Measurements of Charged-Particle Transverse Momentum Spectra in PbPb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV and in pPb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV with the CMS Detector

By

Eric Appelt

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

May-2014

Nashville, Tennessee

Approved:

Professor Julia Velkovska

Professor Charles Maguire

Professor Paul Sheldon

Professor Robert Weller

To Abby.

Acknowledgments

First and foremost, I would like to thank my wife, constant friend, and life-partner Rachel for her ongoing patience, support, love, and encouragement over the last five and a half years as I have worked and traveled to complete the work described in this thesis. I am indebted to my parents, Robin and Carl Appelt, for their unwavering support and belief in me, as well as their tireless efforts to teach me the values of honesty and determined effort. I am grateful for the support of my immediate family members who are always there for me, Sean Appelt and Marie Doyle, including those who became my family nearly five years ago, Barry, Margret, and Becky Buxkamper.

As an experimental physicist, I owe much of what I have learned to Julia Velkovska, who in addition to teaching by example, has opened many doors of opportunity, and has guided and mentored me for the past several years. I also owe much to Charles Maguire and Senta Greene for their mentorship and support. I would like to thank the Relativistic Heavy Ion group at Vanderbilt University for the financial support provided to me.

I would like to thank my previous mentors in years past who have helped me to develop many of the skills needed to succeed in physics. From my tenure as a high school mathematics teacher, Deborah Davies was a wonderful mentor and friend, who taught me by example how to communicate mathematics and physics. From my time at Miami University I was supported and encouraged in the pursuit of discrete mathematics by my adviser Zevi Miller, as well as my professors Daniel Pritikin and Tao Jiang.

I am grateful for the support and friendship of the current and former students and postdoctoral fellows of the Vanderbilt Relativistic Heavy Ion group with whom I have shared the excitement of discovery and traveled around the world: Ron Belmont, Michael Issah, Ravi Janjam, Pelin Kurt, Yaxian Mao, Hong Ni, Dillon Roach, Monika Sharma, Ben Snook, Shengquan Tuo, and Qiao Xu.

It has been a joy to be a part of the CMS Collaboration, and I am humbled to be a part of such an incredible effort and to work with such an intricate machine. I would like to thank a few of the people that I have worked and traveled with for their collaboration: Wit Busza, Ivan Cali, Andres Florez, Bill Gabella, Will Johns, Krisztian Krajczar, Yen-Jie Lee, Wei Li, Andrew Melo, Michael Murray, Matt Nguyen, Christof Roland, Gunther Roland, Steve Sanders, Cathrine Silvestre, Paul Sheldon, Gabor Veres, Edward Wenger, Bolek Wyslouch, and Victoria Zhukova.

I would like to recognize the Vanderbilt Advanced Computing Center for Research and Education (ACCRE), the staff of which have worked tirelessly to provide the computing resources, and who have worked with me as collaborators and friends.

The support of my personal friends has been invaluable through this journey. My close family friends, the Cohens, the Mathurs, and the Brostroms, have been there for as long as I can remember. I have enjoyed the fellowship of my covenant group, with whom I have shared my struggles and successes every other Thursday for the last eight or so years. My good friend John Bucy deserves mention for introducing me to the Linux operating system, as the many weekends spent "goofing around" with Linux have resulted in an extremely useful skill set which has been instrumental in completing this work.

TABLE OF CONTENTS

List of Tables

List of Figures

Abbreviations and Symbols

η The pseudorapidity defined as $\eta = -\ln\left(\tan\left(\theta/2\right)\right)$ where θ is the polar angle with respect to the z-axis

η/s The ratio of shear viscosity to entropy density

γ Photon

$\hat{p}_{\mathrm{T}}$ Minimum transverse component of the 4-momentum exchanged in some $2 \rightarrow 2$ simulated pQCD hard scattering process in simulation

Λ_{QCD} The intrinsic QCD energy scale which is approximately 200 MeV, and below which perturbative methods fail due to the rising coupling constant

$\mathcal{L}$ The instantaneous luminosity, or flux factor produced in a particle accelerator, from which the event rate can be determined by multiplying by the cross section as $dN_{\mathrm{proc}}/dt = \mathcal{L} \cdot \sigma$

μ Muon

μ_B Baryon chemical potential

ν_l Neutrino corresponding to lepton flavor l

ϕ The azimuthal angle, $\phi = \arctan\left(y/x\right)$, generally of a particle produced in a collision

σ The cross section, or probability of the occurrence of a given scattering process, with units of area

$\sqrt{s_{NN}}$ Center-of-mass energy per nucleon pair

$\sqrt{s}$ The center-of-mass energy of a collision system

τ Tau lepton

b Impact parameter between two colliding nuclei. May also be used as an abbreviation for barn or bottom quark, or an abbreviation for a barn, where $1\ \mathrm{b} = 10^{-24}\ \mathrm{cm}^2$

$d\ (\bar{d})$ Down quark (down antiquark)

e Electron

E_{T} The transverse energy, defined $E_{\mathrm{T}} = \sqrt{m^2 + p_{\mathrm{T}}^2}$, where m is the mass of the object

g Gluon

L_{T} The p_{T} of the reconstructed track with the highest p_{T} of any reconstructed track in a given event in some specified η range

N_{coll} Number of binary collisions between nucleons in a nucleus-nucleus or proton-nucleus collision

N_{part} Number of participating nucleons in a nucleus-nucleus collision

R Jet cone radius defined $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$

R_{AA} The nuclear modification factor for some collision between two nuclei of species A, for example Pb or Au. Note that R_{dA} may be used refer to the nuclear modification factor of deuteron-nucleus collisions, and R_{pPb} may be used to refer to the nuclear modification factor of proton-lead collisions

s ($\bar{s}$)   Strange quark (strange antiquark)

u ($\bar{u}$)   Up quark (up antiquark)

v_n nth Fourier coefficient of the azimuthal distribution of particles produced in a collision, often taken with respect to the angle which maximizes its value, and sometimes interpreted as indicative of collective motion

$W^{\pm}$ W boson

x_{T} A scaling variable defined as $x_{\mathrm{T}} = 2p_{\mathrm{T}}/\sqrt{s}$

y The rapidity of a particle or reference frame, defined $y = \tanh^{-1}\beta_z$, where $\beta_z = v_z/c$ and v_z is the velocity along the z-axis

Y_{asym} The forwards-backwards asymmetry, defined as the ratio of yields in corresponding positive and negative η ratios

Z Z boson

p_{T} The transverse momentum, $\sqrt{p_x^2 + p_y^2}$, of a particle produced in a collision where the z-axis is generally taken as the direction of motion of the incoming particles

GEANT4 GEometry ANd Tracking, version 4, a CERN software package to simulate the passage of particles through matter

AA   Generally refers to a collision between two nuclei of species A

ADC   Analog-to-digital converter

AGS   Alternating Gradient Synchrotron

ALICE A Large Ion Collider Experiment, one of the four large experiments at the LHC, and the one experiment dedicated to heavy ion physics

APV25 Read-out chip designed for the CMS silicon strip detector

ATLAS Originally stood for A Toroidal LHC ApparatuS, one of the four large LHC experiments

AVF   Adaptive Vertex Fitter algorithm

BHC   Beam Halo Counter

BNL   Brookhaven National Laboratory

BRAHMS Broad RAnge Hadron Magnetic Spectrometers, one of the large experiments at RHIC

BSC   Beam Scintillator Counter

CDF   The Collider Detector at Fermilab, a large experiment at the Tevatron collider

CERN European Organization for Nuclear Research

CGC Color Glass Condensate, an effective field theory describing QCD in the saturation regime

CKF Combinatorial Kalman Filter algorithm

CMS Compact Muon Solenoid, one of the four main LHC experiments

CMSSW Software application for the reconstruction, analysis, and simulation of CMS data

CPU Central Processing Unit

DD Double diffractive

DGLAP The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equation

DIS Deep inelastic scattering

DY Drell-Yan process

ECAL The Electromagnetic Calorimeter of the CMS detector

EDM Event Data Model, format for storing CMS collision data

EMC The European Muon Collaboration

eV Electron Volt

FED Front-End Driver, system for recording collision event data from the CMS detector systems

FF Fragmentation function

HCAL The Hadronic Calorimeter of the CMS detector

HERA Hadron Elektron Ring Anlage, a lepton-proton collider in Hamburg

HF Forward Hadron Calorimeter

HIP highly-ionizing particle, or a particle traversing the CMS silicon strip detector and leaving an anomalously large energy deposit

HLT High Level Trigger, the second stage of the CMS trigger system consisting of commercial processors

L1 Level-1 (Trigger System), the initial stage of the CMS trigger system consisting of custom programmable electronics

LEP The Large Electron Positron Collider, collider formerly occupying the LHC tunnel

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty, one of the four large LHC experiments

MC Monte Carlo, a type of computer algorithm relying on repeated random sampling to obtain numerical results

nPDF The bound nucleon parton distribution function, to be distinguished from the free proton parton distribution function, simply written as the PDF

NSD   Non-single diffractive, i.e. any hard inelastic or double diffractive collision

PDF   Parton distribution function

PHENIX Pioneering High Energy Nuclear Interaction eXperiment, one of the large experiments at RHIC

pp   Proton-proton

pQCD   Perturbative Quantum Chromodynamics

QCD   Quantum Chromodynamics

QED   Quantum Electrodynamics

QGP   Quark-gluon plasma

RAW   Format for storing CMS collision event data that is uncalibrated and not reconstructed

RECO   Format for storing reconstructed physics objects from CMS collision data

RHIC   Relativistic Heavy Ion Collider

S/N   Signal-to-noise

SD   Single diffractive

SPS   Super Proton Synchrotron

ST   A type of collision event defined by the presence of a reconstructed pixel track with $p_\mathrm{T} >$ 400 MeV/c

STAR   Solenoidal Tracker at RHIC, one of the large experiments at RHIC

TEC   Tracker Endcap, one of the CMS silicon strip detector components

TIB   Tracker Inner Barrel, one of the CMS silicon strip detector components

TID   Tracker Inner Disc, one of the CMS silicon strip detector components

TOB   Tracker Outer Barrel, one of the CMS silicon strip detector components

Trk12   A CMS high-level trigger path accepting events containing a reconstructed track with $p_\mathrm{T} \geq$ 12 GeV/c

Trk20   A CMS high-level trigger path accepting events containing a reconstructed track with $p_\mathrm{T} \geq$ 20 GeV/c

Trk30   A CMS high-level trigger path accepting events containing a reconstructed track with $p_\mathrm{T} \geq$ 30 GeV/c

UPC   Ultra-Peripheral Collision, an electromagnetic interaction between nuclei colliding with a very large impact parameter

Chapter 1

Introduction

1.1 Quantum Chromodynamics and proton-proton Collisions

1.1.1 Overview of QCD and the Standard Model

The Standard Model of Particle Physics is an enormously successful model of describing electromagnetic, weak, and strong nuclear interactions as a gauge theory consisting of fundamental spin-1/2 fermions that interact via a set of spin-1 gauge bosons, and gain an effective mass via interaction with a spin-0 boson field.

The fermions are categorized into two distinct types: leptons which do not participate in the strong interaction, and quarks which do participate. Each type of fermion consists of three generations of electroweak doublets, for a total of 12 different fermion species, or six flavors of quarks and leptons. The properties of the 12 fermions are summarized in Table 1.1.

Charged fermions interact with the familiar massless spin-1 photon field, and the left-handed chirality component of the fermion fields interact with the massive spin-1 $W^{\pm}$ and Z bosons. The electromagnetic and weak interactions are theoretically unified into a common framework by the Higgs Mechanism, in which the spin-0 Higgs boson field generates interaction terms in the Standard Model Lagrangian responsible for the effective masses of the $W^{\pm}$ and Z bosons as well as the fermions. For a basic introduction and description of the Higgs Mechanism and the electroweak sector of the Standard Model see, for example, Ref. [2–5].

The strong interaction is described by the theory of Quantum Chromodynamics (QCD), in which the non-abelian gauge symmetry of the theory is an invariance under the transformations of the group of 3x3 unitary matrices, SU(3). This requires that for each flavor of quark there is a triplet of fields, and that there are 8 different gauge boson fields, called gluons, which mediate the strong interaction between the quarks. The properties of the bosons are given in Table 1.2. Although the focus of this chapter concerns the physics of the QCD sector of the Standard Model, particularly at high temperature, neither a complete mathematical description nor a history of the success of the theory will be presented here. Instead, a few key features of the theory will be briefly described. The basic mathematical structure of the QCD Lagrangian is presented in Appendix A, and some of the early experimental evidence for QCD in the context of deep inelastic scattering (DIS) experiments is described in Appendix B. For a thorough description of the theory see, for example, Ref. [6, 7]. For a lively account of the history and eventual acceptance of QCD, see Ref. [8].

One interesting property of QCD, and non-abelian gauge theories in general, is known as asymptotic freedom. The coupling constant determining the strength of the interactions between color charged particles is found to become asymptotically weaker at large energy or small distance scales [9]. This makes the theory amenable to perturbative approximation in the same manner as Quantum

Fermion	Symbol	Charge	Mass (GeV/c^2)
up quark	u	+2/3	$2.3\,^{+0.7}_{-0.5} \times 10^{-3}$
down quark	d	-1/3	$4.8\,^{+0.5}_{-0.3} \times 10^{-3}$
strange quark	s	-1/3	$95 \pm 5 \times 10^{-3}$
charm quark	c	+2/3	1.275 ± 0.025
bottom quark	b	-1/3	$4.18 \pm 0.03\ (\overline{\text{MS}})$
top quark	t	+2/3	$173.07 \pm 0.52 \pm 0.72$
electron	e	-1	0.510999×10^{-3}
electron neutrino	ν_e	0	$< 2 \times 10^{-6}$
muon	μ	-1	105.658×10^{-3}
muon neutrino	ν_μ	0	$< 0.19 \times 10^{-3}$
tau	τ	-1	1.77682 ± 0.00016
tau neutrino	ν_τ	0	$< 18.2 \times 10^{-3}$

Table 1.1: Basic properties of the elementary fermions. Values for the particle masses are taken from the Particle Data Group combined estimates [1].

Boson	Symbol	Charge	Mass (GeV/c^2)
photon	γ	0	0
$W^\pm$	$W^\pm$	± 1	80.385 ± 0.015
Z	Z	0	91.1876 ± 0.0021
gluon	g	0	0
Higgs	H	0	125.9 ± 0.4

Table 1.2: Basic properties of the elementary bosons. Values for the particle masses are taken from the Particle Data Group combined estimates [1].

Electrodynamics (QED) at high-energy, but due to the growth of the coupling constant, perturbative calculations fail at low energy.

Quarks and gluons are also not observed as free, asymptotic states. Instead, they are confined to complicated color-singlet bound states known as hadrons, which may take the form of mesons with the quantum numbers of a quark and antiquark, or (anti-)baryons with the quantum numbers of three (anti-)quarks. In this manner the myriad hadrons discovered in the 1950s and 1960s can be classified by these quantum numbers, referred to as their valence quark content. These are, for example, π^+ ($u\bar{d}$), K^- ($s\bar{u}$), n (udd), Ω^- (sss), Λ_c^+ (udc), and so forth. This confinement property cannot be understood in terms of perturbation theory, which predicts the production of quarks and gluons in every order [10]. As of this writing, there is no fully analytic proof explaining the emergence of confinement from first-principles in QCD.

QCD is largely an unsolved problem. High energy scattering processes may be partially calculated using perturbative methods, as described in Sec. 1.1.3, but many phenomenological features of the theory remain elusive and are inaccessible to perturbative calculations. An alternative approximation method, lattice gauge theory, works by reformulating QCD on a discrete space-time lattice, with lattice points separated by spacings of length a. The quarks reside on the lattice points, and the gluons on the links between them. This approximation allows a controlled error but requires a large amount of computational resources. The lattice QCD method has been successful in predicting the spectrum of hadron masses from first principles [11], and as described in Sec. 1.2.1, predicting a phase transition of QCD matter at high-temperature.

Experimentally, the interactions of quarks and gluons in the perturbative regime may be explored by producing high energy collisions of particles such as electrons or protons.

1.1.2 Generalities of pp Collisions

A collision between two protons is typically described as occurring at the origin of a coordinate system with the incoming protons moving along the z-axis. The momentum of outgoing particles is generally described in terms of the transverse momentum, $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$, the azimuthal angle of the particle trajectory, $\phi = \arctan{(y/x)}$, and the pseudorapidity, $\eta = -\ln{(\tan{(\theta/2)})}$, where θ is the polar angle measured from the positive z-axis. The choice of variable η is taken for its similarity to rapidity,

$$y = \tanh^{-1}\beta_z = \frac{1}{2}\ln\left(\frac{1 + \beta\cos(\theta)}{1 - \beta\cos(\theta)}\right) = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{1.1}$$

where β_z the longitudinal velocity of the proton.[1] For comparison,

$$\eta = -\ln{(\tan{(\theta/2)})} = \frac{1}{2}\ln\left(\frac{1 + \cos(\theta)}{1 - \cos(\theta)}\right) = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right). \tag{1.2}$$

Thus, for a particle with large p_{T} compared to its mass, $y \approx \eta$.

---

[1]Note that in this document, unless otherwise noted, natural units are employed with $\hbar = c = 1$, and factors of c may be suppressed in equations as appropriate.

The energy of the colliding pair is described in terms of the square root of the Mandelstam variable $s = (p_1 + p_2)^2$, where p_1 and p_2 are the 4-momenta of the incoming protons, and $\sqrt{s}$ is then the center-of-mass energy of the colliding pair [12].

The probability of the occurrence of a given scattering process is given by the cross section, σ, which has units of area, often given in "barns" where $1\ \mathrm{b} = 10^{-24}\ \mathrm{cm}^2$. From the specifics of the collider (number of colliding particle bunches, beam intensity, geometry of the bunch profile) one obtains the flux factor, called luminosity $\mathcal{L}$, in units of $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The rate at which a scattering process occurs is then given as

$$\frac{dN_{\mathrm{proc}}}{dt} = \mathcal{L} \cdot \sigma, \qquad \text{or} \qquad N_{\mathrm{proc}} = \mathcal{L}_{\mathrm{int}} \cdot \sigma \tag{1.3}$$

where N_{proc} is the total number of expected processes of the specified type, and $\mathcal{L}_{\mathrm{int}}$ is called the integrated luminosity taken over some finite time span.

For a kinematically differential process, such as the production of a charged hadron at some particular momentum p, the corresponding differential cross section $d^3\sigma_{\mathrm{had}}/dp^3$ is dependent on the frame of reference. The effect of the Lorentz transformation on $d^3\sigma_{\mathrm{had}}/dp^3$ is opposite to that of the energy, E, and so one can construct a Lorentz invariant cross section as

$$E\frac{d^3\sigma_{\mathrm{had}}}{dp^3} = \frac{d^3\sigma_{\mathrm{had}}}{p_{\mathrm{T}}dp_{\mathrm{T}}dyd\phi} = \frac{d^2\sigma'_{\mathrm{had}}}{2\pi p_{\mathrm{T}}dp_{\mathrm{T}}dy} \approx \frac{d^2\sigma'_{\mathrm{had}}}{2\pi p_{\mathrm{T}}dp_{\mathrm{T}}d\eta} \tag{1.4}$$

where σ'_{had} is the cross section averaged over the azimuthal angle, although assuming that the colliding particles are unpolarized, σ_{had} should be ϕ-independent. Note that the approximation replacing dy with $d\eta$ is only valid for high-p_{T}, and integrating the differential cross section over p_{T} this would yield either $d\sigma'_{\mathrm{had}}/dy$ or $d\sigma'_{\mathrm{had}}/d\eta$ which are different quantities. Lorentz invariant charged-particle yields per collision may be given in the same manner.

Proton-proton (pp) collision processes may be broadly categorized into elastic scatterings, where the protons simply exchange momentum, and inelastic scatterings, where additional processes occur. Perturbative QCD (pQCD) methods may describe the cross sections of inelastic processes in which large momentum transfers take place, as described in Sec. 1.1.3. However, many features of inelastic collisions and even the total inelastic cross section cannot be calculated using pQCD.

Around 25% of all inelastic pp collisions at high energy undergo what is called a diffractive process where a large rapidity gap exists between particles produced in the collision. These processes have not been described in the framework of pQCD, but have been successfully predicted in the effective Regge theory [13], where the interaction is mediated by a color-singlet exchange carrying vacuum quantum numbers, generally called a Pomeron. Diffractive processes are largely classified into single-diffractive (SD) collisions, where one proton dissociates into multiple particles and the other remains intact; and double-diffractive (DD) collisions, where both protons disassociate. The term non-single diffractive (NSD) is often used to describe inelastic collisions which are either non-diffractive or double-diffractive. In many cases, the measurement of the total NSD cross section for some collision energy $\sqrt{s}$ may be more accurately measured than the total inelastic cross section,

depending on the capabilities of the experiment.

1.1.3 High-p_T Charged-Particle Production in pp Collisions

At high-p_T, charged-particle production in pp collisions is dominated by jets of hadrons which are collimated to a kinematic cone of some radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. The challenge is then to understand this collision system in terms of the scattering matrix of perturbative QCD, which is written in terms of incoming and outgoing asymptotic quark and gluon states, not in incoming protons and outgoing jets of hadrons. Several factorization theorems [14] make the connection possible between the quarks and gluons of pQCD, and the experimentally observed jets and hadrons.

This factorization process may be pictured as a separation of scales. In the incoming protons, the partons can be envisioned as effectively frozen, as time dilation renders the time-scale of their interactions in the center-of-mass frame of the collision extremely long. Thus a colliding parton may be viewed as an approximately asymptotic state. Partons from each proton then undergo a "hard" interaction on a large momentum scale which can be calculated in pQCD. The outgoing partons then hadronize into a jet of hadrons, but do so on a long time scale compared to the hard interaction. This allows the outgoing partons to also be approximated as asymptotic states, justifying the pQCD approximation. This picture, although perhaps naively compelling, should be taken with caution as factorization theorems are formally proved in only e^+e^- annihilation, DIS, and Drell-Yan processes, and not for general hadroproduction in hadron-hadron collisions. Counterexamples to factorization exist in some cases [15].

In the case that it holds, the factorization approach allows one to write the invariant cross section of the production of some hadron C with sufficiently large momentum p_C in some collision system $A + B \rightarrow C + X$ as [16]:

$$
\begin{aligned}
E_C \frac{d\sigma}{d^3 p_C} = \sum_{abc} \int dx_A dx_B \frac{dz}{z} \quad & f_{a/A}(x_A, Q_F^2) f_{b/B}(x_B, Q_F^2) \\
& \times |\vec{k}_c| \frac{d\hat\sigma}{d^3 k_c}(p_c/z\sqrt{s}, Q_F^2) \\
& \times D_{C/c}(z, Q^2)
\end{aligned}
\tag{1.5}
$$

Here a and b are partons carrying a fraction x_A or x_B of the total longitudinal momentum of proton A or B respectively. The factorization scale, Q_F is dependent on the details of the calculation and the approximations used, but in general must be well above $\Lambda_{\text{QCD}} \sim 200\,\text{MeV}$ for the pQCD component of the calculation to be valid. The factor $|\vec{k}_c|\, d\hat\sigma/d^3 k_c$ represents the partonic cross section $a + b \rightarrow c + x$, where x represents one or more partons, and $\hat\sigma$ is calculated in pQCD. The sum runs over all possible quark and gluon states.

The factor $f_{a/A}(x_A, Q_F^2)$ is called the parton distribution function (PDF) and is often interpreted as the probability of finding a quark or gluon of species a carrying some fraction x_A of the momentum of proton A when viewed at an energy scale of Q_F. As noted in Ref. [17], such an interpretation is

not fully valid, as the PDF is affected in part by interference effects in final state interactions. The PDFs cannot be determined from pQCD, and must be approximated using experimental data. Using results from fixed target DIS experiments, the HERA electron-proton collider, and the Tevatron pp̄ collider, several PDFs produced from global fits to data, such as CT10 [18], MSTW [19], and NNPDF2.0 [20]. An example of the MSTW calculation at two values of Q^2 for various partonic species can be seen in Fig. 1.1. Appendix B gives a basic description of DIS and the formal definition of the x and Q^2 variables in this context.

MSTW 2008 NLO PDFs (68% C.L.)
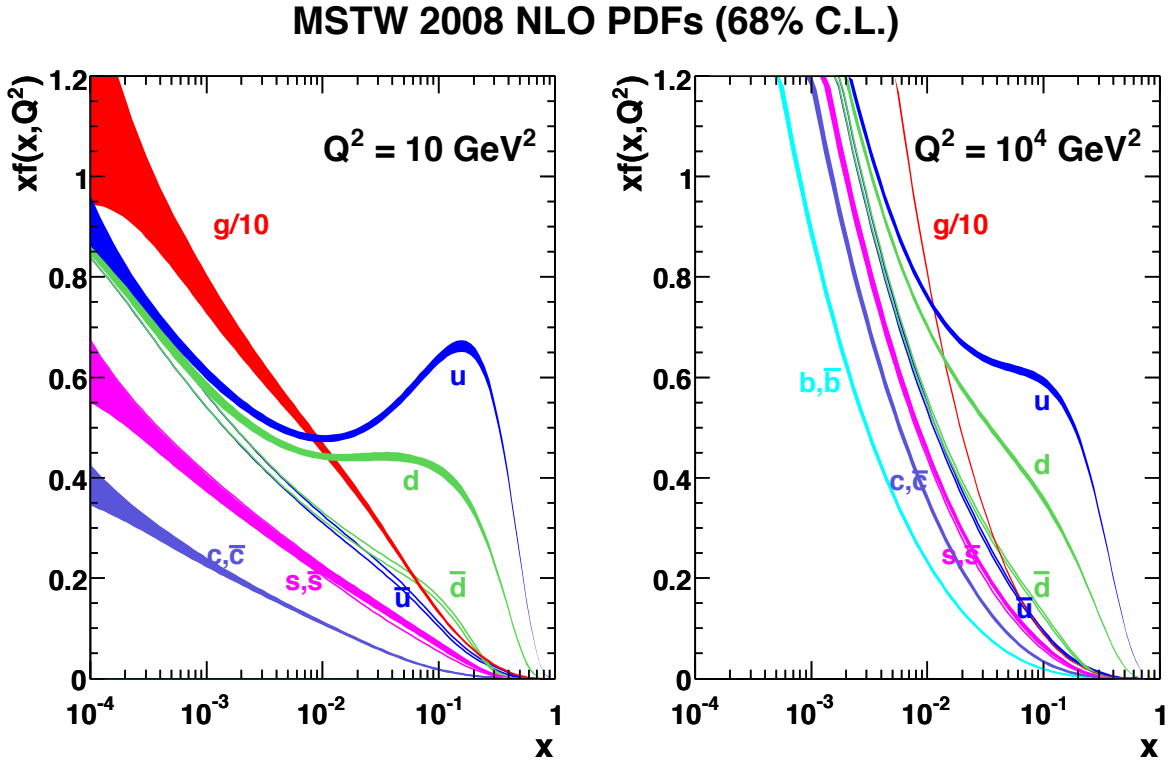
Figure 1.1: Free proton PDFs for gluons, quarks and antiquarks as determined by MSTW NLO 2008 [19]. Note that at smaller Q^2, the momentum of the proton is found mostly in the valence (u,d) quarks, whereas at higher Q^2 a larger number of small-x gluons and sea quarks/anti-quarks are predicted.

Although the PDF, $f_{a/A}(x_A, Q^2)$, cannot be calculated perturbatively, the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equation [21–23] provides a perturbative method for calculating $f_{a/A}(x_A, Q^2)$ at some new value of Q^2 with $Q > \Lambda_{\text{QCD}} \approx 200\,\text{MeV}$ given a known value of $f_{a/A}(x_A, Q_0^2)$ at a fixed value of x_A.

Finally, the factor $D_{C/c}(z, Q_F^2)$ is called the fragmentation function (FF) and represents the probability of the parton c hadronizing into the hadron C carrying a fraction z of its total momentum, at a factorization scale Q_F. As with the PDFs, the FFs cannot be calculated using known pQCD approximations or other methods from first principles, and must be inferred from experimental data. The available FF parametrizations, such as KKP [24], AKK [25], and fDSS [26], typically extend to pions, kaons, and protons, and are drawn primarily from e^+e^- collision data at the LEP

accelerator with additional contributions from $p\bar{p}$ collision data at the Tevatron and pp collision data at the Relativistic Heavy Ion Collider (RHIC).
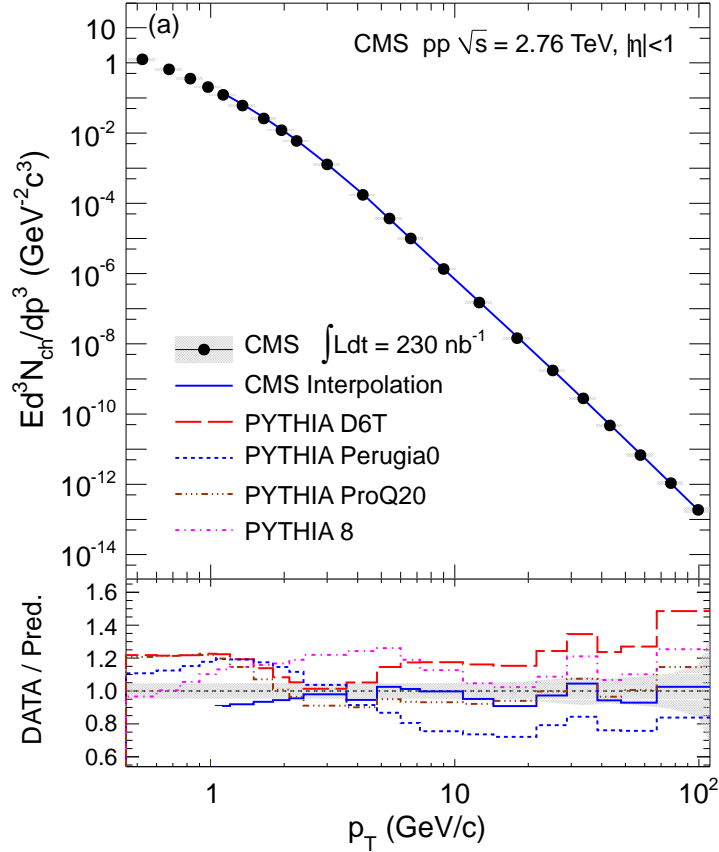
Figure 1.2: (Top) Invariant charged-particle yield per NSD event in pp collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ for $|\eta| < 1.0$ measured with the CMS detector. These data are compared with the predictions of four tunes of the PYTHIA MC generator and with the CMS interpolated spectrum using data at 0.9 and 7 TeV. (Bottom) the ratio of the measured spectrum to the predictions of the four PYTHIA tunes and to the interpolated spectrum. The gray band corresponds to the statistical and systematic uncertainties of the data added in quadrature [27].

The charged-particle invariant yield at midrapidity per NSD pp collision at $\sqrt{s} = 2.76\,\text{TeV}$ [27] is shown in Fig. 1.2. Note the steeply falling shape of the spectrum and the relative rarity of events producing charged hadrons with $p_T > 10\,\text{GeV}/c$. The theoretical predictions from various tunes of the PYTHIA MC generator, described in Appendix C, are compared with these data and are accurate to better than 50% even as the spectrum traverses over 13 orders of magnitude. The PYTHIA generator incorporates PDF sets similar to those described here, perturbative cross section calculations, and handles fragmentation using an effective model. In order to reproduce the low-p_T portion of the spectrum, additional non-perturbative processes are incorporated which are not described above.

1.2 Finite Temperature QCD and Heavy Ion Collisions

1.2.1 The QCD Phase Diagram and the Quark-Gluon Plasma

Suggestions of an exotic phase of bulk nuclear matter consisting of free quarks at some very high temperature or density were discussed early in the history of the QCD model, and were first published by Collins and Perry, who predicted a "quark soup" at very high density behaving as a free-quark gas [28]. Shuryak is credited for coining the term Quark-Gluon Plasma (QGP) to describe this potentially plasma-like phase that could exist at high temperature or density [29].

More recently, lattice QCD calculations have predicted a rapid rise in energy density and pressure in a QCD medium for temperatures in the range of $T \sim 100\text{-}300$ MeV, giving a rigorous theoretical foundation for the hypothesis of deconfinement at high temperature. In addition, lattice QCD has provided predictions for the equation of state of hot nuclear matter. As shown in Figure 1.3, recent lattice calculations favor a deconfinement region at a temperature of $T \sim 185\text{-}195$ MeV [30].
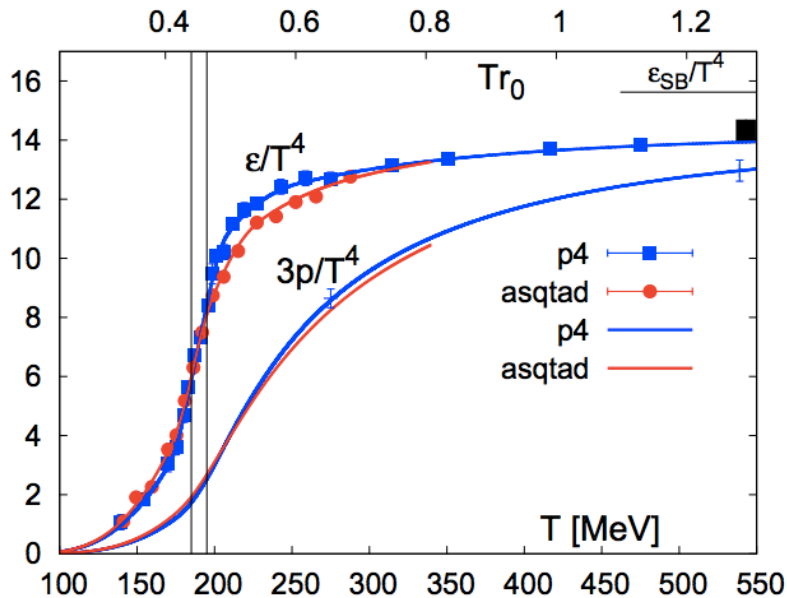
Figure 1.3: Lattice QCD calculation of energy density and pressure scaled by the fourth power of the temperature using two fermion action approximation techniques. On the top horizontal axis is the temperature multiplied by the length scale $r_0 = 0.469(7)$ fm. The vertical band indicates a transition region of $T \sim 185\text{-}195$ MeV, where the sudden increase in energy density represents an increase in the effective number of degrees of freedom [30].

Along with the idea of a phase transition it is natural to discuss a potential phase diagram. This can be constructed with the temperature of the medium taken as one axis. From the fact that baryon number is conserved in QCD interactions, a baryon chemical potential, μ_B may be assigned to the other axis. Unfortunately, lattice QCD is hindered by the infamous sign problem [31] at finite baryon chemical potential and may therefore only explore the region near the temperature axis with $\mu_B < T$. From effective models, at very high baryon chemical potential above $\mu_B \approx 900$ MeV

as with atomic nuclei, and at low temperature, QCD matter is predicted to transition to a color superconducting phase, which is a degenerate Fermi gas of quarks with a condensate of Cooper pairs near the Fermi surface [32, 33].

The phase transition from gas of confined hadrons to a QGP state is predicted to be a crossover transition by lattice calculations. However, effective models at low temperature and high baryon chemical potential suggest a first-order phase transition [34]. This suggests that there exists a critical point at some unknown temperature and baryon chemical potential. These features of the possible QCD phase diagram are graphically summarized in Fig. 1.4.
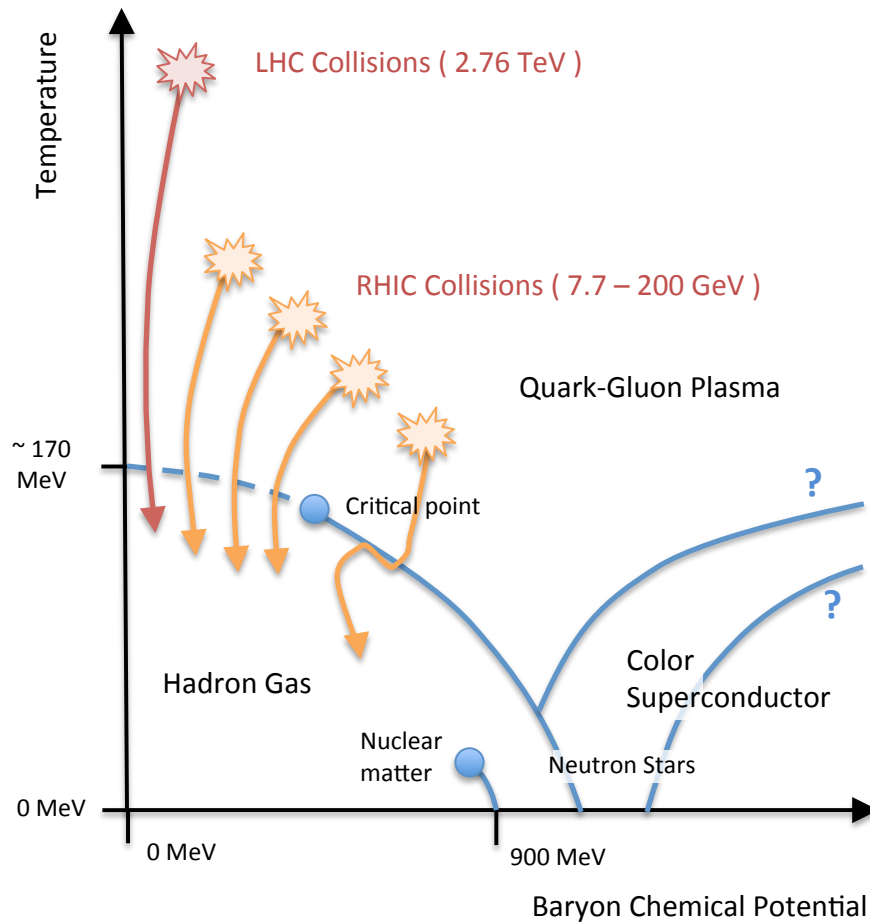
Figure 1.4: Cartoon of a possible QCD phase diagram, with the positions of RHIC and LHC heavy ion collisions indicated.

Although intriguing, the prospect of exploring the exotic states of predicted nuclear matter is daunting, given that in our everyday world, including nuclear reactors and the center of the sun, the temperature scale is nowhere near the order of 100 MeV, or 10^{12} K. It is only in large particle colliders that these energies are achieved, and so to create, for a brief span of time, a bulk medium for which a thermodynamic description would be reasonable, one may consider collisions between heavy nuclei at relativistic energies.

1.2.2 Experimental Production of the QGP

The first experiments that might be described as relativistic heavy ion collisions began in 1974 when the Bevatron at Berkeley was linked with a transfer line to the SuperHILAC to form the Bevalac. However, at these low energies nuclear breakup dominates over particle production. By 1982 this machine was capable of accelerating ions up to uranium at energies up to 2.1 GeV per nucleon. Notable experiments such as the Streamer Chamber and later the Plastic Ball worked to look for evidence of collective flow and thermal equilibrium of the medium produced in these collisions [35].

In 1992 the AGS Booster was added to the Alternating Gradient Synchrotron (AGS) at Brookhaven National Laboratory (BNL), allowing the acceleration of nuclei as large as Au at energies of up to 11.6 GeV per nucleon. Collisions with fixed targets were observed in several experiments. For example, the E864 experiment measured the momentum spectrum of particles produced in the collision, and the E985 experiment looked at the collective motion of the medium. Additional experiments looked at many other observables, such as strangeness enhancement and possible exotic metastable strange particles.

Starting in 1994, the CERN SPS accelerator produced beams of nuclei as large as Pb with energies of up to 160 GeV per nucleon. Several experiments participated in the lead beam program to look at both rare signals and bulk observables. A major focus of the project was to look for several proposed signatures of QGP formation. These signatures include the suppression of the J/ψ meson (NA50 experiment), the broadening of the ρ meson spectral function (NA45 experiment), the enchanced production of hadrons including strange quarks (WA97, NA45, and NA50 experiments), and the indication of a state of thermal equilibrium from measured hadron abundances (NA44, NA45, NA49, NA50, NA52, WA97, and WA98 experiments). Their findings culminated in a press release in 2000 indicating that the accumulated data for several of these signatures provided substantial indirect evidence for deconfinement and a new state of nuclear matter [36]. However, the understanding of some of these signals would be immediately challenged by surprising results from higher energy collisions at the RHIC accelerator over the next few years.

The Relativistic Heavy Ion Collider began its first run of heavy ion collisions in 2000. Unlike previous experiments, RHIC collides two beams of ions at an energy of 100 GeV per nucleon, for a center-of-mass energy per nucleon pair of $\sqrt{s_{NN}} = 200$ GeV. Note that for ultrarelativistic collisions, the center-of-mass energy is much greater for two colliding beams than it is for a single beam hitting a fixed target. One may approximate the equivalent center-of-mass energy as $E_{\mathrm{cm}} = \sqrt{2M_{\mathrm{target}}E_{\mathrm{projectile}}}$. In the equivalent center-of-mass frame, AGS collisions are at approximately $\sqrt{s_{NN}} = 5$ GeV, and SPS collisions are at about $\sqrt{s_{NN}} = 19$ GeV. The collisions at RHIC therefore represented an entire order of magnitude more energy per nucleon than previous experiments.

Four major experiments, BRAHMS, PHOBOS, PHENIX, and STAR, were set up to observe the collisions and today PHENIX and STAR continue to take data and upgrade the detector systems. The first three years of physics runs for the project were summarized in a set of four experimental reviews known as the RHIC "White Papers" as well as a theoretical interpretation of the combined results [37–41]. In this summary, a strong case was made that the hot and dense medium observed in AuAu collisions at high energy did in fact consist of deconfined quark and gluon degrees of

freedom. However, the measured properties of this medium are consistent with a model where the quarks and gluons are strongly coupled and that the medium has the properties of a nearly perfect fluid.

More recently, RHIC has performed a search for the critical point of the QGP phase diagram by examining AuAu collisions at lower energies of $\sqrt{s_{NN}}$ = 7.7, 19.6, 27, and 39 GeV, producing nuclear matter at lower temperature and higher baryon chemical potential as the center-of-mass energy decreases [42, 43].

The Large Hadron Collider (LHC) was designed with the capability to accelerate both protons and ions as heavy as lead. Of the four major LHC experiments, the ALICE detector was designed primarily or the analysis of heavy ion collisions, and the CMS and ALTAS detectors, while designed primarily to analyze pp collisions, are also capable machines for analyzing heavy ion collision events. The LHC has already produced PbPb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV, which is over an order of magnitude higher than those produced at RHIC, and is therefore expected to explore a higher temperature region of the QCD phase diagram. These collisions produced at the LHC and measured with the CMS detector are the focus of this document, and are described in detail in Chapter 2.

The possible regions of the QCD phase diagram which may be explored by the RHIC and LHC collider programs are shown on Fig. 1.4.

1.2.3 Experimental Observation of the QGP

The medium produced in a heavy ion collision is not directly observed, and all of its properties must be inferred from the distributions of long-lived particles which are ultimately produced after the medium cools and undergoes a phase transition to a hadronic gas. Observable quantities must be measured from these distributions, typically averaged over many collision events, and compared with theoretical predictions of various signatures of QGP formation. Although there are many such proposed and measured signatures, this document will focus only on two such signatures: collective motion and partonic energy loss. In this section, general features of the collision systems as well as the observation of collective motion are discussed. Partonic energy loss is described in Sec. 1.2.4.

Understanding the evolution and properties of an expanding QGP medium requires an understanding of the initial state in which the medium is formed, and therefore the initial geometry of the collision system. For the nuclear species typically collided at RHIC and the LHC, Au and Pb, the nucleus is roughly spherical in shape. In the laboratory frame of reference, the colliding nuclei are Lorentz-contracted into thin pancakes, and collide with some impact parameter b. As shown in the left of Fig. 1.5, the impact parameter defines a lenticular overlap region. Nucleons appearing in this region which collide are called participant nucleons, and nucleons appearing outside this region do not collide and are called spectator nucleons.

Femtoscopic quantities such as the impact parameter and number of participating nucleons are typically determined by a model-dependent estimate based on some experimental observable such as the total number of charged particles or energy produced in some kinematic range. As shown in
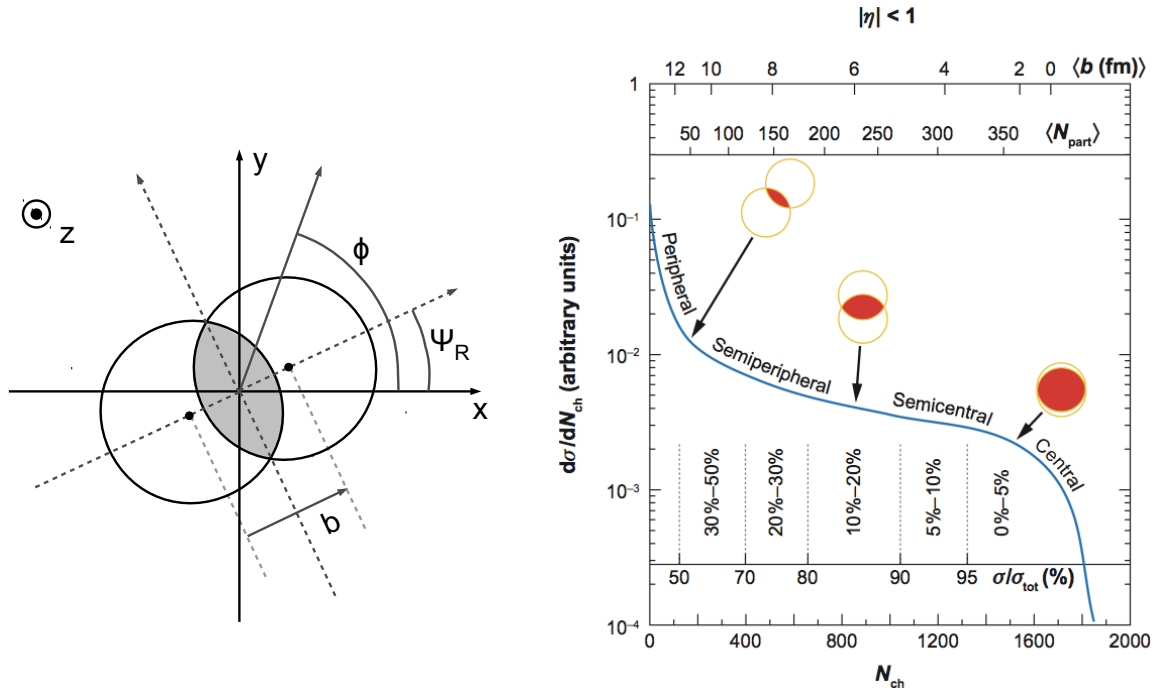
Figure 1.5: (Left) Schematic of a heavy ion collision in the transverse plane indicating the plane of the reaction, Ψ_R, the impact parameter, b, and the overlap region as shown in gray [44]. (Right) Illustration of the relationship between observed centrality estimated by the number of charged particles produced in a collision, N_{ch}, and the impact parameter, b, and average number of participating nucleons, N_{part} [45]. The plotted distribution is illustrative and not an actual measurement.

the right of Fig. 1.5, the total inelastic hadronic cross section can be divided into classes based on the observable quantity. These classes, called centrality classes, are then mapped to average values of the impact parameter, approximate shape of the nuclear overlap area, and number of participating nucleons. The details of the commonly used Glauber model that performs this mapping are described in Sec. 5.5.

As proposed by Bjorken in 1983 [46], one may treat the expansion of the QGP medium as a hydrodynamic system. To practically determine the accuracy of such a hydrodynamic model, one must begin with some understanding of the initial conditions of the collision system, which can be roughly determined from the centrality class. The model must also contain a prescription for the temperature and conditions under which the QGP changes phase to a hadronic gas and hydrodynamic evolution halts, as well as how the energy of the medium is transformed into hadrons. With these ingredients in place, the results of a hydrodynamic model can be compared with the yield of charged particles in a given centrality class to look for evidence of collective motion, or "flow".

Outward collective expansion of the medium may result in a change in the shape of the transverse momentum spectrum of charged hadrons produced in the collision. This could be seen as an overall increase in the average transverse momentum of the particles produced in the collision, and as a

broadening which is dependent on the mass of the hadrons being produced. Such effects are often referred to as "radial flow".

Due to the natural lenticular shape of non-central collisions, transverse pressure gradients along the reaction plane angle, Ψ_R, are expected to be greater than those perpendicular to the plane, which may result in an azimuthal anisotropy in the distribution of charged particles. This effect can be measured by performing a Fourier decomposition on the charged-particle invariant yield in the azimuthal angle, ϕ:

$$E\frac{d^3N}{dp^3} = \frac{d^3N}{p_\mathrm{T}dp_\mathrm{T}dyd\phi} = \frac{d^2N_0}{p_\mathrm{T}dp_\mathrm{T}dy}\left(1 + \sum_{n=1}^{\infty} v_n \cos\left[n(\phi - \Psi_k)\right]\right) \tag{1.6}$$

Here Ψ_k represents the flow angle defined as the angle which maximizes the value of the Fourier coefficient v_k. When the expansion is performed in terms of Ψ_2, the coefficient v_2 identified with collective motion termed "elliptic flow". Due to event-by-event fluctuations in the shape of the the the overlap area, higher-order flow may also be identified with the other coefficients, such as v_3 [47]. One must be careful in equating a positive v_2 signal with collective flow, as various effects unrelated to collective motion may contribute to the measured value. Additionally, the method of reconstructing v_n from the finite particle production in each event, and event-by-event fluctuations in the magnitude of the flow may distort the signal. For an explanation of the various methods for determining the v_n coefficients, and the effects of non-flow and event-by-event flow fluctuations on the measurement, see Ref. [48] and references therein.

Early measurements of particle spectra and anisotropies in AuAu collisions at $\sqrt{s_{NN}} = 130$ and 200 GeV taken at RHIC were well described by ideal relativistic hydrodynamic models with the shear viscosity, η, taken as zero, providing strong evidence that the bulk medium shows efficient thermalization and behaves hydrodynamically [49]. The discrepancies that remained between the data and these models have been explored using additional precise measurements and improved models. Current models may incorporate features such as pre-thermalization dynamics, event-by-event fluctuations of the initial state, finite viscosity, and hadronic rescattering after the phase transition from a QGP to a hadron gas. A recent review of the current state of hydrodynamic modeling may be found in Ref. [50].

As an example of the agreement of a hydrodynamic model to collision data, a comparison of the mean transverse momentum of identified pions, kaons, and protons measured in AuAu collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ [51] is compared to the hydrodynamic model calculation of Bożek and Broniowski [52] in Fig. 1.6. These transverse momenta constrain the values of the shear viscosity, bulk viscosity, and freeze-out temperature used in the model.

The shear viscosity to entropy density, η/s, is of particular interest due to its extremely low value. From available data and model calculations, this quantity may vary as the medium evolves but can be conservatively estimated as $0.07 < \eta/s < 0.43$ when expressed in units of $\hbar/k_B$ [52]. For comparison, in water $\eta/s \sim 2-5$. The only other known physical system to approach such a low value of η/s is a degenerate Fermi gas of ^{6}Li atoms prepared at a temperature on the order of 10^{-6} K, which has been measured as $\eta/s \lesssim 0.5$ [53]. Both of these systems approach a conjectured universal lower bound of $\eta/s = 1/4\pi$ [54].
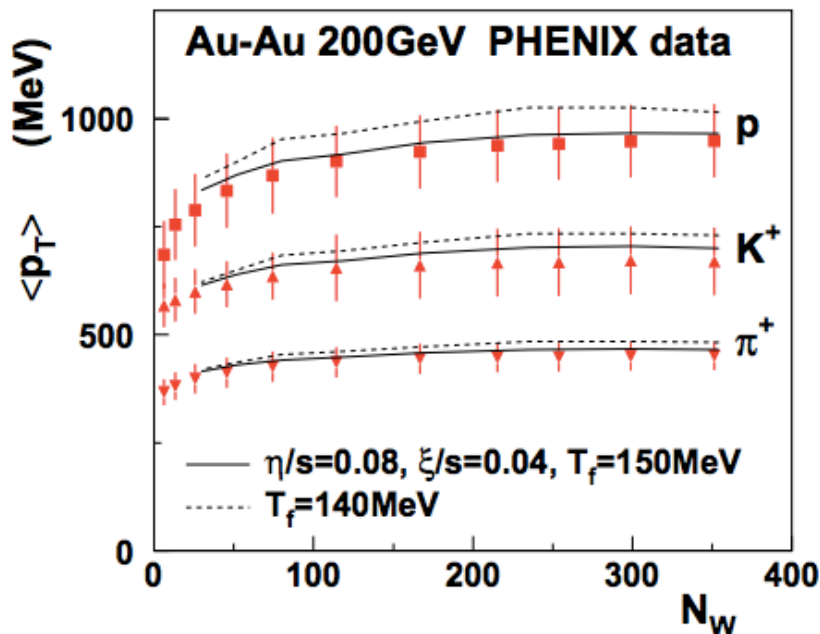
Figure 1.6: The mean transverse momentum of identified charged pions, kaons, and protons in $\sqrt{s_{NN}} = 200$ GeV AuAu collisions shown as a function of N_w, the number of wounded (participant) nucleons in the collision [51]. These $\langle p_T \rangle$ values are compared with a hydrodynamic model with the shear viscosity to entropy density, η/s, and bulk viscosity to entropy density, ξ/s, set to values as shown in the figure in units of $\hbar/k_B$. Two values of the freeze-out temperature, T_f, are used [52].

With the advent of PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, it is now possible to try to compare and understand how the properties of the medium may differ when produced at a much higher temperature and energy density. The analysis presented in this document extends the measurement of the charged-particle transverse momentum spectrum in these PbPb collisions to low-p_T, allowing the mean transverse momentum to be calculated. This measurement will help to constrain the parameters of hydrodynamic models and therefore yield information about the properties of the QGP medium at even higher temperatures than explored at RHIC.

1.2.4 Particle production in Heavy Ion Collisions and Jet Quenching

Most of the charged-particle production in heavy ion collisions occurs at an energy scale too low to understand using perturbative QCD methods. As with pp collisions, rare collisions between nucleon pairs occur at high energy scale and produce "hard" outgoing partons with transverse momenta of several GeV/c. These outgoing partons may interact with the medium, potentially losing energy by collisions or gluonic radiation. This energy loss phenomenon is known as jet quenching, and it may reduce the observed yield of high-p_T jets in heavy ion collisions. The hard-scattered parton may

also be viewed an effective probe of the medium, as the energy loss of the parton is dependent on the properties of the medium and its expanding geometry. For this reason such high-p_T particles are often called hard probes of the medium.

Several theoretical approaches exist for the modeling of this energy loss, and for a broad description of these approaches, a recent review may be found in Ref. [55]. All of these approaches share in common a dependence on the conjectured factorization of the energy loss mechanism from the hard scattering cross section and the PDF. The prediction of the invariant cross section of some hadron C with momentum p_C in a heavy ion collision using any of these approaches may be described in a similar manner to Eqn. 1.5 [56]:

$$
\begin{aligned}
E_C \frac{d\sigma}{d^3 p_C} = \sum_{abc} \int dx_A dx_B \frac{dz}{z} \quad & f_{a/A}^N(x_A, Q_F^2) f_{b/B}^N(x_B, Q_F^2) \\
& \times |\vec{k}_c| \frac{d\hat{\sigma}}{d^3 k_c}(p_c/z\sqrt{s}, Q_F^2) \\
& \times \mathcal{P}(c \to c' \,|\, T, u^\mu) \\
& \times D_{C/c'}(z, Q^2)
\end{aligned}
\tag{1.7}
$$

Here there are two major differences from the factorized equation for pp collisions. First, the PDF $f_{a/A}^N(x_A, Q_F^2)$ represents the PDF of a nucleon in a large nucleus which is modified relative to the free proton PDF. The nuclear PDFs are further discussed in Sec. 1.3. The second difference is the addition of the factor $\mathcal{P}(c \to c' \,|\, T, u^\mu)$, which is the conditional probability containing the effect of the medium changing the parton c to c'. This factor is dependent on the temperature T and flow velocity u^μ which may change as the medium evolves. The integration must now also be performed over the geometry of the medium.

In order to understand the effect of the medium on charged-particle production in terms of experimental observables, one may picture a heavy ion, or AA, collision as consisting of many binary collisions between individual nucleons. One may then compare the particle production in the heavy ion collision to the particle production in an equivalent number of pp collisions. This comparison is given by the nuclear modification factor, or R_AA, defined as

$$
R_\mathrm{AA} = \frac{d^2 N^\mathrm{AA}/dp_\mathrm{T} d\eta}{\langle T_\mathrm{AA} \rangle \, d^2 \sigma^\mathrm{pp}/dp_\mathrm{T} d\eta},
\tag{1.8}
$$

where N^AA is the per-collision particle yield in AA collisions, σ^pp is the particle production cross section in pp collisions, and $\langle T_\mathrm{AA} \rangle$ is a scaling factor defined as the ratio of the number of binary nucleon-nucleon collisions, N_coll, to the inelastic nucleon-nucleon cross section, $\sigma_\mathrm{inel}^{NN}$. These last two quantities are typically determined using a Glauber model as described in Sec. 5.5.

The nuclear modification factor may be measured differentially in p_T or η, in specific centrality classes, or for specific particle species. If $R_\mathrm{AA} = 1$, one states that production is not modified relative to pp collisions. If $R_\mathrm{AA} > 1$, one states that production is enhanced, and if $R_\mathrm{AA} < 1$, one states that production is suppressed, which is the general expectation for hadrons at high-p_T given

the hypothesis of partonic energy loss in the medium produced in heavy ion collisions.
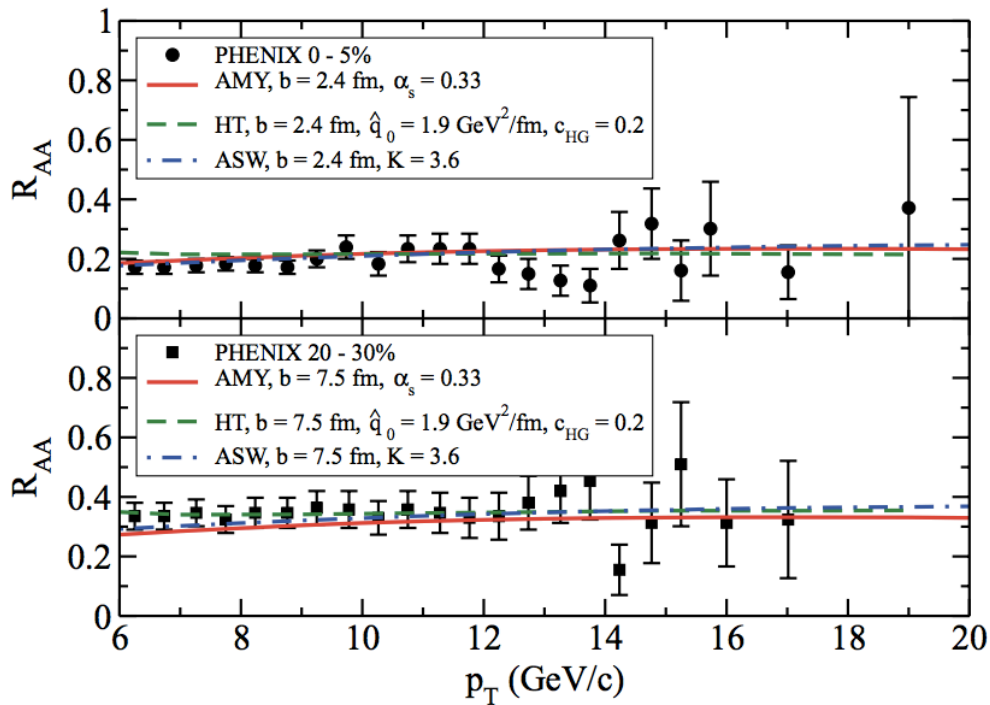
Figure 1.7: The R_{AA} of neutral pions in $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ AuAu collisions as a function of p_{T} as measured with the PHENIX detector in the centrality ranges of 0-5% (Top) and 20-30% (Bottom) [57]. Three different energy loss models are fit to the 0-5% centrality data, and predict the 20-30% centrality data [55].

In Fig. 1.7, the R_{AA} of neutral pions in $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ AuAu collisions [57] is compared with theoretical predictions from three different energy loss models in two different centrality classes [55]. The three approaches used to predict the data are tuned to fit the most central 0-5% events, and predict the 20-30% central events well. Note that the R_{AA} decreases with the most central events, consistent with the picture of a geometrically larger medium, and thus a longer path length that the hard scattered parton must traverse before leaving the medium, resulting in a greater energy loss. In this manner, the comparison of different centrality ranges may aid in the understanding of the path length dependence of partonic energy loss.

The higher energy PbPb collisions produced at the LHC at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$ may also give insight into the temperature dependence of partonic energy loss in the medium. Additionally, the capabilities of the CMS detector, as described in Chapter 2, allow measurements of the charged-particle R_{AA} at much higher p_{T} than previously measured. In Fig. 1.8, the R_{AA} of charged particles in $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$ PbPb collisions is shown as a function of p_{T} for six centrality ranges [27]. As with the lower energy AuAu collisions, strong suppression is noted, especially in the most central events where a larger medium is produced. There is also a rise in R_{AA} for $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$ which was not previously seen due to a lack of precise data in this very high p_{T} range.

Additional studies are required to confirm this general interpretation of $R_{\mathrm{AA}} < 1$ for charged hadrons and neutral pions as an indication of medium induced energy loss. For example, one may
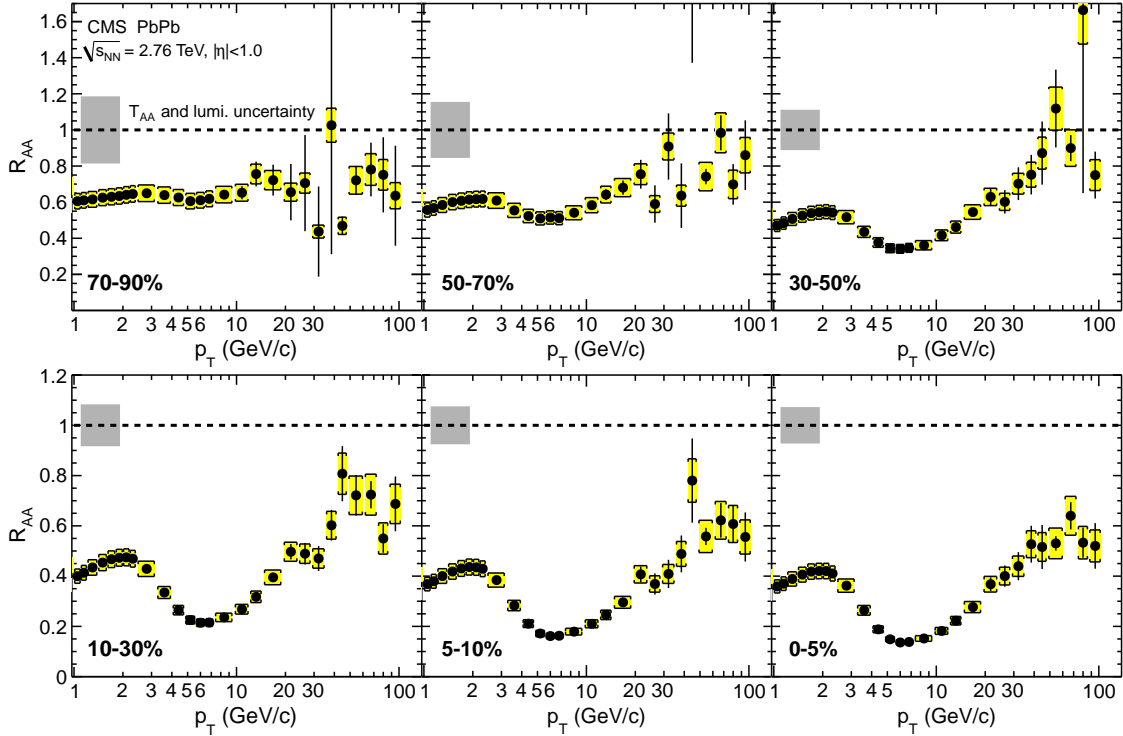
Figure 1.8: The charged-particle R_{AA} in $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ PbPb collisions as a function of p_T in six centrality classes as measured with the CMS detector. The error bars represent the statistical uncertainties and the yellow boxes represent the p_T-dependent systematic uncertainties. The normalization uncertainty on the $\langle T_{AA} \rangle$ calculation is shown by the shaded gray bar in each centrality class [27].

hypothesize that the reduction in the cross section is purely an initial state effect, such as the nuclear modification of the PDF. To test such a hypothesis, one may look at the R_{AA} of colorless probes, such as prompt photons, $W^{\pm}$, or Z bosons decaying into leptons. These particles do not interact strongly and would therefore not be expected to interact significantly with the medium.

In Fig. 1.9, the R_{AA} of Z bosons decaying into e^+e^- and $\mu^+\mu^-$ pairs in $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ PbPb collisions is shown as a function of the average number of participating nucleons in several centrality classes, as well as for inclusive PbPb collision events [58]. For all centrality classes analyzed, R_{AA} is consistent with unity, as expected.
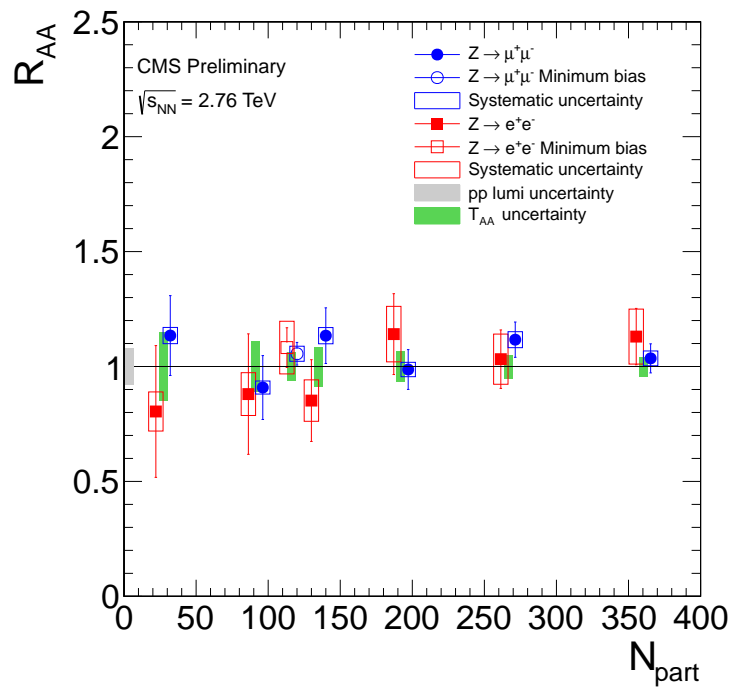
Figure 1.9: Nuclear modification factor for both $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ in $\sqrt{s_{NN}} = 2.76$ TeV PbPb collisions as a function of the number of participating nucleons in the collision [58].

1.3 The Role of Proton-Ion Collisions

The signatures of QGP formation from various observations are now numerous enough to build a compelling argument for the production of this new state of matter. However, in the early years of the RHIC experiments, when the formation of the medium was still in question, it was reasonable to ask if the observed suppression was really an effect of a new state of matter, or simply due to interaction with the hadronic medium of a large nucleus. For example, the suppression of the R_{AA} of neutral pions in AuAu collisions at RHIC can be alternatively modeled as resulting in part due to hadronic rescatterings rather than partonic energy loss in a deconfined medium [59].

To address the alternative explanation of energy loss in a purely hadronic medium, one may proceed by colliding a heavy nucleus, such as Pb or Au, with a proton or deuteron. The working assumption in such a collision is that the hot QGP medium would not form, and that the only effects modifying production of high-p_T hadrons would be initial state effects or hadronic rescatterings.

As shown in the top panel of Fig. 1.10, the PHENIX measurement of the charged-particle R_{dA} in dAu collisions at $\sqrt{s_{NN}} = 200\,\text{GeV}$ [60] does not exhibit any suppression, as is seen in the corresponding central AuAu collisions at the same energy. This further validates the interpretation of partonic energy loss in the QCD medium, but the enhancement of R_{dA} indicates that to more quantitatively understand the nature of partonic energy loss as described by the R_{AA} observable, one must also understand the other initial- and final-state effects that may contribute to this observable. The measurement of R_{dA}, in which the formation of large QGP medium is not expected, provides an observable in which these effects may be more directly studied.

The study of R_{dA} serves a far more important role than as a simple control experiment for AA collision systems. The observed R_{dA} of charged particles is clearly enhanced for $p_T > 2\,\text{GeV}/c$, indicating that other initial- of final-state effects may similarly affect R_{AA}. As seen in the bottom panel of Fig. 1.10, the observed R_{dA} of neutral pions does not exhibit the enhancement present in the charged-particle R_{dA}. This can be explained as an enhancement of baryons relative to mesons and is clearly seen in the spectra of identified protons and pions in dAu collisions [61]. Such effects are interesting due to their potential impact in charged-particle production in AA collision systems, and they are also interesting in their own right. Three such effects resulting in potential nuclear modification are described below: Cronin enhancement, modification of the nuclear PDF, and gluon saturation.

In 1975, experiments at Fermilab of bombarding a Tungsten nucleus with protons created an enhanced number of high p_T hadrons relative to proton-proton collisions. This effect is named the "Cronin effect" or "Cronin enhancement" [62]. Such an effect would make one expect to see $R_{dA} > 1$, as was observed in the charged particles. The Cronin effect is conventionally explained in terms of multiple soft scatterings of an initial state parton by the target nucleus, broadening its momentum prior to hard scattering [63]. This interpretation does not explain the observed enhancement of baryons compared to mesons, and to better incorporate this phenomenon, a model of final-state parton recombination has been alternatively proposed as an explanation of the Cronin effect [64, 65].

A second effect that may modify the observed R_{dA} and R_{AA} is that the bound nucleon parton distribution functions (nPDFs) have been found to differ significantly from the free proton parton
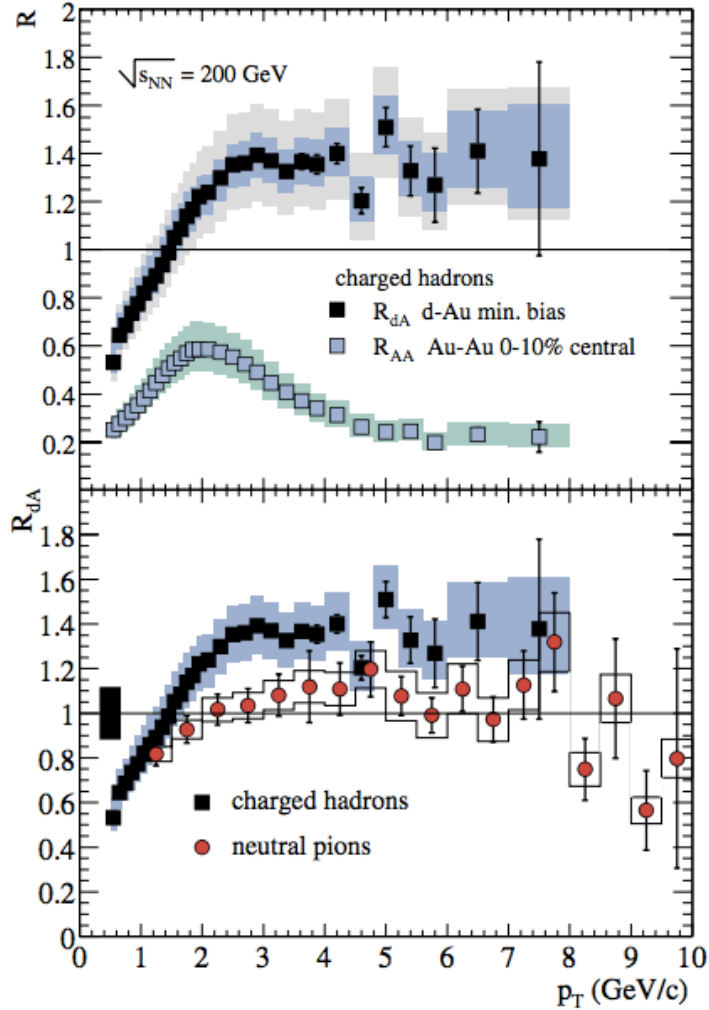
Figure 1.10: (Top) PHENIX measurement of the charged-particle R_{dA} at $\sqrt{s_{NN}} = 200$ GeV for inclusive minimum-bias events compared with the R_{AA} of charged particles at the same energy for 0-10% central events. The inner bands show systematic uncertainties which may vary with p_T, and the outer bands include the normalization uncertainty. (Bottom) Comparison of the R_{dA} of charged hadrons and the R_{dA} of π^0s. The bar at the left represents systematic uncertainties common to both the charged and π^0 measurements [60]

distribution functions by analysis of deeply inelastic lepton-nucleus scattering and Drell-Yan (DY) experiments [68]. In addition to the DIS and DY data, the R_{dA} of neutral pions in dAu collisions at $\sqrt{s_{NN}} = 200$ GeV [69] has been further used to constrain the determination of the nPDFs and their uncertainties [67]. To understand this modification, the ratio of the nuclear PDF for a nucleon in a Pb nucleus to the PDF free proton is shown as a function of x for an energy scale of $Q^2 = 10000$ GeV^2 in Fig. 1.11. In the low-x regime ($x \sim 0.001$), the nPDF is generally found to be suppressed relative to the free proton PDF. This phenomenon is generally referred to as "shadowing". In the moderately high-x regime ($x \sim 0.02$), the nPDF is found to be slightly enhanced, and this effect is known as "anti-shadowing". As the high-p_T particle production observed in pPb collisions is dominated by
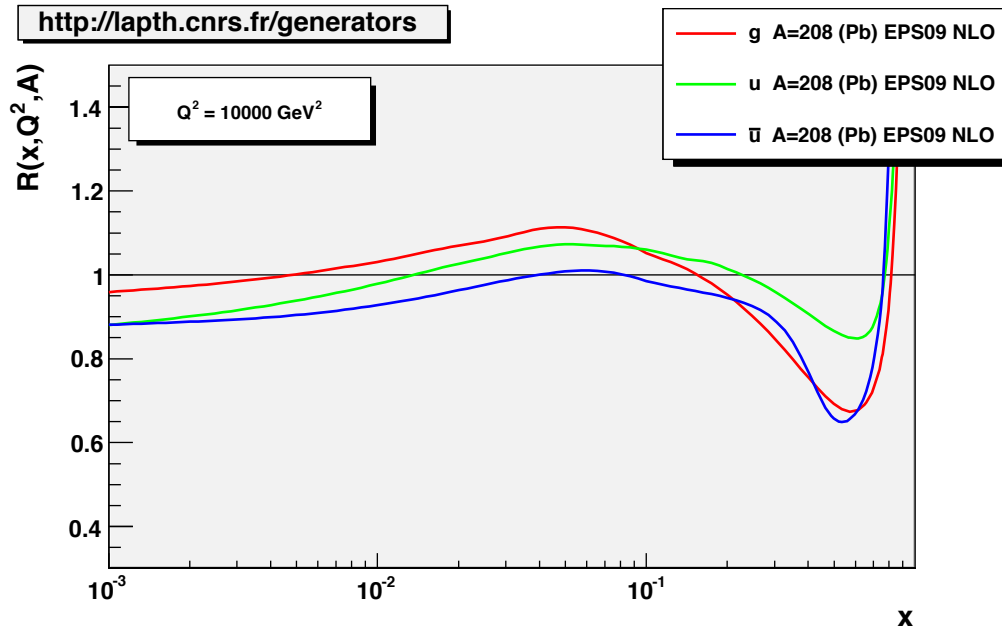
Figure 1.11: The ratio of the nuclear PDF for a nucleon in a Pb nucleus ($A = 208$) to the PDF free proton is shown as a function of x for an energy scale of $Q^2 = 10000 \, \text{GeV}^2$ shown for gluons, u quarks, and $\bar{u}$ antiquarks as calculated using the nPDF generator [66] with the EPS09 NLO parametrization [67].

processes involving high-x partons, the effect of anti-shadowing may enhance the yield of charged particles in pPb collisions relative to the yield in pp collisions. In the very high-x regime, the nPDF is found to be suppressed. This effect is known as the EMC effect, after the European Muon Collaboration first discovered it while measuring the structure functions of iron nuclei [70].

A third effect that may modify both R_{dA} and R_{AA} is partonic saturation. For a fixed energy scale Q^2 and very small Bjorken x, the gluon density increases rapidly, as can be seen in Fig. 1.1. As the gluon number density approaches the unitarity limit for $x \ll 1$, gluon-gluon interactions induce a saturation of the number density at some extremely low value of x [71,72]. This region of $x \ll 1$ with fixed Q^2 is known as the "semi-hard" region of QCD. For heavy nuclei, the gluon density may be expected to scale as $A^{1/3}$, where A is the number of nucleons, so that the onset of saturation effects may be observable at lower energy collisions. The growth of the gluon density and the properties of this regime were found to be calculable using a weak coupling treatment [73,74]. This method an effective field theory called the Color Glass Condensate (CGC), for which a recent review may be found in Ref. [75]. The CGC model provides a potential method to understand the initial state of the heavy ion collision system, which is critical for precisely modeling the bulk dynamics of the QGP medium and the mechanisms of partonic energy loss.

Proton or deuteron collisions with a heavy nucleus at high energy provide a useful environment to test the CGC model and predictions from saturation physics, since saturation effects are amplified in the nucleus, and since large final state effects from the QGP medium are expected to be absent. For example, The transverse momentum spectrum of charged hadrons at forward rapidity in dAu collisions has been well described treating the Au nucleus as a CGC [76]. Even higher energy pPb

21

collisions at the LHC make a smaller x region accessible to experimental observation, and so measurements of charged-particle production, correlations, and p_T spectra may aid in the understanding of the saturation regime.

In addition to these known effects observed in deuteron-ion or proton-ion collisions, recent results from pPb collisions produced at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ have shown correlations between particles produced with large separations in rapidity [77–80]. Similar correlation observations in AuAu and PbPb collisions have commonly been interpreted as evidence of collective flow. Indeed, these new results in pPb collisions have also been described using a hydrodynamic model [81], raising the question of whether a QGP medium could potentially be produced in such a small collision system at such high energy. Alternatively, these correlations have been described in terms of multi-parton interactions in the framework of the CGC model [82].

Whatever the source of these novel correlation phenomena, they highlight the potential for discovery in pPb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$.

1.4   The Structure and Goals of this Thesis

The goals of this thesis are two-fold. The first goal is to present the measurement of the per-event charged-particle invariant yield in PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ as a function of p_T with the CMS detector. Unlike the previous CMS measurement of the charged-particle p_T spectrum [27], this measurement extends below $p_T = 1000\,\text{MeV}/c$ down to $p_T = 300\,\text{MeV}/c$, allowing for a better understanding of the bulk particle production in the medium produced in the collision. This low-p_T measurement is performed in 12 different centrality classes, and is also extended to multiple pseudorapidity ranges. The low-p_T reach of the measurement enables one to perform an extrapolation to $p_T = 0$ and to therefore determine the $\langle p_T \rangle$ of charged particles produced in the collision. This measurement serves to better constrain hydrodynamic models of the evolution of the medium as well as the understanding of the initial state of the collision, where saturation may play a role.

The second goal of this thesis is to present the measurement of R_{pPb} for pPb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ at midrapidity as a function of p_T with the CMS detector. This measurement is performed up to a very high p_T of 100 GeV/c, and therefore significantly extends a previous measurement of R_{pPb} from the ALICE collaboration in the range of $0.5 < p_T < 20\,\text{GeV}/c$ [83]. The result of this measurement is in excellent agreement with the previous ALICE measurement where the p_T ranges overlap, but exhibits a very surprising result at very high-p_T which may challenge the understanding of the nuclear PDF or point towards a novel effect that was not predicted.

In Chapter 2, the LHC machine and CMS detector are described in general. The collision data which have been produced by the LHC and recorded by CMS are presented. Special attention is given to the silicon tracker subsystem at the heart of the CMS detector, which is used to reconstruct the trajectories of charged particles produced in the collisions.

Chapter 3 describes the methods of charged-particle reconstruction for pPb and PbPb collisions from the signals measured in the silicon tracker modules. The charged-particle reconstruction used for the pPb collisions is a standard method used for many different CMS measurements. For the PbPb

collisions, the large flux of charged particles creates a challenge for trajectory reconstruction, and an alternate method is used which was in part developed specifically for this measurement. Chapter 4 details the simulation procedure used to evaluate the performance of the charged-particle trajectory reconstruction.

The selection of collision events to form a sample of collision data is described in Chapter 5. This procedure is different for PbPb and pPb collision events. In the case of the pPb collisions, the phenomenon of "pileup", where multiple collisions are seen by the detector in a single readout, must be considered. The method of defining collision centrality in PbPb collisions and the Glauber model used to connect the centrality classes to the average number of participating nucleons is also described in this chapter.

Chapter 6 details the measurement of the PbPb charged-particle transverse momentum spectrum from the selected data sample and the reconstructed trajectories. The $\langle p_T \rangle$ in each centrality and pseudorapidity range is determined by extrapolation of the measured spectra to $p_T = 0$. The measured $\langle p_T \rangle$ is compared with previous results from lower energy collisions at RHIC, and the significance of the result is discussed.

To date, there does not exist data for pp collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ as such a collision system has not been produced at the LHC or any other facility. Therefore, in order to construct the R_{pPb} of charged particles at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$, the pp collision reference spectrum must be artificially constructed by interpolating between available measurements at different energies. This procedure is detailed in Chapter 7.

In Chapter 8, the measured pPb spectra at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ are presented. Sec. 8.1 details the procedure for creating a data sample with a large enough statistical reach to accurately measure the charged-particle spectrum at very high-p_T in pPb collisions. The CMS detector cannot record every collision event which occurs, and must select interesting collision events via a trigger system described in Sec. 2.5. In order to reach $p_T = 100\,\text{GeV}/c$, a combination of data samples recorded using different trigger criteria is required. Each sample in the combination must then be appropriately normalized. From this combined high-p_T measurement, R_{pPb} is constructed. Additionally, comparisons of the spectra in positive and negative pseudorapidity ranges are constructed. The R_{pPb} is compared with theoretical prediction, and the significance of the results is discussed.

In Chapter 9, concluding remarks are presented relating to both the PbPb and pPb collision measurements and their potential impact on the understanding of the initial state of nuclear collision systems and the hot and dense medium produced in heavy ion collisions.

Chapter 2

The LHC and CMS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider is as of this writing the largest and most powerful two-ring-superconducting-hadron accelerator and collider ever constructed [84]. It resides in a 26.7 km tunnel lying between 45 and 170 m below the ground, settled geographically between *Léman* lake and the *Jura* mountains, crossing the border from Switzerland to France.
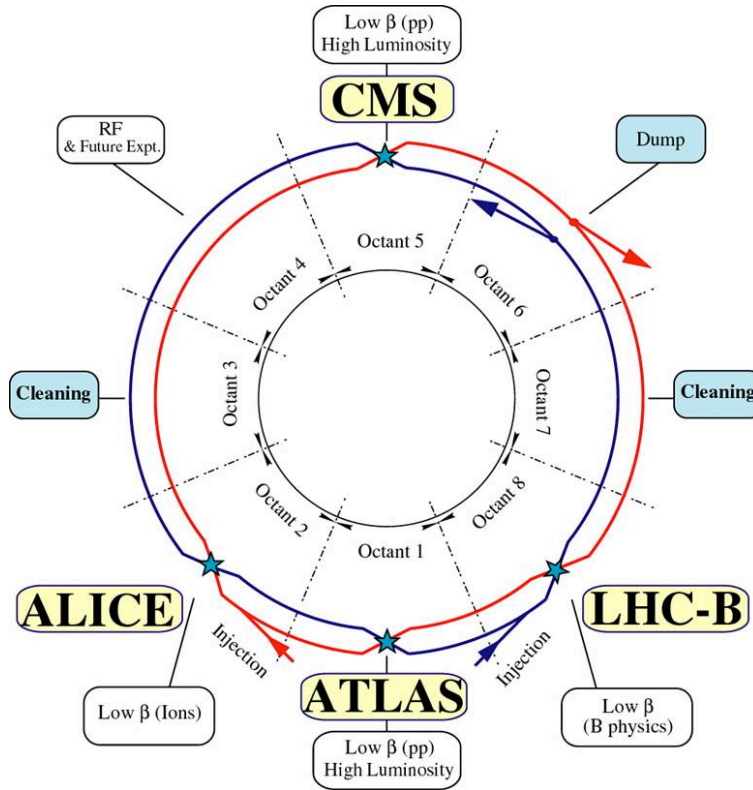
Figure 2.1: Schematic layout of the LHC. [84]

As depicted in Fig. 2.1, the layout of the LHC consists of eight arcs and eight straight sections, with each straight section long enough to serve as an experimental or utility insertion point. Opposite one another across the ring lie the two high-luminosity experiments, the CMS and ATLAS In addition, ALICE and Large Hadron Collider beauty (LHCb) are specialized low-luminosity experiments lying in sections adjacent to ATLAS, and study *b*-physics and heavy ion physics respectively.

The LHC is designed to provide pp collisions at a $\sqrt{s} = 14\,\mathrm{TeV}$ with a nominal luminosity on the order of $\mathcal{L} = 10^{34}\,\mathrm{cm^{-2}s^{-1}}$. In order to achieve this high luminosity, up to 2808 bunches of protons

may be injected into each ring (clockwise and counterclockwise). This results in a narrow spacing of only 25 ns between bunch crossings delivered to the CMS and ATLAS experiments. In addition, the high bunch intensity may result in as many as 20 or more inelastic pp collisions per bunch crossing. Collisions between lead nuclei were also included in the design of the LHC at an early stage, at $\sqrt{s_{NN}} = 5.5$ TeV, and with a nominal luminosity of $\mathcal{L} = 10^{27}$ cm^{-2}s^{-1}.

In November of 2009, the LHC delivered the first pp collision data at center-of-mass energies of 0.9 and 2.36 TeV, breaking the previously held record of $\sqrt{s} = 1.96$ TeV p$\bar{\text{p}}$ collisions produced by the Tevatron accelerator at Fermilab. After a brief winter shutdown, the LHC began its first full year of data taking, delivering over 44 pb^{-1} of pp collision data to the CMS detector at $\sqrt{s} = 7$ TeV. During this period, additional pp collision data was also taken at 0.9 TeV, with a much lower integrated luminosity of several hundred μb^{-1}. The LHC ended 2010 with its first PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, delivering over 7 μb^{-1} to the CMS detector.

In 2011 the LHC increased the intensity of the proton beams by orders of magnitude, and delivered 6.13 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV to CMS. As in 2010, this pp collision run was followed by a PbPb data taking period in which an additional 166.67 μb^{-1} of PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV were delivered to CMS. During the 2011 LHC run, an additional short pp run was performed at $\sqrt{s} = 2.76$ TeV in order to provide a reference for the PbPb data, delivering over 200 nb^{-1}.

In 2012 the energy of each proton beam was increased, and 23.30 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV were delivered to CMS. From these high luminosity pp collision data at $\sqrt{s} = 7$ and 8 TeV, the ATLAS and CMS experiments discovered a new "Higgs-like" resonance with a mass of approximately 125 GeV/c^2, announcing the result in a joint press conference on July 4th, 2012, with subsequent publications quickly following [85, 86]. During this lengthy pp collision run, the LHC was able to deliver a short pilot run of pPb collision data at $\sqrt{s_{NN}} = 5.02$ TeV. In January and February of 2013, a dedicated pPb collision run delivered 31.69 nb^{-1} of pPb collisions also at $\sqrt{s_{NN}} = 5.02$ TeV. This was immediately followed by a few days of pp collision data taking at $\sqrt{s} = 2.76$, in order to provide additional reference data for the 2010 and 2011 PbPb collisions. 5.51 pb^{-1} of reference pp collisions were delivered to CMS in this short period.

Following the 2013 pp run, the LHC entered a 2 year planned shutdown period, after which pp and PbPb collisions will resume at higher energy and instantaneous luminosity.

2.2 Overview of the Compact Muon Solenoid Detector

The Compact Muon Solenoid detector is a multi-purpose apparatus designed primarily to accomplish several goals of the LHC physics program, including the search for the Higgs particle presumed to be responsible for electroweak symmetry breaking, as well as an exploration of possible physics beyond the Standard Model. The CMS detector is therefore designed to measure the results of pp collisions, although the design also allows for a robust heavy ion physics program in which PbPb and pPb collisions are analyzed. The specific major design goals are as follows:

- The ability to reconstruct and identify muons with precise momentum resolution so that the mass of dimuon pairs can be determined with $\approx 1\%$ accuracy at 100 GeV, and the muon charge determined at momenta of up to 1 TeV.

- The ability to reconstruct charged particles with high efficiency, and to determine their distance of closest approach to the collision, so that particles resulting from the weak decays of heavy-flavor hadrons or τ leptons may be identified.

- The ability to reconstruct isolated photons and electrons, so that the mass of diphoton pairs can be determined with $\approx 1\%$ accuracy at 100 GeV, with wide geometric coverage.

- The ability to reconstruct the energy of jets produced in the collision with nearly hermetic geometric coverage, with such precision as to determine missing-transverse-energy and resolve the mass of dijet pairs.

- The ability to select, from an event rate of up to 10^9 inelastic collisions per second, approximately 100 events per second of interest and store the raw data for later analysis.
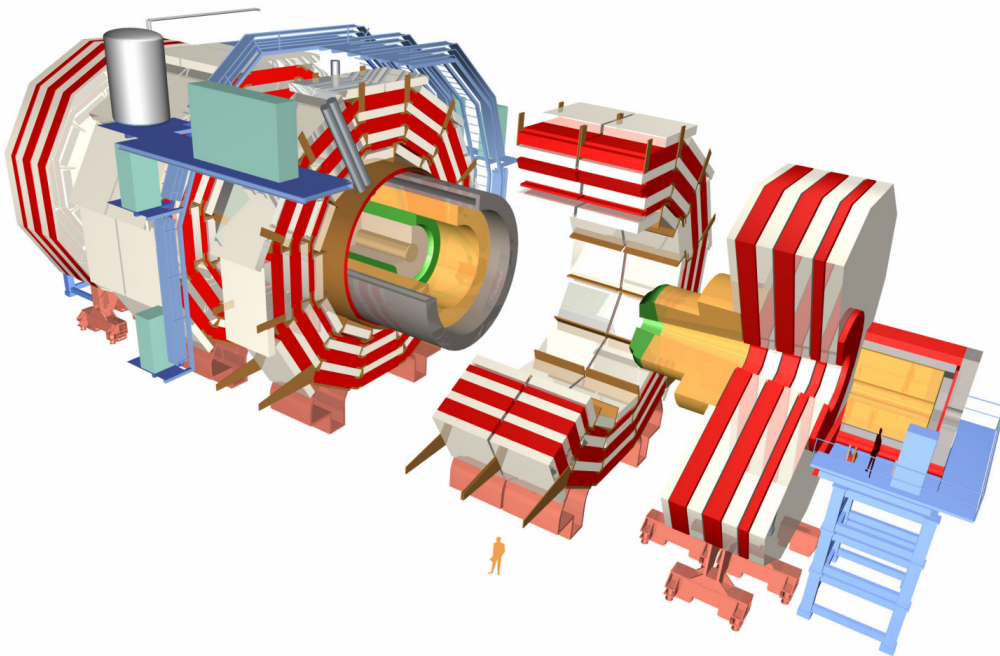
Figure 2.2: Overall layout of the CMS detector, with the silhouette of a person for scale [87]. The beige cylinder in the center of the diagram represents the silicon tracker. The green material represents the electromagnetic calorimeter, and the gold material represents the hadronic calorimetry. The solenoidal magnet is depicted in gray, and the iron return yoke in red. The muon stations are depicted by the white rectangles.

The distinguishing feature of the CMS detector is the 13-m long superconducting solenoidal magnet with a 6-m inner diameter, and a design magnetic field strength of 4 Tesla. The bore of this large magnet coil is sufficiently large to contain compact design of an all-silicon inner tracking system to reconstruct charged particles, an Electromagnetic Calorimeter (ECAL) for the measurement of photon energies, and a Hadronic Calorimeter (HCAL) for the reconstruction of jets. Outside of the magnet coil, the return field is sufficient to saturate the 1.5 m thick iron return yoke, within which 4 layers of muon tracking stations are interspersed, allowing for muon identification with complete geometric coverage. To select collision events of particular interest, a two-stage trigger

system is employed. The Level-1 (L1) Trigger system consists of custom designed programmable electronics, which reduces the event rate to under 100 kHz, and the subsequent High-Level Trigger (HLT) system employs over a thousand commercial processors to further reduce the event rate. The overall layout of the detector is depicted in Fig. 2.2.

The CMS detector is described using a coordinate system with the positive z-axis pointing in the counter-clockwise beam direction, or geographically towards the west. The positive y-axis points up towards the surface, and the positive x-axis is chosen to make the coordinate system right-handed, pointing radially inward towards the center of the LHC. The origin of the coordinate system is the center of the CMS detector, which is the nominal interaction point.

A complete description of the CMS detector can be found in [88]. In the subsequent sections of this chapter, certain details of the detector subsystems relevant to the analysis of charged-particle spectra are described.

2.3 The Inner Tracker System

The goal of the CMS inner tracker system is to reconstruct precisely the trajectories of charged particles and to reconstruct secondary vertices produced from the weak decay of short lived particles. At the design luminosity of the LHC, there were expected to be around 1000 particles produced within the inner tracker acceptance per bunch crossing, where on average 20 pp collisions would occur. This is fortunately not dissimilar to the number of particles produced in a single PbPb collision at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$, which was measured to be 1612 ± 55 per unit pseudorapidity at $\eta = 0$ [89]. The CMS inner tracker is therefore a well designed system for both high-luminosity pp collisions and PbPb collisions.

The inner tracker system occupies a cylindrical volume centered at the nominal interaction point with a length of 5.8 m and a radius of 1.25 m. The solenoidal magnet is designed to provide a nearly uniform 4 Tesla field throughout this volume, although for the first data taking period from 2009 to 2013 it generated an actual field strength of 3.8 Tesla. In order to handle the dense flux of charged particles and brief time between bunch crossings of as little as 25 ns, the system is required to have fine granularity and fast response. This requires a high-power density of on-detector electronics, which in turn requires an efficient cooling system. These demands had to be balanced by the need to minimize the total amount of material in order to minimize the effects of photon conversion, multiple scattering, and nuclear interactions which could degrade the precision of measurements made using the calorimeter systems.

The inner tracker is comprised of two major subsystems, the silicon pixel detector, and the silicon strip detector. The silicon pixel detector resides closest to the beampipe and is comprised of three concentric cylindrical barrel layers at midrapidity, and two disc-shaped endcap layers at positive and negative forward pseudorapidity. Outside of the pixel detector, the silicon strip detector is comprised of 10 barrel layers, comprising the Tracker Inner Barrel (TIB) and Tracker Outer Barrel (TOB). At forward rapidity, the silicon strip detector additionally has 3 smaller disc layers called the Tracker Inner Disc (TID) and 9 larger disc layers called the Tracker Endcap (TEC). A schematic cross section of the inner tracker is shown in Fig. 2.3.
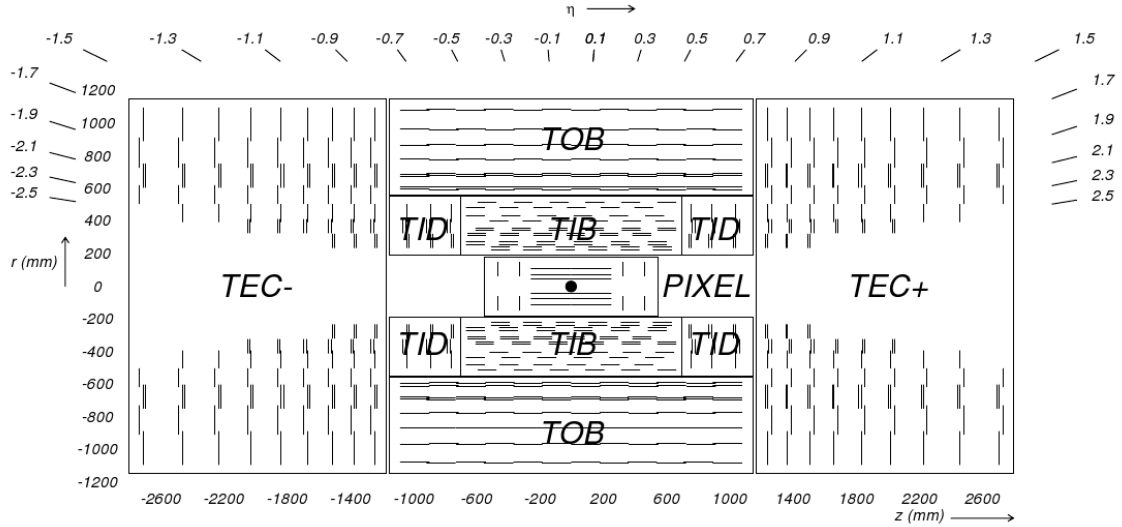
Figure 2.3: Schematic cross section of the CMS silicon inner tracker [88]. Each short line represents a detector module, with double lines indicating back-to-back modules that deliver stereo hits.

2.3.1 The Silicon Pixel Detector

The silicon pixel detector consists of three 53 cm long barrel layers at a distance of 4.4, 7.3, and 10.2 cm from the nominal interaction point, and four endcap layers at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm at a distance of 6–15 cm from the beam axis. This gives complete azimuthal coverage over the pseudorapidity range of $|\eta| < 2.5$, ensuring that nearly every possible trajectory will intersect the detector on some combination of three layers. At low η the three barrel layers will provide the three hits needed to reconstruct a helical trajectory, and at high η the innermost barrel layer combined with the two endcap layers likewise provides three hits. The layout of the pixel detector system is shown in Fig. 2.4.

The detector is comprised of pixel cells with an area of $100 \times 150 \ \mu m^2$. This allows for precise resolution of track trajectories in both the $r - \phi$ and z directions, which in turn allows for the precise reconstruction of secondary vertex positions in 3D space. In total, the system contains over 66 million pixel cells, covering a total area of $1.06 \ m^2$. The pixel cells are bump bonded to read-out chips each of which reads out an array of 52×80 pixels. The system has a zero-suppressed read out scheme with an analog pulse height read out, which allows for a position resolution of $15 - 20 \ \mu m$ in each direction through charge interpolation.

2.3.2 The Silicon Strip Detector

Outside of the pixel detector, the increased distance from the nominal interaction point allows for larger cell size while keeping overall occupancy low. Assuming an average of 1000 charged particles per bunch crossing, the hit density of the innermost pixel barrel layer at approximately 4 cm from the collision point will be 0.025 particles per mm^2. At 22 cm from the collision point, where the

28
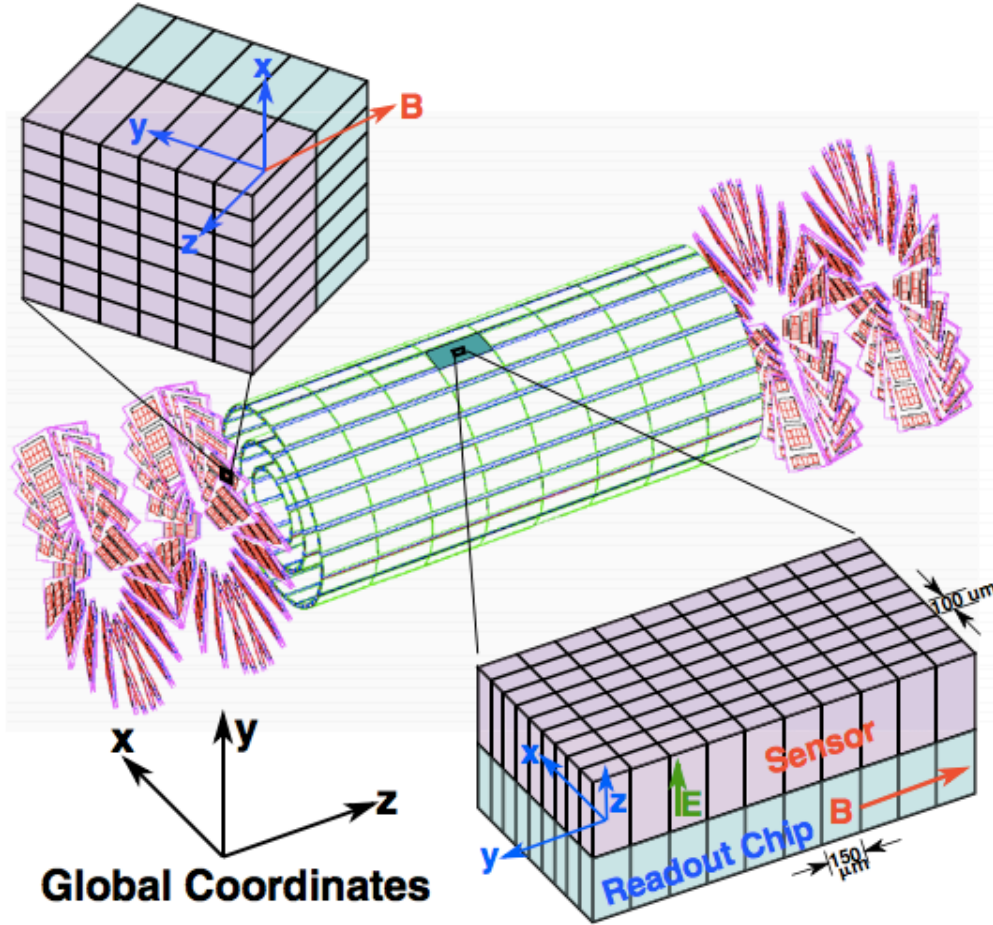
Figure 2.4: The layout of the CMS pixel detector, with example cartoons of individual modules and their local coordinate systems [90].

innermost TIB layer resides, the hit density falls to 0.0015 particles per mm². Finally, at 115 cm, at the outermost layer of the TOB, the hit density has reduced to 7.5×10^{-5} particles per mm². It is therefore possible to achieve 2–3% occupancy on the innermost TIB layer with silicon microstrips with a typical cell size of 10cm × 80μm, and in the outer TOB region with a cell size of 25cm × 180μm. The silicon strip detector contains a total of 9.3 million strips and covers an area of 198 m².

The layout of the TIB and TOB cylinders as well as the TID and TEC rings is such to provide 9 ϕ measurements for a particle with $|\eta| < 2.4$. The first two layers of the TIB, TOB, and TID, as well as the 1st, 2nd, and 5th layers of the TEC are equipped with a second microstrip detector mounted just behind the first with a stereo angle of 100 mrad to provide a measurement of a second coordinate. This ensures 4 two-dimensional measurements for a particle with $|\eta| < 2.4$.

Silicon micro-strips are grouped into sets of 128, and for each set the signals are amplified and stored in the APV25 [91] custom integrated circuit. The signals are only further transmitted upon a positive L1 trigger signal, where they are then digitized to an ADC count. This ADC

count is then corrected by subtracting known pedestal values and common mode noise, which may fluctuate from readout to readout. The signal readout is then zero-suppressed. In the case of PbPb collisions, deformation of the readout signal may happen with significant frequency due to highly ionizing particles. To ensure that these deformations are properly corrected, the non-zero-suppressed readouts are stored for later processing, which is described in Sec. 3.1.1.

2.4   Forward Detectors and Monitoring Systems

2.4.1   The Forward Hadron Calorimeter

The Forward Hadron Calorimeter (HF) is of particular importance to heavy ion physics analysis as the total measured energy in either one or both of the $\pm z$ HF detectors may be used to to estimate the event centrality.

The HF detector subsystem provides full azimuthal coverage over $2.9 < |\eta| < 5.2$, and is required to withstand extremely high particle flux. For each pp collision at design energy, it is estimated that approximately 760 GeV of total energy will be deposited into the HF, compared with only 100 GeV for all of the detector subsystems with $|\eta| < 3$. The HF is comprised of 5 mm thick grooved steel absorber plates, with grooves that are approximately 1 mm wide and deep. Quartz fibers (fused silica core and polymer hard-cladding) are inserted into these grooves. Along with the cladding, the diameter of the fibers is $800 \pm 30\mu$m, and over 1000 km of total fiber length was used to construct the HF.

A signal is generated in the quartz fibers when charged particles with an energy above the Cherenkov threshold (≥ 190 keV for electrons) generate Cherenkov light which is guided to photomultipliers stationed behind a shielding matrix. Half of the quartz fibers run over the full depth of the detector (165 cm) while the other half begin 22 cm from the front of the detector. These two sets of fibers are read out separately, allowing one to distinguish between showers generated by electrons or photons and those generated by hadrons. The hadron showers will on average deposit more energy deeper into the absorber material. Fibers are bundled into 0.175 × 0.15 ($\Delta\eta \times \Delta\phi$) towers.

The HF forms a hollow cylinder with an inner radius of 12.5 cm from the center of the beam line, and an outer radius of 130.0 cm. The front face of the HF is 11.2 m from the nominal interaction point. Azimuthally, the HF consists of 18 modular wedges covering 20° each at both positive and negative pseudorapidity. A diagram of the HF segmentation in the transverse plane is given in Fig. 2.5.

2.4.2   The Beam Scintillator Counters

The Beam Scintillator Counters (BSC) were a series of four scintillator tiles which provided hit and coincidence rates and served as a minimum-bias trigger for inelastic collision events. Two of these tiles were located in front of the HF, at ± 10.9 m from the nominal interaction point, and consisted of hollow discs with a radius of 22–45 cm from the beam line, segmented into 8 slices in azimuth.
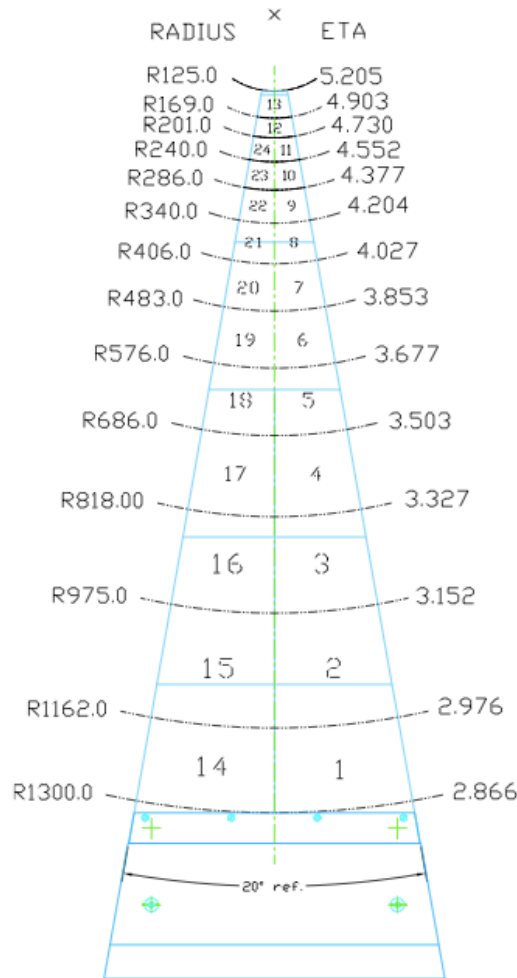
Figure 2.5: Cross-section of a single 20° modular wedge of the HF detector, displaying the radius and pseudorapidity coverage of each segment [88].

Additionally, they featured four large paddles at a radial distance of 55–80 cm. These outer paddles were used to detect beam-halo muons passing through the detector.

The two outer tiles were located at ±14.4 m from the nominal interaction point, and consisted of two tiles with an inner radius of 5 cm and an outer radius of 29 cm. The purpose of these two outer tiles was to distinguish between incoming and outgoing particles along the beamline, as there is a 4 ns timing difference between the inner and outer tiles. This allowed events to be tagged as incoming (towards the interaction point, beam halo only) and outgoing (away from the interaction point, collisions or beam halo).

The BSCs were designed with a short lifetime and functioned during the 2010 and 2011 pp and PbPb runs. In the winter of 2011, the BSCs were removed and replaced with upgraded Beam Halo Counters (BHC) which are also designed to both detect beam-halo events and potentially serve as a minimum bias heavy ion trigger. The performance of the BHC was explored during the 2013 pPb run, but this system was not used in the selection of collisions for this measurement.

2.5 The Level-1 and High Level Trigger System

The CMS Level-1 (L1) [92] and High Level Trigger (HLT) [93] systems provide a two-stage process for reducing the event rate from as high as 40 MHz to under 1 kHz for recording, archival, and subsequent analysis. The L1 trigger is comprised of custom built programmable electronics which are largely integrated with the readout systems of the electromagnetic and hadronic calorimeters and muon stations. The HLT is a software system implemented in a nearby filter farm of approximately one thousand commercial processors.

The L1 trigger must rely on coarsely segmented information from the calorimeters and muon systems, and is designed to reduce the event rate to 30 kHz, with 100 kHz as an absolute maximum allowed value. The decision of whether to drop an event or pass it to the HLT must be made in no more than $3.2\,\mu$s. During this time, the full high-resolution event data is stored in pipelined memories of the front-end readout electronics.

The raw readouts of the heterogeneous detector subsystems are collected into over 600 Front-End Driver (FED) systems which store event data while waiting for the L1 trigger decision. The FEDs provide a common format for the data from the various detector subsystems. The event fragments from the various FEDs are combined into full events through an Event Builder network which then transports full events in parallel to groups of HLT processing nodes.

The HLT processors run a streamlined version of the offline event reconstruction software, which is described generally in Sec. 2.6. For each event a "job" is constructed which unpacks the detector data into a logically convenient format, performs reconstruction of physics objects such as charged-particle tracks and jets, analyzes these objects for relevant physics, and selects these events based on a "menu" of trigger paths.

The HLT menu is composed of hundreds of logically independent paths which may select events based on various criteria such as a pair of muons with $p_\mathrm{T} > 7\,$GeV/c or a single jet with $p_\mathrm{T} > 100\,$GeV/c. If the event rate for a given path is too high to store all such events, a prescale factor N may be applied, selecting only one of every N triggered events for storage. The prescale factors for the various paths may be adjusted over the course of a single LHC fill as the instantaneous luminosity decreases.

2.6 The CMS Computing Model

2.6.1 The CMS Software Application

The simulation, reconstruction, analysis, and HLT filtering of event data is all accomplished in a unified software application known as CMSSW. This flexible application accomplishes these tasks by running a series of precompiled C++ modules, which are configured with a specialized configuration syntax built upon the Python programming language.

All the information ranging from simulated particle trajectories to detector readouts and reconstructed physics objects is stored in a common framework called the Event Data Model (EDM),

which centers around the concept of an event corresponding to a single readout of the detector systems. Events are stored as persistent ROOT files [94].

The simulation of the CMS detector itself is based on the GEANT4 toolkit [95]. This simulation includes the ability to accept generated collision events from a variety of available Monte Carlo (MC) generators, a realistic map of the magnetic field throughout the detector, detailed interactions of the generated particles with the detector material, and the subsequent readout electronics. The simulation is also capable of an event mixing technique, where multiple generated collisions may be combined into a single event. This event mixing system was designed to simulate "pileup", defined as the presence of multiple binary pp (or pPb, PbPb) collisions occurring within a single bunch crossing. The CMS detector simulation is described in more detail in Sec. 4.1.

The reconstruction of event data includes the unpacking of data into a convenient format and applying detector calibration constants, determination of the collision vertex or vertices, reconstruction of charged-particle or muon trajectories (tracks) from the signals in individual tracker layers and muon stations, and reconstruction of jets from the calorimeter signals. Data files are stored in a RAW format, with the uncalibrated signals from each detector subsystem, and in reconstructed, or RECO format, with the physics objects and calibrated signals. The RECO formatted data is then stored, archived, and read for subsequent physics analysis.

2.6.2 CMS Computing Facilities
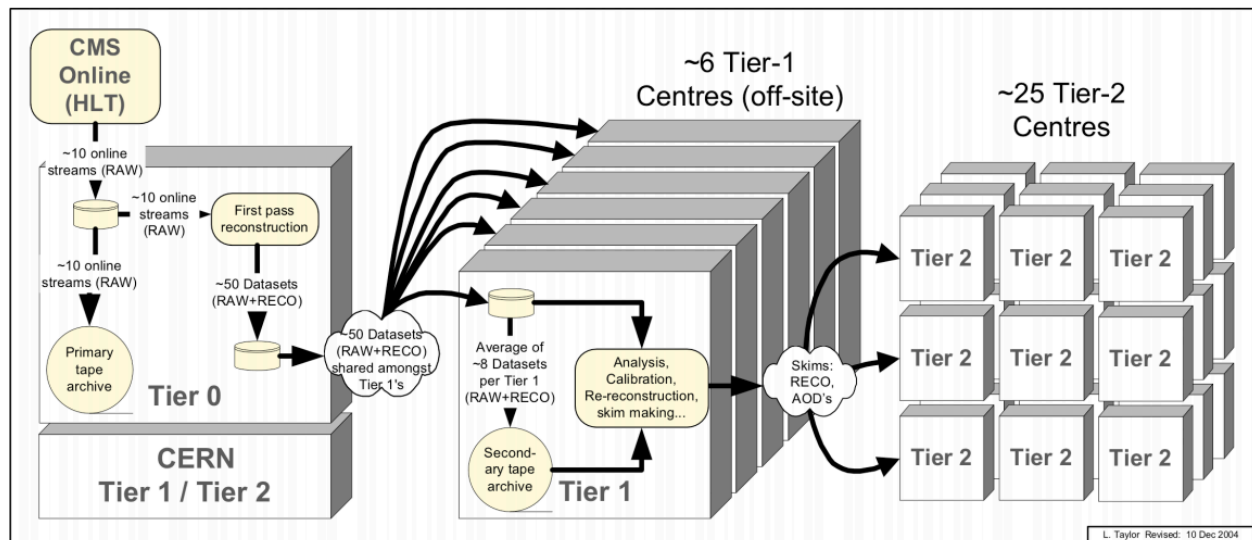
Figure 2.6: Schematic diagram of event data flow from the CMS detector [96].

The simulation, reconstruction, and analysis of CMS detector data is achieved through a global network of tiered computing facilities, as processing power and storage capacity required for all of these tasks exceeds that of any available computing resource.

At the top of the hierarchy, the Tier-0 facility is the common CMS computing facility at CERN. The Tier-0 facility stores and archives incoming collision data, and performs an initial "prompt"

reconstruction of the data. The raw and reconstructed data are copied to a network of Tier-1 computing facilities.

The Tier-1 facilities keep backup copies of both raw and reconstructed data. Additional reconstruction passes over the data may be later performed here as improved calibrations and algorithms become available. The large datasets are also skimmed by selecting events with physics objects of specific interest and storing these in smaller datasets which are copied to Tier-2 facilities.

The Tier-2 facilities keep selected portions of the datasets for physics analysis. These facilities also produce simulated datasets which may be transferred to the Tier-1 facility for wider distribution or analyzed at the site. For the special case of additional reconstruction of pPb and PbPb datasets, the Vanderbilt Tier-2 facility also provides this service.

A schematic diagram of the tiered computing model is shown in Fig. 2.6.

Chapter 3

Charged-Particle Trajectory Reconstruction

3.1 Overview of Charged-Particle Reconstruction in the CMS Detector

A charged particle in a uniform magnetic field follows a helical trajectory that can be completely described by five parameters. For the reconstruction of charged particles in the CMS tracker, the following five parameters are used [97]:

- p_{T}, the transverse momentum

- $\lambda = \cot(\theta)$, where θ is the polar angle

- ϕ, the azimuthal angle

- $d_0 = y_0 \cos\phi - x_0 \sin\phi$, where x_0 and y_0 are the transverse coordinates of the helix at the point of closest approach to the z-axis

- z_0, the longitudinal coordinate of the helix at the point of closest approach to the z-axis

A reconstructed charged-particle trajectory is referred to as a track. As a charged particle traverses a layer of tracker, it will strike the silicon at some angle, resulting in a signal in multiple pixels or strips. The signals that are thought to originate from a single charged particle are combined to form a cluster, which is then analyzed and corrected for the alignment of the detector to give a precise estimate of the charged-particle position as it traversed the detector, which is called a hit. Tracks are reconstructed by applying a pattern recognition algorithm to a sequence of hits on various layers of the silicon pixel and strip detectors.

Track reconstruction begins with the estimate of a helical trajectory called a seed which is formed from a set of three hits on different layers, or in some circumstances, hits on two layers and a measurement of where the collision specifically occurred. A Kalman Filter algorithm [98] is used to propagate this trajectory outward through additional layers, finding hits that are compatible with the trajectory estimate and updating the statistical fit. Tracks may then be selected based on quality criteria such as the number of hits on the track or the χ^2 goodness of fit.

By analyzing the tracks from many events, one can determine the position of the beamspot, which is the position of the beam in the collision region. This is a very precise region in the transverse plane, with a typical radius on the order of 16 μm. The track-based beamspot measurement algorithm can in ideal situations provide a measurement accurate to 2 μm [99]. The interaction region covers a much longer distance along the z-axis, in which collisions follow a Gaussian distribution centered around some point z_0, with a typical spread on the order of $\sigma_{z_0} \sim 7$ cm. Additionally, the beamspot will typically have some small gradient dy/dz and dx/dz. All of these parameters are measured in

an initial reconstruction of a subset of the total data sample and stored for subsequent use in the reconstruction of the full sample.

After tracks have been reconstructed for an event, the tracks and beamspot information may then be used to estimate the precise positions of the inelastic collisions which are called primary vertices. Vertex information may then be used to perform additional, more precise track reconstruction, which in turn may be used to better estimate the positions of the primary vertices.

3.1.1   Pixel and Strip Hit Reconstruction

The signals from the silicon strip detector are corrected for known offsets and common mode noise, and are zero-suppressed for pp and pPb collision runs previous to recording the event in RAW format. However, the non-zero suppressed silicon strip data was recorded in PbPb collision events where the increased particle flux creates a non-trivial amount of highly-ionizing particles (HIPs) which create large signals in the silicon strips through hadronic interactions. This can create a distortion in the common mode noise baseline affecting each 128-strip readout which must be calculated for each event and subtracted. This distortion can take the form of a slope or sag in the baseline across the 128 coupled strips, and an example is shown in Fig. 3.1. For such a case, a specialized baseline restoration algorithm is run which attempts to discern the signal peaks from the sloped baseline. The results of this algorithm in one instance is also shown in Fig. 3.1.
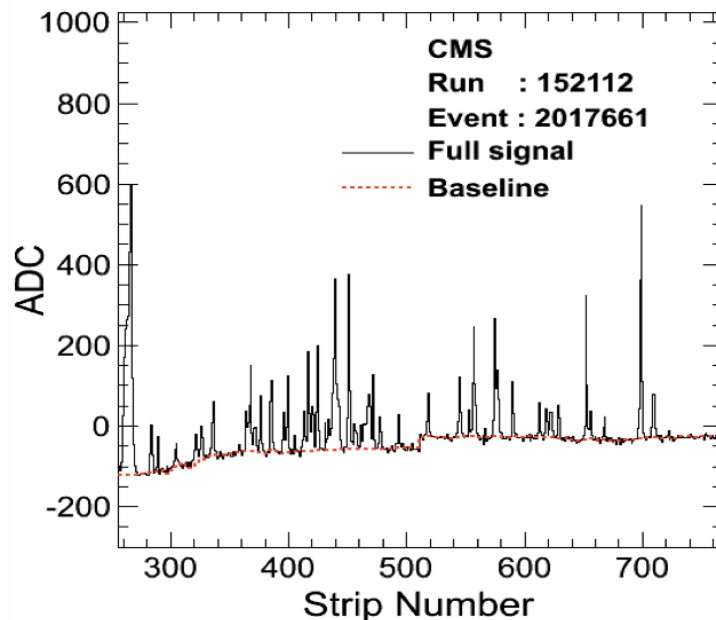
Figure 3.1: Pedestal-subtracted readouts from silicon strips as a function of strip number, where strips are numbered sequentially along the module. Four groups of 128 strips each corresponding to an APV25 readout chip are shown. The spikes are signals from charged particles, and the density of spikes indicates that the module has been affected by a HIP. The red dotted line indicates the reconstructed common-mode-noise baseline. The distortion of the baseline in the first group of strips has been determined by the baseline follower algorithm.

36

Silicon strip clusters are then reconstructed using a three-threshold algorithm. First, a seed strip is selected with a higher signal than neighboring strips, with a threshold of at least 3 times the signal-to-noise (S/N) ratio. Then strips are added to the cluster on either side of the seed if their signal exceeds a second threshold of at least 2 times the S/N ratio. The cluster grows outward until the next strip is below threshold. Once the cluster is complete, it is accepted only if the total signal size of the cluster exceeds a threshold of 5 times the quadratic sum of the noises of all the associated strips. The geometrical center of the cluster in the local coordinate system of the module is determined by taking the centroid of the signal heights in the strips. This is corrected for Lorentz shift of the charge carriers traversing the silicon due to the 3.8 Tesla magnetic field.

Pixel clusters are constructed by starting with a seed pixel containing a readout above a threshold more than 6 times the S/N ratio, and expanded outward by including adjacent pixels including corners that are above threshold. An example pixel cluster from a barrel module at high pseudorapidity is shown in Fig. 3.2.
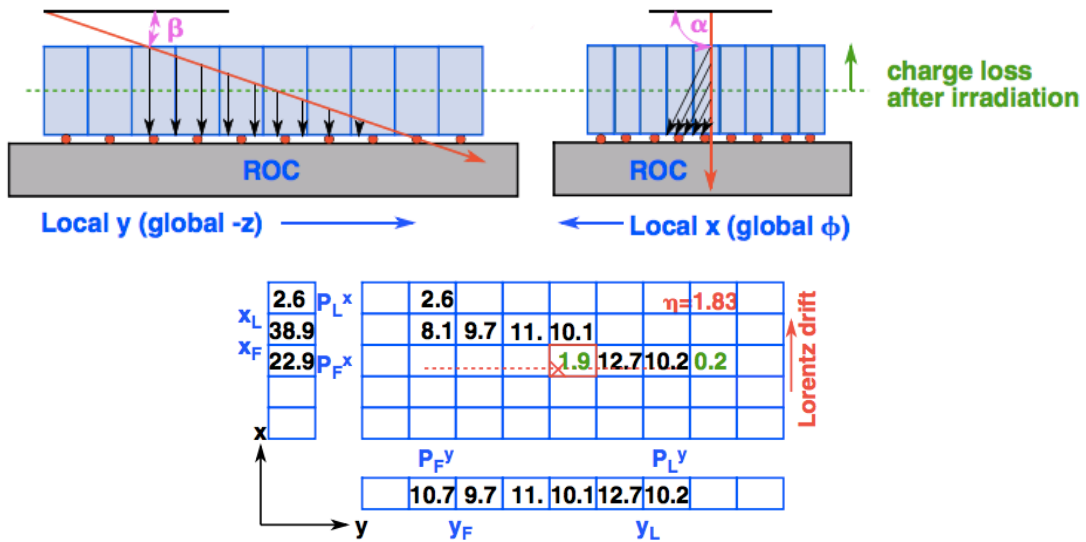
Figure 3.2: (Top) Cartoon of a charged particle (red arrow) traversing a barrel pixel detector module. The motion of deposited charge is indicated with the black arrows, with charge sharing between adjacent pixels along the global ϕ axis resulting from Lorentz-drift from the 3.8 Tesla magnetic field. (Bottom) Charge deposits from the same particle traversal as shown above in each pixel. The numbers represent thousands of electrons. Pixels shown in green are below threshold and are not added to the cluster. The corrected position of the hit, accounting for Lorentz-drift, is shown by the red X [90].

The coordinates of the pixel hit in the local (x, y) coordinate system of the pixel module may be determined separately for x and y using the following expressions [100]

$$x_{rec} = \frac{x_F + x_L}{2} + \frac{P_L^x - P_F^x}{P_L^x + P_F^x} \cdot \frac{W_{\text{eff}}^x(\cot \alpha)}{2} - \frac{\Delta_x}{2} \qquad (3.1)$$

$$y_{rec} = \frac{y_F + y_L}{2} + \frac{P_L^y - P_F^y}{P_L^y + P_F^y} \cdot \frac{W_{\text{eff}}^y(\cot \beta)}{2} - \frac{\Delta_y}{2} \tag{3.2}$$

where the $P_{F/L}^{x/y}$ are the signals in the first and last clustered pixels along the x or y axis; $x_{F/L}$ and $y_{F/L}$ are the x or y coordinates of the boundaries of the first and second or last and next-to-last pixels respectively; α and β are the angles of the charged-particle trajectory along the x or y axis respectively (see Fig. 3.2). The $\Delta_{x/y}$ are the maximum expected Lorentz-drift of the charge carriers along each axis due to the 3.8 Tesla magnetic field. The $W_{\text{eff}}^{x/y}$ are the effective charge widths in each direction given by

$$W_{\text{eff}}^x(\cot \alpha) = |T \cot \alpha + \Delta_x| - (x_L - x_F) \tag{3.3}$$

$$W_{\text{eff}}^y(\cot \beta) = |T \cot \beta + \Delta_y| - (y_L - y_F) \tag{3.4}$$

where T is the sensor thickness. The angles α and β may be estimated from a previously reconstructed track trajectory in a second-pass reconstruction, or approximated using the assumption of a particle originating from the nominal interaction point. To further correct and account for radiation damaged pixel sensors, a template fitting algorithm has been developed and is described in [90].

Once a silicon strip or pixel cluster is determined and the geometrical center of the cluster is identified in the local coordinate system of the detector module, this position must be translated to the global coordinate system of the CMS detector. For tracking resolution on the order of $10-20 \mu$m, this requires an equally precise knowledge of the alignment of the detector module within the CMS apparatus.

Determination of the tracker alignment proceeds from analyzing reconstructed tracks using prior estimates of the alignment starting from optical surveys of the tracker during commissioning, muon tracks from cosmic ray events [101] taken prior to LHC collisions, and continuing as additional collision data is taken. The components of the tracker system may shift over time, so the alignment is recalculated and recorded for each collision run.

The problem of alignment determination can be expressed in terms of a vector $\mathbf{p}$ representing the position of all tracker modules, and a collection of track trajectory vectors $\mathbf{q}_j$. For the ith hit on the jth track, the track residual is defined $\mathbf{r}_{ij} = \mathbf{m}_{ij} - \mathbf{f}_{ij}(\mathbf{p}, \mathbf{q}_j)$, where $\mathbf{m}_{ij}$ is the measured hit position on the module, and $\mathbf{f}_{ij}(\mathbf{p}, \mathbf{q}_j)$ is the trajectory impact point given the alignment $\mathbf{p}$. In the case of a pixel module, $\mathbf{r}_{ij}$ is a 2-component vector, and in the case of a strip module it is a scalar as the position measurement is only 1-dimensional. The alignment problem can then be expressed in terms of minimizing the objective function

$$\chi^2(\mathbf{p}) = \sum_j^{\text{tracks}} \sum_i^{\text{hits}} \mathbf{f}_{ij}^{\mathsf{T}}(\mathbf{p}, \mathbf{q}_j) \mathbf{V}_{ij} \mathbf{f}_{ij}(\mathbf{p}, \mathbf{q}_j), \tag{3.5}$$

38

where $\mathbf{V}_{ij}$ is a covariance matrix in the case of a pixel module, and a squared error in the case of a strip module. As there are tens of thousands of modules, each with six degrees of freedom, translational and rotational, $\mathbf{p}$ has on the order of 10^5 components. This complex minimization problem is approached using two algorithms. The MILLIPEDEII [102, 103] algorithm uses a global approach to minimize χ^2, and the "Hits and Impact Points" algorithm algorithm [104] applies a local approach. From the results of the two algorithms, the final alignment is determined.

With the predetermined alignment of the tracker modules, the position determination in the local coordinate frame of the module can then be reliably translated to the global coordinate frame of the CMS detector. This corrected position based on the signal cluster is the reconstructed hit which may now serve as an input to the pattern recognition algorithms described in the next sections.

3.1.2   Seed Generation, Pixel Tracking, and Pixel Vertexing

Track reconstruction begins with the generation of a seed which should provide an initial constraint on the five parameters of the helical trajectory with small enough uncertainties for the Kalman Filtering algorithm to continue refining the measurement by adding subsequent layers.

Seeds may be generated from hits on three tracker layers called a triplet or from a pair of hits on two layers. A set of layers used to generate the seeds is first chosen, along with some constraint for the origin of the helical trajectory. and minimum p_T of the reconstructed trajectory. This origin constraint may simply be a cylindrical region around the nominal interaction point aligned with the z-axis. The beamspot may be used to provide a tighter constraint in the transverse plane, but only a very loose constraint along the z-axis. If one or more primary vertices have been previously reconstructed, these may be used to provide a tight constraint in all directions.

The three layers of the pixel detector are the most accurate predictors of the helical trajectory, but some particles may not produce reconstructed hits on each layer of the detector that they traverse. These particles may be reconstructed either by seeding from pixel pair layers or using one or more layers of the strip detector as seeding layers.

In the case of triplet seeds from the pixel layers, the helical fit from these hits alone may be sufficiently precise as to accurately reconstruct a track. Such a track is called a pixel track. However, for high-p_T charged particles, the momentum resolution is not very good, owing to the small lever arm of the pixel detector. For a particle with $p_T = 10$ GeV/c, the uncertainty on the reconstructed momentum may be as high as 22%.

Pixel tracks may be of use in cases where the Kalman filtering algorithm is infeasible or not desired due to CPU time constraints. One such application is as a preliminary reconstruction pass in pp or pPb collisions. These tracks are then used to reconstruct primary pixel vertices, which in turn provide seeding constraints for later track reconstruction from pair seeds. Another application is for low-p_T tracking in the high occupancy environment of central PbPb collisions, where the occupancy of the strip detector prevents its use in better constraining the trajectory of the track. The algorithms used to constrain the helical parameters of pixel tracks and reconstruct pixel vertices are described in [105].

3.1.3 Combinatorial Kalman Filter

Kalman Filter algorithms have been used in charged-particle reconstruction in high energy and nuclear physics for over 20 years [106,107]. The approach is fast as there is linear rise in the number of computations with the number of measurements in the track, and it is robust in the presence of multiple scattering. At a given detector layer k, the state of a particle can be modeled as a vector x_k that encodes the position and momentum of the particle. The propagation of the particle to the next outward layer of the detector may be modeled by the linear equation

$$x_{k+1} = \mathbf{F}(k)x_k + \mathbf{G}(k)u_k + w(k) \tag{3.6}$$

where $\mathbf{F}(k)$ encodes the free propagation of the particle, $\mathbf{G}(k)u_k$ encodes the known physics input, in this case the response to the magnetic field, and $w(k)$ is a random source of Gaussian noise with a known variance, in this case multiple scattering. The measurement of the particle trajectory at layer k is modeled as

$$z(k) = \mathbf{H}x_k + v(k) \tag{3.7}$$

where $\mathbf{H}x_k$ represents the measurement process, and $v(k)$ a source of Gaussian noise in the measurement determination. At each stage, the state and convariance matrix of the particle trajectory is updated to include the new measurement information.

The algorithm used here is called a Combinatorial Kalman Filter (CKF), as in many cases there are multiple hits on the next layer, which are compatible with the current trajectory measurement within the error estimate. Beginning with a trajectory and position given from a seed, at each layer, the CKF will select all hits compatible with the trajectory, and subdivide them into a filter for each compatible hit. In addition, another filter will be initiated with the hypothesis that none of the compatible hits belong to the trajectory. These filters are propagated to the next layer where further subdivisions occur with each hit compatible with each trajectory estimate. To avoid a combinatorial explosion of filters as the algorithm propagates outward through several layers, the filters are pruned at each layer by removing trajectories with the highest χ^2 statistic.

In order to remove any biases introduced during the seeding stage, after the track has been built out of hits determined from the first pass of the CKF, the filter is re-initialized at the innermost hit with the trajectory estimate from the seeding, and propagated outward using only the hits determined from the first pass. During this second pass the positions of the hits can be re-evaluated using the angles determined from the first pass, resulting in a more accurate measurement.

Finally, the track is evaluated with a smoothing stage where the filter algorithm is initialized using the results of the previous stage at the outermost detector hit, with the covariance matrix scaled by a large value. The filter is then run inward to the innermost hit, at each layer combining the estimates from the previous outward filter to form a combined final estimate. This procedure yields optimal estimates at each layer. The CKF filter approach to track reconstruction is described in more detail in [108].

3.2 Charged-Particle Reconstruction in pp and pPb Collisions

The reconstruction process in pPb collisions is the same as in pp collisions. As the occupancy of the tracker in a 2013 pPb collision event is similar to a 2012 pp collision event where there may be up to 20 individual collisions in the event, no special adjustment to the tracking algorithm is required.

Three general stages of the pp and pPb reconstruction process are of interest to the spectra measurement. First, a set of pixel tracks are quickly reconstructed to then reconstruct pixel vertices as described in Sec. 3.1.2. These vertices are used in the seeding process for a robust, multi-stage algorithm described below. Finally, these tracks are used for the precise reconstruction of primary vertex positions.

3.2.1 Iterative Tracking

The pp track reconstruction strategy follows an iterative approach consisting of several repetitions of the CKF tracking algorithm on increasingly smaller subsets of the tracker hits. At the time of the pPb data taking, this procedure had expanded into seven iterations. Because the multiplicity of pPb collisions is similar to high-pileup pp events, this procedure could be applied to the pPb data without modification.

Each iteration differs in terms of the layers used to seed the trajectory reconstruction, the minimum p_T of the particle being reconstructed, and the spacial region from which the trajectory may originate. Between iterations, hits that belong to reconstructed tracks of good quality (designated *highPurity* as described below) are removed from the collection of hits used in the next iteration.

In the first two iterations, combinations of three pixel layers from the barrel and/or endcap (pixel triplets) are used as seeds. In both these iterations, the trajectory origin is constrained to a narrow region in the transverse plane around the beamspot. The third iteration utilizes seeds from pairs of pixel hits (pixel pairs) and further constrains the trajectory origin to be very close to a reconstructed pixel vertex along the z-axis. The fourth iteration again uses pixel triplets, but relaxes the origin constraint to search for tracks that may be displaced from the primary vertex, such as weak decay products of D or B mesons. The fifth iteration seeds from a mix of layers from both the pixel and strip detectors to find further displaced tracks or tracks that failed to leave a sufficient number of hits in the pixel detector to be reconstructed in the pixel detector. The final two iterations seed from the strip detector layers, and may reconstruct further displaced tracks. The seeding layers and trajectory origin constraints of the seven iterations are summarized in Table 3.1.

Following each iteration, tracks are filtered to remove tracks that are likely spurious or misreconstructed. Many such misreconstructed tracks can be removed by simply selecting tracks with a good χ^2 or primary vertex compatibility. In order to construct a collection of tracks with minimal contamination of misreconstructed trajectories (i.e. high purity) while maintaining a high efficiency, a more sophisticated set of selections can be applied in which the selection criteria are varied for each track, depending on the p_T, η, and number of layers in which a hit is found. Tracks containing

Iteration	Seeding layers	p_T cut (GeV/c)	d_0 cut (cm)	z_0 cut
Initial Step (0)	pixel triplets	0.6	0.02	4.0σ
Low-p_T Triplet Step (1)	pixel triplets	0.2	0.02	4.0σ
Pixel Pair Step (2)	pixel pairs	0.6	0.015	0.09 cm (*)
Detached Triplet Step (3)	pixel triplets	0.3	1.5	15 cm
Mixed Triplet Step (4)	pixel, TIB, TEC triplets	0.4–0.6	1.5	10 cm
Pixel-less Step (5)	TIB, TID/TEC pairs	0.7	2.0	10 cm
TOB-TEC Step (6)	TOB, TEC pairs	0.6	6.0	30 cm

Table 3.1: Parameters for the seeding of each tracking iteration used for pp and pPb collisions. The p_T cut refers to the minimum allowed p_T of the trajectory, and the d_0 and z_0 are the maximum transverse and longitudinal distances from the beamspot center, respectively. The (*) indicates that the maximum distance is with respect to the nearest reconstructed pixel track rather than the beamspot.

hits on many layers are less likely to be misreconstructed, so the selection criteria may therefore be loosened for those tracks. As described in [97], the following eight quantities are used to select *highPurity* tracks:

- normalized χ^2 of the trajectory fit

- transverse impact parameter d_{xy} with respect to the beamspot

- longitudinal impact parameter d_z with respect to the nearest reconstructed pixel vertex

- significance of the transverse impact parameter $d_{xy}/\delta_{d_{xy}}$, where $\delta_{d_{xy}}$ is the uncertainty on d_{xy} from the track fit

- significance of the longitudinal impact parameter d_z/δ_{d_z}, where δ_{d_z} is the uncertainty on d_z from the track fit

- number of tracker layers with a hit on the track, n_{layers}

- number of tracker "3D" layers with a hit on the track, specifically pixel layers or matched strip layers, $n_{\text{3D layers}}$

- number of layers missing hits between the first and last hit on the track, $n_{\text{lost layers}}$

These selections were optimized for each iteration in simulation, and the following formulas were determined as selection criteria with parameters α_i and β that may be adjusted for each iteration:

- Normalized $\chi^2 < \alpha_0 \cdot n_{\text{layers}}$

- $|d_{xy}| < (\alpha_1 \cdot n_{\text{layers}})^\beta \cdot \sigma_{d_{xy}}$

- $|d_z| < (\alpha_2 \cdot n_{\text{layers}})^\beta \cdot \sigma_{d_z}$

- $d_{xy}/\delta_{d_{xy}} < (\alpha_3 \cdot n_{\text{layers}})^\beta$

- $d_z/\delta_{d_z} < (\alpha_4 \cdot n_{\text{layers}})^\beta$

Iteration	n_{layers}^{min}	$n_{\text{3D layers}}^{min}$	$n_{\text{lost layers}}^{max}$	β	α_0	α_1	α_2	α_3	α_4
0	3	3	2	4	0.70	0.30	0.35	0.40	0.40
1	3	3	2	4	0.70	0.30	0.35	0.40	0.40
2	3	3	2	4	0.70	0.30	0.35	0.40	0.40
3_{VTX}	3	3	1	3	0.70	0.85	0.80	0.90	0.90
3_{TRK}	5	4	1	4	0.40	1.00	1.00	1.00	1.00
4_{VTX}	3	3	1	3	0.40	1.10	1.10	1.20	1.20
4_{TRK}	5	4	0	4	0.30	0.90	0.90	0.90	0.90
5	4	3	0	4	0.20	0.90	0.90	0.90	0.90
6	5	2	0	4	0.20	1.40	1.30	1.40	1.30

Table 3.2: Parameter values used for selecting *highPurity* tracks for each tracking iteration. In iterations 3 and 4, there are two sets of cuts applied, one focusing on track quality and the other focusing on vertex compatibility.

Here $\sigma_{d_{xy}}$ and σ_{d_z} are resolutions on d_z and d_{xy}, which are parametrized as $\sigma_{d_{xy}} = \sqrt{a^2 + (b/p_{\mathrm{T}})^2}$ and $\sigma_{d_z} = \cosh(\eta) \cdot \sigma_{d_{xy}}$, where $a = 30\,\mu\text{m}$ and $b = 10\,\mu\text{m} \cdot \text{GeV}/c$. Additionally, iteration-dependent selections on the minimum n_{layers}, $n_{\text{3D layers}}$, and maximum $n_{\text{lost layers}}$ are applied. The values of the parameters α_i and β and the minimum or maximum layer restrictions for each iteration are given in Table 3.2.

After each iteration, the tracks produced in the new iteration are merged with the tracks from previous iterations. This may result in the duplicate reconstruction of the same charged-particle trajectory. To remove these duplicates, tracks are checked for shared hits, and determined to be duplicate if they share more than 19% of the hits on the track with the least hits. In this case, the tracks are ranked by χ^2, adjusted with a "bonus" of 5 times the number of hits on the track subtracted, and a penalty of 20 times the number of lost hits on the track added. The track with the lower adjusted χ^2 is retained in the final collection before moving to the next iteration.

Note that the seeding layer paramers listed in Table 3.1, the *highPurity* selection parameters listed in Table 3.2, and even the total number of iterations are adjusted with new software revisions. The values presented here are those used for the reconstruction of the 2013 pPb collision events. Earlier CMS measurements of pp collisions that are referenced may use a different set of iterations and cuts. In the future, these parameters may continue to be adjusted.

3.2.2 Primary Vertex Reconstruction

After iterative tracking is performed, an additional primary vertex reconstruction is performed using the collection of iterative tracks. This process begins with a clustering algorithm that assigns tracks to vertex prototypes based on their z coordinate at the point of closest approach to the beamspot. The algorithm proceeds using a simulated annealing technique [109,110] to avoid local minima. The specific algorithm employed is a "mass constrained clustering" described in [111].

Let us designate the z positions of each track at their distances of closest approach to the beamspot as z_t, with a measurement uncertainty of σ_t. The number of vertex prototypes is variable, and

the z coordinate of each prototype may be given as z_v. Let c_{tv} represent a matrix element of the assignment of tracks to vertices, with $c_{tv} = 1$ if track t is assigned to vertex v, and 0 otherwise. To account for the fit quality and displacement of each track, an additional weight, p_t, may be assigned to each track. Additionally, a variable weight is given to each vertex, ρ_v, constrained so that $\sum_v \rho_v = 1$ The figure of merit is then

$$\chi^2 = \sum_{t,v} p_t \rho_v c_{tv} \frac{(z_t - z_v)^2}{\sigma_t^2}. \tag{3.8}$$

Finding the minimum χ^2 by exhaustively searching all possible combinations of c_{tv} is not feasible given that the both the number and values of the z_v can change. Instead the system is reinterpreted as an analogous thermodynamic system, with an energy $E_{tv} = (z_t - z_v)^2/\sigma_t^2$ associated with each pairing. Then χ^2 represents the total energy of the system. One may then look at a statistical ensemble of such systems, replacing the c_{tv} with probabilities p_{tv} that vary from 0 to 1, constrained so that $\sum_v \rho_v p_{tv} = 1$. Then the average energy of the ensemble is

$$\langle \chi^2 \rangle = \langle E \rangle = \sum_{t,v} p_t \rho_v p_{tv} E_{tv} \tag{3.9}$$

Then at thermal equilibrium at some temperature $T = 1/\beta$, the p_{tv} are given by the Boltzmann factor,

$$p_{tv} = \frac{e^{-\beta E_{tv}}}{\sum_{v'} \rho_{v'} e^{-\beta E_{tv'}}}, \tag{3.10}$$

At a given temperature, the p_{tv} is determined for the current vertex positions, z_v, and weights, ρ_v, are updated as

$$z_v \rightarrow \frac{\sum_t p_t \rho_v z_t/\sigma_t^2}{\sum_t p_t p_{tv}/\sigma_t^2}, \quad \rho_v \rightarrow \rho_v \frac{\sum_t p_t p_{tv}}{\sum_t p_t}. \tag{3.11}$$

This process proceeds iteratively until the change in the z_v and ρ_v is below some threshold, and equilibrium is then said to be established.

The algorithm proceeds in three phases. In the first, the initial condition is set with a very high temperature, equal p_{tv} values, and two vertices at some $z_0 \pm \delta$. Equilibrium is established and the temperature is reduced by a small step value, and the process of finding equilibrium repeats.

At any point during this first phase, if the temperature of a vertex drops below a critical temperature, T_c, defined as

$$T_c = 2 \frac{\sum_t \frac{p_t p_{tv}}{\sigma_t^2} \left(\frac{z_t - z_v}{\sigma_t} \right)^2}{\sum_t \frac{p_t p_{tv}}{\sigma_t^2}}, \tag{3.12}$$

the vertex is then split into two new vertices at $z_k \pm \delta$, and the algorithm continues.

If this splitting phase is allowed to continue to $T = 0$, the final result will be a large collection of vertices, each matched with a single track. The spacial resolution of the algorithm increases as temperature decreases, so at some point it is advantageous to halt splitting but continue to anneal to optimize the vertex positions. At some minimum temperature, chosen here as $T_{\min} = 4$, the algorithm enters a second phase where no further vertex splitting is allowed.

In the second phase, all vertex weights are set to be equal as they are only required for splitting determination, the position of the vertices is updated as before, and the Boltzmann factor for the p_{tv} is modified to include a "noise cluster" with an energy of assigning a track to the cluster being set at a fixed $E = \mu_0^2$, with no dependence on z_t. This serves to remove tracks that cannot be well associated with any vertex determined so far. The algorithm now proceeds down to $T = 0$.

The second phase makes the algorithm robust against noise from poorly reconstructed or displaced tracks, but does not prevent poorly determined vertices themselves. In the final assignment phase, each vertex is required to contain at least two tracks with an assignment weight of at least $p_{tv} = 0.9$. After removal of the poor vertices, the system is re-thermalized with the remaining vertices. Final track clusters are selected by associating tracks with $p_{tv} > 0.5$ to the prototype vertex v, with normalization of the p_{Tv} ensuring that tracks are not associated with multiple prototypes.

Once the track clustering algorithm is complete, a set of approximate vertex positions are determined longitudinally along the beamspot, each with a set of associated tracks. To better refine the position of these prototype vertices, an Adaptive Vertex Fitter (AVF) algorithm is employed, as described in detail in [112].

The AVF employs an iterative re-weighted least squares approach that down weights tracks according to their standardized distance from the vertex. In order to avoid local minima, the weight function is determined with a "temperature" according to a geometric annealing schedule, with a Kalman filter formalism used to determine the vertex position at each iteration.

The result of this final algorithm is a set of primary vertices with a precise 3D estimate of the position of the collision. While this algorithm generally performs well in determining each primary collision vertex in an event, the clustering scheme will reconstruct a single collision vertex as a split pair of vertices about 10% of the time in pPb collisions.

3.3 Charged-Particle Reconstruction in PbPb Collisions

Track reconstruction in central PbPb collisions is a challenge due to the extremely high occupancy of the tracker pixels and especially strips. In Fig. 3.3, an event display of a central PbPb collision event taken in November 2010 is shown, which provides a sense of the sheer number of charged-particle trajectories that must be unambiguously reconstructed.

As shown in Fig. 3.4, early simulations of central PbPb collisions predicted that the channel occupancy, defined as the fraction of pixels or strips above threshold in each layer, could reach as high as 1% in the inner pixel layer and 30% in the inner strip layer. This presents a challenge for the seeding of track reconstruction, as the number of potential triplet seeds from the pixel layer
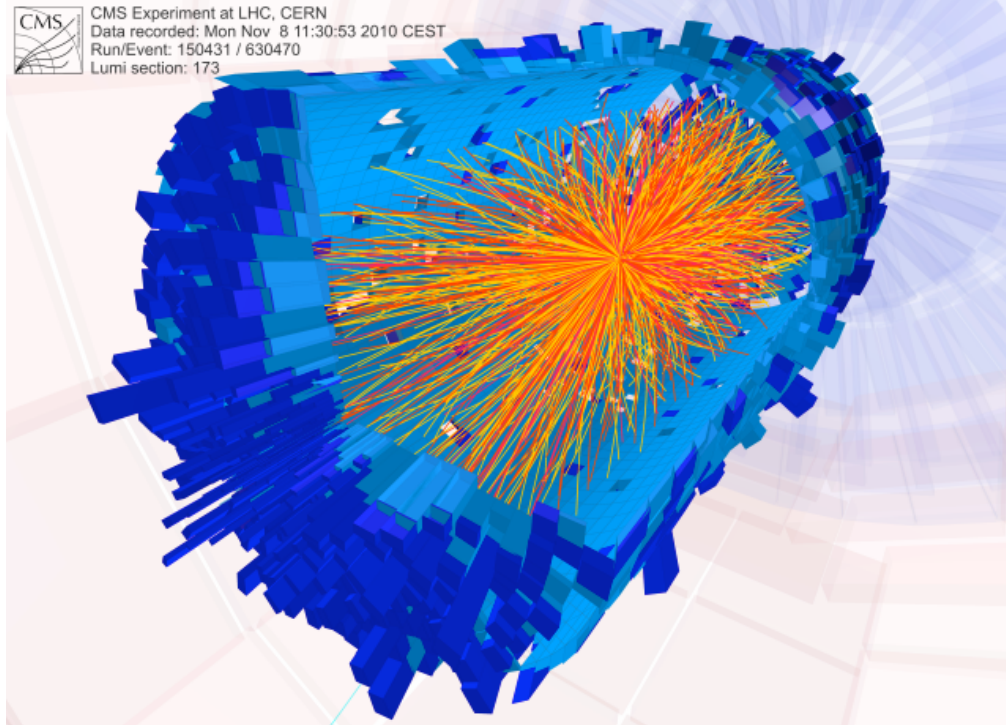
Figure 3.3: Event display of a central PbPb collision event taken in November of 2010. The red and orange lines represent reconstructed tracks with $p_T > 0.9\,\text{GeV}/c$ and the blue boxes represent calorimeter towers. Note that the bulk of the charged-particle production occurs with $p_T < 0.9\,\text{GeV}/c$, so that only a small fraction of all charged-particle trajectories are actually shown [113].

may easily number in the millions even given a reasonable beamspot constraint and minimum p_T threshold. The high occupancy in the inner layers of the strip detector presents a challenge for the CKF algorithm, as the number of compatible hits on the inner strip layer may be very large for a given trajectory estimate.

The iterative reconstruction algorithm used in pp and pPb collisions is not feasible for central PbPb events due to the CPU and memory limitations of the currently available hardware. A different algorithm has been developed in order to overcome the issue of seeding, and for low-p_T reconstruction, the issue of strip occupancy.

In order to begin tracking with a reasonably small number of seeds pixel triplet layers are used in conjunction with an estimate of the primary vertex position along the z-axis. This estimate must be determined without the use of a full set of reconstructed tracks, using a procedure described in Sec. 3.3.1. The minimum trajectory p_T is set to 0.9 GeV/c. These constraints result in a small enough number of seeds with a small enough window of compatible hits on the first strip layer to be propagated successfully using the CKF algorithm.

To reconstruct tracks with $p_T < 0.9\,\text{GeV}/c$, a second tracking iteration is performed with pixel triplet seeds. These tracks are not propagated with the CKF filter, but fitted as pixel tracks as described in Sec. 3.1.2. These two iterations are then merged into a final collection described in
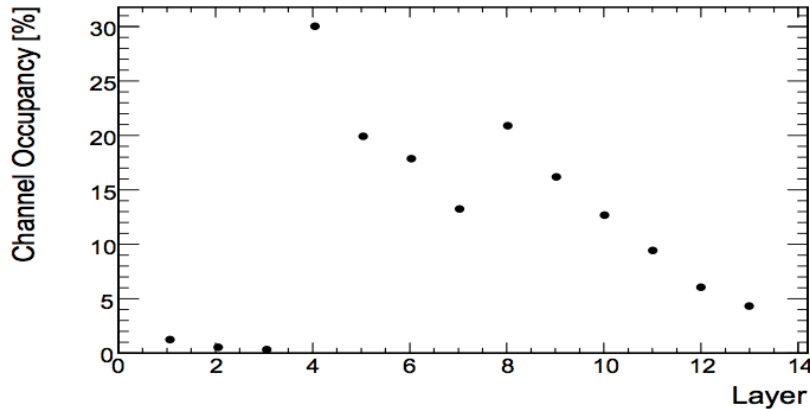
Figure 3.4: The expected occupancy in the barrel pixel detector (layers 1-3), TIB (layers 4-7), and TOB (layers 8-13) expressed as the percent of pixels or strips above threshold for central PbPb collision events [114]. Note that this result is from an early simulation used for the initial determination of a PbPb tracking procedure where particle multiplicities of $dN/dy \sim 3000 - 3500$ were expected. Analysis of the 2010 PbPb collision data at $\sqrt{s_{NN}} = 2.76$ TeV revealed that at midrapidity $dN/d\eta \sim 1600$ for central events [89]. Particle multiplicities may be higher in the future if PbPb collision data is taken at the design energy of $\sqrt{s_{NN}} = 5.5$ TeV.

Sec. 3.3.4.

As of this writing, the tracking reconstruction sequence for PbPb collisions is still undergoing development. The specific two-iteration sequence described here is not common to all published CMS PbPb collision results, but was developed specifically for the measurement of the charged-particle spectrum at low-p_T. Other measurements performed over the same time frame used different sequences to match their physics goals. The following description should not be taken as a general reference for all CMS measurements of the 2010 PbPb collision data.

3.3.1   Primary Vertex Reconstruction

The primary vertex PbPb collision events is reconstructed assuming that only one PbPb collision is present in each event, which is reasonable given the bunch intensity for the PbPb taken in 2010 and used in this measurement. Before any track reconstruction can begin, which is typically required for any vertex reconstruction, an approximate estimate of the vertex position must be determined to limit the origin radius of the seeds along the z-axis.

This first rough approximation is performed by reconstructing a cluster vertex using only the information from the first layer of the pixel barrel detector. The cluster vertex is estimated using a simple algorithm that looks at the length of each pixel cluster along the detector z-axis (see Fig. 3.2 for an example cluster length), and determining if the length of the cluster is compatible with the angle of a charged particle originating from some point z_0 along the z-axis. Starting at $z_0 = -30$

cm from the nominal interaction point, and incrementing by 0.1 cm up to $z_0 = 30$ cm, the algorithm calculates the number of pixel clusters with a length compatible with a particle originating at z_0. The value of z_0 with the most compatible clusters is taken as the initial estimate of the primary vertex position.

Once the cluster vertex is produced, an origin region for pixel triplet seeds is created with a loose constraint around the z-position of the cluster vertex, and a minimum p_T of 0.7 GeV/c. Even with the longitudinal constraint, the number of possible seeds is very large. The number of compatible trajectory seeds is estimated based on the number of pixel barrel hits, and the kinematic region for reconstruction of these first pixel triplet seeds is accordingly narrowed in η around midrapidity in order to keep the number of resulting trajectories small. These trajectory seeds are reconstructed and called proto-tracks.

Proto-tracks are only used in order to obtain a better estimate of the primary vertex position and then discarded. Even with the prior constraints, in a central event there may be many proto-tracks of poor quality with will adversely affect the AVF algorithm described in Sec. 3.2.2. In order to get a more refined estimate of the vertex position a simple median vertex is reconstructed by taking the median of the z coordinates of all proto-tracks at their distance of closest approach to the z-axis. A subset of the proto-tracks are then selected for compatibility with the median vertex.

These selected proto-tracks are then used as inputs to the AVF algorithm. If the AVF algorithm returns more than one vertex, the "best" vertex is selected based on the sum of the p_T^2 of the proto-tracks associated with the vertex. If the AVF algorithm failed to produce a vertex with reasonable statistical errors, the median vertex is used instead. If both the AVF and median algorithms fail, the beamspot position is taken as the collision vertex. In this manner, every reconstructed PbPb collision event is guaranteed to have exactly one reconstructed primary vertex. Once this primary vertex is selected, track seeding can begin in a tightly defined region around the vertex.

3.3.2   First Iteration CKF Tracking

Trajectory seeds for the first iteration are built from pixel triplets, and are constrained to trajectories with $p_T > 0.9$ GeV/c. The transverse distance of closest approach to the beamspot must be within 0.1 cm, and the longitudinal distance of closest approach to the selected primary vertex must be within 0.6 cm.

Before initiating the CKF algorithm, trajectory seeds with a transverse displacement from the selected primary vertex with a statistical significance of more than 6σ, and longitudinal distance of closest approach with an absolute value of more than 0.2 cm.

The selection of good quality tracks was determined as part of the analysis effort for the measurement of charged-particle spectra and is discussed in Sec. 4.4.1.

3.3.3 Second Iteration Pixel Tracking

The second iteration of tracking reconstruction is also built from pixel triplets, except that the constraint is relaxed to include trajectories with p_T as low as $p_T = 0.2\,\mathrm{GeV}/c$. As with the first iteration, the transverse distance of closest approach to the beamspot must be within 0.1 cm, and the longitudinal distance of closest approach to the selected primary vertex must be within 0.6 cm.

Unlike the iterative approach in pp and pPb collisions, the hits associated with the high-quality tracks from the previous iteration are not removed. This is not done as the highly occupied detector may contain merged hits from multiple charged particles, and also because this removal did not result in appreciably improved performance, as the majority of hits are not associated with quality tracks from the first iteration.

These trajectories are not propagated through the strip layers using the CKF algorithm, as the low momentum of the trajectories corresponds to a wide area of compatible hits in a highly occupied detector. As a result, the momentum of these trajectories is poorly determined at high-p_T.

To better constrain the momentum of the trajectory, it is refitted using the transverse position of the beamspot as a fourth constraint along with the three pixel hits. This overdetermines the path of a helix, so one can meaningfully discuss a χ^2 goodness of fit for the trajectory along the three hits and the beamspot.

As before, the selection of good quality tracks was determined as part of the analysis effort for the measurement of charged-particle spectra and is discussed in Sec. 4.4.1.

3.3.4 Merging Procedure for Track Collections

The merging of the track collections from the two iterations is performed as with the pp and pPb iterative tracking, where tracks are considered duplicate if they share more than 19% of their hits. In this case the resolution between duplicates is trivial – a first iteration CKF track is always preferred to a pixel track.

Another important difference is that at high-p_T it is difficult or impossible to distinguish between poor and good quality tracks. For this reason, pixel tracks with $p_T > 1.5\,\mathrm{GeV}/c$ are removed from the final collection. Similarly, first iteration CKF tracks with $p_T < 1.2\,\mathrm{GeV}/c$ are also removed.

Chapter 4

Tracking Performance

4.1 Overview of Full Detector Simulation

In order to properly measure the p_T spectrum of charged particles produced in pPb and PbPb collisions, it is necessary to precisely determine the efficiency of the track reconstruction and correct the measured sample for this inefficiency, as well as performing other corrections. The primary tool for understanding the performance of these complex reconstruction algorithms is a detailed simulation of the interaction of charged particles with the detector, starting from a realistic estimate of the particles produced in the collision, and ending with a full simulated detector readout which may be reconstructed using the same software as in collision events.
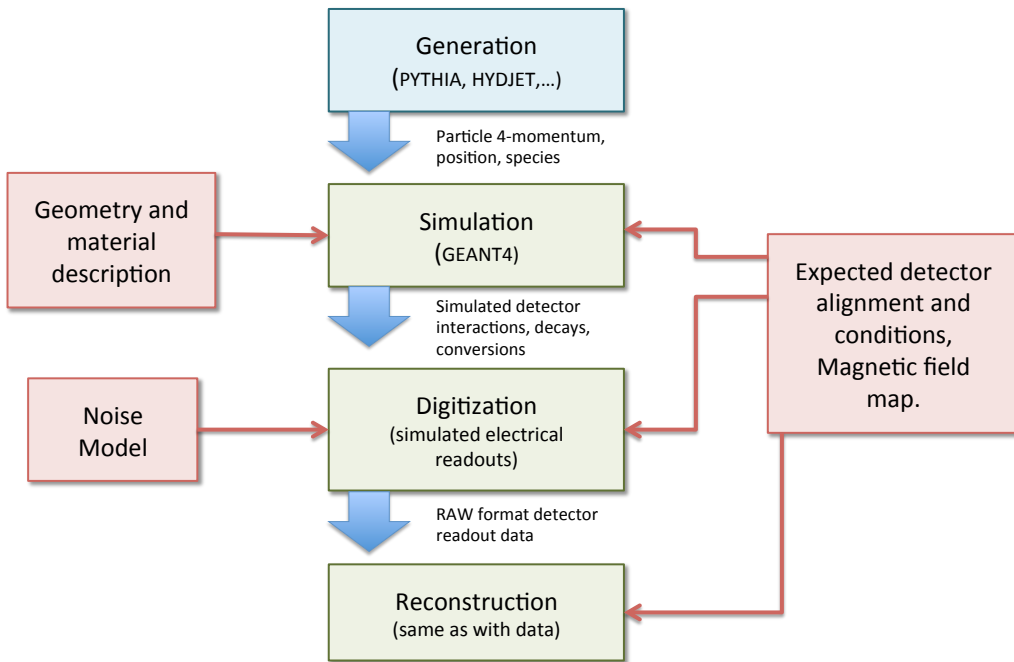
Figure 4.1: Block diagram of the full CMS detector simulation workflow.

The simulation workflow can be divided into four stages, as summarized in Fig. 4.1. The simulation begins with the generation step in which one of several possible MC event generators is used to create a list of particles produced in a random collision. These particles are then propagated to a simulation step where the particles are propagated through a detailed model of the detector material and the interactions with the material are recorded. The digitization step takes these simulated interactions, and uses a realistic model of the detector response and noise to produce a simulated readout of the detector elements. This simulated data is stored in the same RAW format as actual

collision data. Finally, the reconstruction step utilizes the same software modules and configuration to produce simulated physics objects for analysis and comparison to collision data.

Any available event generator can in principle be used in the CMSSW framework for the generation of events, provided that output of the MC event generator is stored in the HEPMC format, encoding the 4-momentum, position, and species of each particle produced in the collision [115]. The simplest possible generator is called a particle gun which generates a specific number of particles of a given species with either fixed or randomly distributed kinematic variables, (p_T, η, ϕ). For realistic generation of pp collisions, the PYTHIA generator [116] (version 6.423, tune Z2 [117] and ProQ20 [118]) is used. PbPb collisions are generated with the HYDJET [119] generator, and pPb collisions are produced using HIJING [120], version 1.383 [121], and EPOS [122]. The physics models of the generators referenced here are described further in Appendix C.

The simulation framework is flexible enough to accept input from multiple generators and combine results. One may, for example, simulate events with multiple collisions by embedding the results of one generator output into a simulated event with another. Hard-scattering events resulting in very high p_T jets may be simulated in pPb collisions by embedding the appropriate PYTHIA di-jet event into minimum-bias HIJING background events. It is even possible to mix data and simulation; for example one may take the RAW formatted data from a PbPb collision event, and add a generated pion from a particle gun which is combined to form a mixed readout of simulation and data.

The simulation step is based on a complete GEANT4 [95] based map of all of the detector material, including all structures, electronics, and cooling systems. A precise map of the magnetic field is required, along with a reasonably accurate estimate of the alignment and calibration of all detector components. Additionally, for the digitization step, a detailed understanding of non-functioning components, such as noisy or dead channels, must be provided to match the data taking period for which the simulation is to represent. The development of this simulation architecture is an ongoing effort of incremental refinements, starting before the construction of the detector, extending to the 2008 calibration using cosmic rays [123], and continuing through each collision data taking period.

The result of the full simulation is RAW formatted data that may be reconstructed using the same software code as with the collision data. The properties of reconstructed tracks from the early 2009 pp collisions were compared in detail to the results of PYTHIA simulations, and were found to be in good agreement [124].

4.2 Determination of Tracking Performance

Tracking performance is evaluated in simulation by attempting to match reconstructed tracks to simulated charged-particle trajectories. Each reconstructed hit on the silicon tracker is associated in simulation with the GEANT4 interaction which produced the detector signal, which is called a simhit. A track is associated with a simulated charged particle if over 75% of the hits on the track are matched with the simhits produced by the charged particle.

The absolute efficiency is defined as the fraction of simulated charged particles which are successfully matched to a track. The acceptance is defined as the fraction of simulated charged particles that

leave hits on any combination of tracker layers used to produce trajectory seeds. The algorithmic efficiency is the probability that a simulated charged particle is matched to a track, given that it left hits on some combination of tracker layers used to produce trajectory seeds. The absolute efficiency is therefore the product of the acceptance and the algorithmic efficiency.

The misreconstruction, or fake rate, is the fraction of reconstructed tracks that are not matched with a simulated charged particle. The multiple reconstruction fraction is the fraction of simulated particles that are matched to more than one reconstructed track. From the fractional difference between the p_T of the reconstructed track and simulated particle, $\delta p_\mathrm{T}/p_\mathrm{T} = (p_{t,\mathrm{rec}} - p_{t,\mathrm{sim}})/p_{t,\mathrm{rec}}$, the momentum resolution is defined as the root mean square of the distribution $\delta p_\mathrm{T}/p_\mathrm{T}$ taken from all reconstructed tracks, $\sigma(\delta p_\mathrm{T}/p_\mathrm{T})$.

Reconstructed charged particles may include the results of weak decays of reasonably long lived particles and electrons from photon conversions in the tracker material or beampipe. For this reason, it is important to define precisely what one means by charged particles produced in the initial collision of protons or ions. For the CMS measurements of charged-particle spectra presented here and in [27, 125], a primary particle is defined as a particle produced in the interaction of the incoming protons or ions with a mean proper lifetime of at least 1 cm/c. Additionally, the decay products of any particle with a mean proper lifetime of less than 1 cm/c are considered primary particles. Any particle failing this definition is called a secondary particle.

This definition of a primary particle effectively includes weakly decaying strange and light quark mesons and baryons such as K_s^0s, Λ^0s, $\Xi^\pm$s, and Ω^-s, but not their decay products. It does not include strong and electromagnetically decaying particles such as $\Delta^\pm$s, weakly decaying heavy flavor mesons such as Ds and Bs, but it *does* include their decay products. The definition is convenient for comparison to event generators such as PYTHIA, which may be set to decay all particles with $c\tau < 1$ cm, and will therefore result in a list of primary charged particles in accordance with this definition.

4.3   Validation of the Efficiency Determination

Even though the properties reconstructed tracks from pp collisions agreed with results of PYTHIA simulations [124], further studies were performed on 2010 pp collision data at $\sqrt{s} = 7$ TeV in order to validate that the efficiency measured in simulation accurately represents the probability of correctly reconstructing the trajectory of a charged particle in collision events [126]. Two of these studies relevant to the reconstruction of charged hadrons which dominate the charged-particle spectrum are described below.

The first method is an event mixing technique called embedding where a simulated charged pion is combined with a collision event, by matching the origin of the pion trajectory to the reconstructed vertex. The detector readouts are modified to include the simulated signals from both the collision event and the simulated pion. The reconstruction efficiency of the embedded pion is then determined and compared with the efficiency of a charged pion embedded into PYTHIA simulation. As shown in Fig. 4.2, the results of pion embedding into either simulation or collision data agree within 1%.

The second method involves the reconstruction of D^0 mesons in pp collision data using two decay
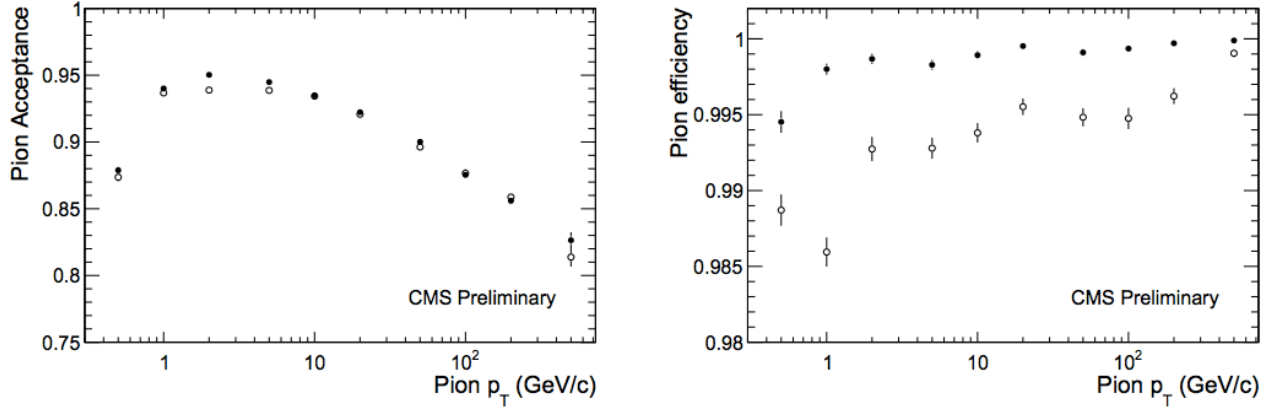
Figure 4.2: The acceptance (Left) and algorithmic efficiency (Right) for pions embedded in minimum-bias pp collision events (open circles) and simulated events (closed circles) [126].

channels. One channel is $D^0 \rightarrow K^- \pi^+$, and the other is $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$. These mesons are reconstructed through the decay chain $D^{*+} \rightarrow D^0 \pi^+$ to provide a common production source. The ratio of the branching fractions as measured should then be given as

$$\mathcal{R} = \frac{N_{K3\pi}}{N_{K\pi}} \cdot \frac{\epsilon_{K\pi}}{\epsilon_{K3\pi}}, \tag{4.1}$$

where the $N_{Kx\pi}$ are the number of D^0 mesons from each channel, and the $\epsilon_{Kx\pi}$ are the reconstruction efficiencies from each channel. Assuming that the pion efficiency measured in simulation is the same as the efficiency in data, $\mathcal{R}$ should be equal to the actual ratio of the branching fractions, which is estimated by the world average of $\mathcal{R}(PDG) = 2.08 \pm 0.05$ [1]. The ratio $\epsilon_{K3\pi}/\epsilon_{K\pi}$ should be equal to the square of the pion reconstruction efficiency, and so

$$\frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}} = \sqrt{\frac{\mathcal{R}}{\mathcal{R}(PDG)}} \tag{4.2}$$

where ϵ_{data} is the tracking efficiency of pions in data, and ϵ_{MC} is the tracking efficiency of pions in simulation. The final result of this technique was $\epsilon_{\text{data}}/\epsilon_{\text{MC}} = 1.007 \pm 0.034 \pm 0.014 \pm 0.012$, where the first uncertainty is statistical, the second systematic, and the third for the value of $\mathcal{R}(PDG)$ [126].

4.4 Tracking Performance in PbPb Collisions

The determination of tracking performance in PbPb presented a number of new challenges relative to pp collisions. Since the iterative tracking scheme was developed specifically for the 2010 PbPb collision data, selection criteria for quality tracks had to be developed by determining variables to be used for selection, and then optimizing the values of these variables in order to maximize

efficiency and minimize fake rate.

The efficiency and fake rate in PbPb was found to vary considerably with event centrality, in addition to p_{T} and η. Therefore the efficiency had to be determined separately for each event centrality class used in the measurement. These classes are described in Sec. 5.4. The reason for this dependence is the high occupancy of the tracker modules.

Due to the high occupancy environment, an additional embedding study of simulated pions into PbPb collision data was required to further validate that the simulation accurately represented the detector readouts in this novel environment.

4.4.1 Determination of Track Selection Criteria

Five variables were used in the determination of selection criteria for the two iterations of PbPb tracks:

- The number of hits associated with each track, N_{hit}

- The normalized statistical χ^2 goodness of fit parameter divided by the number of associated hits, $\chi^2/\mathrm{n.d.o.f.}/N_{\mathrm{hit}}$, or in the case of the pixel tracks, the normalized statistical χ^2 parameter, $\chi^2/\mathrm{n.d.o.f.}$.

- The percent uncertainty on the measured p_{T}, $p_T^{\mathrm{error}}/p_{\mathrm{T}}$, where p_T^{error} is taken from the 5D covariance matrix of the track trajectory fit.

- The approximate normalized transverse distance of closest approach between the track and the selected primary vertex, $d_{xy}/\sigma_{d_{xy}}$, where d_{xy} is calculated using a straight-line approximation of the track trajectory, and $\sigma_{d_{xy}} = \sqrt{\sigma_{d_{xy}}^{TRK} \cdot \sigma_{d_{xy}}^{TRK} + \sigma_{d_x}^{VTX} \cdot \sigma_{d_y}^{VTX}}$, where $\sigma_{d_{xy}}^{TRK}$ is the $\sigma_{d_{xy},d_{xy}}$ element of the 5D covariance matrix of the track trajectory fit, and $\sigma_{d_y}^{VTX}$ and $\sigma_{d_x}^{VTX}$ are the statistical errors in the x and y coordinates of the vertex fit.

- The approximate normalized longitudinal distance of closest approach between the track and the selected primary vertex, d_z/σ_{d_z}, where d_z is calculated using a straight-line approximation of the track trajectory, and $\sigma_{d_z} = \sqrt{\sigma_{d_z}^{TRK} \cdot \sigma_{d_z}^{TRK} + \sigma_{d_z}^{VTX} \cdot \sigma_{d_z}^{VTX}}$, where $\sigma_{d_z}^{TRK}$ is the σ_{d_z,d_z} element of the 5D covariance matrix of the track trajectory fit, and $\sigma_{d_z}^{VTX}$ is the statistical error in the z coordinate of the vertex fit.

Note that the last two variables, $d_{xy}/\sigma_{d_{xy}}$ and d_z/σ_{d_z} can reasonably be interpreted as the approximate statistical significance of the hypothesis that the track did not originate from the selected vertex, expressed in units of standard deviations. This selection variable is useful to remove not only fake tracks, but reconstructed secondary particles as well.

For determining good quality tracks from the first iteration, the variables N_{hit}, $\chi^2/\mathrm{n.d.o.f.}/N_{\mathrm{hit}}$, and $p_T^{\mathrm{error}}/p_{\mathrm{T}}$ were chosen. After studying the track distribution and fake rate as a function of these variables, as shown in Fig. 4.3, values were selected. Similarly, the variables d_z/σ_{d_z} and $\chi^2/\mathrm{n.d.o.f.}$ were chosen to select good quality tracks from the second iteration pixel tracks, and studied as shown in Fig. 4.4. Additionally, pixel tracks are required to pass a cluster shape filter

requiring that the pixel hits have a longitudinal length that is reasonably compatible with the η of the trajectory and the z-position of the primary vertex. The variable $d_{xy}/\sigma_{d_{xy}}$ was not useful as the beamspot constraint of the pixel track fit used in this iteration forced this value to be nearly or identically zero. The selection cuts used for each iteration are listed in Table 4.1.
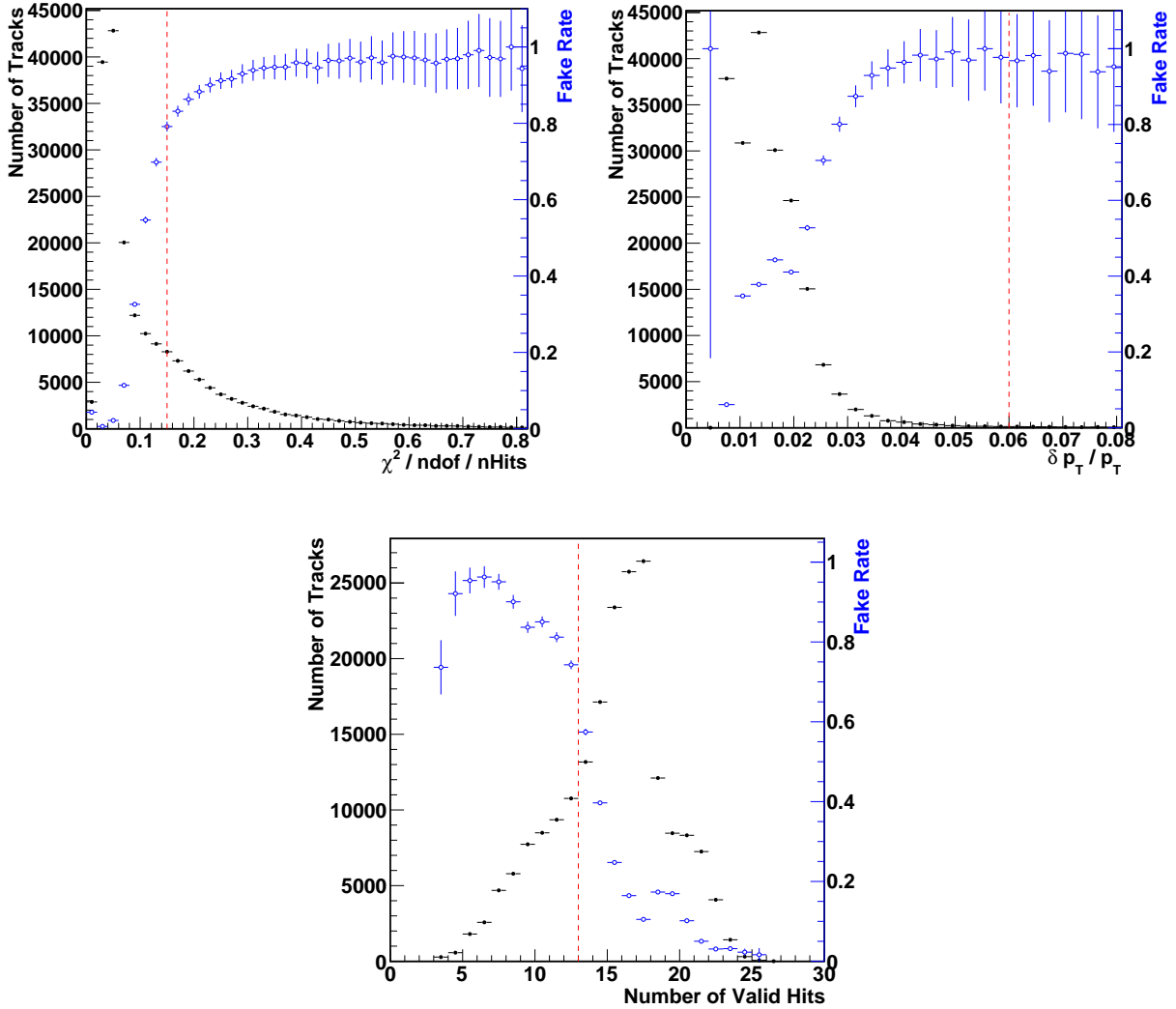
Figure 4.3: The distribution of reconstructed 1st iteration PbPb CKF tracks (black), and the fake rate (blue) in 0-10% central events simulated with HYDJET as a function of $\chi^2/\mathrm{n.d.o.f.}/N_{\mathrm{hit}}$ (Top Right), p_T^{error}/p_T (Top Left), and N_{hit} (Bottom). The dotted red line indicates the value of the parameter used to select quality tracks.

4.4.2 Determination of Track Efficiency and Fake Rate

The efficiency, fake rate, secondary fraction, and multiple reconstruction fraction of selected PbPb tracks was determined using 100,000 minimum-bias events from a HYDJET simulation. Even with

Selection Variable	1st Iteration (CKF tracks)	2nd Iteration (pixel tracks)
N_{hit}	≥ 13	–
$\chi^2/\mathrm{n.d.o.f.}/N_{\mathrm{hit}}$	< 0.15	–
$\chi^2/\mathrm{n.d.o.f.}$	–	< 40
$p_T^{\mathrm{error}}/p_{\mathrm{T}}$	< 0.06	–
$d_{xy}/\sigma_{d_{xy}}$	–	–
d_z/σ_{d_z}	–	< 10
cluster shape filter	no	yes

Table 4.1: Selection criteria used in determining good quality reconstructed tracks in each iteration of the PbPb tracking algorithm.

the division of the sample into centrality classes as described in Sec. 5.4, there are still at least 5,000 events for each subsample, providing ample statistics for determining these properties up to $p_{\mathrm{T}} = 6\,\mathrm{GeV}/c$.

The tracking performance in 0-10% centrality events is shown in Fig. 4.5. The secondary and multiple reconstruction fraction were found to be extremely small, $\sim 0.5\%$. The tracking efficiency and fake rate vary significantly with both p_{T} and η. In Fig. 4.6, the efficiency, fake rate, and momentum resolution is shown as a function of p_{T} in classes of centrality and η. The momentum resolution is below 8% in all classes, and is significantly better at high-p_{T} where the CKF tracks comprise the collection.

Validation that the tracking efficiency measurement in simulation is accurately determined was accomplished using a set of simulated pions embedded into central PbPb collision events. In Fig. 4.7, the efficiency of embedded charged pions is compared with the overall charged-particle efficiency from the HYDJET simulation. For $p_{\mathrm{T}} \lesssim 2\,\mathrm{GeV}/c$, the charged-particle spectrum is dominated by pions and so the efficiencies should be similar. The result of this embedding study shows agreement between the embedding and the HYDJET simulation. From this result as well as those described in Sec. 4.3, one may conclude that the simulation accurately describes the charged-particle efficiency in PbPb collisions.

The tracking efficiency was found to vary with the species of charged particle being reconstructed, as shown in Fig. 4.8. This is expected for several reasons. The antiprotons have a higher cross section for hadronic interactions with the silicon tracker material due to potential annihilation and so the efficiency is expected to be lower at all p_{T}. At very low-p_{T}, the heavier particles will lose more energy through multiple scattering and are more difficult to reconstruct. Due to these effects, the particle composition of the PbPb collisions must be roughly understood, as any difference in the particle composition between the HYDJET simulation and the PbPb collisions may yield a systematic bias on the efficiency determination.
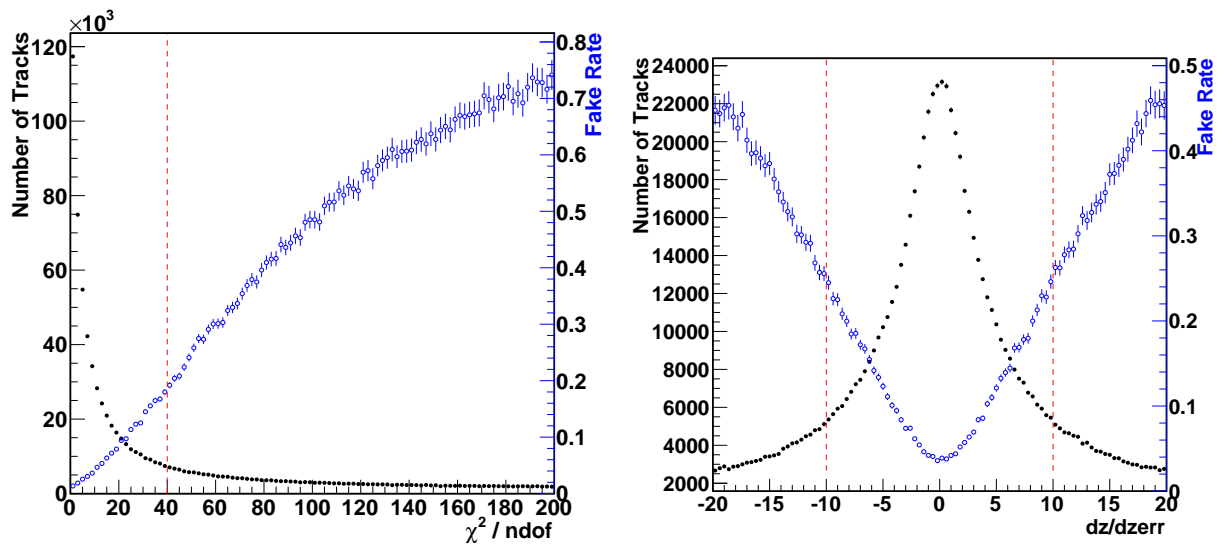
Figure 4.4: The distribution of reconstructed 2nd iteration PbPb pixel tracks (black), and the fake rate (blue) in 0-10% central events simulated with HYDJET as a function of $\chi^2/\text{n.d.o.f.}$ (Right), and d_z/σ_{d_z} (Left). The dotted red line indicates the value of the parameter used to select good quality tracks.
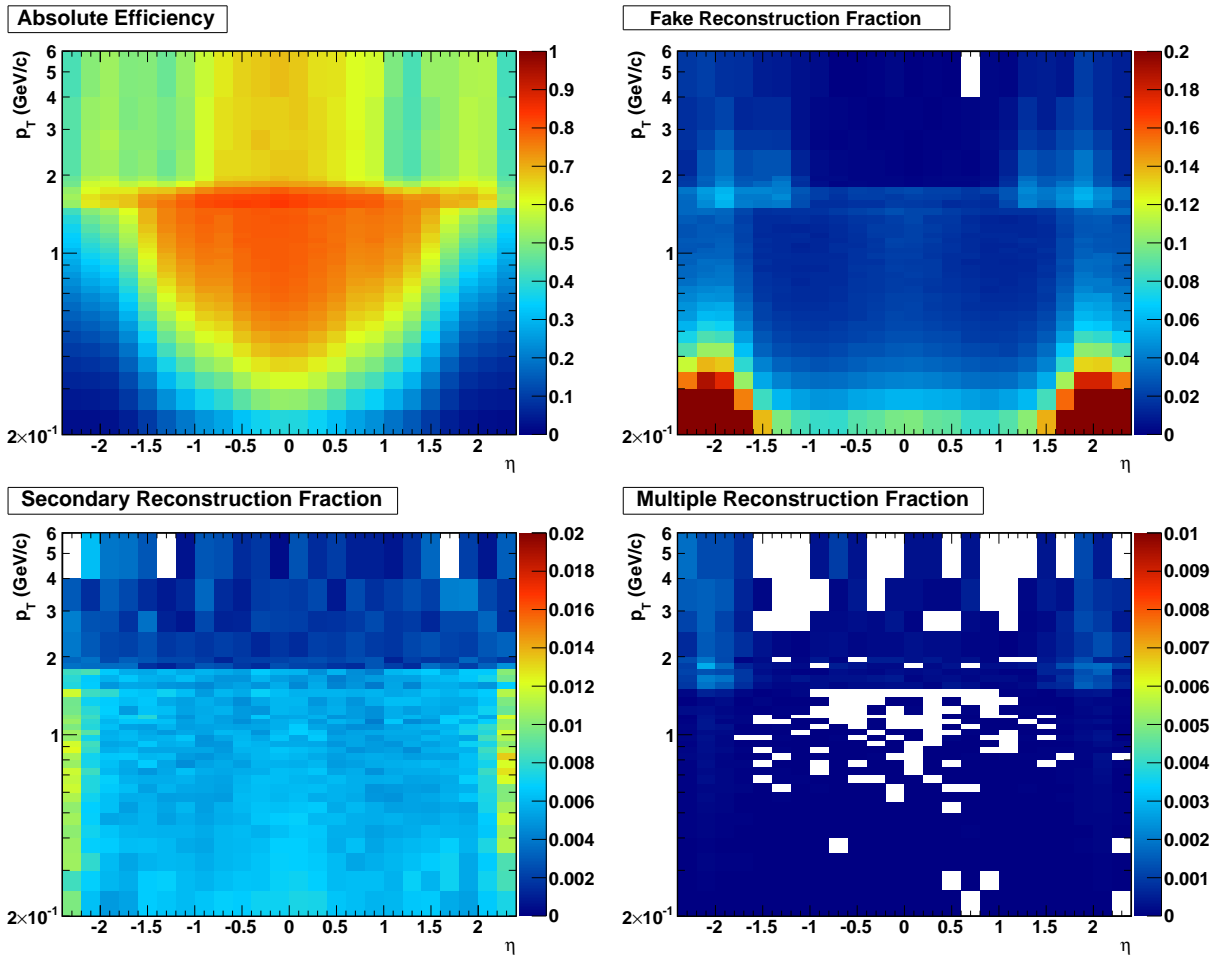
Figure 4.5: The absolute efficiency (Top Left), fake rate (Top Right), secondary reconstruction fraction (Bottom Left), and multiple reconstruction fraction (Bottom Right) in 0-10% central PbPb events as a function of p_T and η. Note that the color scale is different for each of the four figures as indicated.
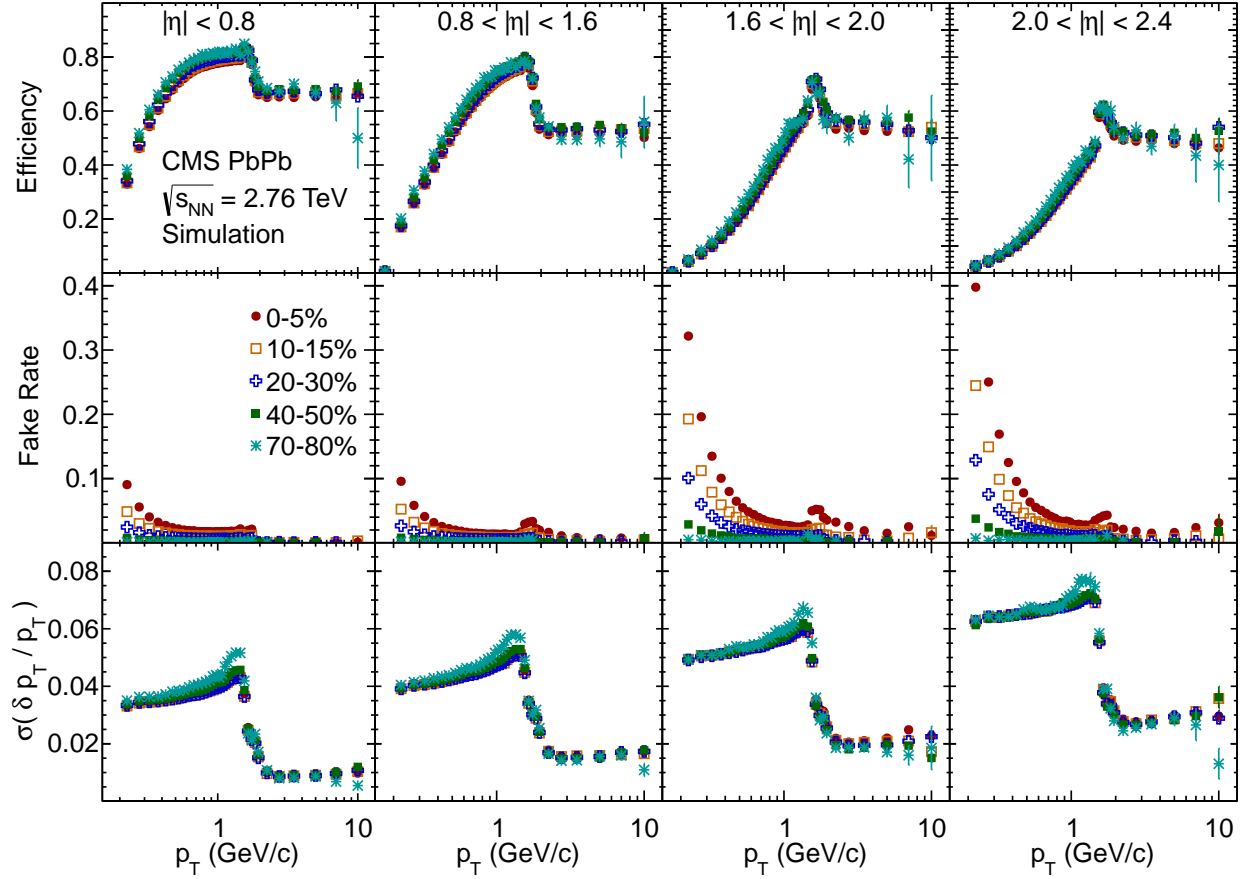
Figure 4.6: Efficiency (top), fake rate (middle), and momentum resolution (bottom) of the charged-particle reconstruction algorithm for PbPb collisions, shown for four pseudorapidity regions: $|\eta| < 0.8$, $0.8 < |\eta| < 1.6$, $1.6 < |\eta| < 2.0$, and $2.0 < |\eta| < 2.4$ displayed from left to right, and shown for the five centrality classes given in the legend [44].
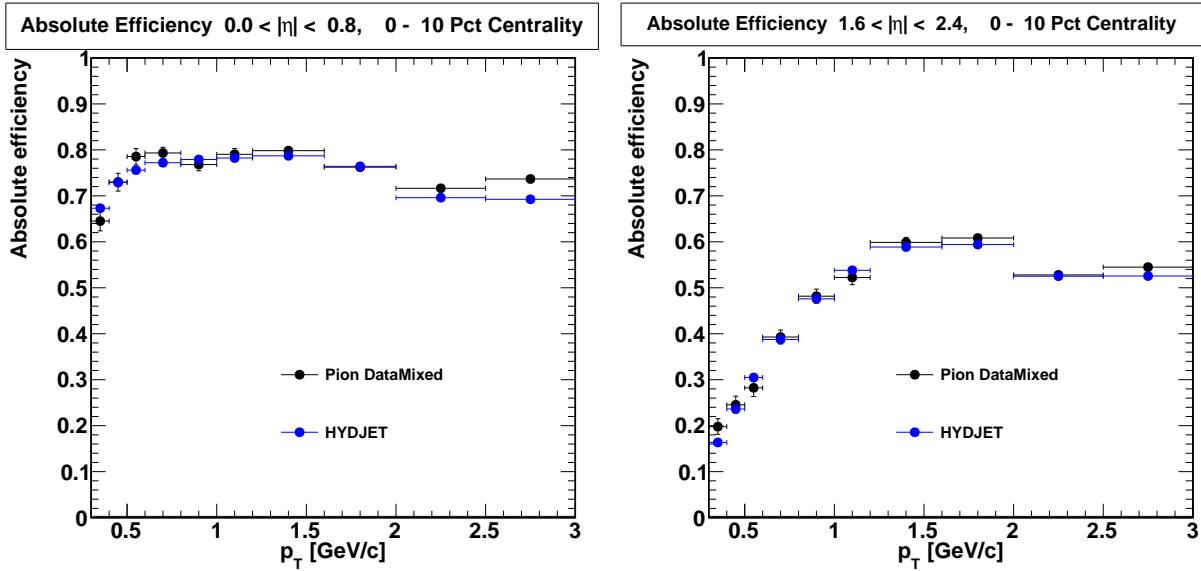
Figure 4.7: The efficiency of charged pions embedded into central PbPb collision events compared with the efficiency of charged particles in HYDJET simulation as a function of p_T at midrapidity (Left) and forward rapidity (Right).
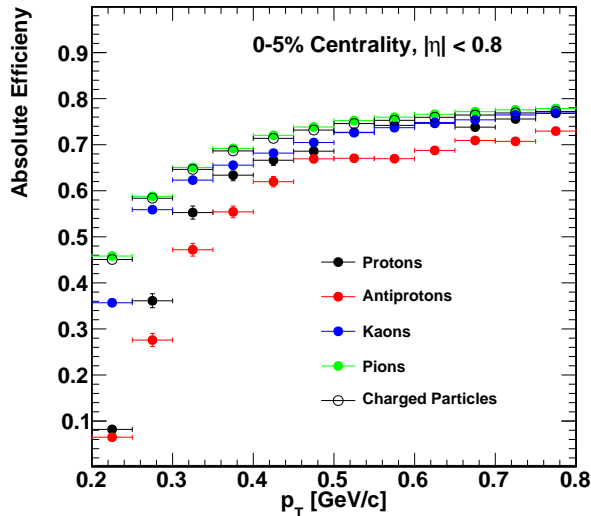
Figure 4.8: The efficiency for different particle species as a function of p_T in 0-5% central simulated HYDJET collision events at midrapidity. Note that this study was performed using an early selection criterion for the pixel tracks and the overall efficiency of charged particles is therefore different than that found using the final selection.

4.5 Tracking Performance in pPb Collisions

The seven-iteration tracking algorithm designed and evaluated in pp collisions works well in pPb collisions. This is to be expected, as the LHC conditions in 2011 and 2012 were such that the average number of pp collisions in one bunch crossing could easily rise to more than 20. For comparison, from Glauber Modeling as described in Sec. 5.5, one may expect that an average pPb collision would result in charged-particle production roughly equivalent in number to $\sim 5 - 10$ inelastic pp collisions.

The *highPurity* selection criterion provides a ready collection of tracks with high efficiency and low fake rate. However, in order to reject most secondary particles and further reduce the fake rate, the following selection criteria are applied:

- The *highPurity* criterion
- Low uncertainty on the measured p_{T}: $p_T^{\mathrm{error}}/p_{\mathrm{T}} < 0.1$
- Transverse vertex compatibility: $d_{xy}/\sigma_{d_{xy}} < 3.0$
- Longitudinal vertex compatibility: $d_z/\sigma_{d_z} < 3.0$

Note that in pPb collisions more than one primary vertex may be reconstructed. Here the quantities $d_{xy}/\sigma_{d_{xy}}$ and d_z/σ_{d_z} are relative to the vertex with the most associated tracks, or in the case of two vertices with an equal number of associated tracks, the one with the best statistical fit.

There major difficulties in evaluating the performance of the pPb tracking only result from the requirement that the spectra measurements are carried to very high-p_{T}. Given the extremely low probability of a charged particle with $p_{\mathrm{T}} \sim 50 - 100\,\mathrm{GeV}/c$ being produced in a collision, one would require billions of simulated minimum-bias events in order to accurately measure the efficiency. Additionally, the efficiency of charged particle reconstruction may be sensitive to the high occupancy of the tracker modules lying within the path of a highly collimated jet of hadrons. A final difficulty emerges due to the uncertainty in the particle composition, and the very plausible theory that for $p_{\mathrm{T}} \gtrsim 5\,\mathrm{GeV}/c$, the fraction of strange and multistrange baryons that comprise the charged particles produced in the collision is considerably larger than in pp events.

4.5.1 Simulation Setup

The cross section of an inelastic pp collision at $\sqrt{s} = 5.02\,\mathrm{TeV}$ in PYTHIA Z2 is approximately 68 mb. In a given jet, the leading fragment will typically have a transverse momentum fraction of $z \approx 0.4$ of the total momentum of the jet. In the same PYTHIA Z2 tune, the cross section of hadronic $2 \to 2$ scatterings producing jets with $p_{\mathrm{T}} > 220\,\mathrm{GeV}/c$ is only 1.0×10^{-5} mb. This means that to potentially simulate a single charged particle with $p_{\mathrm{T}} \approx 100\,\mathrm{GeV}/c$ one would require on the order of millions of simulated inelastic pp collisions. Reasonably assuming a similar amount of charged-particle production in pPb collisions means that one will still need to simulate millions of events just to produce a single high-p_{T} particle for evaluation.

In order to overcome this difficulty, inelastic pPb collision events were simulated using the HIJING event generator, and into these collisions were embedded pp hadronic $2 \to 2$ scattering events from PYTHIA with some minimum transverse momentum of the outgoing partons , $\hat{p}_T$. These are produced at various values of $\hat{p}_T$ in order to produce a sufficient selection of high-p_T charged particles within the cones of jets of various E_T values in order to determine the tracking efficiency as a function of both p_T and jet-E_T. The specific values of $\hat{p}_T$ and their corresponding PYTHIA Z2 cross sections are given in Table 4.2.

Table 4.2: Cross sections and $\hat{p}_T$ for the PYTHIA hadronic $2 \to 2$ scattering events that are embedded into HIJING inelastic collisions

$\hat{p}_T$	cross section (mb)
15	5.335×10^{-01}
30	3.378×10^{-02}
50	3.778×10^{-03}
80	4.412×10^{-04}
120	6.147×10^{-05}
170	1.018×10^{-05}
220	2.477×10^{-06}
280	6.160×10^{-07}
370	1.088×10^{-07}

These embedded simulation events may be analyzed separately in each $\hat{p}_T$ class, or combined together. One may combine the samples weighting each track in each event equally, in which case a sample with an extremely unrealistic jet spectrum is obtained. One may also obtain a sample of simulated events with a reasonably realistic jet-E_T distribution by weighting each charged particle and track in each event by the appropriate cross section as listed in Table 4.2. In order to suppress statistical fluctuations magnified by this weighting procedure, tracks and simulated charged particles with a transverse momentum greater than the corresponding $\hat{p}_T$ of the sample they occur in are removed.

4.5.2 Occupancy Effect in High Energy Jets

In pp and PbPb collisions, the charged-particle efficiency at high-p_T was found to vary with the E_T of the jet that the particle occurred in, even for tracks with the same p_T [27, 125]. Therefore it was expected that similar variation would be found in pPb collisions. In the previous studies, track efficiency was parametrized in terms of the track p_T, η, and the leading jet-E_T in the entire event.

In order to perform such a parametrization, hadronic jets must be reconstructed. This jet reconstruction is performed using the CMS particle flow algorithm [127, 128], which attempts to identify all stable particles in an event, including charged and neutral hadrons, photons, electrons, and muons. These are grouped according to the granularity of the HCAL, and from these groups jets are reconstructed using the anti-k_T sequential algorithm provided in the FASTJET framework [129, 130] with a cone size of $R = 0.3$. As pPb collision events contain a considerable background of particles

produced by other processes not related to the jet, this underlying event background is corrected for using an iterative algorithm described in Ref. [131] using the same implementation as the PbPb jet analysis in Ref. [132].

The absolute tracking efficiency, fake rate, secondary reconstruction, and multiple reconstruction fraction in pPb collisions with different leading reconstructed jet-E_Ts is shown for $-1.46 < \eta < 0.54$ in Fig. 4.9. These results were determined using the combination of pure HIJING events and HIJING with embedded PYTHIA jets as listed in Table 4.2, with equal weighting of events. This η range is important as it represents the $|\eta| < 1.0$ region in the center-of-mass frame of roughly half of the pPb collisions recorded in 2013 (see Sec. 8.2 or a discussion of the lab and center-of-mass frames in these collisions). The fake rate, secondary, and multiple reconstruction fraction have little dependence on leading reconstructed jet-E_T.

The charged-particle efficiency may also be determined using the cross section weighted sample of HIJING events with embedded PYTHIA jets. This is shown in Fig. 4.10. The benefit of using this cross section weighted sample to measure efficiency is that in principle it allows one to correct a measured spectrum for the tracking efficiency without needing to apply a differential correction in bins of leading jet-E_T. If the leading jet-E_T spectrum in pPb collisions is reproduced reasonably well by the PYTHIA Z2 generator, then the result should be the same as the differential correction. The leading jet-E_T differential correction has the drawback of significantly increased statistical uncertainties that may bias the measured spectra.

In order to validate that the PYTHIA Z2 generator produces a reasonable enough leading jet-E_T spectrum to correctly account for the efficiency dependence on leading jet-E_T, measured pPb charged-particle spectra (see Chapter 8) were corrected for efficiency using both a differential leading jet-E_T correction based on the combined samples with equal weighting, and a leading jet-E_T independent correction factor based on the cross-section-weighted sample. The results of using either correction factor were generally compatible within $\sim 3\%$ or less, and showed no systematic differences affecting the overall shape of the measured spectrum. The cross section weighted correction factors produced a spectrum with less point-to-point scatter, and so was chosen to use for the measured result.
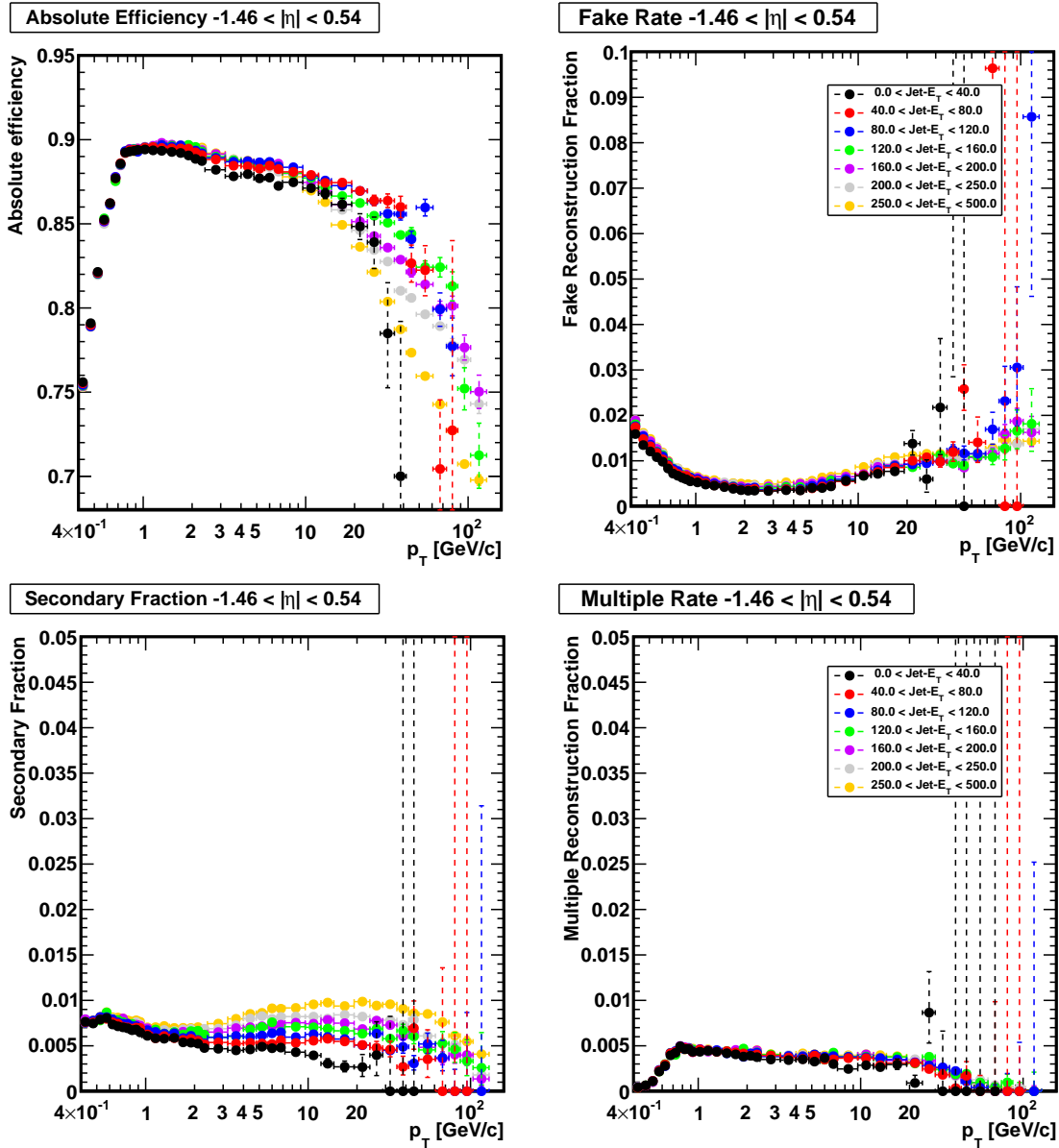
Figure 4.9: The absolute efficiency (Top Left), fake rate (Top Right), secondary reconstruction (Bottom Left), and multiple reconstruction fraction (Bottom Right) for charged particles in events with a leading reconstructed jet-E_T as given in the legend as a function of p_T. These are given in the lab pseudorapidity range of $-1.46 < \eta < 0.54$ and were produced by the combined sample of HIJING collisions and HIJING with embedded PYTHIA jets as listed in Table 4.2, with equal weighting of events.
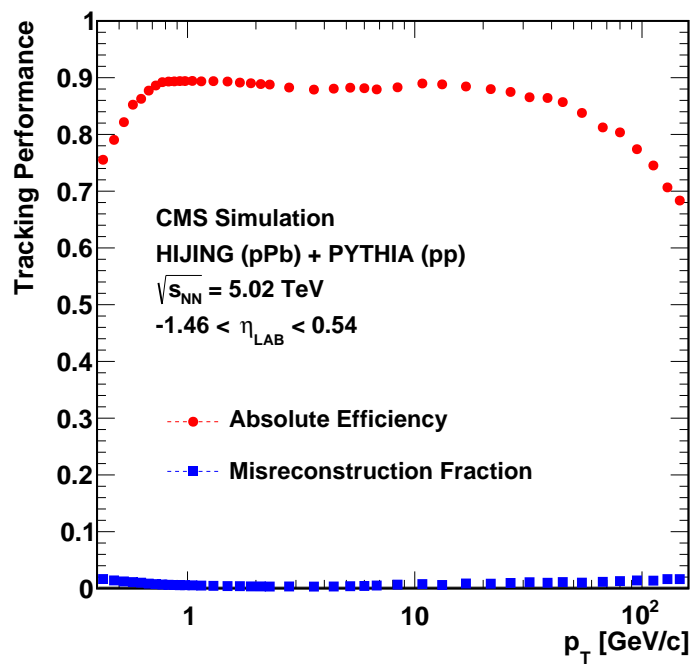
Figure 4.10: The absolute efficiency and fake rate for charged particles as a function of p_T for $-1.46 < \eta < 0.54$ as measured by the cross section weighted sample of HIJING collisions and HIJING with embedded PYTHIA jets with various $\hat{p}_T$ values.

4.5.3 Efficiency of Different Charged-Particle Species

For pions, kaons, and protons, the reconstruction efficiency in pPb events is fairly similar, as is shown in Fig. 4.11, and variations in the relative fractions of these most common hadrons are not expected to have a large effect on the charged-particle efficiency. However, the long-lived strange ($\Sigma^\pm$) and multi-strange ($\Xi^\pm$, $\Omega^\pm$) baryons have a much lower efficiency due to their expected decay after traversing just a few cm of the tracker. In the case of very high-p_T baryons, the particles become more easily reconstructable, as their large Lorentz factor in the lab frame greatly increases their expected lifetime, and they may traverse enough of the detector to be reconstructable. These baryons meet the definition of a primary charged particle, having a mean expected lifetime of greater than 1 cm/c, and so any inefficiency must be corrected for.
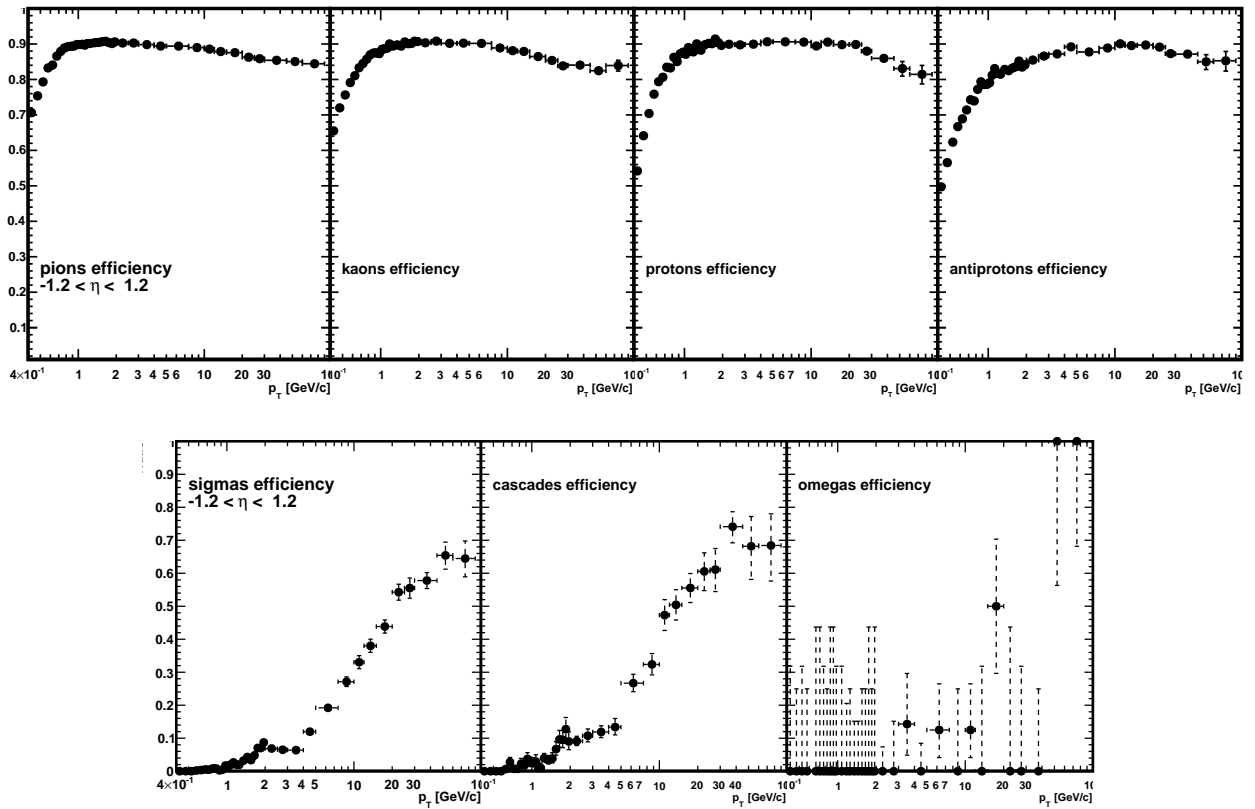
Figure 4.11: The reconstruction efficiency of $\pi^\pm$, $K^\pm$, p, $\bar{p}$, $\Sigma^\pm$, $\Xi^\pm$, and $\Omega^\pm$ (Left to Right, Top to Bottom) from HIJING simulated pPb collision events as a function of p_T.

The fraction of charged particles of a given species ($\pi^\pm$, $K^\pm$, p, $\bar{p}$, $\Sigma^\pm$, $\Xi^\pm$, $\Omega^\pm$) is shown for four different MC generators in Fig. 4.12. Two of the generators, HIJING and EPOS, are simulating pPb collisions, and the others simulate pp or peripheral PbPb collisions and are shown for comparison. The EPOS generator produces a much larger fraction of strange and multi-strange baryons, reaching $\sim 10\%$ of the total charged-particle production in the range $5 < p_\mathrm{T} < 10\,\mathrm{GeV}/c$. Given the low efficiency of these baryons in this p_T range, this discrepancy would result in a large systematic uncertainty in the tracking efficiency.

Although no published measurements of $\Xi^\pm$ or $\Omega^\pm$ production in pPb collisions at any energy close to $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ currently exist, one can look at the performance of EPOS in pp collisions. In the ALICE measurement of $\Xi^\pm$ and $\Omega^\pm$ in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ [133], the PYTHIA Perugia 2011 tune [134] underpredicted these identified particle spectra by a factor of 2 ($\Xi^\pm$) or 4 ($\Omega^\pm$). Predictions from the Gribov-Regge multiple scattering framework of EPOS can describe these measurements to $\sim 10-20\%$ accuracy, while pQCD-based models such as HIJING or PYTHIA generally underpredict this result [135].

One can therefore not ignore the EPOS prediction, and the difference in the efficiencies determined using different generators, however large, is to be taken as a systematic uncertainty on the measured spectra. This difference is shown in Fig. 4.13, and the discrepancy is as much as 10%. This large uncertainty cannot be reduced until measurements of multi-strange baryons in pPb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ have been performed.
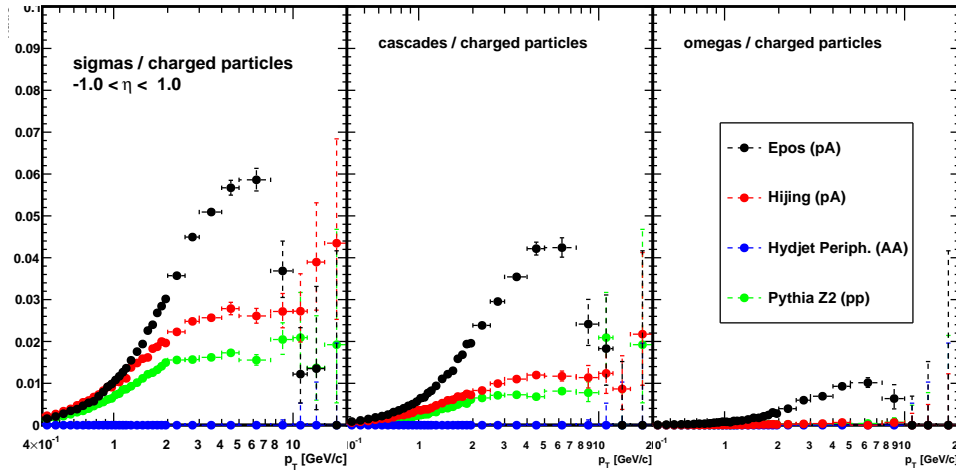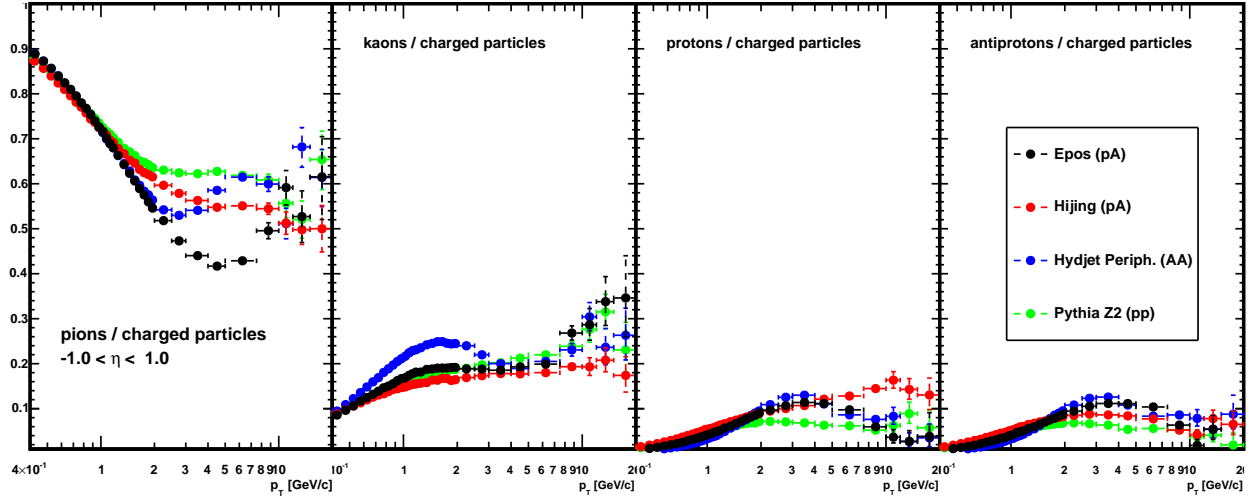
Figure 4.12: The fraction of $\pi^\pm$, $K^\pm$, p, $\bar{p}$, $\Sigma^\pm$, $\Xi^\pm$, and $\Omega^\pm$ (Left to Right, Top to Bottom) in all charged particles in pPb collisions from HIJING and EPOS, pp collisions from PYTHIA, and peripheral PbPb collisions from HYDJET.
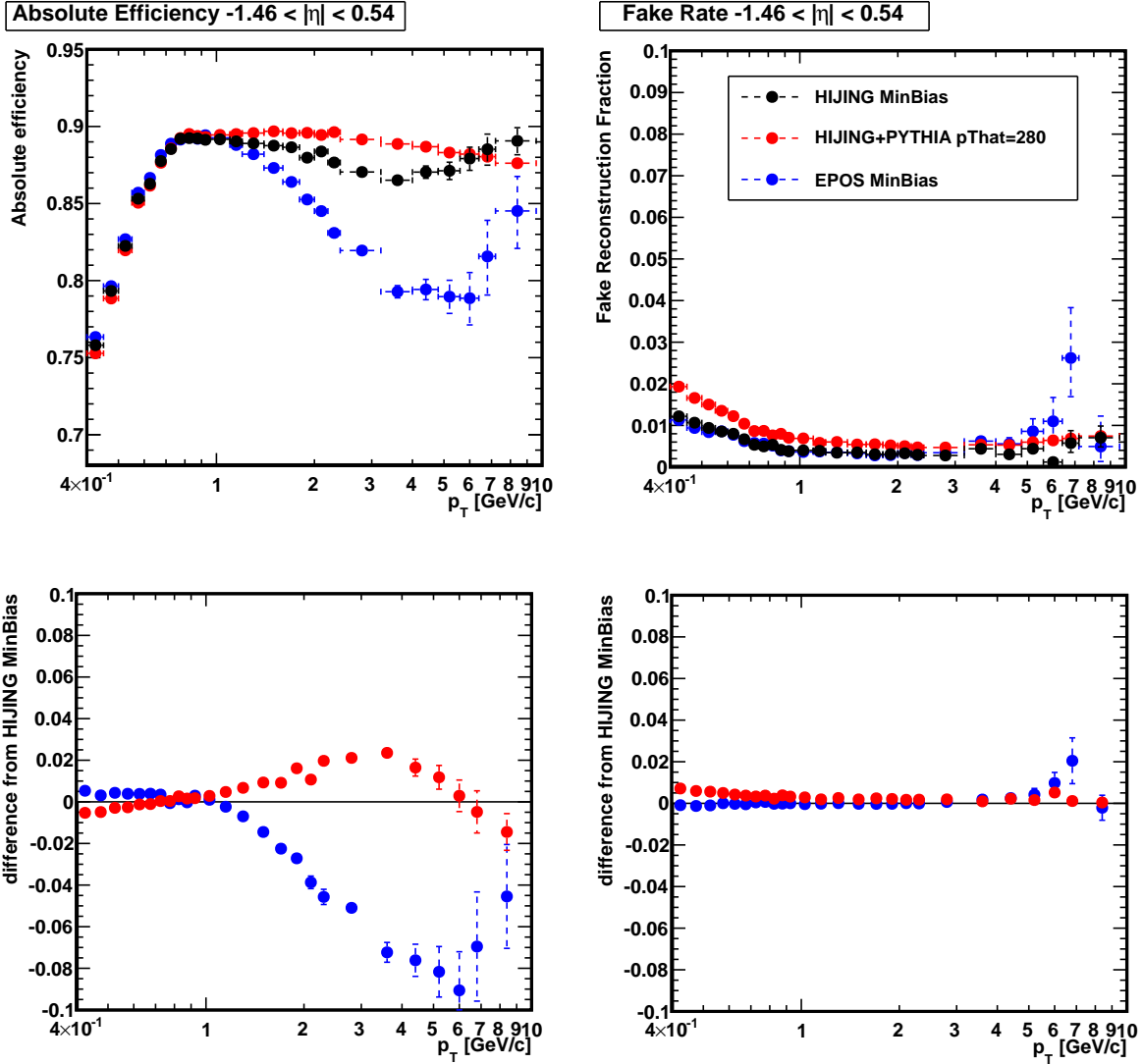
Figure 4.13: The charged-particle efficiency (Top Left) and fake rate (Top Right) in pPb collisions measured using EPOS and HIJING inelastic collision events. The efficiency in HIJING events with embedded PYTHIA jets is also shown for comparison. The ratio of the EPOS to HIJING efficiency (Bottom Left) and fake rate (Bottom Right) is also shown.

Chapter 5

Minimum-Bias Event Selection and Centrality

5.1 Minimum-Bias Triggering and Offline Event Selection

In the measurement of PbPb collisions, one needs in principle a clean and unbiased sample of inelastic collisions in order to determine centrality classes. In practice this is not possible, and so a combination of an L1 or High-Level trigger determined during data taking, as well as a battery of further "offline" event selections performed after reconstruction, are used to create a minimum-bias sample that is as close as possible to a pure and unbiased sample of inelastic collisions.

The goal is similar in pPb collisions, although for the measurement presented, no attempt is made to sample or correct to the full inelastic cross section. Instead, the measurement is made relative to a class of collision events termed Double-Sided (DS). A DS event is one in which at least one particle with an average proper lifetime of $\tau > 10^{-18}$ s and $E > 3\,\mathrm{GeV}$ is produced in the pseudorapidity range of $3 < \eta < 5$ and another such particle produced in the range $-5 < \eta < -3$. For pPb collisions with $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ this definition is nearly equivalent to non-single diffractive (NSD) collisions. This definition is chosen as it is simple to correct the minimum-bias sample back to an unbiased DS sample using simulated events.

The types of events that may contaminate a sample of inelastic or DS collision events are often the result of one of the following sources:

- Beam-halo events, where muons are produced from the collision of protons or ions with metallic beam collimators at some distance from the CMS detector, ultimately producing a muon that travels perpendicular to the beam and enters the detector.

- Large-multiplicity beam-background events, where a spray of particles may be produced and enter the detector via the collision of a proton or ion with residual gas particles in the beampipe, or of a stray ion with a nearby collimator or other component of the beampipe.

- Cosmic ray muons which penetrate through the detector.

- Spurious detector noise causing a given trigger path to fire, potentially due to radioactivation of the detector material.

- In the case of PbPb collisions especially, collisions from purely electromagnetic interactions between Pb nuclei, often with an impact parameter too great to permit nuclear interactions. These events are termed Ultra-Peripheral Collisions (UPC) and are not considered part of the nuclear inelastic cross section [136].

5.1.1 Minimum-Bias Event Selection in PbPb Collisions

To select events with minimum-bias in PbPb collisions, a L1 trigger requirement of a logical OR of two clean and efficient triggers. The first requires a coincidence of signals on at least one segment for both the $+z$ and $-z$ sides of the BSC detectors. The second trigger requires at least one HF tower on each side of the interaction point with a signal that exceeds the readout threshold. Both of these triggers required the presence of two colliding bunches as measured by the LHC beam monitoring devices located close to the detector.

The triggered event sample is further refined by the following offline event selections:

- Beam-halo muon events are identified and excluded by considering the timing difference between the BSC signals located at ± 10.9 m from the interaction point. The flight time between these two detectors is approximately 73 ns, and so to exclude signals produced from muons passing from one side of the detector to the other, the event is excluded if the timing difference is 73 ± 20 ns.

- Large multiplicity beam background events are identified by looking at reconstructed "clusters" of signals in the pixel detector originating from the same charged particle. In events with more than 150 such clusters, the event is rejected if the lengths of the clusters along the z-axis are generally incompatible with the trajectory of a charged particle originating from the interaction point, which is expected to graze the silicon sensor at a given angle resulting in a specific cluster length. This method is the same as was used in [137].

- To further reject non-collision and UPC events, an HF coincidence is imposed, where at least three HF towers with at least 3 GeV of energy each are found on each side of the detector.

- The presence of a reconstructed primary vertex within $z = \pm 10$ cm and $r = 2$ cm of the interaction point is required.

After all selections have been performed, the remaining events show a tight correlation between the total energy deposited in the HF calorimeters and the 1st layer of the barrel pixel detector. This correlation is shown in Fig. 5.1. Very few events deviate significantly from the expectations for hadronic PbPb collisions. This validates both the purity of the selected sample, as well as the use of the HF energy as a measure of event centrality, discussed in Sec. 5.4.

The total hadronic inelastic collision rate varied between 1 Hz and 210 Hz in the 2010 PbPb data used in this measurement. After all selections, 22.6 million minimum-bias events were analyzed, corresponding to an integrated luminosity of 3 μb^{-1}. The minimum-bias trigger and subsequent offline event selection was determined by MC simulation to select $97 \pm 3\%$ of the total inelastic cross section with a negligible fraction of non-collision events contaminating the sample.

5.1.2 Minimum-Bias Event Selection in pPb Collisions

As the BSC detectors had been removed in the winter of 2011, and the performance of the new BHC replacement was still under commissioning study, the minimum-bias trigger consisting of either a signal in the HF or BSC was not suitable for this data taking period. Instead, a HLT
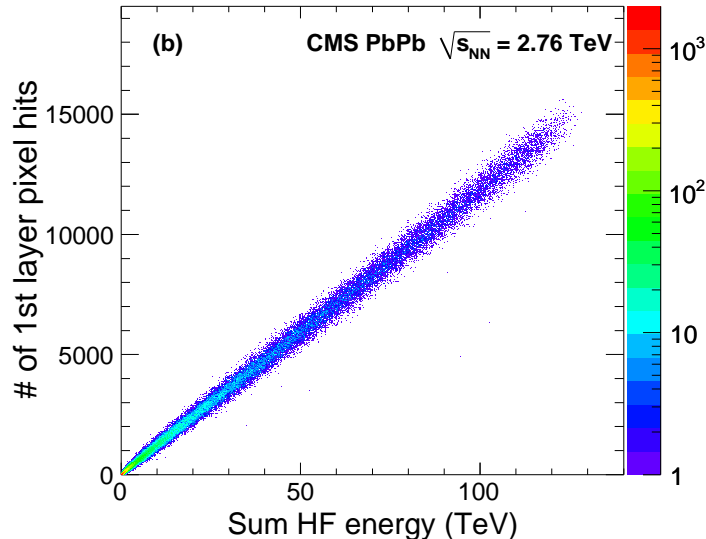
Figure 5.1: Correlation between the total number of hits on the 1st layer of the barrel pixel detector and the total HF energy for a single run containing approximately 60k selected minimum-bias PbPb events [132].

path was selected requiring a single charged particle trajectory with $p_T > 400\,\text{MeV}/c$ reconstructed from signals in the pixel detector. This trigger path was found to be highly efficient, although it was heavily prescaled, typically recording on the order of only one out of every 3000 triggered events.

Several changes were also made to the offline event selection relative to the PbPb configuration. As the BSC was not in operation, the related requirement to remove beam-halo events was not applied. For increased sensitivity given the smaller particle flux in pPb collisions relative to PbPb, the HF coincidence requirement was relaxed to only one tower on each side with $E > 3$ GeV. For the removal of large multiplicity beam-background events, a different approach was utilized, based on the algorithm described in [124]. Here, the collection of fully reconstructed tracks is analyzed, and if more than 10 tracks are found in the event, at least 25% of these tracks must be considered of sufficient quality, specifically meeting the *highPurity* selection.

5.2 Pileup Removal

In a given bunch crossing, there may sometimes be more than one inelastic collision. The term "pileup" is defined as the average number of inelastic collisions per bunch crossing. The term "pileup removal" is the removal of all events in which more than one inelastic collisions is detected from the sample. For pp collisions at design LHC luminosity, pileup removal would result in the removal of nearly the entire sample, as the average number of collisions per bunch crossing may be 20 or more, and only a rare few bunch crossings would result in a single inelastic collisions. At the other extreme, in PbPb collision data taken in 2010 and 2011, the pileup is so low that the fraction of collision events with two or more collisions is negligible.

In the 2013 pPb run, the instantaneous luminosity was such that pileup values of 0.06 were common, meaning that the average bunch crossing would contain only 0.06 collisions. As each inelastic collision is independent of the others, the number of collisions in a bunch crossing is Poisson distributed, with the probability of a given bunch crossing containing k inelastic collisions is

$$P\left(N_{inel} = k\right) = \frac{\mu^k e^{-\mu}}{k!}, \tag{5.1}$$

where μ is the pileup. Therefore, for a pileup of 0.06, the fraction of collision events which contain more than one collision is approximately 3%.

The silicon tracker is easily capable of reconstructing multiple primary vertices corresponding to each inelastic collision, and tracks can usually be definitively associated with the primary vertex corresponding to the inelastic collision that produced them. For the purposes of measuring the charged-particle spectrum of pPb collisions, one could in principle attempt to count the number of collisions in each event, and therefore account for the effect of pileup. Therefore, one may naturally ask why attempt "pileup removal" at all?

The problem with an event containing two collisions is that both collisions will deposit energy into the HF, and the segmentation of the HF is insufficient to determine from which collision the energy deposit originated. Since these HF deposits are used to determine event centrality, as is discussed in detail in Sec. 5.4, it is required that each event contain only one inelastic collision.
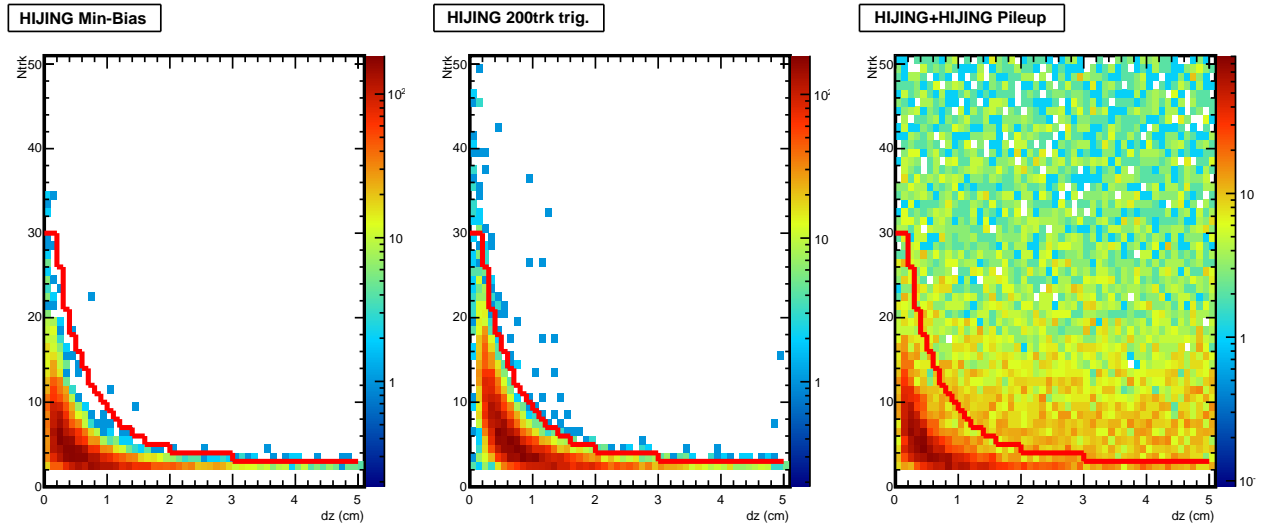
Figure 5.2: The longitudinal displacement between two primary vertices (dz) versus the number of associated tracks (N_{trk}) on the lesser vertex for pPb collision events simulated with the HIJING generator. Events with vertex pairs having (dz, N_{trk}) coordinates above the red line are identified as multiple-collision events and removed. Three types of simulated events are shown: (Left) minimum-bias single-collision events, (Center) high multiplicity single-collision events with over 200 tracks per event, and (Right) events with two independent simulated collisions.

The primary vertexing algorithm used in pp and pPb collisions will attempt to reconstruct a primary vertex for each inelastic collision that occurs in the event. In the case that the two collisions occur within 1 cm or less of one another along the z-axis, the the algorithm may sometimes fail to reconstruct separate vertices for each collision. Alternatively, there are cases where in a single collision event, the algorithm mistakenly produces two reconstructed vertices. For this reason, one cannot simply remove every event where two or more reconstructed vertices are present, and one may also not completely remove all multiple-collision events from a sample.

If two primary vertices are reconstructed sufficiently far apart along the z-axis, they are almost always formed by two different pPb collisions, rather than an algorithmic splitting of a single collision or reconstruction of a secondary decay vertex within the beampipe. The minimum distance required to definitively determine that the vertices correspond to different collisions was studied using HIJING simulations in which only one collision is simulated in the event. In this simulation the longitudinal distance between vertices, dz, was determined for events in which multiple primary vertices were reconstructed.

The maximum distance between two primary vertices spuriously reconstructed from the same collision varies depending on the number of tracks on the vertex with the smaller number of tracks, N_{trk}. The distribution of (dz, N_{trk}) in simulated HIJING events with one collision are shown on the left of Fig. 5.2. The red curve in this figure shows the selection criterion in the (dz, N_{trk}) plane – events with vertex pairs with (dz, N_{trk}) above this line are considered multiple-collision events and are rejected. One small refinement to this method was to additionally look at the transverse displacement between the vertices, d_{xy}. Vertex pairs with $d_{xy} > 500\ \mu$m are usually not the result of multiple collisions, as one vertex is significantly displaced from the beamline, and so such pairs never result in the removal of the event.

On the right of Fig. 5.2 the (dz, N_{trk}) distribution is shown for vertex pairs in HIJING events with two simulated collisions. The (dz, N_{trk}) position of vertex pairs from different collisions is randomly distributed in the plane. From these simulations, it was found that the pileup rejection criterion removed less than 0.01% of single collision events, and removed 88.2% of the two collision events.

An additional data-driven analysis was performed to ensure that the MC based determination of the efficiency (fraction of multiple-collision events removed), and purity (fraction of removed events that are multiple-collision events) is accurate for the actual collision data. The pileup, and hence the multiple collision fraction in selected events, can be estimated in pPb collisions from the rate at which the single-track trigger ($p_\text{T} > 400$ MeV/c) fires.

For such an analysis, one must define a single track collision (ST) as one that fires this single-track trigger. Then one can estimate the probability that a given bunch crossing contains at least one ST collision by looking at the rate at which events pass the ST trigger, R_ST, the prescale factor of the ST trigger, F_prescale, the number of colliding bunches, N_bunch, and the LHC orbit frequency, f_LHC, as follows:

$$P(N_{inel} \geq 1) = \frac{R_\text{ST} \cdot F_\text{prescale}}{N_\text{bunch} \cdot f_\text{LHC}} \tag{5.2}$$
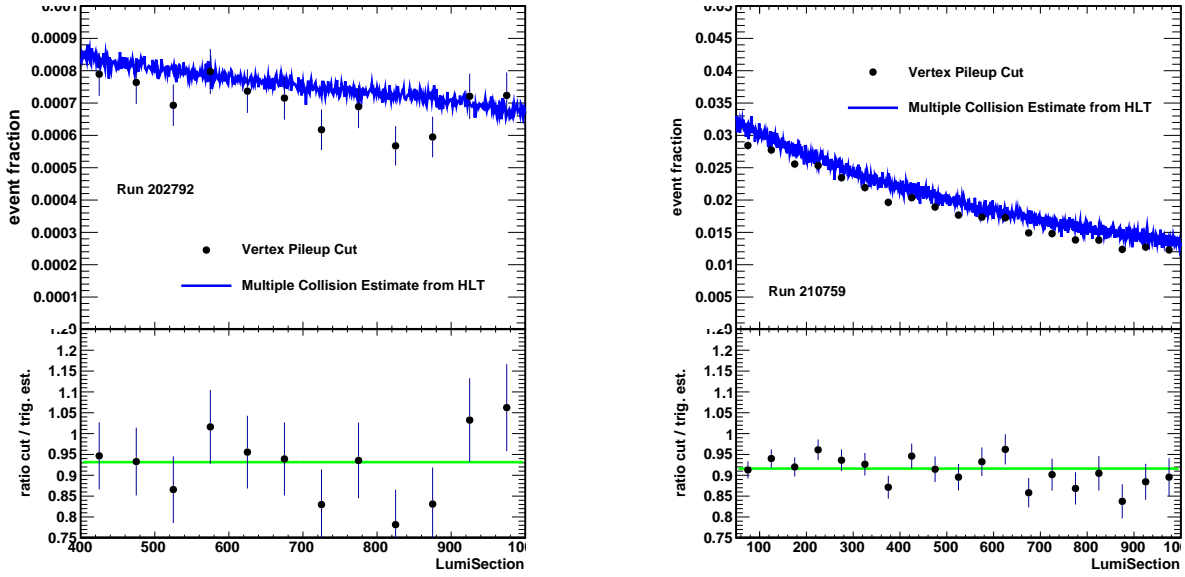
Figure 5.3: (Top) The fraction of single-track triggered events removed by the pileup filter (black) compared with the estimated fraction of multiple-collision events from the trigger rate (blue) as a function of time given in lumisections, or 23 second intervals since the beginning of the run. (Bottom) ratio of the fraction of events removed by the filter to the expected fraction of multiple-collision events. This is shown for an extremely low-intensity pPb run (Left) and a high-intensity pPb run (Right).

Then combining this estimate with equation 5.1, one can determine the pileup, μ:

$$P(N_{inel} \geq 1) = 1 - P(N_{inel} = 0) = 1 - e^{\mu}, \tag{5.3}$$

$$\mu = \ln\left(1 - \frac{R_{ST} \cdot F_{prescale}}{N_{bunch} \cdot f_{LHC}}\right) \tag{5.4}$$

Once μ is determined, then the fraction of triggered events with multiple ST collisions is simply calculated from $P(N_{inel} \geq 2)/P(N_{inel} \geq 1)$. This can be compared with the fraction of triggered events removed by the pileup rejection criterion.

In the right of Fig. 5.3, the fraction of multiple-collision events estimated by the trigger is compared with the amount removed by the selection criterion over the course of a single 2013 pPb run. As the bunch intensity and collision rate decreases over the course of the run, this fraction is estimated as a function of time. The fraction removed is stable over time at $91.6 \pm 0.7\%$ of the total multiple-collision fraction expected. This indicates that the estimate of the efficiency from simulation is accurate.

In the left of Fig. 5.3, an identical comparison is produced using a low-intensity pPb pilot run taken in 2012. In this run, the multiple-collision fraction is extremely low. In this case, the selection

criterion removes $9.19 \pm 2.5\%$ of the (very small) expected fraction of multiple-collision events. This second comparison demonstrates that the selection criterion rarely misidentifies a single collision event, as was seen in the simulated study.

5.3   Trigger Efficiency Determination in pPb

The single track trigger and subsequent offline event selection used for pPb collisions is not fully efficient in terms of selecting double sided (DS) events, as defined in Sec. 4.2. Additionally, this event selection will occasionally select events that do not meet the DS criterion. Part of the reason for the choice of measuring spectra in terms of DS rather than inelastic collisions is that the combination of the single track trigger and HF coincidence filter is extremely efficient at selecting DS and only DS events.

However small, this inefficiency has the potential to distort the measured pPb charged-particle spectrum, as events with more charged particle production are more likely to be selected by the single track trigger. To correct for such a distortion, the trigger efficiency is measured in simulation as a function of the total number of selected tracks that are reconstructed in the event, including tracks for which $p_\mathrm{T} < 0.4\,\mathrm{GeV}/c$. This total number of tracks is termed the event multiplicity, M. The trigger efficiency and fraction of non-DS events in the sample as a function of M is shown in Fig. 5.4 as measured using the HIJING and EPOS event generators, which handle the diffractive portion of the pPb cross section differently.
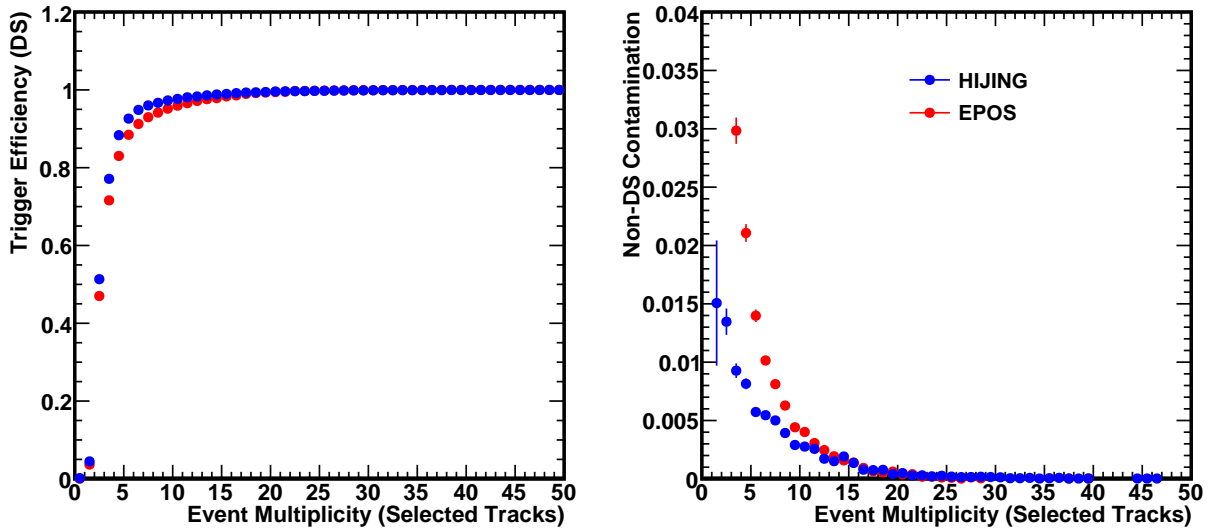
Figure 5.4: (Left) The trigger efficiency in terms of DS events as a function of the number of selected tracks found in each event, evaluated using HIJING and EPOS. (Right) The fraction of the triggered sample consisting of non-DS events as a function of the number of selected tracks in each event.

To correct for any distortion in the measured pPb spectra, the tracks in each of the selected events

are weighted by the trigger efficiency and non-DS contamination as a function of M for the given event:

$$w_{\text{evt}}(M) = \frac{1 - f_{trig}(M)}{\epsilon^{trig}(M)},\tag{5.5}$$

where $\epsilon^{trig}(M)$ is the selection efficiency for DS events with multiplicity M, and $f_{trig}(M)$ is the non-DS contamination rate for selected events with multiplicity M. The total number of events counted in the sample is then adjusted for the proportion of DS events in simulation with $M = 0$.

Although the correction factors used in the measurement are taken from the results from the HIJING simulation, the systematic uncertainty in the trigger efficiency correction factor can be evaluated using the EPOS simulation results.

5.4 Centrality Measurement in PbPb

The variable chosen to estimate the centrality of each PbPb event is the total transverse energy deposited into the HF detectors at both positive and negative pseudorapidity ($2.9 < |\eta| < 5.2$). The distribution of energy per minimum-bias event is shown in Fig. 5.5. This distribution is then used to divide the events into 40 bins, each representing approximately 2.5% of the total inelastic cross section. These bins may be combined to create 5 or 10% centrality classes as used in the PbPb spectra analysis.

Since the minimum-bias trigger and offline event selection for PbPb collisions is only 97±3% efficient, the measured 2.5% centrality bins do not perfectly correspond to 2.5% slices of the total hadronic cross section. However, for collisions producing enough energy in the HF to have a centrality of 0–80%, the minimum-bias trigger is effectively fully efficient. Therefore, as long as the most peripheral events are not considered, one may consider the trigger efficiency to be a simple scaling factor that horizontally stretches the bin boundaries. The bin boundaries are corrected for this efficiency, and can be further adjusted for the assumption of 94% or 100% efficiency in order to estimate the systematic uncertainty on any measurement due to the uncertainty on the minimum-bias trigger efficiency.

5.5 Glauber Modeling

The techniques used to estimate the femtoscopic quantities of the collision, such as the impact parameter, b, or the number of participating nucleons, N_{part}, are generally termed "Glauber models" owing to the pioneering work of Roy Glauber in his systematic treatment of many-body nuclear systems. While traditional Glauber models employed analytic expressions and the optical theorem, where incoming nucleons see the target as a smooth density, more recent approaches use Monte Carlo methods to simulate random distributions of nucleons within each nucleus. A review of both optical and MC Glauber methods can be found in [45]. For the present analysis, a MC Glauber
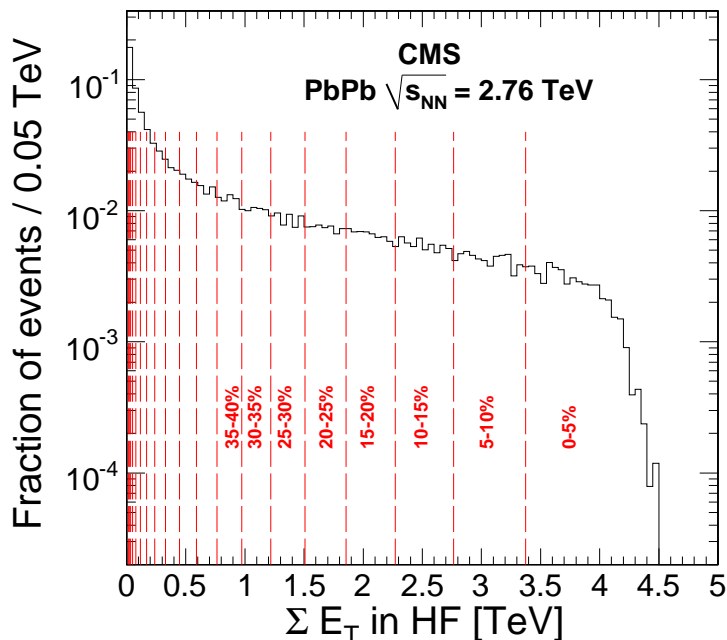
Figure 5.5: The distribution of the total transverse energy in the HF used to determine the centrality of a given PbPb collision. The centrality boundaries for each 5% interval are given by the dashed red lines [89].

model is implemented using the published TGLAUBERSC [138] software package developed by the PHOBOS Collaboration.

In this model, for a roughly spherical nucleus such as Pb, the position of each nucleon is distributed randomly according to a Woods-Saxon distribution [139] of the form

$$\rho(r) = \frac{\rho_0}{1 + e^{(R-r)/a}} \tag{5.6}$$

where R is the nucleon radius, a the surface thickness, and ρ_0 is a normalization factor chosen so that $\int d^3 r \rho(r) = 1$. If a nucleon is randomly chosen to be closer than some minimum internucleon distance, $d_{\min}$, to any other nucleon, then the position of the nucleon is recalculated.

The two randomly generated nuclei are then "collided" by choosing some random impact parameter, b, and shifting the transverse position of the nucleons by $\pm b/2$. The longitudinal coordinate parallel to the beamline plays no role in the collision. Any two nucleons with a transverse distance of less than $D = \sqrt{\sigma_{NN}/\pi}$, where σ_{NN} is the nucleon-nucleon inelastic cross section, are considered to form an inelastic collision, and thus N_{part} and N_{coll} may be tallied from all such "collisions". Examples of collisions from such a MC Glauber model at different impact parameters are shown in Fig. 5.6.

For the PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$, the model parameters are taken as $R = 6.62$ fm, $a = 0.546$ fm, $d_{\min} = 0.4$ fm, and $\sigma_{NN} = 64$ mb [140]. The number of participating nucleons corresponding to each centrality class used in the PbPb analysis as determined using these model parameters is shown in Table 5.1.
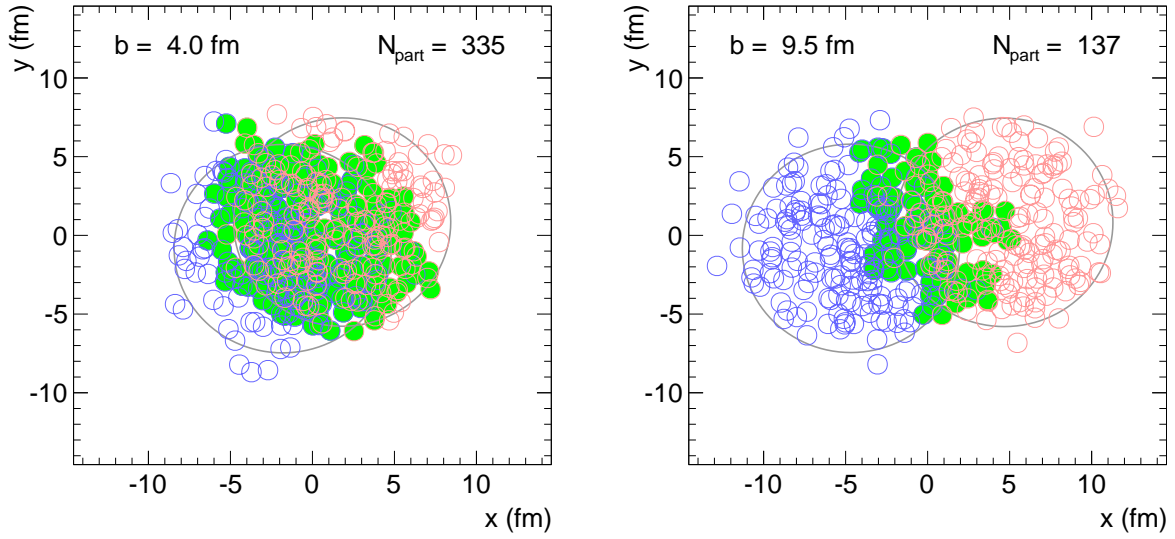
Figure 5.6: Schematic view of PbPb collisions with impact parameters of $b = 4.0$ fm (left) and $b = 9.5$ fm (right) as obtained from the Glauber model. Nucleons from each nucleus are designated by blue or red open circles respectively. The nucleons that participate in inelastic collisions are marked with filled green circles. The number of participating nucleons, N_{part}, is shown for each collision.

For pPb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$, the same parameters are used except for $\sigma_{NN} = 70$ mb. As the pPb spectrum is not measured in centrality classes, but only for the inclusive inelastic cross section, the only variable of interest is the average number of binary nucleon-nucleon collisions, $\langle N_{\mathrm{coll}} \rangle = 6.9 \pm 0.5$.

Centrality	$\langle N_{\mathrm{part}} \rangle$
0–5%	381 ± 2
5–10%	329 ± 3
10–15%	283 ± 3
15–20%	240 ± 3
20–25%	204 ± 3
25–30%	171 ± 3
30–35%	143 ± 3
35–40%	118 ± 3
40–50%	86.2 ± 2.8
50–60%	53.5 ± 2.5
60–70%	30.5 ± 1.8
70–80%	15.7 ± 1.1

Table 5.1: Average value of the number of participating nucleons ($\langle N_{\mathrm{part}} \rangle$) for each centrality class used in the analysis of PbPb collisions, along with the associated systematic uncertainty from the Glauber model. [44].

Chapter 6

Charged-Particle Spectra in PbPb and the Mean Transverse Momentum

6.1 Charged-Particle PbPb Spectra

The charged-particle PbPb spectra are measured in 12 centrality classes: 0–5%, 5–10%, 10–15%, 15–20%, 20–25%, 25–30%, 30–35%, 35–40%, 40–50%, 50–60%, 60–70%, and 70–80%. Additionally, spectra are measured in intervals of $|\eta|$ of $\Delta\eta = 0.4$, although only the spectra at midrapidity, $|\eta| < 0.4$, and the most forward rapidity, $2.0 < |\eta| < 2.4$, are presented.

These results extend previous CMS measurements of the charged-particle spectra [27] to lower-p_T, from 1.0 GeV/c to 0.3 GeV/c, to forward pseudorapidity, and into finer centrality classes.

To arrive at the invariant yield for a given centrality bin, the number of reconstructed tracks from all events in a given centrality class in given p_T and η bin, $N_{track}^{raw}(p_T, \eta, \mathrm{centrality})$, are corrected according to

$$E\frac{d^3 N_{ch}}{d^3 p}(p_\mathrm{T}, \eta, \mathrm{centrality}) = \frac{N_{track}^{raw}(p_\mathrm{T}, \eta, \mathrm{centrality}) \cdot w_{tr}(p_\mathrm{T}, \eta, \mathrm{centrality})}{2\pi p_\mathrm{T} \cdot \delta p_\mathrm{T} \cdot \delta\eta \cdot N^{selected}}, \tag{6.1}$$

where δp_T is the width of the p_T bin, $\delta\eta$ is the width of the η bin, and $N^{selected}$ is the number of selected events in the centrality class. The track weight w_{tr} is applied as a function of p_T, η, and centrality according to

$$w_{tr}(p_\mathrm{T}, \eta, \mathrm{centrality}) = \frac{1 - f}{e} \tag{6.2}$$

where e is the absolute efficiency and f is the fake rate as was determined in the HYDJET simulation. As the secondary and multiple reconstruction were well under 1%, no correction was applied for these small factors.

6.1.1 Evaluation of Systematic Uncertainties

Although the accuracy of the tracking efficiency measurement was validated using simulated pions embedded into PbPb collision events and by the methods listed in Sec. 4.3, a conservative estimate of the uncertainty on the tracking efficiency was determined by varying the selection criteria for the reconstructed tracks, re-evaluating the efficiency and fake rate in the HYDJET simulation, and then measuring the spectra in each case with the corresponding efficiency corrections. From this study,

a systematic uncertainty ranging from 5% at midrapidity to 13% at forward rapidity was assigned as a point-to-point systematic uncertainty on the spectra measurement.

As was shown in Fig. 4.8, the reconstruction efficiency varies depending on the species of the particle being reconstructed. Therefore, if the composition of particle species in the PbPb collisions is different than that of the HYDJET simulation, the measured efficiency will be incorrect. The inclusive charged-particle efficiency may be described as

$$\epsilon_{ch} = \frac{1}{N_{ch}} \left(\epsilon_\pi N_\pi + \epsilon_K N_K + \epsilon_p N_p + \epsilon_{\bar{p}} N_{\bar{p}} \right), \tag{6.3}$$

where ϵ_{ch} is the inclusive charged-particle efficiency, N_{ch} is the number of charged particles, ϵ_q is the efficiency of particle species q, and N_q is the number of particles of species q. Using this relationship, one may vary the N_q and assign a systematic uncertainty on ϵ_{ch} and the measured spectra due to the uncertainty in the particle composition.
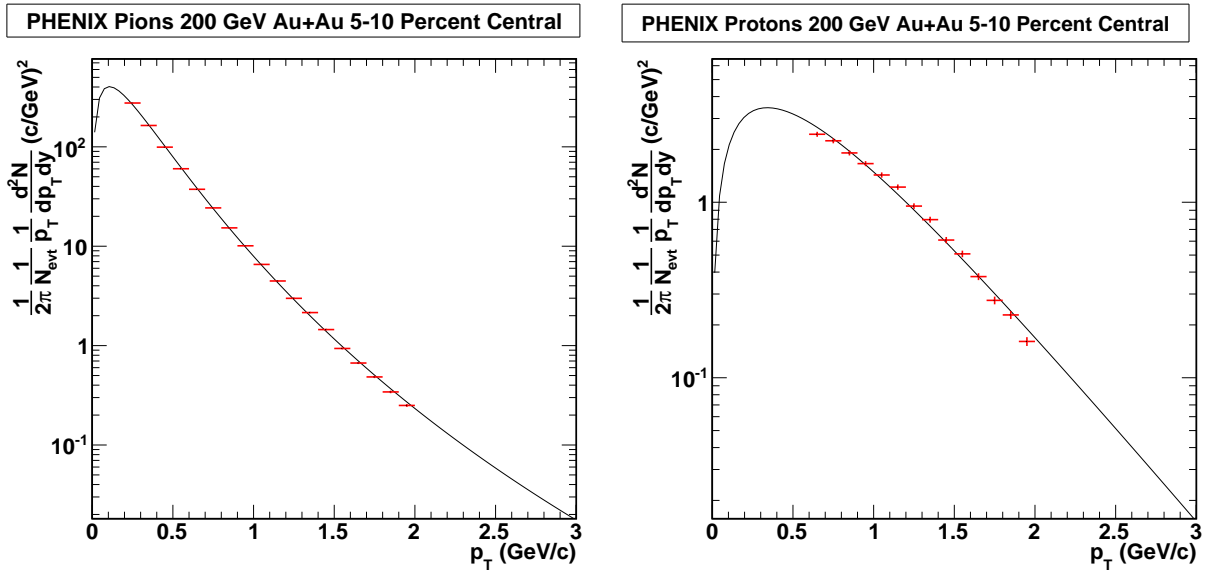
Figure 6.1: Extrapolation fits to pion and proton yields at mid-rapidity. Red points are experimental data from PHENIX AuAu $\sqrt{s_{NN}} = 200 \, \mathrm{GeV}/c$ collisions [141].

To calculate this uncertainty, the efficiencies ϵ_q are taken from Fig. 4.8. At the time when this measurement was performed, no measured spectra of pions, kaons, and protons for PbPb collisions at $\sqrt{s_{NN}} = 2.76 \, \mathrm{TeV}$ had been published. In the absence of this data, the N_q were estimated as a function of p_T by starting with the PHENIX Collaboration measurement of pion, kaon, and proton spectra in AuAu collisions at $\sqrt{s_{NN}} = 200 \, \mathrm{GeV}$ [141]. In order to produce estimates for the N_q for values of p_T to 0.3 GeV/c, these spectra were fit with a Tsallis distribution (see Sec. 6.2.1) and extrapolated to low-p_T, as shown in Fig. 6.1.

The measured spectra were corrected using the efficiency calculated using this procedure, and then again with the fraction of protons and antiprotons artificially increased by a factor of three, and once more with the fraction of protons and antiprotons artificially reduced by a factor of three. Even

with such an extreme variation in the particle composition, the measured spectra did not change very much. As a result of this study, a point-to-point systematic uncertainty of 1% at midrapidity and 2% at forward rapidity was assigned to the spectra measurement.

A final identified source of uncertainty in the spectra measurement relates to the uncertainty of the minimum-bias trigger and event selection used, which was determined to accept $97 \pm 3\%$ of all inelastic PbPb collision events. The centrality classes are determined by ranking all events in terms of total HF energy with the assumption that there is an additional 3% of events that are not selected by the trigger. These events that do not fire the trigger may be reasonably assumed to be in the 90-100% centrality range, and so the effect of the inefficiency is to simply compress the HF energy boundaries between event classes from 80%–0% One may visualize this by imagining the boundaries shown on Fig. 5.5 shifting slightly to the right. The result of this compression is very small for the most central 0–5% centrality class, but has a large effect for the 70–80% centrality class.

The efficiency classes may be recalculated by assuming a trigger efficiency of 94% or 100%, and the spectra may then be determined again by measuring events that fall into the new classes by HF energy. When the trigger efficiency is estimated at 100%, the 70–80% centrality class is shifted to events with a lower total HF energy and less charged particle production. The resulting spectrum is reduced. When the trigger efficiency is estimated at 94%, the 70–80% centrality class is shifted to events with a higher total HF energy, and the resulting spectrum is increased. The change in the measured spectra in each centrality class was found to be mostly an overall scaling factor, with the shape of the spectrum changing by no more than 3%. From this, a point-to-point systematic uncertainty of 3% was assigned to the spectra measurement for all pseudorapidity and centrality ranges. An overall normalization uncertainty was also assigned to the spectra measurement, ranging from 0.4% for the 0–5% centrality class, to 21% for the 70–80% centrality class. The value of this uncertainty for each centrality class is given in Table 6.1.

Table 6.1: Normalization uncertainty in the measurement of the PbPb charged-particle spectra in different centrality intervals resulting from the uncertainty in the minimum-bias trigger efficiency.

Centrality Range	Normalization Uncertainty
0-5%	0.4%
5-10%	1.0%
10-15%	1.7%
15-20%	2.3%
20-25%	3.1%
25-30%	4.1%
30-35%	5.0%
35-40%	6.1%
40-50%	8.0%
50-60%	12%
60-70%	16%
70-80%	21%

The point-to-point systematic uncertainties resulting from the three identified sources are given in Table 6.2 for each pseudorapidity range, along with the combined systematic uncertainty determined by adding the three uncertainties in quadrature. This total point-to-point uncertainty is the same for all centrality classes.

Table 6.2: Point-to-point systematic uncertainties in the measurement of the charged-particle PbPb spectra in different pseudorapidity intervals.

Source of Uncertainty	$|\eta| < 0.8$	$0.8 < |\eta| < 1.2$	$2.0 < |\eta| < 2.4$
Tracking Efficiency	5%	8%	13%
Particle Composition	1%	1%	2%
Trigger Efficiency	3%	3%	3%
Total	6%	9%	14%

6.1.2   Results

The charged-particle spectra of $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ PbPb collisions in each of the 12 centrality classes from 0–5% to 70–80%, at midrapidity ($|\eta| < 0.4$) and forward rapidity ($2.0 < |\eta| < 2.4$), are shown in Fig. 6.2. The combined statistical, point-to-point systematic, and normalization uncertainties are added in quadrature and shown in shaded bands. The spectra are scaled by arbitrary factors for visibility, as marked in the legend.
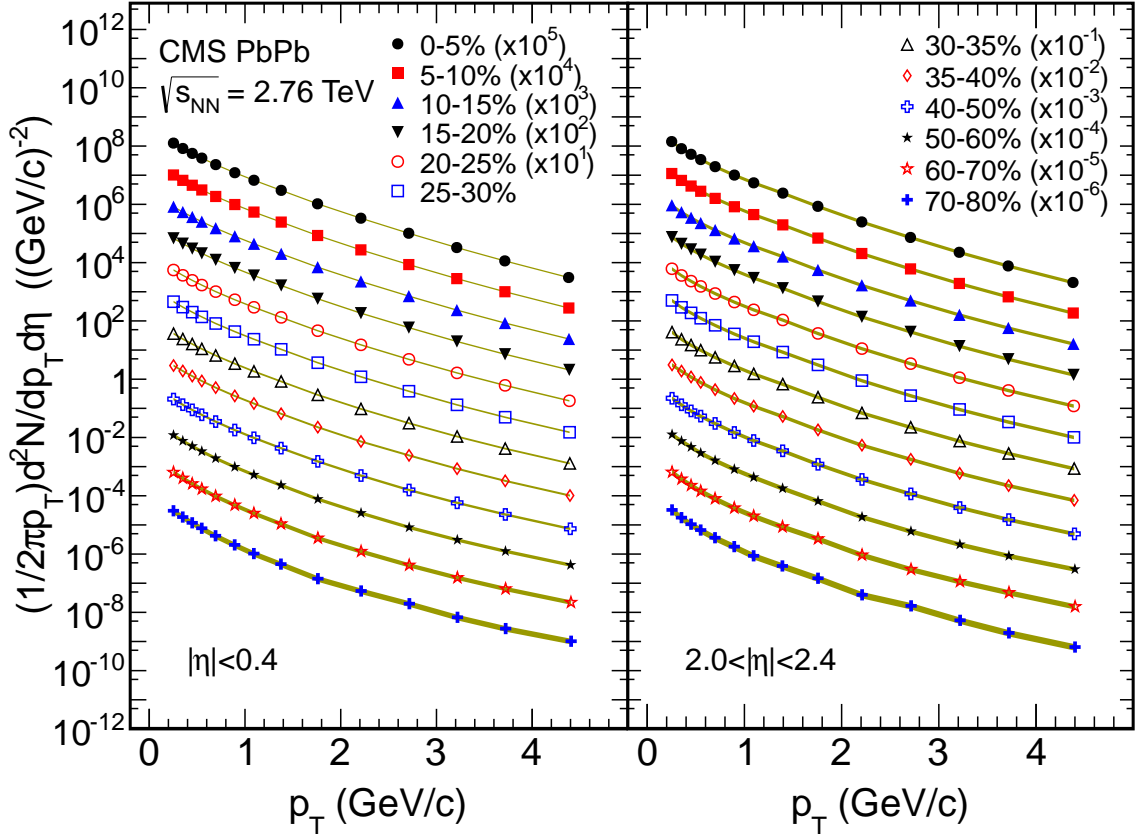
Figure 6.2: Charged-particle spectra at mid-rapidity (left) and forward rapidity (right), for the 12 centrality classes given in the legend. The distributions are offset by arbitrary factors given in the legend for clarity. The shaded bands represent the statistical and systematic uncertainties added in quadrature, including the overall normalization uncertainties from the trigger efficiency estimation [44].

6.2 Mean Transverse Momentum

In order to quantify the evolution of the charged-particle spectra in centrality and pseudorapidity, one may look at the mean p_T of the transverse momentum distributions. As the measured spectra are only determined down to $p_T = 0.3$, it is necessary to extrapolate them down to $p_T = 0$ in order determine $\langle p_T \rangle$. To properly compare these results to previous measurements, the $\langle p_T \rangle$ are presented in terms of the average N_{part} of each centrality class, as given on Table 5.1.

6.2.1 Extrapolation to Zero Transverse Momentum

As was performed in the previous $\langle p_T \rangle$ analysis of charged-particle spectra in pp collisions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV [137, 142], the measured PbPb spectra are fit with a Tsallis Distribution [143–145]:

$$E \frac{\mathrm{d}^3 N_{\mathrm{ch}}}{\mathrm{d}p^3} = \frac{1}{2\pi p_T} \frac{E}{p} \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta \mathrm{d}p_T} = C \left(1 + \frac{E_T}{nT}\right)^{-n}, \tag{6.4}$$

where $E_T = \sqrt{m^2 + p_T^2} - m^2$, and m is taken to be the charged pion mass. The fit is performed over the p_T range of 0.3 to 3.0 GeV/c. As with the previous pp collision analysis, $\langle p_T \rangle$ is calculated by integrating the Tsallis fit from $p_T = 0$ to 0.3 GeV/c, and then using the data points at higher p_T.

In order to evaluate the systematic uncertainty and ensure that the extrapolation is robust, two other functional forms were used to extrapolate to zero. The second fit function used is an exponential form in m_T,

$$\frac{1}{2\pi m_T} \frac{d^2 N}{dm_T d\eta} = \frac{1}{2\pi T (T + m)} \frac{dN}{d\eta} \exp\left[-(m_T - m)/T\right] \tag{6.5}$$

where the fit parameters are the reference mass m and T. The $\langle p_T \rangle$ is then calculated by extrapolating the fit function to $p_T = 0$ and integrating. The exponential fit reproduces the shape of the spectrum poorly, and so could only be fit over the limited range of $0.3 < p_T < 0.7$ in order to reproduce the lowest data point.

The third and final fit is a pQCD inspired power law function over the range $0.3 < p_T < 5.1$. This function, previously used by the STAR collaboration at RHIC [146], is given by the equation:

$$\frac{1}{2\pi p_T} \frac{dN}{dp_T} = C(1 + p_T/p_0)^{-n} \tag{6.6}$$

where C, p_0, and n are fit parameters. This was fitted over the range $0.3 < p_T < 5.0$.

The $\langle p_T \rangle$ at midrapidity determined using each of the three fits and extrapolations to $p_T = 0$ is shown for four representative centrality classes in Fig. 6.3. In order to cross-check the accuracy of
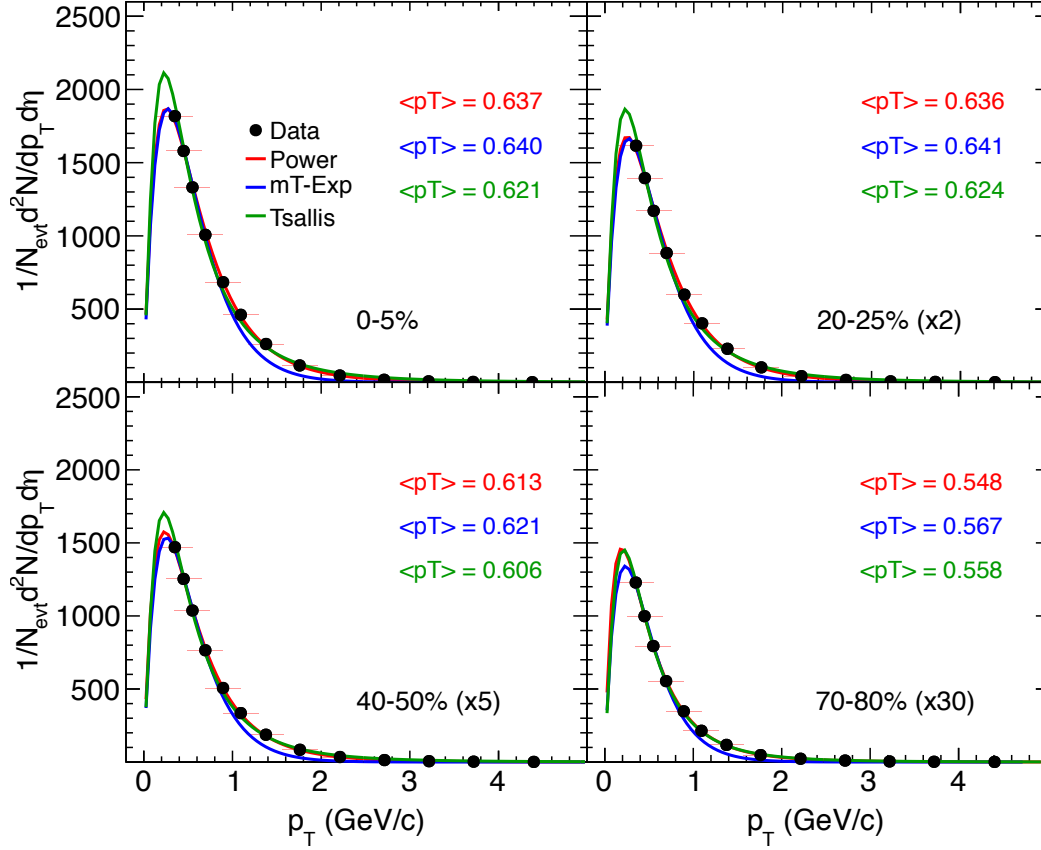
Figure 6.3: Charged-particle spectra in four representative centrality classes at midrapidity ($|\eta| < 0.4$). The Tsallis fit is superimposed in green, and the extracted $\langle p_{\mathrm{T}} \rangle$ and is shown on each plot. Additionally, two alternative fits are shown, specifically a m_T-exponential fit in blue, and a power law fit in red. The details of the three fits are described in the text.

the extrapolation and fit, the fitted spectra for $|\eta| < 0.4$ may be integrated over p_{T} to determine $dN/d\eta$ at midrapidity. At 5-10% centrality, for example, this results in $dN/d\eta = 1300 \pm 110$. This value may then be compared with the published CMS result in Ref. [89], which uses two methods independent of track reconstruction, hit-counting and "tracklets" from hit pairs, on a different set of PbPb collision data in which the magnetic field of the CMS detector was turned off. For 5-10% centrality, this published analysis finds that $dN/d\eta = 1313 \pm 45$. The integrated fit values agree well within systematic uncertainties for all centrality classes, indicating that the extrapolation to $p_{\mathrm{T}} = 0$ is reasonably accurate.

6.2.2   Evaluation of Systematic Uncertainties

Three sources of systematic uncertainty were evaluated. The first is due to the uncertainty on the extrapolation to $p_{\mathrm{T}} = 0$ as determined by the Tsallis fit. As shown in Fig. 6.3, this can be evaluated

by determining $\langle p_T \rangle$ with the results of the different fit and taking the difference between them as an estimate. A point-to-point systematic uncertainty of 3% at midrapidity and 4% at forward rapiditiy is applied to the measured $\langle p_T \rangle$ in all centrality classes from the uncertainty in the fit function used to extrapolate to $p_T = 0$.

The remaining sources of systematic uncertainty result from the uncertainty on the spectra themselves. $\langle p_T \rangle$ was evaluated using tracks reconstructed with the different quality cuts and corresponding correction factors to determine the systematic uncertainty due to tracking efficiency, and a resulting point-to-point uncertainty of 2% at midrapidity and 2.5% at forward rapidity is applied to the measured $\langle p_T \rangle$. A similar study was performed using tracks corrected with the assumption of various proton fractions of the charged-particle composition, but the effect of this variation was negligible and no corresponding uncertainty is applied. Finally, the spectra measured using the 100% and 94% trigger efficiency estimates were evaluated to determine $\langle p_T \rangle$, and from the small resulting differences, a point-to-point systematic uncertainty of 1.5% is applied to all rapidities and centrality classes.

Each of these uncertainties, as well as the total point-to-point systematic uncertainty resulting from the addition of each uncertainty in quadrature is shown in Table 6.3.

Table 6.3: Systematic uncertainty of the mean p_T of charged particles in PbPb collisions from each source and in total as a function of pseudorapidity.

Source of Uncertainty	$|\eta| < 0.4$	$0.8 < |\eta| < 1.2$	$2.0 < |\eta| < 2.4$
Fit function	3%	3%	4%
Trigger Efficiency	1.5%	1.5%	1.5%
Tracking Efficiency	2%	2%	2.5%
Total	3.9%	3.9%	4.9%

6.2.3 Results

The values of $\langle p_T \rangle$ as a function of the average N_{part} for each centrality class in pseudorapidity ranges of $|\eta| < 0.4$, $0.8 < |\eta| < 1.2$, and $2.0 < |\eta| < 2.4$ are shown in Fig. 6.4. In each pseudorapidity range, the values of $\langle p_T \rangle$ increase with N_{part} and then saturate around $N_{\text{part}} \approx 150$.

6.3 Discussion

This measurement of charged-particle spectra extends previous CMS measurements [27] to significantly lower p_T, and finer centrality classes. Additionally, this measurement of the spectra extends to forward rapiditiy. These spectra will help to constrain hydrodynamic models of the evolution of the QGP medium produced in the PbPb collision system, and the extension to forward rapidity may provide additional information about the longitudinal dynamics of the medium.

These spectra have also found utility in the measurement of elliptic flow [44] and higher-order flow [147] in PbPb collisions with CMS. In these measurements, the azimuthal anisotropy coeffi-
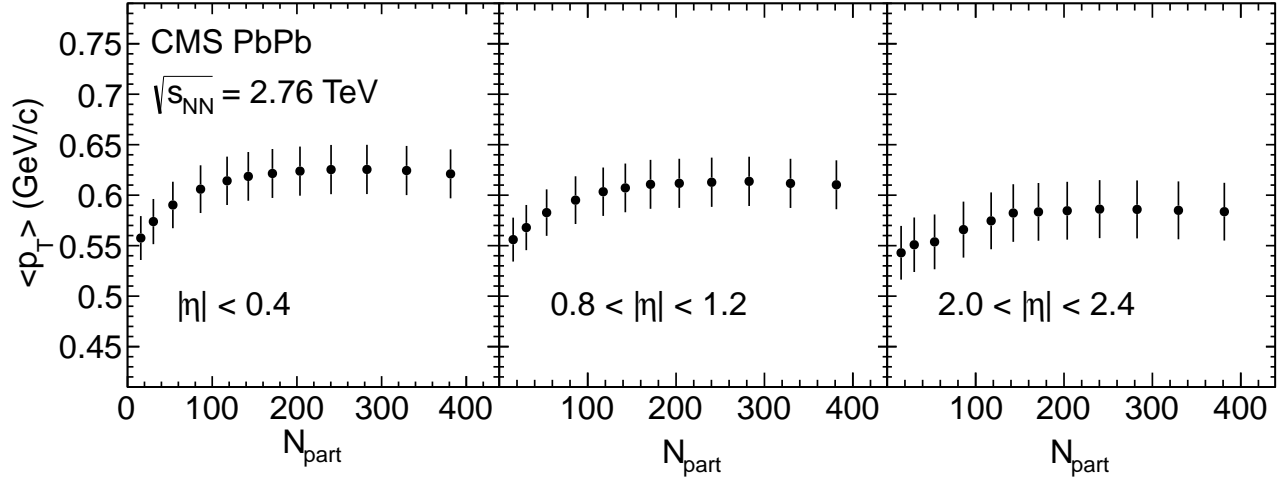
Figure 6.4: The mean transverse momentum of the charged-particle spectra as a function of N_{part} in three pseudorapidity intervals. The error bars represent the quadratic sum of the statistical and systematic uncertainties [44].

cients, v_2, v_3, ..., v_n, are measured as a function of both p_T and η. In order to make comparisons to various theory calculations and previous measurements in other collision systems at different energies, a yield-weighted average of $v_n(p_T, \eta)$ over p_T must be performed. The yield-weighted average v_n, often termed "integrated v_n", may also be determined as a function of pseudorapidity, $v_n(\eta)$, as the spectra measurements presented here extend over a broad η region. In Ref. [44], the integrated $v_2(\eta)$ is compared with similar measurements from the PHOBOS collaboration for AuAu collisions at $\sqrt{s_{NN}} = 200\,\text{GeV}$ [148], in order to test the hypothesis of extended longitudinal scaling at the very high energy of the LHC PbPb collision system. Here, extended longitudinal scaling refers to an observed scaling relationship in $v_2(\eta)$ [149] which is phenomenologically similar to the limiting fragmentation [150] seen in charged particle multiplicities as a function of η, and could be interpreted as an initial-state saturation effect. In Ref. [147], the integrated v_3 at midrapidity determined using two different methods with different sensitivities to event-by-event fluctuations are combined into a comparison with models incorporating potential CGC effects in the initial state [151]. Several other such comparisons are made using the yield-weighted average v_n, which are further described in Ref. [44, 147].

In Fig. 6.5, the value of $\langle p_T \rangle$ measured in PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ is compared with that measured in AuAu collisions at $\sqrt{s_{NN}} = 200\,\text{GeV}$ [152] as a function of the number of participating nucleons, N_{part}. In both collision systems, the same qualitative pattern of a slowing increase with N_{part} which saturates at $N_{part} \sim 150 - 200$ is observed. For all values of N_{part}, the value of $\langle p_T \rangle$ in the PbPb collisions is approximately 20% higher than in the AuAu collisions, indicating some combination of an increase in the initial energy density of the medium and an increase in the radial flow velocity of the medium in the higher energy collisions.

The Bjorken energy density of the medium at the time of formation, ε_{Bj}, [46] can be estimated as
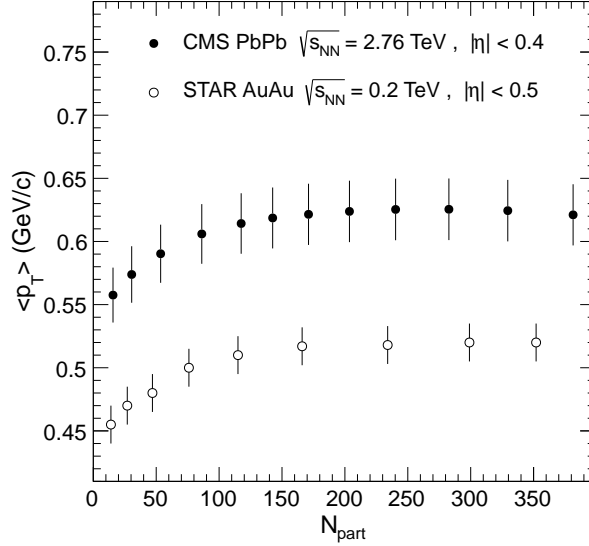
Figure 6.5: Mean transverse momentum of the charged-particle spectra as a function of N_{part} measured by CMS in PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$ (Closed circles) and by STAR [152] in AuAu collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ (Open circles). The error bars represent the quadratic sum of statistical and systematic uncertainties.

$$\varepsilon_{Bj} = \frac{1}{A_{\mathrm{T}}\tau_0}\frac{dE_{\mathrm{T}}}{dy} = \frac{1}{A_{\mathrm{T}}\tau_0}J(y,\eta)\frac{dE_{\mathrm{T}}}{d\eta} = \frac{1}{A_{\mathrm{T}}\tau_0}J(y,\eta)\langle m_{\mathrm{T}}\rangle\frac{dN}{d\eta}, \tag{6.7}$$

where dE_{T}/dy is the total transverse energy produced in the collision per unit rapidity, $J(y,\eta)$ is the Jacobian which depends on the mass and momentum distribution of the produced particles, $m_{\mathrm{T}} = \sqrt{m^2 + p_{\mathrm{T}}^2}$ is the transverse mass, τ_0 is the formation time, and A_{T} is the overlap area of the two nuclei.

Following Ref. [153, 154], we assume $\tau_0 = 1\,\mathrm{fm}/c$, and calculate $A_{\mathrm{T}} = \pi \times (7\,\mathrm{fm})^2$ for 0–5% central PbPb collisions. In order to estimate $dN/d\eta$, the charged particles are taken as comprising 2/3 of the total particle production, and the CMS measurement of $dN_{\mathrm{ch}}/d\eta = 1612\pm55$ at midrapidity [89], which is in excellent agreement with the value of $dN_{\mathrm{ch}}/d\eta$ obtained by integrating the extrapolated spectrum over p_{T}. The Jacobian is taken as $J(y,\eta) = 1.1$. Finally, $\langle m_{\mathrm{T}}\rangle$ can be estimated from the measured value of $\langle p_{\mathrm{T}}\rangle = 0.621\,\mathrm{GeV}/c$ by treating all produced particles as pions with $m = 0.140\,\mathrm{GeV}/c^2$.

With these values, one obtains an initial energy density of $\varepsilon_{Bj} \approx 11\,\mathrm{GeV}/c$ for the most central PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$. This is roughly twice as large as the energy density estimated in central AuAu collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ reported as $\varepsilon_{Bj} = 5.4 \pm 0.6\,\mathrm{GeV}$ by the PHENIX collaboration from the measurement of $dE_{\mathrm{T}}/d\eta$ [153]. From the CMS measurement of $dE_{\mathrm{T}}/d\eta$ [154], a larger estimate of $\varepsilon_{Bj} \approx 14\,\mathrm{GeV}/c$ is obtained for PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$. Although these estimates are only approximate, the larger value of ε_{Bj} obtained from the measurement of $dE_{\mathrm{T}}/d\eta$ may suggest that the fraction of neutral particles produced in the collision is larger than expected.

Chapter 7

Interpolation of pp Charged-Particle Spectra

The charged-particle transverse momentum spectra in pPb collisions measured here are taken at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$. As of 2014, no corresponding pp collisions have ever been produced and measured at $\sqrt{s} = 5.02\,\text{TeV}$, and there are no immediate plans to produce such collisions. In order to determine the nuclear modification factor of charged particles in pPb collisions, an estimate of the pp charged-particle spectrum at $\sqrt{s} = 5.02\,\text{TeV}$ must be somehow estimated from available collision data at different center-of-mass energies. Three general strategies are explored here:

- Using available pp collision data, one can directly interpolate between the measured spectra at each value of p_T. For the high-p_T portion of the spectra, this interpolation may also be performed taking advantage of a scaling relationship by plotting the spectra as a function of $x_\text{T} = 2p_\text{T}/\sqrt{s}$. This technique was used in Ref. [125] to estimate the $\sqrt{s} = 2.76\,\text{TeV}$ pp spectrum before collision data was available.

- From existing MC event generators, such as PYTHIA, one can simply predict the charged-particle pp collision spectrum at $\sqrt{s} = 5.02\,\text{TeV}$. Such a calculation is model-dependent, and is only approximately reliable if the generator has been tuned to collision data at similar energies. However, large discrepancies exist between generator tunes.

- By using an MC event generator, one can estimate the position of the $\sqrt{s} = 5.02\,\text{TeV}$ spectrum relative to those at $\sqrt{s} = 2.76\,\text{TeV}$ and $\sqrt{s} = 7\,\text{TeV}$ by simulating all three energies. The relative placement of the $\sqrt{s} = 5.02\,\text{TeV}$ spectrum determined in simulation is then applied to collision data at $\sqrt{s} = 2.76\,\text{TeV}$ and $\sqrt{s} = 7\,\text{TeV}$ to estimate the $\sqrt{s} = 5.02\,\text{TeV}$ spectrum.

To perform the direct interpolation, charged-particle spectra in pp collisions measured by the CMS Collaboration at $\sqrt{s} = 0.9$, 2.76, and 7 TeV are used [27,125]. To better constrain the interpolation, additional charged-particle spectra in p$\bar{\text{p}}$ collisions measured by the CDF Collaboration at the Tevatron at $\sqrt{s} = 0.63$, 1.8, and 1.96 TeV are also used [155,156].

7.1 Direct Interpolation Method

First step in performing a direct interpolation is to collect the invariant charged-particle cross sections or yields at various energies and transform them into a comparable form. This work was previously completed for Ref. [125], and is here simply expanded to include the $\sqrt{s} = 2.76\,\text{TeV}$ data from Ref. [27]. The selected spectra from CDF at $\sqrt{s} = 0.63$, 1.8, and 1.96 TeV are reported as inelastic cross sections for the average of positive and negative charged particles. The CMS spectra at $\sqrt{s} = 0.9$, 2.76, and 7 TeV are reported as yields of the sum of both positive and negative charged particles, and are taken from non-single diffractive collision events. For the purposes of comparison, the CMS data are scaled by a factor of 2 and the integrated luminosity of the sample to convert to cross sections.

7.1.1 Direct Interpolation by Transverse Momentum
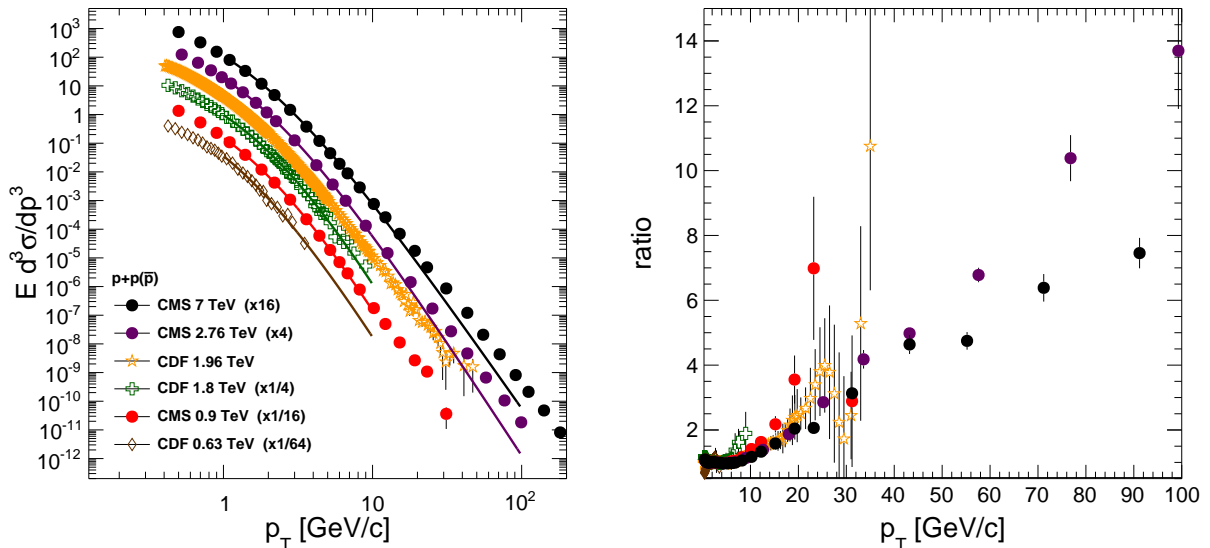
Figure 7.1: (Left) Invariant charged-particle cross sections for the $\sqrt{s} = 0.9, 2.76,$ and 7 TeV CMS pp collision data [27, 125], and then $\sqrt{s} = 0.63, 1.8,$ and 1.96 TeV CDF p$\bar{\text{p}}$ collision data [155, 156] given as a function of p_T. The cross sections are scaled by an arbitrary factor for visibility as given in the legend. Power law fits to the data in the range $1 < p_\mathrm{T} < 100$ GeV/c are shown as solid lines. (Right) Residuals of the data points with respect to the power law fits. The residuals from CDF have been smoothed.

In order to perform the interpolation it is necessary to first describe the cross sections at arbitrary p_T values, since the p_T bins of each measurement are different. The cross sections may be approximately described by a power law fit, given as

$$E\frac{d^3\sigma}{dp^3} = a(1 + p_\mathrm{T}/b)^c. \tag{7.1}$$

The results of such a fit are given in the left side of Fig. 7.1. These fits are not perfectly accurate, and so the residuals of the fit, as shown on the right side of Fig. 7.1, are further described by a 3rd order spline. With this combination of the power law fit and third order spline, the cross sections can now be determined at arbitrary p_T.

The cross sections may now be plotted for various values of p_T as a function of $\sqrt{s}$, as shown in Fig. 7.2 as a log-log plot. To interpolate between these values, a second order polynomial fit, is used. The interpolation was also attempted on a linear-linear plot of cross section versus $\sqrt{s}$, but displayed unphysical non-monotonic behavior and was therefore rejected.

This p_T based interpolation provides a good estimate, but suffers at high-p_T due to the fact that only two of the six measured datasets extend past $p_\mathrm{T} = 30$ GeV/c.
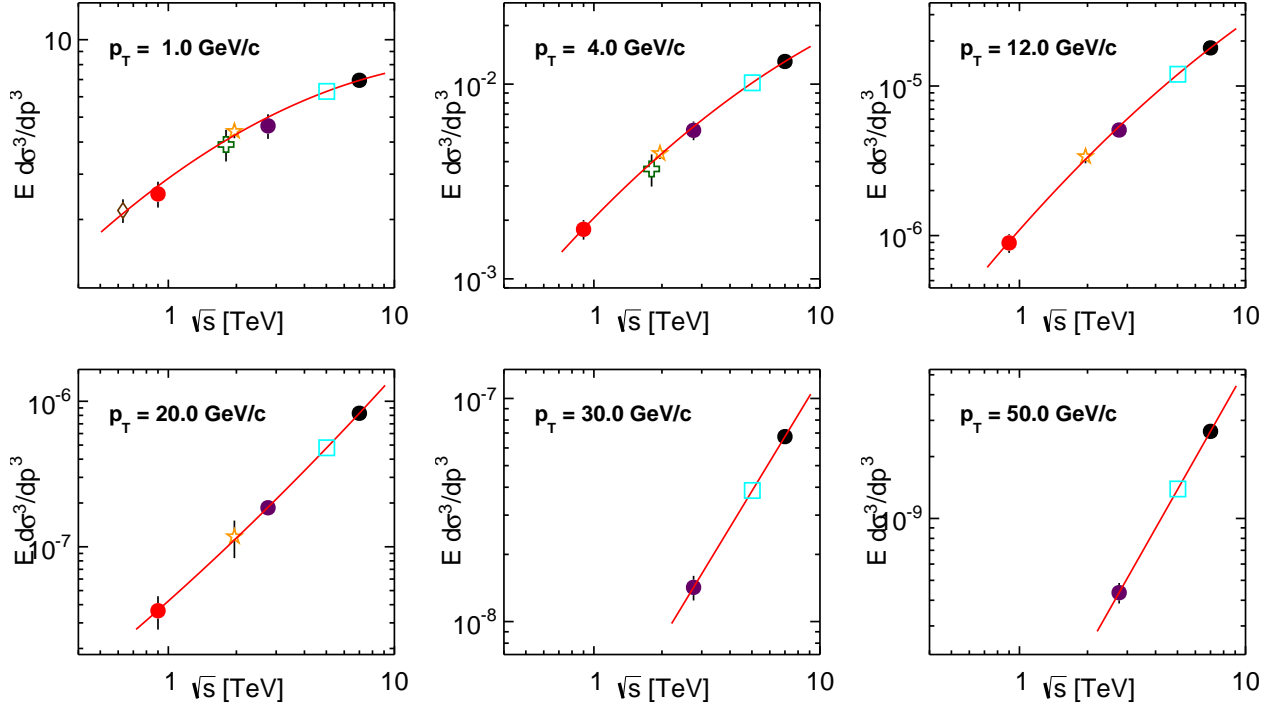
Figure 7.2: Second-order polynomial fits performed in the log-log cross section-$\sqrt{s}$ plane. The data points from CDF and CMS measurements are given using the colors described in Fig. 7.1. The 5.02 TeV prediction is given by the light blue square. A linear fit is performed in the case that there are only two measurements available to constrain the fit.

7.1.2 Direct Interpolation by a Scaling Variable

To provide a better interpolation at high-p_T using more of the measured spectra, one may look at the cross sections as a function of the scaling variable $x_T = 2p_T/\sqrt{s}$. This representation has the advantage that four of the previously measured cross sections, at $\sqrt{s} = 0.9$, 1.96, 2.76, and 7 TeV, extend above $x_T \approx 0.04$ which corresponds to $p_T \approx 100\,\text{GeV}/c$ at $\sqrt{s} = 5.02\,\text{TeV}$. This allows for a better constraint to the interpolation. Unfortunately, the x_T-based interpolation suffers from poor constraint at low-p_T, since few of the measured cross sections extend down to $x_T \approx 0.0004$ which corresponds to $p_T \approx 1\,\text{GeV}/c$ at $\sqrt{s} = 5.02\,\text{TeV}$.

This representation has a further advantage that when the cross sections are multiplied by $\sqrt{s}^{4.9}$, they exhibit an approximate scaling relationship as determined in Ref. [125]. This scaling relationship may result in a better interpolation determined in terms of $\sqrt{s}^{4.9}\,E\,d^3\sigma/dp^3$ as these values are closer together. As before, the first step in performing the interpolation is to describe the cross sections at arbitrary values of x_T. Again, the cross sections are first approximately described by a power law function

$$E\frac{d^3\sigma}{dp^3} = a(1 + x_T/b)^c, \qquad \text{or} \qquad \sqrt{s}^{4.9}\,E\frac{d^3\sigma}{dp^3} = a(1 + x_T/b)^c. \tag{7.2}$$

93

The results of these functional fits are shown in the top left of Fig. 7.3 for $\sqrt{s}^{4.9} E\, d^3\sigma/dp^3$, and in the top right for $E\, d^3\sigma/dp^3$. The residuals of these fits are shown in the bottom left and right, respectively, and are described by a third order polynomial function. This combination of the power law fit and the third order polynomial fit of the residuals describe the cross sections at arbitrary x_{T}.

The cross sections, $E\, d^3\sigma/dp^3$, are then determined at various values of x_{T}, and are shown in Fig. 7.4 for six representative values of x_{T} in a log-log plot. Again, a second order polynomial is used to perform the interpolation in this plane. This representation suffers in that the interpolation must be performed over several orders of magnitude, and because the functional fit is approximately linear in this plane.

The scaled cross sections, $\sqrt{s}^{4.9} E\, d^3\sigma/dp^3$, are also determined as a function of x_{T}, and shown for six representative x_{T} values in log-log plots in Fig. 7.5. Once again, a second order polynomial fit in the log-log plane is used to perform the interpolation. This interpolation is better constrained over a much smaller range than in the $E\, d^3\sigma/dp^3$-x_{T} plane.

The final pp reference spectrum at $\sqrt{s} = 5.02\,\mathrm{TeV}$ is determined using both the x_{T}-based ($\sqrt{s}^{4.9}$) interpolation and the p_{T}-based interpolation. Below a p_{T} of $12.5\,\mathrm{GeV}/c$, the p_{T}-based interpolation is used, and above a p_{T} of $13.5\,\mathrm{GeV}/c$, the x_{T}-based interpolation is used. Between the two values of p_{T} a linear weighting is implemented between the two measurements.

7.2 Relative Placement Method and Systematic Uncertainty

The relative placement method determines the charged-particle cross section at $\sqrt{s} = 5.02\,\mathrm{TeV}$ by using the position of the $\sqrt{s} = 5.02\,\mathrm{TeV}$ in PYTHIA relative to the $\sqrt{s} = 2.76$, and $7\,\mathrm{TeV}$ PYTHIA, spectra. For this method, the ProQ20 tune was chosen due to its relative success in describing the $\sqrt{s} = 2.76$, and $7\,\mathrm{TeV}$ data. The relative position at some value of p_{T}, $R_{\mathrm{MC}}(p_{\mathrm{T}})$, is determined by

$$R_{\mathrm{MC}}(p_{\mathrm{T}}) = \frac{Y_{\mathrm{MC}}(7\,\mathrm{TeV}, p_{\mathrm{T}}) - Y_{\mathrm{MC}}(5.02\,\mathrm{TeV}, p_{\mathrm{T}})}{Y_{\mathrm{MC}}(7\,\mathrm{TeV}, p_{\mathrm{T}}) - Y_{\mathrm{MC}}(2.76\,\mathrm{TeV}, p_{\mathrm{T}})}, \tag{7.3}$$

where $Y_{\mathrm{MC}}(\sqrt{s}, p_{\mathrm{T}})$ represents the cross section or yield for simulated NSD pp collisions at some transverse momentum p_{T} with a center-of-mass energy $\sqrt{s}$, computed by fitting the simulation with a power law function. The results of the PYTHIA simulations and the power law fits at the three center-of-mass energies are shown in Fig. 7.6.

The reference spectrum is then computed as

$$Y_{\mathrm{RP}}(5.02\,\mathrm{TeV}, p_{\mathrm{T}}) = Y_{\mathrm{Data}}(7\,\mathrm{TeV}, p_{\mathrm{T}}) - R_{\mathrm{MC}}(p_{\mathrm{T}}) \cdot (Y_{\mathrm{Data}}(7\,\mathrm{TeV}, p_{\mathrm{T}}) - Y_{\mathrm{Data}}(2.76\,\mathrm{TeV}, p_{\mathrm{T}})) \tag{7.4}$$

where $Y_{\mathrm{Data}}(\sqrt{s}, p_{\mathrm{T}})$ represents the charged-particle invariant yield measured with the CMS detector

at some transverse momentum p_T with a center-of-mass energy $\sqrt{s}$.

The systematic uncertainty of the interpolation is determined by comparing the various alternate interpolation strategies to the final combined p_T and x_T-based interpolated pp reference spectrum. In Fig. 7.7, the following methods are compared with the combined interpolation spectrum:

- The p_T-based interpolation for $p_\mathrm{T} > 12.5\,\mathrm{GeV}/c$

- The x_T-based interpolation performed for the unscaled cross section (as opposed to the $\sqrt{s}^{4.9}$ scaling factor)

- The relative placement method using the PYTHIA ProQ20 tune

- Simulated spectra at $\sqrt{s} = 5.02\,\mathrm{TeV}$ with the PYTHIA ProQ20 tune

- Simulated spectra at $\sqrt{s} = 5.02\,\mathrm{TeV}$ with the PYTHIA Z2 tune

The direct interpolation and relative placement methods generally agree within 10%, which will be taken as a point-to-point systematic uncertainty on the determination of the pPb nuclear modification factor. The pure MC generator result for the Z2 tune also agrees to this level with the interpolations, but the ProQ20 tune does not agree with the interpolations at low to mid-p_T. Because of the large differences between the various PYTHIA tunes, the pure simulation is not considered a reliable estimate of the $\sqrt{s} = 5.02\,\mathrm{TeV}$ pp spectrum at the desired level of accuracy.
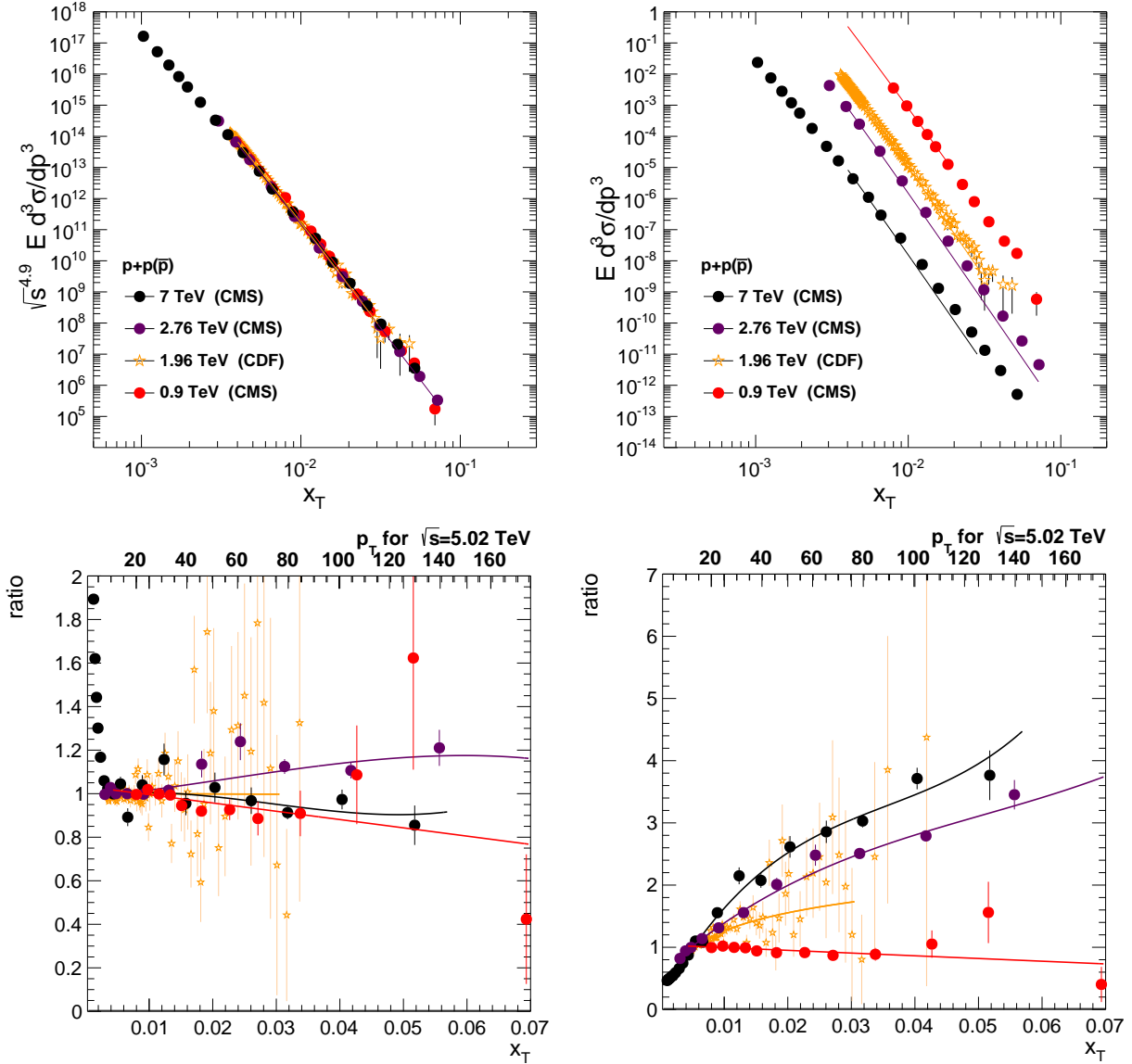
Figure 7.3: (Top) Invariant charged-particle cross sections (Left) and invariant cross sections scaled by $\sqrt{s}^{4.9}$ (Right) for the $\sqrt{s} = 0.9,\ 2.76,$ and $7\,\mathrm{TeV}$ CMS pp collision data [27, 125], and then $\sqrt{s} = 0.63,\ 1.8,$ and $1.96\,\mathrm{TeV}$ CDF p$\bar{\mathrm{p}}$ collision data [155, 156] given as a function of x_T. The cross sections are scaled by an arbitrary factor for visibility as given in the legend. Power law fits to the data are shown as solid lines. (Bottom) Residuals of the data points with respect to the power law fits to the cross sections (Left) and the scaled cross sections (Right) . The residuals from CDF have been smoothed.
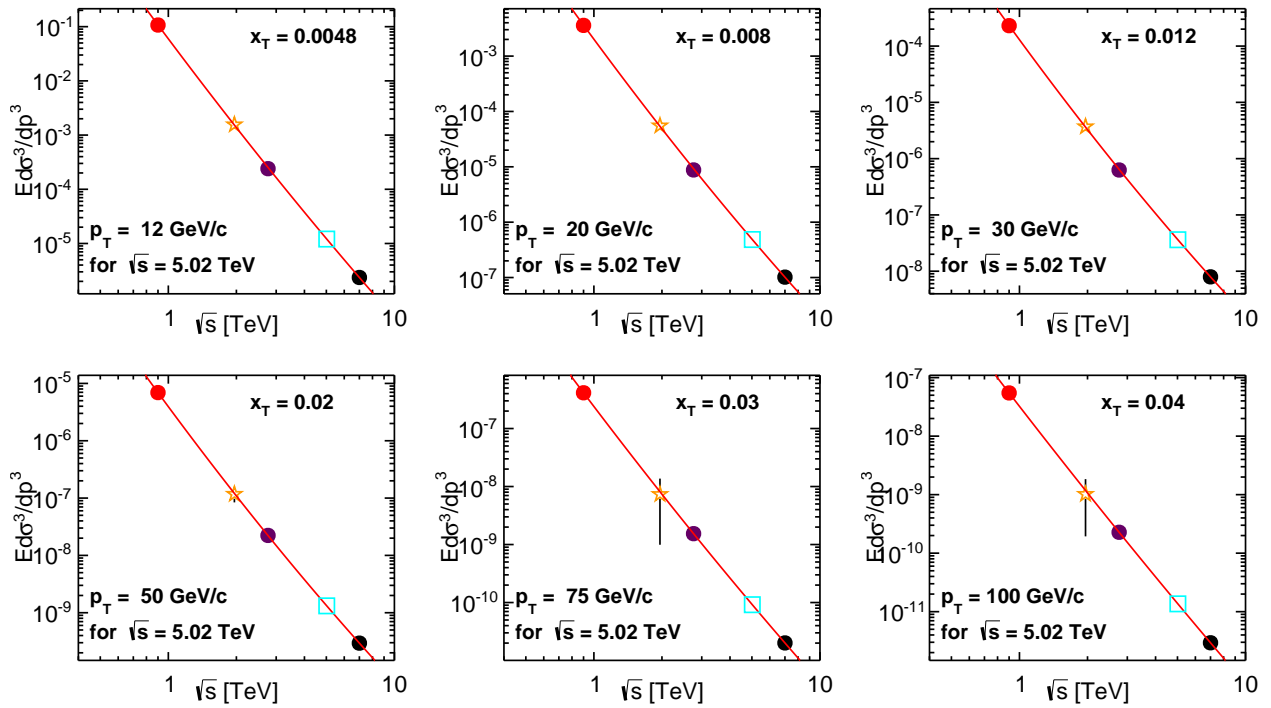
Figure 7.4: Second-order polynomial fits performed in the log-log cross section-$\sqrt{s}$ plane. The data points from CDF and CMS measurements are given using the colors described in Fig. 7.3. The $5.02\,\text{TeV}$ prediction is given by the light blue square.
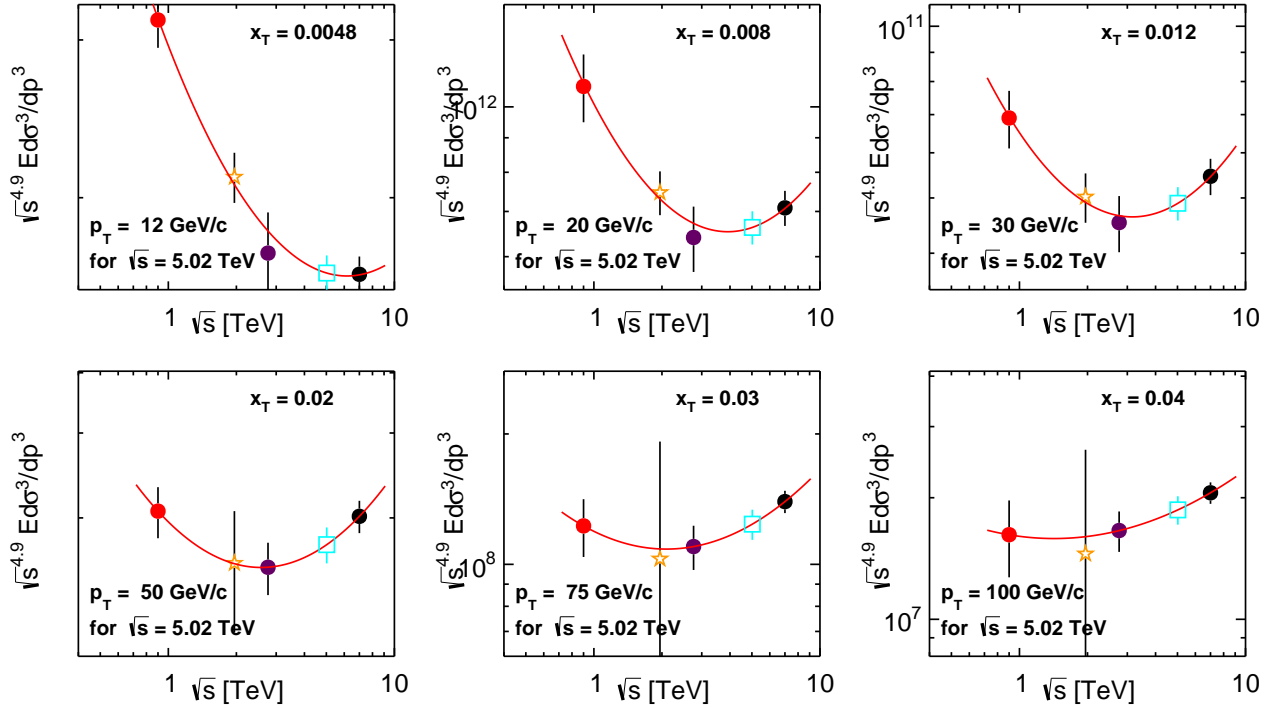
Figure 7.5: Second-order polynomial fits performed in the log-log $\sqrt{s}^{4.9}$ scaled cross section-$\sqrt{s}$ plane. The data points from CDF and CMS measurements are given using the colors described in Fig. 7.3. The 5.02 TeV prediction is given by the light blue square.
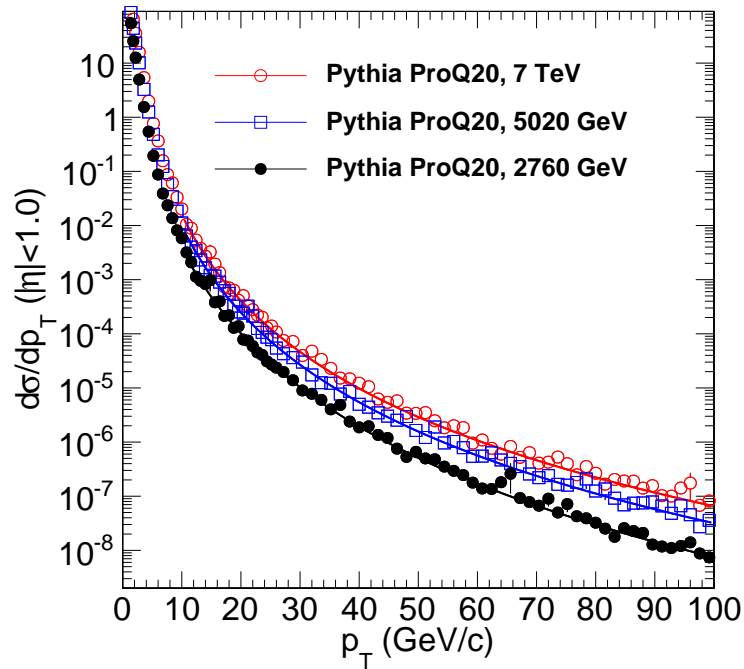
Figure 7.6: Charged-particle invariant cross sections from PYTHIA ProQ20 at $\sqrt{s} = 2060$, 5020 and 7000 GeV/c. The power law fits are shown as the solid lines.
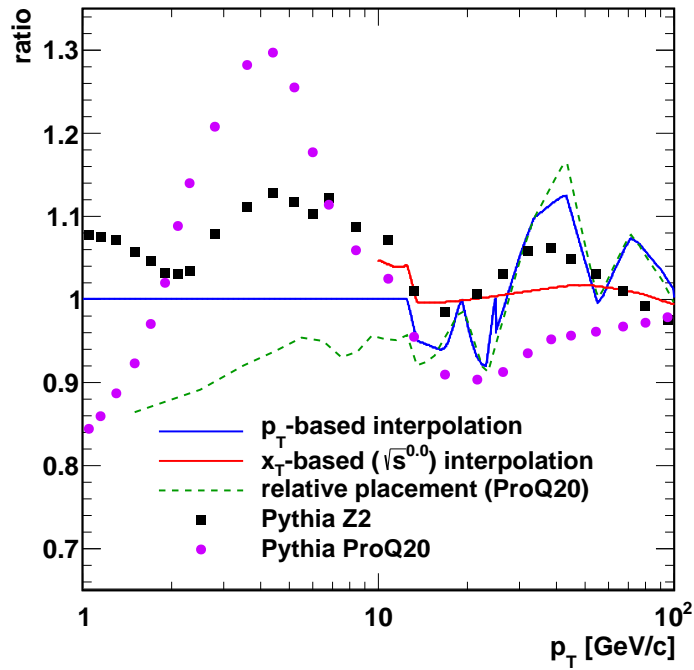
Figure 7.7: The ratio of the charged-particle invariant cross sections at $\sqrt{s} = 5.02\,\mathrm{TeV}$ determined using the various simulations and alternate interpolation methods described in the text to the charged-particle invariant cross section at $\sqrt{s} = 5.02\,\mathrm{TeV}$ determined using the combined p_{T} and x_{T} interpolation method.

Chapter 8

Charged-Particle Spectra in pPb and the Nuclear Modification Factor

8.1 Combination of Triggered pPb Collision Samples

As mentioned in Sec. 5.1.2 the minimum-bias trigger used for selecting pPb DS collision events was prescaled, and so approximately 1 out of every 3000 minimum-bias triggered events was actually recorded by the CMS detector in the 2013 run. With this limited data sample, the statistical reach of the charged-particle spectrum measurement does not extend beyond $p_T \approx 20 - 30 \, \text{GeV}/c$.

In order to achieve the goal of accurately measuring the charged-particle spectrum in pPb collisions to $p_T = 100 \, \text{GeV}/c$, one must incorporate additional samples of recorded events using more sophisticated trigger paths.

The HLT system of the CMS detector is capable of reconstructing full tracks using a similar algorithm to that detailed in Chapter 3. This algorithm is simplified for fast reconstruction, and is capable of providing trigger paths that accept events based on the maximum p_T of all tracks reconstructed in the event, the leading track p_T, which will be denoted by L_T. By using these additional trigger paths, it is possible to extend the statistical reach of the measurement, as is described in the following sections.

8.1.1 High Transverse Momentum Track Triggers

The High Level Trigger implements an iterative tracking procedure similar to the one used in the later, offline, reconstruction of tracks in a recorded event. The HLT algorithm differs in that fewer iterations are performed, leaving out any attempt to reconstruct tracks with low-p_T or substantial displacement from the primary vertex. Additionally, the quality selection criteria for the reconstructed tracks in the HLT are different from those of the offline reconstruction algorithm.

Three HLT trigger paths using these special tracks are of interest. Each one fires on events with at least one reconstructed quality track above a nominal threshold of $p_T = 12$, 20, or 30 GeV/c. These triggers are named Trk12, Trk20, and Trk30, respectively.

The single-track trigger used for minimum-bias DS collision events only requires the reconstruction of a pixel track with $p_T > 400 \, \text{MeV}/c$. These events are extremely numerous, and arrive too quickly during data taking for the HLT computing infrastructure to handle all such events. In order to reduce the number of events to be analyzed, the single track trigger path is gated by an L1 trigger path, meaning it only analyzes the event if a specific L1 trigger has fired. In this case the single-track trigger is gated by a zero-bias L1 trigger which fires on every bunch crossing. The zero-bias trigger is prescaled by a factor of several thousand. This ensures that the single-track trigger is highly efficient, but unfortunately only samples a small subset of the total luminosity.

Similarly, the HLT high-p_T tracking algorithm is too complex to be run on every event that is sampled by the CMS detector and must also follow a L1 gate. However, the L1 trigger system contains no information about the state of the silicon tracker or any capability to reconstruct tracks. As high-p_T charged particles produced in pPb collisions are overwhelmingly dominated by charged hadrons produced in jets, one can assume to a very good approximation that any high-p_T charged particle is part of a jet. Furthermore, the transverse momentum of the track must in principle always be less than that of the jet, and the leading fragment of a jet typically carries less than half the momentum of the jet. Events with high-p_T tracks therefore almost always contain a reconstructed jet with an E_T significantly higher than the p_T of the leading track.

The L1 trigger has implemented paths requiring energy deposits in the ECAL and/or HCAL within some identifiable cone that may be interpreted as a rough jet. One such path requires a cluster of energy deposits with $E_T \geq 12\,\mathrm{GeV}$, and is prescaled. Another path requires $E_T \geq 16\,\mathrm{GeV}$, and is not prescaled.

The Trk12 path is gated by the L1 jet $E_T \geq 12\,\mathrm{GeV}$ trigger, which means that it is already prescaled. A further prescale is required due to the maximum rate at which events may be recorded. There were too many events with tracks of $p_T > 12\,\mathrm{GeV}/c$ to all be recorded. The Trk20 path is also gated by the L1 jet $E_T \geq 12\,\mathrm{GeV}$ trigger, but has a smaller prescale factor. Finally, the rate of events with tracks of $p_T > 30\,\mathrm{GeV}/c$ was small enough that it was feasible to record all such events. To avoid any prescale, the Trk30 path is gated by the L1 jet $E_T \geq 16\,\mathrm{GeV}$ trigger.

8.1.2 Constant Efficiency Regions

The Trk12, Trk20, and Trk30 trigger paths in principle provide a means to extend the statistical reach of the charged-particle spectrum, but their use could also bias the result if the efficiency of the trigger varies with L_T, the p_T of the leading offline reconstructed track in the η range of interest in the event. In a typical analysis using specialized triggers such as these, one would attempt to determine the minimum p_T of the triggered object in which the trigger becomes fully efficient. In the case of the track triggers, due to the imperfect reconstruction efficiency of the tracking algorithm, the trigger never becomes fully efficient. If the Trk12 trigger is significantly more likely to accept an event with $L_T = 18\,\mathrm{GeV}/c$ than it is to accept an event with $L_T = 14\,\mathrm{GeV}/c$, then this could create a biased measurement. With this in mind, one may look for L_T regions where the trigger efficiency is constant.

The constancy of the trigger efficiency can be assessed by looking at the distribution of events with L_T values in some range of interest, and comparing the shape of that distribution to the same distribution for minimum-bias triggered events, which are effectively fully efficient for events with $L_T > 10\,\mathrm{GeV}/c$. In Fig. 8.1, the distribution of leading-tracks, only considering tracks within $|\eta| < 1.0$, is plotted for minimum-bias triggered events. The same distribution is shown for the Trk12 trigger, only it is normalized to match the minimum-bias triggered distribution. This is done by dividing by the number of events with a $L_T > 14\,\mathrm{GeV}/c$ in the Trk12 triggered sample, and multiplying by the number of events with $L_T > 14\,\mathrm{GeV}/c$ in the minimum-bias triggered sample. The ratio of these distributions is shown on the lower plot, and demonstrates that within statistical uncertainty, the relative shape of the Trk12 triggered sample is the same as the minimum-bias
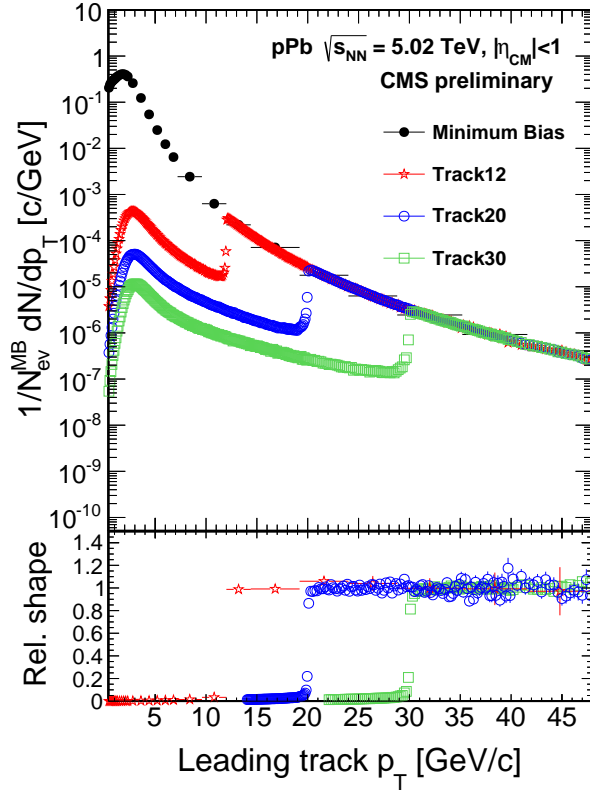
Figure 8.1: Upper panel: Charged-particle yields for various triggers as a function of leading track p_T. The track-triggered distributions are normalized by the number of leading tracks in regions not affected by the turn-on of the triggers. Lower panel: The constant efficiency regions of the various track triggers shown by the ratios of the leading track p_T distributions. The red histogram is the ratio of the 12 GeV/c and the minimum bias, the blue is the 12 GeV/c and the 20 GeV/c, and the green is the 22 GeV/c and 30 GeV/c track-triggered data [157].

sample, meaning that the trigger efficiency is roughly constant as a function of L_T, for $L_T > 14$ GeV/c.

Similarly, the L_T distribution for the Trk20 sample is normalized to the (already normalized) Trk12 sample by dividing by the number of events with $L_T > 22$ GeV/c in the Trk20 triggered sample, and multiplying by the number of events with $L_T > 22$ GeV/c in Trk20 triggered sample. The process is once again performed with the Trk30 sample for events with $L_T > 32$ GeV/c.

From this, it may be concluded that efficiency of the Trk12 trigger does not vary significantly for events with $L_T > 14$ GeV/c. Similarly the efficiencies of the Trk20 and Trk30 trigger do not vary significantly for events with an $L_T > 22$ GeV/c or $L_T > 32$ GeV/c, respectively.
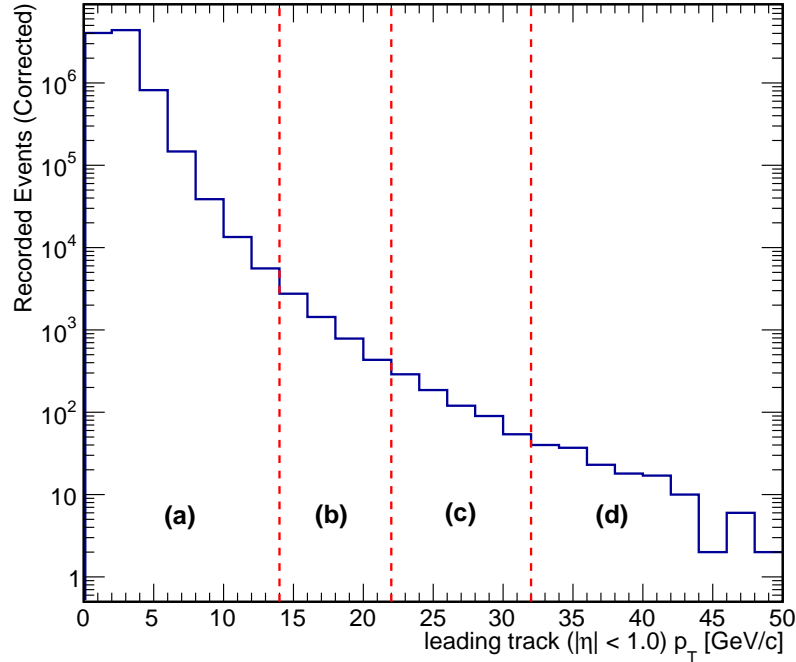
Figure 8.2: The leading-track p_T distribution ($|\eta| < 1.0$), L_T, for minimum-bias triggered events from the portion of the 2013 pPb run with the incoming lead ion traveling towards the positive z axis. Events are weighted in order to correct for trigger inefficiency. The dotted red lines separate different classes of events which are taken from separate trigger paths and then combined to form a full spectrum using the method described in the text. Region (a) is taken from the minimum-bias single-track trigger, (b) from the Trk12 trigger, (c) from the Trk20 trigger, and (d) from the Trk30 trigger.

8.1.3 Trigger Combination Method

In order to use all of the four triggers to obtain a combined and unbiased sample, events were split into four classes based on L_T and events from each class were taken from a different trigger as shown below:

(a) $0 < L_T < 14\,\text{GeV}/c$, taken from the MB (single-track) trigger.

(b) $14 < L_T < 22\,\text{GeV}/c$, taken from the Trk12 trigger.

(c) $22 < L_T < 32\,\text{GeV}/c$, taken from the Trk20 trigger.

(d) $32 < L_T\,\text{GeV}/c$, taken from the Trk30 trigger.

These L_T regions are shown in Fig. 8.2. To meaningfully add together the charged-particle spectra from each of these event classes, the high-p_T track triggered classes must be normalized using some procedure to correct for the fact that each triggered sample corresponds to a different number of DS collision events.

The procedure can be described as follows: First, the charged-particle spectrum is measured for all events in regions (a)–(d) from the MB triggered sample, but is not yet normalized to the number of events in the sample. All tracks from all events in regions (b)–(d) are then removed from the spectrum, and are replaced with those taken from the Trk12 triggered sample. One then divides these replacement tracks by the number of events from the Trk12 triggered sample in the region (b), and multiplies the replacement tracks by the number of events in the MB triggered sample in region (b). With this normalization, the MB triggered tracks in region (a), and the Trk12 triggered tracks in regions (b)–(d) both correspond to the number of events from the MB triggered sample, but the statistical reach of the measurement has been expanded.

This procedure continues with the removal of the tracks from events in regions (c)–(d), and their replacement with tracks from events in the Trk20 triggered sample from regions (c)–(d). These new tracks are then divided by the number of events from the Trk20 triggered sample in region (c), and multiplied by the number of previously normalized events from the removed sample in region (c). Once again, the tracks in regions (a), (b), and (c)–(d) are all normalized to correspond to the number of events from the MB triggered sample. This staged normalization procedure continues one more time, removing the tracks from events in region (d), replacing them with those from the Trk30 trigger, and performing the appropriate normalization.

This procedure can be expressed mathematically as

$$
\frac{1}{N_{\text{evt,MB}}} \frac{dN_{\text{ch}}}{dp_{\text{T}}} = \frac{1}{N_{\text{evt,MB}}} \left[\frac{dN_{\text{ch}}}{dp_{\text{T}}} \right]_{\text{MB}}^{L_{\text{T}}<14} + \frac{1}{N_{\text{evt,Trk12}}} \left[\frac{dN_{\text{ch}}}{dp_{\text{T}}} \right]_{\text{Trk12}}^{14<L_{\text{T}}<22}
$$
$$
+ \frac{1}{N_{\text{evt,Trk20}}} \left[\frac{dN_{\text{ch}}}{dp_{\text{T}}} \right]_{\text{Trk20}}^{22<L_{\text{T}}<32} + \frac{1}{N_{\text{evt,Trk30}}} \left[\frac{dN_{\text{ch}}}{dp_{\text{T}}} \right]_{\text{Trk30}}^{32<L_{\text{T}}}
$$

$$(8.1)$$

where $[dN_{\text{ch}}/dp_{\text{T}}]_{TRG}^{\mathcal{R}}$ is the un-normalized charged-particle p_{T} spectrum taken from events in leading-track p_{T} region $\mathcal{R}$ from the sample taken from trigger path TRG. $N_{\text{evt,MB}}$ is simply the number of events in the minimum-bias triggered sample. The other normalization factors, $N_{\text{evt,Trk12}}$, $N_{\text{evt,Trk20}}$, and $N_{\text{evt,Trk30}}$ are defined as

$$
N_{\text{evt,Trk12}} = N_{\text{evt,MB}} \frac{N_{\text{Trk12}}^{14<L_{\text{T}}<22}}{N_{\text{MB}}^{14<L_{\text{T}}<22}},
$$

$$(8.2)$$

$$
N_{\text{evt,Trk20}} = N_{\text{evt,12}} \frac{N_{\text{Trk20}}^{22<L_{\text{T}}<32}}{N_{\text{Trk12}}^{22<L_{\text{T}}<32}}, \text{ and}
$$

$$(8.3)$$

$$
N_{\text{evt,Trk30}} = N_{\text{evt,20}} \frac{N_{\text{Trk30}}^{32<L_{\text{T}}}}{N_{\text{Trk20}}^{32<L_{\text{T}}}},
$$

$$(8.4)$$

where $N_{TRG}^{\mathcal{R}}$ is the number of events in leading-track p_{T} region $\mathcal{R}$ from the sample taken from trigger path TRG.

8.1.4 Simulation for Trigger Combination Validation

There are a number of potential concerns that may be raised regarding the trigger combination method related to the different tracking algorithms run offline and by the HLT. For example, an event that would have resulted in a reconstructed offline track may fail to be reconstructed in the HLT algorithm. The high-p_T track triggers may naturally be more likely to accept events with more than one track above the p_T threshold. Tracks that are reconstructed in the HLT may not be selected in the offline reconstruction.

In principle, these effects may all be understood by generating a large number of simulated collision events, applying the trigger combination method using the events that were selected by the simulated triggers, and compare the result to the charged-particle spectrum that was generated by the simulation.

Unfortunately, the number of events that would need to be generated is prohibitively large. For example, to generate a pp collision event in PYTHIA Z2 with a charged particle of $p_T \sim 90 - 100\,\mathrm{GeV}/c$, one would typically require a QCD $2 \to 2$ hard scattering process with $\hat{p}_T > 220\,\mathrm{GeV}/c$. As given in Table 4.2, the cross section for such a process is approximately 2.5×10^{-6} mb. The cross section of an inelastic pp collision event is approximately 68 mb. Therefore, to generate just a few charged particles with $p_T \sim 90 - 100\,\mathrm{GeV}/c$, one would need to compute on the order of 100 million inelastic pp collisions. In order to precisely determine the charged-particle per-event yield at $p_T \sim 90 - 100\,\mathrm{GeV}/c$, many billions of events may be required. It is therefore not computationally feasible to perform such a simulation.

In order to reduce the number of simulated collisions required to reach a reasonable yield of charged particles with $p_T \sim 90 - 100\,\mathrm{GeV}/c$, one may simulate PYTHIA Z2 events with a QCD $2 \to 2$ hard scattering process with $\hat{p}_T > 15\,\mathrm{GeV}/c$. Such a simulation produces a charged-particle spectrum that has a reasonably realistic shape for $p_T \gtrsim 3 - 5\,\mathrm{GeV}/c$, and with the overwhelming majority of events not containing any charged particles with $p_T > 14\,\mathrm{GeV}/c$. Such a simulation may still serve to validate that the trigger combination method is sound, with the need to only simulate 500 million events.

While the generation of 500 million such events with PYTHIA Z2 may be reasonably performed using available computing resources, the simulation of the detector and subsequent reconstruction is too resource intensive to be realistically considered. A simplified simulation is therefore required which preserves the relevant features of the offline and HLT track reconstruction and is capable of answering the relevant concerns. The following procedure was developed to simulate the trigger combination method:

- Charged particles are generated using PYTHIA Z2 events with a QCD $2 \to 2$ hard scattering process with $\hat{p}_T > 15\,\mathrm{GeV}/c$. The p_T and η of each charged particle is recorded.

- The charged particle is said to be reconstructed by the offline reconstruction algorithm if a randomly generated number between 0 and 1 is less than the measured offline reconstruction efficiency for the p_T and η of the particle.

- The probability that a charged particle is reconstructed by the HLT algorithm is not independent of the probability that it is reconstructed by the offline algorithm. It was therefore
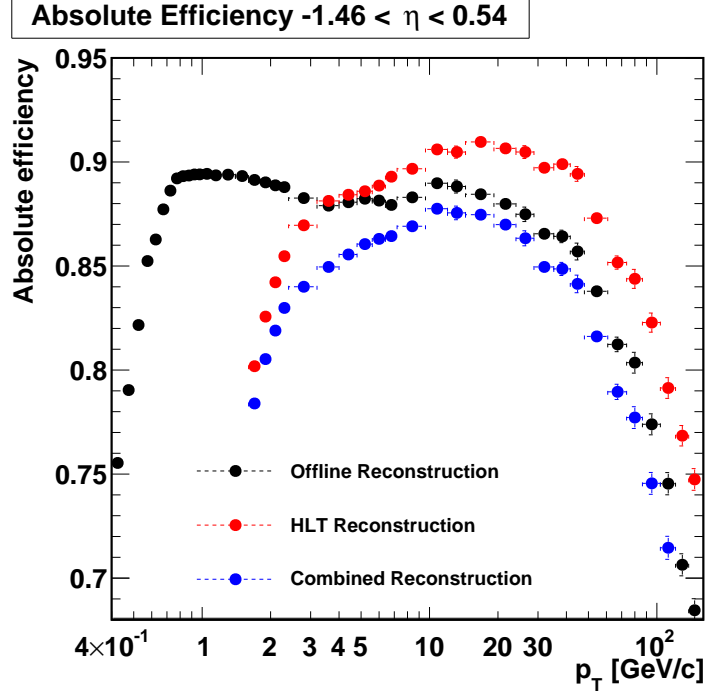
Figure 8.3: Absolute efficiency of charged particles in the offline reconstruction (black), and the HLT reconstruction (red) as a function of p_T measured using the same sample and method as in Fig. 4.10. The joint efficiency of both algorithms, defined as the probability that a given primary charged particle is reconstructed by both the offline and HLT algorithms, is shown in blue.

necessary to measure the joint probability that a particle is reconstructed by both algorithms, as shown in Fig. 8.3. The charged particle is then said to be reconstructed by the HLT algorithm if a randomly generated number between 0 and 1 is less than the conditional probability that the HLT algorithm reconstructs a charged particle given that the offline algorithm did or did not reconstruct the particle.

- If the event has a track above threshold for the Trk12, Trk20, or Trk30 trigger, it is said to be selected for that trigger. The MB trigger selects all events.

- Each trigger is then prescaled by randomly selecting triggered events. The MB trigger is given a prescale factor of 3000, the Trk12 trigger 20, the Trk20 trigger 5, and the Trk30 trigger was unprescaled.

- If a trigger is fired and the event not prescaled away, the charged particles said to be reconstructed by the offline algorithm are recorded into the appropriate trigger sample.

- The triggered samples are then combined according to the trigger combination procedure described in Sec. 8.1.3.

- The combined spectrum is then corrected for the inefficiency of the offline reconstruction algorithm.

- The corrected spectrum is then compared with that of the generated charged particles nor-

malized to the number of generated events.

The result of this 500 million event simulation are shown in Fig. 8.4. The corrected spectra derived from the combined sample agree with the generated charged-particle spectra at very high-p_T to 3% accuracy. The small 3% discrepancy is the result can be attributed to the fact that the difference between the offline reconstruction efficiency and the joint offline and HLT reconstruction efficiency changes slightly with p_T. From this simulated study, it is concluded that the trigger combination method accurately reproduces the charged-particle spectrum, in spite of differences between the offline and HLT reconstruction algorithms.
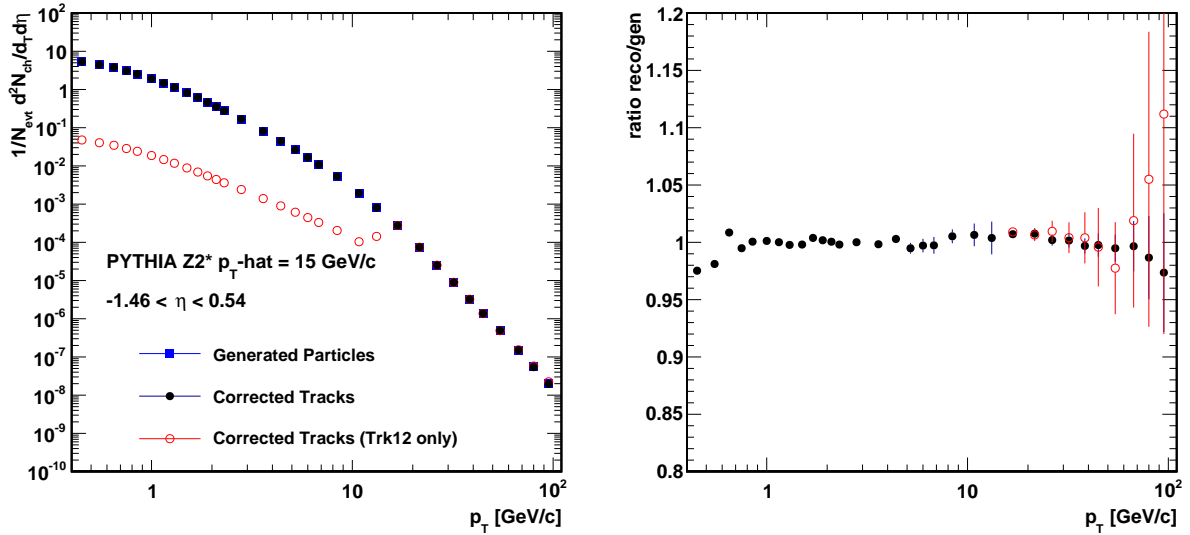
Figure 8.4: (Left) Charged-particle spectra generated with PYTHIA Z2 simulation (blue), reconstructed tracks from the simulation of the trigger combination method described in the text (black), and reconstructed tracks from just the Trk12 sample normalized to the number of equivalent MB events (red). (Right) ratio of the reconstructed tracks from the combined sample (Black), and the normalized Trk12 sample (red) to the generated spectra.

8.2 Rapidity Boost to the Center-of-Mass Frame

In the pPb collisions taken in 2013, the beam energies were not equal, with the $E_{\mathrm{p}} = 4.0\,\mathrm{TeV}$ and $E_{\mathrm{Pb}} = 1.58\,\mathrm{TeV}$, where E_{p} is the energy of the proton and E_{Pb} is the energy of a nucleon in the Pb nucleus. The center-of-mass frame of the collision is therefore boosted relative to the lab frame towards the direction of the p beam. The pp reference spectra are evaluated for $|\eta < 1.0|$, where the center-mass-frame is the same as the lab frame. For the pPb collision, one must distinguish the pseudorapidity in the lab frame, η_{lab}, from the pseudorapidity in the center-of-mass frame η_{CM}, and if possible, measure the spectrum in some kinematic region that corresponds to $|\eta_{\mathrm{CM}}| < 1.0$ in some reasonable approximation. Taking as a convention that the incoming proton moves in the positive z direction, the velocity of the center-of-mass frame, β, and center-of-mass energy, E_{CM}, can be found by Lorentz transforming the total energy and momentum, i.e.

$$\begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} E_{\mathrm{p}} + E_{\mathrm{Pb}} \\ p_{\mathrm{p}} + p_{\mathrm{Pb}} \end{bmatrix} = \begin{bmatrix} E_{\mathrm{CM}} \\ 0 \end{bmatrix}. \tag{8.5}$$

Solving for β and E_{CM}, and approximating the incoming proton and nucleon as massless, one finds

$$\beta = \frac{p_{\mathrm{p}} + p_{\mathrm{Pb}}}{E_{\mathrm{p}} + E_{\mathrm{Pb}}} \approx \frac{4 - 1.58}{4 + 1.58} = 0.434, \text{ and} \tag{8.6}$$

$$E_{\mathrm{COM}} = \sqrt{1 - \beta^2}(E_{\mathrm{p}} + E_{\mathrm{Pb}}) = 5.02\,\mathrm{TeV}. \tag{8.7}$$

Rewriting in terms of rapidity,

$$y = \tanh^{-1}(\beta) = \frac{1}{2} \ln\left(\frac{1 + \beta}{1 - \beta}\right) = 0.465. \tag{8.8}$$

The 2013 pPb data was taken in two periods, with the Pb ion moving towards positive z in the first period, and the proton moving towards z in the second. Therefore, in the second period $y_{\mathrm{CM}} = y_{\mathrm{lab}} + 0.465$, and in the first period, $y_{\mathrm{CM}} = 0.465 - y_{\mathrm{lab}}$. For sufficiently high-$p_{\mathrm{T}}$ particles, $y \approx \eta$, and therefore the range $|\eta_{\mathrm{CM}}| < 1.0$ is equivalent to $-0.535 < \eta_{\mathrm{CM}} < 1.465$ for the second period, or $-1.465 < \eta_{\mathrm{CM}} < 0.535$ for the first.

In order to reliably measure unidentified charged particles in the range $|\eta_{\mathrm{CM}}| < 1.0$, the minimum-$p_{\mathrm{T}}$ in which one may validly approximate the rapidity shift with a pseudorapidity shift must be determined. In Fig. 8.5, the error of such an approximation is explored. For $p_{\mathrm{T}} < 1.0\,\mathrm{GeV}/c$, the charged-particle yield is dominated by low-mass charged pions. As is demonstrated in the figure, for any charged pion with $p_{\mathrm{T}} \gtrsim 0.4\,\mathrm{GeV}/c$ in the tracker acceptance of $|\eta_{\mathrm{lab}}| < 2.4$, the center-of-mass pseudorapidity is well approximated by a shift in the lab pseudorapidity. For charged protons, this approximation is sound $p_{\mathrm{T}} \gtrsim 1.0\,\mathrm{GeV}/c$, where the total charged-particle yield begins to contain a significant fraction of protons. Based on this analysis, it was concluded that the charged-particle p_{T} spectrum for $p_{\mathrm{T}} < 0.4\,\mathrm{GeV}/c$ in $|\eta_{\mathrm{CM}}| < 1.0$ may be determined to an accuracy well below expected
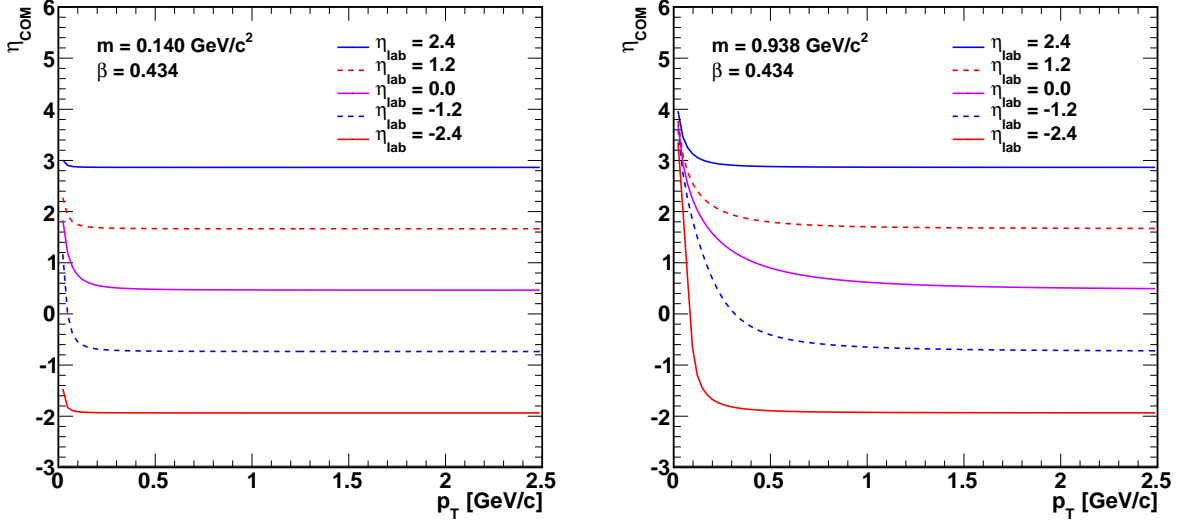
Figure 8.5: The pseudorapidity in the center-of-mass frame, η_{CM}, as a function of p_{T}, for charged pions (Left) and protons (Right), given for five different pseudorapidities in the lab frame, η_{lab} as marked in the legend. Here it is assumed that the colliding proton moves in the positive z direction in the lab frame.

systematic uncertainties simply by measuring the charged-particle p_{T} spectrum in a suitably shifted pseudorapidity range.

8.3 Charged-Particle pPb Spectra

To arrive at the invariant yield for double-sided pPb collisions, the number of reconstructed tracks from each event taken from each trigger path TRG with a total reconstructed track multiplicity M in some p_{T} and η bin, $N_{TRG}^{raw}(p_T, \eta, M)$, are corrected and summed according to

$$E\frac{d^3 N_{ch}}{d^3 p}(p_{\text{T}}, \eta) = \frac{\sum_{TRG}\sum_{\text{events}} N_{TRG}^{raw}(p_{\text{T}}, \eta, M) \cdot w_{tr}(p_{\text{T}}, \eta) \cdot w_{\text{evt}(M)}}{2\pi p_{\text{T}} \cdot \delta p_{\text{T}} \cdot \delta \eta \cdot N_{\text{evt},TRG} \cdot (1 + f_{\text{evt}}^{M=0})} \tag{8.9}$$

where TRG ranges over the MB, Trk12, Trk20, and Trk30 trigger paths, δp_{T} is the width of the p_{T} bin, $\delta \eta$ is the width of the η bin, $N_{\text{evt},TRG}$ is the appropriate scaled number of events for the trigger path TRG as defined in Sec. 8.1.3, $w_{\text{evt}}(M)$ is the trigger efficiency correction to DS collisions as described in Sec. 5.3 for an event with multiplicity M, and $f_{\text{evt}}^{M=0}$ is the fraction of DS events with reconstructed track multiplicity zero. The track weight, $w_{tr}(p_{\text{T}}, \eta)$ is applied as a function of p_{T} and η according to

$$w_{tr}(p_T, \eta) = \frac{(1 - F) \cdot (1 - S)}{A \cdot \varepsilon^{tr} \cdot (1 + D)} \tag{8.10}$$

where $A \cdot \varepsilon^{tr}$ is the absolute efficiency, D is the fraction of tracks corresponding to the same, multiply-reconstructed charged particle, S the fraction of tracks corresponding to a non-primary charged particle, and F is the fraction of fake tracks that do not correspond to any charged particle.

8.3.1 Wide-Binning Effect and Momentum Resolution

Two additional effects that may distort the measured spectrum must be considered, especially when extending the spectrum to high-p_T. The first is what may be described as the wide-binning effect, which may distort a steeply falling spectrum if measured in relatively wide bins in terms of $\delta p_T / p_T$.
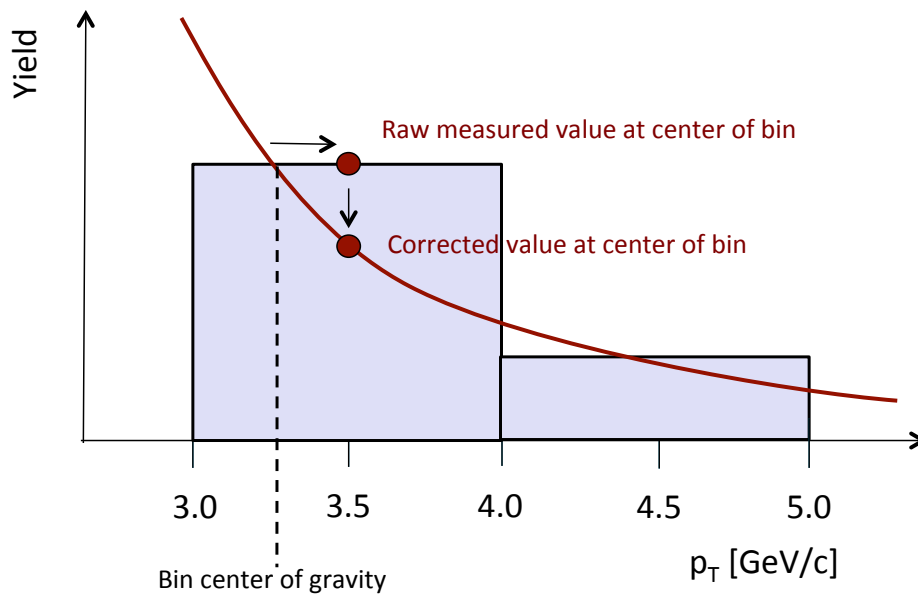
Figure 8.6: Diagram demonstrating the wide-binning effect. The red line indicates a hypothetical charged-particle spectrum, and the blue bars represent the yield in each wide p_T-bin.

As shown in Fig. 8.6, one may imagine a spectrum measured in bins of $\delta p_T = 1 \, \text{GeV}/c$. For the $3 < p_T < 4 \, \text{GeV}/c$ bin, the majority of measured particles may fall below the center value of $p_T = 3.5 \, \text{GeV}/c$. If one simply reports the height of the bin at the center value, then this p_T value is above the true value of the spectrum at $p_T = 3.5 \, \text{GeV}/c$, which would be properly determined in the bins were infinitesimally small. One could report the measured yield in each bin at the average p_T value of all particles in the bin, but this would result in a spectrum with mismatched p_T values relative to the pp reference spectrum or other spectra, making it difficult to construct the nuclear modification factor or compare to spectra in other pseudorapidity ranges.

Instead, a correction factor is determined for each p_T-bin which can reduce the measured value at the bin center to the approximate true value which would have been determined using arbitrarily
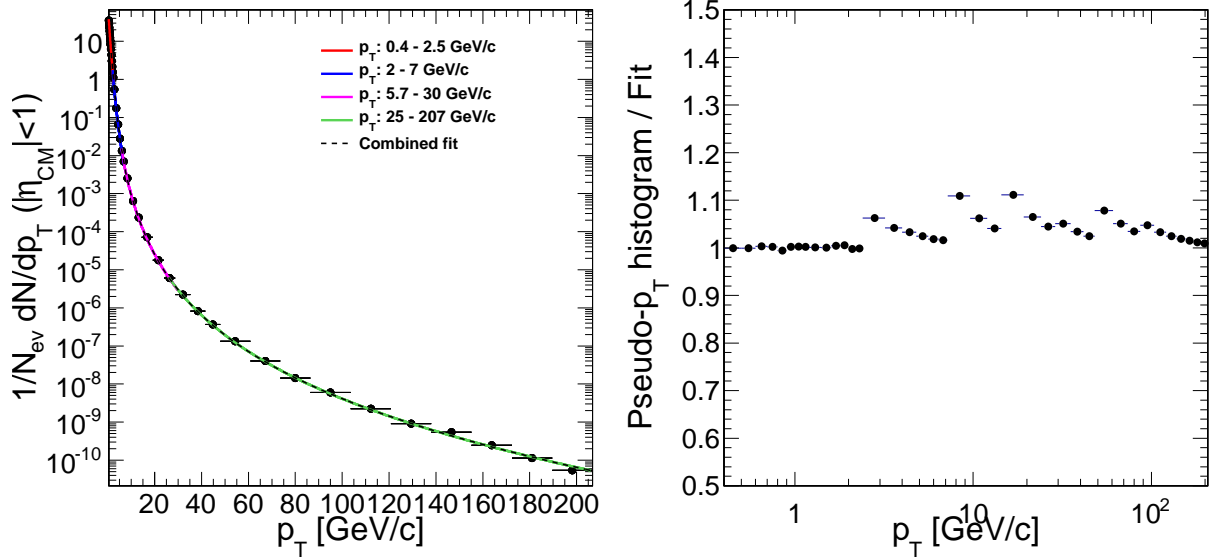
Figure 8.7: (Left) The measured charged-particle spectrum together with the individual and the combined power law fits as described in the text. (Right) The ratio of the pseudo-p_T distribution and the combined fit; this ratio is used as the correction factor for the wide-binning effect [158].

small bins. First the spectrum, having been corrected for tracking efficiency and other effects as specified in equation 8.9, is described using a number of power law functions which are fit to the p_T ranges of 0.4–2.5 GeV/c, 2–7 GeV/c, 5.7–30 GeV/c, and 25-207 GeV/c. Unlike most statistical fits, the goal of these fits is not to match the value of the function to the value of the yield at the bin center, but to match the integral of the function, divided by the bin width δp_T, to the corresponding yield of the given bin. In this manner, the fit functions approximate the spectrum that would have produced the measured yields in the wide bins. A combined fit function is determined by using the results of each fit in a different p_T range. This combined function is not perfectly continuous, but this is not a problem for the determination of the correction factor. The combined fit function is shown on the left of Fig. 8.7.

To then determine the correction factor, the combined fit function is taken as a probability distribution. Random numbers are generated uniformly and weighted using the combined fit function, and assigned to the wide p_T-bins to create a pseudo-p_T distribution. The correction factor is then extracted by dividing the pseudo-p_T distribution by the value of the combined fit function at the center of the p_T bin. This resulting correction factor is shown on the right side of Fig. 8.7, and is applied to the measured spectra.

A second effect that may distort a steeply falling spectrum is due to the finite momentum resolution of the tracking reconstruction. However, some caution is needed in determining the momentum resolution, $\sigma(\delta p_T/p_T)$, from simulation. On the left side of Fig. 8.8, the distribution of the correlation between the p_T of the simulated particle and the p_T of the reconstructed track is shown. The distribution appears to become wider exactly as the p_T-bins become wider, which is due to the wide-binning effect distorting the measurement of the momentum resolution. To avoid this distortion,
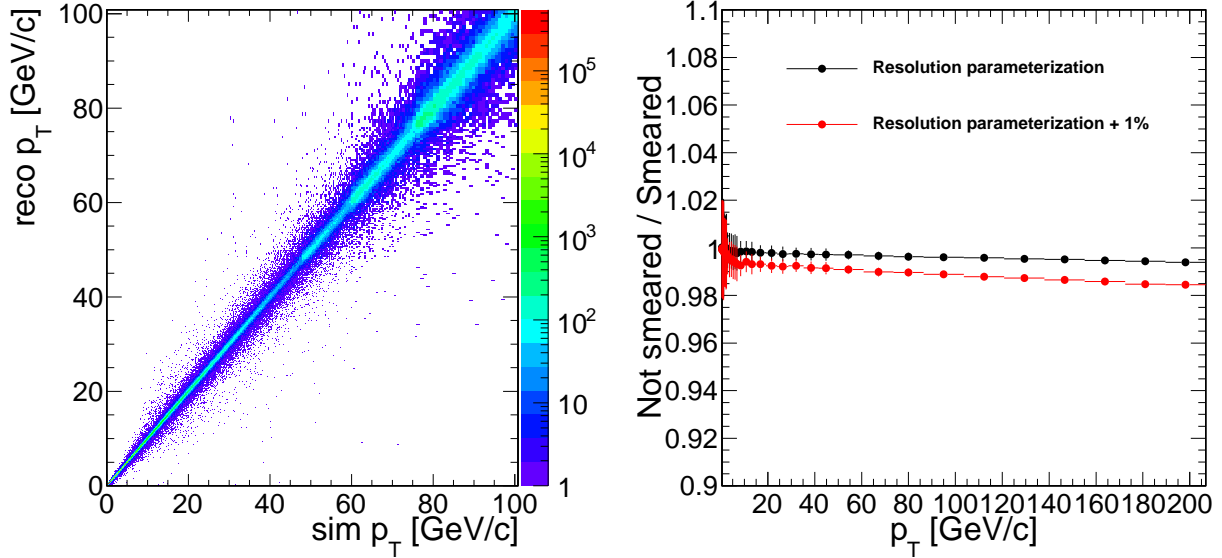
Figure 8.8: (Left) Correlation between simulated and reconstructed p_T values in HIJING simulation with embedded $\hat{p}_T$=370 PYTHIA jets. Artificial structures can be seen where the bin width, δp_T, increases from one p_T-bin to the next. (Right) The effect of the finite momentum resolution as measured using the pseudo-p_T distribution smeared using the momentum resolution as measured in simulation (Black) and as measured in simulation with a 1% increase in $\sigma(\delta p_T/p_T)$ [158].

one must use sufficiently small p_T-bins in the measurement of $\sigma(\delta p_T/p_T)$ from simulation.

With the momentum resolution accurately measured in simulation, one can then quantitatively estimate the distortion in the measured spectrum due to the smearing of the charged-particle p_T in the reconstruction process. The same combined fit distribution used to correct for the wide binning effect can again be employed. Random numbers are again uniformly generated and weighted according to the combined fit function. These pseudo-p_T values are then smeared according to their p_T value with a Gaussian distribution derived from the measurement of the momentum resolution.

The smeared distribution is then compared to the unsmeared distribution to determine the amount of distortion in the spectrum. This is shown on the right of Fig. 8.8 In the p_T range used for the measurement, $0.4 < p_T < 100\,\mathrm{GeV}/c$, the distortion from the finite momentum resolution is too small to accurately determine a correction factor. To conservatively estimate any possible systematic bias from the finite momentum resolution, this smearing procedure was repeated assuming that $\sigma(\delta p_T/p_T)$ was 1% higher than the 1–2% found in simulation. From this study, a 1% point-to-point systematic uncertainty due to finite momentum resolution is applied to the measured spectra.

8.3.2 Evaluation of Systematic Uncertainties

Seven different sources of systematic uncertainty evaluated and applied to the measurement of the charged particle transverse momentum spectra as point-to-point uncorrelated uncertainties.

112

These relate to the minimum-bias trigger efficiency, the finite momentum resolution of the track reconstruction, the uncertainty in the particle species composition, the fake rate of the track reconstruction, the efficiency of the track selection criteria, the inherent uncertainty of the trigger combination method due to statistical limitations, and potential biasing of the trigger combination method.

Table 8.1: Table of systematic uncertainties for the measurement of charged-particle spectra, R_{pPb}, and Y_{asym}.

Uncertainty	$p_{\mathrm{T}}\,[\,\mathrm{GeV}/c\,]$				
	$0.4 - 1.0$	$1.0 - 2.0$	$2.0 - 3.2$	$3.2 - 14.0$	> 14.0
Trigger Efficiency	1.0%	1.0%	1.0%	1.0%	1.0%
Momentum Resolution	1.0%	1.0%	1.0%	1.0%	1.0%
Particle Species Composition	1.0%	3.0%	6.0%	10.0%	5.0%
Tracking Fake Rate	1.0%	1.0%	1.0%	1.0%	1.0%
Track Selection	1.2%	1.6%	2.2%	4.0%	4.0%
Spectra Combination	–	–	–	–	1.0%
Trigger Bias	–	–	–	7.0%	10.0%
pp Interpolation	10.0%	10.0%	10.0%	10.0%	10.0%
Total (Spectra)	2.3%	3.8%	6.6%	13.0%	12.0%
Total (R_{pPb})	10.3%	10.7%	12.0%	16.4%	15.7%
Total (Y_{asym} $0.3 < \eta < 0.8$)	2.0%	2.0%	2.0%	2.0%	2.0-3.0%
Total ($Y_{/\mathrm{asym}}$ $0.8 < \eta < 1.3$)	2.0%	2.0%	2.0%	2.0-3.0%	3.0-5.0%
Total (Y_{asym} $1.3 < \eta < 1.8$)	2.0%	2.0%	2.0%	2.0-4.0%	4.0-10.0%

As described in Sec. 5.3, the minimum bias trigger and offline event selection is not fully efficient in accepting double-sided events, and may additionally accept spurious non-DS events. This effect is corrected for using a weighting factor based on the reconstructed track multiplicity, M, of each event. This correction factor was determined using a HIJING simulation, and alternatively using an EPOS simulation. The difference in the measured spectra corrected using the the weighting from either simulation is only 1%. Most of the difference is simply a change in the normalization of the spectra, but there is also a small, $\sim 0.1\%$, difference in the shape of the spectra.

From the study described in Sec. 8.3.1, a 1% undertainty due to the finite momentum reconstruction of the tracking algorithm is assigned to the measured spectra.

As detailed in Sec. 4.5.3, weak-decaying strange and multistrange baryons are considered primary charged particles, but are difficult to reconstruct at low-p_{T} due to their short mean lifetimes. The charged-particle efficiency is therefore quite sensitive to the fraction of charged particles comprised of these baryons. The difference in efficiency from HIJING and EPOS simulations, as shown in Fig. 4.11, is taken as a p_{T}-dependent systematic ranging in value from 1–10%.

The fake rate of the track reconstruction was evaluated with the assumption that fake tracks are

produced in proportion to properly reconstructed particle trajectories. This assumption may not hold true in general, and so the rate at which fake tracks are produced was also determined on a per-event basis as a function of p_T, classifying events by their reconstructed track multiplicity and leading jet-E_T. The spectra were then determined using this alternative per-event fake rate correction, and were found to agree with the spectra determined using the standard fake rate correction factor. Small point-to-point differences on the order of 1% were noted, and so a 1% systematic uncertainty due to the estimation of the fake rate is applied to the measured spectra.

The choice of tracking selection criteria as described in Sec. 4.5 can introduce a potential bias in the measured spectra, particularly if the reconstructed tracks in simulation have different distributions than the reconstructed tracks in the collision data in the distributions of the variables used to determine track quality. The most sensitive of these variables are the transverse and longitudinal vertex-compatibility significance, $d_{xy}/\sigma_{d_{xy}}$ and d_z/σ_{d_z}. This potential bias was evaluated by changing these selection criteria from $d_{xy/z}/\sigma_{d_{xy/z}} < 3.0$ to $d_{xy/z}/\sigma_{d_{xy/z}} < 2.0$ and $d_{xy/z}/\sigma_{d_{xy/z}} < 6.0$. The track reconstruction correction factors were then re-evaluated in simulation with respect to the altered criteria. The difference in the corrected spectra with the altered and standard criteria were compared, and based on the discrepancy between these measurements, a p_T-dependent systematic uncertainty ranging from 1.2–4.0% is applied to the measured spectra. This uncertainty affects the shape of the spectra at low-p_T, but is largely a change in normalization at high-p_T.

The trigger combination method suffers from a statistical limitation when scaling the number of events in the leading-track p_T range from 14-22 GeV/c taken from the Trk12 trigger to those taken with the MB trigger. This is due to the limited statistical reach of the charged-particle spectrum from the MB spectrum due to the high prescale factor on the single-track trigger. This introduces an effective 1% normalization uncertainty on the portion of the spectrum above $p_T = 14$ GeV/c relative to the portion of the spectrum below $p_T = 14$ GeV/c.

Although the potential bias of the trigger combination method was explored in simulation as presented in Sec. 8.1.4 and found to be approximately less than 3%, an additional check was performed by using an alternate combination of triggers based on reconstructed HLT jets rather than reconstructed HLT tracks. This alternate method was the one used in Ref. [27, 125] and the discrepancy between the two methods was explored both by applying both trigger combinations to measure the 2013 pPb charged-particle spectra, as well as re-measuring the pp charged-particle spectra at $\sqrt{s} = 2.76$ TeV using collision data taken in 2013, and comparing to the measured spectra from the 2011 pp collision data presented in Ref. [27]. Based on these studies, a p_T-dependent systematic uncertainty ranging from 7–10% is applied to the measured pPb spectra for $p_T > 3.2$ GeV/c.

These systematic uncertainties are summarized in Table 8.1 and added in quadrature to determine the final systematic uncertainty assigned to the measured charged-particle pPb spectra as a function of p_T.

8.3.3 Results

The corrected $\sqrt{s_{NN}} = 5.02$ TeV pPb charged-particle yields are presented as a function of p_T for $0.4 < p_T < 100$ GeV/c in Fig. 8.9 for seven center-of-mass pseudorapidity ranges: $-1.8 < \eta_{CM} < -1.3$, $-1.3 < \eta_{CM} < -0.8$, $-0.8 < \eta_{CM} < -0.3$, $-1.0 < \eta_{CM} < 1.0$, $0.3 < \eta_{CM} < 0.8$,

$0.8 < \eta_{\mathrm{CM}} < 1.3$, and $1.3 < \eta_{\mathrm{CM}} < 1.8$. For comparison, the corresponding charged-particle yield from the interpolated pp reference at $\sqrt{s} = 5.02\,\mathrm{TeV}$ is scaled by a factor of $N_{\mathrm{coll}} = 6.9$ and shown alongside the pPb yield.
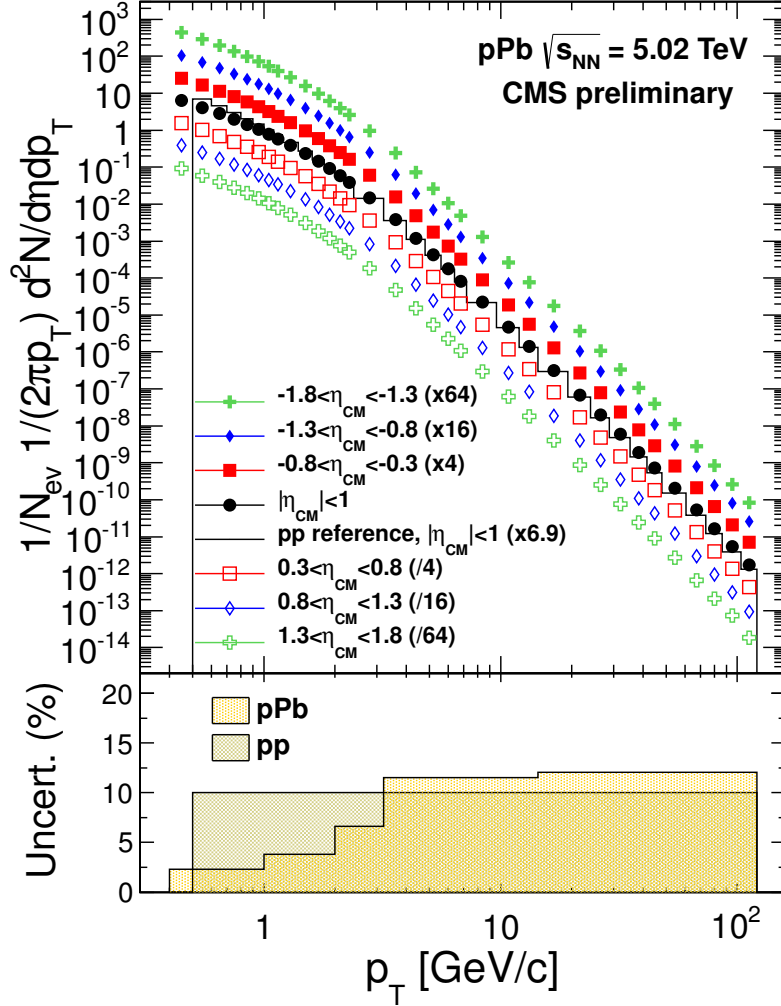
Figure 8.9: Upper panel: Charged-particle yields measured in $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV pPb collisions in the $|\eta_{\mathrm{CM}}| < 1$, $0.3 < \eta_{\mathrm{CM}} < 0.8$, $0.8 < \eta_{\mathrm{CM}} < 1.3$, $1.3 < \eta_{\mathrm{CM}} < 1.8$, $-0.8 < \eta_{\mathrm{CM}} < -0.3$, $-1.3 < \eta_{\mathrm{CM}} < -0.8$, and $-1.8 < \eta_{\mathrm{CM}} < -1.3$, pseudorapidity bins. Positive pseudorapidity values correspond to the p-going side. The yields from the forward and backward η ranges have been scaled by multiples of 4 for better visibility. The charged-particle yield from the interpolated pp reference is also shown, and is scaled by a factor of 6.9, which corresponds to the average N_{coll} estimate from the Glauber model. Lower panel: systematic uncertainties of the measured pPb and interpolated pp spectra as a function of p_{T} [157].

8.4   Forward-Backward Asymmetry

The forward-backward asymmetry is defined as the charged-particle yield in some negative pseudorapidity range $[-b, -a]$ divided by the charged particle yield in some corresponding positive pseudorapidity range $[a, b]$:

$$Y_{\mathrm{asym}}(p_\mathrm{T}) = \frac{d^2 N_{ch}(p_\mathrm{T})/d\eta dp_\mathrm{T}|_{\eta_{\mathrm{CM}} \in [-b, -a]}}{d^2 N_{ch}(p_\mathrm{T})/d\eta dp_\mathrm{T}|_{\eta_{\mathrm{CM}} \in [a, b]}}. \tag{8.11}$$

Note that with the convention that the incoming proton moves in the positive z-direction, the $Y_{\mathrm{asym}}(p_\mathrm{T})$ presented here measures the yield in the pseudorapidity region corresponding to what may be described as the "lead-going side" divided by that of the "proton-going side".

In general, the systematic uncertainties which are applied to the spectra measurement largely cancel when the spectra from one pseudorapidity range are divided by those of another pseudorapidity range. Any bias from the determination of the minimum bias trigger efficiency, trigger combination method, or tracking selection criteria will affect the yield in each pseudorapidity range in the same manner. In order to provide a conservative estimate of any remaining uncertainties, the fact that the pPb collision data was taken in two configurations, where the proton and lead beam directions were interchanged, is exploited.

In the first period of the 2013 pPb collision run at the LHC, the proton moved towards the negative z-axis. In the second period, the proton moved towards the positive z-axis. Therefore, it is possible to look at the same pseudorapidity region in the center-of-mass frame in different portions of the detector. In the absence of any systematic biases, these measurements should result in the same yield since they probe the same pseudorapidity region in the center-of-mass frame. From the small differences determined in these measurements, a systematic uncertainty ranging from 4.0–10% is assigned to the Y_{asym} ratio as a function of both η and p_T. This systematic uncertainty is shown in Table 8.1.

The forward-backward asymmetry is presented for three pseudorapidity ranges as a function of p_T in Fig. 8.10. In each of the three $|\eta|$ ranges, the value of Y_{asym} rises up to a p_T of 3 GeV/c, and then falls to unity at a p_T of 10 GeV/c and remains constant to higher p_T, although in the most forward range of $1.3 < |\eta| < 1.8$ the value of Y_{asym} decreases slightly below unity for $p_\mathrm{T} > 90$ GeV/c. The rise up to $p_\mathrm{T} = 3$ GeV/c is increasingly pronounced in the $0.8 < |\eta| < 1.3$ and $1.3 < |\eta| < 1.8$ ranges. At the lowest measured p_T value of 0.4 GeV/c, Y_{asym} is consistent with unity for the $0.3 < |\eta| < 0.8$ range, but is higher in the more forward regions.

In the most forward region of $1.3 < |\eta| < 1.8$, Y_{asym} is additionally measured for events falling into a specific activity class defined in terms of the total transverse energy measured by the HF in both positive and negative pseudorapidity constrained to $|\eta| > 4$. These measurements are presented for events with $E_\mathrm{T} < 20$ GeV, and events with $E_\mathrm{T} > 40$ GeV in Fig. 8.11. For the $E_\mathrm{T} > 40$ GeV event class, Y_{asym} is higher at low-p_T compared with the inclusive measurement, and exhibits a pronounced rise above unity for $p_\mathrm{T} > 10$ GeV/c. For the $E_\mathrm{T} < 20$ GeV event class, Y_{asym} is lower at low-p_T compared with the inclusive measurement, and exhibits a slight decrease below unity for $p_\mathrm{T} > 10$ GeV/c.
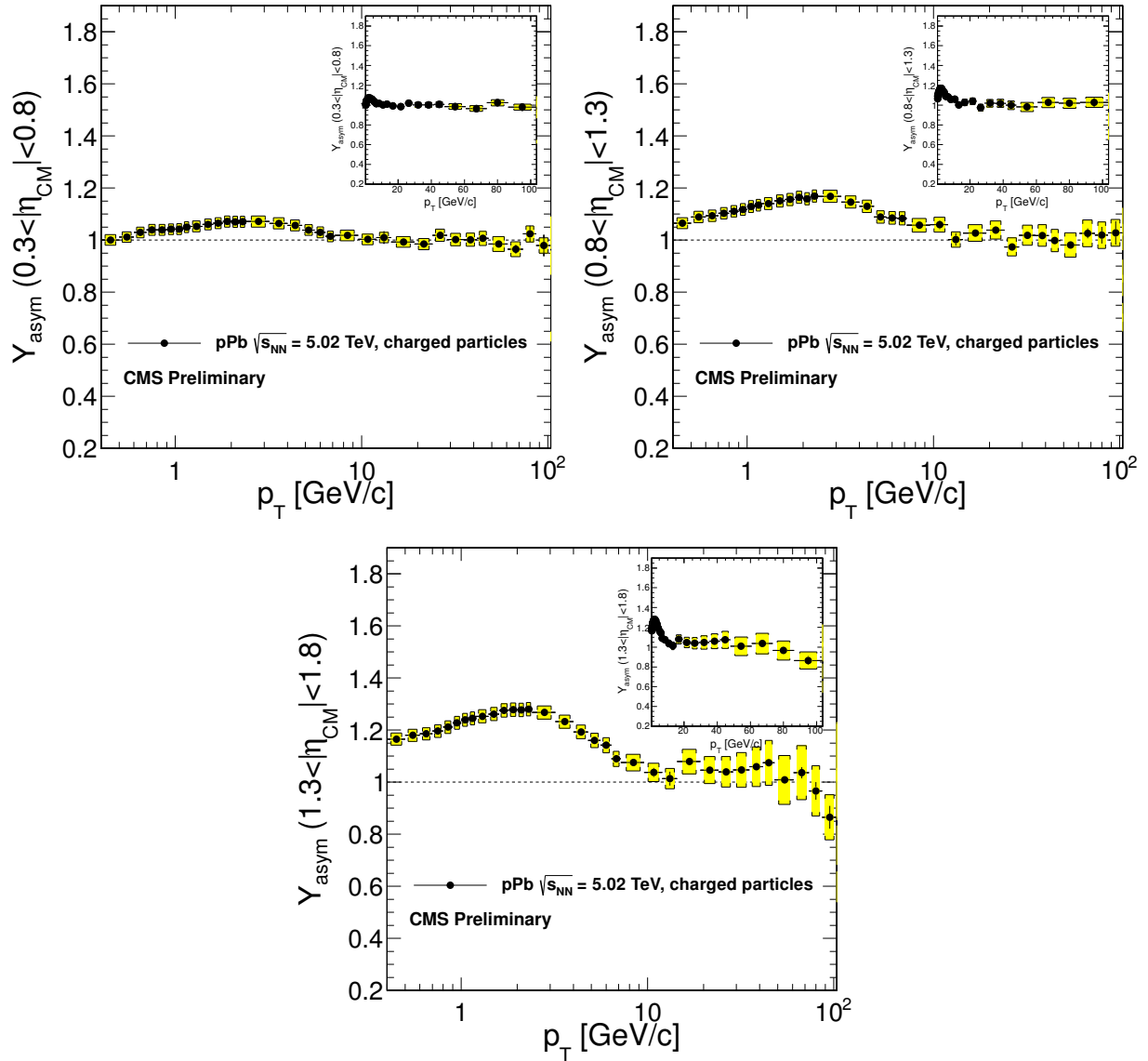
Figure 8.10: Charged-particle asymmetry as a function of p_{T} for $0.3 < |\eta_{\mathrm{CM}}| < 0.8$, $0.8 < |\eta_{\mathrm{CM}}| < 1.3$, and $1.3 < |\eta_{\mathrm{CM}}| < 1.8$. The asymmetry is computed as the charged particle yields on the Pb-going side divided by the those on the p-going side [157]

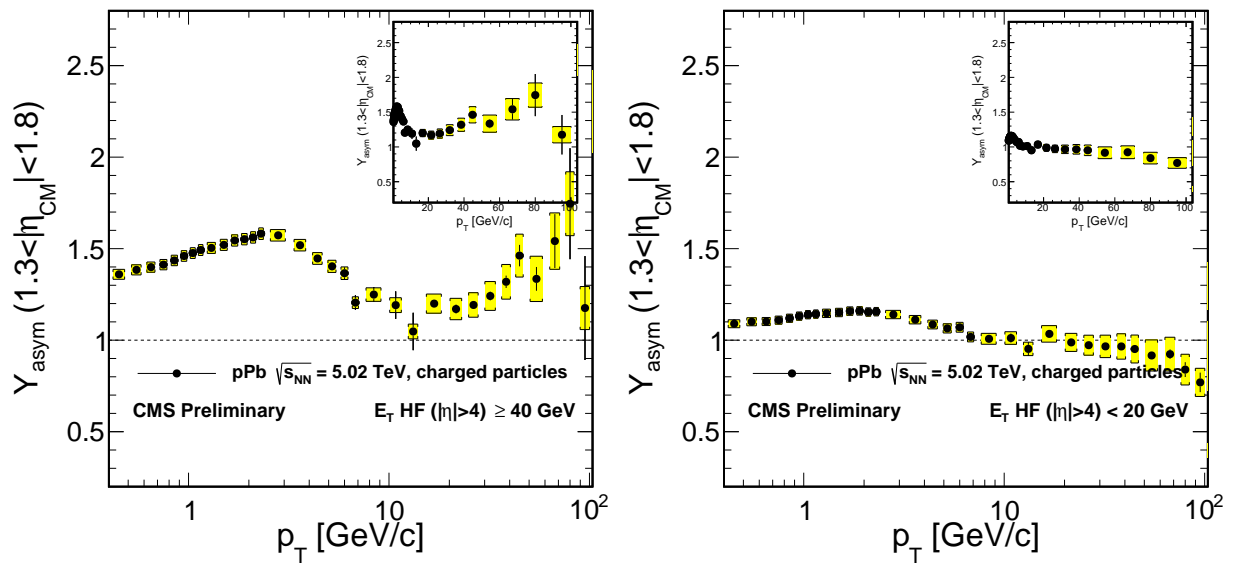**Figure 8.11:** Charged-particle asymmetry as a function of p_T for $1.3 < |\eta_\mathrm{CM}| < 1.8$. In event activity classes determined by the total transverse energy measured in the portion of the HF with $|\eta| > 4$ for events with $E_\mathrm{T} > 40\,\mathrm{GeV}$ (Left) and $E_\mathrm{T} < 20\,\mathrm{GeV}$ (Right) [159]

8.5   Nuclear Modification Factor
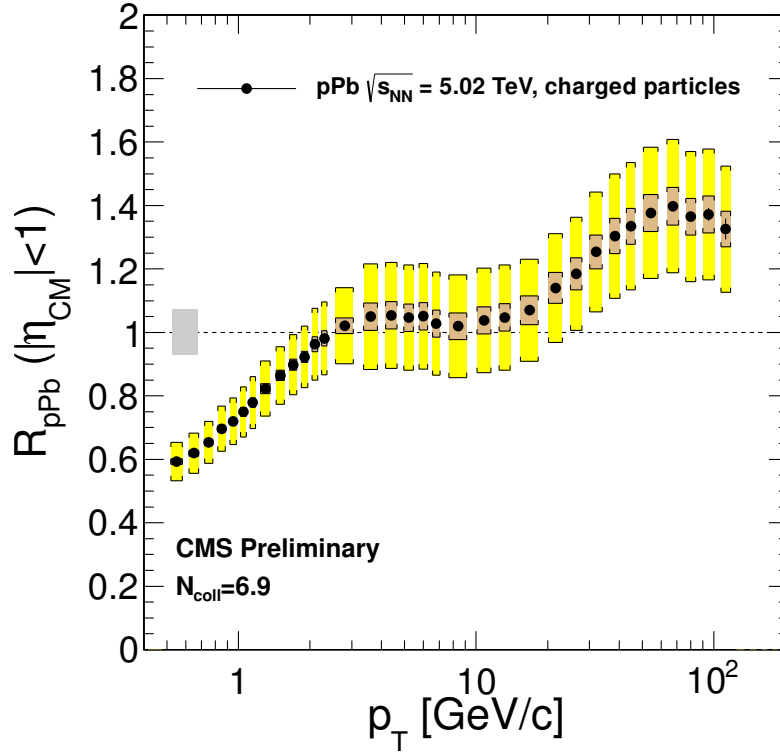
Figure 8.12: Measured nuclear modification factor for charged particles in $|\eta_{\mathrm{CM}}| < 1$. The light gray uncertainty band represents the uncertainty of the Glauber calculation of N_{coll}. The light brown uncertainty band around the measured values shows the uncertainty coming from the following sources that are strongly correlated in specific p_{T} regions: combination of spectra, track selection, and trigger efficiency. All other uncertainties are shown by the yellow band [157].

The charged-particle R_{pPb} for $|\eta| < 1.0$ is shown in Fig. 8.12. The systematic uncertainty from the Glauber calculation of N_{coll}, as given in Sec. 5.5, is shown as a light gray uncertainty band. The point-to-point systematic uncertainties include all seven estimated uncertainties for the pPb spectra measurement as well as the 10% systematic uncertainty on the interpolation of the pp reference spectrum. These uncertainties are divided into two groups, the strongly correlated uncertainties in specific p_{T} regions arising from the combination of spectra, track selection, and trigger efficiency are given by the brown band. All other uncertainties are represented by the yellow band. The uncertainties are summarized in Table 8.1. The measured R_{pPb} rises to unity at $p_{\mathrm{T}} \approx 4\,\mathrm{GeV}/c$, is flat to $p_{\mathrm{T}} \approx 20\,\mathrm{GeV}/c$, and then increases at higher p_{T} to reach a maximum value between 1.3 and 1.4 around $p_{\mathrm{T}} \approx 70\,\mathrm{GeV}/c$.

8.6 Discussion

In the region where they overlap kinematically, the charged-particle spectra measured with CMS are in excellent agreement with those previously measured by ALICE [83]. The CMS measurement extends the kinematic reach of the previous measurement from $|\eta| < 1.3$ to $|\eta| < 1.8$, and from $p_T < 20$ to $p_T < 100$. In Fig. 8.13, the R_{pPb} measured with CMS for $|\eta_{CM}| < 1$ is compared with the ALICE measurement for $|\eta_{CM}| < 0.3$. Despite the difference in pseudorapidity range, these results are also in good agreement.
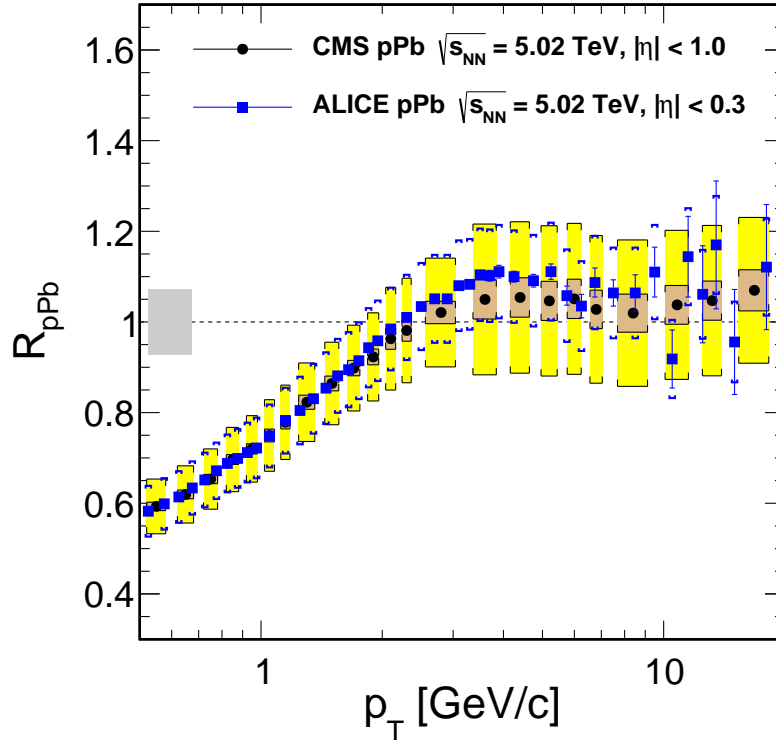
Figure 8.13: CMS measurement of R_{pPb} for charged particles in $|\eta_{CM}| < 1$ (Black) and ALICE measurement of R_{pPb} for charged particles in $|\eta_{CM}| < 0.3$ (Blue) [83]. The statistical and systematic uncertainties for the CMS measurement are given as in Fig. 8.12. The statistical uncertainties of the ALICE measurement are given by the vertical bars, and the systematic uncertainties by the blue brackets.

Previous to the 2013 pPb collision data taking at the LHC, several predictions were made for Y_{asym} at $0.3 < |\eta| < 0.8$ and $p_T < 15\,\text{GeV}/c$ using different MC generator codes and nPDF parametrizations [160]. The first of these, the κTPQCD_v2.0 code, is based on a perturbative QCD model that implements transverse momentum broadening due to nuclear multiple scattering and may generate observed Cronin enhancement [167, 168]. This was initialized and run separately with the following parametrizations: EKS98 [161], EPS08 [162], HKN [163], and HIJING2.0 [164]. As the HIJING2.0 parametrization includes a dependence on the position of the nucleon in the nucleus, it is possible to perform a prediction based on pPb collision events in various centrality ranges. With this parametrization, simulations both for inclusive and 0-20% central events are performed.

Figure 8.14: (Left) The CMS measurement of charged-particle Y_{asym} in the range $0.3 < |\eta| < 0.8$ for $p_{\mathrm{T}} < 15\,\mathrm{GeV}/c$. (Right) Theoretical predictions for Y_{asym} in the same kinematic range [160] with the KTPQCD_V2.0 model using the following parametrizations for inclusive, "minimum-bias" (MB), events: EKS98 [161] / EPS08 [162] (Blue Solid Line), HKN [163] (Green Solid Line), and HIJING2.0 [164] (Red Line). In the case of the HIJING2.0 parametrization, the prediction is made for inclusive events (Solid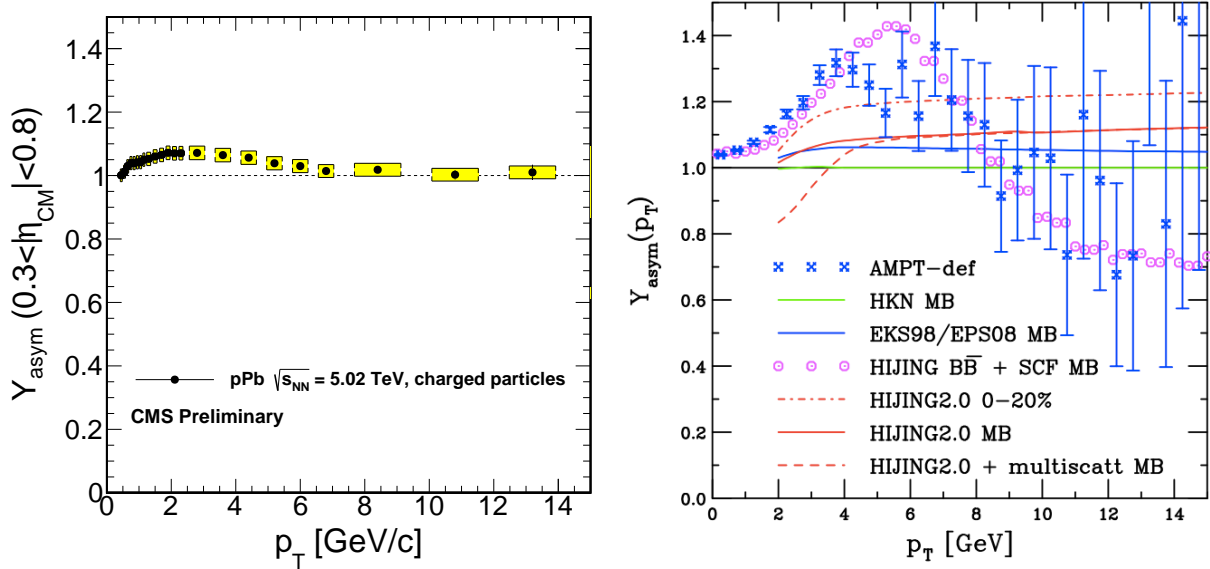 Red Line), for inclusive events with multiple scattering effects (Dashed Red Line), and for 0-20% central events (Dot-Dashed Red Line). Predictions from the HIJING2.0BB̄ [165] (Violet Circles) and AMPT [166] models are also shown.

The prediction was also made using the HIJING2.0BB̄ model, which incorporates strong color field effects [165]. A recent version of the AMPT generator was also used to perform the prediction [166]. These predictions are compared side-by-side to the CMS measurement of Y_{asym} in the same kinematic region in Fig. 8.14. None of the predictions clearly describe both the low-p_{T} rise reaching a maximum around $p_{\mathrm{T}} \approx 2 - 3\,\mathrm{GeV}/c$ and the return to unity above $p_{\mathrm{T}} \approx 10\,\mathrm{GeV}/c$. The AMPT and HIJING2.0BB̄ do qualitatively describe the low-p_{T} rise and fall to a degree, but the magnitude and position of the peak is markedly different in the data. These comparisons illustrate the sensitivity of the Y_{asym} observable to the modeling of the initial conditions of the pPb collision system, and the theoretical challenge of describing the features of the p_{T}-dependence of Y_{asym}. The three different pseudorapidity ranges explored in this measurement constitute a precise set of data from which the initial conditions of the pPb collision system may be better understood.

In the most forward Y_{asym} measurement at $1.3 < |\eta| < 1.8$, the structure at low-p_{T} suggests an overall increase in $dN_{\mathrm{ch}}/d\eta$ in the negative η, or lead-going, side relative to the positive η, or proton-going, side. This increase of approximately 20% is similar to the difference in $dN_{\mathrm{ch}}/d\eta$ in similar positive and negative η regions as observed in dAu collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ [169, 170]. This low-p_{T} increase in the charged-particle yield at negative pseudorapidity is consistent with the suppression of the nPDF in the low-x shadowing region. This supression may also be inferred from the low-p_{T} behavior of the charged-particle R_{pPb}, which is below unity for $p_{\mathrm{T}} < 3\,\mathrm{GeV}/c$.
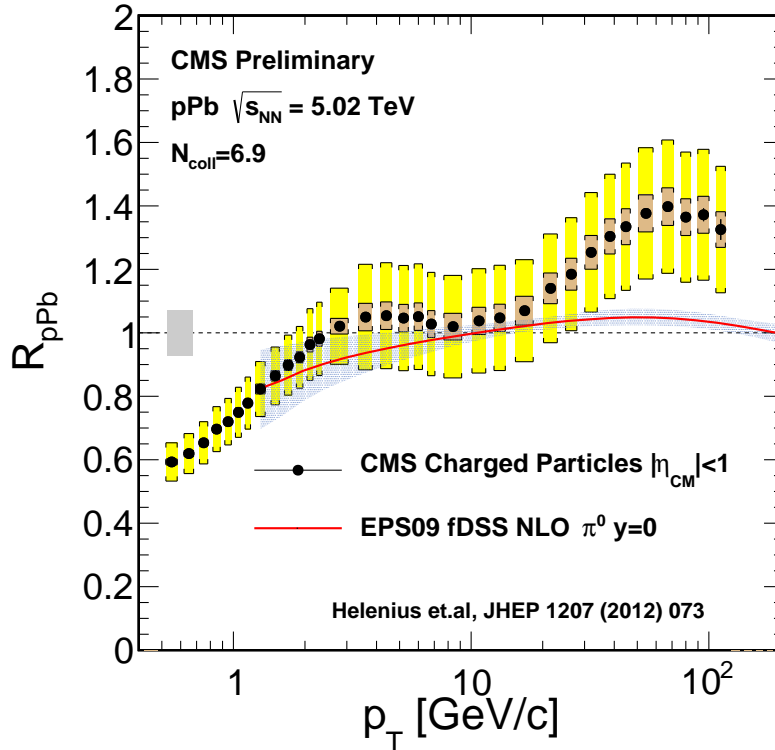
122

Figure 8.15: The CMS measurement of charged-particle R_{pPb} for $|\eta| < 1.0$ compared with a theoretical prediction of π^0 R_{pPb} at $y = 0$ [67].

In Fig. 8.15, the results of the charged-particle R_{pPb} measurement are compared with a theoretical calculation of the R_{pPb} of neutral pions using the EPS09 [67] nPDF parametrization and the fDSS [26] fragmentation functions. The increase of the measured charged-particle R_{pPb} as compared to the π^0 R_{pPb} prediction in the range $2 \lesssim p_{\mathrm{T}} \lesssim 8$ is reminiscent of the difference in charged particle and π^0 R_{dA} seen at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ as shown in Fig. 1.10. The theoretical and measured R_{pPb} agree at low p_{T} and in the range $10 \lesssim p_{\mathrm{T}} \lesssim 20$, but for $p_{\mathrm{T}} > 20$ the enhancement in R_{pPb} is much larger than predicted. As can be seen in Fig. 1.11, some high-p_{T} enhancement in R_{pPb} can be attributed to enhancement in the nPDF in the gluon anti-shadowing region.

It is then natural to consider if the unexpectedly large enhancement in R_{pPb} can be explained in terms of a larger than expected increase in the nPDF in the anti-shadowing region. To explore this potential interpretation, one may begin by determining if the observation of Y_{asym} consistent with unity is possible under the assumption of a large increase in the nPDF in the anti-shadowing region.

For processes resulting in a high-p_{T} charged particle of a given p_{T} value, one may expect that the pseudorapidity of the observed particle is correlated with the Bjorken x of the initial parton in the Pb nucleus which participates in the hard-scattering interaction, which is denoted x_2. As the convention of this measurement is to take the direction of the lead nucleus as towards the negative z-axis, one may expect that for larger x_2 values, an observed hadron at a given p_{T} is likely to have a negative η value, and for smaller x_2 values, the observed hadron at a given p_{T} is likely to have
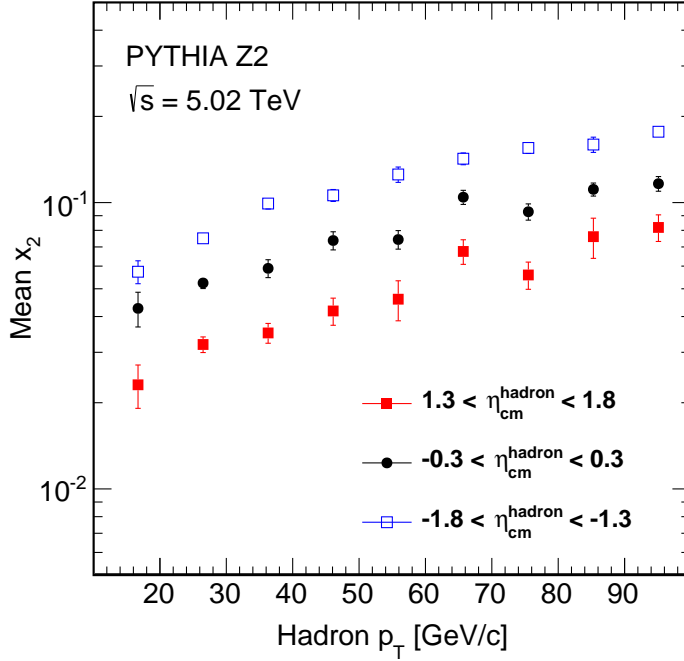
Figure 8.16: Mean Bjorken x of the initial state parton participating in a hard-scattering interaction, denoted x_2, from the nucleon moving in the negative z direction associated with each final-state charged hadron produced in the collision as a function of the p_T of that final-state hadron, as simulated in pp collisions with PYTHIA Z2, for hadrons in the three η regions as given in the legend.

a positive η value. Note that for the initial parton from the proton, denoted x_1, one would expect the opposite relationship to the η of the observed hadron.

This relationship between the x_2 of the parton in the Pb nucleus and the p_T of the observed charged particle should indicate that the p_T value for which the onset of the R_{pPb} enhancement occurs will vary with p_T. If this variance is significant in the η range considered in the measurement, $|\eta| < 1.8$, then Y_{asym} should be seen to deviate from unity above $p_T = 20\,\text{GeV}/c$. In Fig. 8.16, the results of a PYTHIA Z2 simulation are presented, in which the mean value of the x_2 of the initial state parton corresponding to an observed high-p_T hadron in some η interval is shown as a function of the p_T of the observed hadron. In all pseudorapidity regions studied, charged hadrons with $p_T > 20\,\text{GeV}/c$ were found to correspond on average with initial-state partons in the Pb nucleus with $0.02 < x < 0.2$. Referring back to Fig. 1.11, this x region roughly corresponds to the anti-shadowing regime, and so it is not implausible that an approximate plateau of the nPDF to free proton PDF ratio could exist from $0.02 < x < 0.2$, which in turn may explain enhancement in R_{pPb} while not resulting in any deviation from unity for Y_{asym} with $|\eta| < 1.8$. Even with such a plateau, eventually at high enough p_T and high enough x, the regime of the EMC effect should be encountered and the Y_{asym} should therefore fall below unity. Although the measurement of Y_{asym} does fall below unity at very high-p_T for the most forward η range considered, this drop cannot be considered significant given

the uncertainty in the measurement.

Even if larger than expected anti-shadowing can be shown to be an internally consistent explanation given the measurements of R_{pPb} and Y_{asym}, it may still be difficult or even impossible to incorporate this data into a good quality global fit of the nPDF as the evolution of the PDF is constrained by the DGLAP equation. It is therefore reasonable to look for other explanations of the unexpected rise in R_{pPb}, which may very well signal a novel effect beyond enhancement of the nPDF in the anti-shadowing regime. As shown in Fig. 8.11, Y_{asym} varies greatly with the total energy of the particles produced in the event in the forward pseudorapidity region of $4 < |\eta| < 5$. This indicates that the rise in R_{pPb} may perhaps be better understood by further measurements performed differentially in different centrality classes defined by forward energy.

Chapter 9

Conclusion

Over the last three years, the LHC has delivered PbPb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ and pPb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ with enough integrated luminosity to perform measurements with an extremely high p_{T} reach. This accomplishment has ushered in a new and exciting era in the study of heavy ion collisions, in which the QGP medium can be studied at higher temperature than ever produced before, and with new experimental apparatuses capable of providing greater precision and new observables.

The measurement of PbPb spectra at low-p_{T} and in several pseudorapidity regions provides an important constraint to the modeling and description of the evolution of the expanding medium. In addition, these low-p_{T} spectra are essential for determining yield-weighted average v_n coefficients, which may be then be compared with theory and other collision systems to better understand the hydrodynamic properties of the evolving medium and the initial state of the PbPb collision system. When compared with results from lower energy collision systems, the $\langle p_{\text{T}} \rangle$ suggests that the medium produced in these PbPb collisions has a higher initial energy density or stronger radial flow than the medium produced in lower energy collisions.

The measurement of R_{pPb} may in principle serve as a control for the R_{AA} of charged particles in PbPb collisions, providing a stronger argument for the interpretation of partonic energy loss in the medium produced in the collision. With sufficient precision, this measurement may also help to constrain the nuclear parton distribution function and gain insight into the initial state of the Pb nucleus. This measurement of Y_{asym} may also help to understand the initial state, and the low-p_{T} features of the observation may provide insight into saturation effects and the CGC model.

Below $p_{\text{T}} = 20\,\text{GeV}/c$, these pPb spectra and R_{pPb} measurements confirm the previous measurements in a similar kinematic range. Above $p_{\text{T}} = 20\,\text{GeV}/c$, this measurement has been described as "mind-boggling" [171]. Conventional explanations of the R_{pPb} enhancement at high-p_{T} due to anti-shadowing in the nPDF may be insufficient to understand the magnitude of the rise. The presence of this enhancement may also call into question how the corresponding high-p_{T} rise in the R_{AA} of charged particles in PbPb is described.

Further measurements and detailed comparisons may be required to elucidate the mechanism behind the enhancement in R_{pPb} at high-p_{T}. The measurements of Y_{asym} in different event activity classes determined by the total energy measured in the $4 < |\eta| < 5$ region hint that the R_{pPb} of charged particles may also exhibit and interesting dependence on event activity or centrality.

The QCD Lagrangian

The QCD Lagrangian and its mathematical properties can be understood by first inspecting the Quantum Electrodynamic (QED) Lagrangian and recognizing the QCD Lagrangian as a natural non-abelian generalization of this structure. The QED Lagrangian may be written as

$$\mathcal{L}_{\text{QED}} = \bar{\psi}\left(i\gamma^\mu \partial_\mu - e_c \gamma^\mu A_\mu - m_e\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{A.1}$$

where repeated Lorentz indices are summed, ψ is a four-component Dirac bi-spinor representing an electron, the γ^μ are the Dirac matrices, $\bar{\psi}$ is the Dirac adjoint defined $\psi^\dagger \gamma^0$, m_e is the mass of the electron, e_c is the fundamental charge, A_μ is the electromagnetic covariant four-potential or photon field, and $F_{\mu\nu}$ is the electromagnetic field tensor, defined

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\mu A_\nu. \tag{A.2}$$

It can easily be shown that $\mathcal{L}_{\text{QED}}$ is unchanged if the electron and photon fields are simultaneously transformed as

$$\psi(x) \to e^{ie_c\theta(x)}\psi(x), \qquad \text{and} \qquad A_\mu \to A_\mu - \partial_\mu\theta(x), \tag{A.3}$$

where $\theta(x)$ is any differentiable function of x, the position four-vector. This invariance is termed a local gauge symmetry, and by Noether's theorem must correspond to a conserved quantity which may be shown to be the electric charge.

The factor $e^{ie_c\theta}$, which is just a complex number with a magnitude of 1, may be recognized as a trivial 1x1 complex valued matrix with a determinant of 1. In other words, this is a unitary matrix which preserves the magnitude of ψ, viewed as a single component vector. The collection of all such 1x1 transformation matrices along with the operation of multiplication forms the circle group $U(1)$ which is obviously abelian as $e^{ie_c\theta_1} \cdot e^{ie_c\theta_2} = e^{ie_c\theta_2} \cdot e^{ie_c\theta_1}$ for any real valued θ_1 and θ_2.

With this in mind, one may then envision a set of three fermion fields ψ_r, ψ_b, and ψ_g and a set of transformations which mix the fields together while preserving their normalization and orthogonality

$$\begin{aligned}
\psi_r \to \psi_r' &= \alpha_1\psi_r + \alpha_2\psi_g + \alpha_3\psi_b, \\
\psi_g \to \psi_r' &= \beta_1\psi_r + \beta_2\psi_g + \beta_3\psi_b, \\
\psi_b \to \psi_r' &= \gamma_1\psi_r + \gamma_2\psi_g + \gamma_3\psi_b.
\end{aligned} \tag{A.4}$$

The fields can be arranged into a 3-component vector, in which case the transformation becomes

$$
\boldsymbol{\Psi} = \begin{pmatrix} \psi_r' \\ \psi_g' \\ \psi_b' \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \tag{A.5}
$$

This transformation matrix is a 3x3 unitary matrix, and along with the operation of multiplication forms the group $U(3)$. Any 3x3 matrix which is a member of this group may be written as [1]

$$
U = \exp\left(i\alpha_0 I_3 + i \sum_{k=1}^{8} \alpha_k \lambda_k / 2 \right) = e^{i\alpha_0} \exp\left(i \sum_{k=1}^{8} \alpha_k \lambda_k / 2 \right) = e^{i\alpha_0} V, \tag{A.6}
$$

where the α_k are real numbers and the λ_k are the Gell-Mann matrices,

$$
\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}
$$
$$
\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \tag{A.7}
$$
$$
\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
$$

Throwing away the factor $e^{i\alpha_0}$, which only adds an overall phase to the three fermion fields, restricts one to the unitary matrices with a determinant of 1, V, which form the special unitary group $SU(3)$. Defining $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_8)$ and $\mathbf{t} = (\lambda_1/2, \lambda_2/2, \ldots, \lambda_8/2)$, one may write

$$
V = \exp\left(i \sum_{k=1}^{8} \alpha_k \lambda_k / 2 \right) = \exp\left(i\boldsymbol{\alpha} \cdot \mathbf{t} \right). \tag{A.8}
$$

Just as the QED Lagrangian is invariant under $U(1)$ transformations of the field ψ, one may now attempt to construct a Lagrangian for the fields $\boldsymbol{\Psi}$ which is invariant under local $SU(3)$ transformations, i.e. $\boldsymbol{\Psi}' = \exp(i\boldsymbol{\alpha}(x) \cdot \mathbf{t})\boldsymbol{\Psi}$. In QED, this is accomplished by introducing the photon field replacing the derivative in the free field Lagrangian with the covariant derivative $D_\mu = \partial_\mu + ie_c A_\mu$, thus obtaining the minimally coupled gauge-covariant Lagrangian. By analogy, one may introduce eight gluon fields, $\mathbf{G}_\mu$, and replace the free field derivative with

---

[1] The exponential function for a square matrix A is defined $\exp(A) = I + \sum_{n=1}^{\infty} A^n / n!$

$$D_\mu = I_3 \partial_\mu + i g_s \mathbf{G}_\mu \cdot \mathbf{t}, \qquad \text{where} \qquad \mathbf{G}_\mu = (G_\mu^1, G_\mu^2, \dots, G_\mu^8), \tag{A.9}$$

and one requires that each of the gluon fields transform as

$$G_\mu^{k\prime} = G_\mu^k + \partial_\mu \alpha_k(x) + \sum_{a,b,c=1}^{8} f_{abc} \alpha_b(x) G_\mu^c, \tag{A.10}$$

where the f_{abc} are the structure constants of $SU(3)$, defined by the commutator of the Gell-Mann matrices as

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2} \right] = \frac{\lambda_a \lambda_b}{4} - \frac{\lambda_b \lambda_a}{4} = i \sum_{c=1}^{8} f_{abc} \frac{\lambda_c}{2} \tag{A.11}$$

If the multiplication of the $SU(3)$ transformation matrices was commutative, then the structure constants would all vanish. As they do not, in order to cancel all terms arising from this transformation in the covariant derivative D_μ, the gluon field tensors must be defined as

$$F_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s \sum_{b,c=1}^{8} f_{abc} G_\mu^b G_\nu^c \tag{A.12}$$

It may then be shown that the following QCD Lagrangian is invariant under local $SU(3)$ transformations:

$$\mathcal{L} = \sum_q \overline{\mathbf{\Psi}}_q \left(i I_3 \gamma^\mu \partial_\mu - g_s \gamma^\mu \mathbf{G}_\mu \cdot \mathbf{t} - m_q I_3 \right) \mathbf{\Psi}_q - \frac{1}{4} \sum_{a=1}^{8} F_{\mu\nu}^a F^{a\mu\nu} \tag{A.13}$$

where the initial sum is over all quark flavors q, and m_q is the mass of each quark flavor. One may apply Noether's theorem to obtain three conserved charges corresponding to the $SU(3)$ symmetry, which are referred to as "red", "blue", and "green". As with QED, the minimally gauge covariant derivative introduces an interaction term between the quarks and the gluons. Unlike QED, the introduction of the nonvanishing structure functions into the term $F_{\mu\nu}^a F^{a\mu\nu}$ introduces self-interaction terms for the gluon fields, which in the perturbative formulation of QCD correspond to interaction vertices with three or four gluon lines. Unlike the photon in QED, the QCD gluon fields themselves carry a double color charge, (i.e red-antigreen, green-antiblue, etc.).

Appendix B

Deep Inelastic Scattering

Deep inelastic scattering (DIS) refers to a process in which a lepton scatters off of a nucleon in an inelastic collision in which the nucleon does not remain intact. In the example of an electron-proton collision, this may take the form of a neutral current interaction mediated by a photon or Z boson given as

$$e^- + p \to e^- + X, \tag{B.1}$$

where X is some ensemble of hadrons. These interactions also include charged current interaction, mediated by a $W^\pm$ boson, as for example a muon neutrino-proton collision

$$\nu_\mu + p \to \mu^- + X. \tag{B.2}$$

Loosely following the presentation in Ref. [172], a simplified picture of DIS in the context of an electron-proton collision in a fixed target experiment is presented here, ignoring contributions from heavy weak bosons. The purpose of this chapter is to simply provide a brief overview of DIS, to explain the meaning of the variables x and Q^2, to describe how a few features of the QCD model may be inferred from DIS data, and to explain the relationship between the parton distribution functions and the DIS structure functions. For a more comprehensive review of DIS including recent results, see for example Chapter 18 of Ref. [1].

Consider the collision of an electron with beam energy E with a fixed-target proton with the mass denoted M_p. The outgoing electron has energy E' and leaves with an angle θ relative to the beamline. The outgoing hadrons have a combined energy of $\nu = E - E'$. These quantities may be measured in the laboratory.

A description of this events may also be given in terms of the variables s^2, x, and y, where s^2 is the square of the overall center-of-mass energy $s^2 = 2M_p E$; x and y are the dimensionless scaling variables, which are defined in terms of a Lorentz frame in which the proton four-momentum P has a very large three-momentum $|\vec{P}|$, i.e. $|\vec{P}| \gg M_p$, as shown below.

The kinematics of an e^-p collision in such a reference frame are shown in Fig. B.1. The incoming electron has four-momentum k, and transfers four-momentum q to the struck parton. The outgoing electron has four-momentum k'. The following Lorentz invariant quantities can now be described:

- $Q^2 = -q^2 = 2EE' \cos(\theta)$
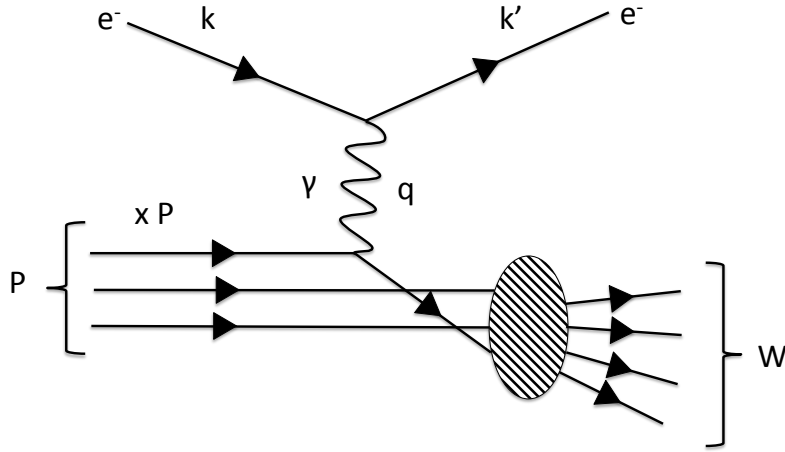- $x = Q^2/2P \cdot q = Q^2/2M_p\nu$
- $y = 2P \cdot q/s^2 = \nu/E$

Figure B.1: Diagram of an e^-p collision in a reference frame where the proton has some large momentum P. The incoming electron transfers 4-momentum $q = k - k'$ to a parton with a fraction x of the total proton momentum. The total invariant mass of the hadronic products of the inelastic collision is W.

- $W^2 = (P + q)^2 = M_p^2 + 2M\nu - Q^2$

Here W is the invariant mass of the hadronic system recoiling off of the scattered lepton, and Px is the momentum of the struck parton. The variable x can then be interpreted as the momentum fraction of the proton carried by the struck parton, when viewed in a frame where the proton has a high momentum.

The differential inelastic cross section may then be given in terms of x and y as

$$\frac{d^2\sigma^{ep}}{dxdy} = \frac{8\pi\alpha^2 M_p E}{Q^4} \left[F_2(x, Q^2)(1 - y) + xy^2 F_1(x, Q^2) \right], \tag{B.3}$$

where F_1 and F_2 are the proton structure functions, which are derived from a Fourier transform of the the proton charge density. At very low, Q^2, their behavior is consistent with the view of the proton as a point particle.

Deviation from pointlike behavior was first observed by the SLAC-MIT Experiment in the measurement of the inelastic cross section at $\theta = 6°$, and $10°$ with incident electron energies of nearly 20 GeV [173]. The structure functions were found to have little to no dependence on Q^2, which was consistent with the interpretation of the proton as comprised of small pointlike particles [174]. This Q^2 independence is called "Bjorken scaling", and the variable x is referred to as the Bjorken x. The observed structure functions in early DIS experiments were also found to approximately satisfy the Callan-Gross relationship [175]

$$R = (F_2 - 2xF_1)/2xF_1 \approx 0, \tag{B.4}$$

which indicates that the charged partons have a spin of $1/2$.

Assuming $R = 0$ and ignoring heavy-quark contributions, one can then relate the momentum densities of the partons, called the parton density functions (PDFs) for the u, d, s quarks, and $\bar{u}$, $\bar{d}$, $\bar{s}$ antiquarks to the structure function F_2 as

$$
\begin{aligned}
\frac{F_2(x, Q^2)}{x} = \quad & \frac{4}{9} \left[f_u(x, Q^2) + f_{\bar{u}}(x, Q^2) \right] \\
& + \frac{1}{9} \left[f_d(x, Q^2) + f_{\bar{d}}(x, Q^2) \right] \\
& + \frac{1}{9} \left[f_s(x, Q^2) + f_{\bar{s}}(x, Q^2) \right].
\end{aligned} \tag{B.5}
$$

The differences between the quark and antiquark densities can be determined using the additional F_3 structure function arising in parity-violating charged current interactions which are not described here.

In QCD, the radiation of hard gluons from quarks leads to logarithmic violations of Bjorken scaling, which become large for small x. This prediction of scaling violation is clearly seen in the combined DIS cross section data from the early fixed-target experiments and the H1 and ZEUS experiments at the HERA accelerator [177], where $e^{\pm}$p collisions probed a much larger region of the (Q^2, x) phase space. These measured cross sections are shown in Fig. B.2 [176]. The QCD prediction for HERAPDF 1.0 is also shown, which provides a measurement of the PDFs of the quarks and gluons. Note that as the Q^2 dependence of the PDF at each value of x is constrained by the DGLAP equation [21–23], these fits to the combined data also demonstrate an impressive confirmation of the theory of QCD.
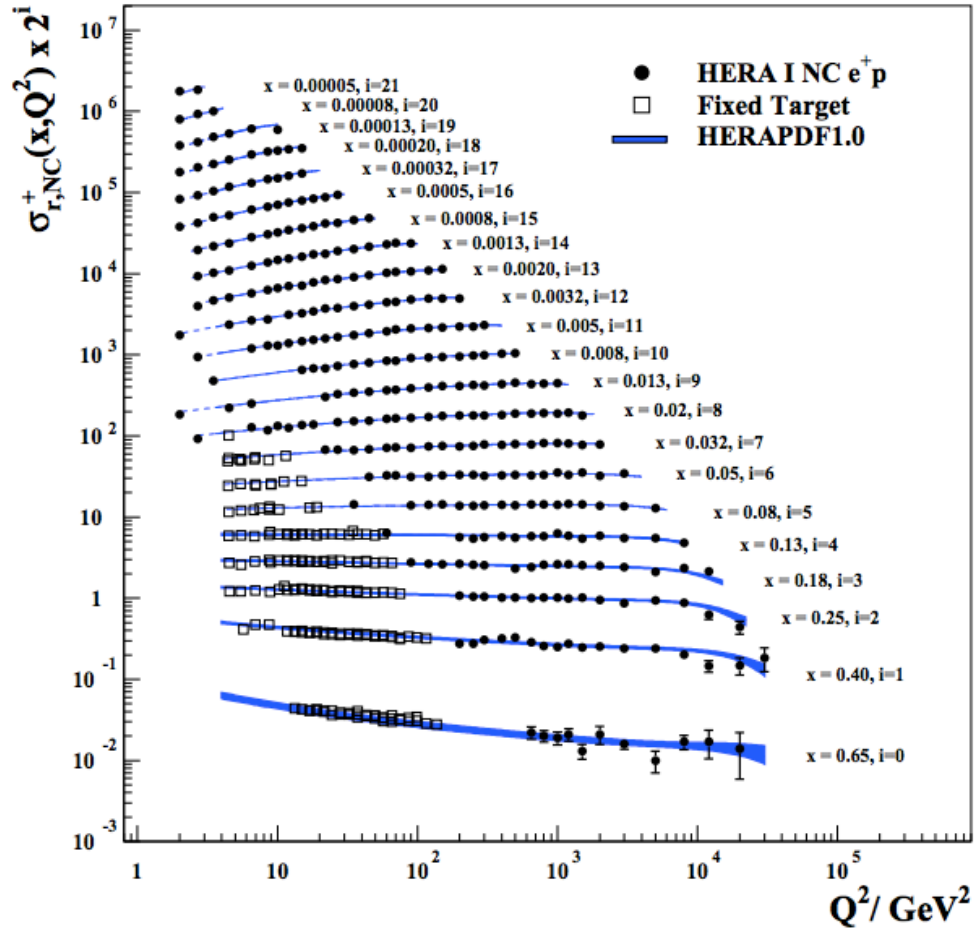
Figure B.2: The HERA combined neutral current $e^{\pm}p$ reduced cross section and fixed-target data as a function of Q^2 shown for various values of x. The HERAPDF 1.0 fit is superimposed, with the band representing the uncertainty in the fit. Dashed lines are shown for Q^2 values not included in the QCD analysis [176].

Appendix C

Monte Carlo Generators

The goal of an event generator is to produce as realistically as possible a collision event as would be produced in a real detector experiment predicting the same average behavior and fluctuations that would be seen in the actual data. In the actual data, such fluctuations arise from quantum mechanics. In the event generator, the Monte Carlo technique is used to select random variables based on probability distributions.

In practice, each event may take the form as a list of particles, their positions, and momenta, which were randomly produced in a given simulated collision. In this Appendix, the MC event generators which were used in conjunction with the GEANT4 simulation of the CMS detector to describe the detector response to pp, pPb, and PbPb collisions are described.

C.1   The PYTHIA Event Generator

The PYTHIA generator is a general purpose event generator used frequently in high energy physics which focuses on hard processes in a variety of possible collision systems, such as e^+e^-, e^-p, or pp. The program attempts to produce complete events using as best as possible the current understanding of all underlying physics.

This flexible program is capable of producing events focusing on a number of possible processes in the context of pp collisions, ranging from non-perturbative diffractive events, to the production of jets from hard partonic scattering, to Higgs boson production and decays, and even to many theorized processes beyond the standard model. For the measurements presented in this thesis, version 6.423 of the program is used. For the purpose of measuring inclusive hadron production in pp collisions, the program is configured to produce inelastic, diffractive, low-p_T production, and QCD jet events with a probability of each type determined by the expected cross section. In order to extend the statistical reach of the simulation for high-p_T hadrons produced in the final state the program is also set to simulate collision events with some minimum momentum transfer in the $2 \rightarrow 2$ hard scattering process.

Hard scattering processes are simulated using a factorization scheme, starting with user-definable parton distribution function parametrizations, next-to-leading order $2 \rightarrow 2$ pQCD scattering matrix computations, and finally with jet fragmentation described by the Lund string fragmentation model [178]. Initial- and final-state gluon radiation is incorporated as a partonic shower, allowing for the production of multi-jet $2 \rightarrow 3$, $2 \rightarrow 4$, ..., $2 \rightarrow n$ processes from basic $2 \rightarrow 2$ hard scattering QCD processes. The fragmentation of the beam-remnants consisting of the partonic content of the protons not directly entering into the hard scattering process is also modeled.

The parametrization of the PDF, inputs into the string fragmentation model, and variables relating to the other features of the PYTHIA simulation may be "tuned" in order to accurately reproduce

recent collision data. In this case, from the available 2010 and 2011 results in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV from the ATLAS and CMS experiments have been used to produce the Z2 [117] and ProQ20 [118] tunes used in the simulations described in this thesis. A full description of the physics models incorporated into PYTHIA 6.4 is given in Ref. [116].

C.2   The HYDJET Event Generator

The HYDJET event generator incorporates the early RHIC findings of jet quenching and elliptic flow of the QGP medium into a model of AA collision systems. The simulation includes a superposition of the "soft" or low-p_T hadron production described by a hydrodynamical model, and "hard" production of hadrons produced in jets from partons which lose energy as they traverse the medium. Each collision is performed with some finite impact parameter b which defines both the shape of the overlap area and the expected number of binary collisions N_{coll}.

For the hard portion of the simulation, the spectrum of hard scattered partons from the binary collisions is computed using PYTHIA version 6.2, and fragmentation is not initially performed. The spacial position of the scattering vertex is placed randomly according to a probability distribution taken from the expected nucleon density in the overlap region. The scattering cross section of the parton traversing the medium is then calculated, and the generation of displacements between scatterings is computed, with length-dependent energy loss of the parton by gluon radiation. This rescattering halts when the parton exits the dense portion of the medium or is reduced in p_T to twice the effective temperature of the medium. The quenched parton and radiated gluons are then fragmented into hadrons using the PYTHIA model.

The 4-momenta and masses of "soft" hadrons from the bulk medium are generated according to an isotropic Boltzmann distribution for each "liquid element", where liquid elements are generated with a random position and 4-velocity u_μ in accordance with the motion and spacial extent of the medium at freeze-out, under the assumption of an initial elliptical eccentricity given by the geometrical overlap determined from the impact parameter b. The 4-momentum of the randomly produced hadrons is then boosted according to u_μ.

A more detailed description of the HYDJET model may be found in Ref. [119].

C.3   The HIJING Event Generator

The Heavy-Ion Jet Interaction Generator, or HIJING, version 1.383 may be used to produce pp, pA, or AA collision events. For the measurements in this thesis, it is used to generate pA collisions. The model treats each nucleon-nucleon collision separately, with the number of collisions is probabilistically generated based on the overlap area between the two colliding nuclei with some random impact parameter b. Production of jets is simulated with a factorization scheme, where shadowing effects are incorporated into the impact-parameter dependent parton distribution functions of the nucleons in the large nucleus, and PYTHIA subroutines are used for the determination of hard process cross sections and subsequent radiation. Fragmentation is handled by a Lund-inspired model.

Jet quenching is simulated using a length-dependent energy loss parameter dE/dx. A detailed description of the model and software can be found in Ref. [120, 121]

C.4 The EPOS Event Generator

The EPOS [122] generator describes pp and pA collisions in the same framework, using a phenomenological approach that departs from the usual approach towards hard production in term of a single $2 \to 2$ interaction between initial partons described in terms of the parton distribution functions, in which the other initial-state partons do not play a further role in the interaction.
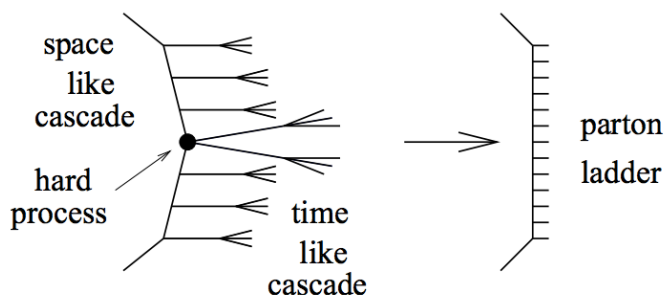

Figure C.1: Cartoon of elementary parton-parton scattering, where the perturbative hard scattering process is preceded by an initial space-like cascade of partons, which emit further partons in a final time-like cascade. This process is described symbolically by the parton ladder [122].

The dynamical process that underlies the factorized $2 \to 2$ approach may be viewed as a cascade of partonic emissions, which is shown in Fig C.1 and represented symbolically as a "parton ladder".

In an approach where multiple such scatterings between initial-state partons may occur, or multiple ladders, one must be careful to ensure overall energy conservation as the total energy must be shared among the individual partonic interactions. Furthermore, quantum mechanical interference should be considered.

This requires not only a consideration of inelastic scatterings between partons which result in particle production, called "open ladders", but also purely elastic interactions in which the partons exchange momentum, called "closed ladders". Examples of these processes are shown diagramatically in Fig. C.2.

The main feature of the EPOS model is to incorporate the possibility that a parton produced in one ladder may interact with another initial-state parton in the target nucleus, producing a new ladder. This phenomenon of "ladder splitting" is shown in Fig. C.3, and may be expected to be more prevalent in a heavy target nucleus, as the increased partonic densitiy may result in more available partons to be found in a given phase space.
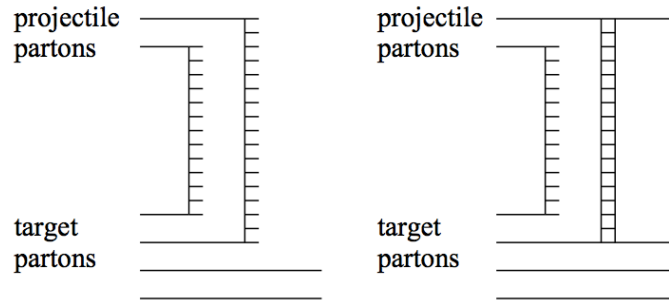
Figure C.2: Basic parton-parton interactions in a collision event with multiple partonic scatterings. Each initial-state parton from the projectile interacts with exactly one parton in the target, either elastically as a "closed ladder" or inelastically as an "open ladder" [122].

Figure C.3: Inelastic (Left) and elastic (Right) splitting of a parton ladder by interaction with a second target parton [122].

The effects of ladder splitting in terms of accurately describing dAu data from RHIC and the mathematical details of the EPOS model may be found in Ref. [122].

Bibliography

[1] J. Beringer et al. (Particle Data Group). The review of particle physics. *Phys. Rev. D*, **86**:01001, 2012. and 2013 partial update for the 2014 edition.
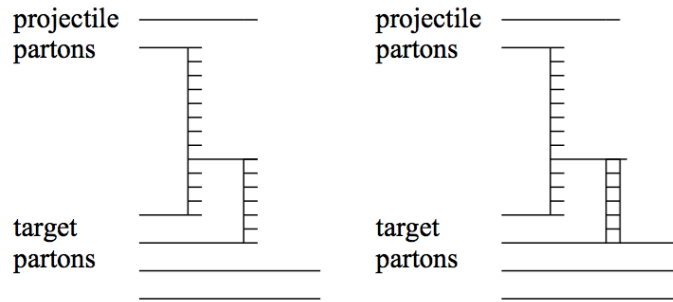
[2] G. Kane. *Modern elemetary particle physics*. Perseus Publishing, Cambridge, MA, 1993.

[3] D. Griffiths. *Introduction to elementary particles*. Weinheim : Wiley-VCH Verlag, New York, NY, 2004.

[4] S. Novaes. Standard model: an introduction, 1999. IFT-P-010-2000, arXiv:hep-ph/0001283 (http://arxiv.org/abs/hep-ph/0001283).

[5] A. Pich. The Standard model of electroweak interactions, 1994. FTUV-94-62, IFIC-94-59, C94-02-07, arXiv:hep-ph/9412274 (http://arxiv.org/abs/hep-ph/9412274).

[6] W. Greiner et al. *Quantum chromodynamics*. Springer-Verlag, Berlin Heidelberg, 2002.

[7] A. Pich. Quantum chromodynamics, 1995. FTUV-95-19, IFIC-95-19, C94-08-29, arXiv:hep-ph/9505231 (http://arxiv.org/abs/hepph/9505231).

[8] D. Gross. Twenty five years of asymptotic freedom. *Nucl. Phys. Proc. Suppl.*, **74**:426–446, 1999.

[9] D. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, **30**:1343–1346, Jun 1973.

[10] E. Poggio et al. Smearing the quark model. *Phys. Rev. D*, **13**:1958, 1976.

[11] A. Bazavov et al. Nonperturbative QCD simulations with 2+1 flavors of improved staggered quarks. *Rev. Mod. Phys.*, **82**:1349, 2010.

[12] S. Mandelstam. Determination of the pion-nucleon scattering amplitude from dispersion relations and unitarity. General theory. *Phys. Rev.*, **112**:1344–1360, 1958.

[13] V. Gribov. *The theory of complex angular momenta: Gribov lectures on theoretical physics*. Cambridge University Press, Cambridge, MA, 2003.

[14] J. Collins et al. Factorization of hard processes in QCD. *Adv. Ser. Direct. High Energy Phys.*, **5**:1–91, 1988.

[15] J. Collins and J.-W. Qiu. $k_{\rm T}$ factorization is violated in production of high-transverse-momentum particles in hadron-hadron collisions. *Phys. Rev. D*, **75**:114014, 2007.

[16] R. Brock et al. (CTEQ Collaboration). Handbook of perturbative QCD: Version 1.0. *Rev. Mod. Phys.*, **67**:157–248, 1995.

[17] S. Brodsky et al. Structure functions are not parton probabilities. *Phys. Rev. D*, **65**:114025, 2002.

[18] L. Hung-Liang et al. New parton distributions for collider physics. *Phys. Rev. D*, **82**:074024, 2010.

[19] A. Martin et al. Parton distributions for the LHC. *Eur. Phys. J. C*, **63**:189–285, 2009.

[20] R. Ball et al. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl. Phys. B*, **838**:136–206, 2010.

[21] V. Gribov and L. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, **15**:438–450, 1972.

[22] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B*, **126**:298–318, 1977.

[23] Y. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and ee-annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JTEP*, **46**:641, 1977.

[24] B. Kniehl et al. Fragmentation functions for pions, kaons, and protons at next-to-leading order. *Nucl. Phys. B*, **582**:514–536, 2000.

[25] S. Albino et al. AKK update: improvements from new theoretical input and experimental data. *Nucl. Phys. B*, **803**:42–104, 2008.

[26] D. de Florian et al. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev. D*, **75**:114010, 2007.

[27] S. Chatrchyan et al. (CMS Collaboration). Study of high-p_T charged particle suppression in PbPb compared to *pp* collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Eur. Phys. J. C*, **72**:1945, 2012.

[28] J. C. Collins and M. J. Perry. Superdense matter: Neutrons or asymptotically free quarks? *Phys. Rev. Lett.*, **34**:1353–1356, 1975.

[29] Edward V. Shuryak. Quantum chromodynamics and the theory of superdense matter. *Physics Reports*, **61**:71 – 158, 1980.

[30] A. Bazavov et al. Equation of state and QCD transition at finite temperature. *Phys. Rev. D*, **80**:014504, 2009.

[31] M. Troyer and U. Wiese. Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations. *Phys. Rev. Lett.*, **94**:170201, 2005.

[32] R. Rapp et al. Diquark Bose condensates in high density matter and instantons. *Phys. Rev. Lett.*, **81**:53–56, 1998.

[33] M. Alford et al. Color superconductivity in dense quark matter. *Rev. Mod. Phys.*, **80**:1455–1515, 2008.

[34] K. Fukushima and T. Hatsuda. The phase diagram of dense QCD. *Rept. Prog. Phys.*, **74**:014001, 2011.

[35] H. Gutbrod et al. Plastic ball experiments. *Rep. Prog. Phys.*, **52**:1267, 1989.

[36] U. Heinz and M. Jacob. Evidence for a new state of matter: An assessment of the results from the CERN lead beam programme, 2000. arXiv:nucl-th/0002042 (http://arxiv.org/abs/nucl-th/0002042).

[37] I. Arsene et al. (BRAHMS Collaboration). Quark-gluon plasma and color glass condensate at RHIC? The perspective from the BRAHMS experiment. *Nucl. Phys. A*, **757**:1–27, 2005.

[38] K. Adcox et al. (PHENIX Collaboration). Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the phenix collaboration. *Nucl. Phys. A*, **757**:184–283, 2005.

[39] B. Back et al. (PHOBOS Collaboration). The PHOBOS perspective on discoveries at RHIC. *Nucl. Phys. A*, **757**:28–101, 2005.

[40] J. Adams et al. (STAR Collaboration). Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions. *Nucl. Phys. A*, **757**:102, 2005.

[41] D. Rischke (ed.) and G. Levin (ed.). *Quark-gluon plasma. New discoveries at RHIC: Case for the strongly interacting quark-gluon plasma. Contributions from the RBRC Workshop held May 14-15, 2004*. *Nucl. Phys. A*, **750**:1–171, 2005.

[42] L. Kumar for the STAR Collaboration. STAR Results from the RHIC Beam Energy Scan-I. *Nucl. Phys. A*, **904-905**:256c–263c, 2013.

[43] T. Mitchell for the PHENIX Collaboration. The RHIC Beam Energy Scan Program: Results from the PHENIX Experiment. *Nucl. Phys. A*, **904-905**:903c–906c, 2013.

[44] S. Chatrchyan et al. (CMS Collaboration). Measurement of the elliptic anisotropy of charged particles produced in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Phys. Rev. C*, **87**:014902, 2013.

[45] M. Miller et al. Glauber modeling in high-energy nuclear collisions. *Ann. Rev. Nucl. Part. Sci.*, **57**:205–243, 2007.

[46] J. Bjorken. Highly relativistic nucleus-nucleus collisions: The central rapidity region. *Phys. Rev. D*, **27**:140–151, 1983.

[47] B. Alver and G. Roland. Collision-geometry fluctuations and triangular flow in heavy-ion collisions. *Phys. Rev. C*, **81**:054905, 2010.

[48] J.-Y. Ollitrault et al. Effect of flow fluctuations and nonflow on elliptic flow methods. *Phys. Rev. C*, **80**:014904, 2009.

[49] P. Kolb and U. Heinz. Hydrodynamic description of ultrarelativistic heavy ion collisions, 2003. arXiv:nucl-th/0305084 (http://arxiv.org/abs/nucl-th/0305084).

[50] U. Heinz and R. Snellings. Collective flow and viscosity in relativistic heavy-ion collisions. *Ann. Rev. Nucl. Part. Sci.*, **63**:123–151, 2013.

[51] S. Adler et al. (PHENIX Collaboration). Identified charged particle spectra and yields in Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. C*, **69**:034909, 2004.

[52] P. Bożek and W. Broniowski. Transverse-momentum fluctuations in relativistic heavy-ion collisions from event-by-event viscous hydrodynamics. *Phys. Rev. C*, **85**:044910, 2012.

[53] A. Turlapov et al. Is a gas of strongly interacting atomic fermions a nearly perfect fluid? *J. Low Temp. Phys.*, **150**:567–576, 2008.

[54] P. Kovtun et al. Holography and hydrodynamics: Diffusion on stretched horizons. *JHEP*, **0310**:064, 2003.

[55] S. Bass et al. Systematic comparison of jet energy-loss schemes in a realistic hydrodynamic medium. *Phys. Rev. C*, **79**:024901, 2009.

[56] S. Jeon. Jet quenching in evolving QGP medium. *Nucl. Phys. A*, **830**:107C–114C, 2009.

[57] M. Shimomura for the PHENIX Collaboration. High-p_T π^0, η identified and inclusive charged hadron spectra from PHENIX. *Nucl. Phys. A*, **774**:457–460, 2006.

[58] CMS Collaboration. Z boson production with the 2011 data in PbPb collisions, 2013. CMS-PAS-HIN-13-004 (http://cds.cern.ch/record/1564677/).

[59] K. Gallmeister et al. Quenching of high transverse momentum hadron spectra by hadronic interactions in heavy ion collisions at RHIC. *Phys. Rev. C*, **67**:044905, 2003.

[60] S. Adler et al. (PHENIX Collaboration). Absence of suppression in particle production at large transverse momentum in $\sqrt{s_{NN}} = 200$ GeV d+Au collisions. *Phys. Rev. Lett.*, **91**:072303, 2003.

[61] A. Adare et al. (PHENIX Collaboration). Spectra and ratios of identified particles in Au+Au and d+Au collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$. *Phys. Rev. C*, **88**:024906, 2013.

[62] J. Cronin et al. Production of hadrons at large transverse momentum at 200, 300, and 400 GeV. *Phys. Rev. D*, **11**:3105–3123, 1975.

[63] A. Accardi. Cronin effect in proton nucleus collisions: A Survey of theoretical models, 2002. arXiv:hep-ph/0212148 (http://arxiv.org/abs/hep-ph/0212148).

[64] R. Hwa, and C.B. Yang. Proton production in d+Au collisions and the Cronin effect. *Phys. Rev. C*, **70**:037901, 2004.

[65] R. Hwa, and C.B. Yang. Final state interaction as the origin of the Cronin effect. *Phys. Rev. Lett.*, **93**:082302, 2004.

[66] F. Arleo and J.-P. Guillet. nPDF generator, 2010. http://lapth.cnrs.fr/npdfgenerator/.

[67] I. Helenius et al. Impact-parameter dependent nuclear parton distribution functions: EPS09s and EKS98s and their applications in nuclear hard processes. *JHEP*, **1207**:073, 2012.

[68] M. Arneodo. Nuclear effects in structure functions. *Phys. Reports*, **240**:301–393, 1994.

[69] S. Adler et al. (PHENIX Collaboration). Centrality dependence of pi0 and eta production at large transverse momentum in $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ d+Au collisions. *Phys. Rev. Lett.*, **98**:172302, 2007.

[70] J. Aubert et al. (EMC Collaboration). The ratio of the nucleon structure functions $F2_n$ for iron and deuterium. *Phys. Lett. B*, **123**:275, 1983.

[71] L. Gribov et al. Semihard processes in qcd. *Phys. Reports*, **100**:1–150, 1983.

[72] A. Mueller and J. Qiu. Gluon recombination and shadowing at small values of x. *Nucl. Phys. B*, **268**:427–452, 1986.

[73] L. McLerran and R. Venugopalan. Gluon distribution functions for very large nuclei at small transverse momentum. *Phys. Rev. D*, **49**:3352–3355, 1994.

[74] L. McLerran and R. Venugopalan. Computing quark and gluon distribution functions for very large nuclei. *Phys. Rev. D*, **49**:2233–2241, 1994.

[75] F. Gelis et al. The color glass condensate. *Ann. Rev. Nucl. Part. Sci.*, **60**:463–489, 2010.

[76] A. Dumitru et al. The Color glass condensate and hadron production in the forward region. *Nucl. Phys. A*, **765**:464–482, 2006.

[77] S. Chatrchyan et al. (CMS Collaboration). Observation of long-range near-side angular correlations in proton-lead collisions at the LHC. *Phys. Lett. B*, **718**:795–814, 2013.

[78] B. Abelev et al. (ALICE Collaboration). Long-range angular correlations on the near and away side in p-Pb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$. *Phys. Lett. B*, **719**:29–41, 2013.

[79] G. Aad et al. (ATLAS Collaboration). Observation of Associated Near-side and Away-side Long-range Correlations in $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ Proton-lead Collisions with the ATLAS Detector. *Phys. Rev. Lett.*, **110**:182302, 2013.

[80] S. Chatrchyan et al. (CMS Collaboration). Multiplicity and transverse momentum dependence of two- and four-particle correlations in pPb and PbPb collisions. *Phys. Lett. B*, **724**:213–240, 2013.

[81] P. Bozek and W. Broniowski. Correlations from hydrodynamic flow in p-Pb collisions. *Phys. Lett. B*, **718**:1557–1561, 2013.

[82] K. Dusling and R. Venugopalan. Comparison of the color glass condensate to dihadron correlations in proton-proton and proton-nucleus collisions. *Phys. Rev. D*, **87**:094034, 2013.

[83] B. Abelev et al. (ALICE Collaboration). Transverse momentum distribution and nuclear modification factor of charged particles in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *Phys. Rev. Lett.*, **110**:082302, 2013.

[84] L. Evans and P. Bryant. LHC machine. *JINST*, **03**(08):S08001, 2008.

[85] G. Aad et al. (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**:1 – 29, 2012.

[86] S. Chatrchyan et al. (CMS Collaboration). Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, **716**:30 – 61, 2012.

[87] CERN. CMS detector drawings (CMS-PHO-GEN-2012-002), 2012. https://cds.cern.ch/record/1433717.

[88] S. Chatrchyan et al. (CMS Collaboration). The CMS experiment at the CERN LHC. *JINST*, **03**:S08004, 2008.

[89] S. Chatrchyan et al. (CMS Collaboration). Dependence on pseudorapidity and on centrality of charged hadron production in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *JHEP*, **08**:141, 2011.

[90] M. Swartz et al. A new technique for the reconstruction, validation, and simulation of hits in the CMS Pixel Detector. *PoS*, **Vertex 2007**:035, 2007. CMS-NOTE-2007-033 (https://cds.cern.ch/record/1073691).

[91] M. Raymond et al. The CMS Tracker APV25 0.25 μmCMOS readout chip. In *Proceedings of the 6th Workshop on Electronics for LHC Experiments*, Cracow Poland, 2000.

[92] G. Bayatyan et al. (CMS Collaboration). *CMS TriDAS project : Technical design report, volume 1: The Trigger Systems*, 2004. CERN-LHCC-2000-038 ; CMS-TDR-6-1 (http://cds.cern.ch/record/706847).

[93] G. Bayatyan et al. (CMS Collaboration). *CMS The TriDAS Project : Technical design report, volume 2: Data Acquisition and High-Level Trigger*, 2002. CERN-LHCC-2002-026 ; CMS-TDR-6 (http://cds.cern.ch/record/578006).

[94] R. Brun and F. Rademakers. Root: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, **389**:81–86, 1997. see also http://root.cern.ch.

[95] J. Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, **53**:270–278, 2006.

[96] C. Grandi et al. CMS computing model: The "CMS Computing Model RTAG", 2004. CMS-NOTE-2004-031. CERN-LHCC-2004-035. LHCC-G-083 (http://cds.cern.ch/record/814248).

[97] CMS Collaboration. Tracking and vertexing results from first collisions, 2010. CMS-PAS-TRK-10-001 (http://cds.cern.ch/record/1258204).

[98] G. Welch and G. Bishop. An introduction to the Kalman filter, 2006. TR 95-014 University of North Carolina at Chapel Hill, Department of Computer Science.

[99] T. Miao et al. Beam position determination using tracks, 2007. CMS-NOTE-2007-021 (http://cds.cern.ch/record/1061285).

[100] S. Cucciarelli et al. Position determination of pixel hits, 2002. CMS-NOTE-2002-049 (https://cds.cern.ch/record/687475).

[101] CMS Collaboration. Alignment of the cms silicon tracker during commissioning with cosmic rays. *JINST*, **5**:T03009, 2010.

[102] V. Blobel. Software alignment for tracking detectors. *Nucl. Instrumen. Meth. A*, **566**:5–13, 2006.

[103] G. Flucke et al. CMS silicon tracker alignment strategy with the Millepede II algorithm. *JINST*, **3**:P09002, 2008.

[104] V. Karimäki et al. The HIP algorithm for track based alignment and its application to the CMS Pixel Detector, 2006. CMS-NOTE-2006-018 (http://cds.cern.ch/record/926537).

[105] S. Cucciarelli et al. Track reconstruction, primary vertex finding and seed generation with the Pixel Detector, 2006. CMS-NOTE-2006-026 (http://cds.cern.ch/record/927384).

[106] P. Billoir and S. Qian. Simultaneous pattern recognition and track fitting by the kalman filtering method. *Nucl. Instr. Meth. A*, **294**:219–228, 1990.

[107] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nucl. Instr. Meth. A*, **262**:444–450, 1987.

[108] "W. Adam et al.". Track reconstruction in the CMS tracker, 2006. CMS-NOTE-2006-041 (https://cds.cern.ch/record/934067).

[109] S. Kirkpatrick et al. Optimization by simulated annealing. *Science*, **220**:671–680, 1983.

[110] V. Černý. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *J. Opt. Theor. App.*, **45**:41–51, 1985.

[111] K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proc. IEEE*, **86**:2210–2239, 1998.

[112] T. Speer et al. Vertex fitting in the CMS Tracker, 2006. CMS-NOTE-2006-032 (http://cds.cern.ch/record/927395/).

[113] CERN. CMS collision events: First lead ion collisions (CMS-PHO-EVENTS-2010-002), 2010. https://cds.cern.ch/record/1305179.

[114] C. Roland. Track reconstruction in heavy ion events using the CMS Tracker, 2006. CMS-NOTE-2006-031 (http://cds.cern.ch/record/927403).

[115] M. Dobbs and J. Hansen. The HepMC C++ Monte Carlo event record for high energy physics. *Comp. Phys. Comm.*, **134**:41–46, 2001.

[116] T. Sjöstrand et al. PYTHIA 6.4 Physics and Manual. *JHEP*, **05**:026, 2006.

[117] R. Field. Early LHC underlying event data - findings and surprises, 2010. arXiv:hep-ph/1010.3558 (http://arxiv.org/abs/1010.3558).

[118] A. Buckley et al. Systematic event generator tuning for the LHC. *Eur. Phys. J. C*, **65**:331–357, 2010.

[119] I. Lokhtin and A. Snigirev. A model of jet quenching in ultrarelativistic heavy ion collisions and high-p_T hadron spectra at RHIC. *Eur. Phys. J. C*, **45**:211–217, 2006.

[120] X. Wang and M. Gyulassy. HIJING: A Monte Carlo model for multiple jet production in p p, p A and A A collisions. *Phys. Rev. D*, **44**:3501–3516, 1991.

[121] M. Gyulassy and X. Wang. HIJING 1.0: A Monte Carlo program for parton and particle production in high-energy hadronic and nuclear collisions. *Comput. Phys. Commun.*, **83**:307, 1994.

[122] K. Werner et al. Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron-gold collisions at RHIC. *Phys. Rev. C*, **74**:044902, 2006.

[123] S. Chatrchyan et al. (CMS Collaboration). Commissioning of the CMS experiment and the cosmic run at four tesla. *JINST*, **5**:T03001, 2010.

[124] V. Khachatryan et al. (CMS Collaboration). CMS tracking performance results from early LHC operation. *Eur. Phys. J. C*, **70**:1165–1192, 2010.

[125] S. Chatrchyan et al. (CMS Collaboration). Charged particle transverse momentum spectra in *pp* collisions at $\sqrt{s} = 0.9$ and 7 TeV. *JHEP*, **1108**:086, 2011.

[126] CMS Collaboration. Measurement of tracking efficiency, 2010. CMS-PAS-TRK-10-002 (`http://cds.cern.ch/record/1279139`).

[127] CMS Collaboration. Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_\mathrm{T}^\mathrm{miss}$, 2009. CMS-PAS-PFT-09-001 (`http://cdsweb.cern.ch/record/1194487`).

[128] CMS Collaboration. Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector, 2010. CMS-PAS-PFT-10-001 (`http://cdsweb.cern.ch/record/1247373`).

[129] M. Cacciari et al. FastJet user manual. *Eur. Phys. J. C*, **72**:1896, 2012.

[130] M. Cacciari et al. The anti-k_t jet clustering algorithm. *JHEP*, **04**:063, 2008.

[131] O. Kodolova et al. The performance of the jet identification and reconstruction in heavy ions collisions with CMS detector. *Eur. Phys. J. C*, **50**:117, 2007.

[132] S. Chatrchyan et al. (CMS Collaboration). Observation and studies of jet quenching in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV . *Phys. Rev. C*, **84**:024906, 2011.

[133] B. Abelev et al. (ALICE Collaboration). Multi-strange baryon production in pp collisions at $\sqrt{s} = 7$ TeV with ALICE. *Phys. Lett. B*, **712**:309–318, 2012.

[134] P. Skands. Tuning Monte Carlo generators: The Perugia tunes. *Phys. Rev. D*, **82**:074018, 2010.

[135] K. Werner et al. Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3, 2013. arXiv:1312.1233 (`http://arxiv.org/abs/1312.1233`).

[136] Ø. Djuvsland, Øystein and J. Nystrand. Single and double photonuclear excitations in Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV at the CERN Large Hadron Collider. *Phys. Rev. C*, **83**:041901, 2011.

[137] V. Khachatryan et al. (CMS Collaboration). Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV. *JHEP*, **1002**:041, 2010.

[138] B. Alver et al. The PHOBOS Glauber Monte Carlo, 2008. arXiv:0805.4411 (`http://arxiv.org/abs/0805.4411`).

[139] R. Woods and D. Saxon. Diffuse surface optical model for nucleon-nuclei scattering. *Phys. Rev.*, **95**:577–578, 1954.

[140] D. d'Enterria and K. Reygers. Glauber Monte Carlo results for nucleus-nucleus and proton-nucleus collisions at the Large Hadron Collider and the role of the neutron skin. To be submitted.

[141] S. Adler et al. (PHENIX Collaboration). Identified charged particle spectra and yields in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. C*, **69**:034909, 2004.

[142] V. Khachatryan et al. (CMS Collaboration). Transverse-momentum and pseudorapidity distributions of charged hadrons in *pp* collisions at $\sqrt{s}$ = 7 TeV. *Phys. Rev. Lett.*, **105**:022002, 2010.

[143] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, **52**:479, 1988.

[144] G. Wilk and Z. Włodarczyk. Power laws in elementary and heavy-ion collisions: A story of fluctuations and nonextensivity? *Eur. Phys. J. A*, **40**:299, 2009.

[145] T. Biró et al. Non-extensive approach to quark matter. *Eur. Phys. J. A*, **40**:325, 2009.

[146] C. Adler et al. (STAR Collaboration). Centrality dependence of high-p_T hadron suppression in Au + Au collisions at $\sqrt{s_{NN}}$ = 130 GeV. *Phys. Rev. Lett.*, **89**:202301, 2002.

[147] S. Chatrchyan et al. (CMS Collaboration). Measurement of higher-order harmonic azimuthal anisotropy in PbPb collisions at a nucleon-nucleon center-of-mass energy of 2.76 TeV, 2013. arXiv:nucl-ex/1310.8651 (http://arxiv.org/abs/1310.8651), submitted to *Phys. Rev. C*.

[148] B. Back et al. (PHOBOS Collaboration). Centrality and pseudorapidity dependence of elliptic flow for charged hadrons in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. C*, **72**:051901, 2005.

[149] W. Busza. Extended longitudinal scaling: Direct evidence of saturation. *Nucl. Phys. A*, **854**:57–63, 2011.

[150] J. Benecke et al. Hypothesis of limiting fragmentation in high-energy collisions. *Phys. Rev.*, **188**:2159–2169, 1969.

[151] R. Bhalerao et al. New flow observables. *J. Phys. G*, **38**:124055, 2011.

[152] J. Adams, et al. (STAR Collaboration). Measurements of transverse energy distributions in Au + Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. C*, **70**:054907, 2004.

[153] S. Adler et al. (PHENIX Collaboration). Systematic studies of the centrality and $\sqrt{s_{NN}}$ dependence of the $dE_{\mathrm{T}}/d\eta$ and $dN_{\mathrm{ch}}/d\eta$ in heavy ion collisions at mid-rapidity. *Phys. Rev. C*, **71**:034908, 2005.

[154] S. Chatrchyan et al. (CMS Collaboration). Measurement of the pseudorapidity and centrality dependence of the transverse energy density in PbPb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV. *Phys. Rev. Lett.*, **109**:152303, 2012.

[155] F. Abe et al. (CDF Collaboration). Transverse momentum distributions of charged particles produced in $\bar{p}p$ interactions at $\sqrt{s}$ = 630 GeV and 1800 GeV. *Phys. Rev. Lett.*, **61**:1819, 1988.

[156] T. Aaltonen et al. (CDF Collaboration). Measurement of particle production and inclusive differential cross sections in p anti-p Collisions at $\sqrt{s}$ = 1.96 TeV. *Phys. Rev. D*, **79**:112005, 2009.

[157] CMS Collaboration. Charged particle nuclear modification factor and pseudorapidity asymmetry in pPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV with CMS, 2013. CMS-PAS-HIN-12-017 (http://cds.cern.ch/record/1625865).

[158] CMS Collaboration. Transverse momentum and pseudorapidity dependence of charged particle production and nuclear modification factor in pPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV with CMS, 2013. CMS AN-12-377 (internal CMS collaboration document).

[159] CMS Collaboration. Supplimentary plots to accompany CMS-PAS-HIN-12-017, 2013. (https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsHIN12017).

[160] J. Albacete et al. Predictions for p+Pb collisions at $\sqrt{s_{NN}}$ = 5 TeV. *Int. J. Mod. Phys. E*, **22**:1330007, 2013.

[161] K. Eskola et al. The scale dependent nuclear effects in parton distributions for practical applications. *Eur. Phys. J. C*, **9**:61–68, 1999.

[162] K. Eskola et al. An improved global analysis of nuclear parton distribution functions including RHIC data. *JHEP*, **0807**:102, 2008.

[163] M. Hirai et al. Determination of nuclear parton distributions. *Phys. Rev. D*, **64**:034003, 2001.

[164] S.-Y. Li and X.-N. Wang. Gluon shadowing and hadron production at RHIC. *Phys. Lett. B*, **527**:85–91, 2002.

[165] V. Topor Pop et al. Strong longitudinal color field effects in pp collisions at energies available at the Large Hadron Collider. *Phys. Rev. C*, **83**:024902, 2011.

[166] Z. Lin et al. A multi-phase transport model for relativistic heavy ion collisions. *Phys. Rev. C*, **72**:064901, 2005.

[167] Y. Zhang et al. High p_T pion and kaon production in relativistic nuclear collisions. *Phys. Rev. C*, **65**:034903, 2002.

[168] G. Papp et al. Intrinsic parton transverse momentum in NLO pion production. 2002. arXiv:hep-ph/0212249 (http://arxiv.org/abs/hep-ph/0212249).

[169] B. Back et al. (PHOBOS Collaboration). Pseudorapidity distribution of charged particles in d + Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys.Rev.Lett.*, **93**:082301, 2004.

[170] I. Arsene et al. (BRAHMS Collaboration). Centrality dependence of charged particle pseudorapidity distributions from d+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev. Lett.*, **94**:032301, 2005.

[171] F. Gelis. Theory wrap-up, 2013. Presentation given at 6th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions (Hard Probes).

[172] S. Mishra and F. Sciulli. Deep inelastic lepton-nucleon scattering. *Ann. Rev. Nucl. Part. Sci.*, **39**:259–310, 1989.

[173] E. Bloom et al. High-energy inelastic e p scattering at 6-degrees and 10-degrees. *Phys. Rev. Lett.*, **23**:930–934, 1969.

[174] J. Bjorken and E. Paschos. Inelastic electron-proton and γ-proton scattering and the structure of the nucleon. *Phys. Rev.*, **185**:1975–1982, 1969.

[175] C. Callan and D. Gross. High-energy electroproduction and the constitution of the electric current. *Phys. Rev. Lett.*, **22**:156, 1969.

[176] F. Aaron et al. (H1 and ZEUS Collaborations). Combined measurement and QCD analysis of the inclusive $e^{\pm}$p scattering cross sections at HERA. *JHEP*, **01**:109, 2010.

[177] G. Voss and B. Wiik. The electron proton collider HERA. *Ann. Rev. Nucl. Part. Sci.*, **44**:413–452, 1994.

[178] B. Andersson et al. Parton fragmentation and string dynamics. *Phys. Reports*, **97**:31–145, 1983.