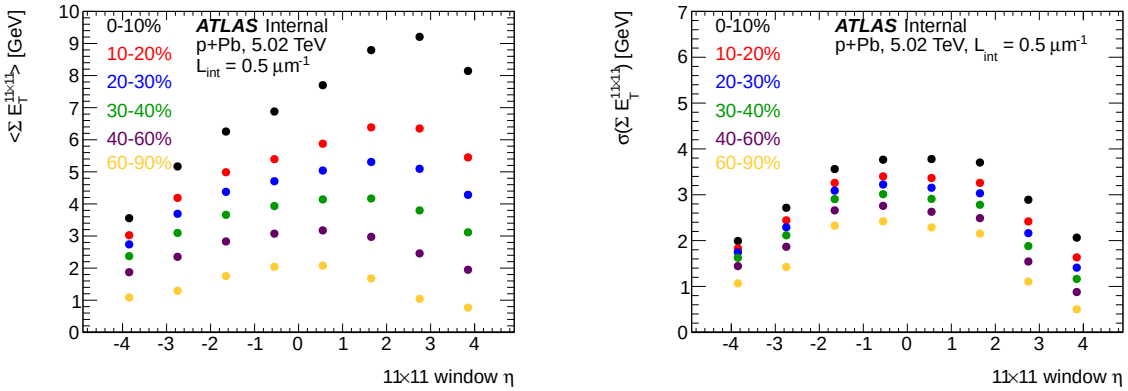


Figure 8.36: Mean (left) and standard deviation (right) of the E_{T} distribution in 11×11 tower windows as a function of the window η position, shown for different selections on the p+Pb event centrality.

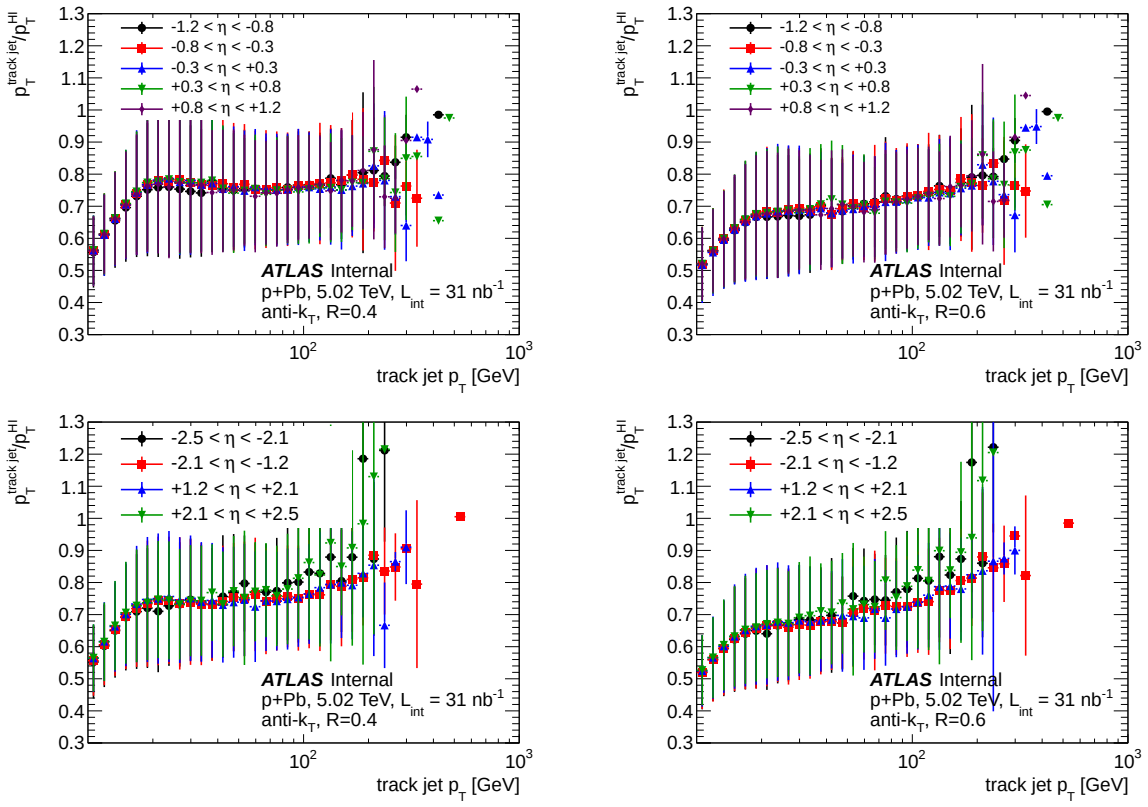


Figure 8.37: Mean ratio of the p_{T} in R=0.4 track jets to HI jets, ${p_{\mathrm{T}}}^{\text{track jet}}/{p_{\mathrm{T}}}^{\mathrm{HI}}$ in p+Pb data. The vertical error bars are the standard deviation of the ratio, and not the statistical uncertainty on the central value. Results are shown for R=0.4 HI jets (left column) and R=0.6 HI jets (right column) and separately for the five bins closest to mid-rapidity (top row), four intermediate-rapidity bins where track jet reconstruction is possible (bottom row).

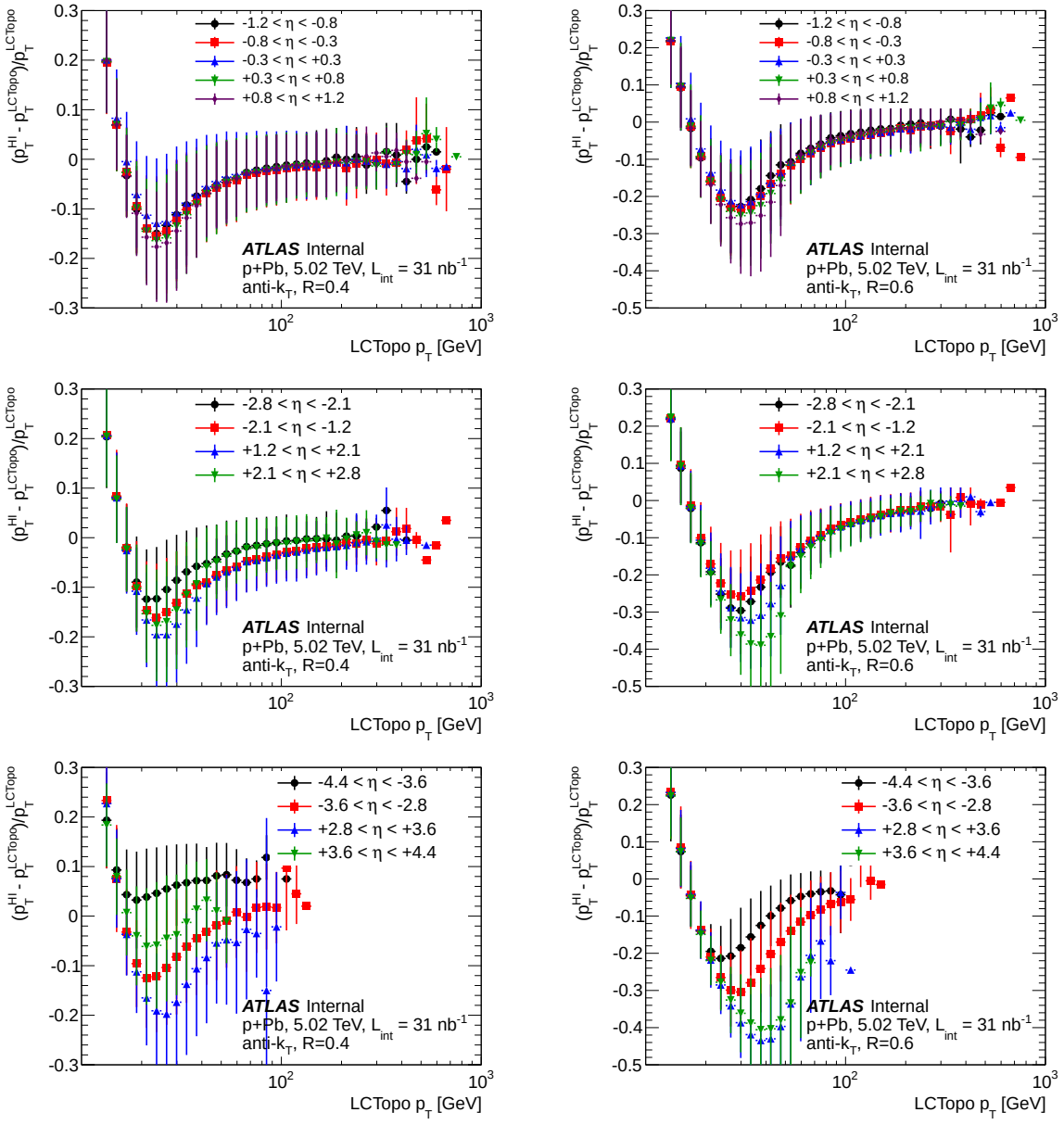


Figure 8.38: Relative jet energy scale between HI and LCTopo jets, $\Delta p_{\mathrm{T}}/p_{\mathrm{T}} = ({p_{\mathrm{T}}}^{\mathrm{HI}} - {p_{\mathrm{T}}}^{\mathrm{LCTopo}})/{p_{\mathrm{T}}}^{\mathrm{LCTopo}}$ in p+Pb data. The vertical error bars are the standard deviation of the $\Delta p_{\mathrm{T}}/p_{\mathrm{T}}$ and not the statistical uncertainty on the central value. Results are shown for R=0.4 jets (left column) and R=0.6 jets (right column) and separately for the five bins closest to mid-rapidity (top row), four intermediate-rapidity bins (middle row) and four most forward/backward bins (bottom row).

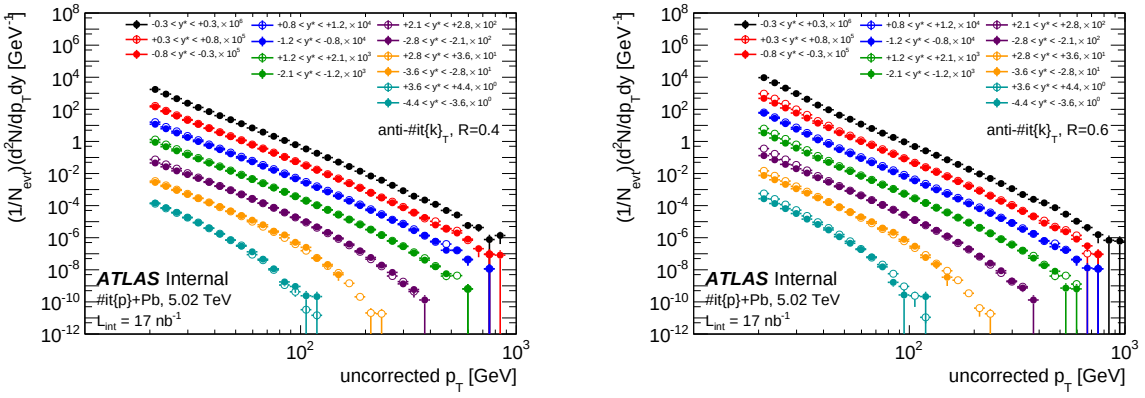


Figure 8.39: Per-event jet yields in minimum bias p+Pb collisions in y^* bins for R=0.4 (left) and R=0.6 (right) jets. Jet yields shown are at the reconstructed level, corrected for selection efficiency but before corrections for detector resolution and other effects.

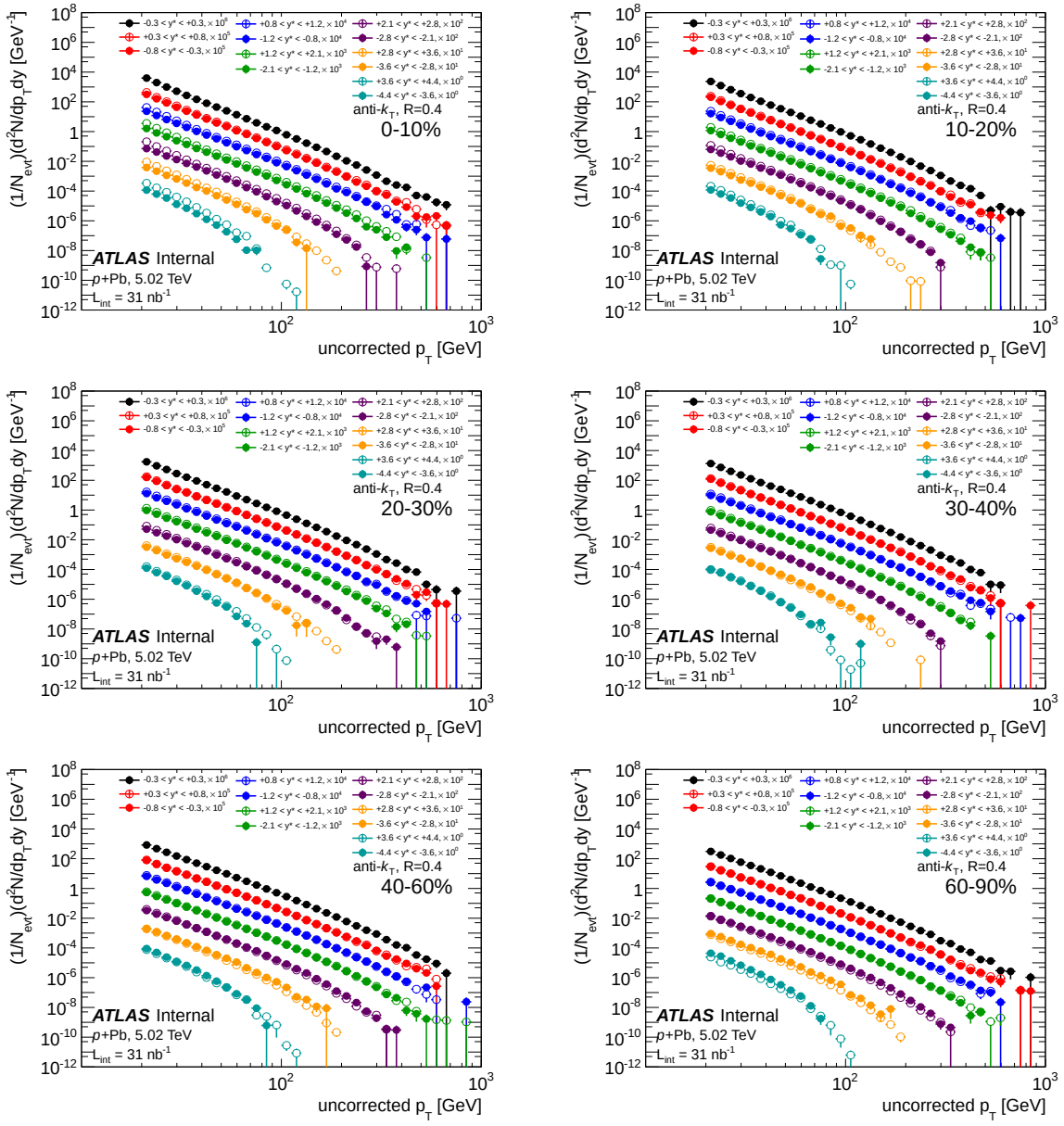


Figure 8.40: Per-event jet yields in minimum bias p+Pb collisions in y^* bins for R=0.4 jets. Each plot shows the yields in a different selection on the p+Pb event centrality. Jet yields shown are at the reconstructed level, corrected for selection efficiency but before corrections for detector resolution and other effects.

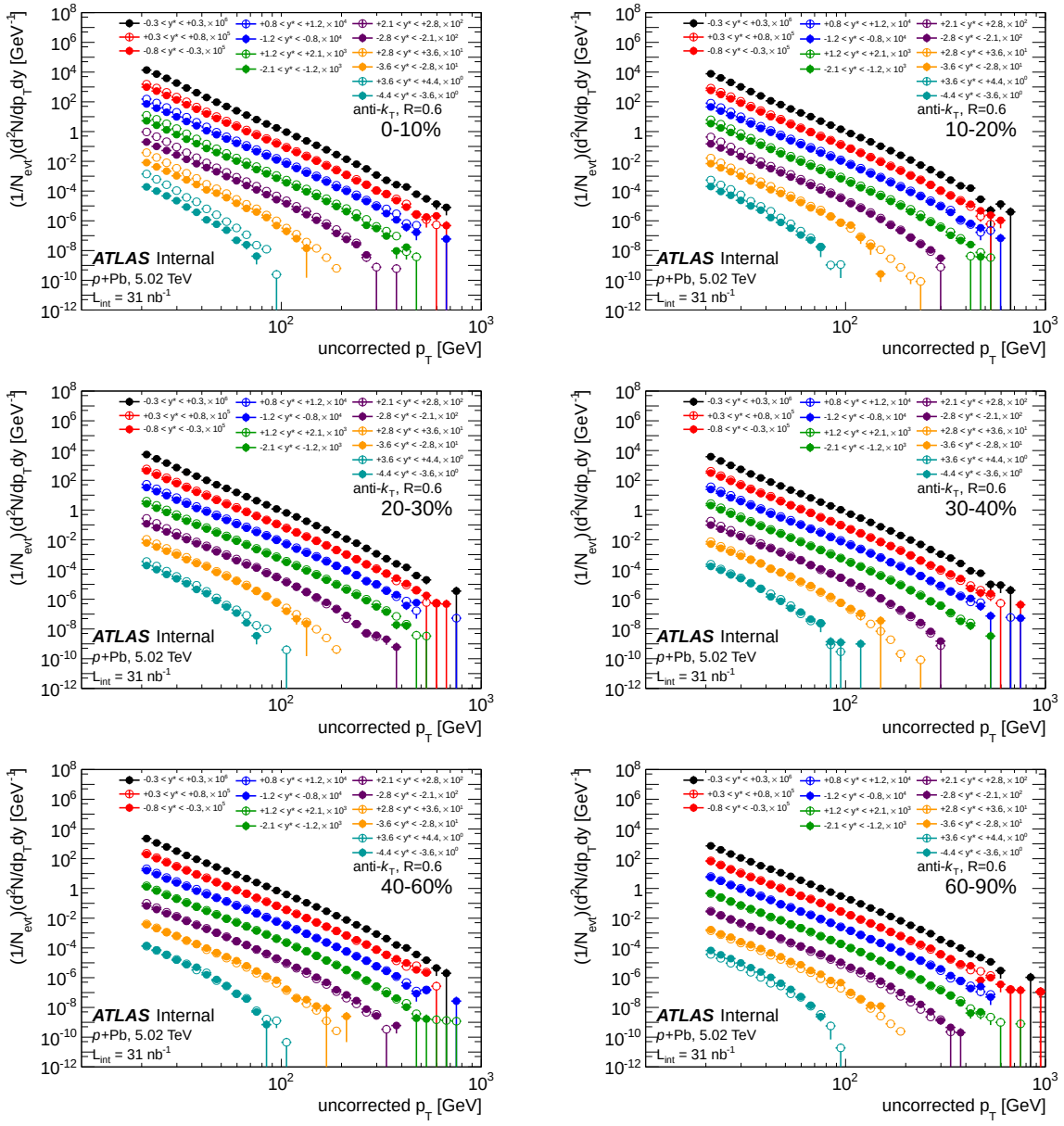


Figure 8.41: Per-event jet yields in minimum bias p+Pb collisions in y^* bins for R=0.R jets. Each plot shows the yields in a different selection on the p+Pb event centrality. Jet yields shown are at the reconstructed level, corrected for selection efficiency but before corrections for detector resolution and other effects.

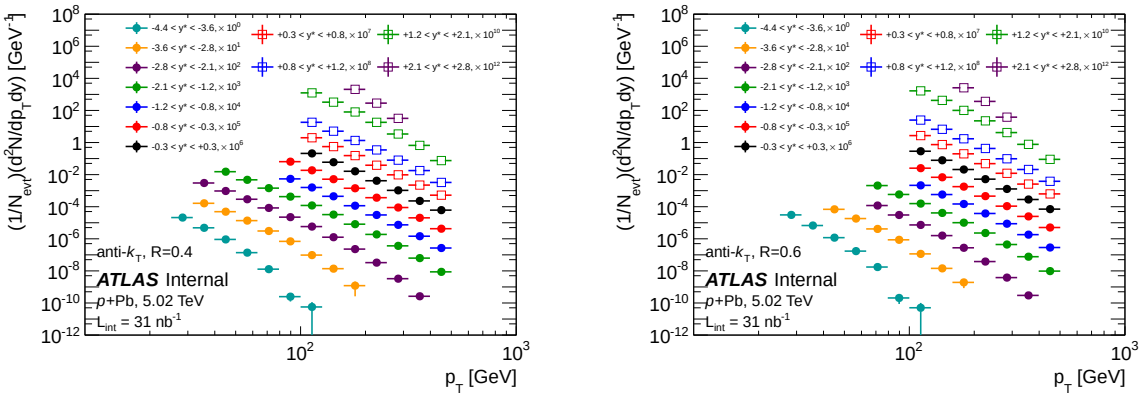


Figure 8.42: Per-event jet yields in minimum bias p+Pb collisions in y^* bins for R=0.4 (left) and R=0.6 (right) jets. Jet yields shown are corrected for all detector effects.

8.7.3 Cross-check on raw R_{CP} with track jets and pp-reconstruction jets

As a cross-check on the HI jet R_{CP} results (especially on possible centrality-dependent problems with the energy scale), RCP is also constructed as a function of two other jet collections:

- `AntiKt4LCTopo` and `AntiKt6LCTopo`, which are the anti-k_{T} R=0.4 and R=0.6 topocluster based jet collection used in pp jet analyses
- `antikt4Track`, which are the result of anti-k_{T} R=0.4 jet reconstruction run on charged tracks with $p_{\mathrm{T}} > 0.5$ GeV

Each `HIItrEM` jet is associated with the nearest `LCTopo` jet within $\Delta R < 0.4$ of the HI jet and also the nearest `Track` jet within $\Delta R < 0.4$ of the HI jet, where R=0.4 HI jets are associated with R=0.4 LCTopo and track jets, and R=0.6 HI jets with R=0.6 LCTopo and R=0.4 track jets (since there was no R=0.6 track jet reconstruction).

The R_{CP} is replotted, but as a function of the associated `LCTopo` jet p_{T}, if any, in Figures B.18 and B.19 for R=0.4 jets and Figures B.20 and B.21 for R=0.6 jets. These can be found in Appendix B.4 That is, the same collection of objects is shown (since the LCTopo jets are associated to the HI jets), but at a different energy scale. Results are shown for LCTopo jets with $p_{\mathrm{T}} > 30$ GeV, where the relative energy scale between LCTopo and HI jets is well understood (see Section 8.6.3).

Since the LCTopo jets do not have an underlying event subtraction, the R_{CP} suffers from an artificial increase due to the centrality-dependent p_{T}-feeding. Thus, the R_{CP} is somewhat higher than that in the HI jets, especially at low-p_{T}. However, the same suppression, which is systematic with η and p_{T}, can be unambiguously seen in the LCTopo jets.

The R_{CP} is replotted, but as a function of the associated R=0.4 track jet p_{T}, if any, in Figure B.22 for R=0.4 jets and Figure B.23 for R=0.6 jets, which can be found in Appendix B.4. In this case, the y^* bins are restricted to those with tracking acceptance ($-2.1 < y^*$ and $y^* < 2.8$, which approximately correspond to $-2.55 \lesssim \eta^{\mathrm{lab}}$ and $\eta^{\mathrm{lab}} < 2.35$). Although the range of y^* does not allow a comparison of the full y^* dependence of the suppression, the same p_{T}-, centrality- and y^*-dependent suppression in the R_{CP} can be observed.

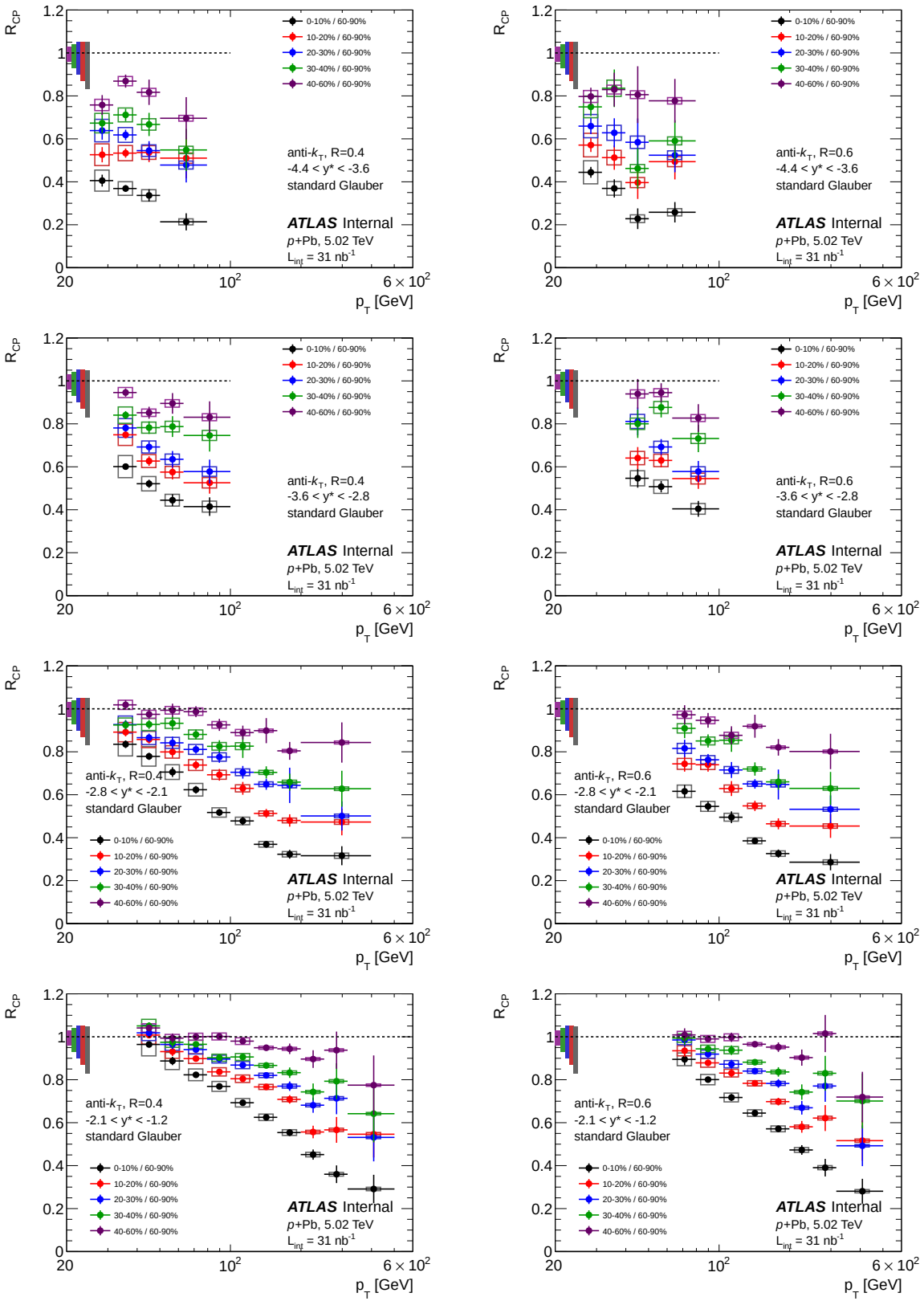


Figure 8.43: R_{CP} for R=0.4 jets (left column) and R=0.6 jets (right column) in the four rapidity bins with $-4.4 < y^* < -1.2$. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

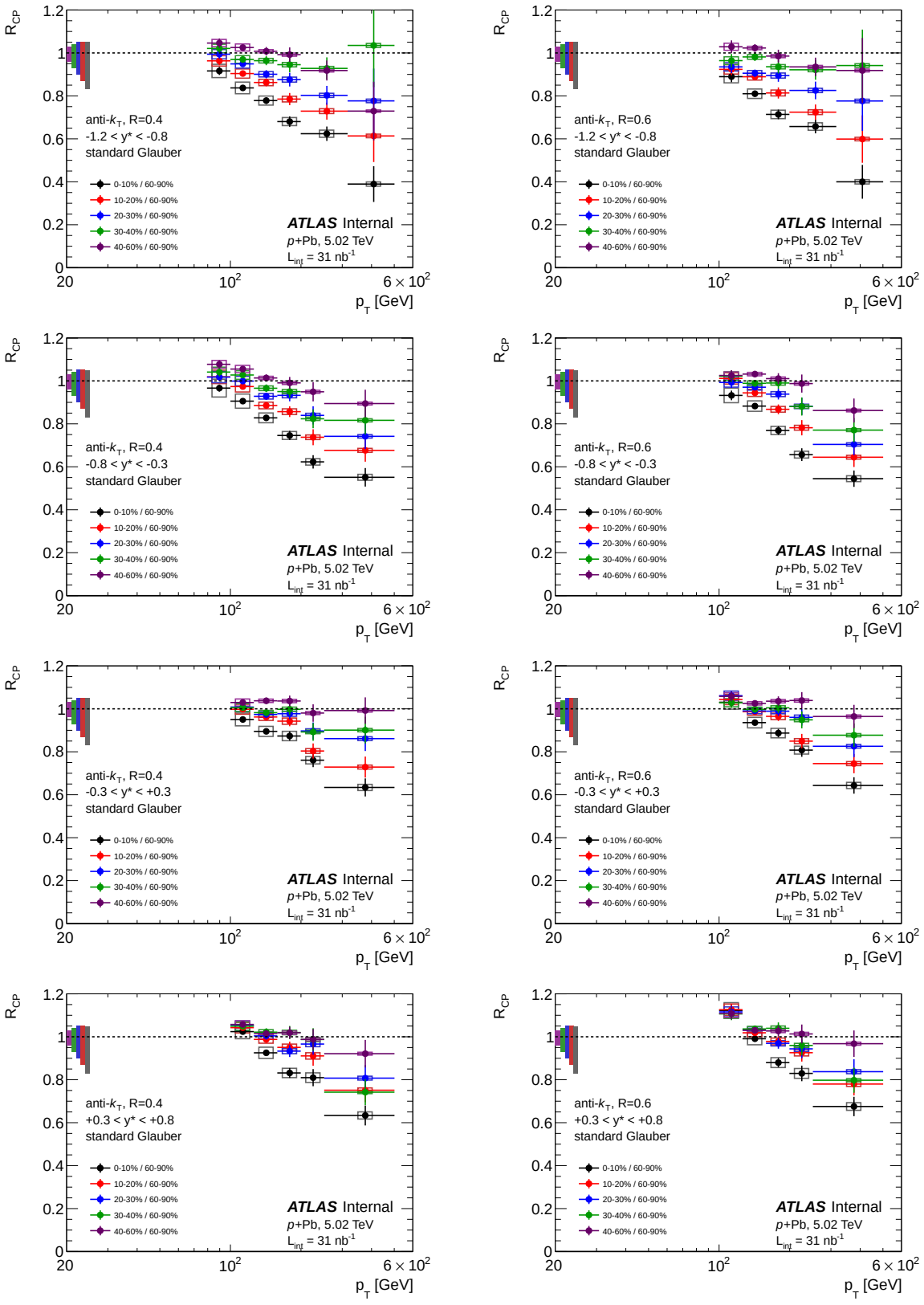


Figure 8.44: R_{CP} for R=0.4 jets (left column) and R=0.6 jets (right column) in the four rapidity bins with $-1.2 < y^* < +0.8$. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

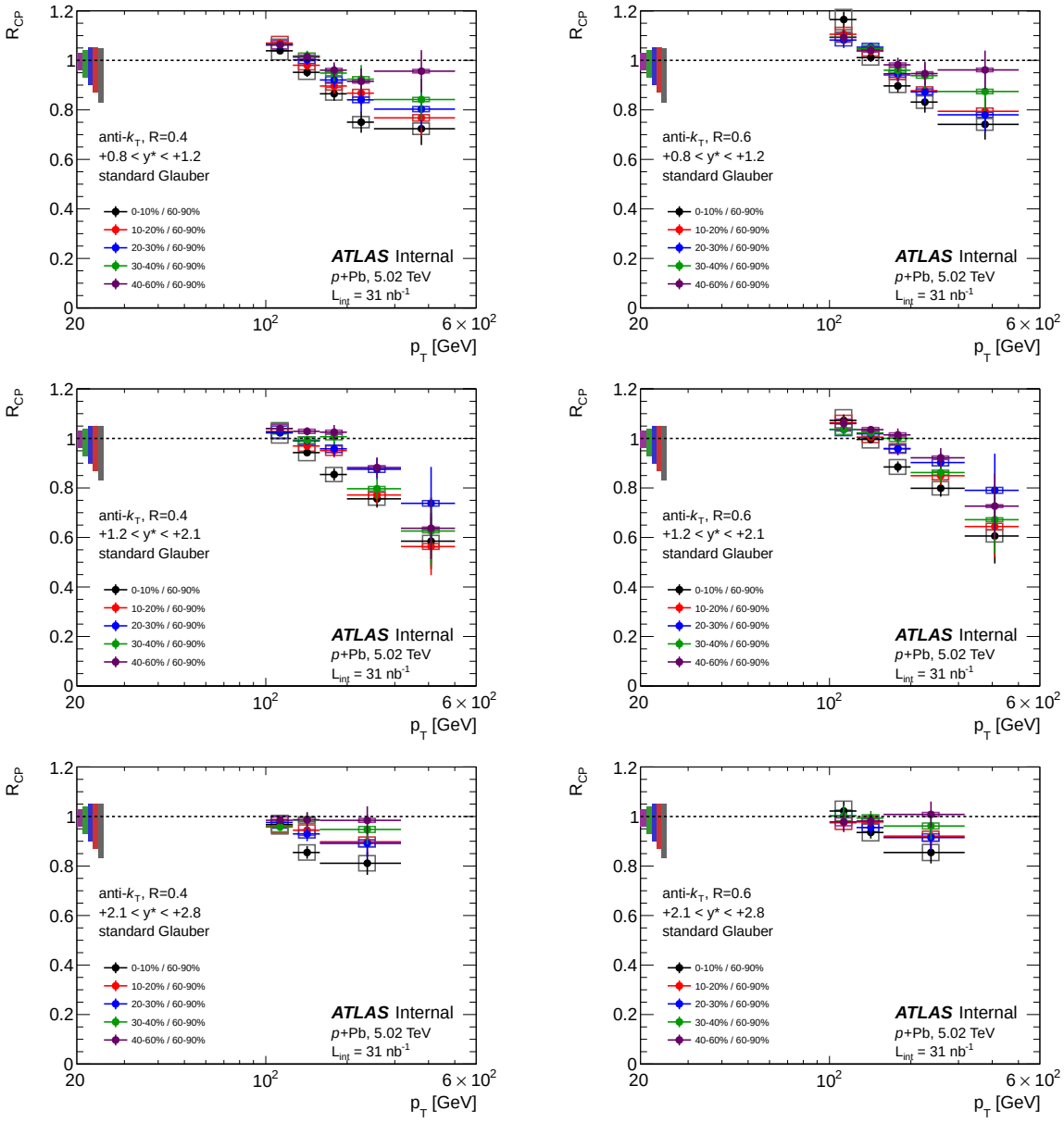


Figure 8.45: R_{CP} for R=0.4 jets (left column) and R=0.6 jets (right column) in the three rapidity bins with $+0.8 < y^* < +2.8$. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

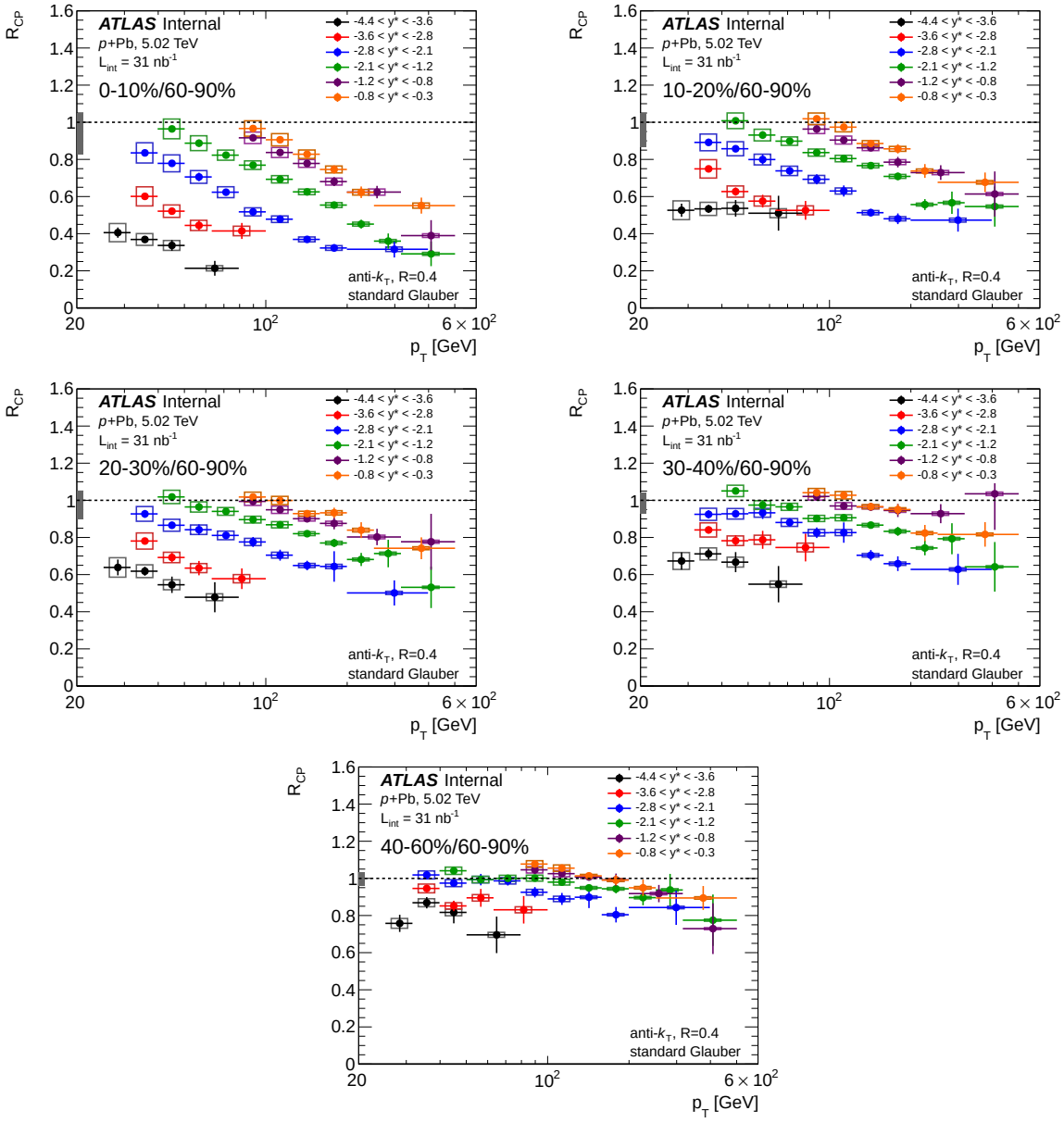


Figure 8.46: R_{CP} for R=0.4 jets in the five centrality selections, showing the six rapidity bins with $y^* < -0.3$ (the p-going direction). The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

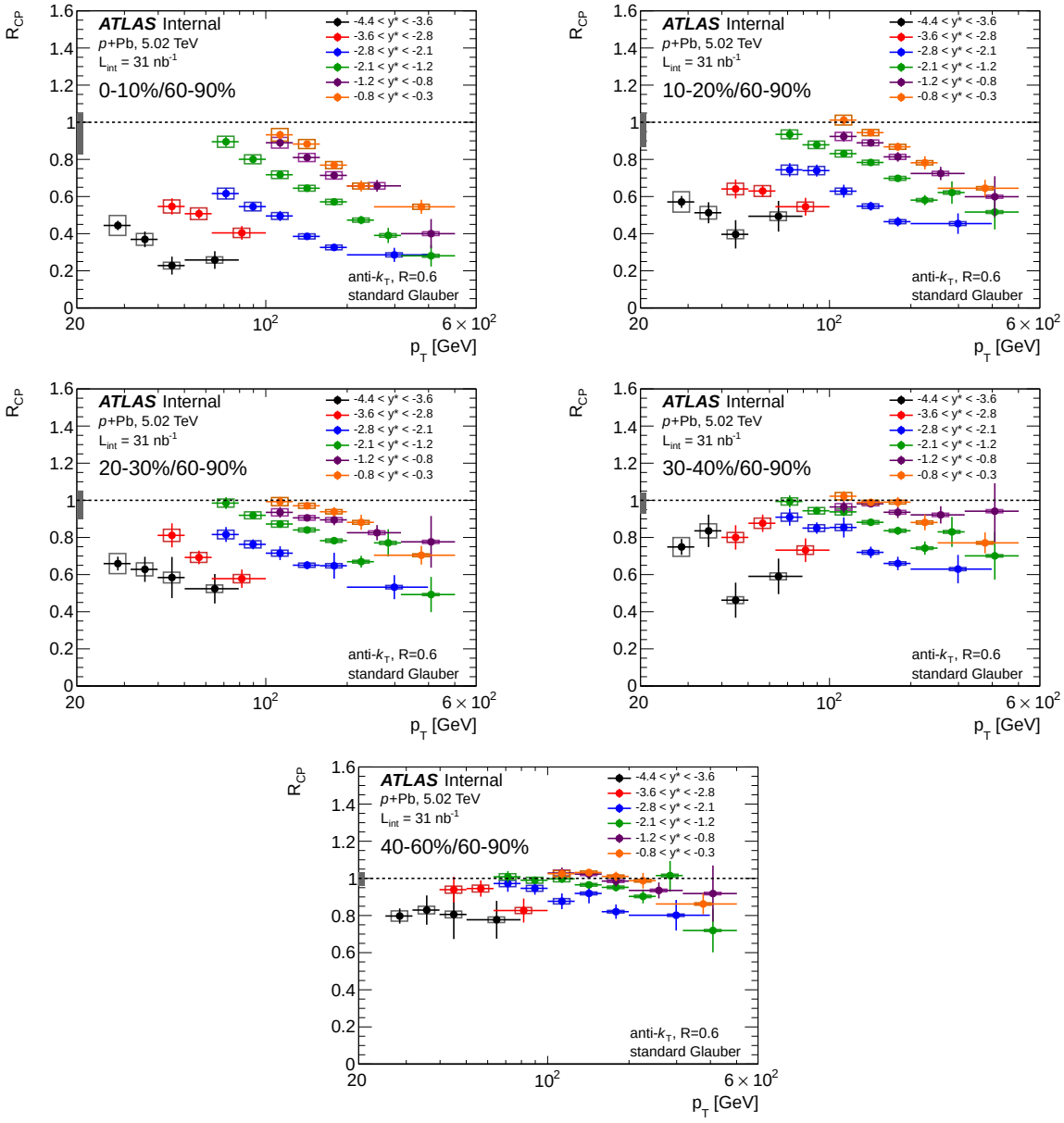


Figure 8.47: R_{CP} for R=0.6 jets in the five centrality selections, showing the six rapidity bins with $y^* < -0.3$ (the p-going direction). The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

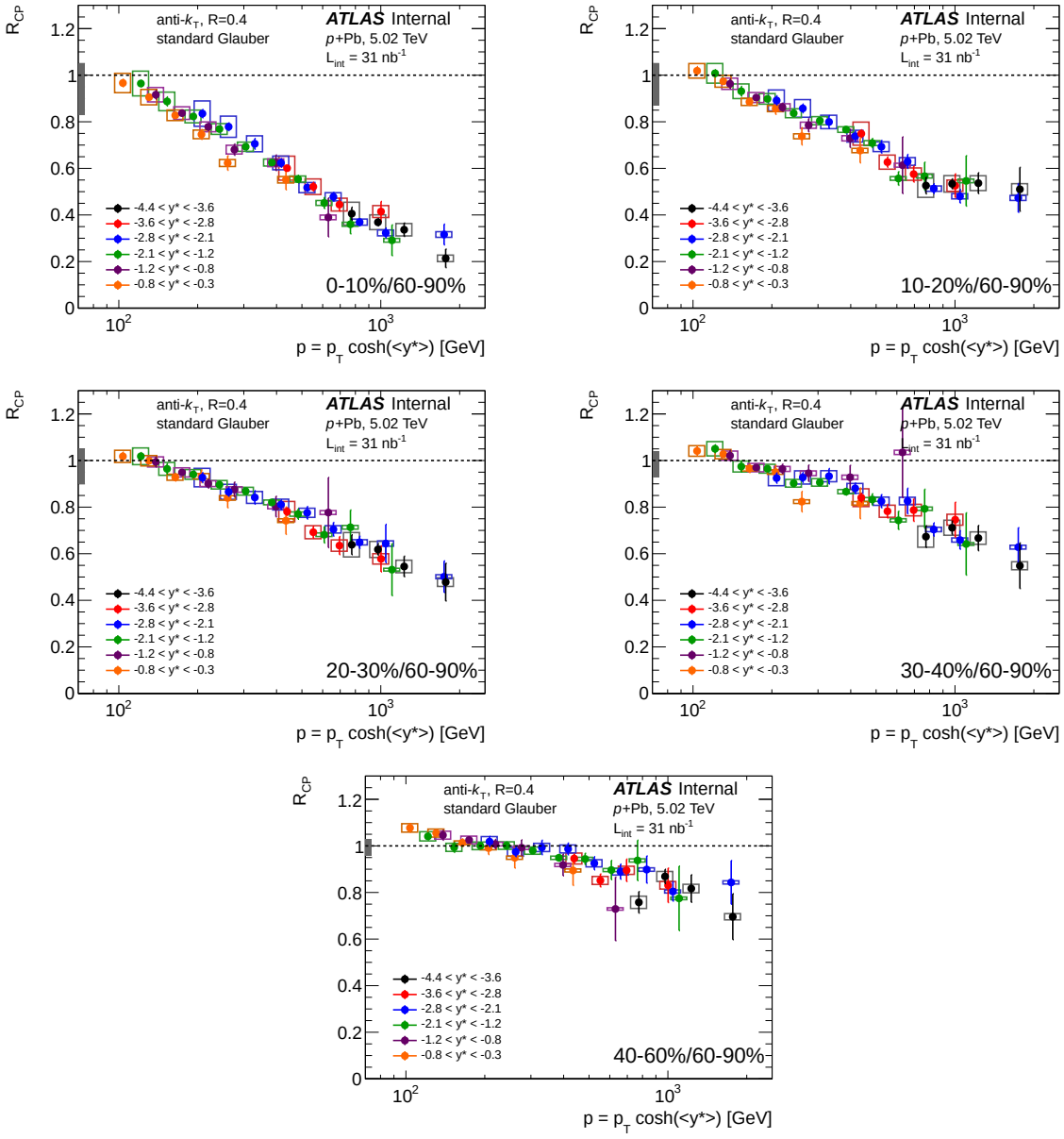


Figure 8.48: R_{CP} for R=0.4 jets in the five centrality selections, showing the six rapidity bins with $y^* < -0.3$ (the p-going direction). The R_{CP} is plotted as a function of $p = p_{\mathrm{T}} \cosh \langle y^* \rangle$, where $\langle y^* \rangle$ is the center of the y^* rapidity bin. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

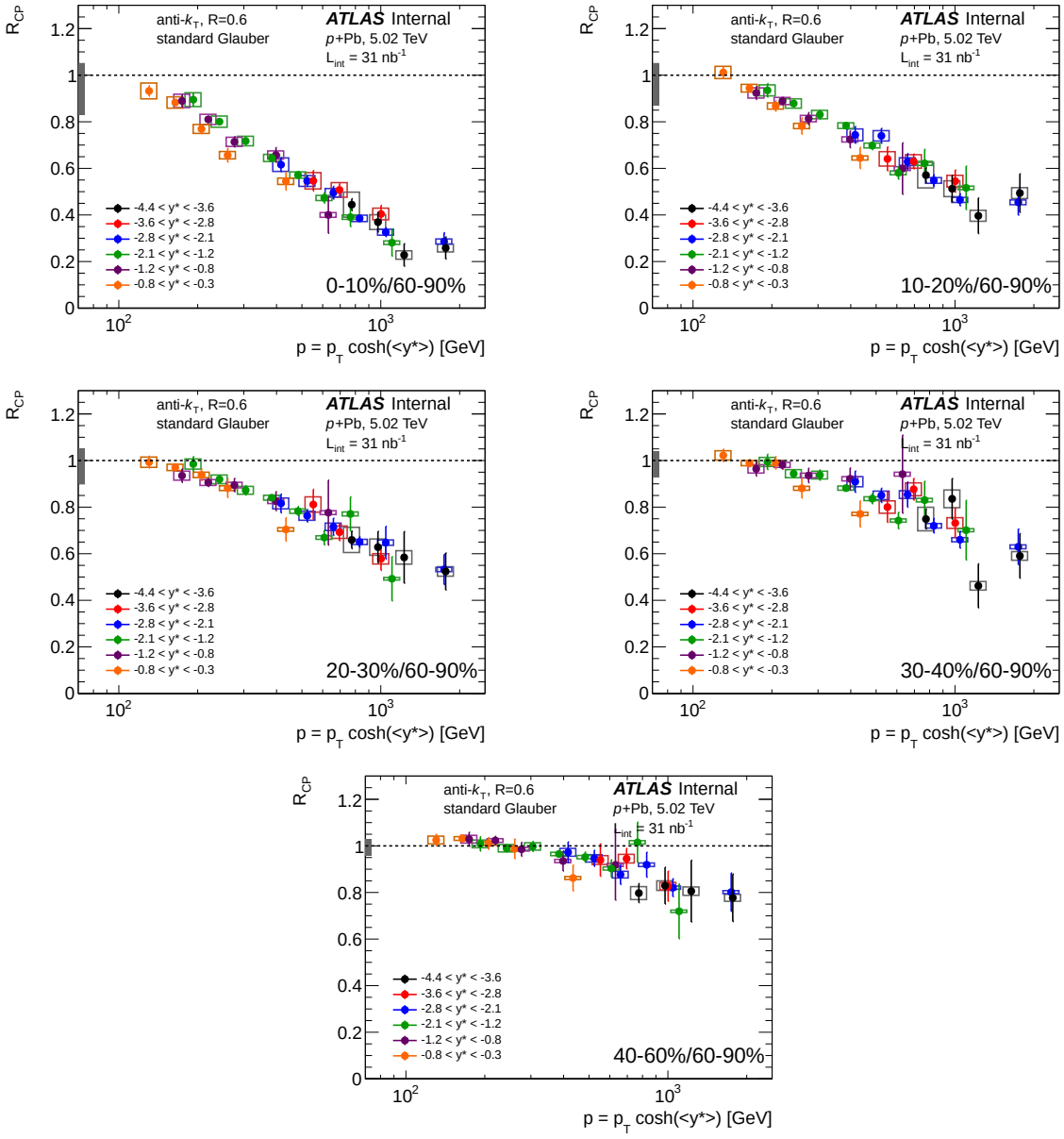


Figure 8.49: R_{CP} for R=0.6 jets in the five centrality selections, showing the six rapidity bins with $y^* < -0.3$ (the p-going direction). The R_{CP} is plotted as a function of $p = p_{\mathrm{T}} \cosh \langle y^* \rangle$, where $\langle y^* \rangle$ is the center of the y^* rapidity bin. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{CP}} = 1$ denote the systematic uncertainty from R_{coll}.

8.7.4 Jet cross-section in 2.76 TeV *pp* collisions

The jet cross-section in *pp* collisions is somewhat simpler to construct than the per-event yields in p+Pb. For a p_{T} bin populated by a given trigger, the cross-section is

$$\frac{d^2\sigma}{dp_{\mathrm{T}}dy}(p_{\mathrm{T}}, y) = \frac{1}{\mathcal{L}_{int}^{\mathrm{trig}}} N^{\mathrm{jet,trig}}(p_{\mathrm{T}} \in \Delta p_{\mathrm{T}}, y \in \Delta y) \frac{1}{\Delta p_{\mathrm{T}} \Delta y} \tag{8.49}$$

where $N^{\mathrm{jet,trig}}$ is the total yield of jets selected by the trigger, $L_{\mathrm{int}}^{\mathrm{trig}}$ is the luminosity for the trigger, and Δp_{T} and Δy are the width of the (p_{T}, y) bin. Figure 8.50 shows the *pp* jet cross-section at the uncorrected p_{T} scale for R=0.4 and R=0.6 jets, for three choices of binning.

Figures B.14 and B.15 in Appendix B.3 shows power-law fits of the form $c_0 \cdot {p_{\mathrm{T}}}^{-c_1+c_2 p_{\mathrm{T}}}$ to data at the uncorrected scale. Interestingly, this simple form is able to describe the data very well in all pseudorapidity bins. Furthermore, the ratio of the data to fits show no unusual jumps that are not consistent with statistical fluctuations.

Figures B.16 and B.17 in Appendix B.3 show the comparison of data to the 2.76 TeV PYTHIA MC, with both at the detector scale. While the ratio is near unity in all bins, a small but consistent slope is observed in the ratio in most of the bins, with a slope in $\log p_{\mathrm{T}}$-space of approximately $\approx 0.1\text{-}0.2 \cdot log p_{\mathrm{T}}$.

As a check on the *pp* cross-section results before they are used in the R_{pPb}, the present results are compared to the $\sqrt{s} = 2.76$ TeV jet cross-section with the 2011 data[275; 272]. The ratio of the 2013 result (the present one) to the published ATLAS 2011 result for the $\sqrt{s} = 2.76$ TeV cross-section is shown in Figure 8.51 for R=0.4 jets and Figure 8.52 for R=0.6 jets. The black vertical lines are the statistical uncertainty on the 2013 data, the red boxes are the statistical uncertainty on the 2011 data, and the blue boxes are the total systematic uncertainty on the 2011 data. There is an additional 2.7% on the systematic uncertainty from the luminosity in the 2011 data which is not plotted.

There are a few features in the comparison. First, the ratio is consitent with flat for 20 GeV $< p_{\mathrm{T}} <$ 200 GeV. Thus, the shapes of the corrected spectra are generally consistent. Second, the ratio seems to be systematically below 1 by $\approx 10\%$ in a way that is nearly independent of rapidity. This implies there may be an overall normalization effect (such as from the lack of a preliminary calibration for the 2013 $\sqrt{s} = 2.76$ TeV luminosity) remaining.

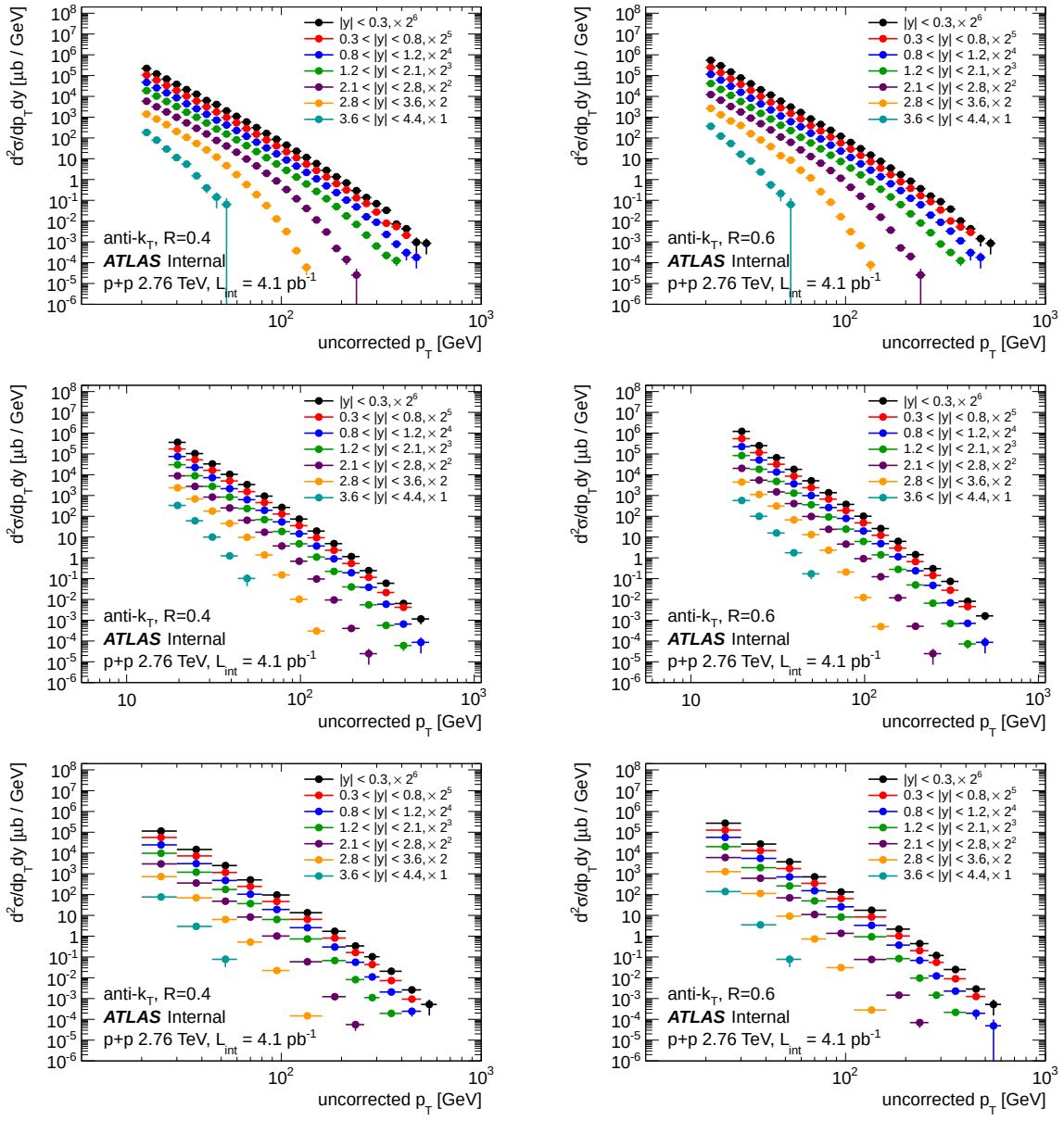


Figure 8.50: Jet cross-section in pp collisions at $\sqrt{s} = 2.76$ TeV in y bins for R=0.4 (left) and R=0.6 (right) jets. Results are shown with three separate binnings: the nominal reconstructed-level binning (top row), the $\sqrt{s}$-binning to be used in the x_{T}-interpolation (middle row) and the standard ATLAS jet binning (bottom row). The jet cross-section is shown at the reconstructed level, corrected for selection efficiency but before corrections for detector resolution and other effects.

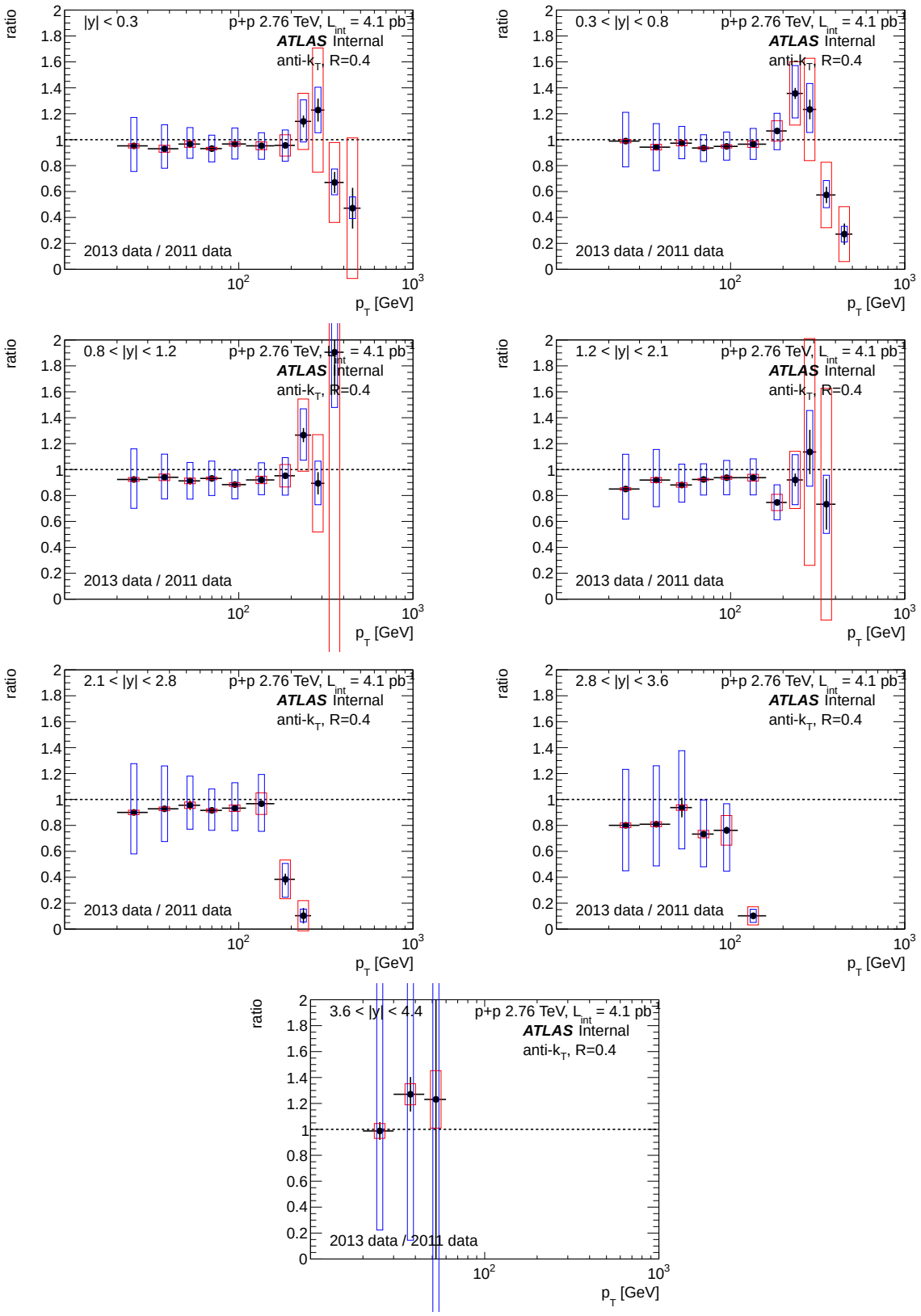


Figure 8.51: Comparison of the present results to the 2011 ATLAS jet cross-section at $\sqrt{s} = 2.76$ TeV, for R=0.4 jets. The black vertical lines and red boxes are the statistical uncertainty on the current and 2011 result, respectively, and the blue boxes show the total systematic uncertainty on

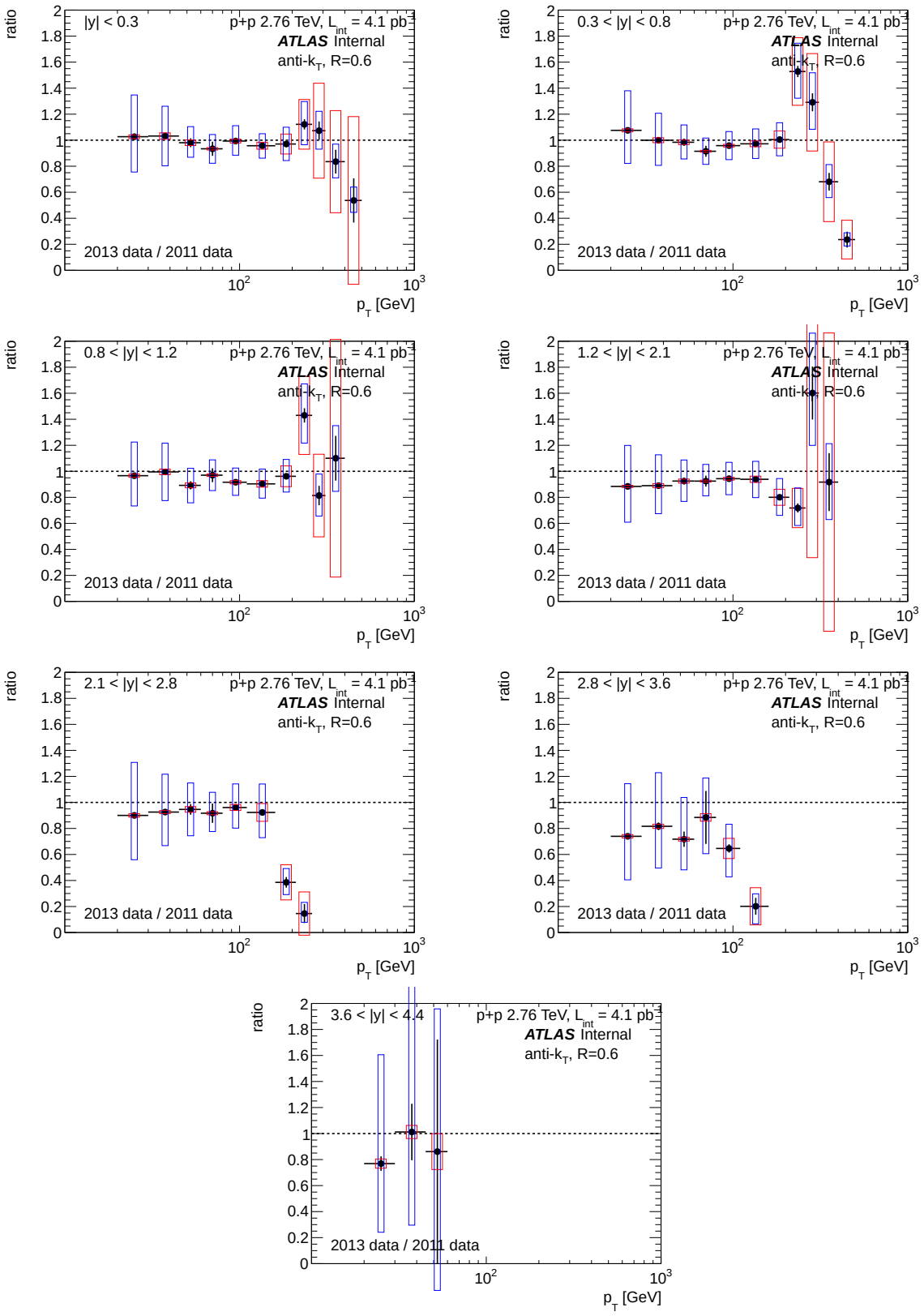


Figure 8.52: Comparison of the present results to the 2011 ATLAS jet cross-section at $\sqrt{s} = 2.76$ TeV, for R=0.6 jets. The black vertical lines and red boxes are the statistical uncertainty on the current and 2011 result, respectively, and the blue boxes show the total systematic uncertainty on

Third, in the three bins closest to mid-rapidty, there is spike in the ratio between 200-300 GeV, and at very high-p_{T} (> 300 GeV in $0.3 < |y| < 0.8$, and > 100 GeV in the two $2.1 < |y| < 3.6$ bins), the ratio suddenly drops. However, these are regions where the 2011 measurement was running out of statistics and the central value of 2011 cross-section results had deviations from the pQCD prediction.

8.7.5 Interpolated 5.02 TeV *pp* reference

The *pp* cross-section at $\sqrt{s} = 2.76$ TeV is interpolated to the reference center of mass energy of $\sqrt{s} = 5.02$ TeV needed to construct the R_{pPb}. The interpolation is based on an ATLAS measurement of the x_{T} scaling in jets between 2.76 TeV and 7 TeV[275]. In this measurement, the variable $\rho(x_{\mathrm{T}}; y)$ is defined as

$$\rho(x_{\mathrm{T}}; y) = \left(\frac{2.76 \text{ TeV}}{7 \text{ TeV}}\right)^3 \cdot \frac{\sigma(x_{\mathrm{T}}; y, 2.76 \text{ TeV})}{\sigma(x_{\mathrm{T}}; y, 7 \text{ TeV})} \tag{8.50}$$

where $\sigma(x_{\mathrm{T}}; y, \sqrt{s}) = d^2\sigma^{\sqrt{s}}/dp_{\mathrm{T}}dy$ in the p_{T} bin corresponding to $p_{\mathrm{T}} = x_{\mathrm{T}}\sqrt{s}/2$. (That is, ρ in the bins $x_{\mathrm{T}} = 0.1$ to $x_{\mathrm{T}} = 0.2$ is the ratio of $d^2\sigma/dp_{\mathrm{T}}dy$ at 2.76 TeV in the range $p_{\mathrm{T}} = 138$ GeV to 276 GeV to $d^2\sigma/dp_{\mathrm{T}}dy$ at 7 TeV in the range $p_{\mathrm{T}} = 350$ GeV to 700 GeV.)

$\rho(x_{\mathrm{T}}; y)$ can be thought of as an effective correction to the power law behavior of the jet cross-section ${p_{\mathrm{T}}}^{-n}$ which varies slowly with x_{T}. The form in Equation 8.50 can be rewritten to constuct the *pp* cross-section at another $\sqrt{s}$, given the spectrum at $\sqrt{s} = 2.76$ TeV, via

$$\sigma(x_{\mathrm{T}}; y, \sqrt{s}) = \left(\frac{2.76 \text{ TeV}}{\sqrt{s}}\right)^{n(x_{\mathrm{T}})} \cdot \sigma(x_{\mathrm{T}}; y, 2.76 \text{ TeV}) \tag{8.51}$$

This equation requires one other value of $\sqrt{s}$ to be constrain $n(x_{\mathrm{T}})$. The values of ρ published in [275] are used for this purpose. Rewriting Equation 8.51 with the definition of ρ in Equation 8.50,

$$\sigma(x_{\mathrm{T}}; y, 7 \text{ TeV}) = \frac{1}{\rho(x_{\mathrm{T}})}\left(\frac{2.76 \text{ TeV}}{7 \text{ TeV}}\right)^3 \cdot \sigma(x_{\mathrm{T}}; y, 2.76 \text{ TeV}) \tag{8.52}$$

Thus,

$$1/\rho(x_{\mathrm{T}}) = \left(\frac{2.76 \text{ TeV}}{7 \text{ TeV}}\right)^{n(x_{\mathrm{T}})-3} \tag{8.53}$$

$$n(x_{\mathrm{T}}) = 3 + \log \rho(x_{\mathrm{T}}) / \log(7/2.76) \tag{8.54}$$

Thus, the $\sqrt{s} = 5.02$ TeV pp cross-section can be constructed according to Equation 8.51 with the measurement of $n(x_{\mathrm{T}})$ given in Equation 8.54.

A linear fit in $\log(x_{\mathrm{T}})$ space is performed to the $\rho(x_{\mathrm{T}})$ data, and the result of the fit is evaluated at the mid-point of chosen x_{T} bins. Figures 8.53 and 8.54 shows the ρ data and fits for R=0.4 and R=0.6 jets. The fits do a good job of describing the data.

The 2.76 TeV pp data are rebinned in the (2.76 TeV/ 5.02 TeV)-scaled equivalent of the p_{T} bins used in the p+Pb analysis, and corrected with bin-by-bin correction factors. Then, the 5.02 TeV pp reference is created following Equation 8.51. An important systematic uncertainty that enters directly into the final results is the systematic uncertainty on ρ. This will be discussed in Section 8.8.6. The x_{T}-interpolated pp results are shown in Figures 8.55 and 8.56 for R=0.4 and R=0.6 jets, respectively.

8.7.6 Jet R_{pPb}

The nuclear modification factor R_{pPb} is presented here.

The p+Pb yields above $p_{\mathrm{T}} > {p_{\mathrm{T}}}^{\mathrm{cut}}$ (where the bin migration effects are independent of the p+Pb underlying event) are corrected for detector effects with the bin by bin correction factors. They are divided by the T_{AB}-scaled pp cross-section at 5.02 TeV. The results are shown in Figures 8.57 and 8.58 for R=0.4 jets and Figures 8.59 and 8.60 for R=0.6 jets. The 0-90% R_{pPb}, which is expected to be ≈ 1, appears to sit systematically high at $\approx$ 1.1-1.2. This may be related to the fact that the 2013 pp data appears to be systematically lower than the 2011 pp result (since the T_{AB}-scaled pp reference is in the denominator, this would manifest itself as a systematically *high* R_{pPb}). However, the R_{pPb} is generally flat in p_{T}. Furthermore, the R_{pPb} for each centrality selection is interesting. The peripheral R_{pPb} appears to rise, while the most central R_{pPb} decreases, with the other centrality selections in between these. As in the R_{CP}, the centrality dependence increases with p_{T} and onsets earlier the more forward the rapidity bin.

As an alternate check on the results, the R_{pPb} is also constructed using the T_{AB}-scaled truth jet cross-section in the 5.02 TeV MC Signal as the denominator, denoted ${R_{\mathrm{pPb}}}^{\mathrm{PYTHIA}}$. The ${R_{\mathrm{pPb}}}^{\mathrm{PYTHIA}}$ is shown in Figures 8.63 and 8.64 for R=0.4 jets and Figures 8.65 and 8.66 for R=0.6 jets. The 0-90% ${R_{\mathrm{pPb}}}^{\mathrm{PYTHIA}}$ is near 1 and almost flat but for a very gradual slope with p_{T}. This slope was also seen in the comparison of the pp jet cross-section to MC, where the ratio had a

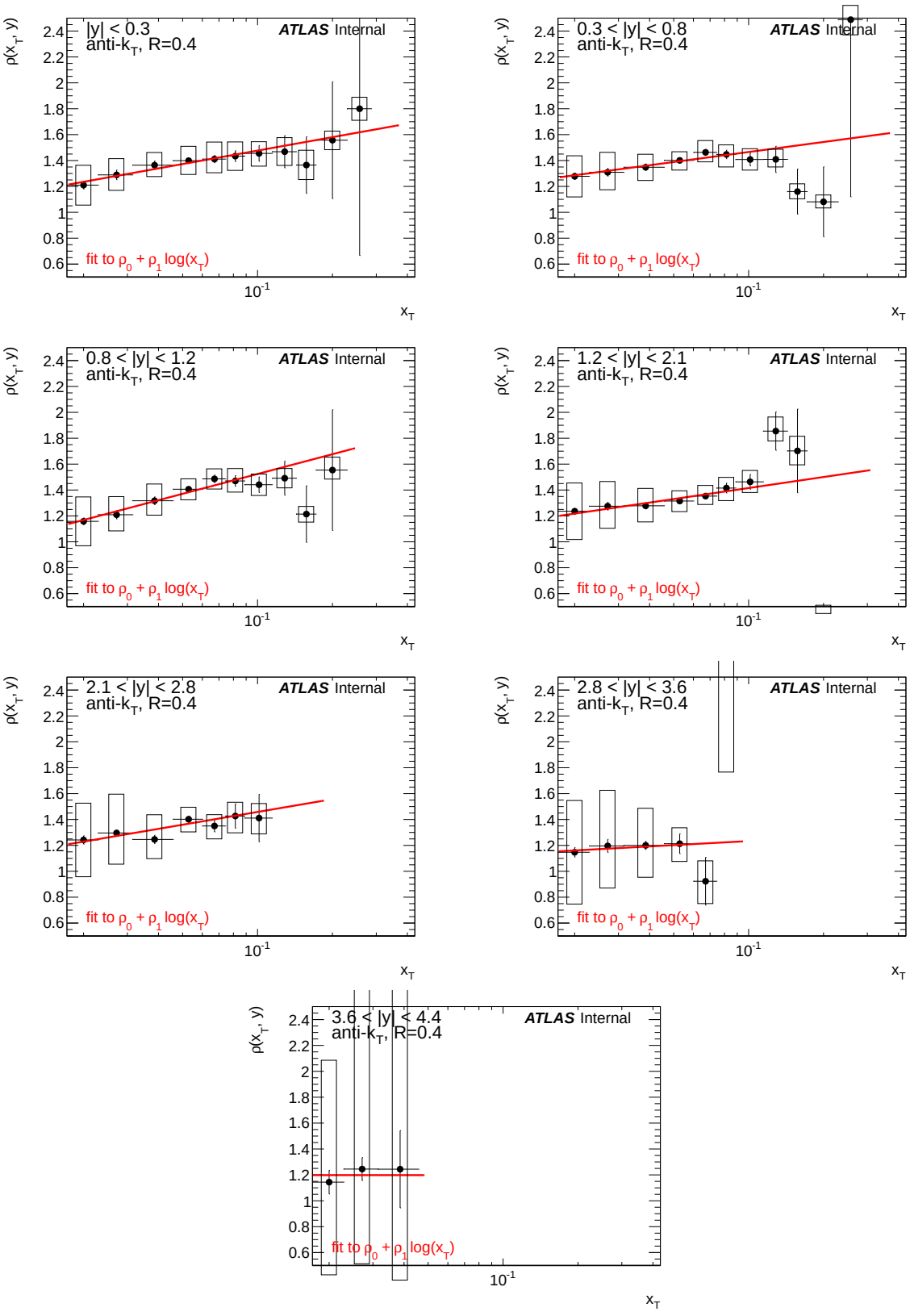


Figure 8.53: Fits to the ATLAS measurement of ρ, which is the ratio of 2.76 TeV and 7 TeV pp jet cross-sections in $x_{\mathrm{T}} = p_{\mathrm{T}}/(\sqrt{s}/2)$ bins, shown for R=0.4 jets. Vertical error bars are the statistical uncertainty on the measurement, while the height of the boxes is the total systematic uncertainty,

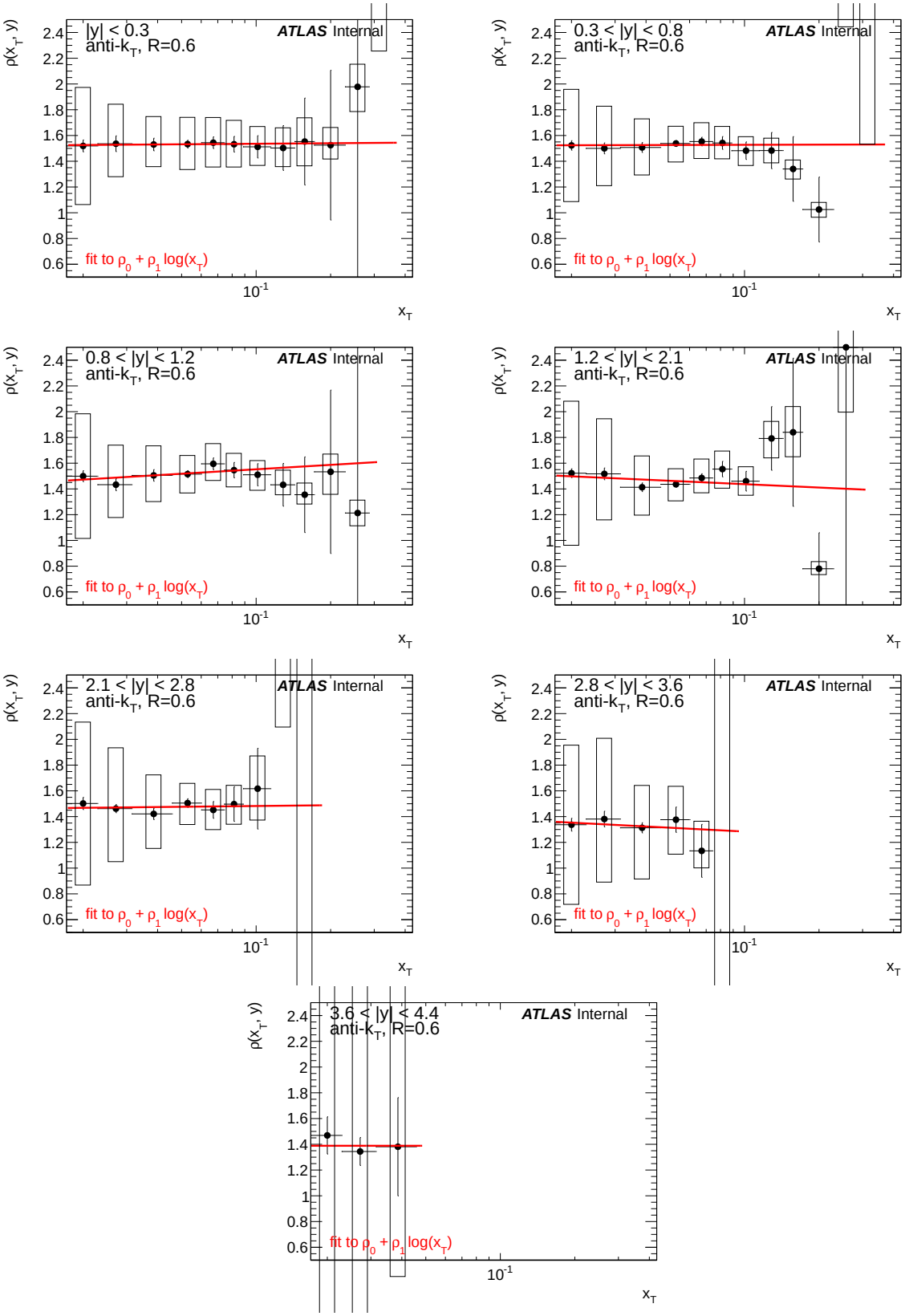


Figure 8.54: Fits to the ATLAS measurement of ρ, which is the ratio of 2.76 TeV and 7 TeV pp jet cross-sections in $x_{\mathrm{T}} = p_{\mathrm{T}}/(\sqrt{s}/2)$ bins, shown for R=0.6 jets. Vertical error bars are the statistical uncertainty on the measurement, while the height of the boxes is the total systematic uncertainty,

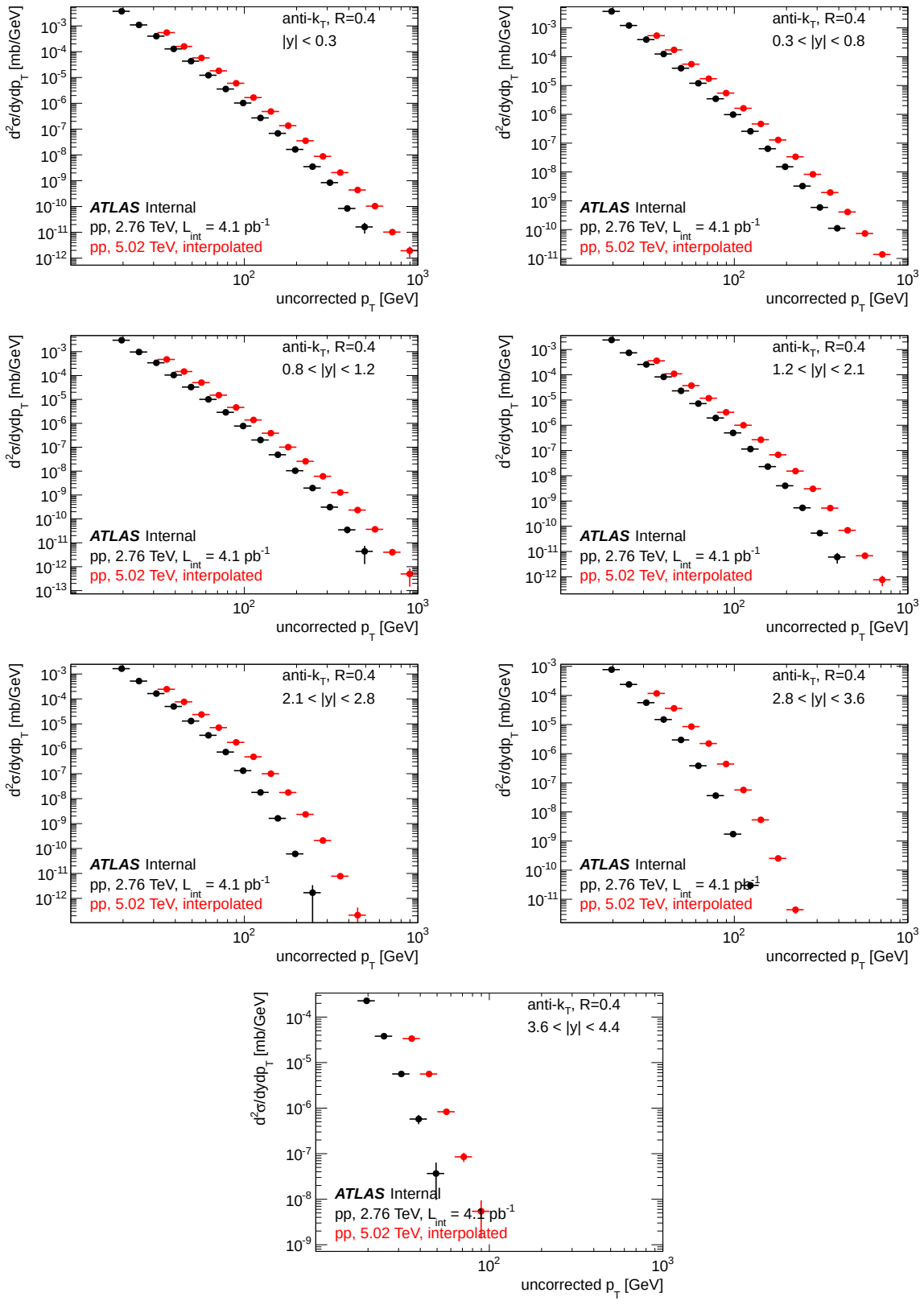


Figure 8.55: Construction of the $\sqrt{s} = 5.02$ TeV pp reference using an x_{T}-scaling based interpolation of the $\sqrt{s} = 2.76$ TeV results, for $R = 0.4$ jets.

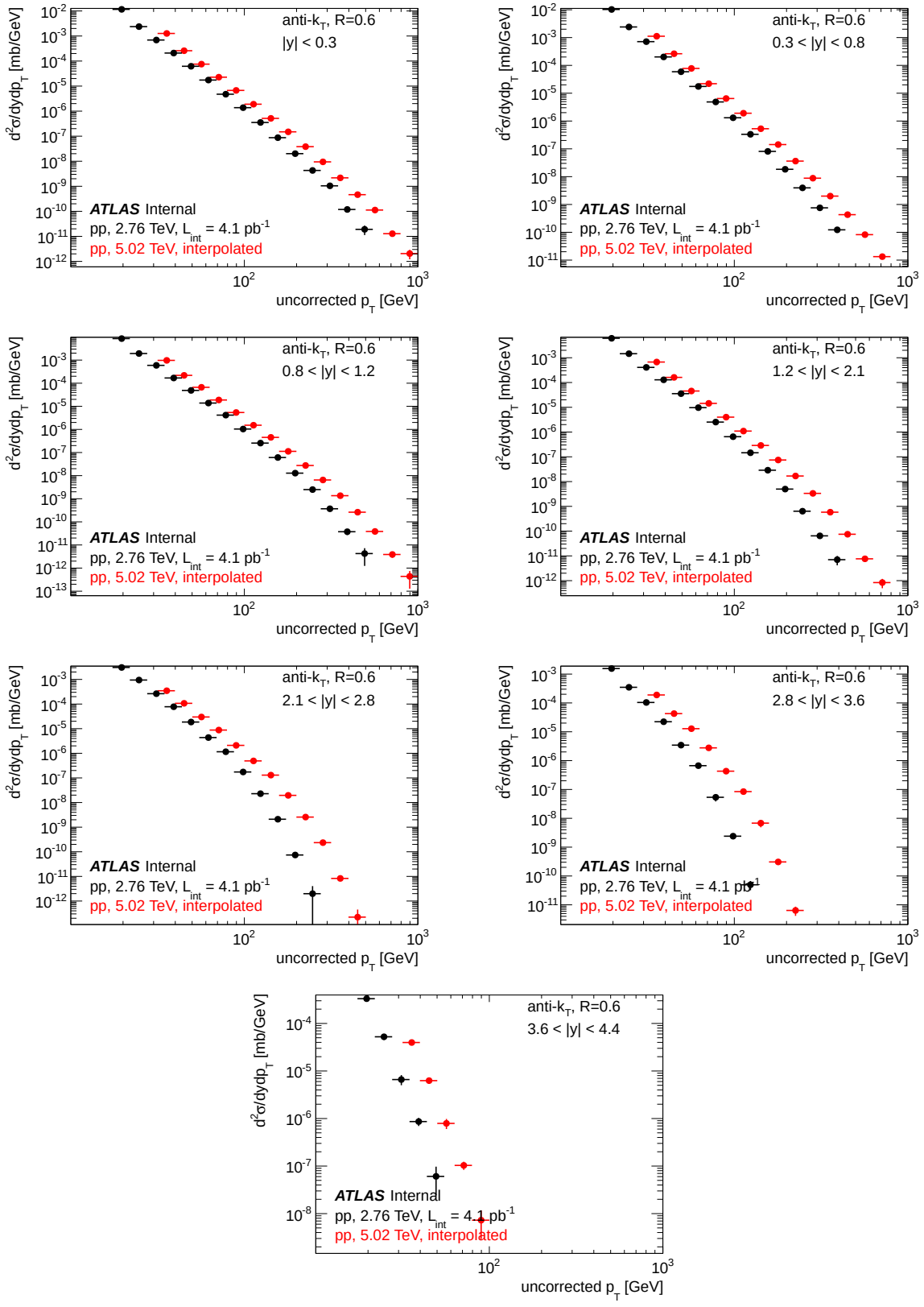


Figure 8.56: Construction of the $\sqrt{s} = 5.02$ TeV pp reference using an x_{T}-scaling based interpolation of the $\sqrt{s} = 2.76$ TeV results, for $R = 0.6$ jets.

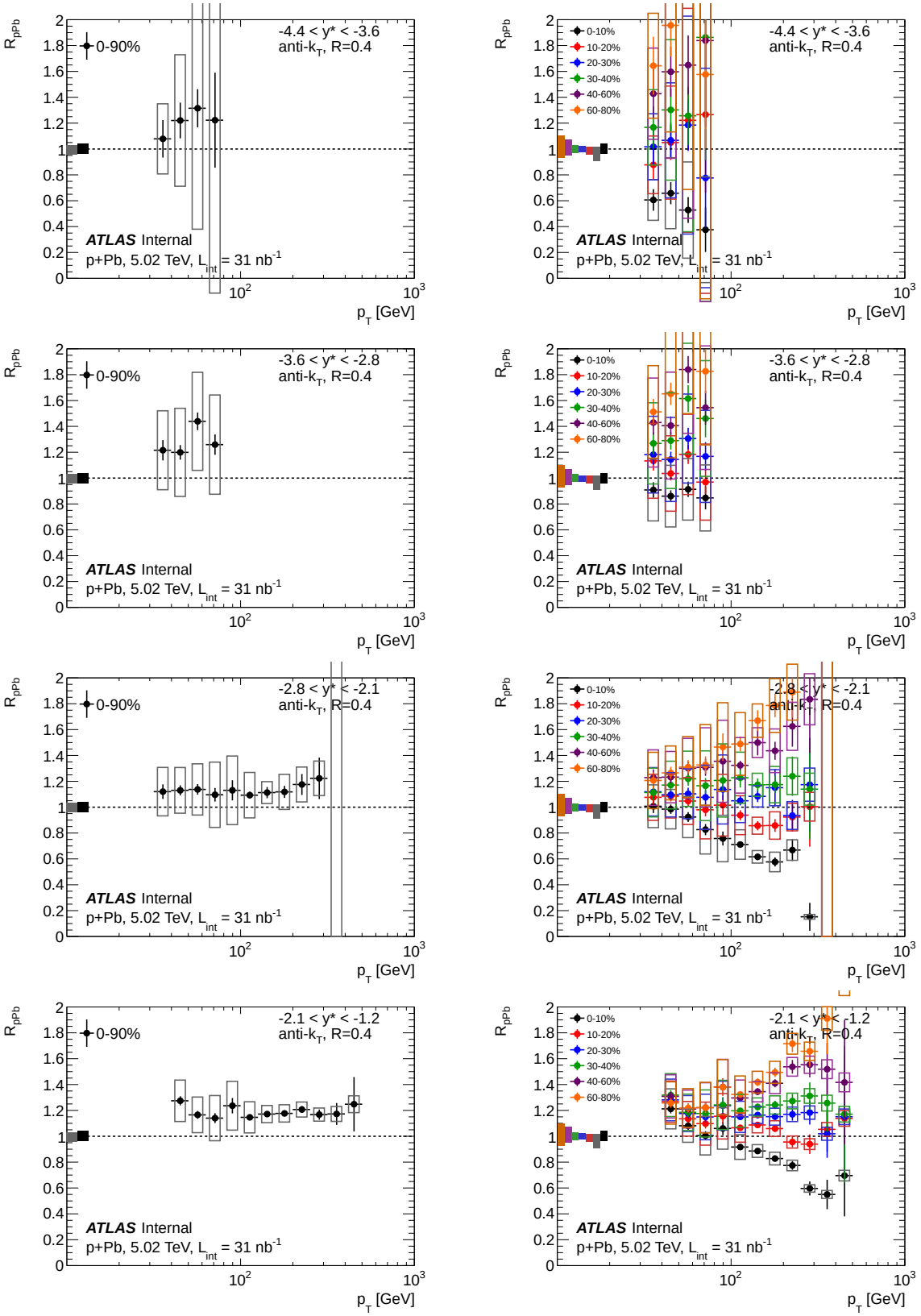


Figure 8.57: R_{pPb} for R=0.4 jets. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins $y^* < -1.2$. Additional 4.3% and 3% systematic uncertainties

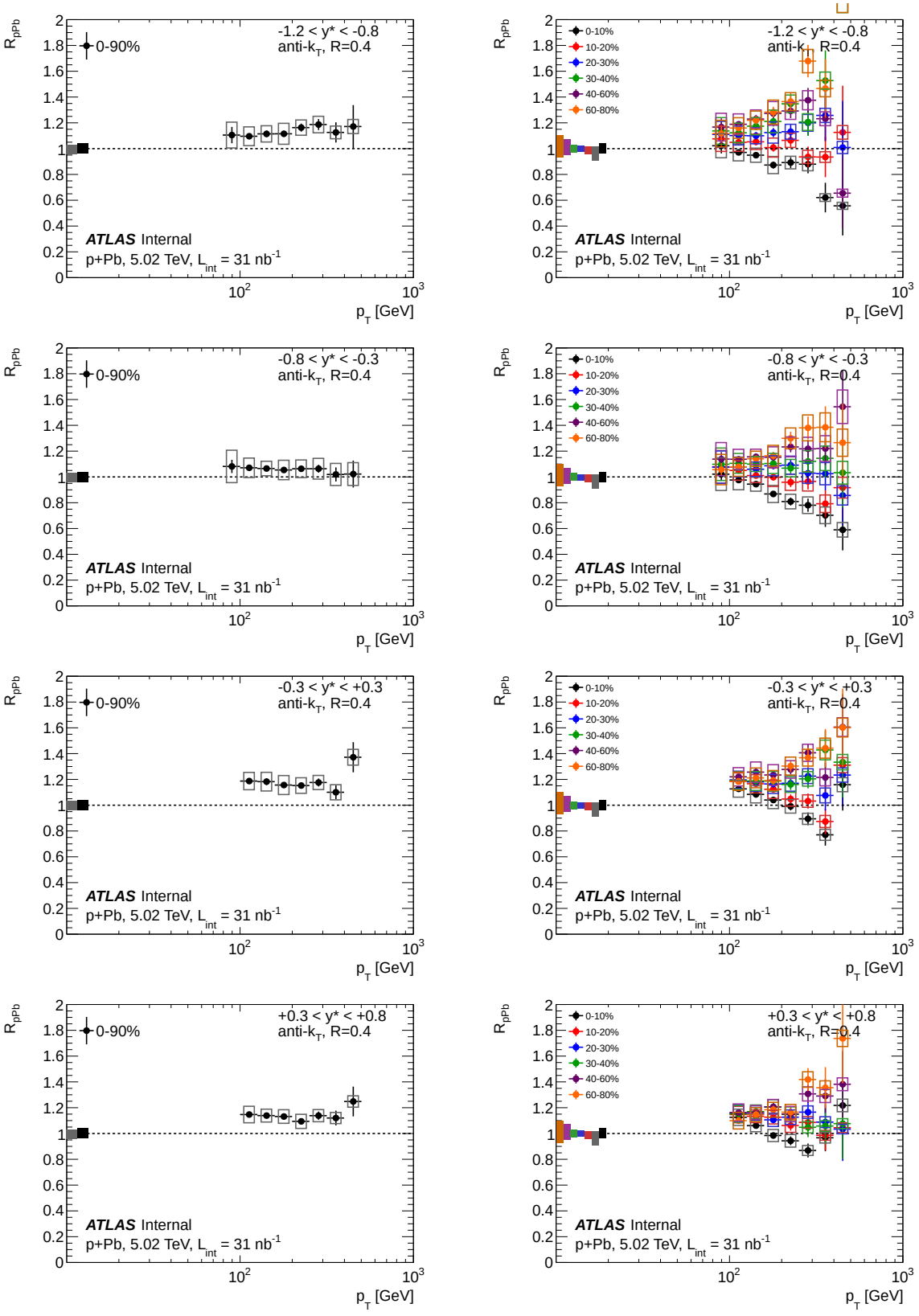


Figure 8.58: R_{pPb} for R=0.4 jets. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins $-1.2 < y^* < +0.8$. Additional 4.3% and 3% systematic

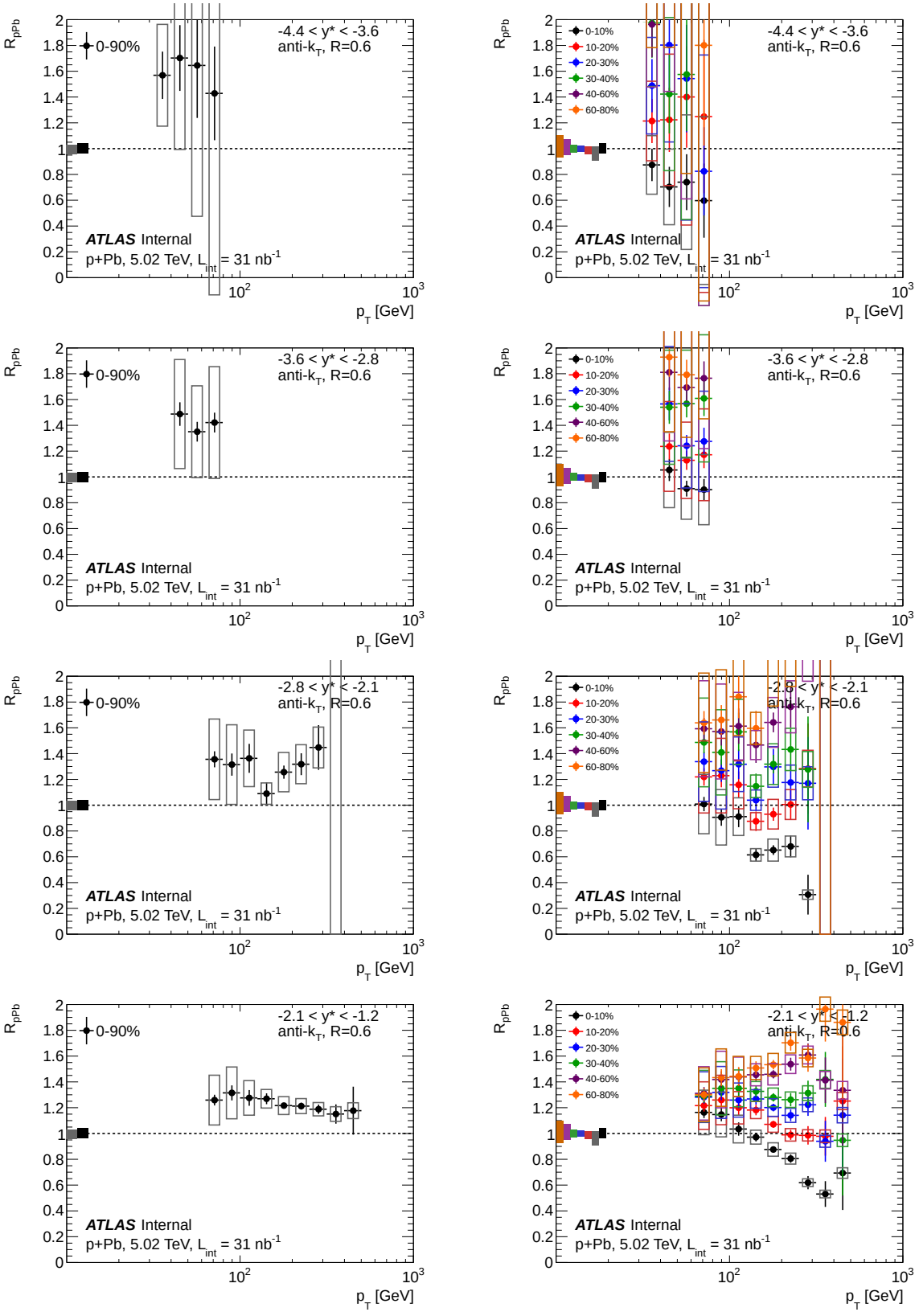


Figure 8.59: R_{pPb} for R=0.6 jets. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins $y^* < -1.2$. Additional 4.3% and 3% systematic uncertainties

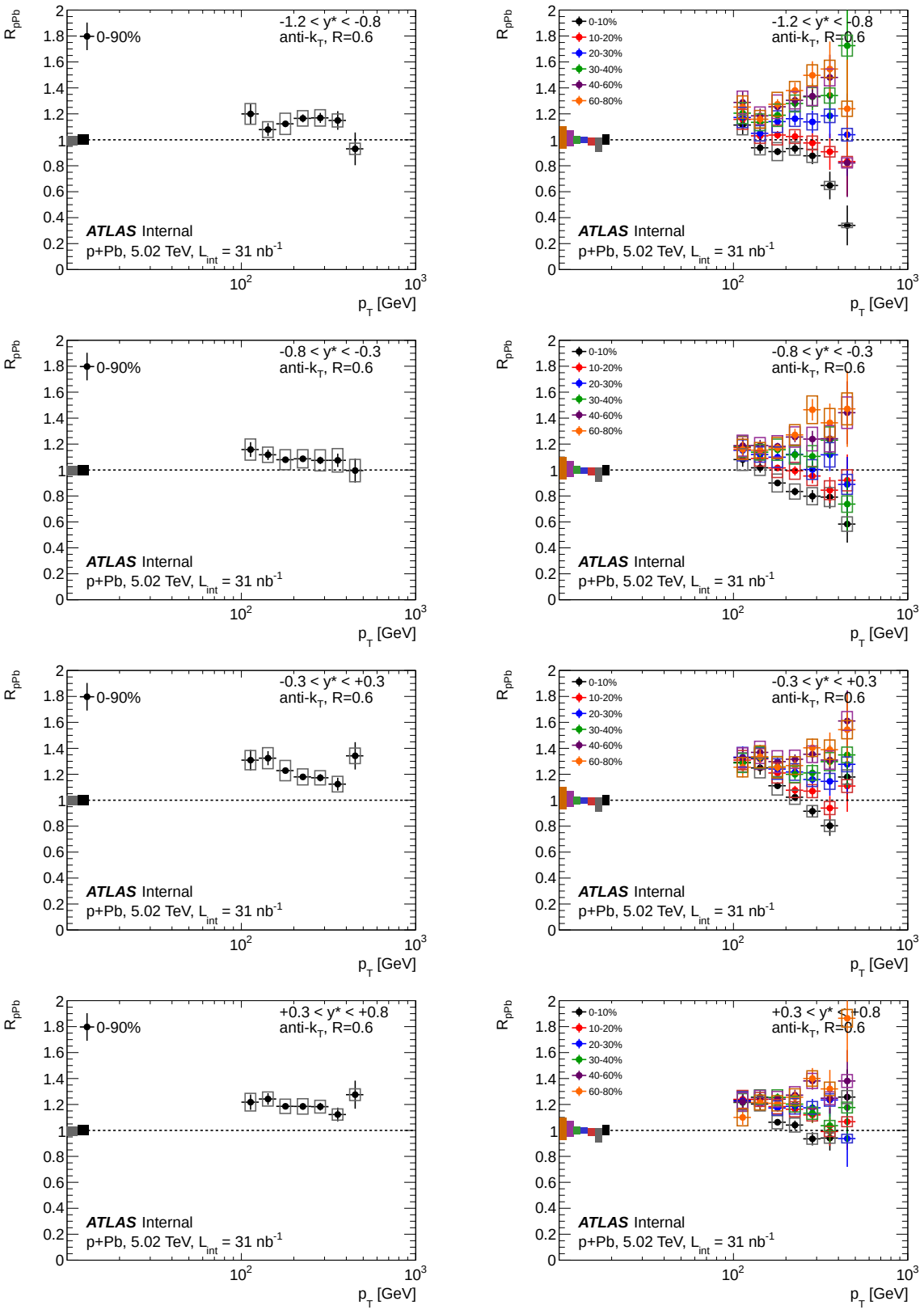


Figure 8.60: R_{pPb} for R=0.6 jets. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins $-1.2 < y^* < +0.8$. Additional 4.3% and 3% systematic

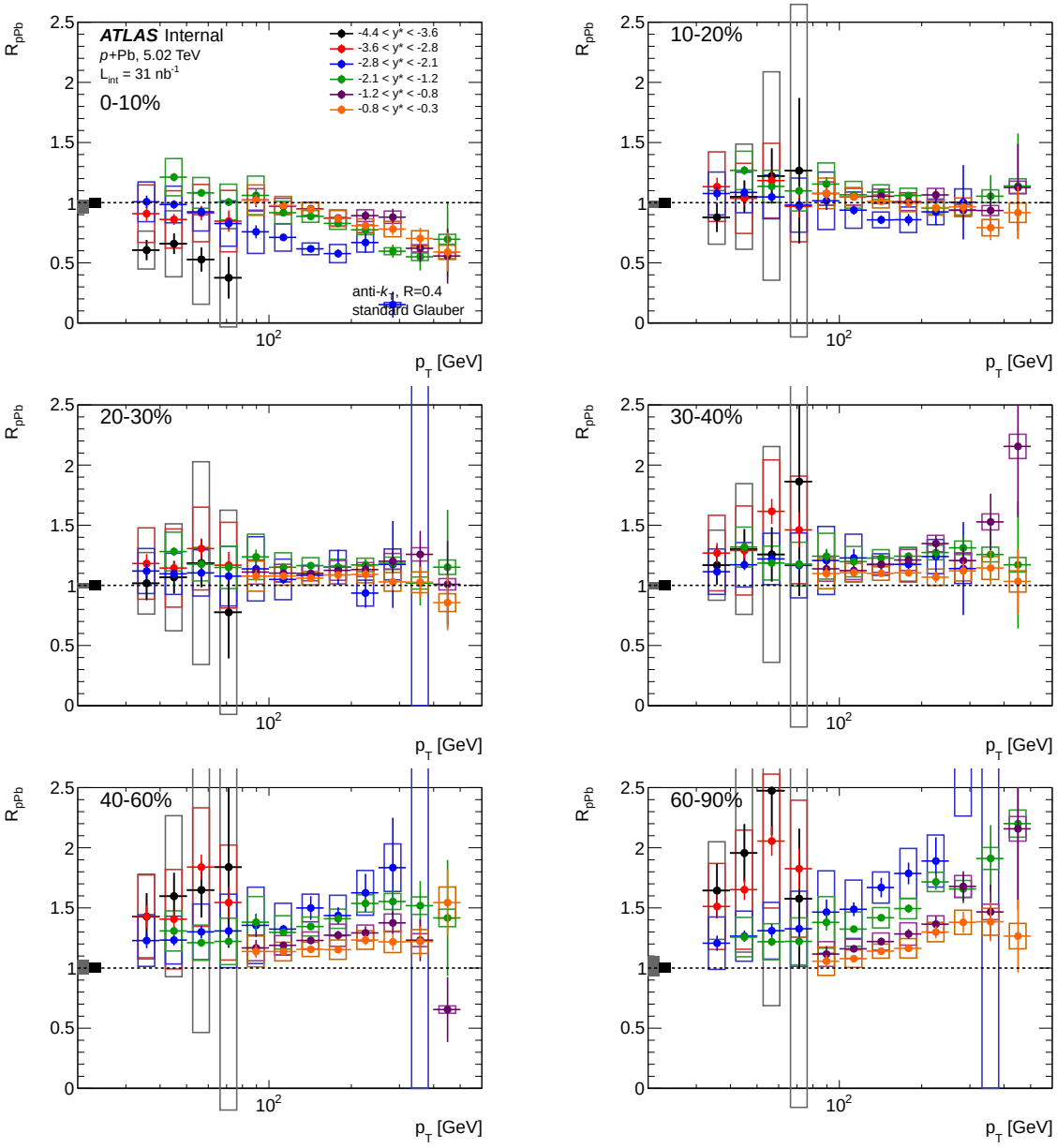


Figure 8.61: R_{pPb} for R=0.4 jets in the six centrality selections, showing the six rapidity bins with $y^* < -0.3$. The R_{CP} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{pPb}} = 1$ denote the systematic uncertainty from T_{AB}.

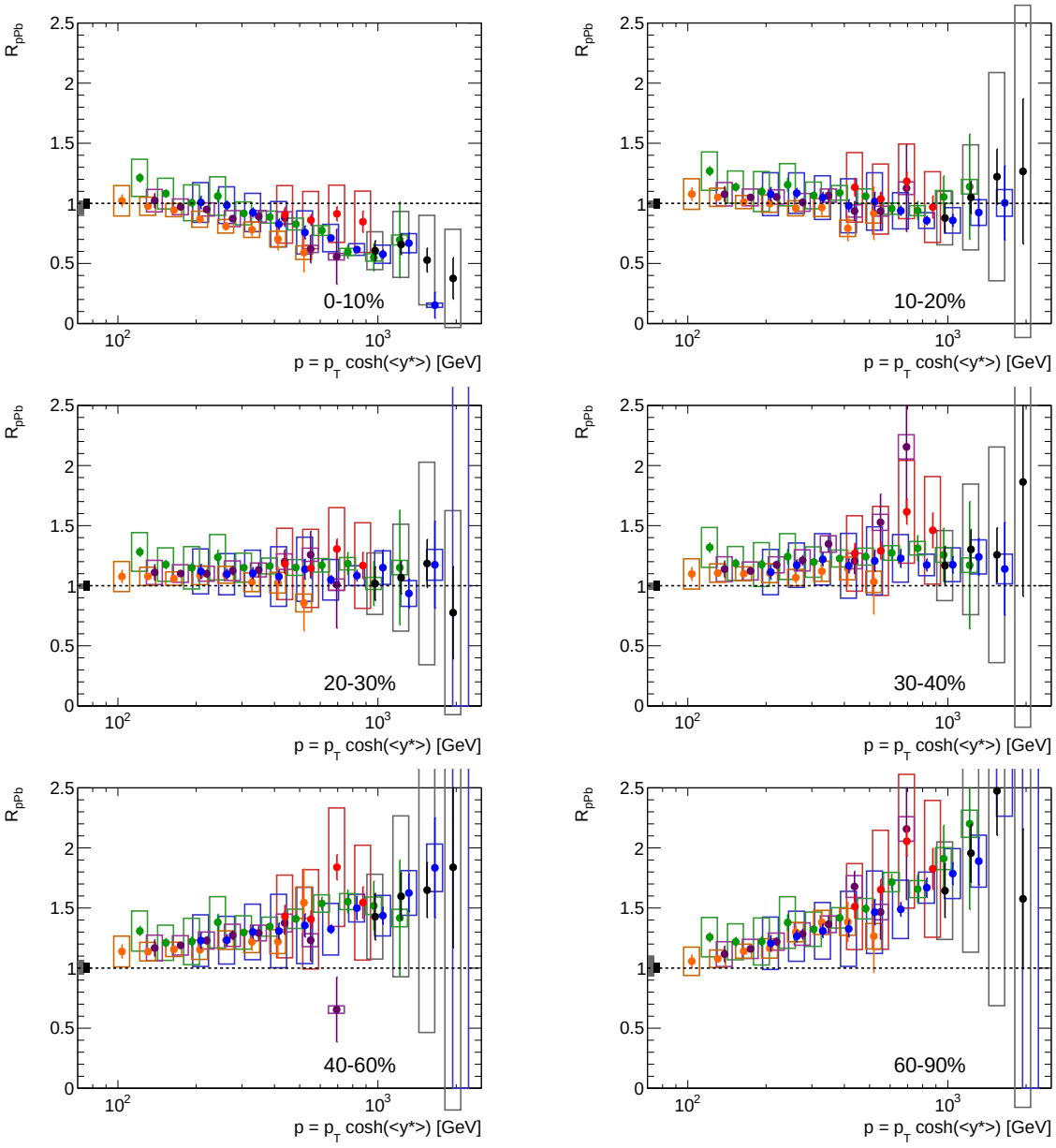


Figure 8.62: R_{pPb} for R=0.4 jets in the sixe centrality selections, showing the six rapidity bins with $y^* < -0.3$. The R_{pPb} is plotted as a function of $p = p_{\mathrm{T}} \cosh \langle y^* \rangle$, where $\langle y^* \rangle$ is the center of the y^* rapidity bin. The R_{pPb} is corrected for detector effects. The vertical bars along the left side of the dotted line $R_{\mathrm{pPb}} = 1$ denote the systematic uncertainty from T_{AB} and the luminosity.

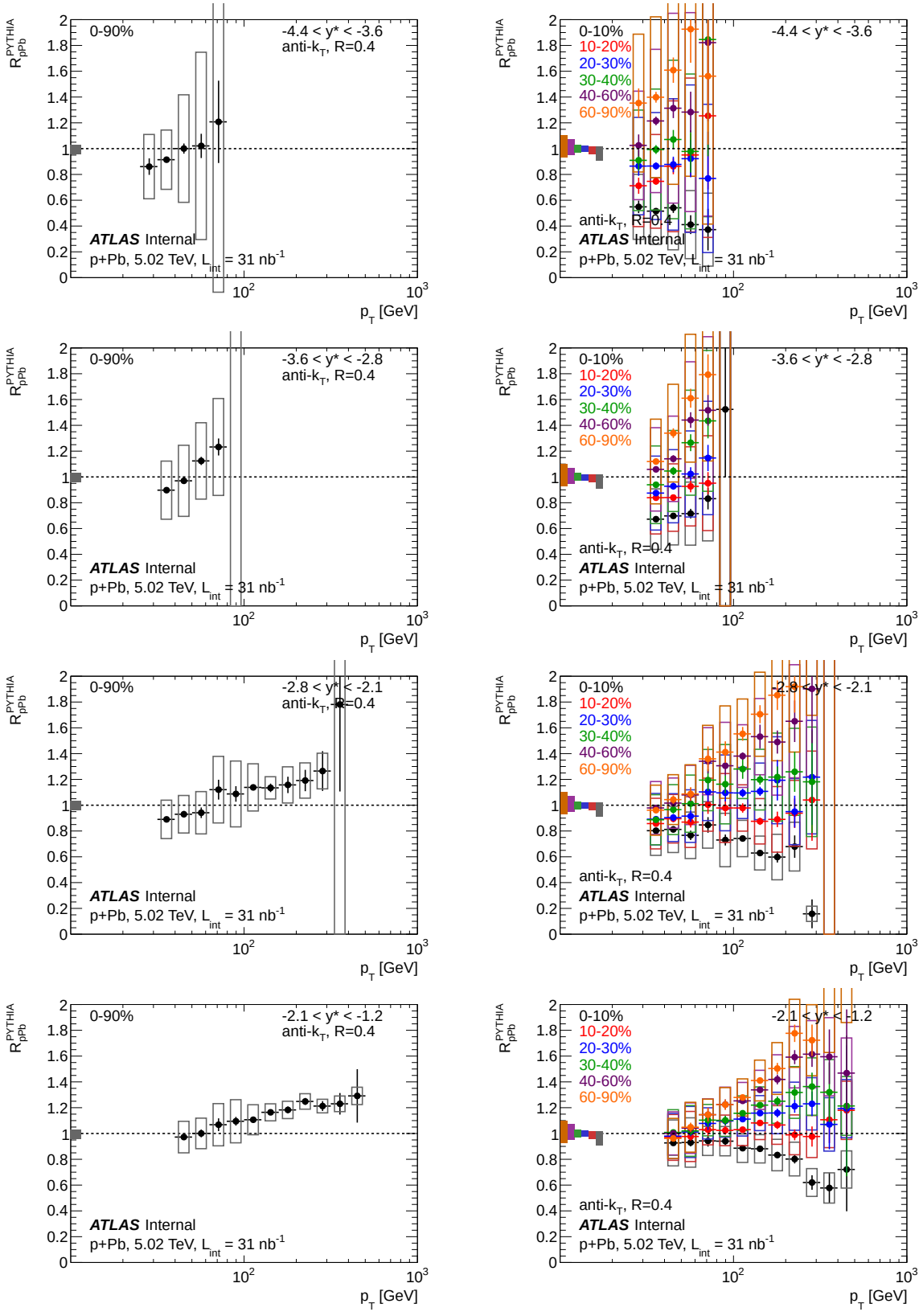


Figure 8.63: $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$ for R=0.4 jets made with a PYTHIA jet cross-section as the *pp* reference. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins

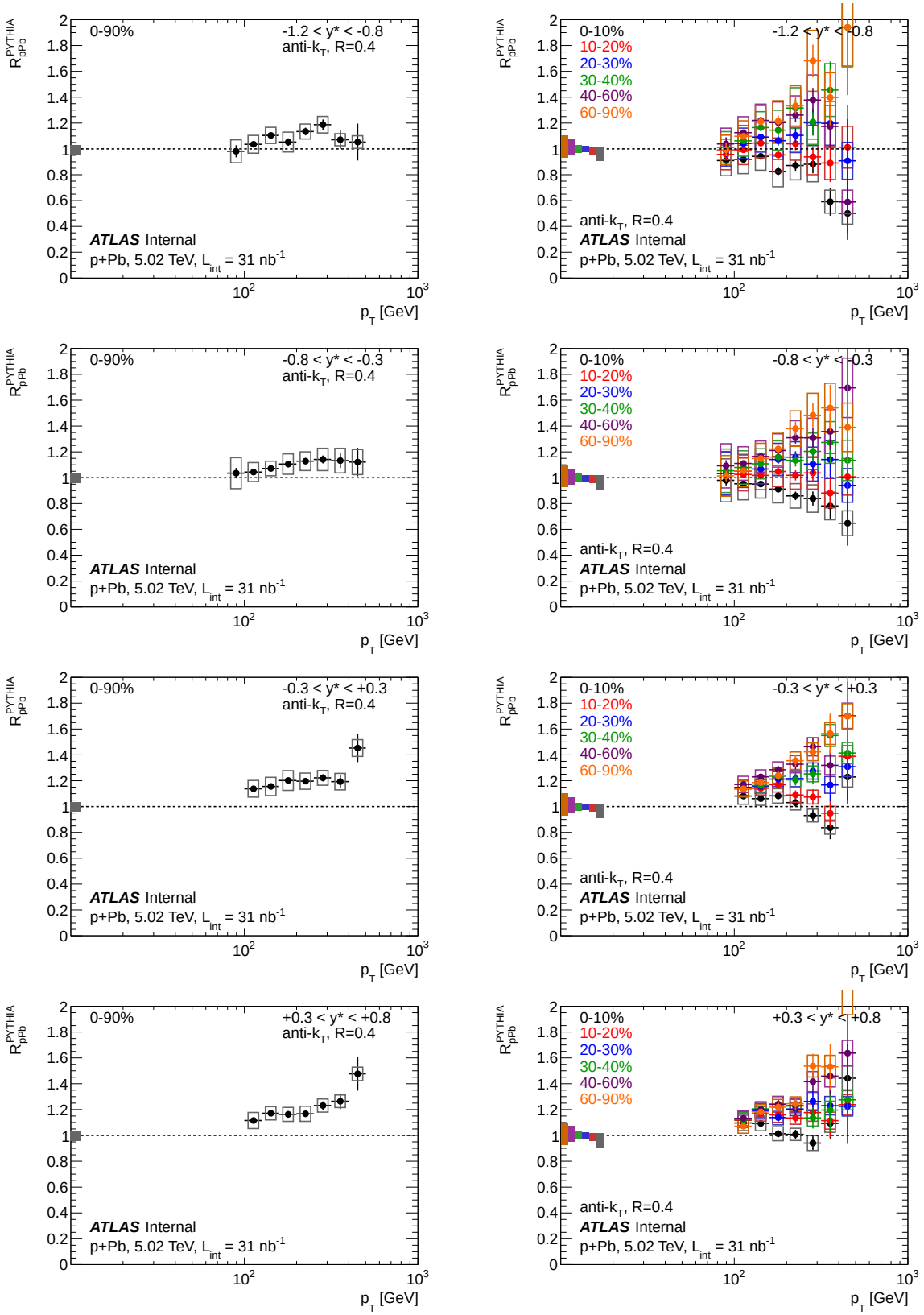


Figure 8.64: $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$ for R=0.4 jets made with a PYTHIA jet cross-section as the *pp* reference. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins

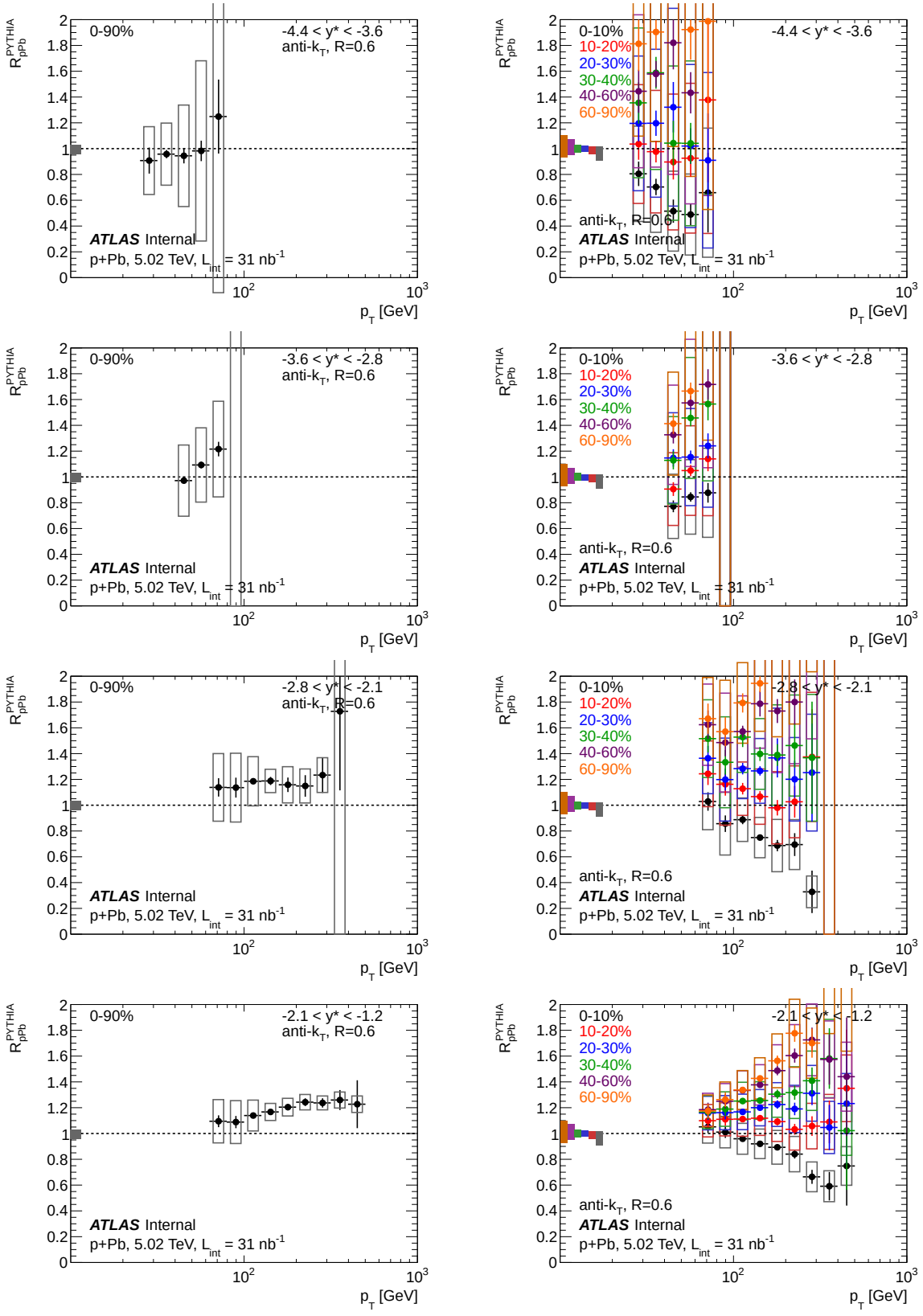


Figure 8.65: $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$ for R=0.6 jets made with a PYTHIA jet cross-section as the *pp* reference. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins

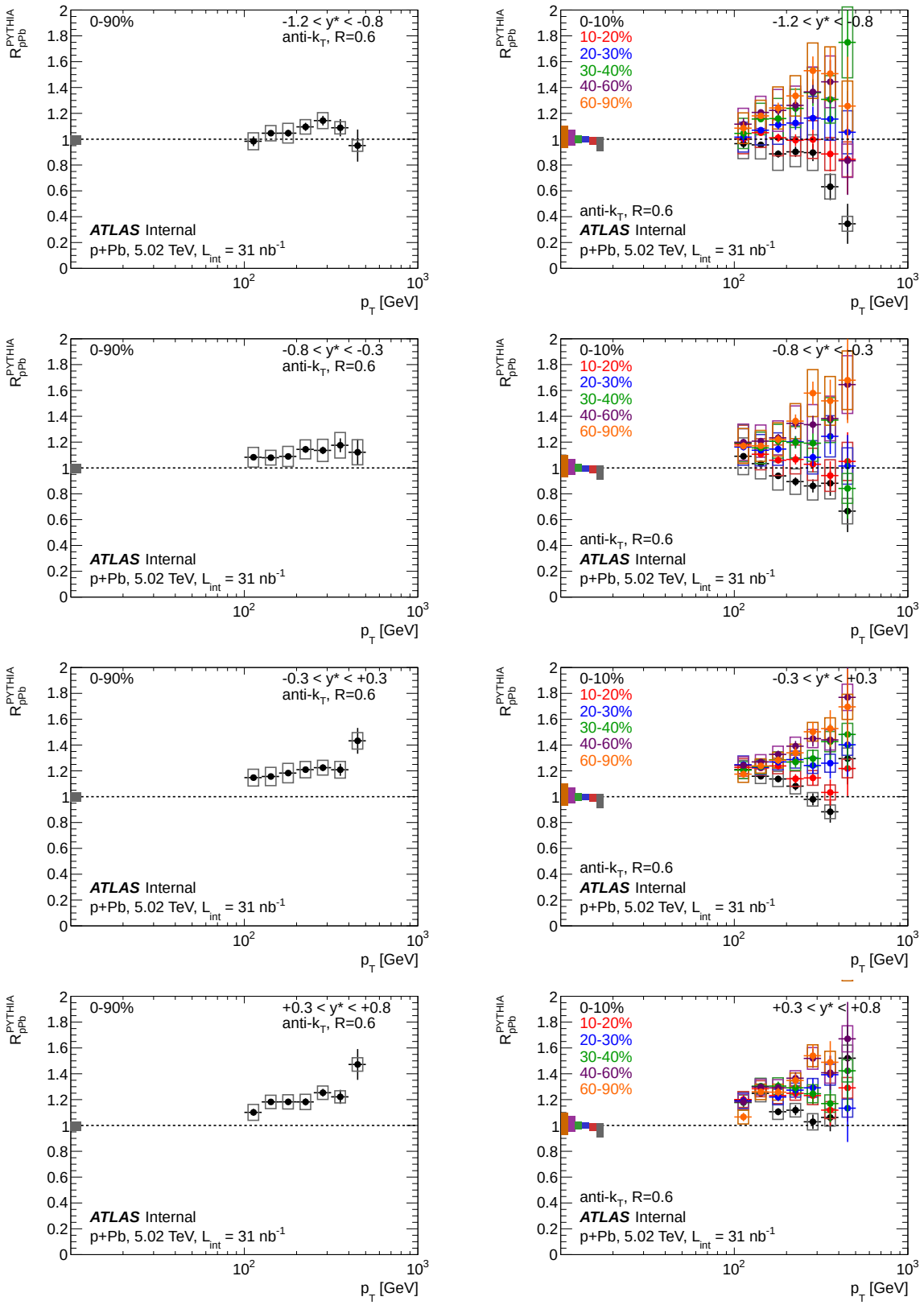


Figure 8.66: ${R_{\mathrm{pPb}}}^{\mathrm{PYTHIA}}$ for R=0.6 jets made with a PYTHIA jet cross-section as the *pp* reference. Each row is a different rapidity bin. The 0-90% R_{pPb} is shown in the column on the left, while the centrality-dependent R_{pPb} is shown on the right. Results are shown here for the four rapidity bins

slope that was approximately linear in $\log p_{\mathrm{T}}$. On the other hand, the $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$ for the different centrality selections has a centrality dependence that increases with p_{T} and is stronger at more forward rapidities. This is a reflection of the centrality dependence seen in the R_{CP} above. Figure 8.61 and 8.62 show the R_{pPb} at fixed centrality, with multiple rapidities overlaid on the same plot, plotted as a function of p_{T} and $p = p_{\mathrm{T}} \cosh(y^*)$. Although the systematic uncertainties are somewhat larger, a similar scaling behavior is observed.

8.8 Systematic Uncertainties

This section gives an overview of the major sources of systematic uncertainty on the p+Pb and pp jet yields and the R_{CP} and R_{pPb}. Figure 8.67 summarizes the systematic uncertainty breakdown by source in the R_{CP}, Figure 8.68 summarizes the systematic uncertainty breakdown by source in the R_{pPb}, and Figure 8.69 summarizes the systematic uncertainty breakdown by source in the $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$. The different sources are described below.

8.8.1 Uncertainties due to trigger efficiency

A systematic uncertainty of 1% from the trigger efficiency is applied to the p+Pb and pp jet yields in the region of $20 < p_{\mathrm{T}} < 125$ GeV, where results from different EF jet triggers are patched together (described further in Section 8.4.2). This systematic uncertainty is taken to be uncorrelated in the different y^* and centrality bins. Thus, it shows up in quadrature in the R_{CP} (which is the ratio of p+Pb yields) and the R_{pPb} (which is the ratio of p+Pb yields to the pp cross-section).

8.8.2 Uncertainties due to cleaning efficiency

A systematic uncertainty of 0.5% on the correction due to the jet selection efficiency is applied, as is detailed in Section 8.3.5. The uncertainty arises from the maximum possible deviation from the efficiency observed in data from the fits used to parameterize the efficiency. Conservatively, the uncertainty is treated as uncorrelated between the centrality bins and between the p+Pb and the pp yields, and is taken to be the same in all y^* bins. Thus, it shows up in quadrature in the R_{CP} and the R_{pPb}.

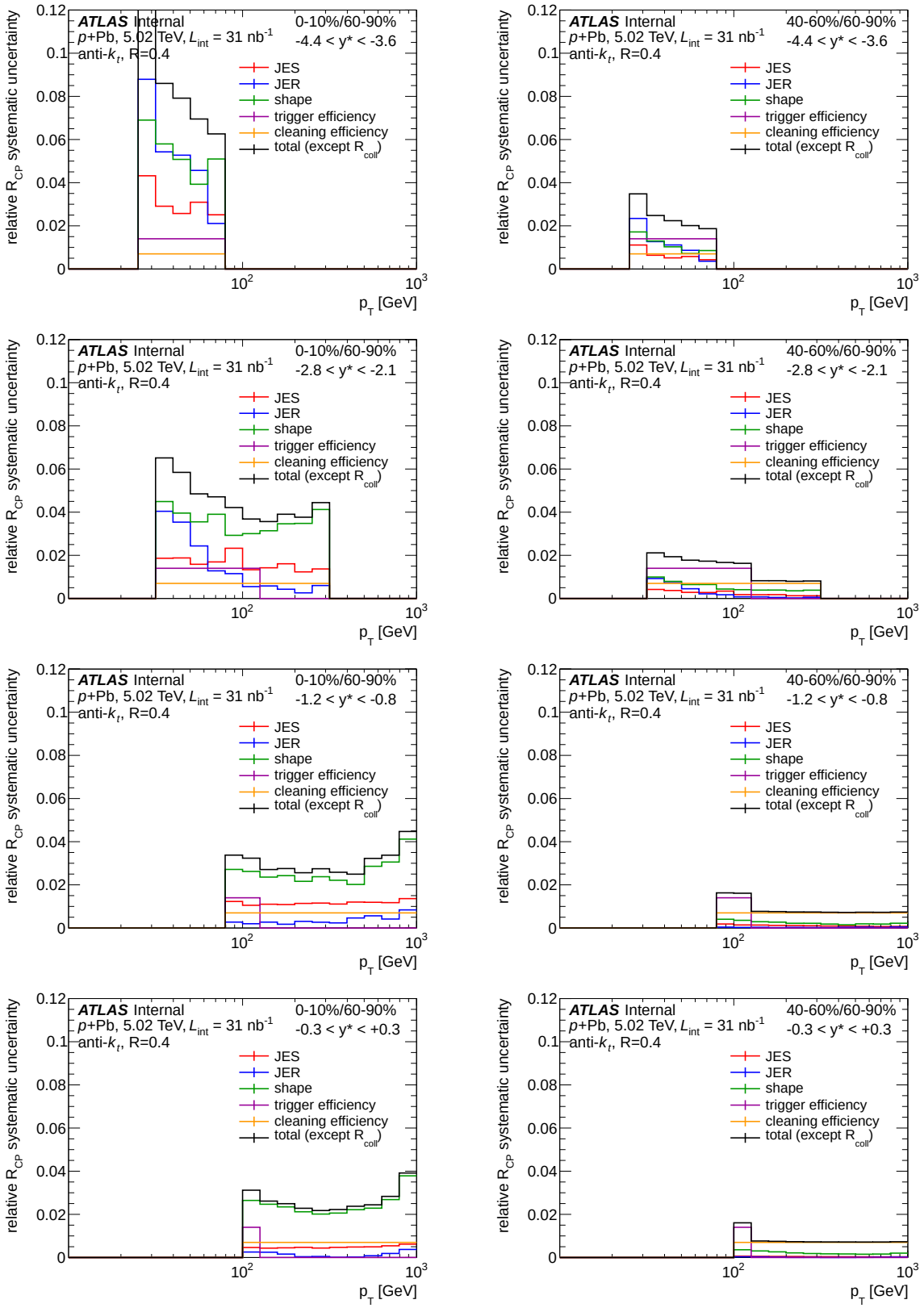


Figure 8.67: Breakdown of the systematic uncertainty on the R_{CP} from different sources, shown for a selection of rapidities in each row and the 0-10%/60-90% R_{CP} (left column) and 30-40%/60-90% R_{CP} (right column). The overall R_{coll} uncertainties are p_{T} and y^{*} independent, and are not shown.

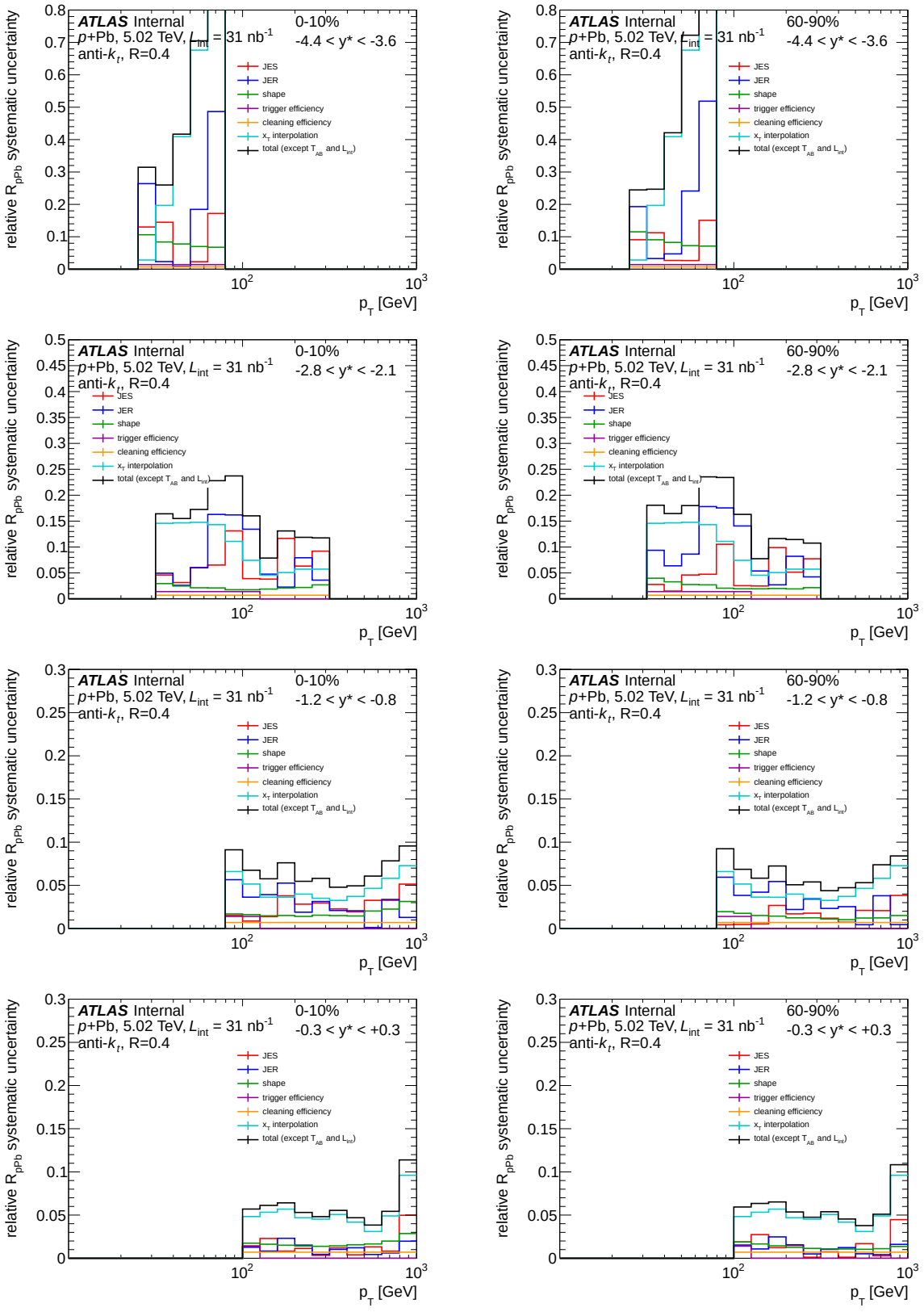


Figure 8.68: Breakdown of the systematic uncertainty on the R_{pPb} from different sources, shown for a selection of rapidities in each row and the 0-10% R_{pPb} (left column) and 60-90% R_{pPb} (right column). The overall T_{AB} uncertainties and luminosity uncertainties are p_{T} and y^* independent,

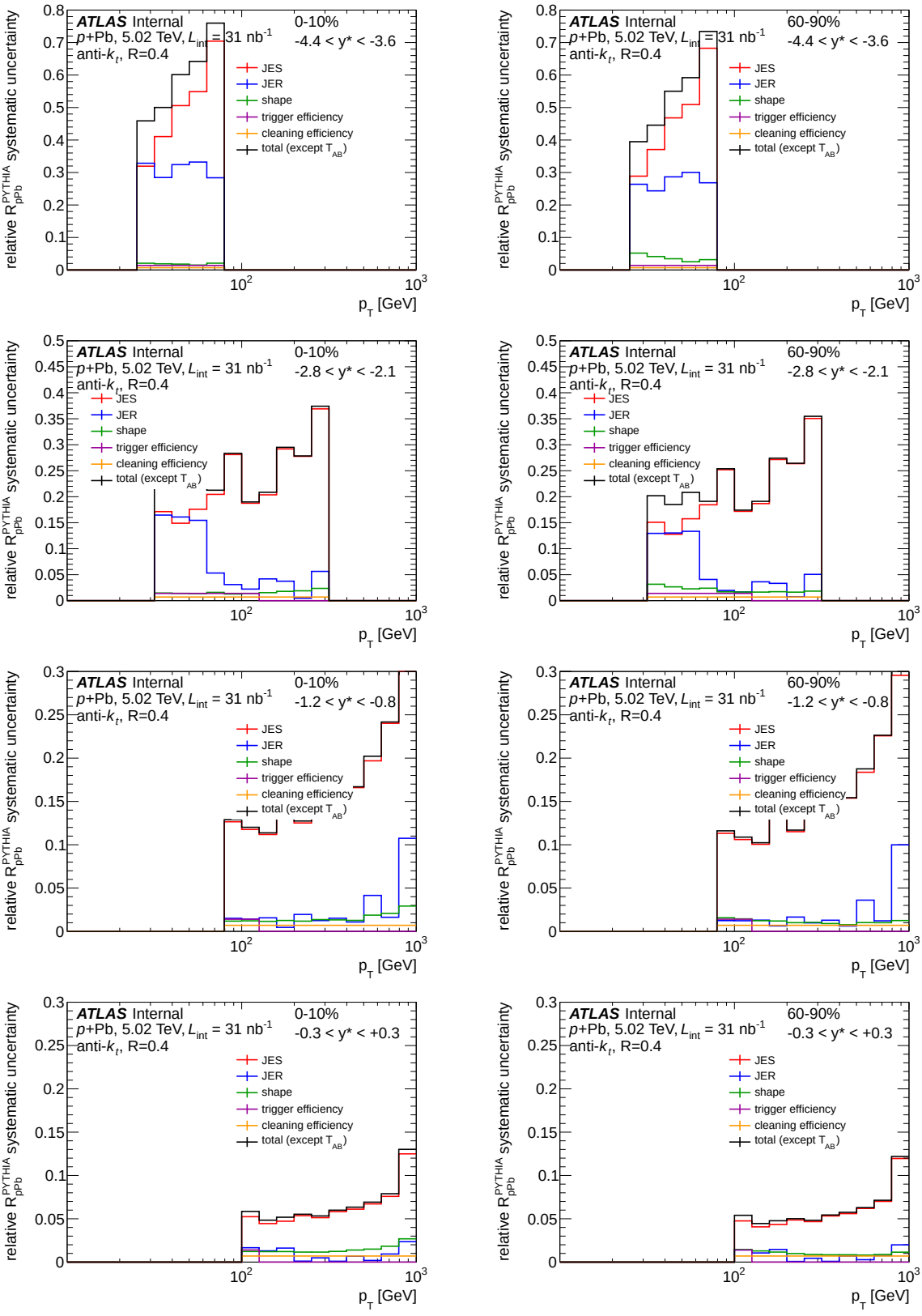


Figure 8.69: Breakdown of the systematic uncertainty on the $R_{\mathrm{pPb}}{}^{\mathrm{PYTHIA}}$ from different sources, shown for a selection of rapidities in each row and the 0-10% R_{pPb} (left column) and 60-90% R_{pPb} (right column). The overall T_{AB} uncertainties are p_{T} and y^* independent and are not shown.

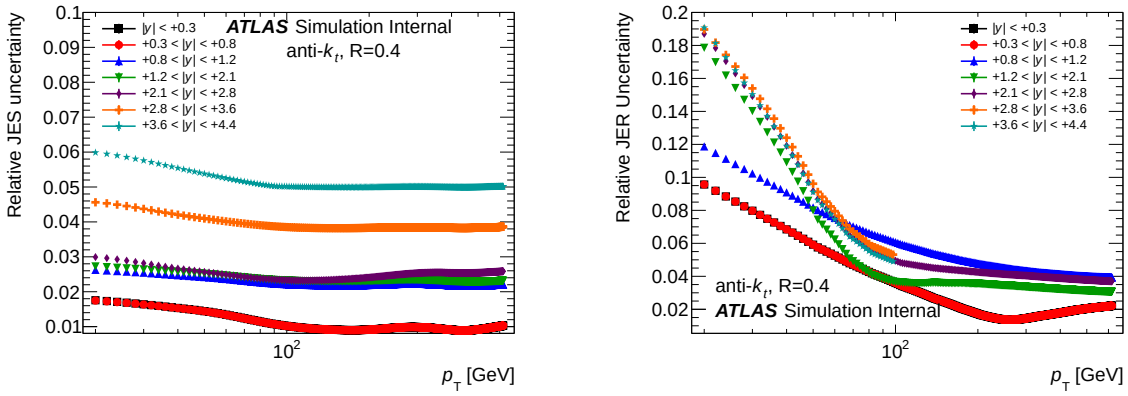


Figure 8.70: Systematic uncertainty in the JES (left) and JER (right) as a function of jet p_{T} and y.

8.8.3 Uncertainties in the jet energy scale and resolution

A systematic uncertainty is assigned to the R_{CP} and R_{pPb} from the possible variation in the jet energy scale and resolution.

First, there is the uncertainty from the jet energy scale. This quantifies how the bin-by-bin correction factors (described in detail in Section 8.5.4) for the p+Pb yields and the pp cross-section are affected by the assumption that the jet energy scale is higher or lower than what is used in the MC. This uncertainty is a p_{T}, and y-dependent quantity, and is determined from *in situ* ATLAS studies of the calorimeter response and systematic variations in the jet response in MC simulations[276]. It is determined with the `00-08-06` tag of the `JetUncertainties` package. The left plot in Figure 8.70 shows the JES uncertainty as a function of p_{T} and y.

The scale in MC is varied by this uncertainty to generate alternative correction factors. For each truth jet, the JES uncertainty X is calculated at the truth jet's p_{T} and y. The correction factors are determined again, with the $p_{\mathrm{T}}^{\mathrm{reco}}$ of each truth-matched reconstructed jet in the Signal MC changed according to

$$\frac{p_{\mathrm{T}}^{\mathrm{reco,new}} - p_{\mathrm{T}}^{\mathrm{truth}}}{p_{\mathrm{T}}^{\mathrm{truth}}} = \frac{p_{\mathrm{T}}^{\mathrm{reco,old}} - p_{\mathrm{T}}^{\mathrm{truth}}}{p_{\mathrm{T}}^{\mathrm{truth}}} \pm X \tag{8.55}$$

$$p_{\mathrm{T}}^{\mathrm{reco,new}} = p_{\mathrm{T}}^{\mathrm{reco,old}} \pm X p_{\mathrm{T}}^{\mathrm{truth}} \tag{8.56}$$

with the + and − systematic changes in the correction factors recorded separately. While a variation in the JES of $\approx X\%$ typically results in a $\approx 5X\%$ change in the yield (since it is a scale error on a steeply-falling power law spectrum), the JES uncertainty largely cancels in the ratio between p+Pb yields at different centralities. It also largely cancels in the $R_{\rm pPb}$ (the p+Pb/ pp comparison), but less so, for two reasons. First, the pp results are $x_{\rm T}$ scaled and thus the change in the $R_{\rm pPb}$ at fixed $p_{\rm T}$ is caused by changes at different $p_{\rm T}$ values in the 2.76 TeV pp and 5.02 TeV p+Pb spectra. Second, the pp and p+Pb systems have a relative shift between their centers of mass, so that a given value of y^* in one system corresponds to different values of $y^{\rm lab}$ between the two. Thus, the cancellation in the $R_{\rm pPb}$ is not as complete as in the $R_{\rm CP}$.

Second, there is the uncertainty from the jet energy resolution. This quantifies how the bin-by-bin correction factors are affected by the assumption that the jet energy resolution is systematically higher than what it described by the MC. This is modeled as an additional resolution $\sigma_{\rm syst}$, and is determined in data from *in situ* studies of dijet energy balance[277]. It is provided by the `02-00-02` tag of the `JetResolution` package. The right plot in Figure 8.70 shows the JER uncertainty as a function of jet $p_{\rm T}$ and y.

For each truth jet, the $\sigma_{\rm syst}$ is determined. Then, a smearing factor Y is determined from sampling a Gaussian distribution with mean 0 and standard deviation $\sigma_{\rm syst}$. The correction factors are determined again, with the $p_{\rm T}^{\rm reco}$ of each truth-matched reconstructed jet in the Signal MC changed according to

$$p_{\rm T}^{\rm reco,new} = p_{\rm T}^{\rm reco,old} + Y p_{\rm T}^{\rm truth} \tag{8.57}$$

This has the effect of increasing the effective resolution from the original resolution σ to $\sqrt{\sigma^2 + \sigma_{\rm syst}^2}$. Similarly, the JER uncertainty mostly cancels in the $R_{\rm CP}$, and largely but not completely cancels in the $R_{\rm pPb}$.

8.8.4 Uncertainties in the unfolding procedure from the spectral shape

This quantifies the uncertainty arising from a mismatch between the shape of the truth spectrum used to generate the correction factors and the data to which they are applied. In all cases where correction factors are used, the MC truth is reweighted by the ratio of data/reconstructed MC so that the MC better captures the spectral shape observed in data.

This effect is small for the *pp* cross-section and the 0-90% p+Pb yields, where the data and MC only have small differences in their shape. Thus, the scale of the correction is $<$ 2-3%. However, the change in shape of the centrality-dependent p+Pb yields (a physics effect) is larger. The correction factors can change by as much as 5% in the R_{CP} for the most central to the most peripheral, with the change larger at small p_{T}, up to 10% at the lowest p_{T} values used. On the other hand, this uncertainty is *smaller* in the R_{pPb}, since the shape differences between any individual centrality-selected p+Pb yield and the *pp* spectrum are smaller than they are between the yields in the two most outlying p+Pb centralities.

For the results presented here, the whole magnitude of the change in the correction factors from reweighing the truth is conservatively taken as the systematic uncertainty in the unfolding procedure due to the shape uncertainty.

8.8.5 Uncertainties in the determination of geometric quantities (T_{AB}, R_{coll})

The systematic uncertainties on the geometric quantities R_{coll} and T_{AB} are described in detail in [263]. R_{coll}, which is used in the R_{CP}, is the ratio of the number of binary nucleon-nucleon collisions in the given centrality selection relative to the 60-90% centrality (Equation 8.4), and the nuclear thickness function T_{AB}, which is used in the R_{pPb}, is the nucleonic luminosity for p+Pb collisions in a given centrality (Equation 8.6).

The uncertainty is derived from a number of sources in the centrality analysis, including the energy scale in MC, the total event selection efficiency, the description of the inelastic *pp* reference, variations on the Glauber model (including the uncertainty on the assumed inelastic *pp* cross-section at $\sqrt{s} = 5.02$ TeV), and the uncertainty from modeling the response in the FCal as a function of the number of nucleon participants. The uncertainties are generally asymmetric and are summarized in Table 8.2. They are centrality-dependent, but independent of p_{T} and y.

8.8.6 Uncertainties due to x_{T} interpolation

The *pp* cross-section reference at 5.02 TeV is sensitive to the systematics of the x_{T} interpolation procedure, which is detailed in Section 8.7.5. In particular, the measurements of ρ used for the interpolation has systematic uncertainties. Since ρ is the ratio of two jet cross-sections measured in ATLAS, there is a partial cancellation of systematic uncertainties between the two measurements.

Thus, the remaining uncertainty is treated as completely uncorrelated with the present $\sqrt{s} = 2.76$ TeV measurement. Furthermore, the systematic uncertainties are extremely correlated in x_{T} (and therefore p_{T}) point to point.

From Equations 8.51 and 8.54, the dependence of the $\sqrt{s} = 5.02$ TeV pp cross-section on ρ is

$$\sigma^{pp} \propto \left(\frac{2.76}{5.02}\right)^{\log \rho(x_{\mathrm{T}})/\log(7/2.76)} \tag{8.58}$$

$$\sigma^{pp} \propto \rho(x_{\mathrm{T}})^{\log(2.76/5.02)/\log(7/2.76)} \approx \rho(x_{\mathrm{T}})^{-0.643} \tag{8.59}$$

Thus, the relative systematic uncertainty is simple to propagate via

$$\frac{\sigma_{\sigma^{pp}}}{\sigma^{pp}} = 0.643\frac{\sigma_{\rho}}{\rho} \tag{8.60}$$

Since ρ is measured in different bins than the $\sqrt{s} = 2.76$ TeV pp cross-section used here, the uncertainty on the x_{T}-scaled cross-section at a given x_{T} is taken by interpolating the systematic uncertainty from the neighboring bins used in the measurement. The uncertainties on ρ are p_{T} and y^* dependent, being larger in the more forward rapidities and at lower p_{T}. ρ is one of the dominant systematics in the measurement. At mid-rapidity and high-p_{T}, where the uncertainty is smallest, it is 5%.

In addition to the systematic uncertainty visible on the plots of ρ, there is a 4.3% overall uncertainty from the uncertainties on the luminosity for the 7 TeV and 2.76 TeV datasets. Since this is an overall luminosity uncertainty, this 4.3% is stated separately in the Figure caption rather than shown on the plot.

8.8.7 Uncertainties due to luminosity

The determination of the pp cross-section depends directly ($d^2\sigma/dp_{\mathrm{T}}dy \propto 1/\mathcal{L}_{int}$) on the luminosity calibration, which unfortunately is in a preliminary stage. On advice of ATLAS luminosity analyzers, a systematic uncertainty of 3% is applied to the pp cross-section and the R_{pPb}. This uncertainty is independent of p_{T} and y.

Chapter 9

Discussion

If we shadows have offended,
Think but this, and all is mended,
That you have but slumber'd here
While these visions did appear.
And this weak and idle theme,
No more yielding but a dream,
Gentles, do not reprehend:
if you pardon, we will mend:

Midsummer Night's Dream, Act V, Scene I

The results presented here represent the first attempt to systematically map out the partonic structure of the nucleus when probed at LHC energies. Several distinct measurements have been presented which, when told together, paint a surprising but consistent picture of cold nuclear matter effects on hard scattered partons. First, fully reconstructed jets were measured in deuteron-gold (d+Au) collisions at mid-rapidity at RHIC. The RHIC data showed an unexpected centrality (impact parameter) dependence in the rate of hard processes. When compared against the expected rate from the geometric overlap of the deuteron with the local nuclear thickness, the apparent rate of jets from the periphery of the nucleus (tagged as low-activity events) was found to be increased, while the rate of jets in the center of the nucleus (tagged as high-activity events) was found to be decreased. On the other hand, when integrating over impact parameter, no modification of

the jet rate relative to the total partonic content in the nucleus was observed. Such a centrality dependence in the modification factor for jets was unexpected.

Shortly after the observation of this anomalous centrality dependence, the LHC recorded proton-lead data with a factor of 25 increase in the center of mass energy and substantially more accessible kinematic range for jet measurements. Before any impact parameter-dependent measurements could be performed, the proton-lead (p+Pb) data had to be understood in a way that made the different geometric configurations of proton-nucleus collisions experimentally accessible. Thus, the quantities of interest were extracted by mapping the range of p+Pb events in a geometric model to the data. Finally, the rate of fully reconstructed jets was measured as a function of p+Pb event centrality (and thus, apparent impact parameter) over many units of rapidity and to nearly 1 TeV in the transverse momentum of the jets. Just as was seen at RHIC, the relative rate of jets coming from events with small impact parameter was smaller than the geometric expectation, while the rate of jets from events with a large impact parameter was larger. On the other hand, the impact parameter integrated rate of jets was not found to be strongly modified. Furthermore, the suppression and enhancement patterns in all rapidities were found to obey a simple scaling behavior in the total jet energy.

We discuss each of the results here in turn.

9.1 Jets in d+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV

The geometric-scaled rate of jets in central d+Au collisions compared to that in peripheral d+Au collisions (R_{CP}) is shown in Figure 9.1. At $p_{\mathrm{T}} = 10$ GeV, the R_{CP} in all centrality selections is consistent with 1, signifying a jet rate in line with the geometric expectation. However, the R_{CP} monotonically decreases with increasing p_{T} in all the centrality selections. The most central to most peripheral ratio 0-20%/60-88% reaches ≈ 0.6 at $p_{\mathrm{T}} \approx 40$ GeV. This is a strong effect – it represents a relative increase of 2.5× more jets in peripheral d+Au events relative to central d+Au events when all of the geometric factors have been taken out. Figure 9.2 shows the geometric-scaled rate of jets in inclusive, central and peripheral collisions, but this time compared to a pp collision reference. The left side of the Figure shows the R_{dAu} for 0-88% (minimum bias) d+Au collisions, which is everywhere flat and consistent with 1. Despite the significant modification observed in

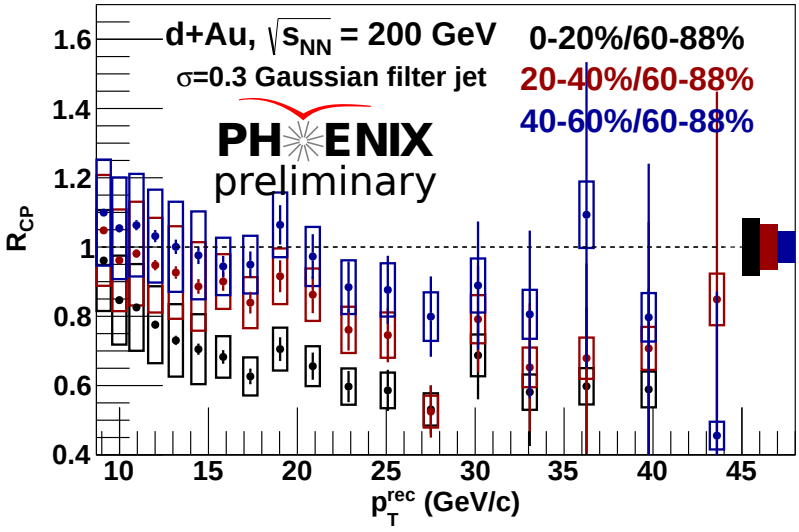


Figure 9.1: Nuclear modification factor, R_{CP}, for jets in d+Au collisions at RHIC, from Chapter 6.

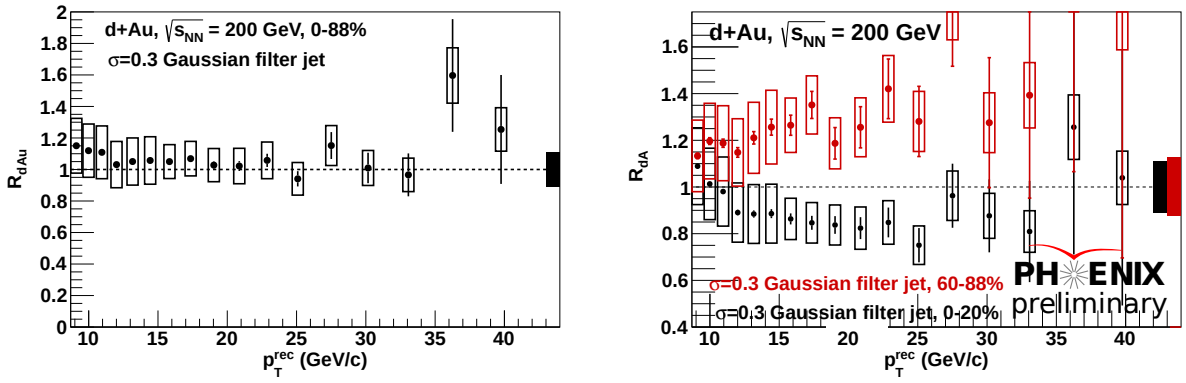


Figure 9.2: Nuclear modification factor, R_{dAu}, for jets in inclusive (left) and centrality-selected (right) d+Au collisions at RHIC, from Chapter 6.

the R_{CP}, the minimum bias R_{dAu} is consistent with no effect. Thus, the absolute modification in the central and peripheral R_{dAu} must be such that together they cause the observed effect in the R_{CP} while giving no effect when averaged over. At $p_{\mathrm{T}} = 10$ GeV, the R_{dAu} in both central and peripheral events is consistent with 1, signifying no change in the jet rate from cold nuclear matter effects. However, the central R_{dAu} decreases, and the peripheral R_{dAu} increases, with p_{T} up to the

limit of statistics, where they reach ≈ 0.85 and ≈ 1.3, respectively. The data can be interpreted as the total number of jets from the geometric expectation being preserved, but moved around in centrality.

The data are challenging to explain. A slight suppression in central events could probably be made consistent with initial state energy loss calculations without tension. However, the central $R_{\rm dAu}$ is actually consistent with 1, just barely within the full systematic uncertainties of the measurement. On the other hand, the rise in the peripheral $R_{\rm dAu}$ is unambiguous. Furthermore, the uncertainties in the $R_{\rm dAu}$ arising from the pp reference are completely correlated, and the centrality dependence in the jet rate is definite and large. In fact, the suppression in the $R_{\rm CP}$ can be seen by eye before any corrections for the trigger efficiency or bin migration in jet $p_{\rm T}$ from the presence of the d+Au underlying event (UE). Furthermore, a recent measurement of neutral pions in the same data set also sees a similarly suppressed $R_{\rm CP}$, albeit with a different sensitivity to the energy scale. Thus, the measurement is experimentally robust.

This relatively large centrality dependence is difficult to explain within theoretical frameworks. At such high-Q^2, nPDF effects are expected to be very small, and nPDF fits which model the impact parameter dependence of the modification cannot accomodate such large centrality-dependent changes. Furthermore, it is difficult to have a physical picture of the Cronin effect and/or transverse momentum broadening which causes a stronger increase in peripheral collisions than it does the central. Several other explanations can be suggested but none are immediately satisfactory. For example, perhaps the deuteron structure function at high x is not well known[247; 248]. Actually within the kinematic range measured here ($x \sim$ 0.1-0.5), this not the case. Another suggestion is that perhaps the centrality determination is just wrong. But the geometric-scaling in the d+Au centralities works very well for single particles with $p_{\rm T} <$ 10-15 GeV, and there are cross-checks on the Glauber model results available in data. A third suggestion is that while for the minimum bias d+Au events, the centrality categorization works, perhaps there is a bias in the centrality selection in which jets are preferentially associated with low-activity events not reflective of the d+Au event geometry. Actually, studies of the possible "centrality bias" in the presence of a jet in PHENIX[244] show that if anything the measured $R_{\rm CP}$ is systematically higher than it really is (e.g. the effect goes the other way). However, we will return to this point.

It should also be noted that this data does not challenge the picture of jet quenching observed

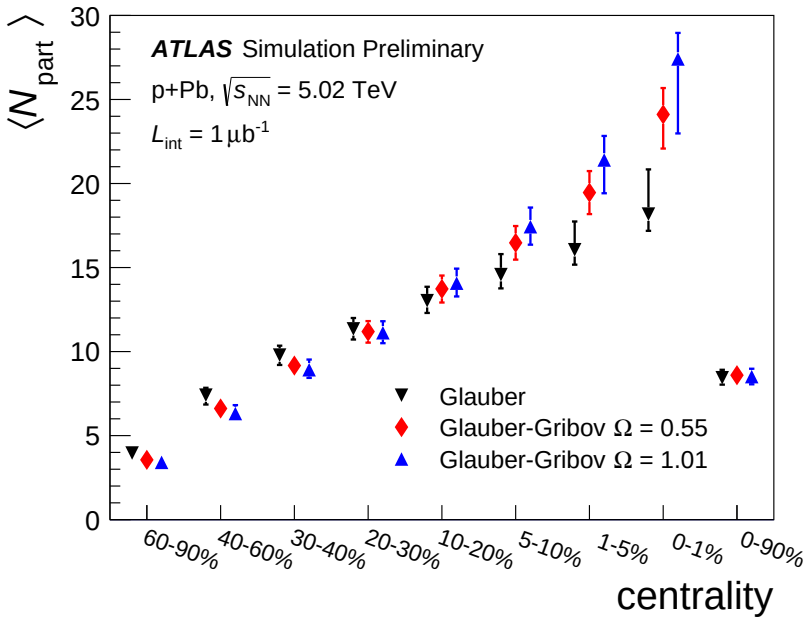


Figure 9.3: Mean number of nucleon participants, $\langle N_{\rm part}\rangle$, in the p+Pb centrality selections. The three Glauber model variants are each shown, from [264].

at RHIC. The $R_{\rm dAu}$ and $R_{\rm CP}$ for jets is consistent with no CNM effect at $p_{\rm T} = 10$ GeV, and measurement of single hadrons in Au+Au collisions show that they are suppressed by a factor of approximately 5 below this $p_{\rm T}$. The data do, however, complicate the use of centrality-selected data in the extraction of impact parameter dependent nPDF's.

9.2 Centrality determination in p+Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV

The mean number of nucleon participants in each p+Pb centrality selection are shown in Figure 9.3, with comparisons between the default Glauber model and Glauber-Gribov variants. It can be seen that while the three models agree very well on the mean number of participants in minimum bias (0-90%) p+Pb collisions, they give substantially different results for the number of participants in extremely central (for example, 0-1%) or peripheral (for example, 60-90%) events.

The centrality analysis included a few features, born out of necessity, which were not used in the centrality categorization of d+Au collisions at RHIC, where the mapping of the Glauber $N_{\rm part}$ distribution onto the observed data has historically been more straightforward. First, the p+Pb event

topologies and contributions to the minimum bias event sample from diffractive and photonuclear excitation were investigated with a study of the reconstructed pseudorapidity distributions on either pseudorapidity edge of the detector. Second, the centrality signal in p+Pb events was explicitly corrected for effects such as the vertex-dependent changes in acceptance and small time-dependent variations in electronic pedestals. In this way, no additional decorrelation between the $N_{\rm part}$ of p+Pb events and the centrality, as measured in the Pb-going forward calorimeter, is introduced. Third, two alternatives to the Glauber model were used to generate the probability distribution for a p+Pb event to have $N_{\rm part}$ total nucleon participants, based on the idea that in high energy p+A collisions, there must surely be event by event fluctuations in the partonic configuration of the projectile. Thus, at fixed impact parameter, fluctuations in the number of participants arise not only from event-by-event fluctuations in the nucleon positions but also from different "snapshots" of proton wavefunction.

Although the data was able to be reproduced by each of the three models, there were several indications that the Glauber-Gribov models, especially the model with the smaller $\sigma_{\rm NN}$ fluctuations ($\Omega = 0.55$), may be a more appropriate model of high-energy proton-nucleus collisions. First, while the results in the default Glauber model suffered from large systematic changes in response to the choice of fit model, the Glauber-Gribov model with $\Omega = 0.55$ required fewer non-linear deviations from the "wounded nucleon" assumptions in the best global fits. That is, the Glauber-Gribov model could reproduce the p+Pb data with a more natural description of the $N_{\rm part}$ dependence of the soft underlying event $E_{\rm T}$. Second, the Glauber-Gribov model better describes the data in the low event activity (peripheral) region, where the composition of p+Pb events selected by a small FCal $E_{\rm T}$ are not well understood and may easily have a different composition than what it suggested by the default Glauber model. Interestingly, the best Glauber-Gribov fits also include a contribution of events that are consistent with diffractive excitation of the Pb nucleus, underlining the growing need to understand the role of diffraction in hadronic collisions at LHC energies.

Third, a recent measurement of the ATLAS soft charged particle multiplicity in p+Pb collisions[264] could help interpretating which Glauber model should be taken as a more appropriate model of the p+Pb geometries. Figure 9.4 shows the multiplicity per participant pair plotted as a function of the number of participants for each of the three models. It can be seen that using the default Glauber model to estimate the mean number of participants results in a substantial rise of multiplicity per

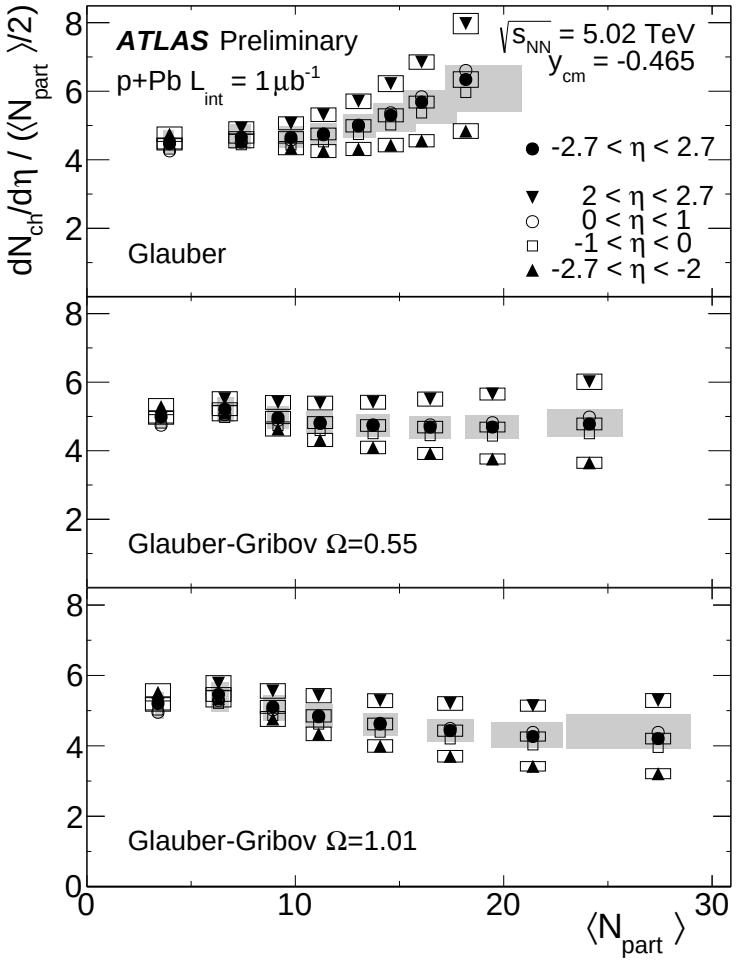


Figure 9.4: ATLAS measurement of the charged particle multiplicity per participant pair, plotted in rapidity bins as a function of the number of participants. Results with N_{part} extracted from the different Glauber model variants are shown, from [264].

nucleon pair, increashing 50% *per pair* when the number of participants changes from $N_{\text{part}} \approx 12$ to $N_{\text{part}} \approx 18$. On the other hand, the Glauber-Gribov models show either a flat N_{part} dependence ($\Omega = 0.55$) or a small decrease in the per pair multiplicity after a few collisions ($\Omega = 1.01$), both of which are more physically plausible.

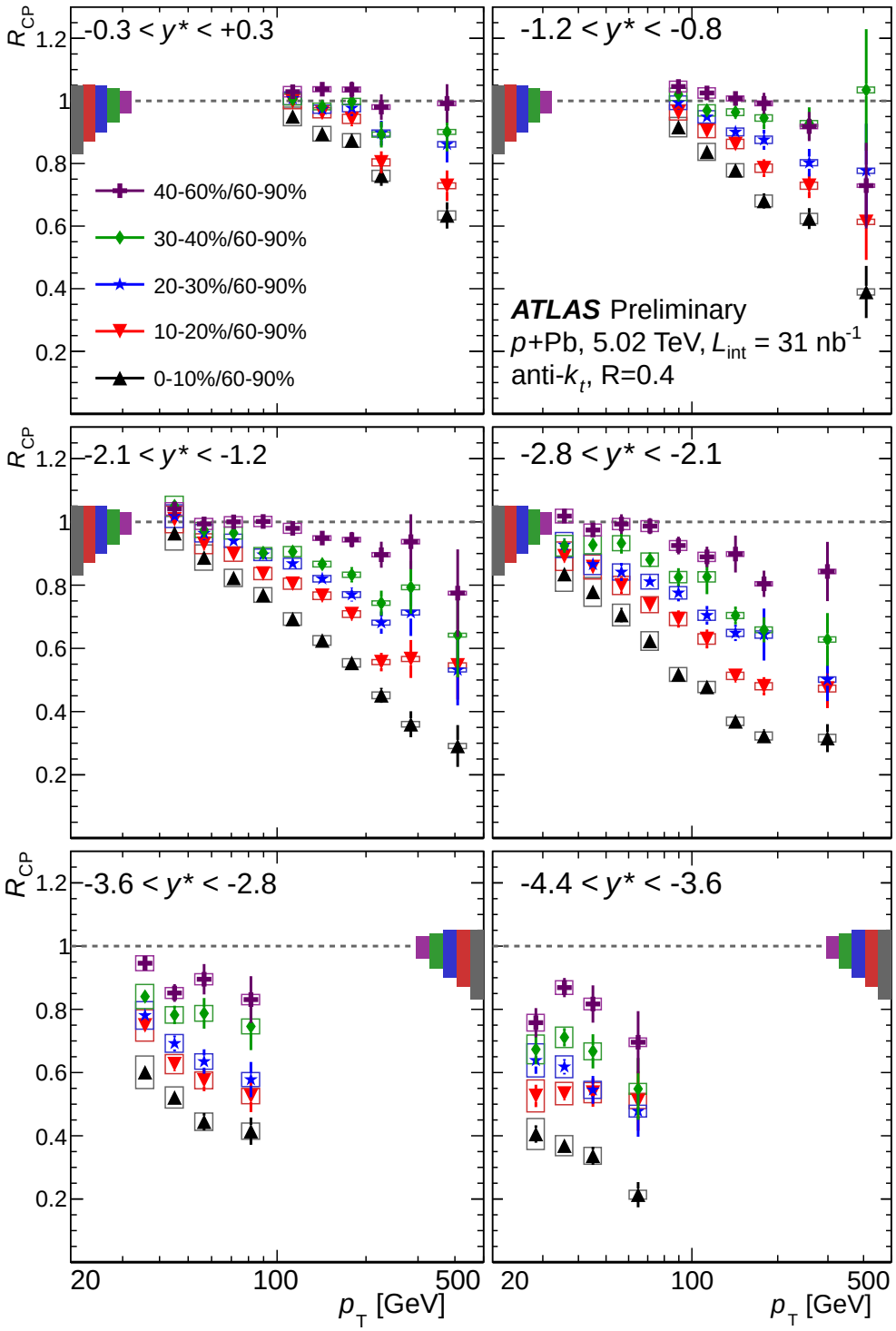


Figure 9.5: Nuclear modification factor, R_{CP}, for jets in p+Pb collisions at the LHC, six rapidity bins, from [278].

9.3 Jets in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV

The geometric-scaled rate of jets in different p+Pb centrality selections compared to that in peripheral collisions is shown in Figure 9.5 for two rapidity selections: near mid-rapidity ($-0.8 < y^* < -0.3$) and at mid-forward rapidity ($-2.1 < y^* < -1.2$, where $y^* < 0$ corresponds to the p-going direction and thus small nuclear-x and large proton-x). Within the Glauber model, the centrality selections correspond to probing different regions of the Pb nucleus. The R_{CP} is constructed in the region where the detector response to jets has been determined to be independent of the p+Pb centrality. The R_{CP} is monotonically decreasing with increasing p_{T} in all centralities and, furthermore, is also monotonically decreasing (with the exception of statistical fluctuations) with increasingly central events. All of the centrality selected yields, even the 40-60%, appear to have a steeper p_{T} distribution than the 60-90% reference. Furthermore, the R_{CP} appears to be broadly consistent with a linear decrease as a function of $\log p_{\mathrm{T}}$. Another way to think about this pattern is that the nuclear modification results in each centrality-selected spectrum being successively steeper with more central selections. The suppression is quite strong. For example, for $p_{\mathrm{T}} = 300$ GeV jets in $-2.1 < y^* < -1.2$, the $R_{\mathrm{CP}} \approx 0.3$ in the 0-10%/60-90% selection, implying a factor of 3 difference in the relative rates of jet production between the core and periphery of the Pb nucleus, once the geometric factors have been scaled out.

Furthermore, the suppression pattern has an interesting rapidity dependence. The left side of Figure 9.6 shows the R_{CP} in the six most forward rapidity bins, separately in two centrality selections. Thus, at fixed centrality and p_{T}, the R_{CP} also evolves monotonically in rapidity, with a systematically larger suppression in increasingly forward ($y^* < 0$) bins. Interestingly, the linear trend of all the R_{CP}'s on a logarithmic p_{T} scale suggests that there may be a scale transformation relating them all. In fact, the horizontal axis of the data can be rescaled by a factor $\cosh(y^*)$ so that all the R_{CP} series are plotted as a function of function of $p \approx p_{\mathrm{T}} \cosh(y^*)$, where y^* is the midpoint of the rapidity bin and p is thus the approximate total jet energy. The R_{CP} plotted against p is shown on the right side of Figure 9.6. Remarkably, when plotted this way the R_{CP} in the different rapidity bins seem to line up together, suggesting that

$$R_{\mathrm{CP}}(p_{\mathrm{T}}, y^*) = R_{\mathrm{CP}}(p) \tag{9.1}$$

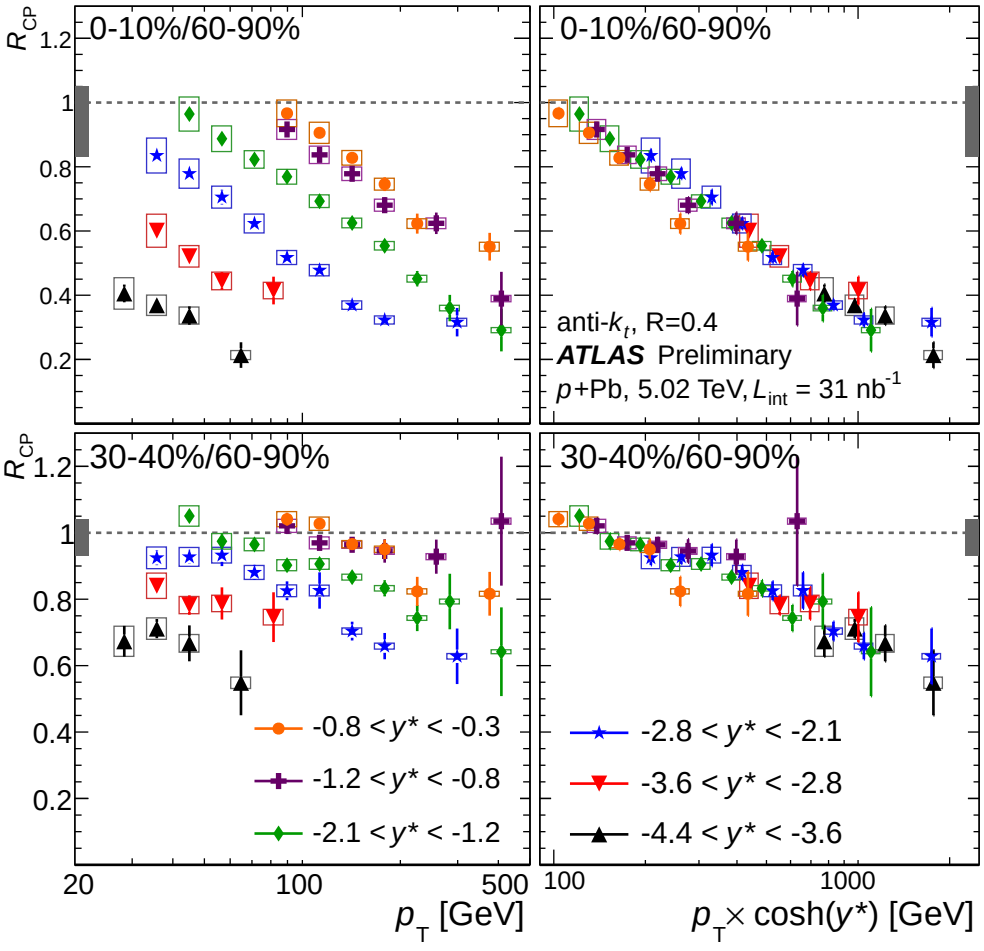


Figure 9.6: Nuclear modification factor, R_{CP}, for jets in p+Pb collisions, for 0-10%/60-90% (top) and 30-40%/60-90% (right) centrality selection, showing multiple rapidity bins on the same plot. The panels on the right show the R_{CP} plotted as a function of p, from [278].

This phenomenon is present in each centrality bins, with a different slope for the $R_{\mathrm{CP}}(p)$ in each. This scaling behavior suggests that the mechanism responsible for the suppression may have a common origin and only be a function of a simple kinematic variable – the total jet energy.

While the R_{CP} indicates that the relative prevalence of jets is a function of the impact parameter in an x-dependent way, only the R_{pPb} can measure changes in the absolute rate of jets. The R_{pPb} for 0-90% p+Pb events is shown in Figure 9.7 for several rapidity bins. The central value of the inclusive R_{pPb} are consistent with a slight enhancement (≈ 1.1) in the jet yield at fixed p_{T} relative

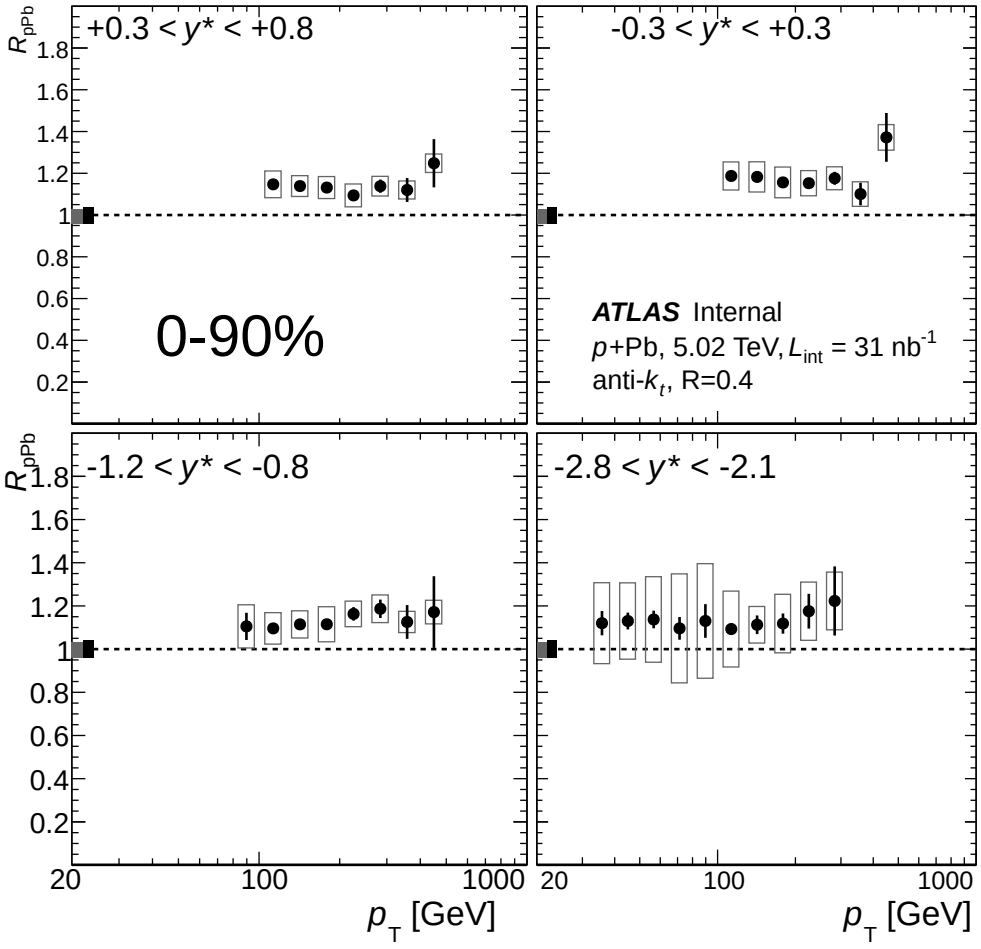


Figure 9.7: Nuclear modification factor R_{pPb} for 0-90% events, shown in four rapidity bins.

to pp collisions scaled by a geometric factor accounting for the thickness of the nucleus. However, the systematic uncertainties preclude a strong statement about the size of the enhancement. In all rapidity bins the R_{pPb} is flat and in most bins it is consistent with 1 within uncertainties. Certainly, any large modifications in the jet yield are ruled out. Thus, when p+Pb collisions are considered integrating over impact parameter, the jet rate is consistent with the geometric scaling expectation (or only very slightly above it). By contrast, the R_{CP} has a very strong centrality dependence. By the definition of the R_{pPb} and the R_{CP}, no effect in the inclusive R_{pPb} and a strong effect in the R_{CP} constrains the central and peripheral-selected R_{pPb} results to lie on opposite sides of $R_{\mathrm{pPb}} = 1$.

Figure 9.8 shows the centrality-dependent R_{pPb} for the same rapidity bins as the previous figure.

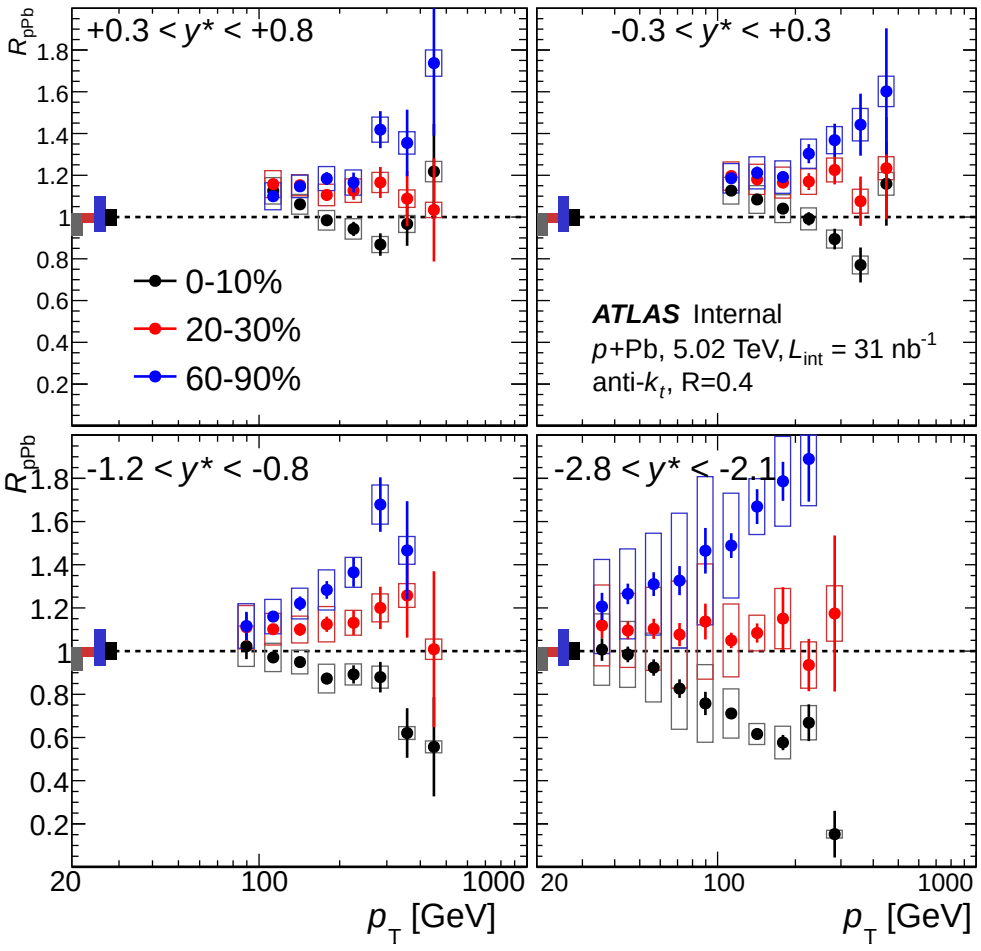


Figure 9.8: Nuclear modification factor, R_{pPb}, for different centrality selections, shown in four rapidity bins.

The most central (0-10%) R_{pPb} has a negative slope and is consistent with an overall suppression at high-p_{T}. On the other hand, the peripheral R_{pPb} (60-90%) has a positive slope and is consistent with an overall enhancement. Generally, R_{pPb} values in the other centrality selections fall smoothly between them. Interestingly, the R_{pPb} in a mid-central bin (20-30% centrality) is the closest to flat. Furthermore, the centrality-dependent R_{pPb}'s seem to monotonically increase or decrease up to the limit of statistics. At no point in p_{T} do they appear to saturate. Finally, the rapidity dependence of the most central and most peripheral R_{pPb} is plotted in Figure 9.9. There is a strong rapidity dependence in the size of the modification, but when plotted as a function of p, the central and

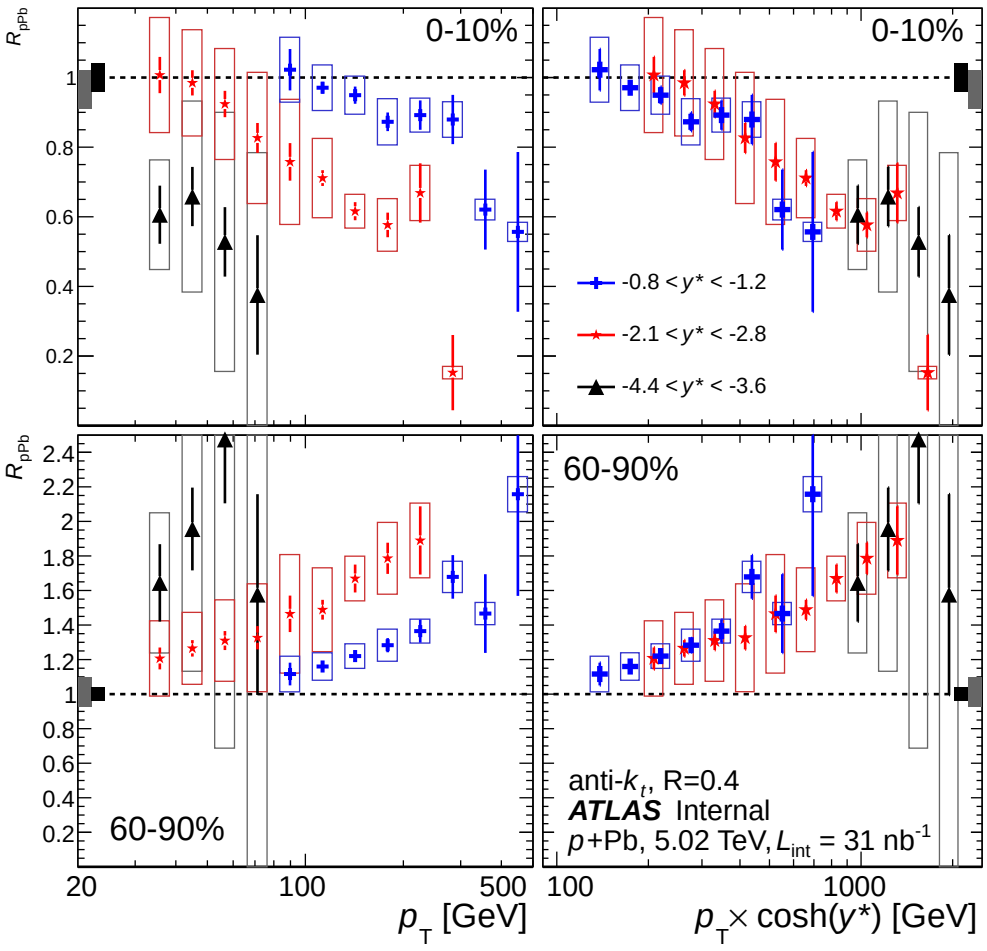


Figure 9.9: Nuclear modification factor, R_{pPb}, for jets in p+Pb collisions, for 0-10% (top) and 60-90% (right) centrality selection, showing multiple rapidity bins on the same plot. The panels on the right show the R_{pPb} plotted as a function of p.

peripheral R_{pPb} each appear to be a simple function of the total jet energy.

As with the PHENIX data, these results are not yet fully understood. Qualitatively, the mid-rapidity results are remarkably similar to PHENIX, in which the large effect in the R_{CP} is split between an enhancement in peripheral collisions and a suppression in central one, with no net modification in inclusive p+Pb collisions. Also as with the PHENIX result, it is difficult to come up with an easy theoretical explanation for the data. For one, the magnitude of the effect actually increases at higher p_{T}, whereas all nuclear-enhanced leading twist contributions are suppressed

by powers of Q^{-2}. Second, it is challenging to come up with a physics effect (for example, such as transverse momentum broadening) which would cause an apparent increase of jets at fixed in peripheral jets while simultaneously causing a supression in central jets, *while also having a larger effect at higher* p_{T}. Third, the impact parameter dependence of nPDF sets is very small compared with the size of the effect in data.

It has been suggested that the data can be explained by an effective "rapidity shift", in which selecting central or peripheral events systematically boosts the jet dN/dy distributions in the nucleus-going or proton-going direction, respectively. However, this explanation is not theoretically well founded, since impact-parameter-dependent rapidity shifts of the parton-parton center of mass system do not naturally enter into a pQCD/factorization approach. Second, an examination of the data shows that this observation is not correct. If, whatever the physical mechanism, the results corresponded to an overall shift of the jet rapidity in the different centralities, then the trend in the R_{CP} should be reversed for backward rapidity $y^* > 0$, such that the R_{CP} and central R_{pPb} would rise above 1 and the peripheral R_{pPb} would fall below. In fact, data in the backwards bins used in this analysis show precisely the opposite effect.

While one of the original goals of this measurement was to explore the transition from the dilute to saturated partonic regime at low-x using low-p_{T} forward jets, the large modification observed at forward rapidities complicates any measurement of non-linear QCD effects, which are not expected to survive at such high p_{T}. Finally, the geometric factors in the R_{CP} and R_{pPb} shown here have been derived exclusively from the default Glauber model. The main feature of the Glauber-Gribov model, which is the fluctuating transverse size of the low-x partons in the proton, may not be applicable in a straightforward way to p+A events with a hard scattering. Thus, only the default Glauber model is used although, if the Glauber-Gribov variant models were used to estimate R_{coll} and T_{AB}, the observed suppression and enhancement in central and peripheral events, respectively, would actually *increase* in magnitude.

9.4 Conclusion

Taken together, the PHENIX and ATLAS data are intriguing. Qualitatively, they demonstrate similar patterns in the modification of the jet rate in central and peripheral p+A (in this case,

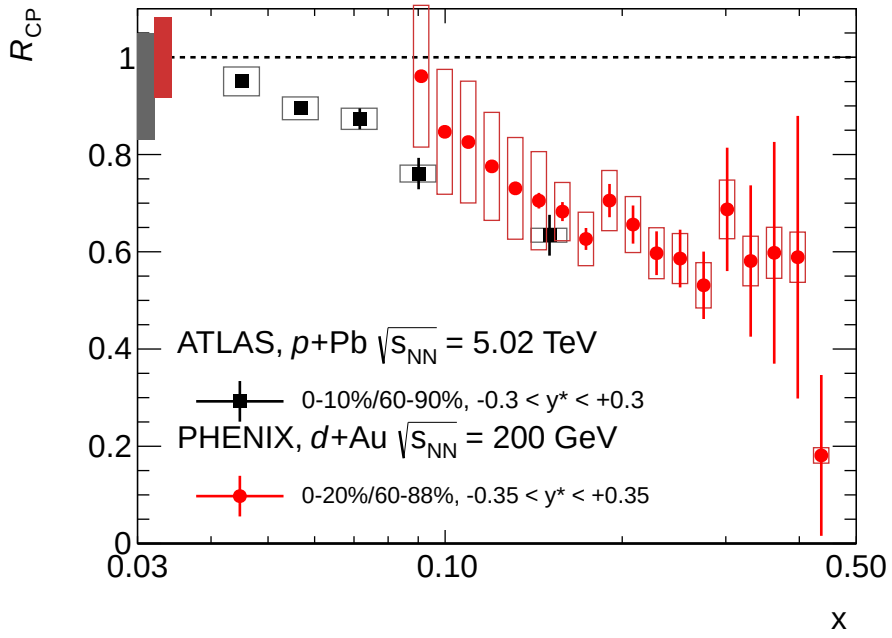


Figure 9.10: Comparison of the $R_{\rm CP}$ for mid-rapidity jets in d+Au collisions at $\sqrt{s_{\rm NN}} = 200$ GeV and p+Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV, as a function of x.

d+Au and p+Pb) collisions, in two collision systems which differ by a factor of 25 in the center of mass energy. In both systems, the high-$p_{\rm T}$ jet yield in central events is suppressed relative to geometric expectations, and enhanced in peripheral events. Since the rapidity dependence of the results in ATLAS appear to be explained by a single kinematic variable, it may be possible that the PHENIX and ATLAS results could be related through a simple scaling, as well. The most natural choice for such a scaling is to compare the results at mid-rapidity (where PHENIX has acceptance) and scale the horizontal axes by the nucleon beam energy ($E_{\rm beam} = 100$ GeV in d+Au collisions at RHIC, and $= 2.5$ TeV in p+Pb collisions at the LHC). Thus, Figure 9.10 plots the most central to most peripheral $R_{\rm CP}$ in both systems as a function of the variable $p_{\rm T}/E_{\rm beam}$ (which at mid-rapidity $\approx x$). The $R_{\rm CP}$ is qualitatively similar at RHIC and the LHC. The central values of the data points imply that the modification at the LHC begins at a lower x than it does at RHIC. However, the data as a function of x are actually consistent at the two energies within uncertainties. Figure 9.11 shows the inclusive $R_{\rm dAu}$ or $R_{\rm pPb}$ in both systems. Figure 9.12 shows the most central and most peripheral $R_{\rm dAu}$ and $R_{\rm pPb}$ plotted as a function of x. The PHENIX

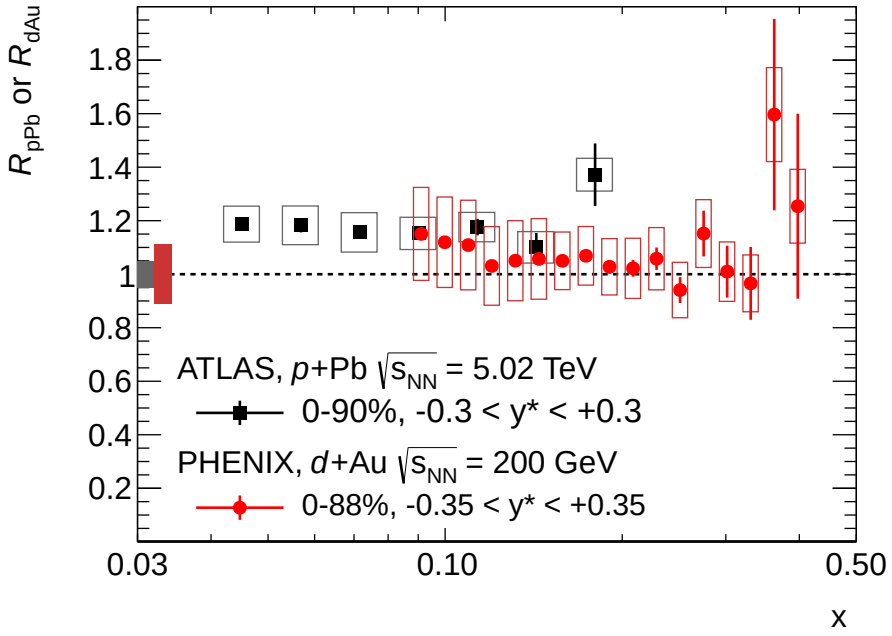


Figure 9.11: Comparison of the inclusive R_{dAu} and R_{pPb} for mid-rapidity jets in d+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV and p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, as a function of x.

and ATLAS data have qualitatively similar shapes. Again, although the central values of the data suggest that the modifications at the LHC begin at a lower x relative to those at RHIC, the data are actually consistent within uncertainties.

One early interpretation of the ATLAS and PHENIX data could be that there is a not yet understood correlation between the production of a jet and the underlying event far away from the jet in hadronic collisions. In fact, it is known that the mean event multiplicity and transverse energy in pp collisions increase with the highest-p_{T} object in the event. Assuming the same effect exists in p+A collisions, it is almost certain that events with jets would be categorized as having a higher centrality than they normally do, resulting in an enhanced (decreased) jet rate above the geometric expectation when selecting on central (peripheral) collisions. However, this effect has the opposite sign as what is observed in data.

Conceivably, requiring even higher p_{T} jets in the event could at some point suppress the underlying event activity, resulting in the opposite effect. For example, perhaps selecting high-x partons in the proton biases the distribution of the low-x partons in such a way that the soft multiplicity

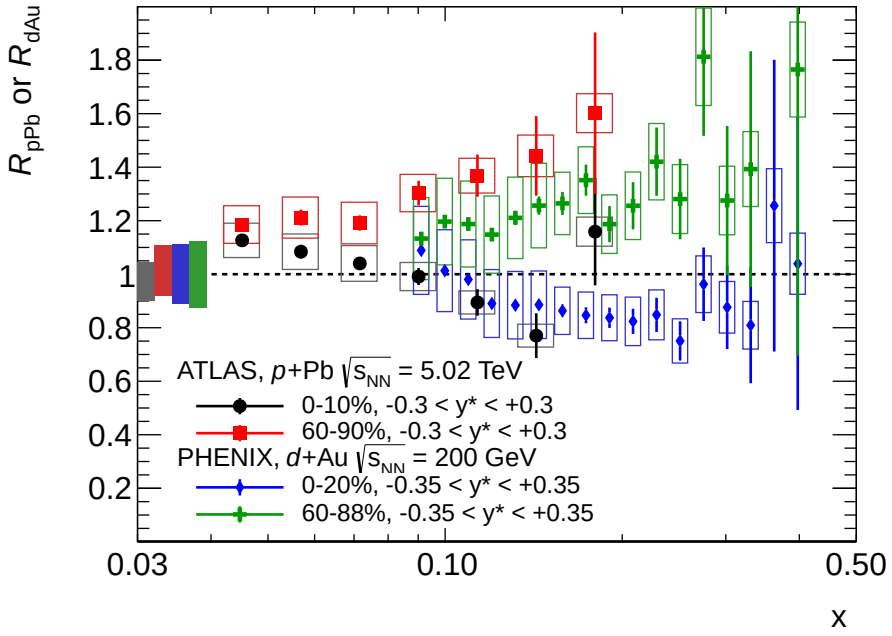


Figure 9.12: Comparison of the most central and most peripheral R_{dAu} and R_{pPb} for mid-rapidity jets in d+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV and p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, as a function of x.

arising from multiple parton interactions (MPIs) besides the high-x parton participating in the hard scattering is lowered. That is, the configurations of the proton wavefunction with a high-x parton tend to be more localized in space or otherwise have a smaller effective cross-section. If this is the case, it is happening at a fairly small value of proton-x at the LHC (since $p_{\mathrm{T}} = 100$ GeV jets at mid-rapidity are $x = 0.04$).

The simple scaling in the total jet energy suggests that the observed effects may be related to initial state effects arising from interactions of the partons in the nucleus before the hard parton-parton scattering. If this is the case, the data may be evidence of a heretofore unknown initial state mechanism. Ultimately any explanation of the underlying physics processes would have to address the smooth centrality dependence, the scaling behavior across the entire rapidity range and the p_{T} dependence of the observed modification. In particular, the underlying mechanism must address the enhanced rate of jets in peripheral collisions, where nuclear effects have naïvely been expected to be negligible.

An exciting possible implications of the ATLAS data is that nuclear effects over a wide range of rapidity could have a common origin. Within the ATLAS results, scaling behavior suggests that the relatively modest cold nuclear matter effects at $p_{\mathrm{T}} > 100$ GeV at mid-rapidity are actually connected to the strong effects for $p_{\mathrm{T}} > 20$ GeV jets four units of rapidity away. If the same is true at RHIC energies, this implies that the modification observed in high-p_{T} jets at mid-rapidity measured in PHENIX may be caused by the same mechanism responsible for (some of) the strong suppression of forward hadrons observed by all the RHIC experiments.

This work achieves several experimental milestones in modern heavy ion physics. It is the first measurement of the centrality-dependent yields of fully reconstructed jets in deuteron-nucleus collisions at RHIC and, separately, in proton-nucleus collisions at the LHC. It also represents the first experimental determination of the proton-nucleus event centrality at the LHC.

Despite this success, the measurements presented here are intended to be only the beginning of a detailed and systematic study of how hard processes are affected by the cold nuclear medium. It is important to investigate how, if at all, the jets in different centrality selections are modified. For example, the nuclear modification may be more pronounced for quark jets than gluon jets. Additionally, examining the energy balance of jet-photon or dijet pairs might better isolate what, if any, part of the suppression in central events could be consistent with initial state energy loss. Furthermore, other hard probes are available, which have different sensitivity to cold nuclear matter effects. Thus, it may be that precise and differential measurements of the rates of heavy flavor quarks, quarkonia and electroweak bosons, as well as reconstructed jets at lower p_{T}, are needed for a fuller understanding of the results presented here.

In the future, an Electron-Ion Collider (EIC) may help shed light on the same nuclear effects explored here through e+A collisions, which offer the advantage of a clean initial state in the projectile and the ability to precisely reconstruct the (x, Q^2) kinematics event by event by measuring the outgoing electon. Furthermore, using leptons to probe the nucleus could serve as a control against effects which are speculated to arise from fluctuations in the hadronic wavefunction of protons in p+A collisions.

References

[1] J. Beringer et al., (Particle Data Group), *2012 Review of Particle Physics*, Phys. Rev. D **86** (2012) 010001.

[2] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[3] F. Halzen and A. Martin, *An Introductory Course in Modern Particle Physics*. Wiley, 1984.

[4] T.-P. Cheng and L.-F. Li, *Gauge theory of elementary particle physics*. Oxford University Press, 1984.

[5] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346.

[6] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346–1349.

[7] D. Gross, D. Politzer, and F. Wilczek, *Nobel Lectures in Physics 2004*, Rev. Mod. Phys. **77** (2005).

[8] J. C. Street and E. C. Stevenson, *New Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron*, Phys. Rev. **52** (1937) 1003–1004.

[9] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (1962) 1067–1084.

[10] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nucl. Phys. **26** (1961) 222 – 229.

[11] V. E. Barnes et al., *Observation of a Hyperon with Strangeness Minus Three*, Phys. Rev. Lett. **12** (1964) 204–206.

[12] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett. **8** (1964) 214 – 215.

[13] G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and its Breaking*, CERN Report No.8182/TH.401 (1964).

[14] B. Bjorken and S. Glashow, *Elementary particles and SU(4)*, Phys. Lett. **11** (1964) 255 – 257.

[15] J. J. Aubert, U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart, P. Goldhagen, J. Leong, T. McCorriston, T. G. Rhoades, M. Rohde, S. C. C. Ting, S. L. Wu, and Y. Y. Lee, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404–1406.

[16] J. E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406–1408.

[17] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652–657.

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138–140.

[19] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh, J. A. Appel, B. C. Brown, C. N. Brown, W. R. Innes, K. Ueno, T. Yamanouchi, A. S. Ito, H. Jöstlein, D. M. Kaplan, and R. D. Kephart, *Observation of a Dimuon Resonance at 9.5* GeV *in 400-*GeV *Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252–255.

[20] F. Abe et al., (CDF Collaboration), *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626–2631.

[21] S. Abachi et al., (D0 Collaboration), *Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} =$ 1.8* TeV, Phys. Rev. Lett. **74** (1995) 2422–2426.

[22] D. P. Barber et al., *Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA*, Phys. Rev. Lett. **43** (1979) 830–833.

[23] R. Brandelik et al., *Evidence for planar events in e^+e^- annihilation at high energies*, Phys. Lett. B **86** (1979) 243 – 249.

[24] C. Berger et al., *Evidence for gluon bremsstrahlung in e^+e^- annihilations at high energies*, Phys. Lett. B **86** (1979) 418 – 425.

[25] R. Brandelik et al., *Evidence for a spin-1 gluon in three-jet events*, Phys. Lett. B **97** (1980) 453 – 458.

[26] H. Harari, *A new quark model for hadrons*, Phys. Lett. B **57** (1975) 265 – 269.

[27] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1988) 1–91, `arXiv:hep-ph/0409313 [hep-ph]`.

[28] J. F. Owens, E. Reya, and M. Glück, *Detailed quantum-chromodynamic predictions for high-p_T processes*, Phys. Rev. D **18** (1978) 1501–1514.

[29] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, *High-Energy Inelastic $e-p$ Scattering at 6° and 10°*, Phys. Rev. Lett. **23** (1969) 930–934.

[30] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor, *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969) 935–939.

[31] J. D. Bjorken and E. A. Paschos, *Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon*, Phys. Rev. **185** (1969) 1975–1982.

[32] L. Whitlow, E. Riordan, S. Dasu, S. Rock, and A. Bodek, *Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross sections*, Phys. Lett. B **282** (1992) 475 – 482.

[33] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298–318.

[34] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641–653.

[35] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450.

[36] J. F. Owens, A. Accardi, and W. Melnitchouk, *Global parton distributions with nuclear and finite-Q^2 corrections*, Phys. Rev. D **87** (2013) 094012.

[37] `http://wwwold.jlab.org/theory/cj/pdfs.html`.

[38] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, et al., *The CT10 NNLO Global Analysis of QCD*, `arXiv:1302.6246 [hep-ph]`.

[39] C. Adloff et al., *Deep-Inelastic Inclusive ep Scattering at Low x and a Determination of α_s*, Eur. Phys. J. C **21** (2001) 33–61.

[40] J. Breitweg et al., *ZEUS results on the measurement and phenomenology of F_2 at low x and low Q^2*, Eur. Phys. J **7** (1999) 609–630.

[41] I. I. Balitsky and L. N. Lipatov Sov. J. Nucl. Phys. **28** (1978) 822.

[42] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin Sov. Phys. JETP **45** (1977) 199.

[43] M. Strikman and D. Treleani, *Measuring Double-Parton Distributions in Nucleons at Proton-Nucleus Colliders*, Phys. Rev. Lett. **88** (2002) 031801.

[44] X. Ji, *Generalized Parton Distributions*, Ann. Rev. Nucl. Part. Sci. **54** (2004) 413–450.

[45] J. Huth et al., *Toward a standardization of jet definitions*, in *Research Directions for the Decade, Proceedings of the Summer Study on High Energy Physics, Snowmass, Colorado.* 1990.

[46] Y.-S. Lai and B. A. Cole, *Jet reconstruction in hadronic collisions by Gaussian filtering*, `arXiv:0806.1499 [nucl-ex]`.

[47] Y.-S. Lai, *Direct Jet Reconstruction in Proton-Proton and Copper-Copper Collisions at* $\sqrt{s_{\rm NN}} = 200$ GeV. Ph.D thesis, Columbia University, 2012.

[48] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **2008** (2008) 063.

[49] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896.

[50] `http://fastjet.fr/`.

[51] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. **B641** (2006) 57–61.

[52] R. Ellis and J. Sexton, *QCD radiative corrections to parton-parton scattering*, Nucl. Phys. B **269** (1986) 445 – 484.

[53] S. D. Ellis, Z. Kunszt, and D. E. Soper, *One-jet inclusive cross section at order α_s^3: Quarks and gluons*, Phys. Rev. Lett. **64** (1990) 2121–2124.

[54] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys. Rev. **D68** (2003) 094002.

[55] `http://www.desy.de/~znagy/Site/NLOJet++.html`.

[56] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67** (2010) 637–686.

[57] M. Dasgupta, L. Magnea, and G. P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, JHEP **0802** (2008) 055.

[58] ATLAS Collaboration, *Measurement of inclusive jet and dijet production in pp collisions at* $\sqrt{s} = 7$ TeV *using the ATLAS detector*, Phys. Rev. **D86** (2012) 014022.

[59] ATLAS Collaboration, *Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at* $\sqrt{s_{\rm NN}} = 2.76$ TeV *with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **105** (2010) 252303.

[60] ATLAS Collaboration, *Measurement of the jet radius and transverse momentum dependence of inclusive jet suppression in lead-lead collisions at with the ATLAS detector*, Phys. Lett. B **719** (2013) 220 – 241.

[61] W. Busza, J. E. Elias, D. F. Jacobs, P. A. Swartz, C. C. Young, and M. R. Sogard, *Charged-Particle Multiplicity in π^--Nucleus Interactions at 100 and 175* GeV/c, Phys. Rev. Lett. **34** (1975) 836–839.

[62] D. H. Brick et al., *Multiparticle production by 200-*GeV/c *hadrons on gold, silver, and magnesium targets*, Phys. Rev. D **39** (1989) 2484–2493.

[63] E. Stenlund and I. Otterlund, *On slow particle production in hadron-nucleus interactions*, Nucl. Phys. B **198** (1982) 407 – 426.

[64] M. Hegab and J. Hüfner, *How often does a high-energy hadron collide inside a nucleus?*, Phys. Lett. B **105** (1981) 103 – 106.

[65] I. Chemakin et al., (E910 Collaboration), *Measuring centrality with slow protons in proton-nucleus collisions at 18* GeV/c, Phys. Rev. C **60** (1999) 024902.

[66] R. Glauber, *Lectures in Theoretical Physics (Interscience, New York, 1958)*, Vol. I (1971) 315.

[67] R. Glauber and G. Matthiae, *High-energy scattering of protons by nuclei*, Nucl. Phys. B **21** (1970) 135 – 157.

[68] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, *Glauber Modeling in High-Energy Nuclear Collisions*, Ann. Rev. Nucl. Part. Sci. **57** (2007) 205–243.

[69] H. D. Vries, C. D. Jager, and C. D. Vries, *Nuclear charge-density-distribution parameters from elastic electron scattering*, Atomic Data and Nuclear Data Tables **36** (1987) 495 – 536.

[70] L. Hulthen and M. Sugawara, *Handbuch der Physik, vol. 39*, Berlin, Göttingen, Heidelberg. Springer (1957) 1.

[71] S. S. Adler et al., (PHENIX Collaboration), *Absence of Suppression in Particle Production at Large Transverse Momentum in* $\sqrt{s_{\rm NN}} = 200$ GeV $d+$Au *Collisions*, Phys. Rev. Lett. **91** (2003) 072303.

[72] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross-Section at* $\sqrt{s} = 7$ TeV *with the ATLAS Detector*, Nature Commun. **2** (2011) 463.

[73] G. Antchev et al., (TOTEM Collaboration), *Luminosity-independent measurements of total, elastic and inelastic cross-sections at* $\sqrt{s} = 7$ TeV, Eur. Phys. Lett. **101** (2013) 21004.

[74] A. Donnachie and P. Landshoff, *Total cross sections*, Phys. Lett. B **296** (1992) 227 – 232.

[75] B. Alver, M. Baker, C. Loizides, and P. Steinberg, *The PHOBOS Glauber Monte Carlo*, `arXiv:0805.4411 [nucl-ex]`.

[76] `http://www.hepforge.org/downloads/tglaubermc`.

[77] A. Białłas, M. Bleszyński, and W. Czyż, *Multiplicity distributions in nucleus-nucleus collisions at high energies*, Nucl. Phys. B **111** (1976) 461 – 476.

[78] X.-N. Wang and M. Gyulassy, *Energy and Centrality Dependence of Rapidity Densities at RHIC Energies*, Phys. Rev. Lett. **86** (2001) 3496–3499.

[79] D. Kharzeev and M. Nardi, *Hadron production in nuclear collisions at RHIC and high-density QCD*, Phys. Lett. B **507** (2001) 121 – 128.

[80] S. Eremin and S. Voloshin, *Nucleon participants or quark participants?*, Phys. Rev. C **67** (2003) 064905.

[81] A. Bazilevsky, J. Mitchell, K. Reygers, R. Seto, and M. Tannenbaum, *Participant Quark Scaling of Transverse Energy Production in 200* GeV *Au+Au, d+Au, p+p and 62.4* GeV *Au+Au Collisions*, PHENIX Analysis Note #1091.

[82] V. Guzey and M. Strikman, *Proton-nucleus scattering and cross section fluctuations at RHIC and LHC*, Phys. Lett. B **633** (2006) 245 – 252.

[83] V. Guzey and M. Strikman, *Erratum to: Proton-nucleus scattering and cross section fluctuations at RHIC and LHC*, Phys. Lett. B **663** (2008) 456.

[84] M. Alvioli and M. Strikman, *Color fluctuation effects in proton-nucleus collisions*, Phys. Lett. B **722** (2013) 347 – 354.

[85] P. Mulders, A. Schreiber, and H. Meyer, *The convolution approach in deep-inelastic scattering*, Nucl. Phys. A **549** (1992) 498 – 520.

[86] Z.-B. Kang, I. Vitev, and H. Xing, *Nuclear modification of high transverse momentum particle production in p+A collisions at RHIC and LHC*, Phys. Lett. B **718** (2012) 482 – 487.

[87] Y. L. Dokshitzer, V. Khoze, A. Mueller, and S. Troyan, *Basics of perturbative QCD*. Edition Frontières, 1991.

[88] N. Christ, B. Hasslacher, and A. H. Mueller, *Light-Cone Behavior of Perturbation Theory*, Phys. Rev. D **6** (1972) 3543–3562.

[89] M. Luo, J. Qiu, and G. Sterman, *Nuclear dependence at large transverse momentum*, Phys. Lett. B **279** (1992) 377 – 383.

[90] M. Luo, J. Qiu, and G. Sterman, *Twist-4 nuclear parton distributions from photoproduction*, Phys. Rev. D **49** (1994) 4493–4502.

[91] M. Luo, J. Qiu, and G. Sterman, *Anomalous nuclear enhancement in deeply inelastic scattering and photoproduction*, Phys. Rev. D **50** (1994) 1951–1971.

[92] J. W. Cronin, H. J. Frisch, M. J. Shochet, J. P. Boymond, P. A. Piroué, and R. L. Sumner, *Production of hadrons at large transverse momentum at 200, 300, and 400* GeV, Phys. Rev. D **11** (1975) 3105–3123.

[93] D. Antreasyan, J. W. Cronin, H. J. Frisch, M. J. Shochet, L. Kluberg, P. A. Piroué, and R. L. Sumner, *Production of hadrons at large transverse momentum in 200-, 300-, and 400*-GeV $p-p$ *and p-nucleus collisions*, Phys. Rev. D **19** (1979) 764–778.

[94] A. Accardi, *Cronin effect in proton nucleus collisions: A Survey of theoretical models*, `arXiv:hep-ph/0212148 [hep-ph]`.

[95] Y. Zhang, G. Fai, G. Papp, G. G. Barnaföldi, and P. Lévai, *High-p_T pion and kaon production in relativistic nuclear collisions*, Phys. Rev. C **65** (2002) 034903.

[96] A. Majumder and B. Müller, *Higher twist jet broadening and classical propagation*, Phys. Rev. C **77** (2008) 054903.

[97] S. S. Adler et al., (PHENIX Collaboration), *Centrality Dependence of π^0 and η Production at Large Transverse Momentum in $\sqrt{s_{\rm NN}} = 200$ GeV $d+$Au Collisions*, Phys. Rev. Lett. **98** (2007) 172302.

[98] B. Abelev et al., (ALICE Collaboration), *Transverse Momentum Distribution and Nuclear Modification Factor of Charged Particles in p+*Pb *Collisions at $\sqrt{s_{\rm NN}}$=5.02 TeV*, Phys. Rev. Lett. **110** (2013) 082302.

[99] S. S. Adler et al., (PHENIX Collaboration), *Nuclear effects on hadron production in $d+$Au collisions at $\sqrt{s_{NN}} = 200$ GeV revealed by comparison with $p+p$ data*, Phys. Rev. C **74** (2006) 024904.

[100] D. S. Barton, G. W. Brandenburg, W. Busza, T. Dobrowolski, J. I. Friedman, C. Halliwell, H. W. Kendall, T. Lyons, B. Nelson, L. Rosenson, R. Verdier, M. T. Chiaradia, C. DeMarzo, C. Favuzzi, G. Germinario, L. Guerriero, P. LaVopa, G. Maggi, F. Posa, G. Selvaggi, P. Spinelli, F. Waldner, D. Cutts, R. S. Dulude, B. W. Hughlock, R. E. Lanou, J. T. Massimo, A. E. Brenner, D. C. Carey, J. E. Elias, P. H. Garbincius, V. A. Polychronakos, J. Nassalski, and T. Siemiarczuk, *Experimental study of the A dependence of inclusive hadron fragmentation*, Phys. Rev. D **27** (1983) 2580–2599.

[101] S. J. Brodsky and P. Hoyer, *A bound on the energy loss of partons in nuclei*, Phys. Lett. B **298** (1993) 165 – 170.

[102] Y. Berdnikov, V. Kim, V. Kosmach, M. Ryzhinskiy, V. Samsonov, et al., *Initial-state nuclear effects in proton-nucleus collisions*, Eur. Phys. J. **A26** (2005) 179–184.

[103] I. Vitev, *Non-Abelian energy loss in cold nuclear matter*, Phys. Rev. C **75** (2007) 064906.

[104] X. Guo and X.-N. Wang, *Multiple Scattering, Parton Energy Loss, and Modified Fragmentation Functions in Deeply Inelastic eA Scattering*, Phys. Rev. Lett. **85** (2000) 3591–3594.

[105] G. Ovanesyan and I. Vitev, *An effective theory for jet propagation in dense QCD matter: jet broadening and medium-induced bremsstrahlung*, JHEP **1106** (2011) 080.

[106] B. Kopeliovich, J. Nemchik, E. Predazzi, and A. Hayashigaki, *Nuclear hadronization: within or without?*, Nucl. Phys. A **740** (2004) 211 – 245.

[107] A. Accardi, D. Grnewald, V. Muccifora, and H. Pirner, *Atomic mass dependence of hadron production in deep inelastic scattering on nuclei*, Nucl. Phys. A **761** (2005) 67 – 91.

[108] G. Barnafoldi, J. Barrette, M. Gyulassy, P. Levai, G. Papp, et al., *Predictions for p+Pb at 5.02A* TeV *to test initial state nuclear shadowing at the Large Hadron Collider*, `arXiv:1211.2256 [nucl-th]`.

[109] S. J. Brodsky and H. J. Lu, *Shadowing and antishadowing of nuclear structure functions*, Phys. Rev. Lett. **64** (1990) 1342–1345.

[110] L. Frankfurt, V. Guzey, and M. Strikman, *Leading twist nuclear shadowing phenomena in hard processes with nuclei*, Phys. Rep. **512** (2012) 255 – 393.

[111] J.-W. Qiu and I. Vitev, *Coherent QCD multiple scattering in protonnucleus collisions*, Phys. Lett. B **632** (2006) 507 – 511.

[112] S. yuan Li and X.-N. Wang, *Gluon shadowing and hadron production at RHIC*, Phys. Lett. B **527** (2002) 85 – 91.

[113] A. Accardi, N. Armesto, M. Botje, S. Brodsky, B. Cole, et al., *Hard probes in heavy ion collisions at the LHC: PDFs, shadowing and pA collisions*, `arXiv:hep-ph/0308248 [hep-ph]`.

[114] C. Salgado et al., *Proton-Nucleus Collisions at the LHC: Scientific Opportunities and Requirements*, J. Phys. **G39** (2012) 015010.

[115] K. Eskola, H. Paukkunen, and C. Salgado, *EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions*, JHEP **0904** (2009) 065.

[116] M. Hirai, S. Kumano, and T.-H. Nagai, *Determination of nuclear parton distribution functions and their uncertainties at next-to-leading order*, Phys. Rev. C **76** (2007) 065207.

[117] M. Adams, S. Aid, P. Anthony, D. Averill, M. Baker, B. Baller, A. Banerjee, A. Bhatti, U. Bratzler, H. Braun, et al., *Shadowing in inelastic scattering of muons on carbon, calcium and lead at low* x_{Bj}, Z. Phys. C **67** (1995) 403–410.

[118] M. Arneodo et al., *The* Q^2 *dependence of the structure function ratio* $F_2^{\mathrm{Sn}}/F_2^{\mathrm{C}}$ *and the difference* $R^{\mathrm{Sn}} - R^{\mathrm{C}}$ *deep inelastic muon scattering*, Nucl. Phys. B **481** (1996) 23 – 39.

[119] I. Arsene et al., (BRAHMS Collaboration), *Evolution of the Nuclear Modification Factors with Rapidity and Centrality in* $d+\mathrm{Au}$ *Collisions at* $\sqrt{s_{\mathrm{NN}}} = 200$ GeV, Phys. Rev. Lett. **93** (2004) 242303.

[120] S. S. Adler et al., (PHENIX Collaboration), *Nuclear Modification Factors for Hadrons at Forward and Backward Rapidities in Deuteron-Gold Collisions at* $\sqrt{s_{\mathrm{NN}}} = 200$ GeV, Phys. Rev. Lett. **94** (2005) 082302.

[121] J. Qiu, *Nuclear shadowing at small values of* x, Nucl. Phys. B **291** (1987) 746 – 764.

[122] S. J. Brodsky and H. J. Lu, *Shadowing and antishadowing of nuclear structure functions*, Phys. Rev. Lett. **64** (1990) 1342–1345.

[123] L. Frankfurt and M. Strikman, *Hard nuclear processes and microscopic nuclear structure*, Phys. Rep. **160** (1988) 235 – 427.

[124] J. Aubert et al., *The ratio of the nucleon structure functions* F_2^{N} *for iron and deuterium*, Phys. Lett. B **123** (1983) 275 – 278.

[125] J. Seely et al., *New Measurements of the European Muon Collaboration Effect in Very Light Nuclei*, Phys. Rev. Lett. **103** (2009) 202301.

[126] L. B. Weinstein, E. Piasetzky, D. W. Higinbotham, J. Gomez, O. Hen, and R. Shneor, *Short Range Correlations and the EMC Effect*, Phys. Rev. Lett. **106** (2011) 052301.

[127] A. Bodek and J. L. Ritchie, *Fermi-motion effects in deep-inelastic lepton scattering from nuclear targets*, Phys. Rev. D **23** (1981) 1070–1091.

[128] I. Helenius, K. J. Eskola, H. Honkanen, and C. A. Salgado, *Impact-Parameter Dependent Nuclear Parton Distribution Functions: EPS09s and EKS98s and Their Applications in Nuclear Hard Processes*, JHEP **1207** (2012) 073.

[129] P. Quiroga-Arias, J. G. Milhano, and U. A. Wiedemann, *Testing nuclear parton distributions with pA collisions at the* TeV *scale*, Phys. Rev. C **82** (2010) 034903.

[130] CMS Collaboration, *Study of dijet momentum balance and pseudorapidity distributions in pPb collisions at* $\sqrt{s_{\rm NN}} = 5.02$ TeV, CMS Physics Analysis Summary HIN-13-001.

[131] C. Adloff et al., (H1), *On the rise of the proton structure function* F_2 *towards low* x, Phys. Lett. B **520** (2001) 183 – 190.

[132] M. Froissart, *Asymptotic Behavior and Subtractions in the Mandelstam Representation*, Phys. Rev. **123** (1961) 1053–1057.

[133] A. Mueller and J. Qiu, *Gluon recombination and shadowing at small values of* x, Nucl. Phys. B **268** (1986) 427 – 452.

[134] I. Balitsky, *Operator expansion for high-energy scattering*, Nucl. Phys. B **463** (1996) 99 – 157.

[135] Y. V. Kovchegov, *Small* x F_2 *structure function of a nucleus including multiple pomeron exchanges*, Phys. Rev. **D60** (1999) 034008.

[136] Y. V. Kovchegov, *Unitarization of the BFKL Pomeron on a nucleus*, Phys. Rev. D **61** (2000) 074018.

[137] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *The BFKL equation from the Wilson renormalization group*, Nucl. Phys. B **504** (1997) 415 – 431.

[138] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *Wilson renormalization group for low x physics: Towards the high density regime*, Phys. Rev. D **59** (1998) 014014.

[139] E. Iancu, A. Leonidov, and L. McLerran, *Nonlinear gluon evolution in the color glass condensate: I*, Nucl. Phys. A **692** (2001) 583 – 645.

[140] E. Iancu, A. Leonidov, and L. McLerran, *The renormalization group equation for the color glass condensate*, Phys. Lett. B **510** (2001) 133 – 144.

[141] J. L. Albacete and Y. V. Kovchegov, *Solving the high energy evolution equation including running coupling corrections*, Phys. Rev. D **75** (2007) 125021.

[142] E. Gardi, J. Kuokkanen, K. Rummukainen, and H. Weigert, *Running coupling and power corrections in non-linear evolution at the high-energy limit*, Nucl. Phys. A **784** (2007) 282 – 340.

[143] A. Accardi, J. Albacete, M. Anselmino, N. Armesto, E. Aschenauer, et al., *Electron Ion Collider: The Next QCD Frontier - Understanding the glue that binds us all*, `arXiv:1212.1701 [nucl-ex]`.

[144] L. Gribov, E. Levin, and M. Ryskin, *Semihard processes in QCD*, Phys. Rep. **100** (1983) 1 – 150.

[145] L. McLerran and R. Venugopalan, *Computing quark and gluon distribution functions for very large nuclei*, Phys. Rev. D **49** (1994) 2233–2241.

[146] L. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, Phys. Rev. D **49** (1994) 3352–3355.

[147] L. McLerran and R. Venugopalan, *Green's function in the color field of a large nucleus*, Phys. Rev. D **50** (1994) 2225–2233.

[148] K. Golec-Biernat and M. Wüsthoff, *Saturation effects in deep inelastic scattering at low Q^2 and its implications on diffraction*, Phys. Rev. D **59** (1998) 014017.

[149] D. Kharzeev, E. Levin, and L. McLerran, *Parton saturation and Npart scaling of semi-hard processes in QCD*, Phys. Lett. B **561** (2003) 93 – 101.

[150] D. Kharzeev, Y. V. Kovchegov, and K. Tuchin, *Cronin effect and high-p_T suppression in pA collisions*, Phys. Rev. D **68** (2003) 094013.

[151] D. Kharzeev, Y. V. Kovchegov, and K. Tuchin, *Nuclear modification factor in collisions: onset of suppression in the color glass condensate*, Phys. Lett. B **599** (2004) 23 – 31.

[152] V. Guzey, M. Strikman, and W. Vogelsang, *Observations on dA scattering at forward rapidities*, Phys. Lett. B **603** (2004) 173 – 183.

[153] L. Frankfurt and M. Strikman, *Energy losses in the black disc regime and correlation effects in the STAR forward pion production in dAu collisions*, Phys. Lett. B **645** (2007) 412 – 421.

[154] J. Qiu and I. Vitev, *Resummed QCD Power Corrections to Nuclear Shadowing*, Phys. Rev. Lett. **93** (2004) 262301.

[155] B. Z. Kopeliovich, J. Nemchik, I. K. Potashnikova, M. B. Johnson, and I. Schmidt, *Breakdown of QCD factorization at large Feynman x*, Phys. Rev. C **72** (2005) 054606.

[156] A. Adare et al., (PHENIX Collaboration), *Suppression of Back-to-Back Hadron Pairs at Forward Rapidity in $d+$Au Collisions at $\sqrt{s_{\rm NN}} = 200$ GeV*, Phys. Rev. Lett. **107** (2011) 172301.

[157] E. Braidot, *Two-particle correlations at forward rapidity in STAR*, Nucl. Phys. A **854** (2011) 168 – 174.

[158] J. Adams et al., (STAR Collaboration), *Forward Neutral Pion Production in $p+p$ and $d+$Au Collisions at $\sqrt{s_{\rm NN}} = 200$ GeV*, Phys. Rev. Lett. **97** (2006) 152302.

[159] J. L. Albacete and C. Marquet, *Azimuthal Correlations of Forward Dihadrons in $d+$Au Collisions at RHIC in the Color Glass Condensate*, Phys. Rev. Lett. **105** (2010) 162301.

[160] M. Strikman and W. Vogelsang, *Multiple parton interactions and forward double pion production in pp and dA scattering*, Phys. Rev. D **83** (2011) 034029.

[161] B. Abelev et al., (ALICE Collaboration), *Pseudorapidity Density of Charged Particles in p+Pb Collisions at $\sqrt{s_{\rm NN}}$=5.02 TeV*, Phys. Rev. Lett. **110** (2013) 032301.

[162] ATLAS Collaboration, *Observation of Associated Near-Side and Away-Side Long-Range Correlations in $\sqrt{s_{\rm NN}}$=5.02 TeV Proton-Lead Collisions with the ATLAS Detector*, Phys. Rev. Lett. **110** (2013) 182302.

[163] S. Chatrchyan et al., (CMS Collaboration), *Observation of long-range, near-side angular correlations in pPb collisions at the LHC*, Phys. Lett. B **718** (2013) 795 – 814.

[164] B. Abelev et al., (ALICE Collaboration), *Long-range angular correlations on the near and away side in p-Pb collisions at*, Phys. Lett. B **719** (2013) 29 – 41.

[165] K. Dusling and R. Venugopalan, *Evidence for BFKL and saturation dynamics from dihadron spectra at the LHC*, Phys. Rev. D **87** (2013) 051502.

[166] S. Catani, M. Ciafaloni, and F. Hautmann, *High energy factorization and small-x heavy flavour production*, Nucl. Phys. B **366** (1991) 135 – 188.

[167] K. Kutak and S. Sapeta, *Gluon saturation in dijet production in p-Pb collisions at the Large Hadron Collider*, Phys. Rev. D **86** (2012) 094043.

[168] J. L. Albacete and C. Marquet, *Single inclusive hadron production at RHIC and the LHC from the color glass condensate*, Phys. Lett. B **687** (2010) 174 – 179.

[169] P. Quiroga-Arias, J. G. Milhano, and U. A. Wiedemann, *Testing factorization in pA collisions at the LHC*, Nucl. Phys. A **855** (2011) 498 – 501. Proceedings of the 4th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions - HP2010.

[170] E. O. Lawrence, *Method and apparatus for the acceleration of ions*, US Patent 1948384 (1934).

[171] E. D. Courant, M. S. Livingston, and H. S. Snyder, *The Strong-Focusing Synchroton: A New High Energy Accelerator*, Phys. Rev. **88** (1952) 1190–1196.

[172] E. Courant and H. Snyder, *Theory of the alternating-gradient synchrotron*, Ann. Phys. **3** (1958) 1 – 48.

[173] G. W. Hill, *On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon*, Acta Mathematica **8** (1886) 1–36.

[174] S. Van der Meer, *Calibration of the effective beam height in the ISR*, internal report CERN-ISR-PO/68-31 (1968).

[175] D. Perepelitsa and B. Cole, *Luminosity determination in p+p collisions at* $\sqrt{s} = 2.76$ TeV *with the ATLAS detector in 2011*, ATL-COM-LUM-2013-004.

[176] G. Aad, B. Abbott, J. Abdallah, A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, et al., *Luminosity determination in pp collisions at* $\sqrt{s} = 7$ TeV *using the ATLAS detector at the LHC*, The European Physical Journal C **71** (2011) 1–37.

[177] M. Harrison, T. Ludlam, and S. Ozaki, *RHIC project overview*, NIM A **499** (2003) 235–244.

[178] H. Hahn, E. Forsyth, H. Foelsche, M. Harrison, J. Kewisch, G. Parzen, S. Peggs, E. Raka, A. Ruggiero, A. Stevens, S. Tepikian, P. Thieberger, D. Trbojevic, J. Wei, E. Willen, S. Ozaki, and S. Lee, *The RHIC design overview*, NIM A **499** (2003) 245 – 263.

[179] M. Harrison, S. Peggs, and T. Roser, *The RHIC accelerator*, Annual Review of Nuclear and Particle Science **52** (2002) 425–469.

[180] I. Alekseev, C. Allgower, M. Bai, Y. Batygin, L. Bozano, K. Brown, G. Bunce, P. Cameron, E. Courant, S. Erin, et al., *Polarized proton collider at RHIC*, NIM A **499** (2003) 392–414.

[181] I. Arsene et al., *Quark-gluon plasma and color glass condensate at RHIC? The perspective from the BRAHMS experiment*, Nucl. Phys. A **757** (2005) 1 – 27.

[182] K. Adcox et al., *Formation of dense partonic matter in relativistic nucleusnucleus collisions at RHIC: Experimental evaluation by the PHENIX Collaboration*, Nucl. Phys. A **757** (2005) 184 – 283.

[183] B. Back et al., *The PHOBOS perspective on discoveries at RHIC*, Nucl. Phys. A **757** (2005) 28 – 101.

[184] J. Adams et al., *Experimental and theoretical challenges in the search for the quark-gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions*, Nucl. Phys. A **757** (2005) 102 – 183.

[185] C. Gardner, N. Abreu, L. Ahrens, J. Alessi, and M. Bai, *Setup and performance of RHIC for the 2008 run with deuteron-gold collisions*, in *Proceedings of the European Particle Accelerator Conference, 2008.* 2008.

[186] T. Satogata, L. Ahrens, M. Bai, J. Beebe-Wang, M. Blaskiewicz, J. Brennan, K. Brown, D. Bruno, P. Cameron, J. Cardona, et al., *Commissioning of RHIC deuteron-gold collisions*, in *Proceedings of the Particle Accelerator Conference, 2003.*, vol. 3, IEEE. 2003.

[187] C. Montag, M. Bai, W. MacKay, T. Roser, N. Abreu, L. Ahrens, D. Barton, J. Beebe-Wang, M. Blaskiewicz, J. Brennan, et al., *RHIC Polarized Proton Performance in Run-8*, in *AIP Conference Proceedings*. 2009.

[188] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1 – 29.

[189] S. Chatrchyan et al., *Observation of a new boson at a mass of 125* GeV *with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30 – 61.

[190] ATLAS Collaboration, *Coupling properties of the new Higgs-like boson observed with the ATLAS detector at the LHC*, ATLAS-CONF-2012-127. `http://cds.cern.ch/record/1476765`.

[191] ATLAS Collaboration, *An update of combined measurements of the new Higgs-like boson with high mass resolution channels*, ATLAS-CONF-2012-170. `http://cds.cern.ch/record/1499629`.

[192] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[193] O. S. Brning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report v.1 : the LHC Main Ring*. CERN, Geneva, 2004.

[194] O. S. Brning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report v.2 : the LHC Infrastructure and General Services*. CERN, Geneva, 2004.

[195] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report v.3 : the LHC Injector Chain*. CERN, Geneva, 2004.

[196] K. Adcox et al., *PHENIX detector overview*, NIM A **499** (2003) 469 – 479.

[197] R. Nouicer, (PHENIX Collaboration), *Silicon Vertex Tracker for PHENIX Upgrade at RHIC: Capabilities and Detector Technology*, PoS **VERTEX2007** (2007) 042, arXiv:0801.2947 [nucl-ex].

[198] D. E. Fields et al., (PHENIX Collaboration), *Forward Silicon Vertex Tracker for the PHENIX experiment at RHIC*, Journal of Physics: Conference Series **316** (2011) 012032.

[199] S. Campbell et al., (PHENIX Collaboration), *A Proposal for the Muon Piston Calorimeter Extension (MPC-EX) to the PHENIX Experiment at RHIC*, arXiv:1301.1096 [nucl-ex].

[200] C. Aidala, N. Ajitanand, Y. Akiba, Y. Akiba, R. Akimoto, et al., *sPHENIX: An Upgrade Concept from the PHENIX Collaboration*, arXiv:1207.6378 [nucl-ex].

[201] K. Adcox et al., (PHENIX Collaboration), *Construction and performance of the PHENIX pad chambers*, NIM A **497** (2003) 263 – 293.

[202] S. Aronson et al., (PHENIX Collaboration), *PHENIX magnet system*, NIM A **499** (2003) 480–488.

[203] K. Adcox et al., (PHENIX Collaboration), *PHENIX central arm tracking detectors*, NIM A **499** (2003) 489–507.

[204] R. Veenhof, *Garfield, recent developments*, NIM A **419** (1998) 726 – 730.

[205] http://garfield.web.cern.ch/garfield/.

[206] M. Aizawa et al., (PHENIX Collaboration), *PHENIX central arm particle ID detectors*, NIM A **499** (2003) 508–520.

[207] L. Aphecetche et al., (PHENIX Collaboration), *PHENIX calorimeter*, NIM A **499** (2003) 521 – 536.

[208] G. David, Y. Goto, E. Kistenev, S. Stoll, S. White, C. Woody, A. Bazilevsky, S. Belikov, S. Chernichenkov, A. Denisov, et al., *The PHENIX lead-scintillator electromagnetic calorimeter: Test beam and construction experience*, Nuclear Science, IEEE Transactions **45** (1998) 692–697.

[209] H. Fessler, P. Freund, J. Gebauer, K. Glas, K. Pretzl, P. Seyboth, J. Seyerlein, and J. Thevenin, *A scintillator-lead photon calorimeter using optical fiber readout systems*, NIM A **240** (1985) 284 – 288.

[210] T. Peitzmann et al., *A new monitoring system for the photon spectrometer LEDA in the WA98 experiment*, NIM A **376** (1996) 368 – 374.

[211] G. David, T. Sakaguchi, and R. Vertesi, *Simulation Verification in the EMCal PbSc Shower Characteristics of Hadrons*, PHENIX Analysis Note #330.

[212] S. S. Adler et al., (PHENIX Collaboration), *PHENIX on-line systems*, NIM A **499** (2003) 560–592.

[213] S. Adler et al., (PHENIX Collaboration), *PHENIX on-line and off-line computing*, NIM A **499** (2003) 593–602.

[214] J. Haggerty, *The JSEB PCI Interface*, PHENIX Technical Note #407.

[215] `http://www.corba.org`.

[216] M. Allen et al., (PHENIX Collaboration), *PHENIX inner detectors*, NIM A **499** (2003) 549–559.

[217] K. Ikematsu et al., *A start-timing detector for the collider experiment PHENIX at RHIC-BNL*, NIM A **411** (1998) 238 – 248.

[218] A. Bazilevsky, R. Bennett, A. Deshpande, Y. Goto, D. Kawall, and J. Seele, *Absolute Luminosity Determination Using the Vernier Scan Technique: Run4-6 Analysis and Preliminary Results at* $\sqrt{s}$*=62.4* GeV, PHENIX Analysis Note #597.

[219] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[220] A. Yamamoto, Y. Makida, R. Ruber, Y. Doi, T. Haruyama, F. Haug, H. ten Kate, M. Kawai, T. Kondo, Y. Kondo, J. Metselaar, S. Mizumaki, G. Olesen, O. Pavlov, S. Ravat, E. Sbrissa, K. Tanaka, T. Taylor, and H. Yamaoka, *The ATLAS central solenoid*, NIM A **584** (2008) 53 – 74.

[221] M. Aleksa, F. Bergsma, P. A. Giudici, A. Kehrli, M. Losasso, X. Pons, H. Sandaker, P. S. Miyagawa, S. W. Snow, J. C. Hart, and L. Chevalier, *Measurement of the ATLAS solenoid magnetic field*, JINST **3** (2008) P04003.

[222] ATLAS Collaboration, *Performance of the Minimum Bias Trigger in pp Collisions at* $\sqrt{s} = 900$ GeV, ATLAS-CONF-2010-025. `http://cds.cern.ch/record/1277657`.

[223] ATLAS Collaboration, *Performance of the Minimum Bias Trigger in pp Collisions at* $\sqrt{s} = 7$ TeV, ATLAS-CONF-2010-068. `http://cds.cern.ch/record/1281343`.

[224] P. Jenni and M. Nessi, *ATLAS Forward Detectors for Luminosity Measurement and Monitoring*, Tech. Rep. CERN-LHCC-2004-010. LHCC-I-014, CERN, Geneva, Mar, 2004.

[225] V. Cindro, D. Dobos, I. Dolenc, H. Frais-Klbl, A. Goriek, E. Griesmayer, H. Kagan, G. Kramberger, B. Macek, I. Mandi, M. Miku, M. Niegl, H. Pernegger, D. Tardif, W. Trischuk, P. Weilhammer, and M. Zavrtanik, *The ATLAS Beam Conditions Monitor*, JINST **3** (2008) P02004.

[226] *ATLAS liquid-argon calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.

[227] B. Aubert et al., *Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter*, NIM A **558** (2006) 388 – 418.

[228] B. Aubert et al., *Development and construction of large size signal electrodes for the ATLAS electromagnetic calorimeter*, NIM A **539** (2005) 558 – 594.

[229] J. Abdallah et al., *The Production and Qualification of Scintillator Tiles for the ATLAS Hadronic Calorimeter*, tech. rep., CERN, Geneva, Dec, 2007.

[230] A. Artamonov, D. Bailey, G. Belanger, M. Cadabeschi, T. Y. Chen, V. Epshteyn, P. Gorbounov, K. K. Joo, M. Khakzad, V. Khovanskiy, P. Krieger, P. Loch, J. Mayer, E. Neuheimer, F. G. Oakham, M. O'Neill, R. S. Orr, M. Qi, J. Rutherfoord, A. Savine, M. Schram, P. Shatalov, L. Shaver, M. Shupe, G. Stairs, V. Strickland, D. Tompkins, I. Tsukerman, and K. Vincent, *The ATLAS Forward Calorimeter*, JINST **3** (2008) P02010.

[231] J. P. Archambault, A. Artamonov, M. Cadabeschi, V. Epshteyn, C. Galt, P. Gorbounov, L. Heelan, M. Khakzad, V. Khovanskiy, P. Krieger, P. Loch, F. G. Oakham, R. S. Orr, J. Rutherfoord, A. Savine, M. Schram, P. Shatalov, L. Shaver, M. Shupe, V. Strickland, and I. Tsukerman, *Energy calibration of the ATLAS Liquid Argon Forward Calorimeter*, JINST **3** (2008) P02002.

[232] ATLAS Collaboration, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New Journal of Physics **13** (2011) 053033.

[233] N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[234] J. Jackson, *The ATLAS semiconductor tracker (SCT)*, NIM A **541** (2005) 89 – 95.

[235] A. Ahmad et al., *The silicon microstrip sensors of the ATLAS semiconductor tracker*, NIM A **578** (2007) 98 – 118.

[236] P. A. Coe, D. F. Howell, and R. B. Nickerson, *Frequency scanning interferometry in ATLAS: remote, multiple, simultaneous and precise distance measurements in a hostile environment*, Measurement Science and Technology **15** (2004) 2175.

[237] E. Vazquez, *Control study of two-particle correlations in heavy ion collisions at RHIC-PHENIX*. Ph.D thesis, Columbia University, 2013.

[238] J. Nagle, *PHENIX Run 8 d+Au @ 200* GeV *Centrality Recalibrator and QA*, PHENIX Analysis Note #844.

[239] `http://www.phenix.bnl.gov/WWW/trigger/pp/c-arm/Run3/run8.html`.

[240] `http://www.phenix.bnl.gov/WWW/trigger/pp/c-arm/Run3/run8_dAu/performance.html`.

[241] `http://logbook2.phenix.bnl.gov:7815/ERT/593`.

[242] J. Nagle, *PHENIX Run-08 d-Au @ 200* GeV *Centrality Categorization (in support of PHENIX Final Results)*, PHENIX Analysis Note #900.

[243] J. Nagle, S. Beckman, D. McGlinchey, A. Sickles, and J. Orjuela-Koop, *Run 8 d+Au 200* GeV *Centrality Categorize Addendum*, PHENIX Analysis Note #1087.

[244] J. Orjuela-Koop, D. McGlinchey, and J. Nagle, *Transverse Momentum Dependence of Centrality Bias Effects in HIJING and AMPT Simulations of d+Au Collisions*, PHENIX Analysis Note #1106.

[245] `http://www.phys.ufl.edu/~rfield/cdf/tunes/py_tuneA.html`.

[246] S. S. Adler et al., (PHENIX Collaboration), *Centrality dependence of charged hadron production in deuteron+gold and nucleon+gold collisions at* $\sqrt{s_{NN}} = 200$ GeV, Phys. Rev. C **77** (2008) 014905.

[247] A. Accardi, W. Melnitchouk, J. F. Owens, M. E. Christy, C. E. Keppel, L. Zhu, and J. G. Morfín, *Uncertainties in determining parton distributions at large* x, Phys. Rev. D **84** (2011) 014008.

[248] J. Arrington, J. G. Rubin, and W. Melnitchouk, *How Well Do We Know the Neutron Structure Function?*, Phys. Rev. Lett. **108** (2012) 252001.

[249] A. Angerami, B. Cole, A. Milov, and P. Steinberg, *Centrality Determination in the 2010 Pb+Pb physics data*, ATL-COM-PHYS-2011-427.

[250] ATLAS Collaboration, *Measurement of the centrality dependence of the charged particle pseudorapidity distribution in lead-lead collisions at with the ATLAS detector*, Phys. Lett. B **710** (2012) 363 – 382.

[251] ATLAS Collaboration, *Measurement of high-*p_{T} *isolated prompt photons in lead-lead collisions at* $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV *with the ATLAS detector at the LHC*, ATLAS-CONF-2012-051. `https://cds.cern.ch/record/1451913`.

[252] ATLAS Collaboration, *Measurement of Z Boson Production in Pb-Pb Collisions at* $\sqrt{s_{\mathrm{NN}}}$=2.76 TeV *with the ATLAS Detector*, Phys. Rev. Lett. **110** (2013) 022301.

[253] P. Steinberg, *Inclusive pseudorapidity distributions in p(d) + A collisions modeled with shifted rapidity distributions*, `arXiv:0703002 [nucl-ex]`.

[254] ATLAS Collaboration, *Rapidity gap cross sections measured with the ATLAS detector in pp collisions at* $\sqrt{s} = 7$ TeV, Eur. Phys. J. **C72** (2012) 1926.

[255] L. Frankfurt and M. Strikman, *Diffractive phenomena in high energy processes*, `arXiv:1304.4308 [hep-ph]`.

[256] G. Aad, P. Behera, W. H. Bell, U. Bitenc, G. Brandt, T. Eifert, H. M. Gray, J. Katzy, A. Kastanas, M. Kayl, O. Kepka, T. Kuhl, R. Kwee, M. Leyton, M. Limper, A. Lister, R. Meera-Lebbai, S. Nektarijevi, E. Nurse, K. Rosbach, A. Salzburger, M. Warsinsky, and R. Zaidan, *Track Reconstruction Efficiency in* $\sqrt{s} = 7$ TeV *Data for Tracks with* $p_{\mathrm{T}} > 100$ MeV, ATL-COM-PHYS-2010-682.

[257] ATLAS Collaboration, *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7* TeV *centre-of-mass energy with the ATLAS detector*, Eur. Phys. J. **C71** (2011) 1512.

[258] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep., CERN, Geneva, Apr, 2008.

[259] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **2006** (2006) 026.

[260] ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, ATL-PHYS-PUB-2011-009. `http://cds.cern.ch/record/1363300`.

[261] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, `arXiv:hep-ph/0201195 [hep-ph]`.

[262] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003. `http://cds.cern.ch/record/1474107`.

[263] D. Perepelitsa, B. Cole, and P. Steinberg, *Centrality Determination in the 2012 and 2013 p+Pb data in ATLAS*, ATL-COM-PHYS-2013-588.

[264] ATLAS Collaboration, *Measurement of the centrality dependence of the charged particle pseudorapidity distribution in proton-lead collisions at* $\sqrt{s_{\rm NN}} = 5.02$ TeV *with the ATLAS detector*, ATLAS-CONF-2013-096. `http://cds.cern.ch/record/1599773`.

[265] ATLAS Collaboration, *Centrality dependence of Jet Yields and Jet Fragmentation in Lead-Lead Collisions at* $\sqrt{s}_{\rm NN} = 2.76$ TeV *with the ATLAS detector at the LHC*, ATLAS-CONF-2011-075. `http://cds.cern.ch/record/1353220`.

[266] B. Cole, M. Spousta, A. Angerami, D. Perepelitsa, T. Kosek, P. Balek, M. Rybar, and H. Santos, *Supporting note for measurement of inclusive jet charged particle fragmentation functions in Pb+Pb collisions*, ATL-COM-PHYS-2012-1653.

[267] A. Angerami, *Jet Quenching in Relativistic Heavy Ion Collisions at the LHC*. Ph.D thesis, Columbia University, 2012.

[268] A. Angerami, B. Cole, D. Perepelitsa, P. Balek, T. Kosek, M. Rybar, M. Spousta, A. Galan, and H. Santos, *Performance of Jet Reconstruction in Heavy Ion Collisions*, ATL-PHYS-INT-2012-058.

[269] A. Angerami, B. A. Cole, J. Dolejsi, N. Grau, M. Leltchouck, M. Spousta, and W. G. Holzmann, *Detailed Study of Jet Seed Finding Performance for Background Subtraction in Heavy Ion Events*, ATL-COM-PHYS-2008-206.

[270] A. Angerami, B. Cole, R. Debbe, J. Dolejsi, N. Grau, W. Holzmann, M. Leltchouck, and M. Spousta, *The HIJetRec Package for Jet Reconstruction in a Heavy Ion Environment in ATLAS*, ATL-PHYS-INT-2008-039, ATL-COM-PHYS-2008-201.

[271] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ TeV *with ATLAS 2011 data*, ATLAS-CONF-2013-004. `http://cds.cern.ch/record/1509552`.

[272] T. Carli, A. Cooper-Sarker, C. Meyer, V. Kus, B. Malaescu, F. Mueller, S. Shimizu, P. Starovoitov, S. Yanush, and G. Vardanyan, *Measurement of the inclusive jet cross section in proton-proton collision data at 2.76* TeV *centre-of-mass energy using the ATLAS detector*, ATL-COM-PHYS-2012-498.

[273] A. Angerami, P. Balek, B. Cole, M. Rybar, and M. Spousta, *The Performance of Jet Trigger in Heavy Ion Collisions*, ATL-COM-DAQ-2012-209.

[274] ATLAS Collaboration, *Transverse energy fluctuations in Pb+Pb collisions at* $\sqrt{s_{\rm NN}} = 2.76$ TeV *with the ATLAS detector at the LHC*, ATLAS-CONF-2012-045. `http://cds.cern.ch/record/1440894`.

[275] ATLAS Collaboration, *Measurement of the inclusive jet cross section in pp collisions at* $\sqrt{s} = 2.76$ TeV *and comparison to the inclusive jet cross section at* $\sqrt{s} = 7$ TeV *using the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2509.

[276] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ TeV, Eur. Phys. J. **C73** (2013) 2304.

[277] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ TeV *recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2306.

[278] ATLAS Collaboration, *Measurement of the centrality dependence of inclusive jet production in p+Pb data at* $\sqrt{s_{\rm NN}} = 5.02$ TeV *with the ATLAS detector*, ATLAS-CONF-2013-105. `http://cds.cern.ch/record/1624014`.

Appendix A

p+Pb centrality fits and systematics

A.1 Centrality fits to 2013 p+Pb data

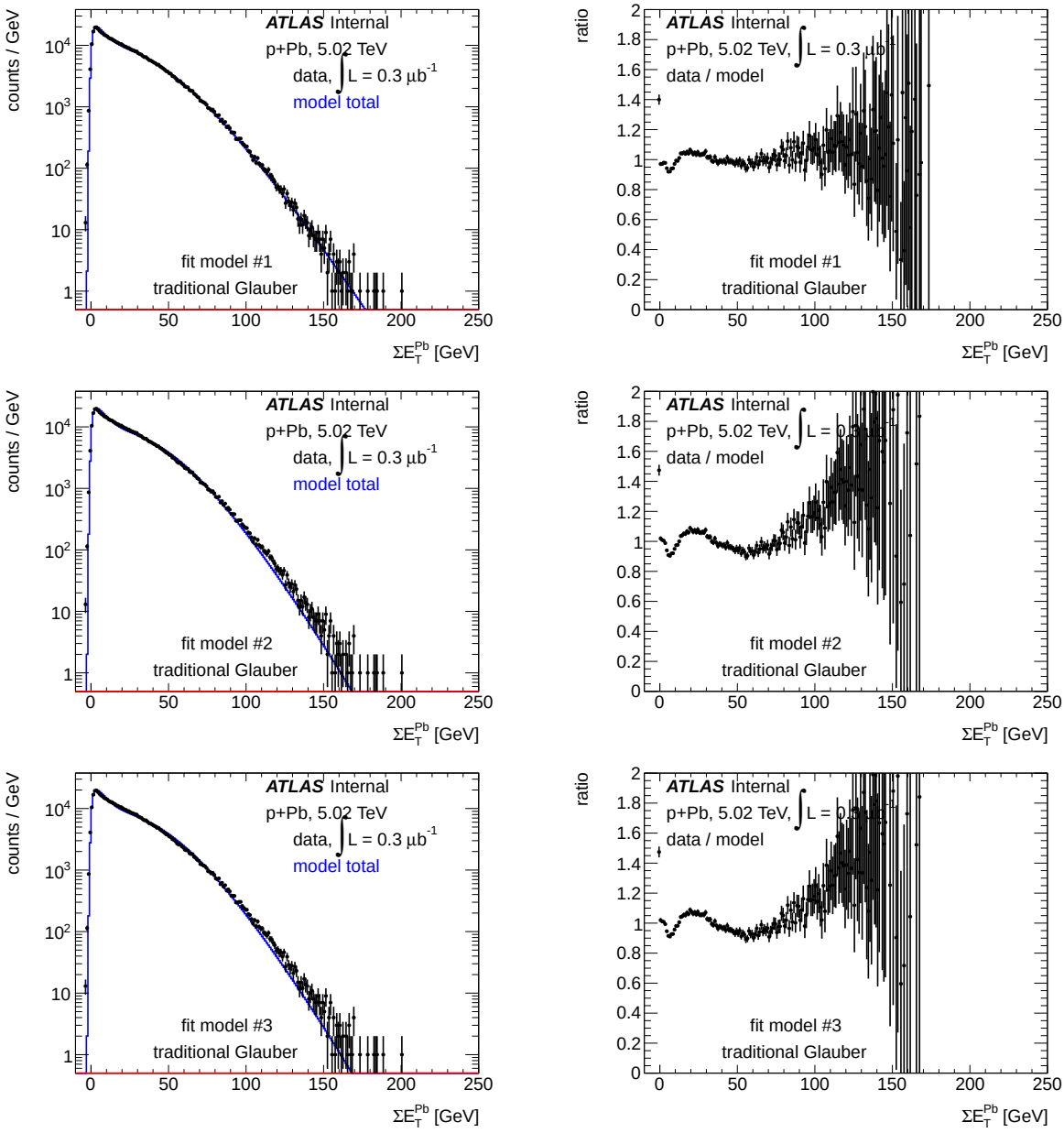


Figure A.1: Fits to 2013 p+Pb data using the default Glauber model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

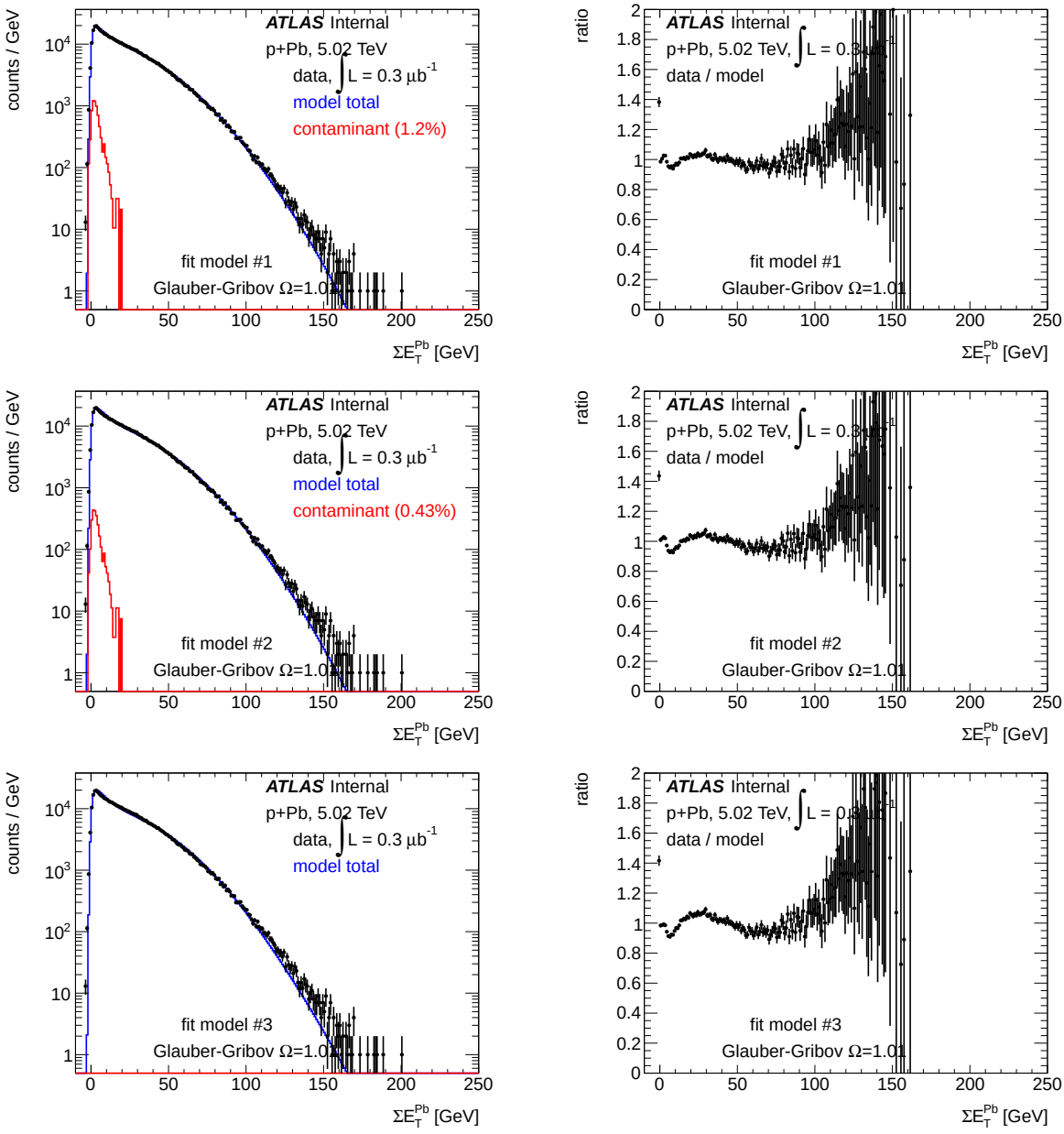


Figure A.2: Fits to 2013 p+Pb data using the Glauber-Gribov $\Omega = 1.01$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

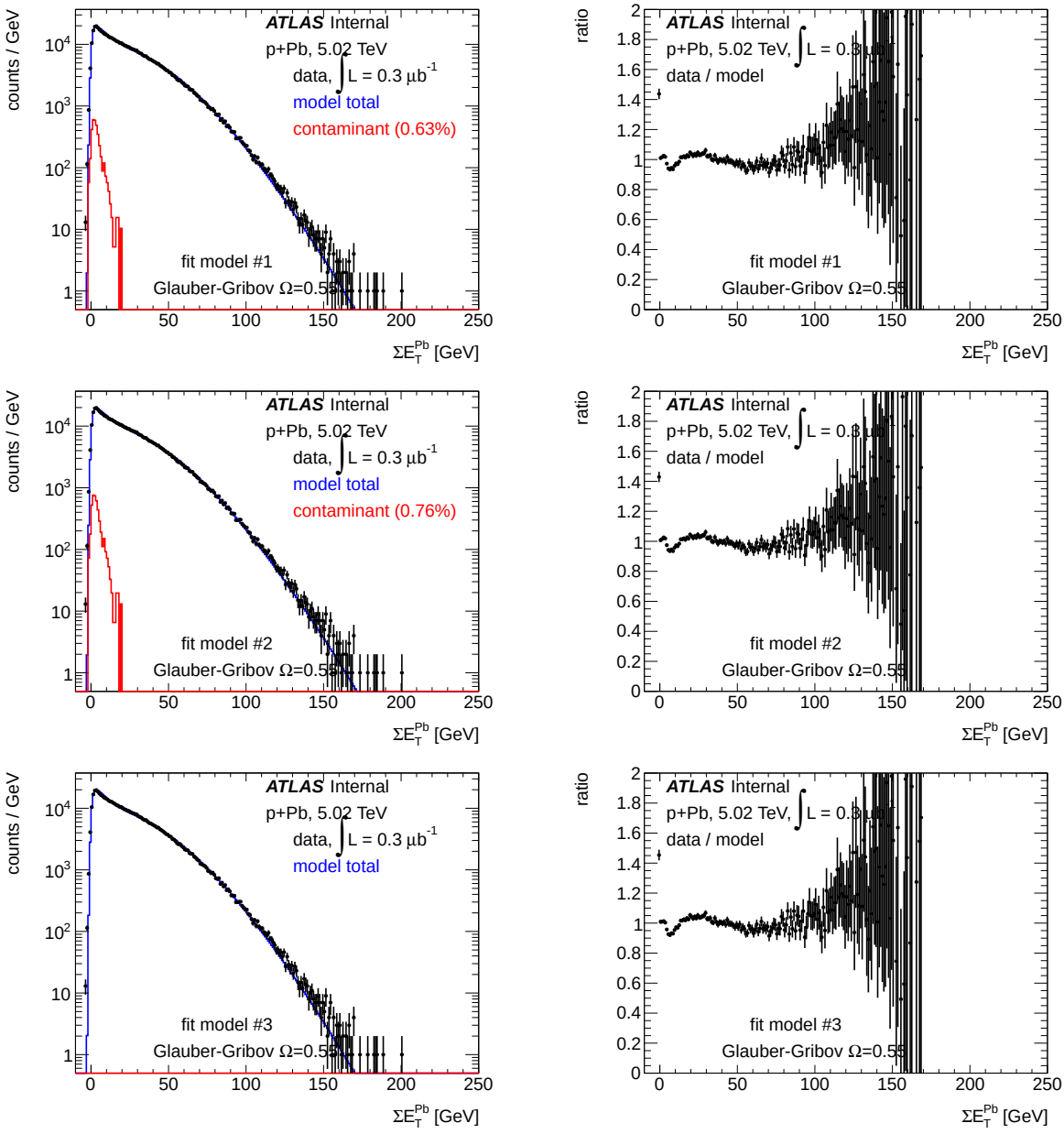


Figure A.3: Fits to 2013 p+Pb data using the Glauber-Gribov $\Omega = 0.55$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

A.2 Centrality fits with PYTHIA 6 reference

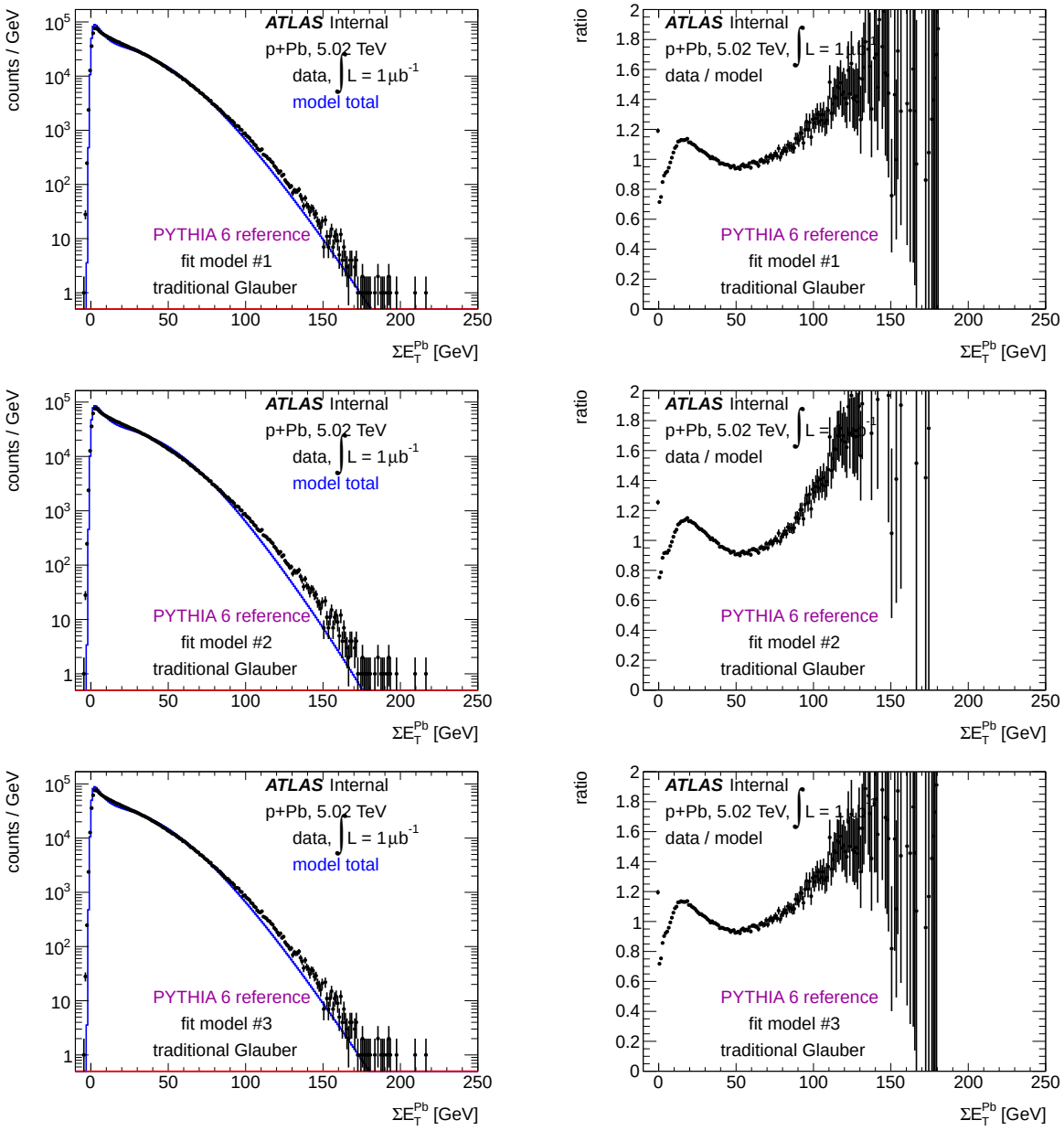


Figure A.4: Fits to 2012 p+Pb data using the default Glauber model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

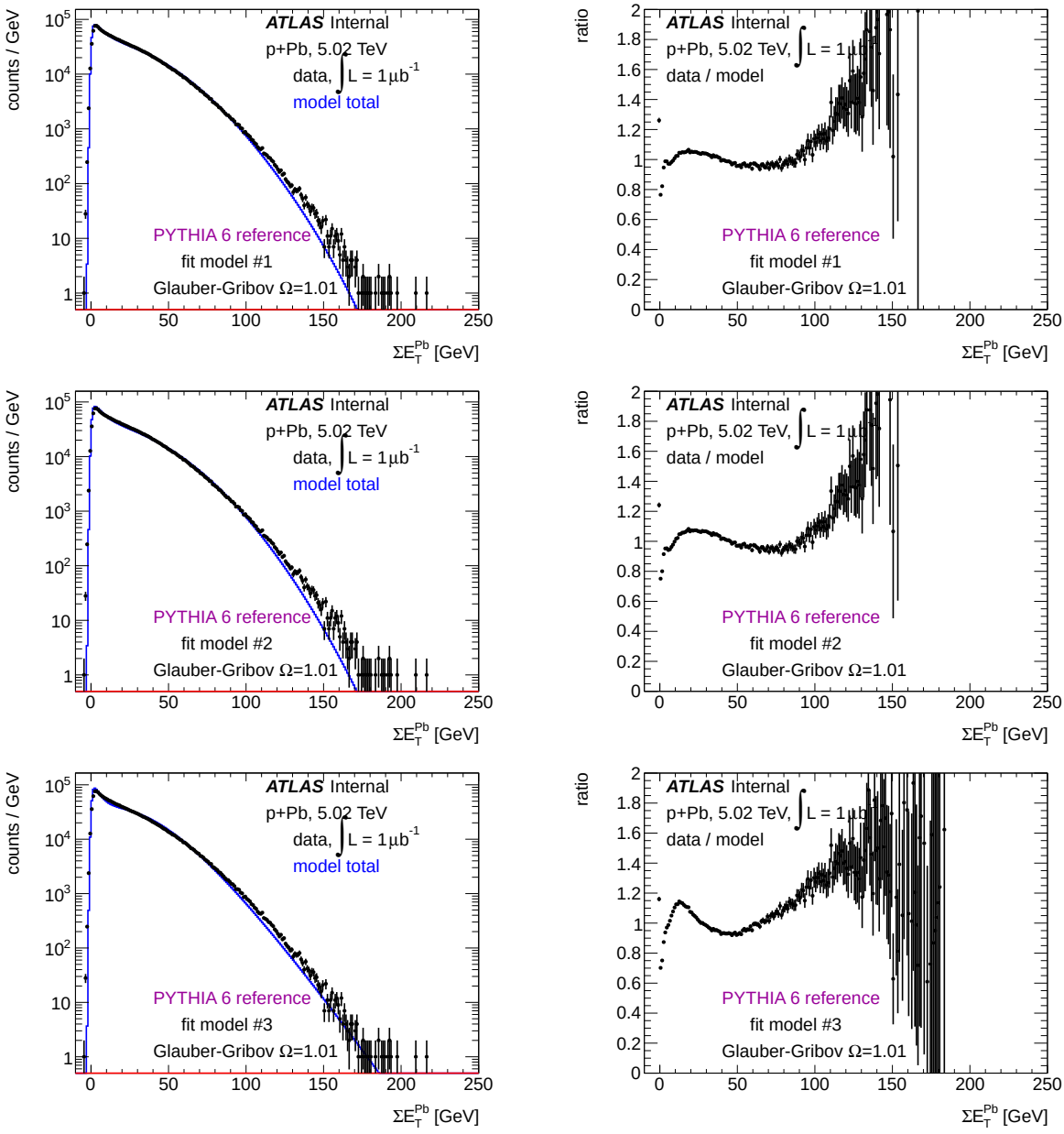


Figure A.5: Fits to 2012 p+Pb data using the Glauber-Gribov $\Omega = 1.01$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

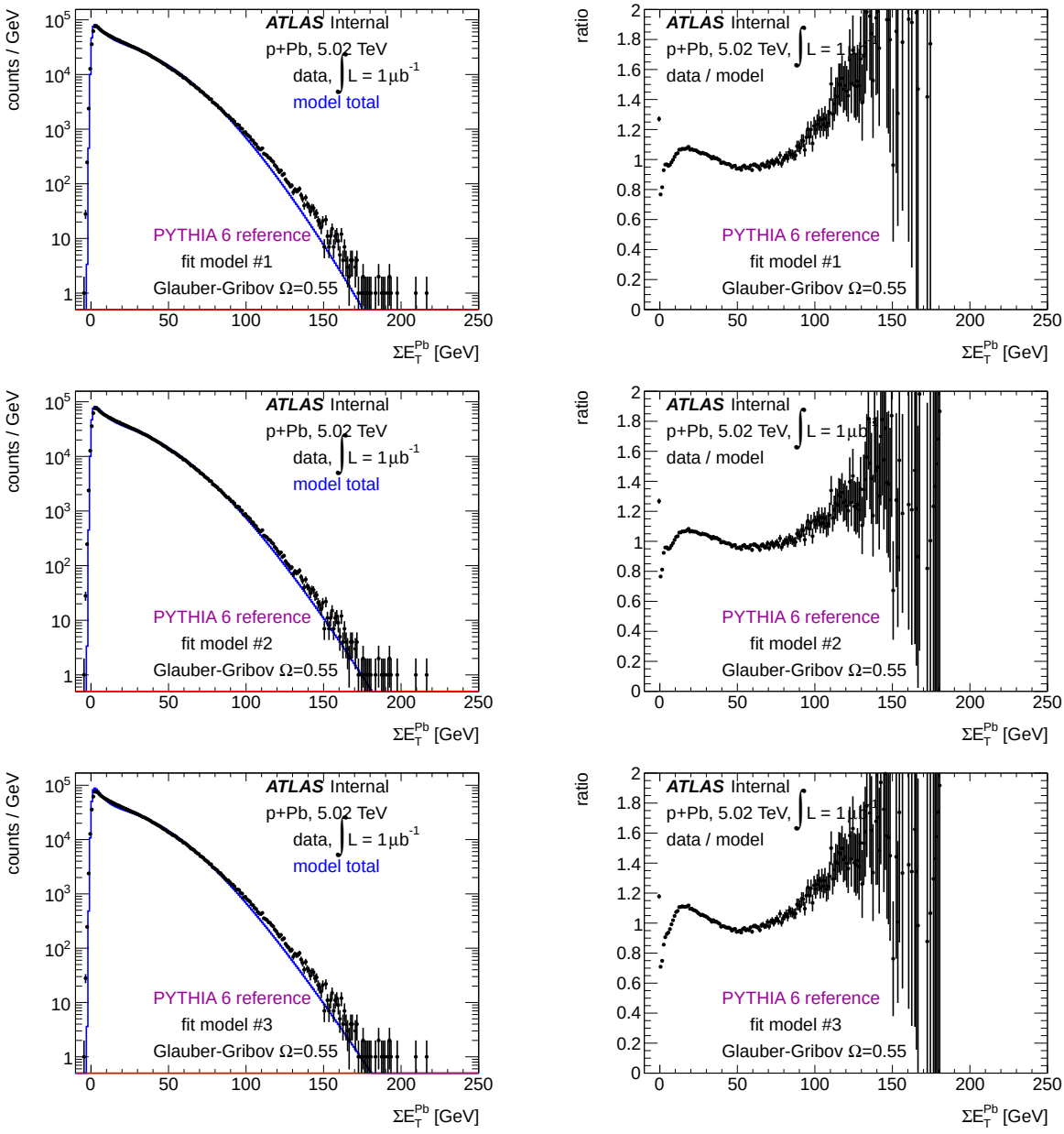


Figure A.6: Fits to 2012 p+Pb data using the Glauber-Gribov $\Omega = 0.55$ model, showing fit model #1 (first row), fit model #2 (middle row) and fit model #3 (bottom row). The left column shows the model prediction and $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ distribution in data, and the right column shows the data/model ratio.

Centrality selection	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (default)**	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (0-96%)**	$\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ **range (0-100%)**
2013 data			
0-1%	$>$ 91.01 GeV	$>$ 90.71 GeV	$>$ 91.30 GeV
0-5%	$>$ 65.95 GeV	$>$ 65.60 GeV	$>$ 66.29 GeV
0-10%	$>$ 53.69 GeV	$>$ 53.30 GeV	$>$ 54.07 GeV
10-20%	40.01-53.69 GeV	39.57-53.30 GeV	40.43-54.07 GeV
20-30%	31.03-40.01 GeV	30.55-39.57 GeV	31.50-40.43 GeV
30-40%	24.07-31.03 GeV	23.55-30.55 GeV	24.58-31.50 GeV
40-60%	13.40-24.07 GeV	12.83-23.55 GeV	13.95-24.58 GeV
60-90%	2.55-13.40 GeV	2.00-12.83 GeV	3.07-13.95 GeV

Table A.1: $\Sigma E_{\mathrm{T}}^{\mathrm{Pb}}$ ranges in 2013 p+Pb data for given centrality selections.

A.3 Tables of results for centrality fits to 2013 data

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2013 data			
60-90%	$3.976^{+0.206}_{-0.293}$ $\binom{+5.19\%}{-7.36\%}$	$3.434^{+0.26}_{-0.156}$ $\binom{+7.57\%}{-4.55\%}$	$3.591^{+0.192}_{-0.174}$ $\binom{+5.35\%}{-4.85\%}$
40-60%	$7.433^{+0.446}_{-0.529}$ $\binom{+6\%}{-7.12\%}$	$6.366^{+0.493}_{-0.339}$ $\binom{+7.74\%}{-5.32\%}$	$6.675^{+0.389}_{-0.378}$ $\binom{+5.83\%}{-5.66\%}$
30-40%	$9.806^{+0.565}_{-0.574}$ $\binom{+5.76\%}{-5.85\%}$	$8.968^{+0.596}_{-0.49}$ $\binom{+6.65\%}{-5.46\%}$	$9.22^{+0.498}_{-0.536}$ $\binom{+5.4\%}{-5.81\%}$
20-30%	$11.37^{+0.65}_{-0.631}$ $\binom{+5.72\%}{-5.55\%}$	$11.17^{+0.68}_{-0.625}$ $\binom{+6.09\%}{-5.59\%}$	$11.23^{+0.619}_{-0.665}$ $\binom{+5.51\%}{-5.92\%}$
10-20%	$13.05^{+0.821}_{-0.734}$ $\binom{+6.3\%}{-5.62\%}$	$14.11^{+0.862}_{-0.792}$ $\binom{+6.11\%}{-5.61\%}$	$13.77^{+0.79}_{-0.814}$ $\binom{+5.74\%}{-5.91\%}$
0-10%	$15.56^{+1.48}_{-0.874}$ $\binom{+9.54\%}{-5.62\%}$	$19.98^{+1.29}_{-1.7}$ $\binom{+6.45\%}{-8.53\%}$	$18.42^{+1.14}_{-1.2}$ $\binom{+6.18\%}{-6.52\%}$
5-10%	$14.61^{+1.21}_{-0.824}$ $\binom{+8.27\%}{-5.64\%}$	$17.46^{+1.13}_{-1.08}$ $\binom{+6.49\%}{-6.18\%}$	16.49^{+1}_{-1} $\binom{+6.07\%}{-6.09\%}$
0-5%	$16.53^{+1.85}_{-0.927}$ $\binom{+11.2\%}{-5.61\%}$	$22.61^{+1.48}_{-2.46}$ $\binom{+6.52\%}{-10.9\%}$	$20.45^{+1.29}_{-1.45}$ $\binom{+6.3\%}{-7.1\%}$
1-5%	$16.1^{+1.66}_{-0.905}$ $\binom{+10.3\%}{-5.62\%}$	$21.41^{+1.46}_{-1.98}$ $\binom{+6.84\%}{-9.25\%}$	$19.52^{+1.24}_{-1.32}$ $\binom{+6.34\%}{-6.76\%}$
0-1%	$18.22^{+2.62}_{-1.01}$ $\binom{+14.4\%}{-5.56\%}$	$27.42^{+1.58}_{-4.48}$ $\binom{+5.77\%}{-16.3\%}$	$24.18^{+1.52}_{-2.07}$ $\binom{+6.3\%}{-8.55\%}$
0-90%	$8.467^{+0.48}_{-0.414}$ $\binom{+5.66\%}{-4.89\%}$	$8.619^{+0.506}_{-0.446}$ $\binom{+5.87\%}{-5.17\%}$	$8.535^{+0.474}_{-0.461}$ $\binom{+5.55\%}{-5.4\%}$

Table A.2: Mean N_{part} values for all centrality bins in 2013 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

A.4 Summary of systematic uncertainties for Glauber-Gribov results

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2013 data			
60-90%	$0.04251^{+0.00297}_{-0.00412}$ $\binom{+6.97\%}{-9.7\%}$	$0.03477^{+0.0037}_{-0.00211}$ $\binom{+10.6\%}{-6.06\%}$	$0.03701^{+0.00266}_{-0.00227}$ $\binom{+7.18\%}{-6.12\%}$
40-60%	$0.0919^{+0.00428}_{-0.0062}$ $\binom{+4.66\%}{-6.75\%}$	$0.07666^{+0.00635}_{-0.00328}$ $\binom{+8.28\%}{-4.28\%}$	$0.08107^{+0.00428}_{-0.0034}$ $\binom{+5.28\%}{-4.19\%}$
30-40%	$0.1258^{+0.00333}_{-0.00413}$ $\binom{+2.65\%}{-3.28\%}$	$0.1138^{+0.00641}_{-0.00337}$ $\binom{+5.63\%}{-2.96\%}$	$0.1174^{+0.00344}_{-0.00323}$ $\binom{+2.93\%}{-2.75\%}$
20-30%	$0.1481^{+0.00359}_{-0.00244}$ $\binom{+2.42\%}{-1.65\%}$	$0.1453^{+0.00538}_{-0.00316}$ $\binom{+3.7\%}{-2.18\%}$	$0.1462^{+0.00333}_{-0.0031}$ $\binom{+2.28\%}{-2.12\%}$
10-20%	$0.1721^{+0.00709}_{-0.003}$ $\binom{+4.12\%}{-1.74\%}$	$0.1873^{+0.00483}_{-0.00294}$ $\binom{+2.58\%}{-1.57\%}$	$0.1824^{+0.00415}_{-0.00317}$ $\binom{+2.27\%}{-1.74\%}$
0-10%	$0.208^{+0.0188}_{-0.00487}$ $\binom{+9.05\%}{-2.34\%}$	$0.2711^{+0.00741}_{-0.0188}$ $\binom{+2.73\%}{-6.94\%}$	$0.2489^{+0.00787}_{-0.00864}$ $\binom{+3.16\%}{-3.47\%}$
5-10%	$0.1944^{+0.0143}_{-0.00409}$ $\binom{+7.35\%}{-2.11\%}$	$0.2351^{+0.00647}_{-0.00717}$ $\binom{+2.75\%}{-3.05\%}$	$0.2213^{+0.00601}_{-0.00502}$ $\binom{+2.72\%}{-2.27\%}$
0-5%	$0.2218^{+0.0245}_{-0.00563}$ $\binom{+11.1\%}{-2.54\%}$	$0.3087^{+0.00852}_{-0.0306}$ $\binom{+2.76\%}{-9.92\%}$	$0.2779^{+0.00957}_{-0.0128}$ $\binom{+3.44\%}{-4.61\%}$
1-5%	$0.2157^{+0.0216}_{-0.00528}$ $\binom{+10\%}{-2.45\%}$	$0.2916^{+0.00797}_{-0.023}$ $\binom{+2.73\%}{-7.89\%}$	$0.2646^{+0.00841}_{-0.0103}$ $\binom{+3.18\%}{-3.89\%}$
0-1%	$0.246^{+0.0361}_{-0.00699}$ $\binom{+14.7\%}{-2.84\%}$	$0.3775^{+0.012}_{-0.0611}$ $\binom{+3.18\%}{-16.2\%}$	$0.3311^{+0.0145}_{-0.0234}$ $\binom{+4.38\%}{-7.07\%}$
0-90%	$0.1067^{+0.00454}_{-0.00267}$ $\binom{+4.26\%}{-2.5\%}$	$0.1088^{+0.00442}_{-0.00245}$ $\binom{+4.06\%}{-2.26\%}$	$0.1076^{+0.00402}_{-0.00263}$ $\binom{+3.73\%}{-2.44\%}$

Table A.3: Mean T_{AB} values for all centrality bins in 2013 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

Centrality selection	**standard Glauber**	**Glauber-Gribov** $\Omega = 1.01$	**Glauber-Gribov** $\Omega = 0.55$
2013 data			
40-60% / 60-90%	$2.162^{+0.0833}_{-0.0693}$ $\binom{+3.85\%}{-3.21\%}$	$2.205^{+0.0529}_{-0.0624}$ $\binom{+2.4\%}{-2.83\%}$	$2.191^{+0.057}_{-0.0602}$ $\binom{+2.6\%}{-2.75\%}$
30-40% / 60-90%	$2.959^{+0.21}_{-0.135}$ $\binom{+7.09\%}{-4.55\%}$	$3.274^{+0.121}_{-0.161}$ $\binom{+3.68\%}{-4.9\%}$	$3.173^{+0.125}_{-0.133}$ $\binom{+3.95\%}{-4.2\%}$
20-30% / 60-90%	$3.484^{+0.329}_{-0.175}$ $\binom{+9.43\%}{-5.03\%}$	$4.178^{+0.185}_{-0.273}$ $\binom{+4.42\%}{-6.54\%}$	$3.95^{+0.18}_{-0.204}$ $\binom{+4.56\%}{-5.17\%}$
10-20% / 60-90%	$4.048^{+0.493}_{-0.214}$ $\binom{+12.2\%}{-5.28\%}$	$5.385^{+0.272}_{-0.466}$ $\binom{+5.05\%}{-8.65\%}$	$4.929^{+0.248}_{-0.307}$ $\binom{+5.03\%}{-6.23\%}$
0-10% / 60-90%	$4.892^{+0.827}_{-0.267}$ $\binom{+16.9\%}{-5.46\%}$	$7.798^{+0.43}_{-1.17}$ $\binom{+5.51\%}{-15.1\%}$	$6.725^{+0.368}_{-0.571}$ $\binom{+5.47\%}{-8.48\%}$
5-10% / 60-90%	$4.572^{+0.69}_{-0.247}$ $\binom{+15.1\%}{-5.4\%}$	$6.762^{+0.368}_{-0.78}$ $\binom{+5.45\%}{-11.5\%}$	$5.979^{+0.321}_{-0.441}$ $\binom{+5.36\%}{-7.37\%}$
0-5% / 60-90%	$5.217^{+0.986}_{-0.294}$ $\binom{+18.9\%}{-5.63\%}$	$8.878^{+0.501}_{-1.58}$ $\binom{+5.64\%}{-17.7\%}$	$7.508^{+0.42}_{-0.714}$ $\binom{+5.59\%}{-9.51\%}$
1-5% / 60-90%	$5.073^{+0.908}_{-0.282}$ $\binom{+17.9\%}{-5.55\%}$	$8.385^{+0.488}_{-1.34}$ $\binom{+5.82\%}{-16\%}$	$7.149^{+0.401}_{-0.633}$ $\binom{+5.61\%}{-8.86\%}$
0-1% / 60-90%	$5.787^{+1.3}_{-0.342}$ $\binom{+22.4\%}{-5.91\%}$	$10.86^{+0.575}_{-2.54}$ $\binom{+5.3\%}{-23.4\%}$	$8.947^{+0.503}_{-1.04}$ $\binom{+5.62\%}{-11.7\%}$

Table A.4: Mean R_{coll} values for all centrality bins in 2013 data, along with total systematic uncertainty (the uncertainty is asymmetric and shown for positive and negative uncertainties separately, and in absolute and relative terms).

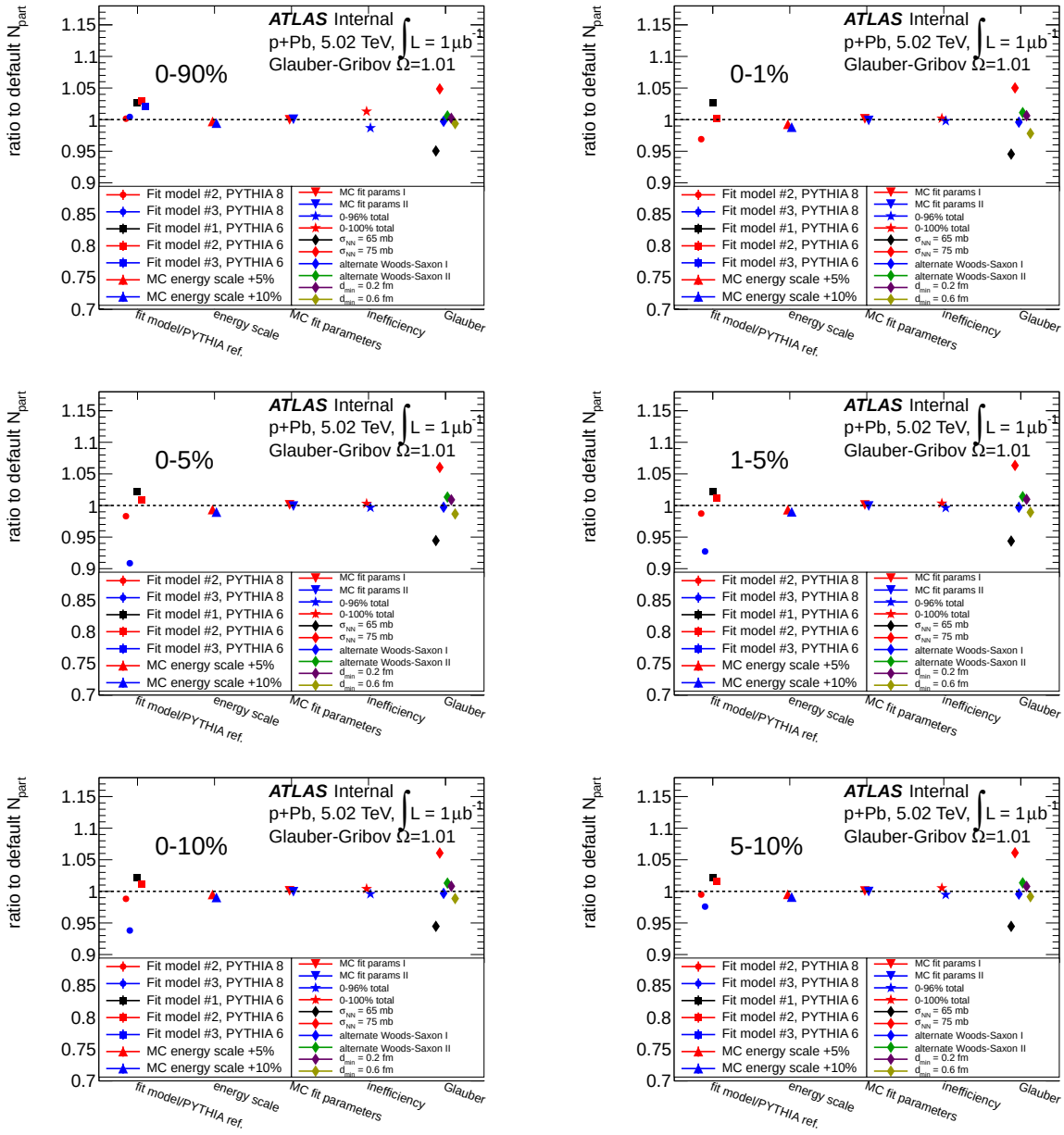


Figure A.7: Overview of systematic uncertainty for the Glauber-Gribov $\Omega = 1.01$ model. The change in $\langle N_{\rm part} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on $N_{\rm part}$ in the 0-90% (top left), 0-1% (top right), 0-5% (middle left), 1-5% (middle right), 0-10% (bottom left) and 5-10% (bottom right) centrality bins.

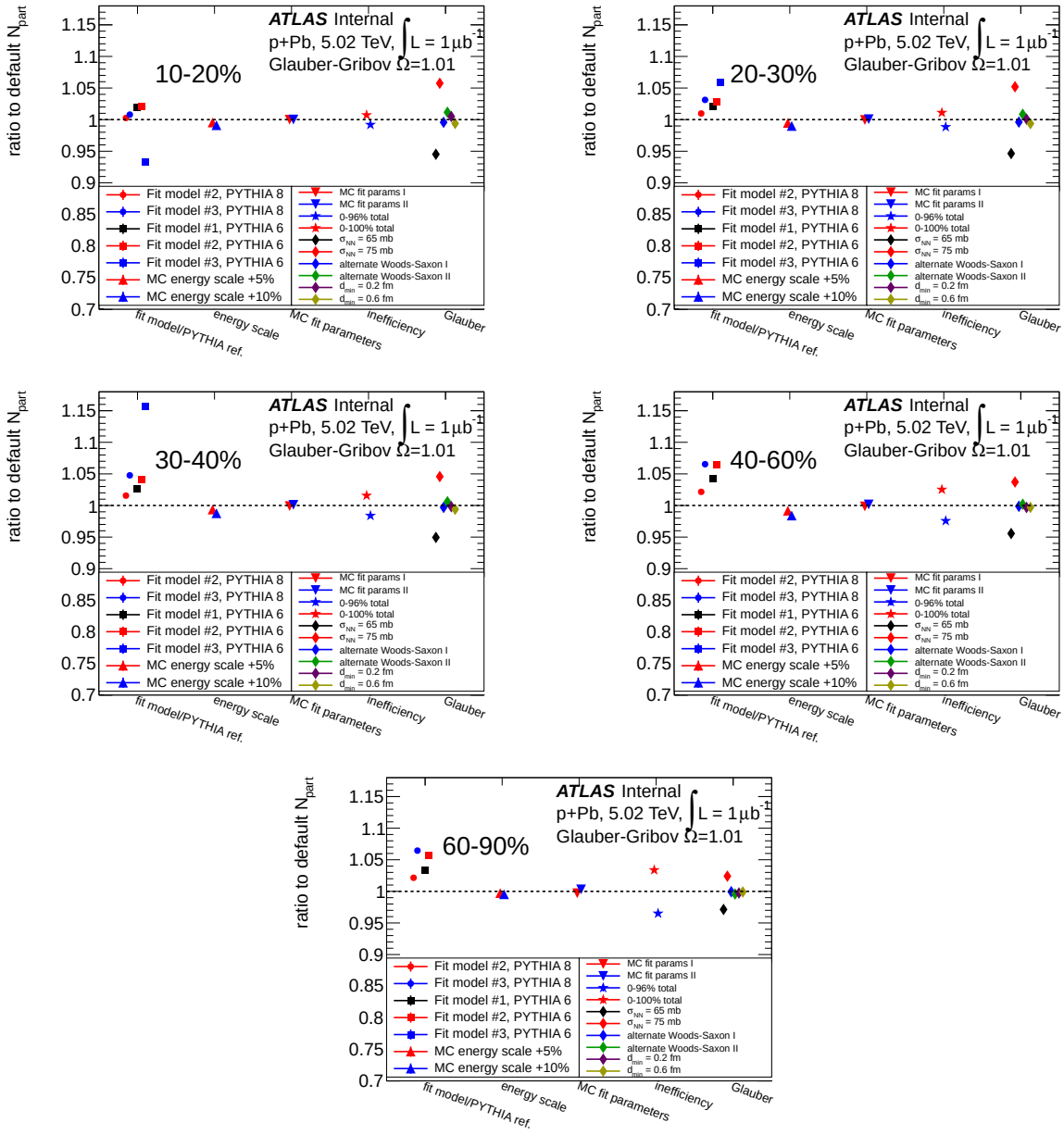


Figure A.8: Overview of systematic uncertainty for the Glauber-Gribov $\Omega = 1.01$ model. The change in $\langle N_{\rm part} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on $N_{\rm part}$ in the 10-20% (top left), 20-30% (top right), 30-40% (middle left), 40-60% (middle right) and 60-90% (bottom) centrality bins.

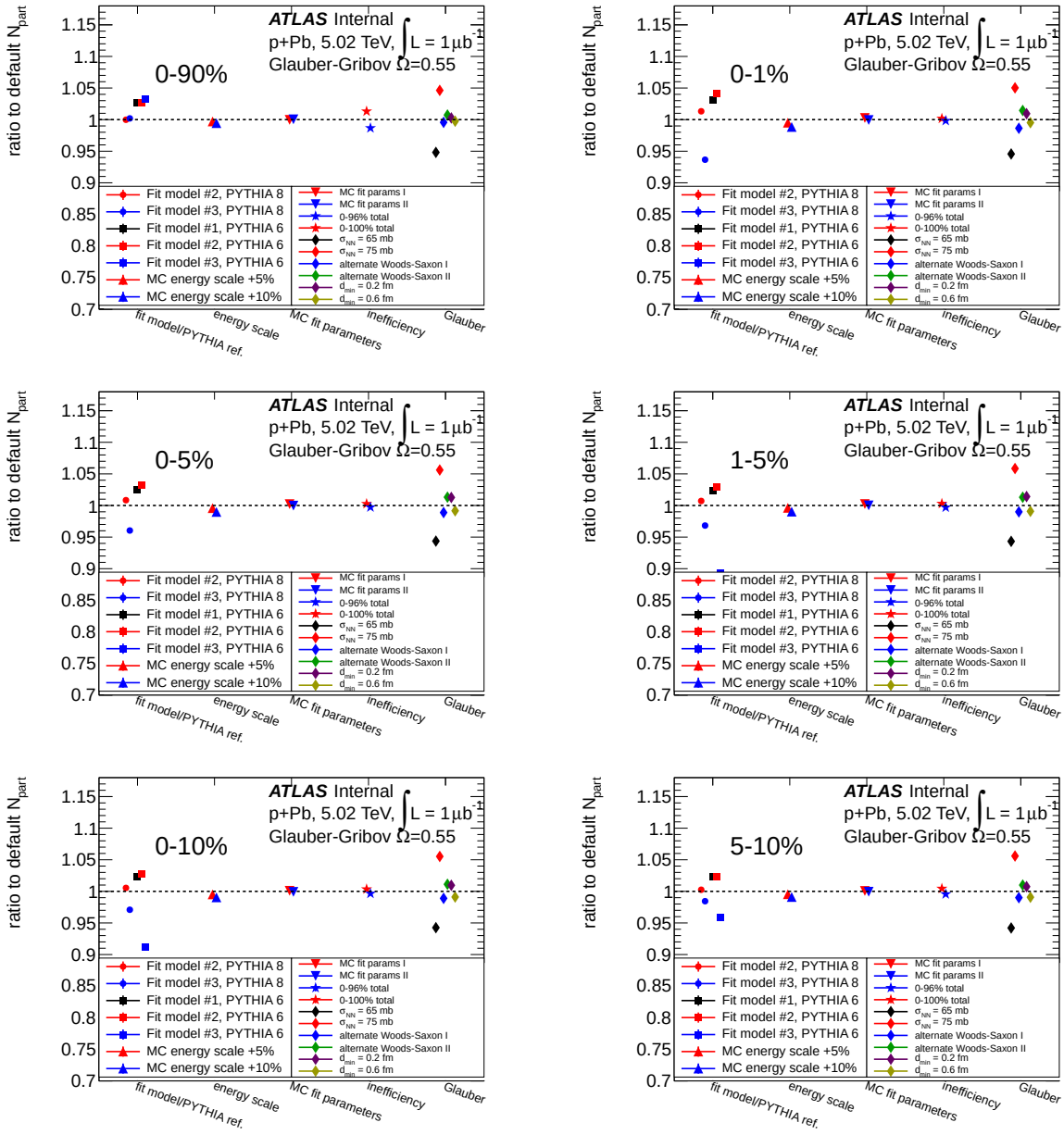


Figure A.9: Overview of systematic uncertainty for the Glauber-Gribov $\Omega = 0.55$ model. The change in $\langle N_{\rm part} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on $N_{\rm part}$ in the 0-90% (top left), 0-1% (top right), 0-5% (middle left), 1-5% (middle right), 0-10% (bottom left) and 5-10% (bottom right) centrality bins.

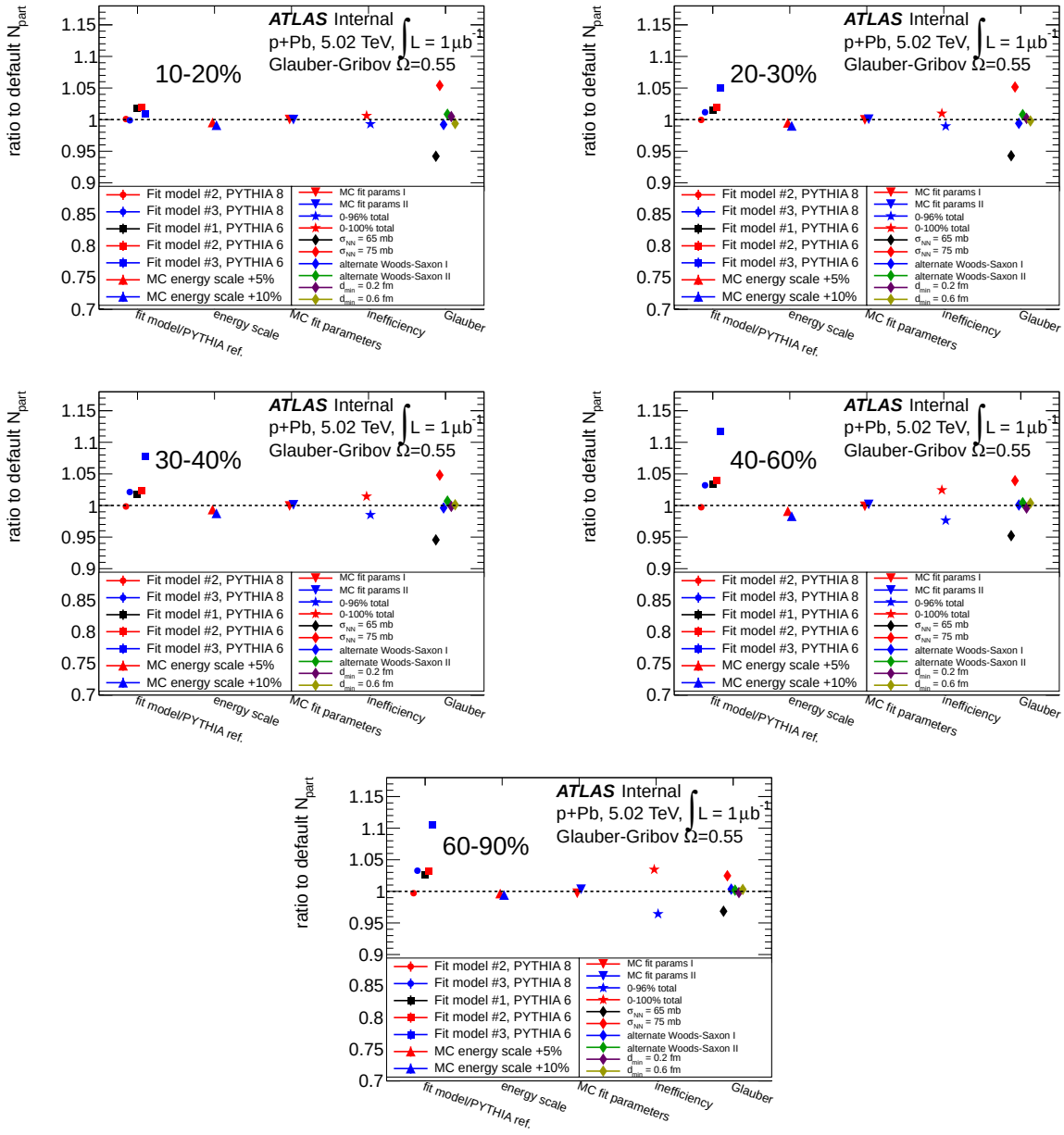


Figure A.10: Overview of systematic uncertainty for the Glauber-Gribov $\Omega = 0.55$ model. The change in $\langle N_{\rm part} \rangle$ relative to the default value as a function of possible systematic variations is plotted. The figures here show the effects on $N_{\rm part}$ in the 10-20% (top left), 20-30% (top right), 30-40% (middle left), 40-60% (middle right) and 60-90% (bottom) centrality bins.

Appendix B

p+Pb and pp Jet results

B.1 Data-driven cleaning efficiency in p+Pb and pp

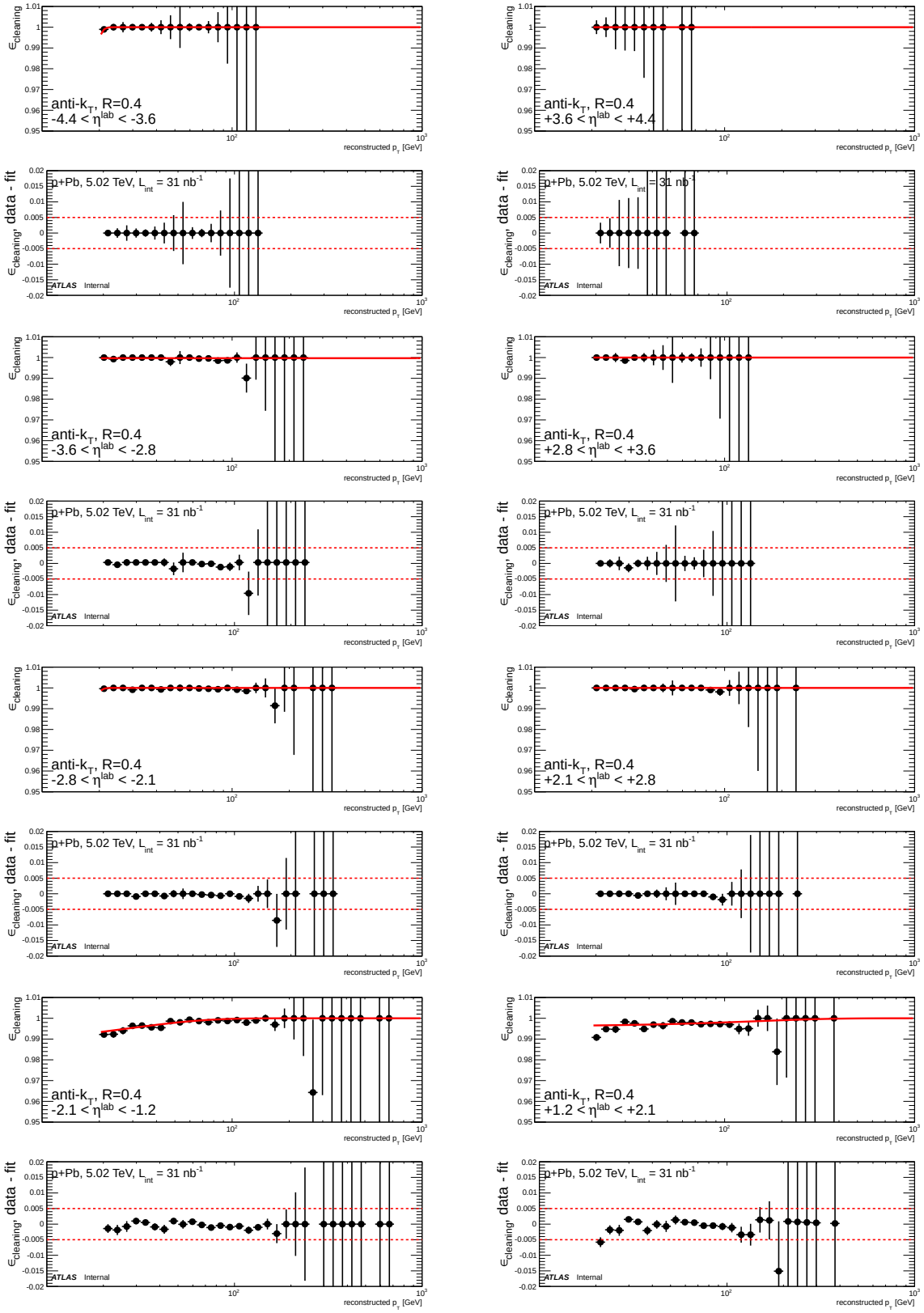


Figure B.1: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.4 jets in p+Pb. This figure shows pseudorapidity bins with the range $\left|\eta^{\mathrm{lab}}\right| < 1.2$.

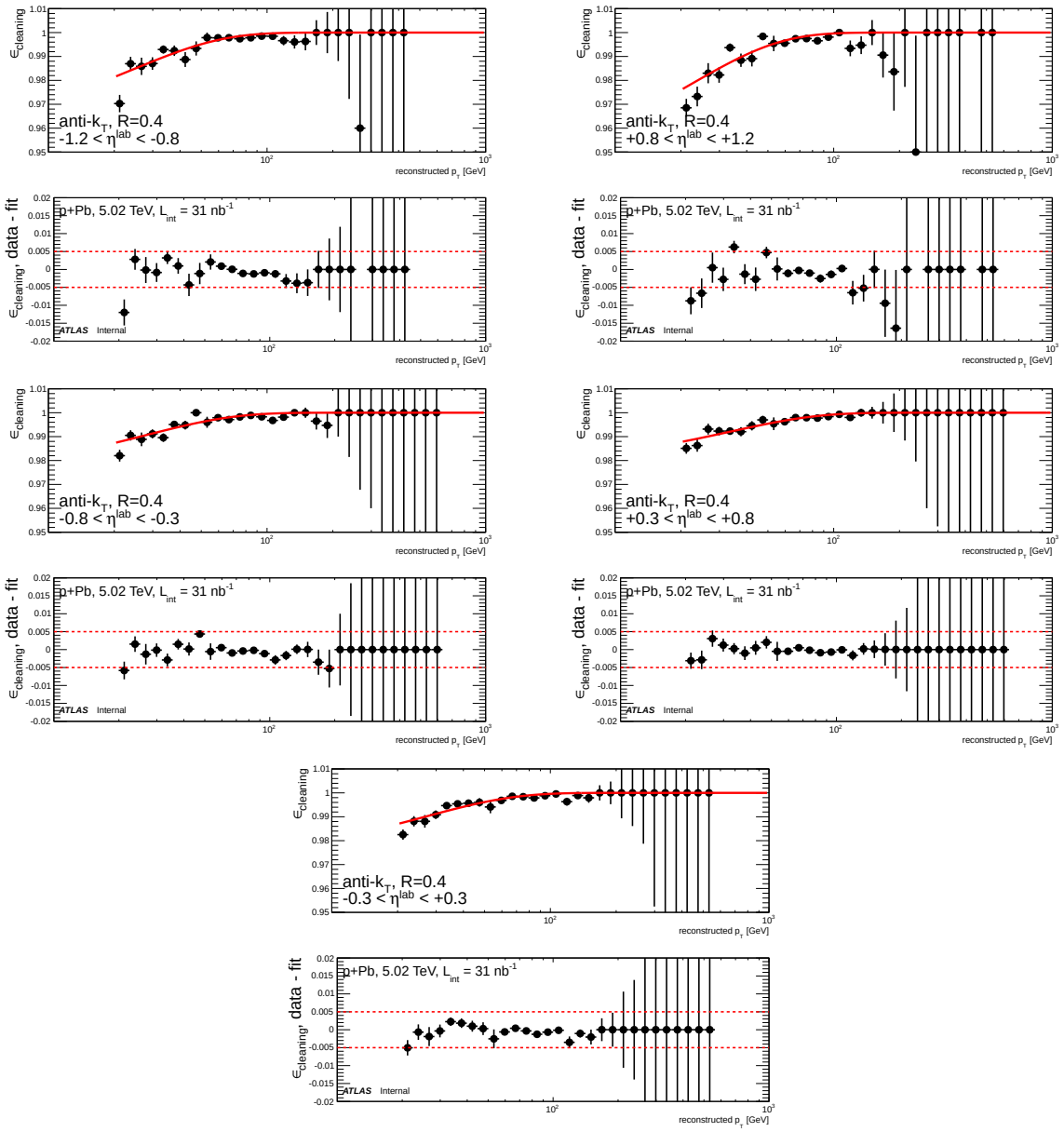


Figure B.2: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.4 jets in p+Pb. This figure shows pseudorapidity bins with the range $\left|\eta^{\mathrm{lab}}\right| > 1.2$.

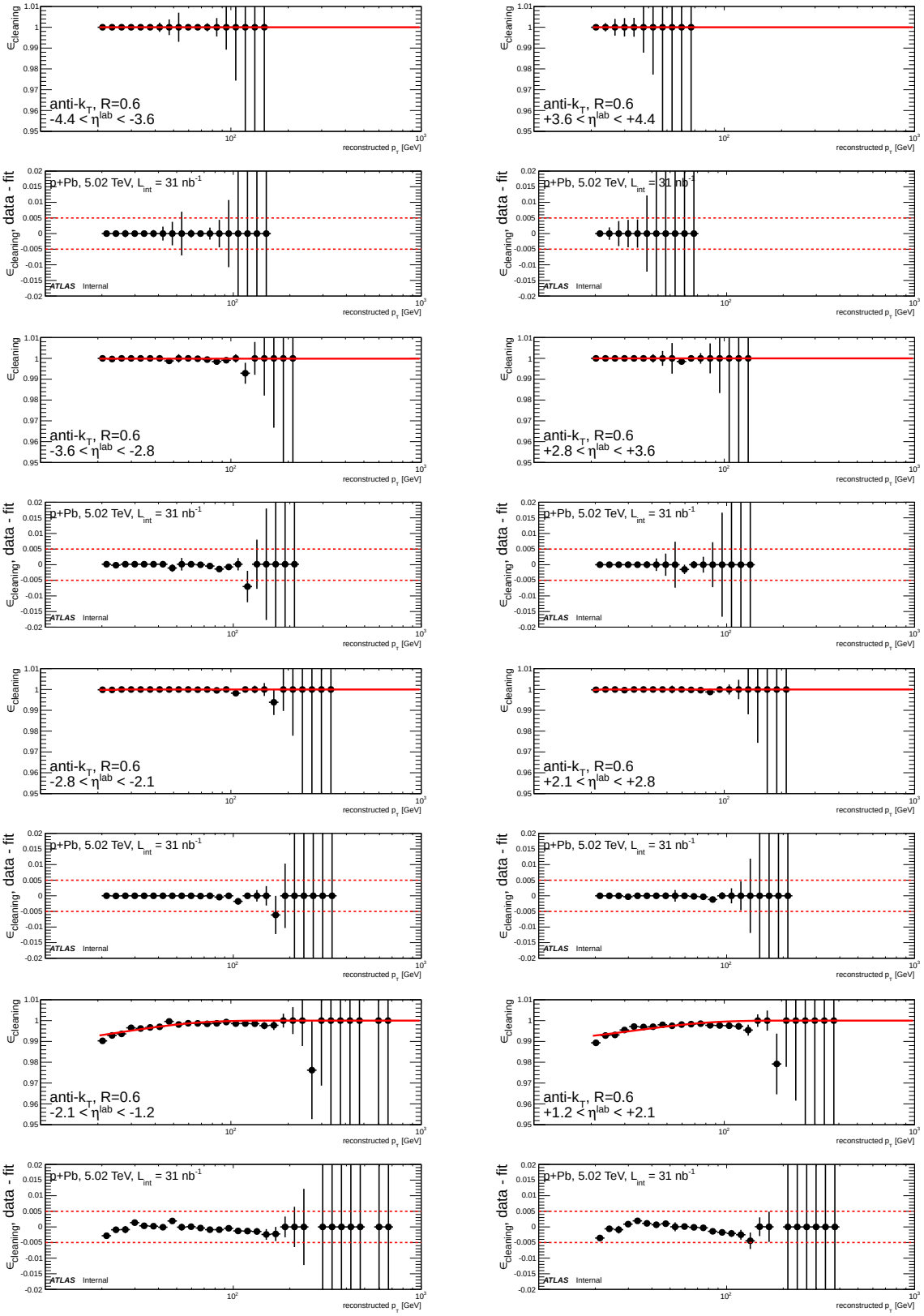


Figure B.3: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.6 jets in p+Pb. This figure shows pseudorapidity bins with the range $\left|\eta^{\mathrm{lab}}\right| < 1.2$.

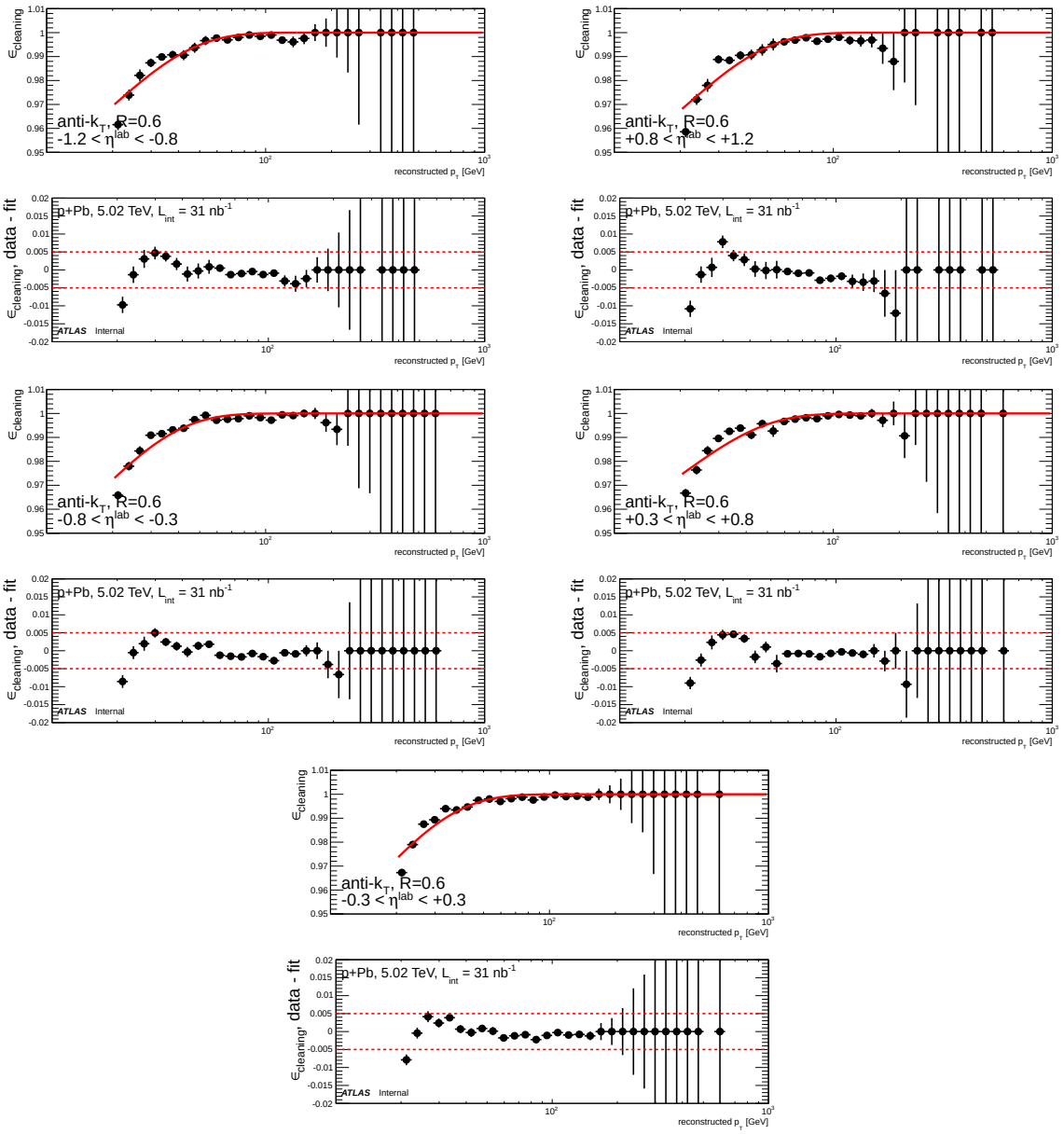


Figure B.4: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.6 jets in p+Pb. This figure shows pseudorapidity bins with the range $\left|\eta^{\mathrm{lab}}\right| > 1.2$.

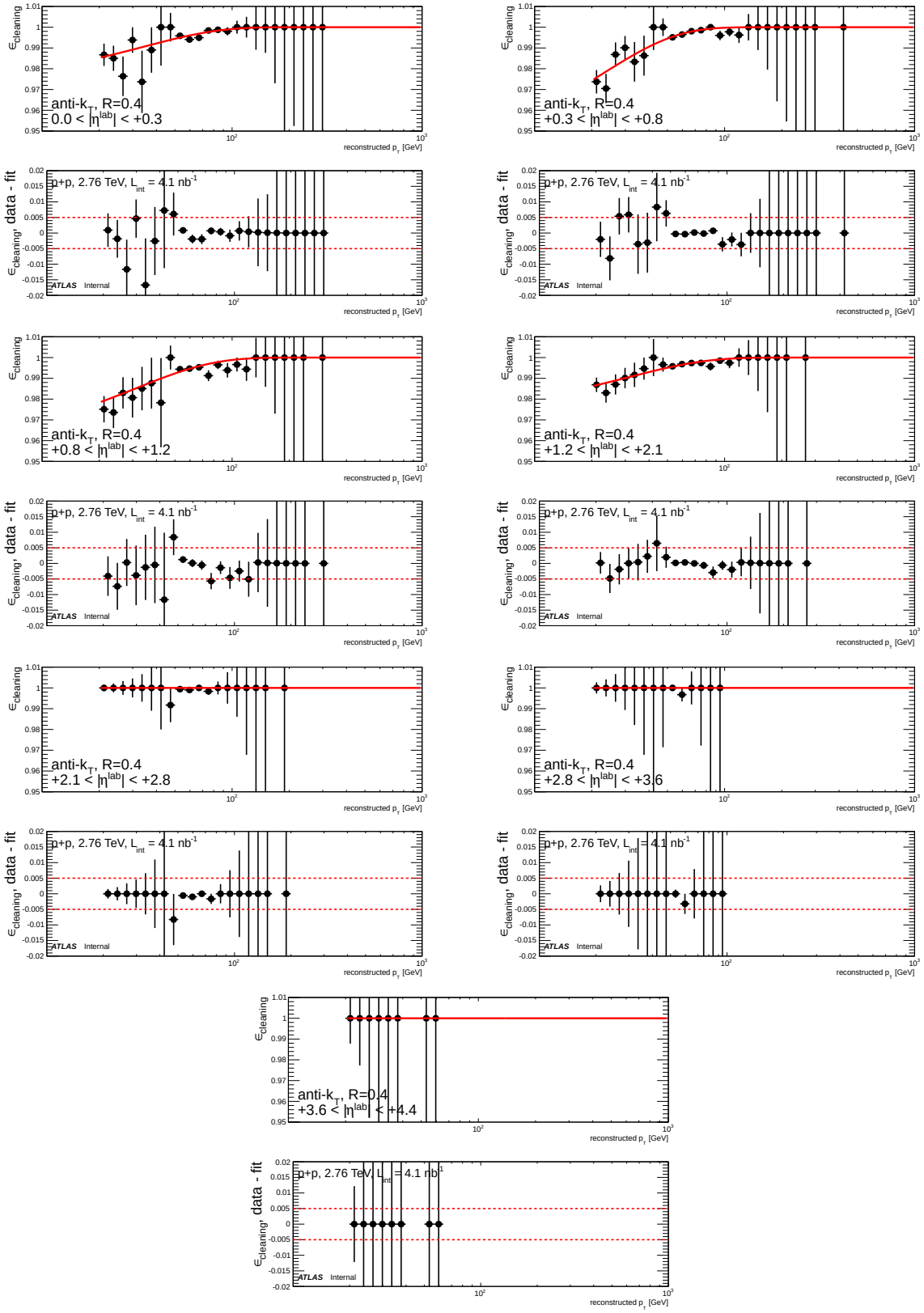


Figure B.5: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.4 jets in *pp*. Each of the seven pseudorapidity bins are shown separately.

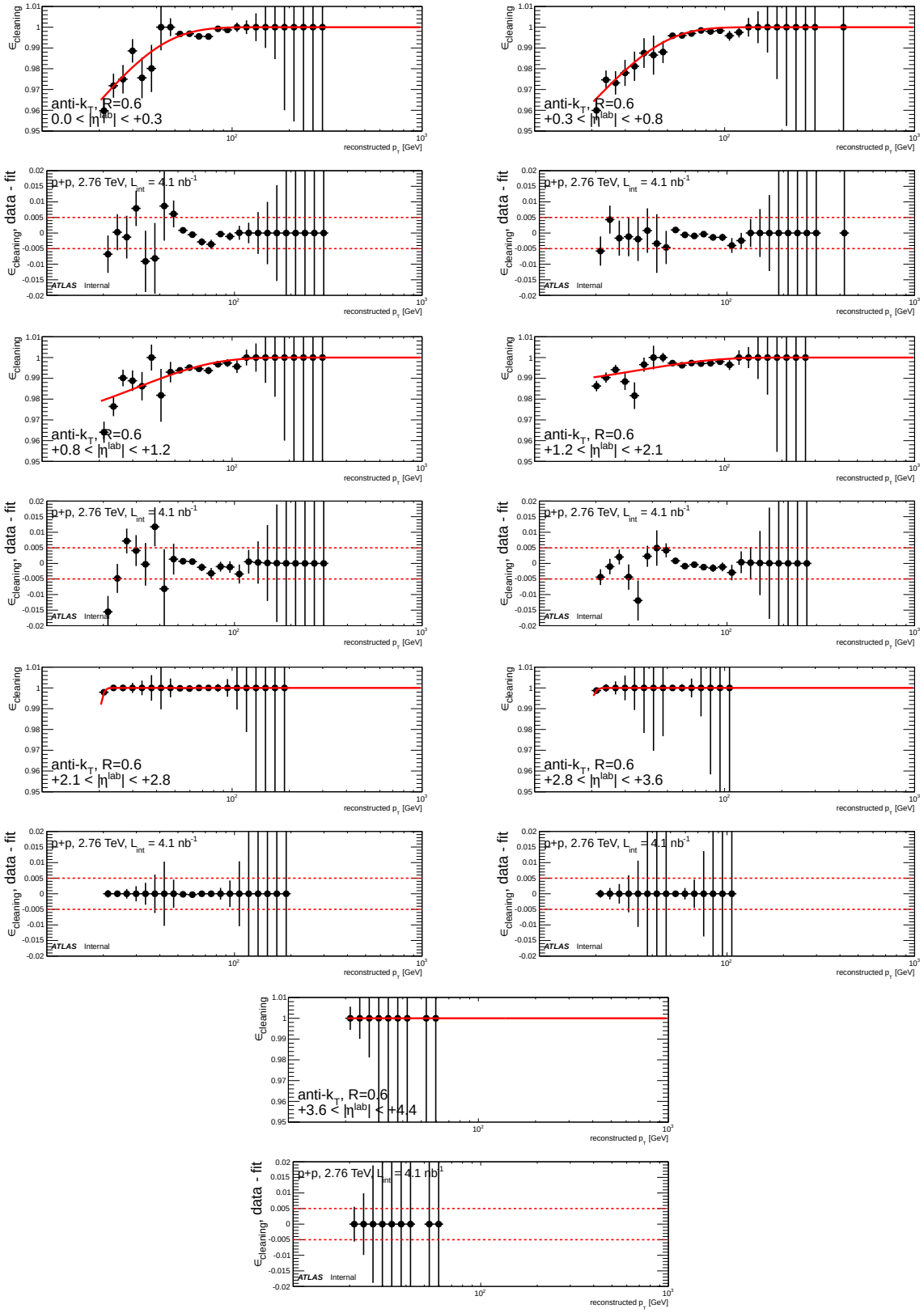


Figure B.6: Efficiency of cleaning cuts determined from the good jet sample, as a function of p_{T}, for R=0.6 jets in *pp*. Each of the seven pseudorapidity bins are shown separately.

B.2 Trigger efficiency in p+Pb and pp

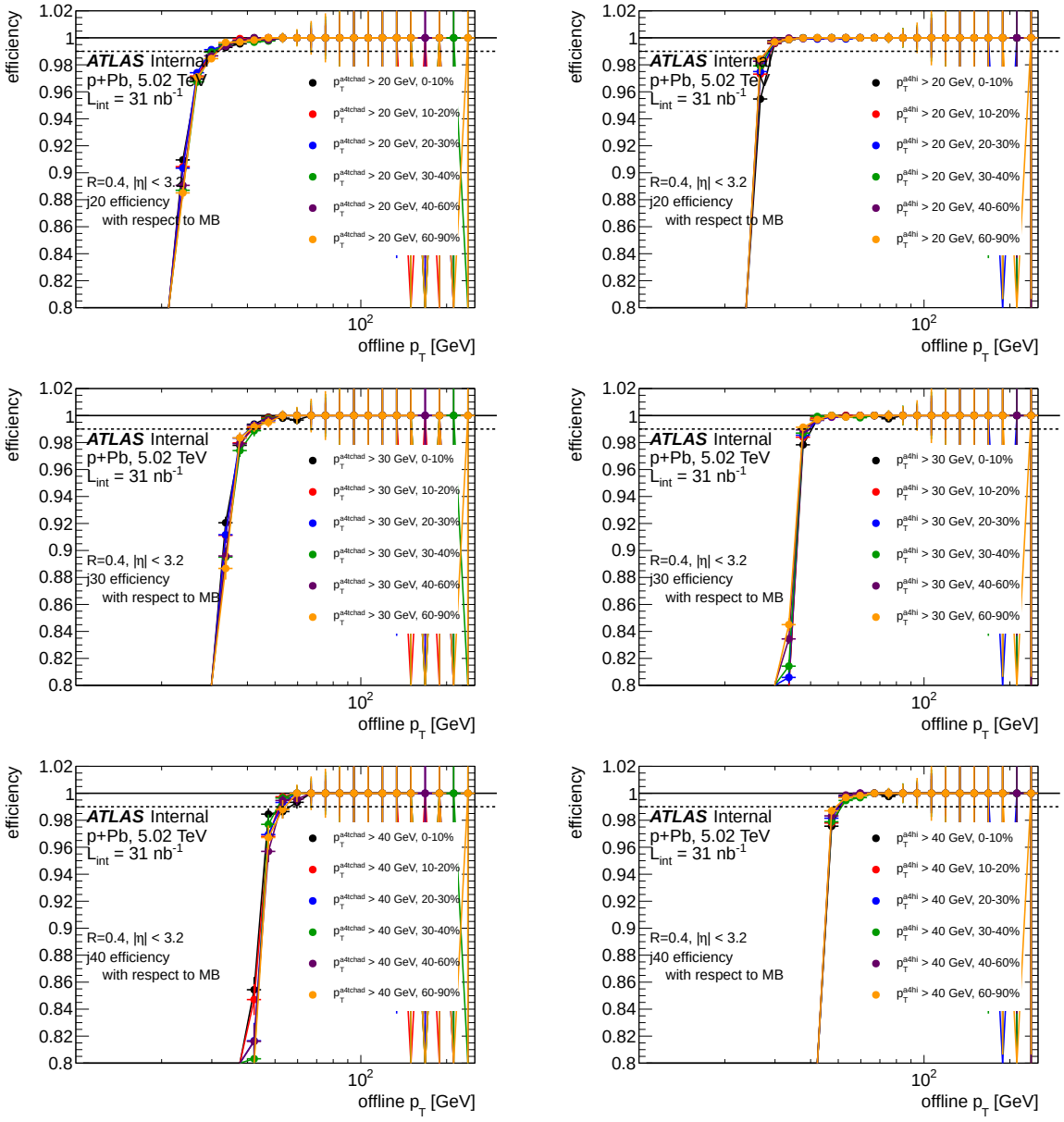


Figure B.7: Centrality dependence of the R=0.4 jet trigger efficiency, shown for j20 triggers (first plot), j30 triggers (second plot) and j40 triggers (third plot). Results for `a4tchad` (left column) and `a4hi` (right column) triggers are shown.

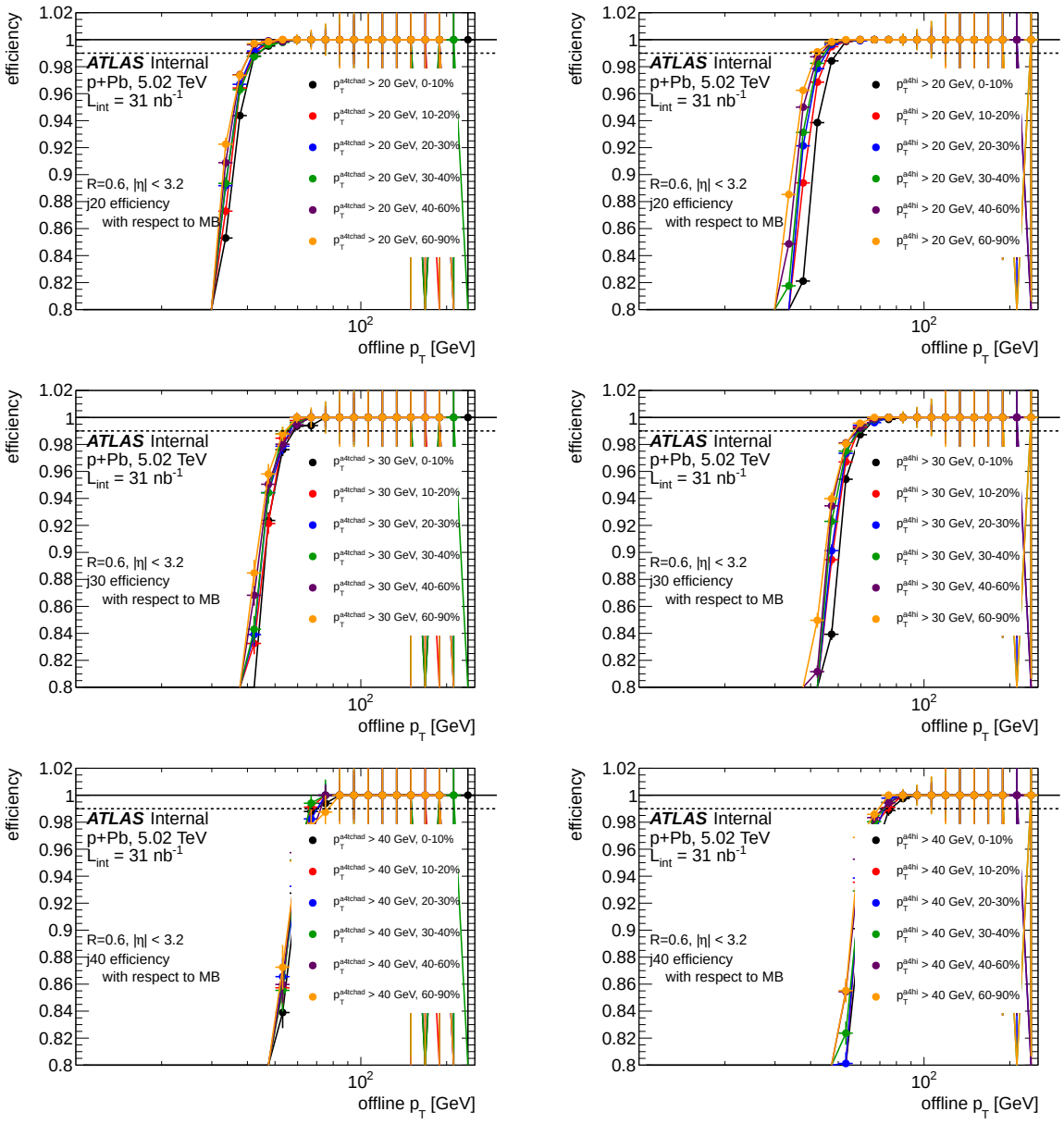


Figure B.8: Centrality dependence of the R=0.6 jet trigger efficiency, shown for j20 triggers (first plot), j30 triggers (second plot) and j40 triggers (third plot). Results for `a4tchad` (left column) and `a4hi` (right column) triggers are shown.

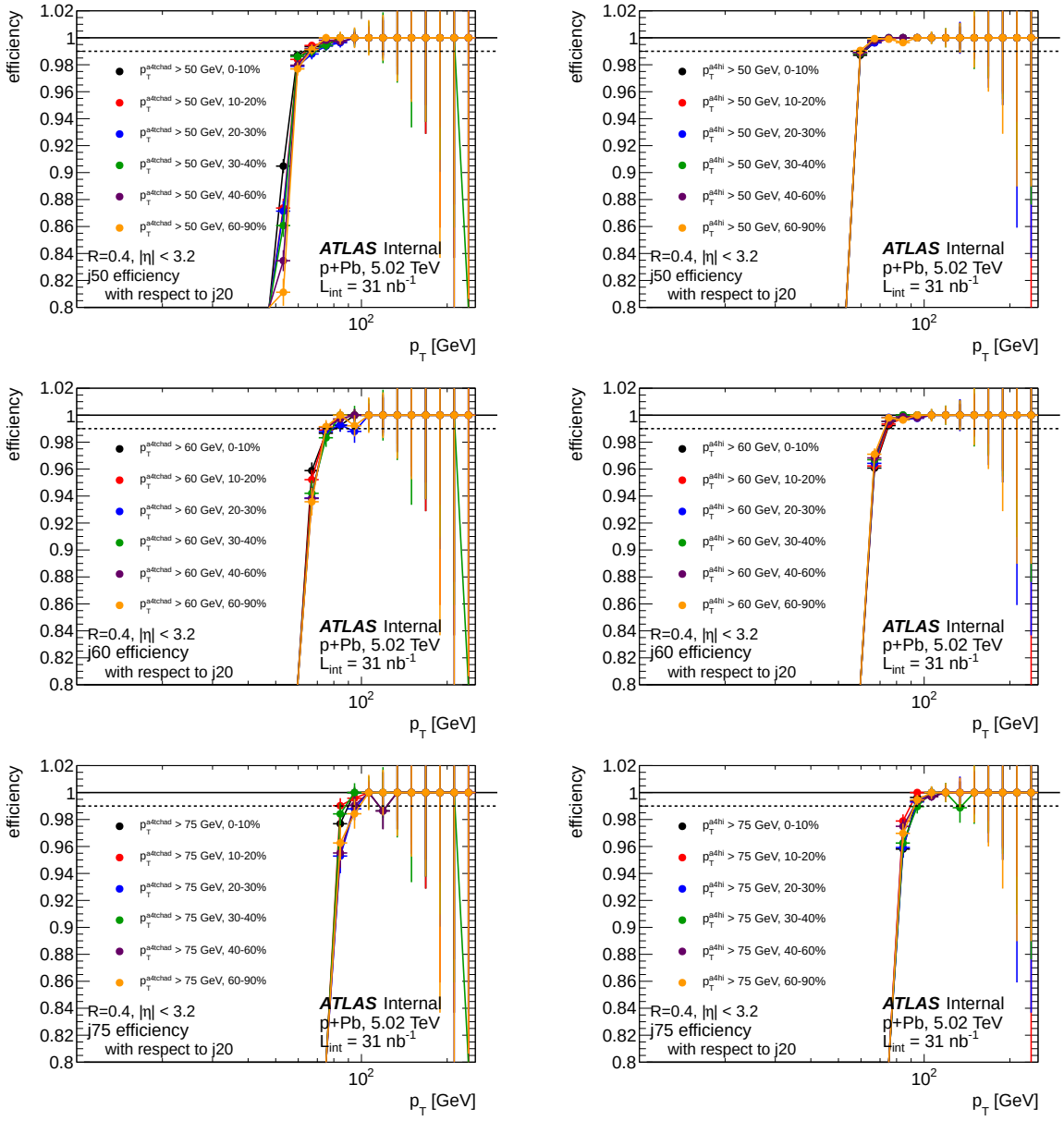


Figure B.9: Centrality dependence of the R=0.4 jet trigger efficiency, shown for j50 triggers (first plot), j60 triggers (second plot) and j75 triggers (third plot). Results for `a4tchad` (left column) and `a4hi` (right column) triggers are shown.

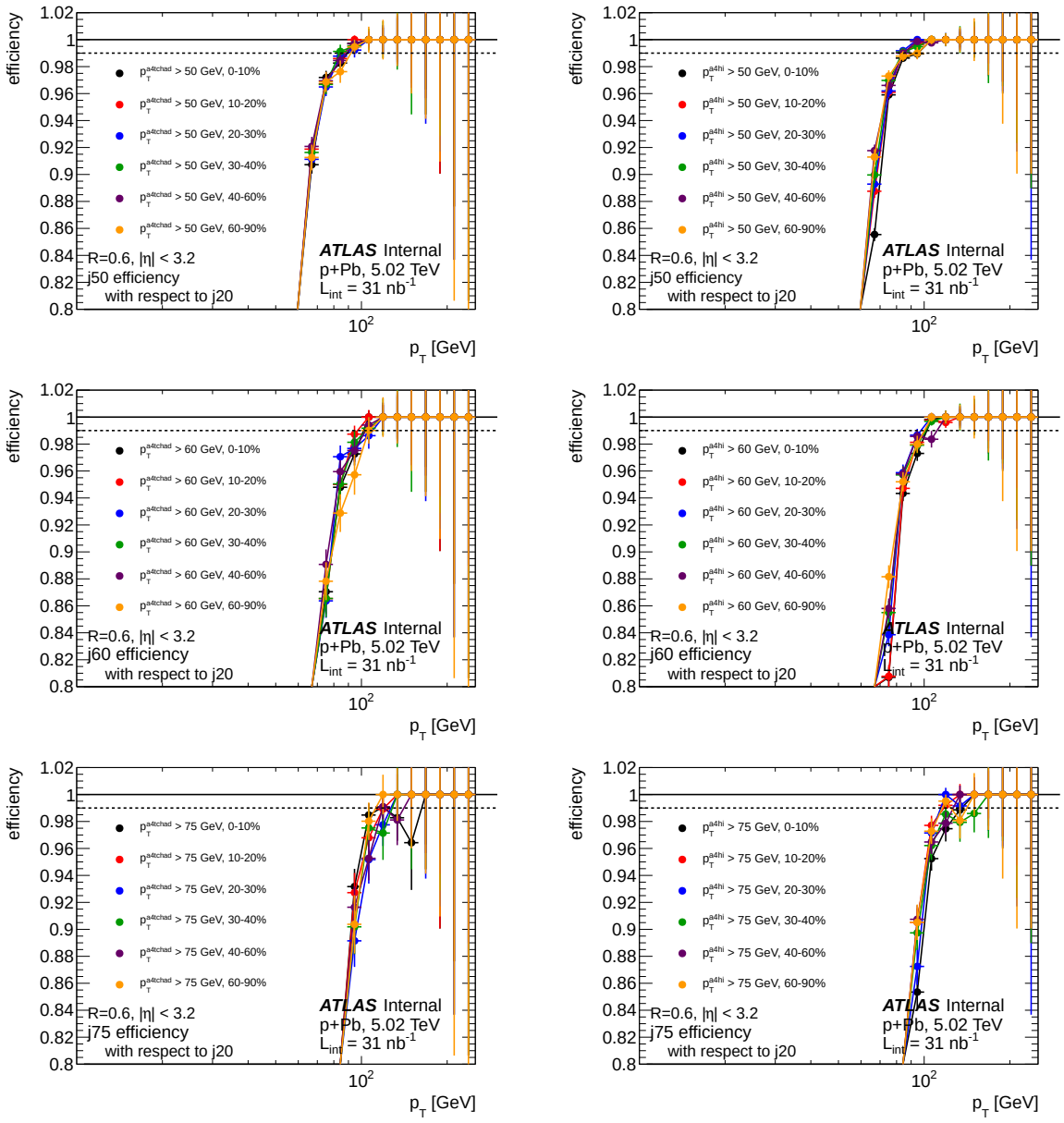


Figure B.10: Centrality dependence of the R=0.6 jet trigger efficiency, shown for j50 triggers (first plot), j60 triggers (second plot) and j75 triggers (third plot). Results for `a4tchad` (left column) and `a4hi` (right column) triggers are shown.

B.3 Ratio of p+Pb and pp uncorrected data to power law fits and MC

B.4 R_{CP} with associated Track and LCTopo jets

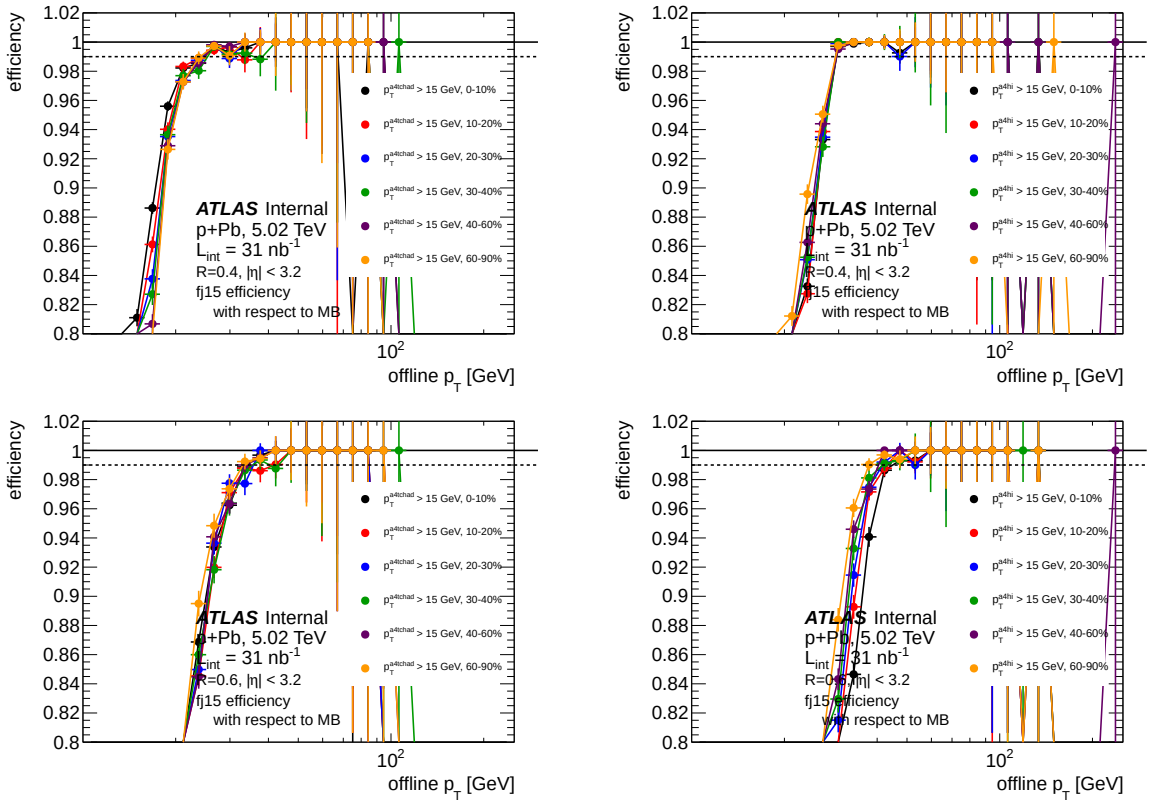


Figure B.11: Centrality dependence of the R=0.6 jet trigger efficiency, shown for fj15 triggers in R=0.4 jets (top row) and R=0.6 jets (bottom row). Results for `a4tchad` (left column) and `a4hi` (right column) triggers are shown.

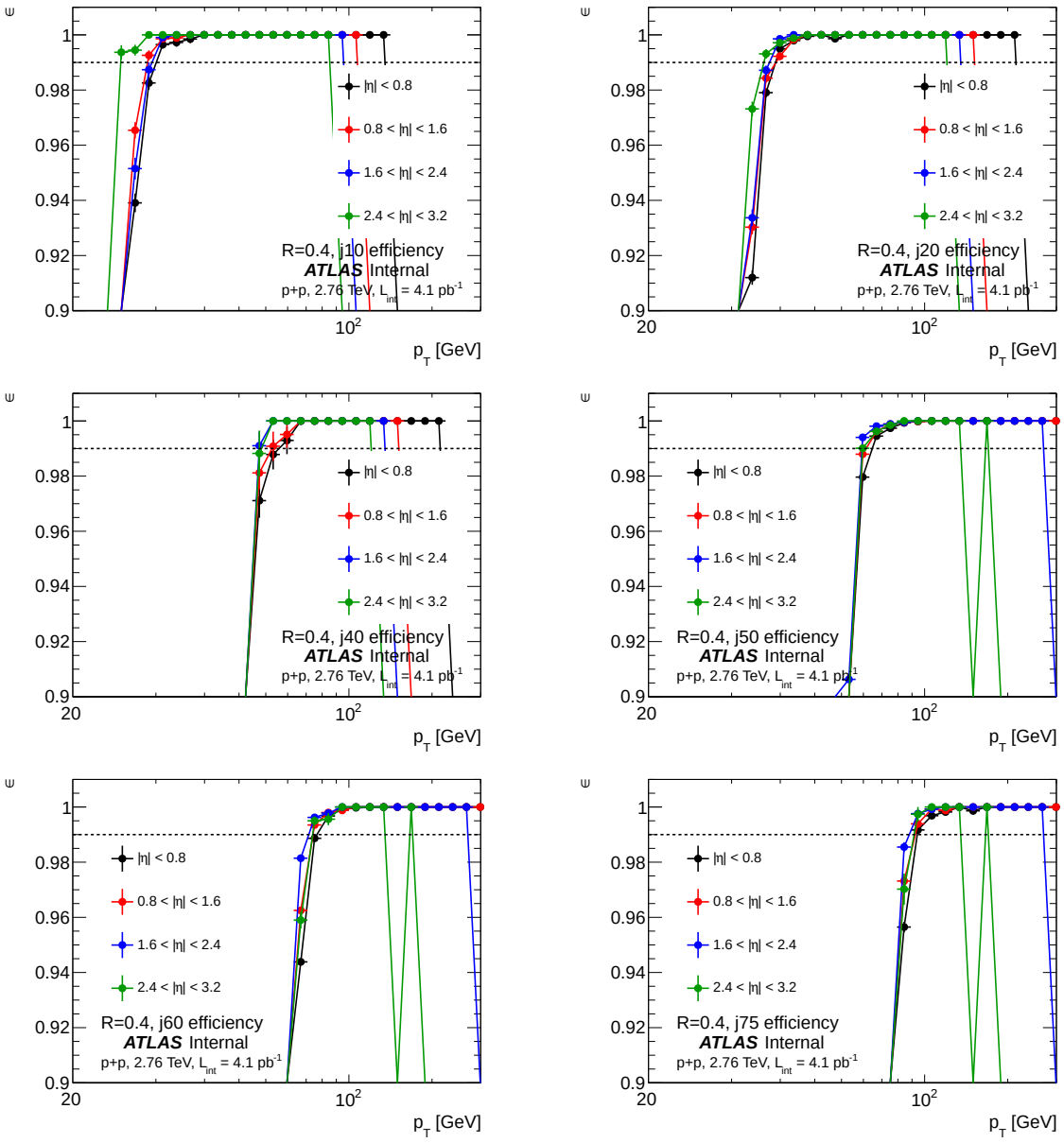


Figure B.12: η dependence of the jet trigger efficiencies in pp data, for R=0.4 jets. The efficiency in $\Delta\eta = 0.8$-sized bins is shown for j10 triggers (top left), j20 triggers (top right), j40 triggers (middle left), j50 triggers (middle right), j60 triggers (bottom left) and j75 triggers (bottom right).

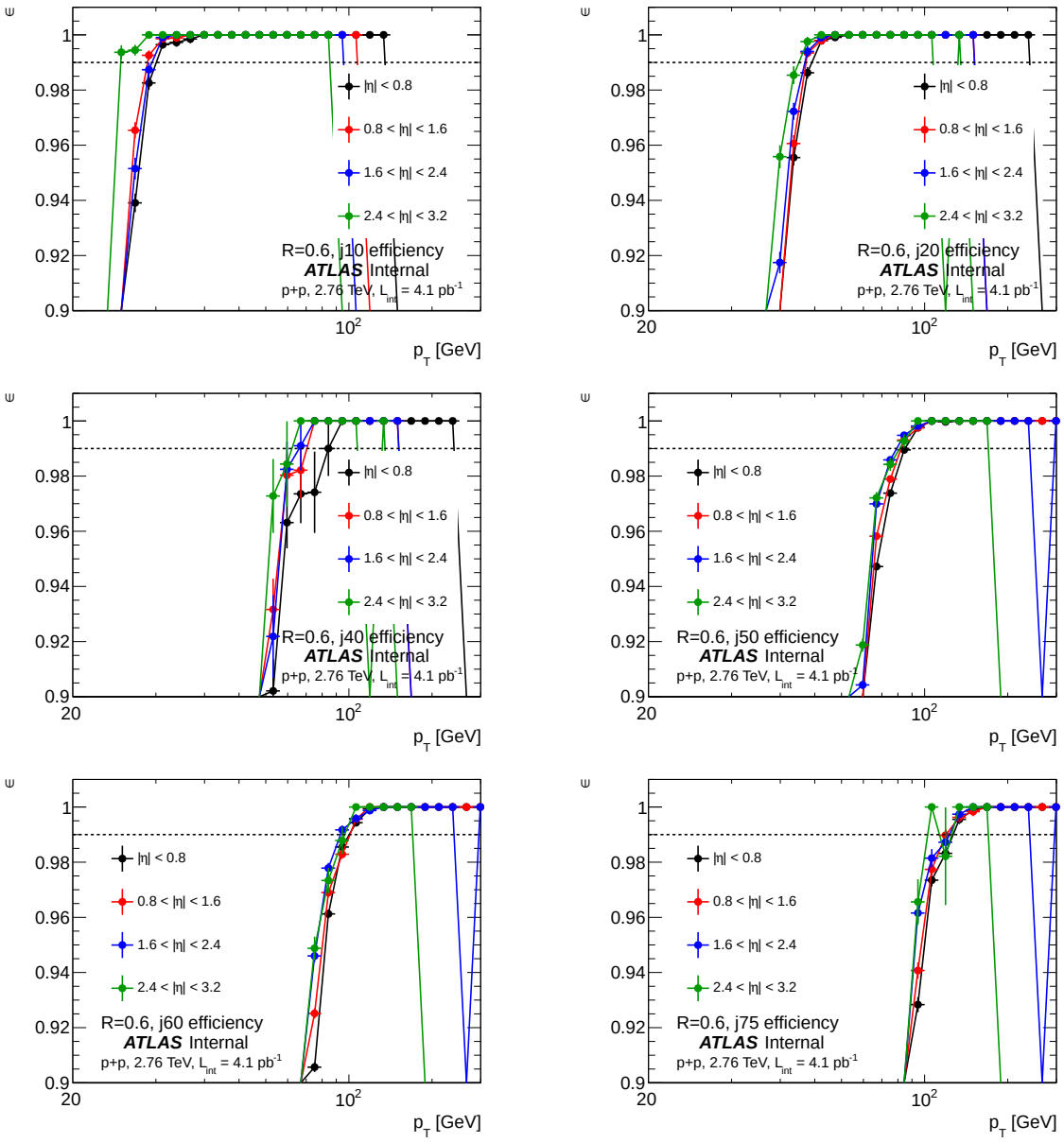


Figure B.13: η dependence of the jet trigger efficiencies in pp data, for R=0.6 jets. The efficiency in $\Delta\eta = 0.8$-sized bins is shown for j10 triggers (top left), j20 triggers (top right), j40 triggers (middle left), j50 triggers (middle right), j60 triggers (bottom left) and j75 triggers (bottom right).

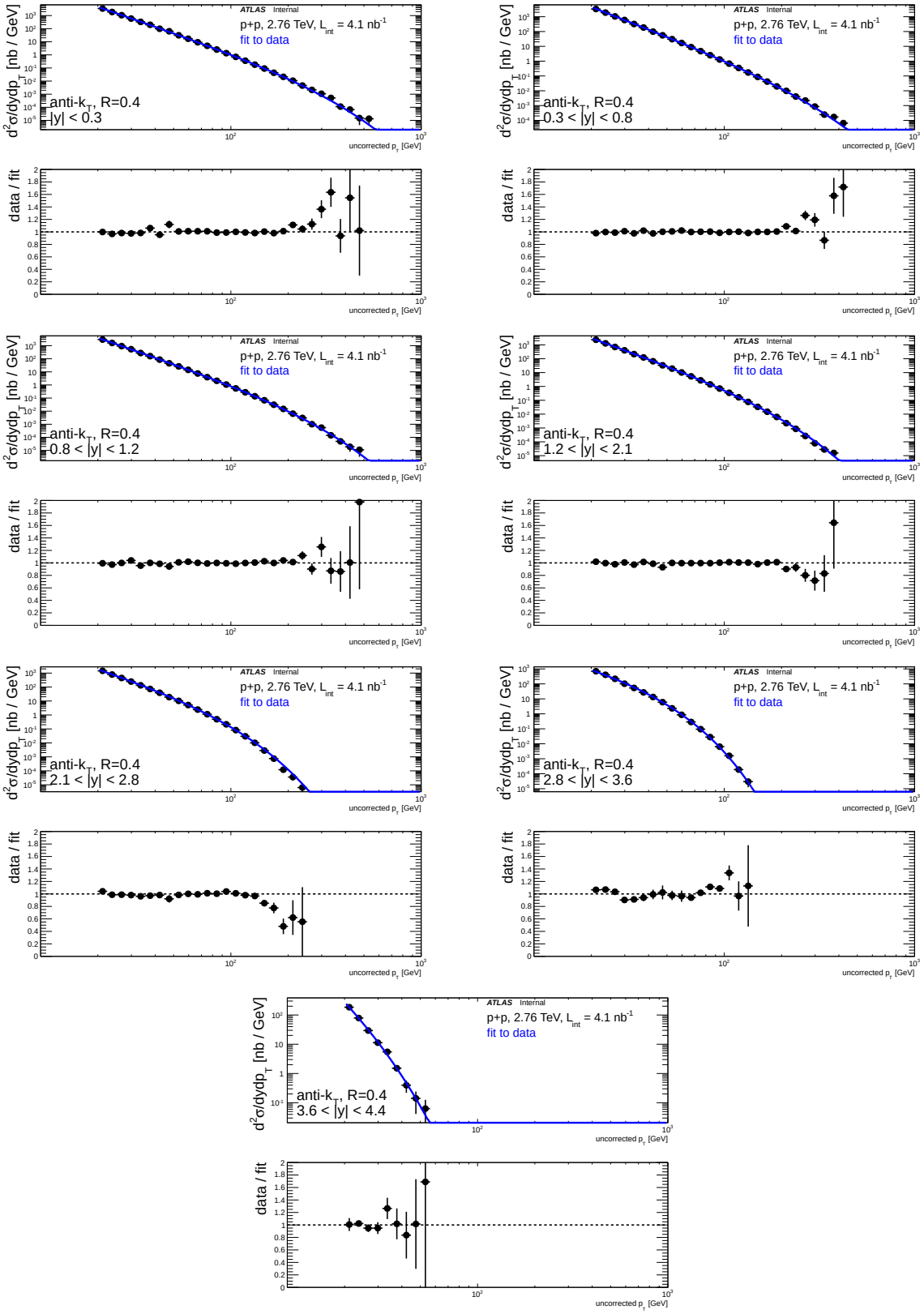


Figure B.14: Fits of the function $c_0 \cdot {p_{\mathrm{T}}}^{-c_1+c_2 p_{\mathrm{T}}}$ to pp jet spectra, for R=0.4 jets, showing the ratio of data to the fit. Each $|\eta|$ bin is shown separately.

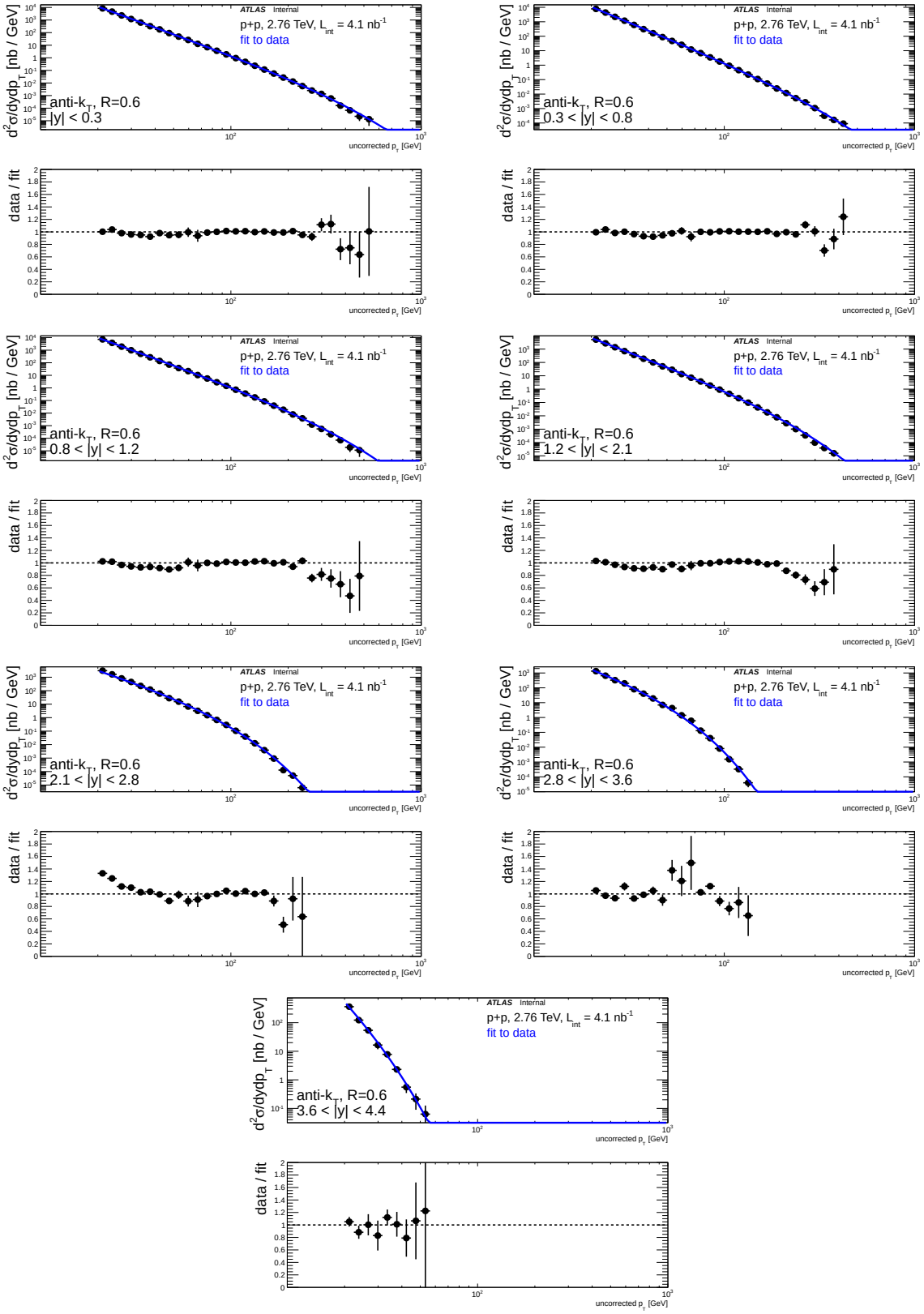


Figure B.15: Fits of the function $c_0 {p_{\mathrm{T}}}^{-c_1+c_2 p_{\mathrm{T}}}$ to pp jet spectra, for R=0.6 jets, showing the ratio of data to the fit. Each $|\eta|$ bin is shown separately.

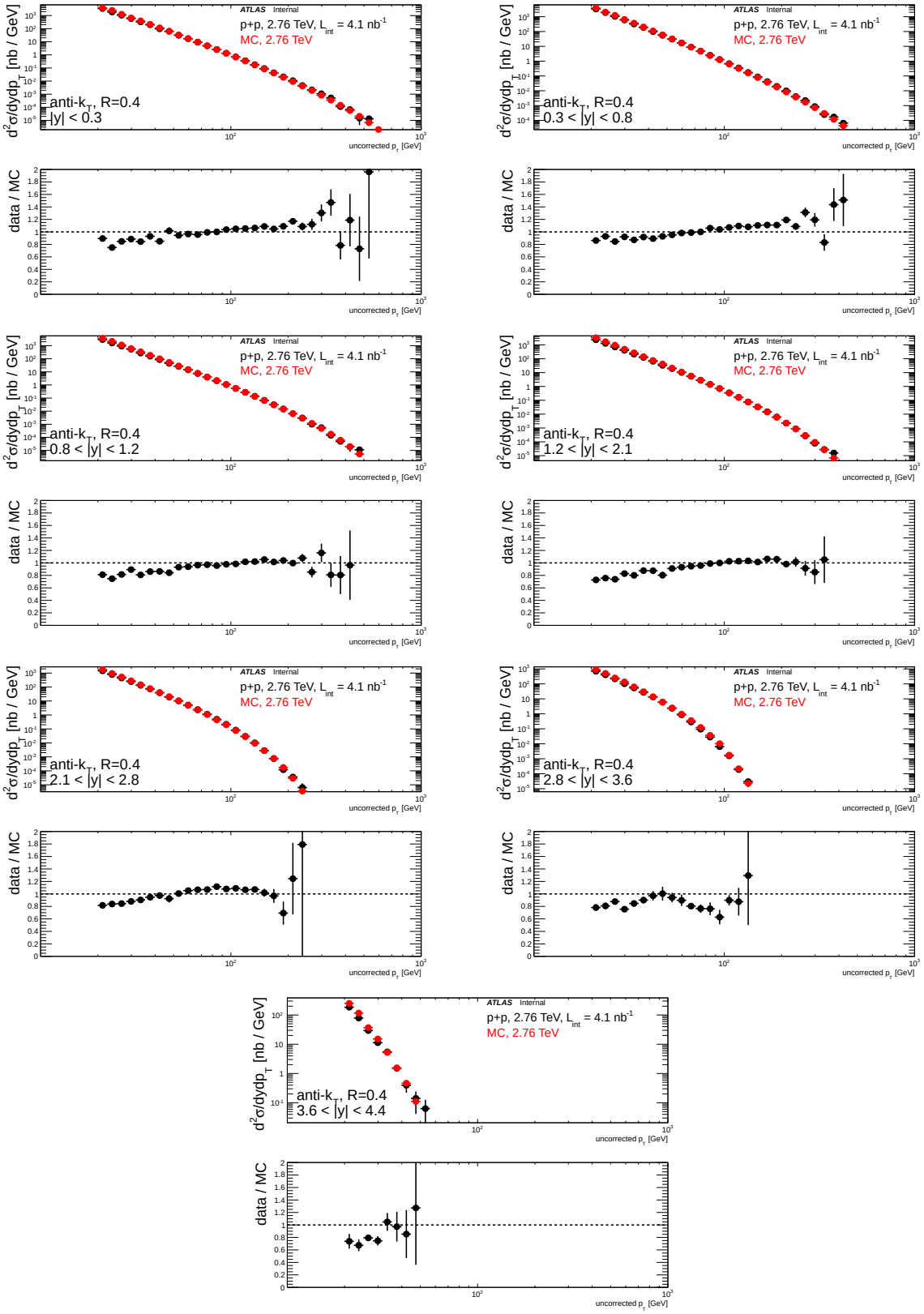


Figure B.16: Ratio of *pp* jet spectra to PYTHIA, at the uncorrected detector scale for R=0.4 jets, showing the ratio of data to MC. Each $|\eta|$ bin is shown separately.

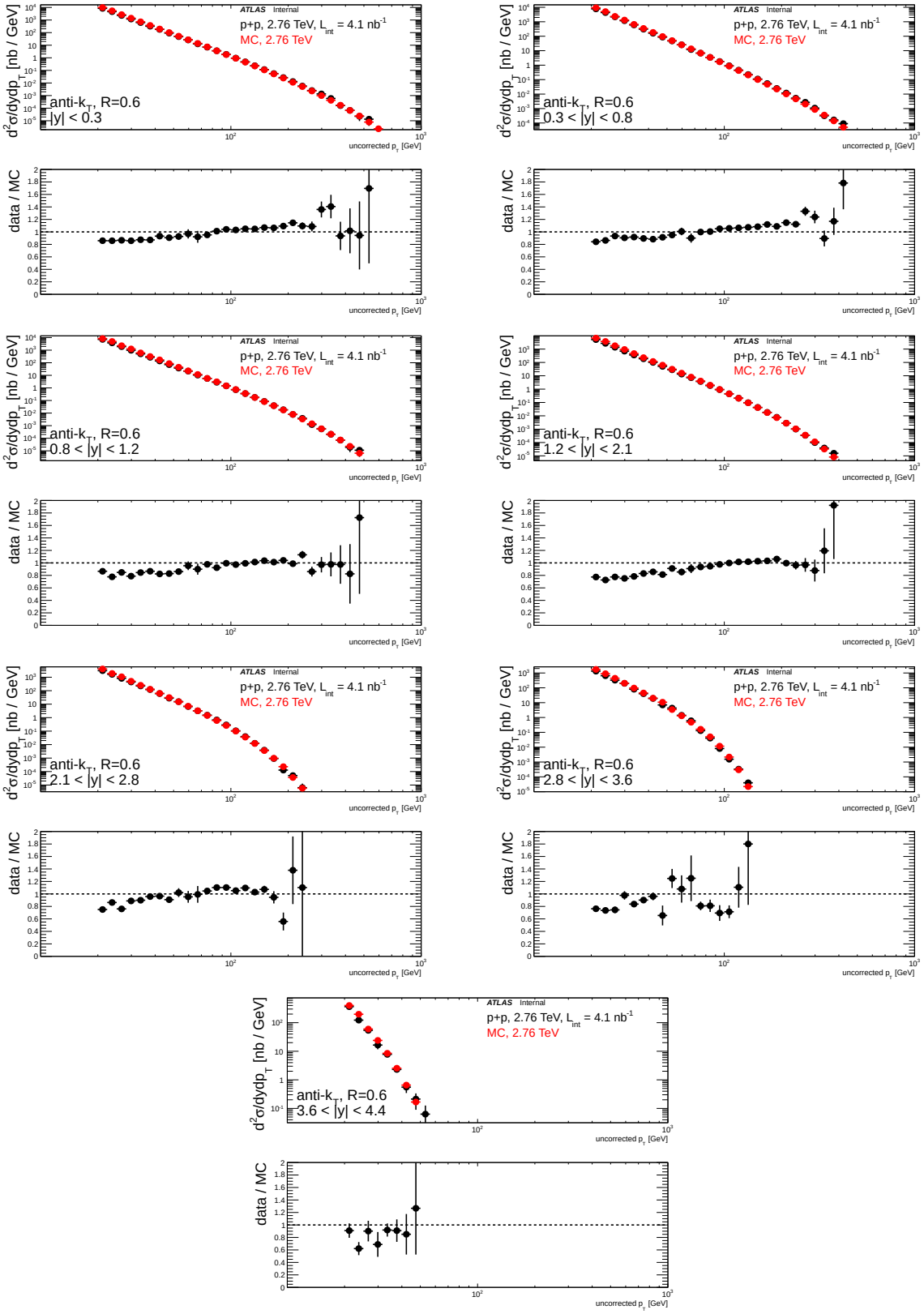


Figure B.17: Ratio of *pp* jet spectra to PYTHIA, at the uncorrected detector scale for R=0.6 jets, showing the ratio of data to MC. Each $|\eta|$ bin is shown separately.

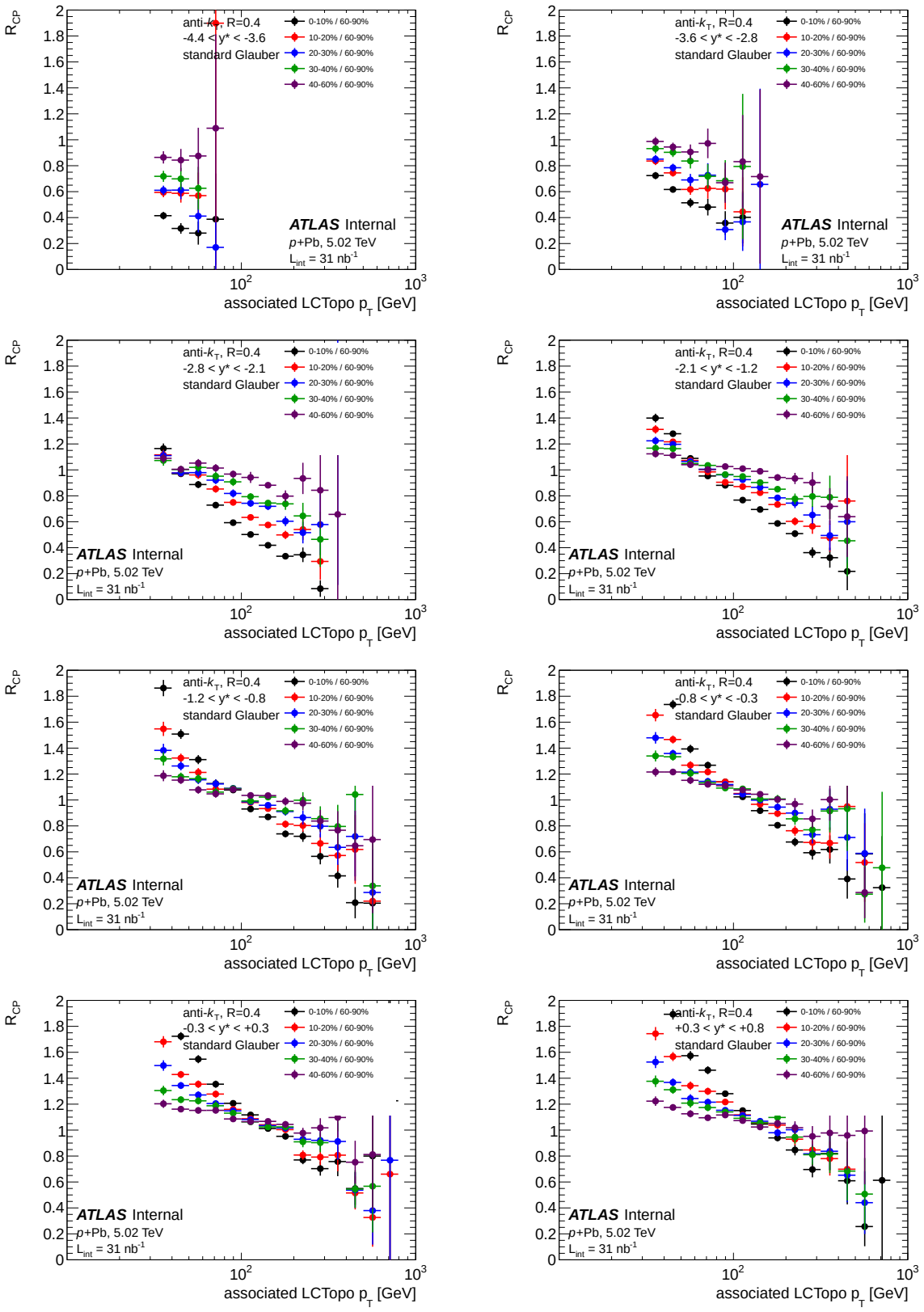


Figure B.18: R_{CP} for R=0.4 jets, but plotted as a function of the associated LCTopo jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the eight rapidity bins with $y^* < 0.8$.

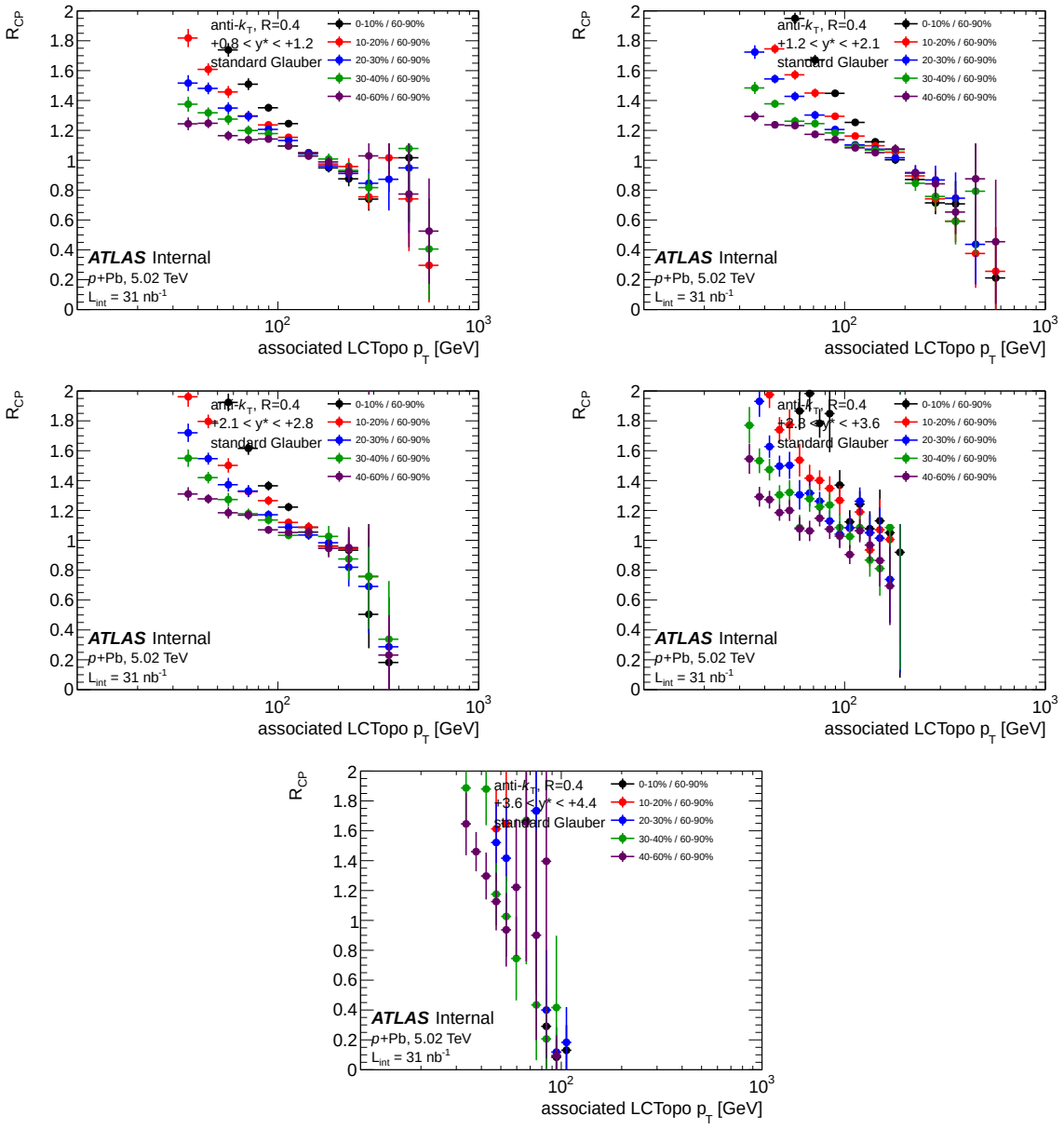


Figure B.19: R_{CP} for R=0.4 jets, but plotted as a function of the associated LCTopo jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the five rapidity bins with $y^* > 0.8$.

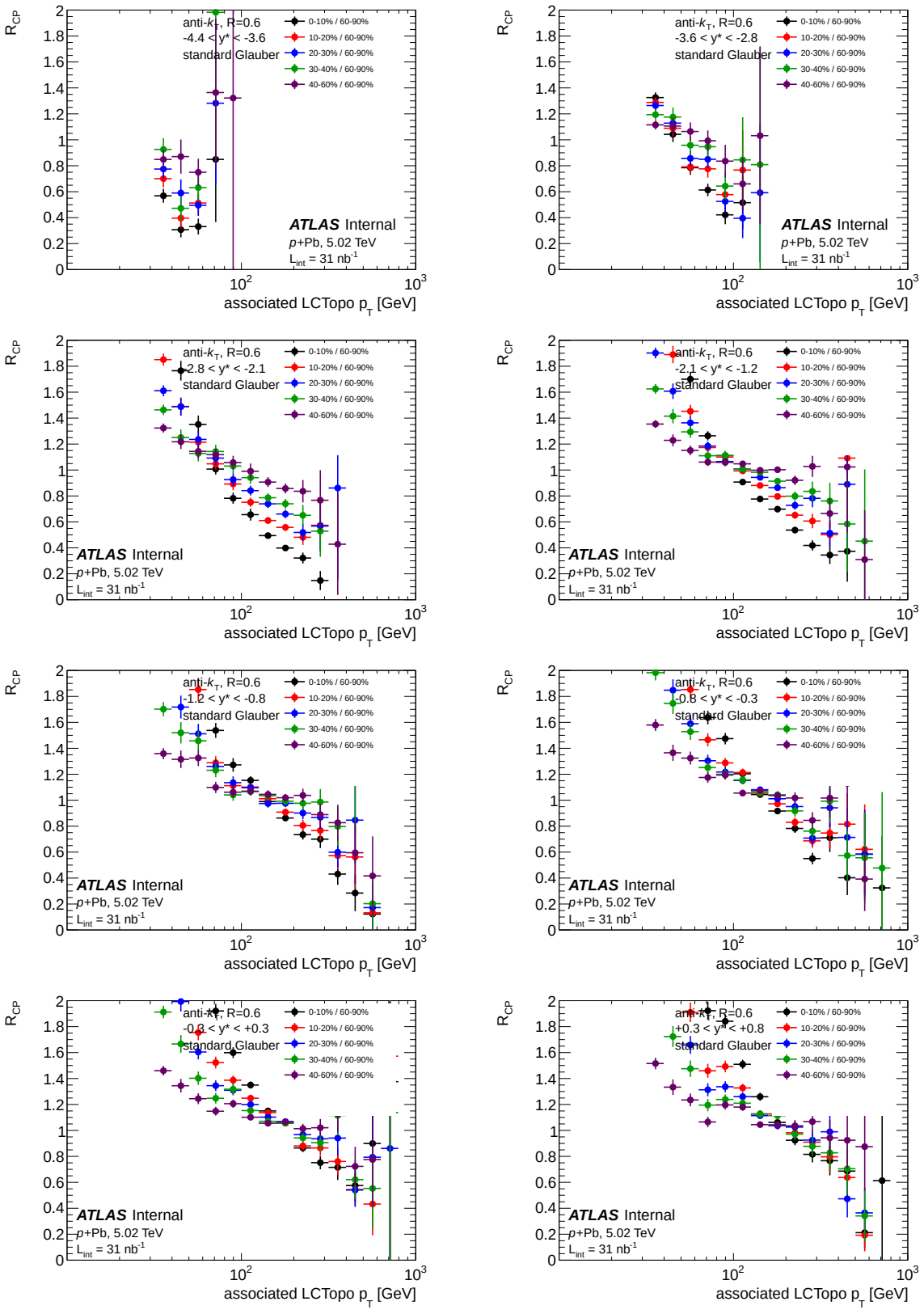


Figure B.20: R_{CP} for R=0.6 jets, but plotted as a function of the associated LCTopo jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the eight rapidity bins with $y^* < 0.8$.

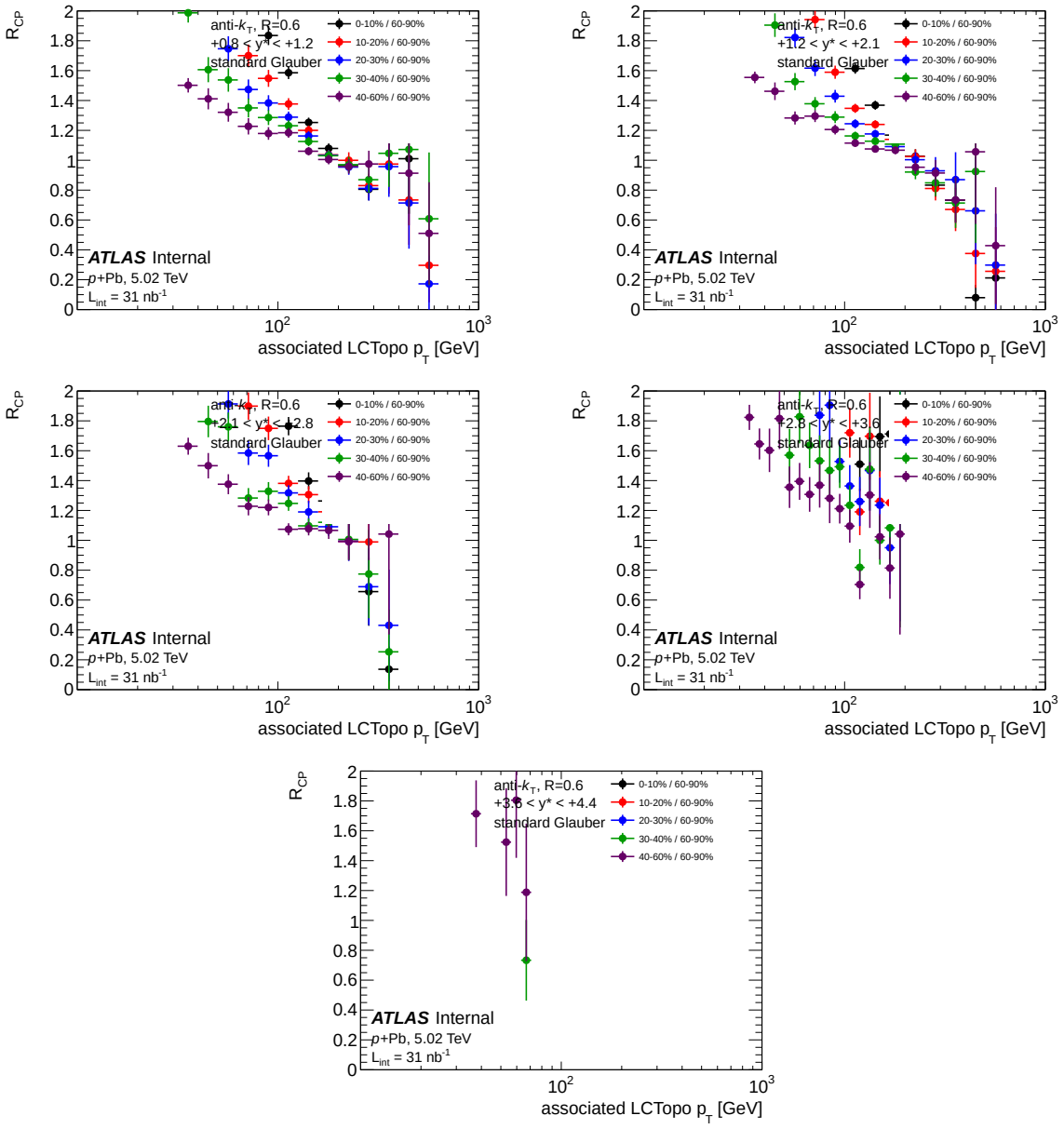


Figure B.21: R_{CP} for R=0.6 jets, but plotted as a function of the associated LCTopo jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the five rapidity bins with $y^* > 0.8$.

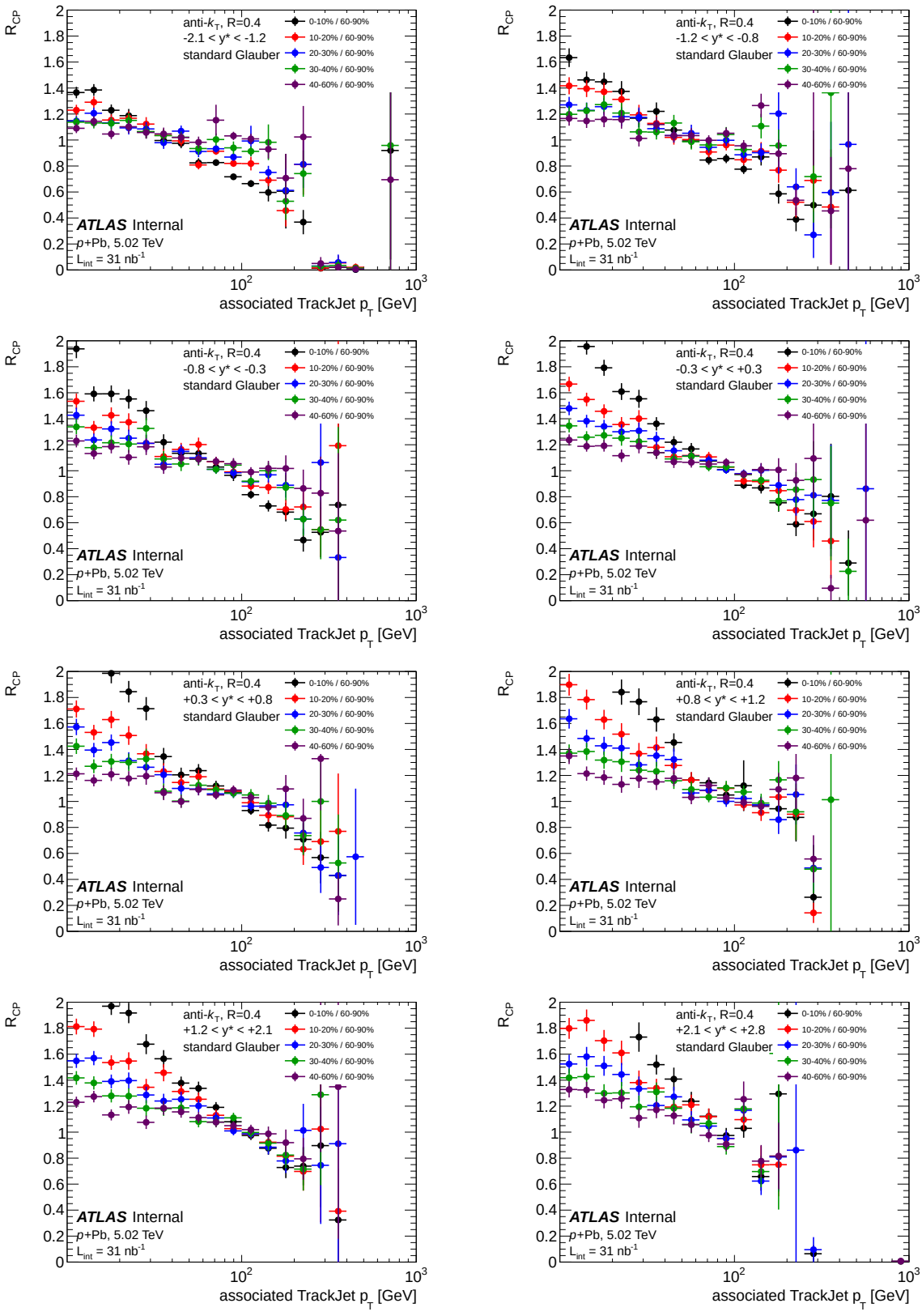


Figure B.22: R_{CP} for R=0.4 jets, but plotted as a function of the associated track jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the eight rapidity bins with tracking acceptance, $-2.1 < y^* < 2.8$.

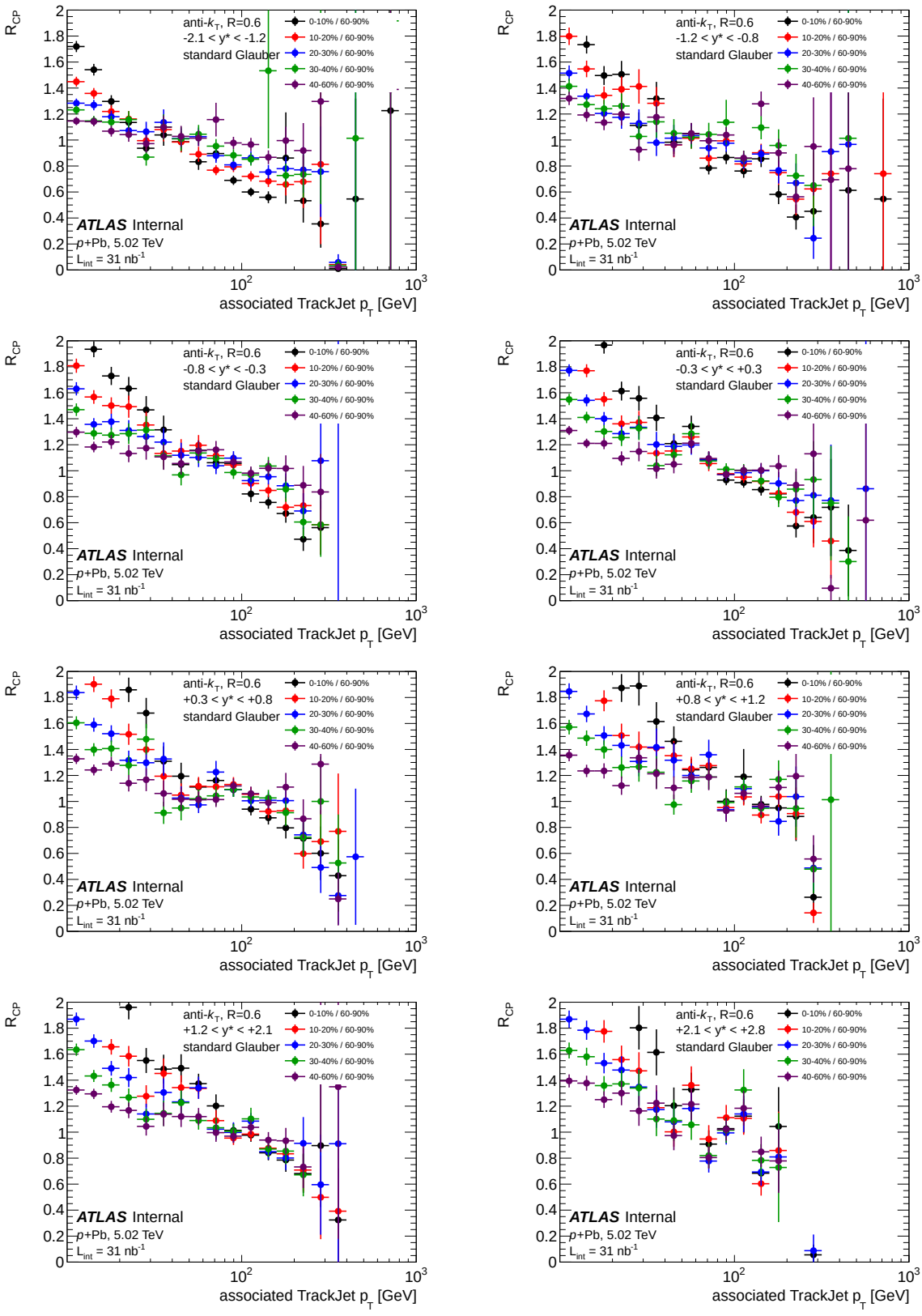


Figure B.23: R_{CP} for R=0.6 jets, but plotted as a function of the associated track jet p_{T}. The R_{CP} in this cross-check is not corrected for detector effects. Results are shown in the eight rapidity bins with tracking acceptance, $-2.1 < y^* < 2.8$.