Rivet Implementation of Invariant Yields in Relativistic Heavy Ion Collisions

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Arthur Raber

August 2026

To Mrs. Mary Tiedeman, my high school math teacher, who bullied me into going to college.

Acknowledgments

I would like to thank Dr. Christine Nattrass and my fellow research colleauges, Dr. Grzywacz, Dr. Madurga, the Rivet Developers, the PHENIX collaboration, and the 2024 UTK Physiscs Graduate Cohort. It would have been physically impossible for me to have made it this far or do any of this work without everyone single one of y'all.

Abstract

RIVET's framework provides a systematic approach to the comparison between model simulations and experimental data. This work will finalize analyses from the first five physics runs of the PHENIX experiment at RHIC, which are compared against the model PYTHIA8-Angantyr, with the intention of submitting these analyses to the PHENIX collaboration for confirmation. Confirmed analyses can then, in turn, be used for model tuning and the development of RIVET. In addition, it will develop a more methodical procedure for future students to produce further analyses.

Table of Contents

1 Introduction 1

1.1 Formulation of QGP . . . 2

1.2 Rivet Framework . . . 2

1.2.1 Analyses Definitions . . . 6

1.3 Pythia8-Angantyr . . . 7

1.4 Measurable and Extrapolated Values . . . 9

1.4.1 Centrality . . . 9

1.4.2 Particle Yields . . . 9

1.4.3 Nuclear Modification Factor(s) . . . 11

1.5 PHENIX . . . 12

1.5.1 Measurement Detector Upgrades . . . 14

2 Methodology 21

2.1 Setting Up Docker . . . 21

2.2 Required Files . . . 22

2.3 Primary Analyses Issues . . . 23

2.4 Bin Shift Correction . . . 25

2.5 Divide By Zero Protection . . . 26

2.6 Updates by Analysis . . . 27

2.6.1 PHENIX 2004 I623413 . . . 27

2.6.2 PHENIX 2007 I731133 . . . 28

2.6.3 PHENIX 2008 I778168 . . . 29

3 Results 30

3.1 Results by Analysis . . . 30
3.1.1 PHENIX 2004 I623413 . . . 30
3.1.2 PHENIX 2007 I731133 . . . 31
3.1.3 PHENIX 2008 I778168 . . . 33
3.2 Discussion . . . 34

4 Conclusions 35

Bibliography 36

Appendix 39
A Summary of Equations & Definitions . . . 39
A.1 Pseudorapidity and Rapidity . . . 39
B Docker Start-Up . . . 39
C Code . . . 42
C.1 Bin Shift . . . 42
D PHENIX_2004_I623413 . . . 44
E PHENIX_2007_I731133 . . . 59
F PHENIX_2008_I778168 . . . 63

Vita 66

List of Tables

1.1 Angular and Pseudorapidity Coverage of PHENIX 13

3.1 Particle Masses 34

List of Figures

1.1 Evolution of matter after the Big Bang [15]. 3

1.2 Chemical (T_{ch}) and Thermal Freeze-out (T_{fo}) of hadronic gas. Image snippets courtesy of the MADAI collaboration, Hannah Petersen, and Jonah Bernhard. 3

1.3 Timescale of experimental QGP production [20]. 4

1.4 Workflow and usage of data obtained from the PHENIX collisions. 10

1.5 Schematic of collision centrality for interacting particles in RHIC collisions. . 10

1.6 Diagram of PHENIX Calorimeter and Detector [2]. 15

1.7 Central arms of the Run 1 PHENIX detector, relevant to I623413 16

1.8 Inner detectors of the Run 1 PHENIX detector, relevant to I623413 18

1.9 Central arms of Runs 2 - 5 of the PHENIX detector, relevant to I731133 and I778168 19

1.10 Inner detectors of the Runs 2 - 5 PHENIX detector, relevant to I731133 and I778168 20

2.1 Complete set of files in the RIVETAnalyses repository 24

3.1 Invariant yields of the top 0-5% most central collision of (top to bottom, left to right) $\pi^{+/-}$, $K^{+/-}$, and p & $\bar{p}$ from PHENIX 2004 I623413. 32

3.2 Invarient Yields of π^0 (left) and the ratio of η to π (right) for the Au+Au collision system at min bias ranges of PHENIX 2007 I731133. 33

Chapter 1

Introduction

Fractions of a second after the Big Bang, a state of matter known as the Quark-Gluon Plasma (QGP) existed before cooling under rapid expansion, leading to the hadrons that constitute observable matter today, as shown in Figure 1.1 [1]. Before cooling, quarks existed deconfined as a 'sea' of particles. Intuitively, we can visualize the high-energy collisions used to replicate the environment of the QGP as two ice cubes slamming together with sufficient momentum in outer space such that they momentarily melt upon impact. The rapid expansion results in a quick solidification of the nuclear matter. The properties of the re-frozen Nuclear Matter give us information on the briefly existing liquid.

Chapter 1 introduces characteristics, conditions, and lifespan of QGP, the state of matter that PHENIX set out to study. We break down the formatting of a generic RIVET analysis we used to implement the experimental analysis logic, such as kinematic selections and specific analysis choices. The code is broken down into sections explaining specific choices. The extension from p+p collisions to heavy-ion collisions in the model PYTHIA8-Angantyr is described. The observables used to compare the model to the experiment are defined. Finally, the PHENIX collaboration and the associated detectors used in each experiment that we will be comparing the model against are specified.

1.1 Formulation of QGP

Under immense heat and/or baryon density, hadronic matter undergoes a phase change that is described by quark deconfinement predicted by QCD, where color confinement is lost, and quarks and gluons are no longer bound as hadrons. When this occurs, that hadronic matter is then defined by a different state of matter. The phrase QGP was first coined and defined in 1980:

> *When the energy density exceeds some typical hadronic value ($\sim$1 GeV/fm^3), matter no longer consists of separate hadrons (protons, neutrons, etc.), but of their fundamental constituents, quarks and gluons. Because of the apparent analogy with similar phenomena in atomic physics, we may call this phase of matter the QCD (or quark-gluon) plasma* [21].

The phase transition to QGP is expected to be around a temperature of T $\approx$ 170 MeV $\approx$ 10^{12} K [17]. Figure 1.3 illustrates the flow of states of matter in a collision producing QGP. After QGP reaches an approximate thermal equilibrium, the system undergoes a phase transition to a hadronic gas. From here, cooling continues across two further stages:

- Chemical Freeze-out occurs once the quarks within the hadronic gas stop interacting inelastically with each other to form new hadrons. Here, the particle yield ratios become fixed.
- Thermal Freeze-Out takes place when formed hadrons no longer collide elastically and then travel out toward detectors.

Depending on the produced hadron, it may decay before reaching the detectors. In that case, the primary particle is reconstructed from its decay products. Figure 1.2 depicts the freeze-out stages.

1.2 Rivet Framework

To validate models describing collision systems, rigorous comparisons between experimental data and models must be performed. A systematic approach to comparisons between

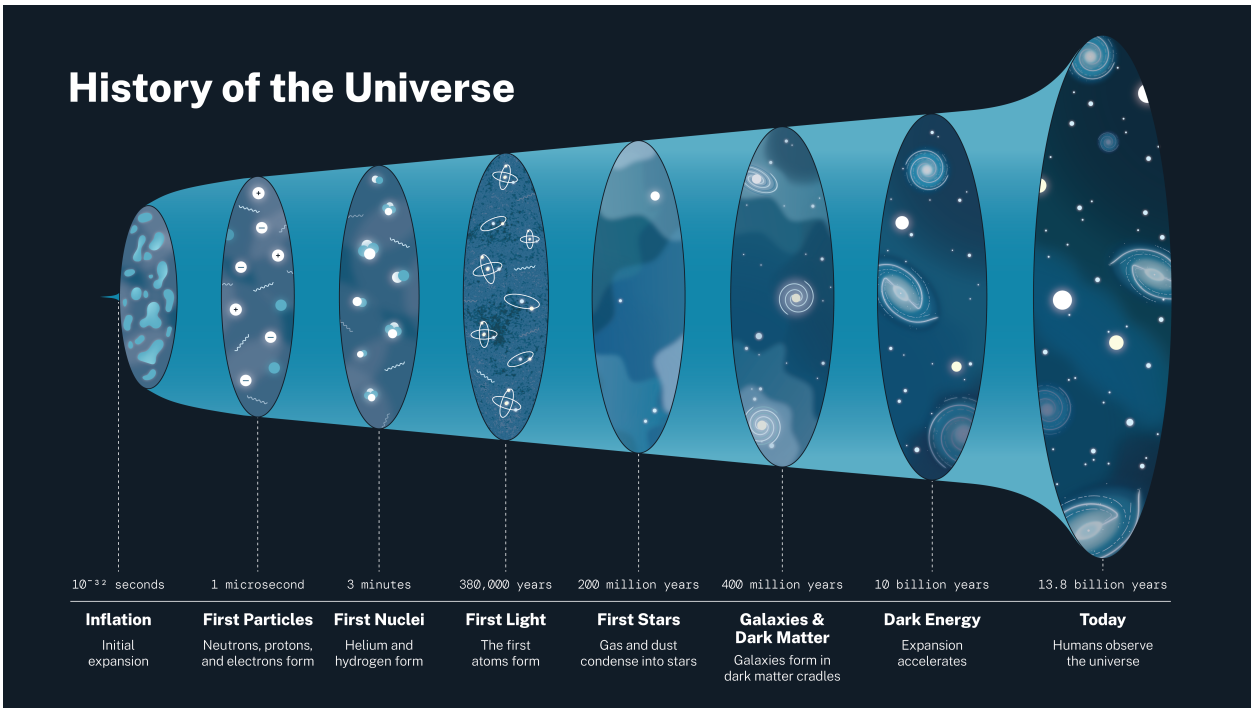


Figure 1.1: Evolution of matter after the Big Bang [15].

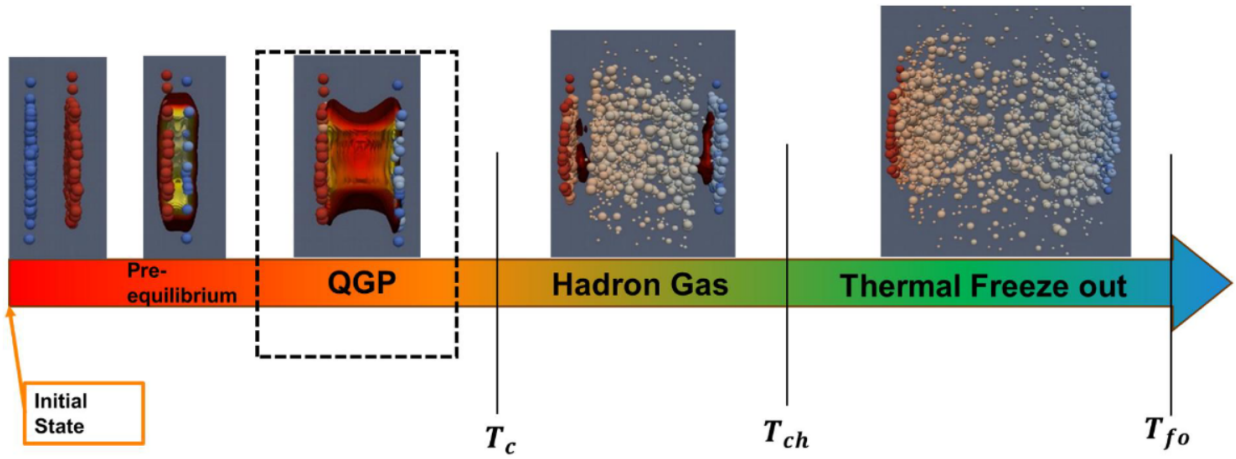


Figure 1.2: Chemical (T_{ch}) and Thermal Freeze-out (T_{fo}) of hadronic gas. Image snippets courtesy of the MADAI collaboration, Hannah Petersen, and Jonah Bernhard.

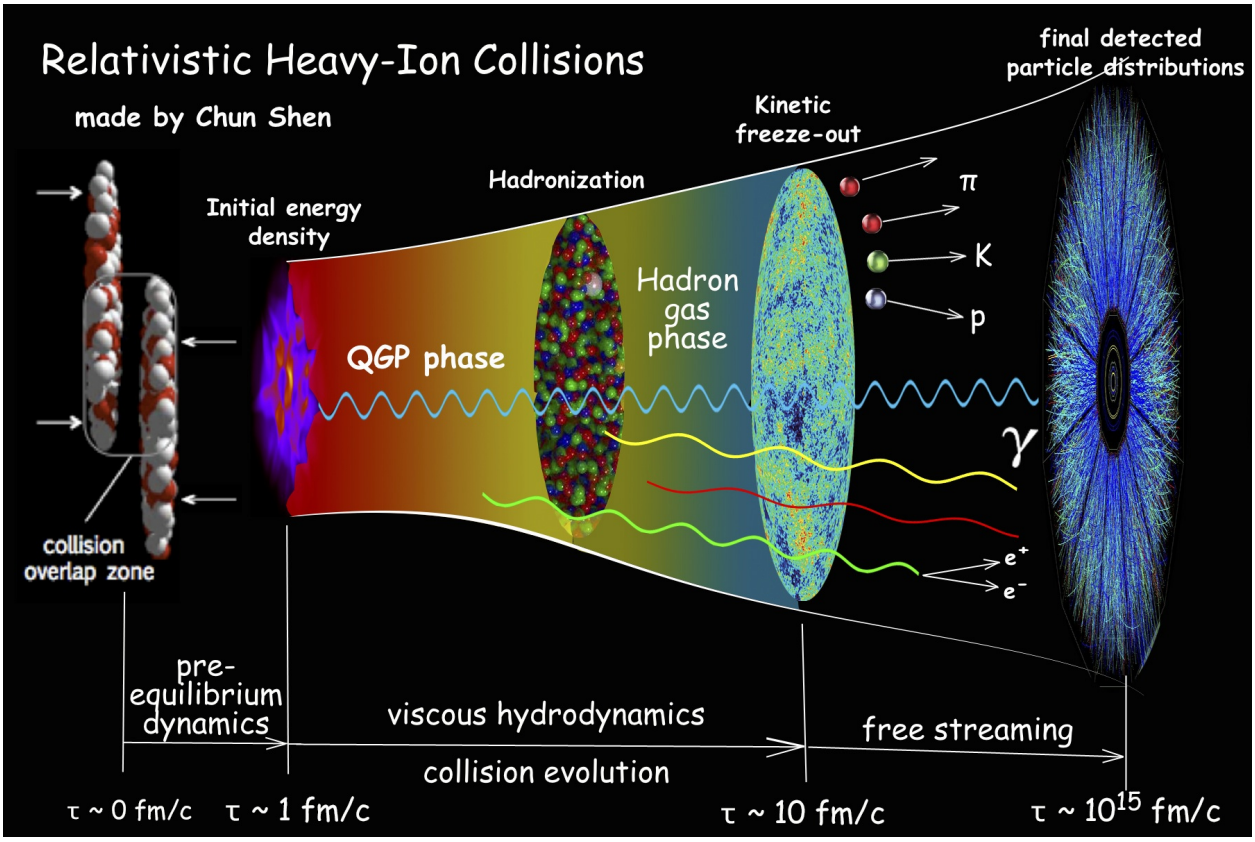


Figure 1.3: Timescale of experimental QGP production [20].

data and models ensures that the model describes the data. Simulated collisions from the most high-energy Monte Carlo event generators produce a standardized output format called HEPMC. The RIVET (Robust Independent Validation of Experiment and Theory) Toolkit takes HEPMC events as input to compare simulated events to experimental data, analyze final state particles, and apply physics analysis routines to reproduce experimental measurements [19]. Figure 1.4 details the flow of data from PHENIX to RIVET, to further work.

A fully written RIVET analysis compares the model to data. Each RIVET analysis consists of three key stages:

- **init ()**: Projections defining the final-state objects (e.g., charged particles, jets) are declared and initialized. The way the experimental data are plotted in HEPData is used to create objects, such as histograms, with binning identical to the experimental data. These will store results from the simulation. This ensures ease of comparison of the simulation to experimental data.

- **analyze ()**: For each generated event, the defined projections are applied to construct particle-level observables. Selected observables are used to fill the corresponding histograms. Events are categorized into centrality bins of a given collision, where the (5-15)% bin corresponds to the *most* central collisions. The (60-82)% bin corresponds to the least central collisions.

- **finalize ()**: Operations that are only applied once to a given data set are performed here. This often consists of normalizing the accumulated histograms and finalizing them for direct comparison with experimental data.

Analyses are written using specific data analysis techniques to ensure accurate representation of the data and physics; these techniques will be discussed further in Chapter 2.

The flow chart in Figure 1.4 contains projects that are utilized by Rivet, and that utilize Rivet. HEPData is the 'High-Energy Physics Database' that functions as an open-access repository for collision scattering experiments [22]. The reference binning and experimental data used are stored here. The output files from Rivet are also extremely useful for model

tuning. Previously approved analyses are currently being used as a reference for training the Gaussian process emulator within Bayesian parameter estimations within the group. This furthers the importance of continuing to update, produce, and fix these analyses.

1.2.1 Analyses Definitions

A RIVET analysis applies the analysis logic used in the data so that a single code can make comparisons between data and any Monte Carlo model that produces a HEPMC output. Projections and cuts are the first initialization of this process, which are defined as follows:

- Projections are typically used to define the particles that we want to keep from the simulation. This consists of selecting all stable final-state particles, charged particles, specific particles, or excluding daughters from long-lived particles relevant to the experiment. For example, the reported proton yields from PHENIX do not include protons produced from weak decays, such as $\Lambda \rightarrow p\ \pi$. The STAR collaboration includes these protons from weak decays and reports the total proton yield [4] [7]. Alternatively, specific particles can be selected for identification by using their 'particle identification number'.

- Cuts are used to replicate detector positions azimuthally and for given rapidity ranges. This narrows down the analysis to only record the particles that would have been measured by the PHENIX detector.

Following hadronization, not all of the particles produced are stable. To obtain a consistent definition of final-state observables, a distinction is made between primary particles and those originating from secondary decays. The ALICE definition of a primary particle is used in the PHENIX analyses. 'A primary particle is a particle with a mean proper lifetime larger than 1 cm/c,' produced either directly from the interactions or from produced particles with a mean proper lifetime shorter than 1 cm/c [13]. PHENIX uses the same primary particle definition.

1.3 Pythia8-Angantyr

The physics model PYTHIA is best described by its developers:

> *PYTHIA 8.3 is a scientific code library that is widely used for the generation of events in high-energy collisions between particles, where effects of the strong nuclear force, governed by Quantum Chromodynamics (QCD), are of high importance. It is written mainly in C++ and interweaves a comprehensive set of detailed physics models for the evolution from a few-body hard-scattering process to a complex multi-particle final state* [11].

A heavy-ion extension to the PYTHIA8 package is the Angantyr model, which is used for extending the dynamics of proton-proton (pp) collisions to collisions between two nuclei (AA) and proton-Nucleus (pA), by overlaying nuclear models onto p+p collisions [12]. Angantyr uses a minimal number of free parameters. Our reasoning for using PYTHIA8-Angantyr stems from its ease of use and extremely fast simulation speeds. While not perfect, the need for higher statistics, on the order of 10^6 events in some cases, outweighs the need for perfection in the model. Further, the output YODA and HEPMC files produced allow for easy comparison to data and simulation. This makes for effective C++ RIVET code development, which is the focus of this work.

This extrapolation uses several scattering characteristics of high-energy p+p collisions:

- At high energies, p+p collisions are dominated by inelastic scattering, but elastic and diffractive processes still arise from fluctuations in the proton's internal structure.
- Further, p+p collisions require accounting for multi-parton interactions in non-diffractive events. Diffractive events may include hard scatterings, which can be described using the Ingelman–Schlein model of hard diffraction.

Omitting these would misrepresent particle production and fluctuations.

Several considerations are important for this generalization to be applicable:

1. **Nucleon Position inside Nuclei:** The spatial distribution of the nucleons inside each nucleus must be considered. Monte Carlo generators are typically used to sample

nucleon positions according to nuclear density distributions (such as Woods–Saxon profiles). This step determines the initial geometry of the collision, which strongly influences observables such as collision centrality and particle multiplicity.

2. **Calculating Nucleon-Nucleon Collisions (Glauber Formalism):** The nucleons that do interact must be determined. The Glauber model uses the eikonal approximation, assuming that they travel along straight lines until the collision.

   In Angantyr, once the interacting nucleons are determined, nucleon–nucleon collisions are mapped onto PYTHIA events. The state of the nucleon target and projectile is fixed at the moment of interaction.

   Fluctuations in the projectile nucleon (proton) in the generalized p+A collisions estimate the number of Nucleon-Nucleon (N+N) collisions that occur. Angantyr averages over the target nucleon states to calculate the cross section and number of wounded nucleons in the p+A collisions. The A+A collisions consider the superposition of nucleon–nucleon interactions of both the target and projectile.

3. **Interaction Contributions:** With the interacting nucleons identified, each interaction contribution is considered for the final-state particle production.

   - **Soft Particle Production from Wounded Nucleons:** For p+A and A+A collisions, particle production is determined by the number of wounded (participating) nucleons.
   - **Hard Process and Multiple Nucleon Collisions:** The Hard partonic subcollisions scale with the number of NN sub-collisions. For p+A and A+A collisions, cascading N+N collisions are included using PYTHIA8's multi-parton interaction system to get to A+A collision final states.

1.4 Measurable and Extrapolated Values

Implementing a Rivet analysis allows experimental-style definitions of observables to be reproduced through binned histograms, enabling direct comparison with experimental measurements.

1.4.1 Centrality

Centrality characterizes the geometry of a heavy-ion collision and is closely related to the impact parameter 'b', the transverse distance between the centers of the colliding nuclei. Small impact parameters correspond, on average, to a larger overlap region and therefore to more participating nucleons, while large impact parameters produce more peripheral collisions. Figure 1.5 gives a physical representation of the geometry.

Since the impact parameter cannot be measured directly, experiments determine centrality using the charged-particle multiplicity. Events are ordered and divided into percentiles of the total inelastic cross section, with the highest multiplicity events corresponding to the most central collisions and the lowest multiplicity events corresponding to the most peripheral collisions.

Rivet reproduces this procedure through its heavy-ion centrality framework. A calibration distribution is first constructed from the charged-particle multiplicity within the same detector acceptance used by the experiment. The generated events are then assigned centrality percentiles from this calibrated distribution, ensuring that the model and experimental data are classified according to the same centrality definition.

1.4.2 Particle Yields

The particle yields (N) define the count of produced particles as a dimensionless quantity per transverse momentum (p_T) and pseudorapidity (η) bin. The transverse momentum refers to particle's component of movement perpendicular to the beamline. Pseudorapidity (and rapidity) are defined in Appendix A.1. The differential yield is a normalized density of

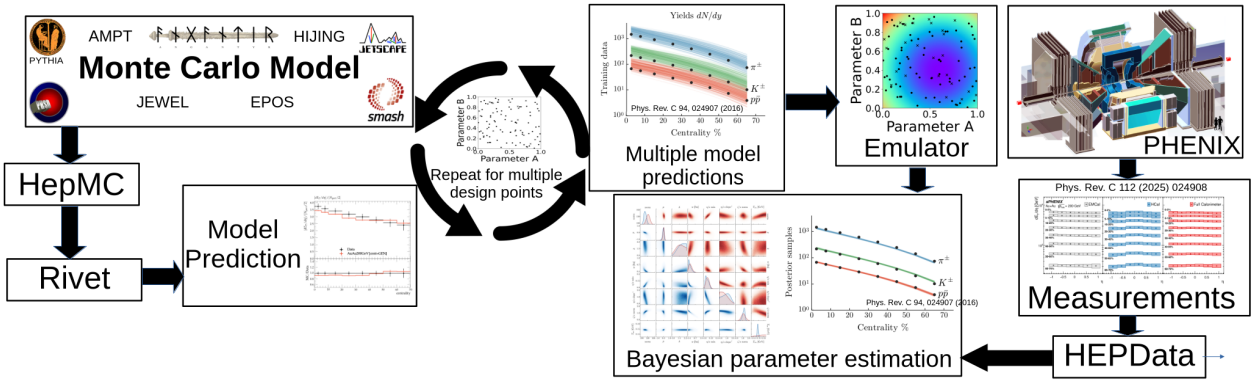


Figure 1.4: Workflow and usage of data obtained from the PHENIX collisions.

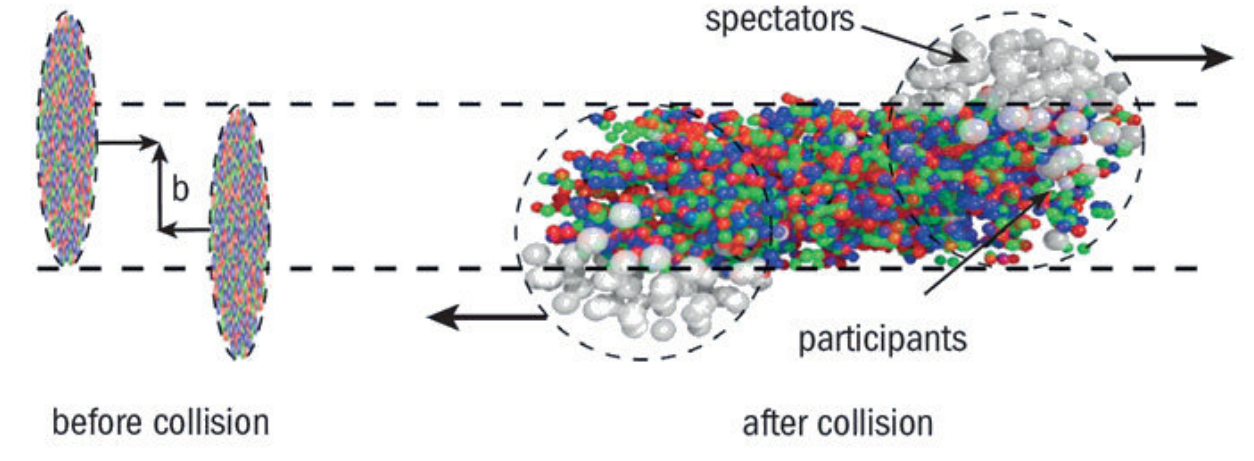


Figure 1.5: Schematic of collision centrality for interacting particles in RHIC collisions.

particles per event per phase space of the kinematic variables. It is defined as:

$$\left[\frac{1}{N_{event}}\frac{d^2N}{dp_T\ d\eta}\right]_i \approx \frac{N^{(i)}_{particles}}{N_{event}\Delta p_T^{(i)}\Delta\eta^{(i)}} \tag{1.1}$$

where definitions for each variable are given here: The 'i' index refers to a specific p_T and η bin. N_{event} refers to the number of analyzed events of a given collision system in that centrality class. $N^{(i)}_{particles}$ is the number of particles produced, summed over all events. The centrality class is imposed through the event selection of N_{event} and $N^{(i)}_{particles}$. The data is stored in histograms produced within RIVET. The integrated yield is then defined by:

$$< N >= \int \frac{1}{N_{event}}\frac{d^2N}{dp_T\ d\eta}dp_T\ d\eta \tag{1.2}$$

1.4.3 Nuclear Modification Factor(s)

The Nuclear Modification Factor describes how nuclear effects modify particle production. For two particles, A_1 and A_2, per p_T and η bin, we define it as follows:

$$R_{A_1A_2} = \frac{B}{C < T_{A_1A_2} >} \tag{1.3}$$

$$B = \frac{1}{N^{A_1A_2}_{event}}\frac{d^2N^{A_1A_2}}{dp_T\ d\eta}, \quad C = \frac{d^2\sigma_{pp}}{dp_T\ d\eta}, \quad < T_{A_1A_2} >= \frac{< N_{coll} >}{\sigma_{pp}} \tag{1.4}$$

Note that B is exactly the differential yield from 1.1. In the denominator, C is the differential cross section of the p-p baseline data, taken under the same conditions. $< T_{A_1A_2} >$ is defined as the Nuclear Thickness factor. In a given collision, each nucleus is capable of interacting with multiple nucleons of the incoming particle, vice versa. This scales the p-p baseline based on the expected number of n-n collisions.

Values for B and C are stored as histograms for a given centrality. In this work, we use the nuclear thickness used by the experiment. Since the PHENIX experiment calculated T_{AA} in a particular implementation of a Glauber model variation, using the same T_{AA} as the collaboration removes a model-dependent factor in the comparison. If the pre-existing

constant is not used, T_{AA} can be calculated using the number of binary nucleon collisions from the simulation.

Values of $R_{A_1A_2}$ have 3 interpretations:

- $R_{A_1A_2} = 1$ means the collisions behave like an independent superposition of p-p collisions.
- $R_{A_1A_2} < 1$ is defined as suppression, where fewer particles are produced.
- $R_{A_1A_2} > 1$ enhancement of particle production is expected.

On occasion, the Ratio of Central to Peripheral (RCP) of p-A or A-A collisions was used to approximate $R_{A_1A_2}$. This was sometimes done when p-p collision data were not available, but also done due to cancellation of some uncertainties.

1.5 PHENIX

The PHENIX Collaboration coalesced from several existing groups to measure rare processes in hadronic, leptonic, and photonic channels [1]. Following installation at RHIC, data from various collision types were collected with the intent to address the following key physics goals: [23]:

- To confirm that QGP is best defined as a thermalized quark-gluon plasma and to quantify both the equilibrium and non-equilibrium features of the produced medium.
- To investigate the gluon structure of the proton, including the spin contributions from gluons and sea quarks, via polarized proton–proton collisions.
- To measure the gluon distribution in cold nuclear matter using proton-nucleus collisions.

The initial PHENIX proposal consisted of three main detector 'arms' [18]. Figure 1.6 shows a to-scale model of the detector geometry. Table 1.1 summarizes the angular and pseudorapidity coverage.

Table 1.1: Angular and Pseudorapidity Coverage of PHENIX

Element	θ (deg)	η	ϕ (deg)	$\Delta\Omega$ (sr)
Central Arms	70-110	-0.35 $< \eta <$ 0.35	$\pm$ (22.5 $< \phi <$ 112.5)	1.07
Forward Arm	10-35	1.15 $< \eta <$ 2.44	0-360	1.17
Inner Detectors	2-178	-4.0 $< \eta <$ -4.0	0-360	$\sim 4\pi$

The central arms are responsible for the measurement and tracking of charged hadrons, electrons, and photons. The forward arms are dedicated to muon detection. The inner detectors measure the start time, vertex, and multiplicity of an interaction.

Initially, PHENIX only included an outer coil around the collision point. The inner coil and Dalitz-tuned detectors were later installed to detect π^0 decay channels. Initial tracking detectors were installed at r$\approx$2m and r$\approx$4.5m.

Particle Identification duties were split into several categories due to the wide range of momenta from electrons and hadrons, ranging from $0.2 - 5.0\ GeV/c$ or greater:

- **Time-of-Flight (TOF):** High-resolution scintillators take up one-third of the central arm for pion and kaon distinction. The PbSc EMCal time resolutions also aided in electron and hadron discrimination.

- **Cherenkov Counter:** Ring-imaging Cherenkov (RICH) counters are positioned between the inner & outer trackers to provide electron identification via Cherenkov radiation.

- **Time Expansion Chamber (TEC):** Measures track positions and energy deposition along a given track.

The inner detector subsystem measures the start time and vertex location of an event, as well as event multiplicity. Two Cherenkov beam-beam counters (BBC) cover a pseudorapidity interval of $3.1 \leq |\eta| \leq 4$, providing a vertex location of $\sigma_z \approx 2cm$ and a start time of $\sigma_t < 100ps$. A silicon detector (MVD) array surrounds the collision point, covering a pseudorapidity range of $-2.7 \leq \eta \leq 2.7$ at $z = 0$, covering all of the azimuthal detection range. The MVD measures total multiplicity, $d^2N/d\eta d\phi$, and provides a vertex location of $\sigma_z < 500\mu m$.

The charged particle tracking chambers consisted of three pad chambers and two drift chambers.

1.5.1 Measurement Detector Upgrades

This work contains data taken during various revisions of the PHENIX detector. These revisions are necessary to ensure proper replication of simulated collisions. Specific updates of the detectors from various runs are detailed chronologically:

1. **PHENIX 2004 I623413**: A cross-sectional view of the PHENIX detectors used in the $\sqrt{S_{NN}} = 130$ GeV run is shown in figure 1.7 [9]. This analysis uses the drift chamber (DC), pad chamber (PC), and the TOF of the East Arm for pion, kaon, and proton measurements. The $70 - 110°$ polar angle θ and pseudorapitity range of $|\eta| < .35$ are utilized here.

   The drift chambers are the first interaction point for tracking after a collision. They are responsible for measuring the momentum and direction vector of charged particles. In both (East and West) arms, they run 1.8m parallel to the beam, with 90° of coverage, centered at 2.2m from the beam line.

   Three pad chambers are utilized with pixel detectors with an effective readout size of 8.45mm parallel to the beamline and 8.40mm perpendicular to the beam. On either arm, pad chamber one (PC1) is positioned on the outer radial edge of the respective DC, 2.49m from the beam line. PC3 is positioned only in the East Arm, 4.98m from the beamline.

   Along the beamline, the timing and event trigger system is as follows: The BBC covers a forward rapidity angle of $3.0 \leq \eta \leq 3.9$ with full azimuthal coverage, and is positioned north and south of the MVD on either side of the collision chamber. Zero-Degree Calorimeters are positioned 18m north and south of the collision center and aid in event triggering. Figure 1.8 shows a side view of the detectors along the beamline.

2. **PHENIX 2007 I731133**: The cross section and side views of the PHENIX detectors for Runs 2 and 3 can be seen in Figures 1.9 and 1.10, respectively [10]. The data

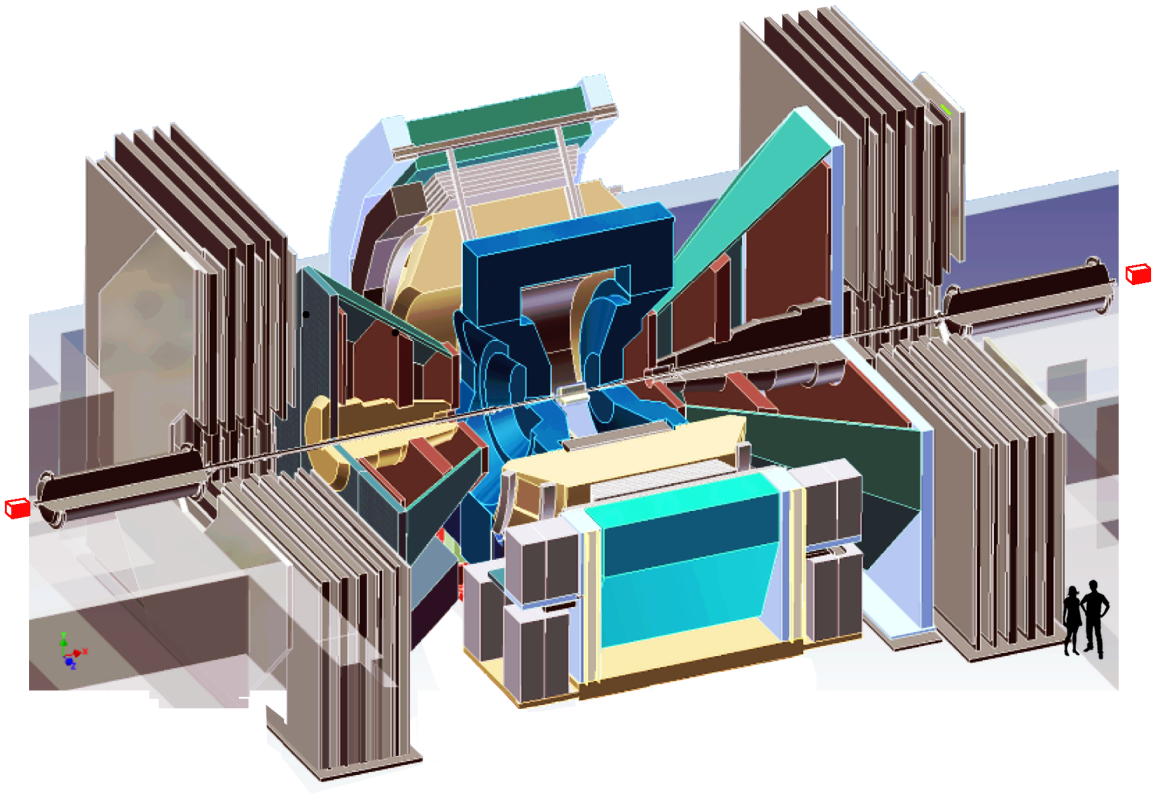


Figure 1.6: Diagram of PHENIX Calorimeter and Detector [2].

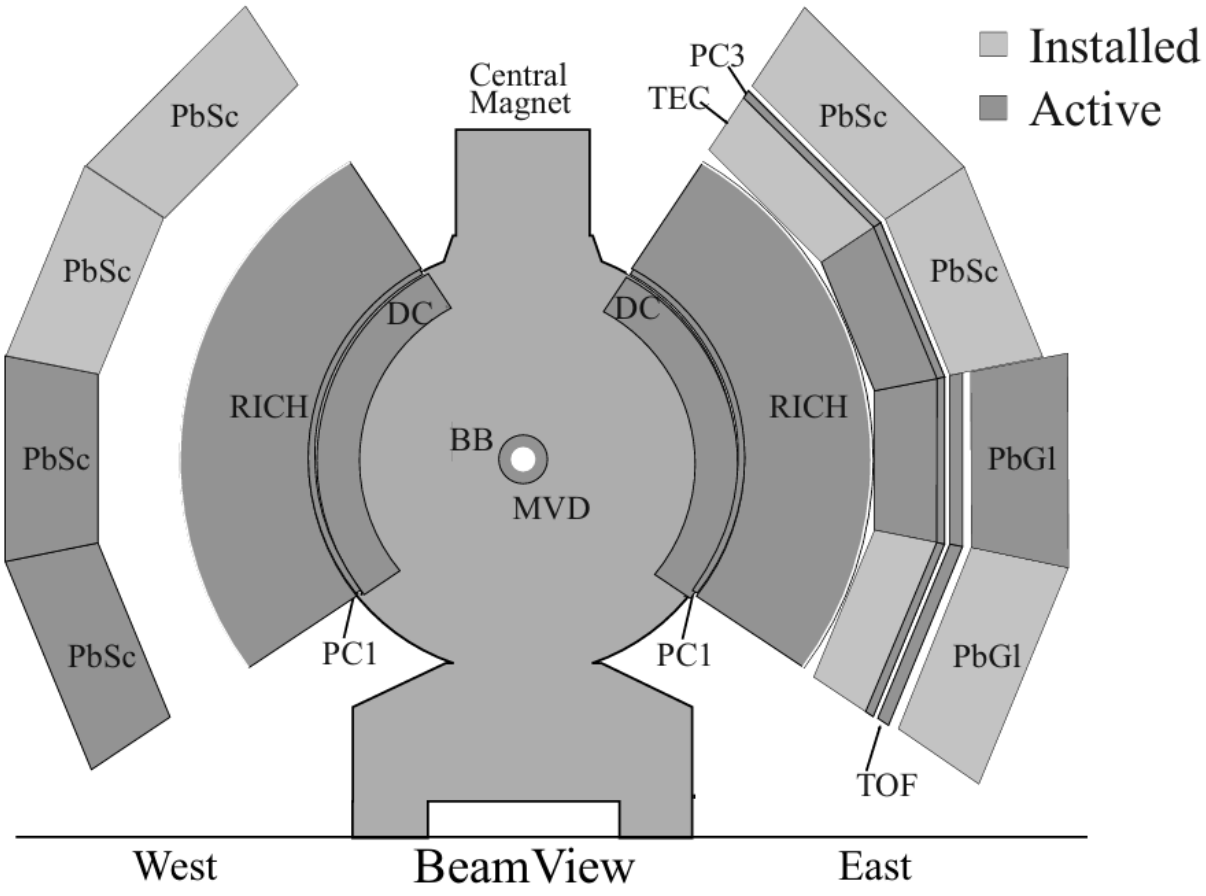


Figure 1.7: Central arms of the Run 1 PHENIX detector, relevant to I623413

compared in this work were obtained during Run 2 (2001–2002, Au+Au) and Run 3 (2003, d+Au, p+p) at $\sqrt{S_{NN}} = 200$ GeV.

The Drift Chamber and Pad Chambers are utilized for tracking, with the addition of the electromagnetic calorimeters (EMCal) (Pbxx), which were not active in Run 1. The EMCal covers a total pseudorapidity range of $\Delta\eta \approx 0.7$ and approximately 180° azimuthally.

The BBC and ZDC are used for triggering and event characterization. Together, they are used in the centrality measurements of Au+Au and d+Au collisions.

3. **PHENIX 2008 I778168**: The data compared in this work were obtained during Run 4 (Au+Au) and Run 5 (p+p) at $\sqrt{S_{NN}} = 200$ GeV [6]. Here the BBC is also used for measuring charged particle multiplicity. Together, the BBC and ZDC were used in a similar fashion as in Runs 3 and 4. The EMCal (which saw no upgrades) was used for reconstructing decayed particles. Further specifications of the detector from Run 2 forward are detailed elsewhere [8].

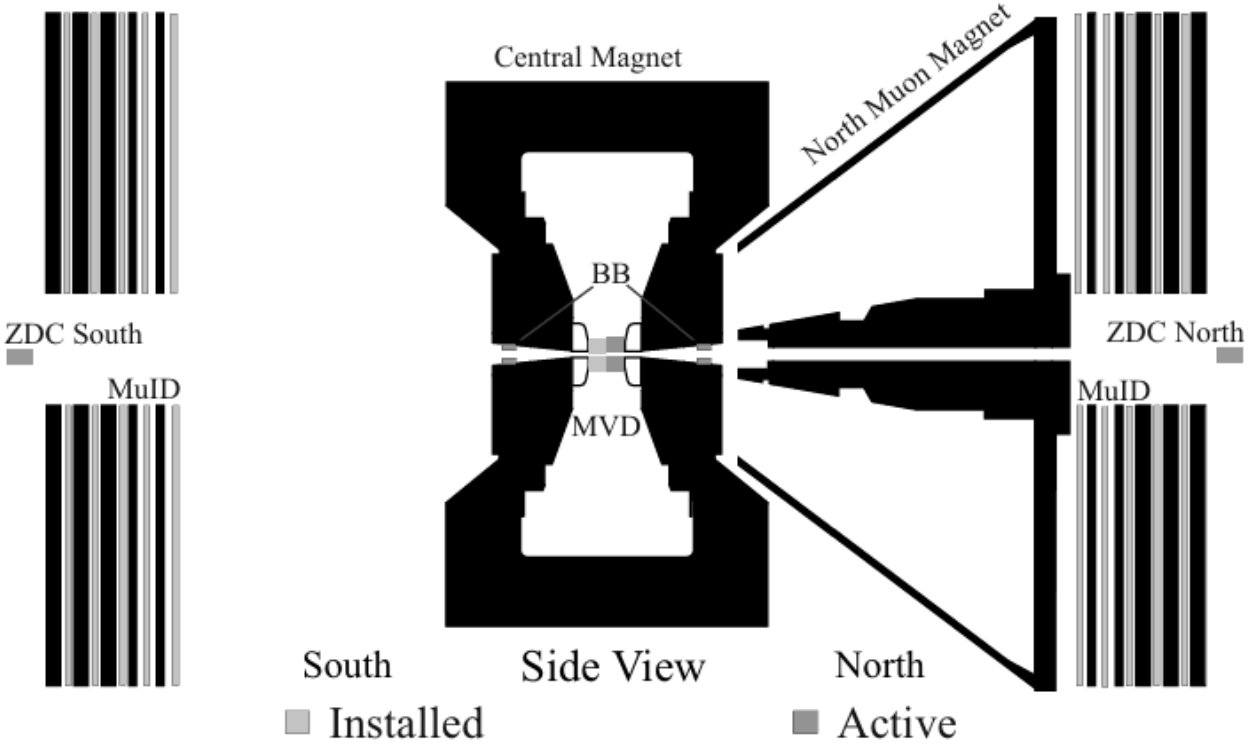


Figure 1.8: Inner detectors of the Run 1 PHENIX detector, relevant to I623413

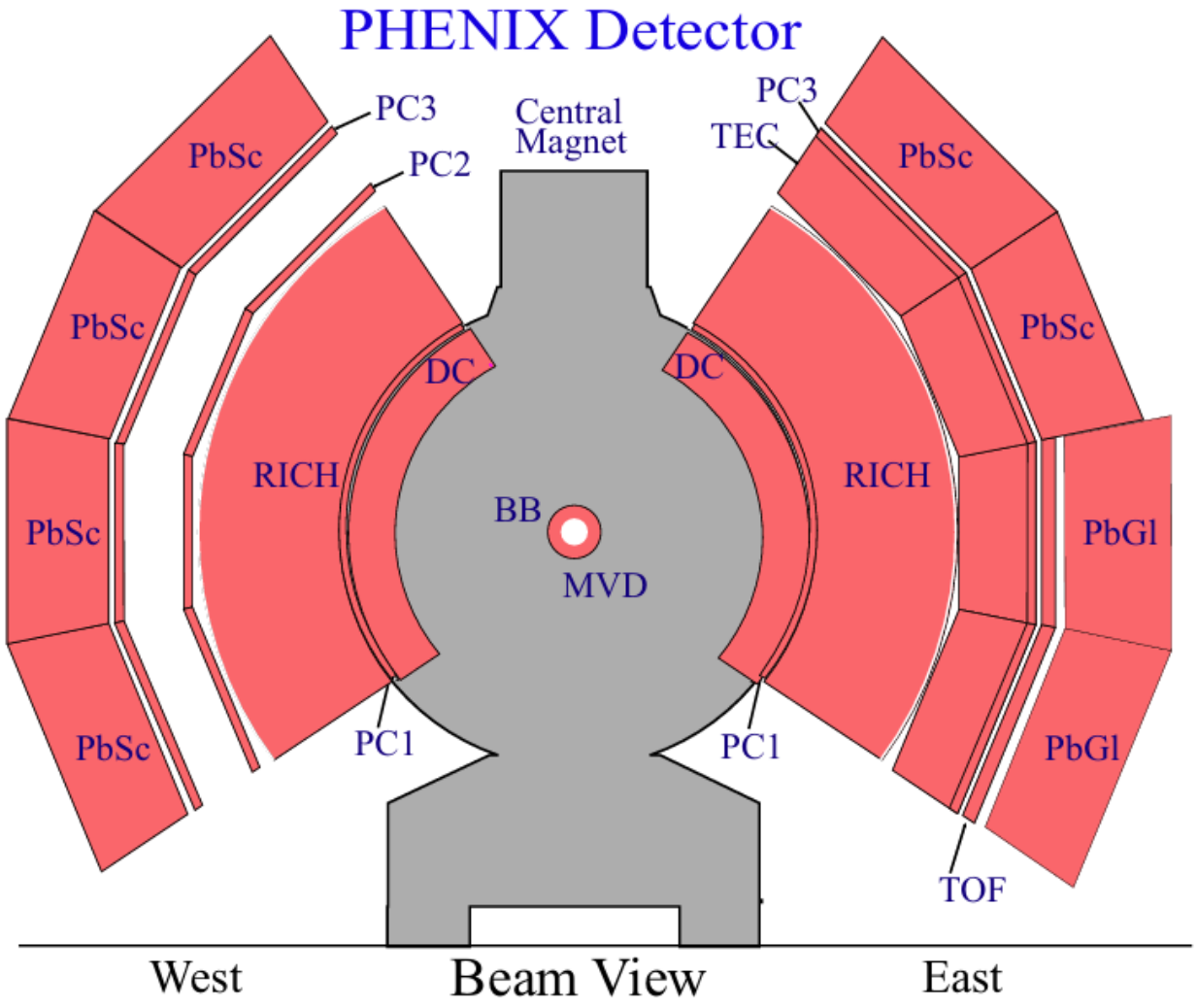


Figure 1.9: Central arms of Runs 2 - 5 of the PHENIX detector, relevant to I731133 and I778168

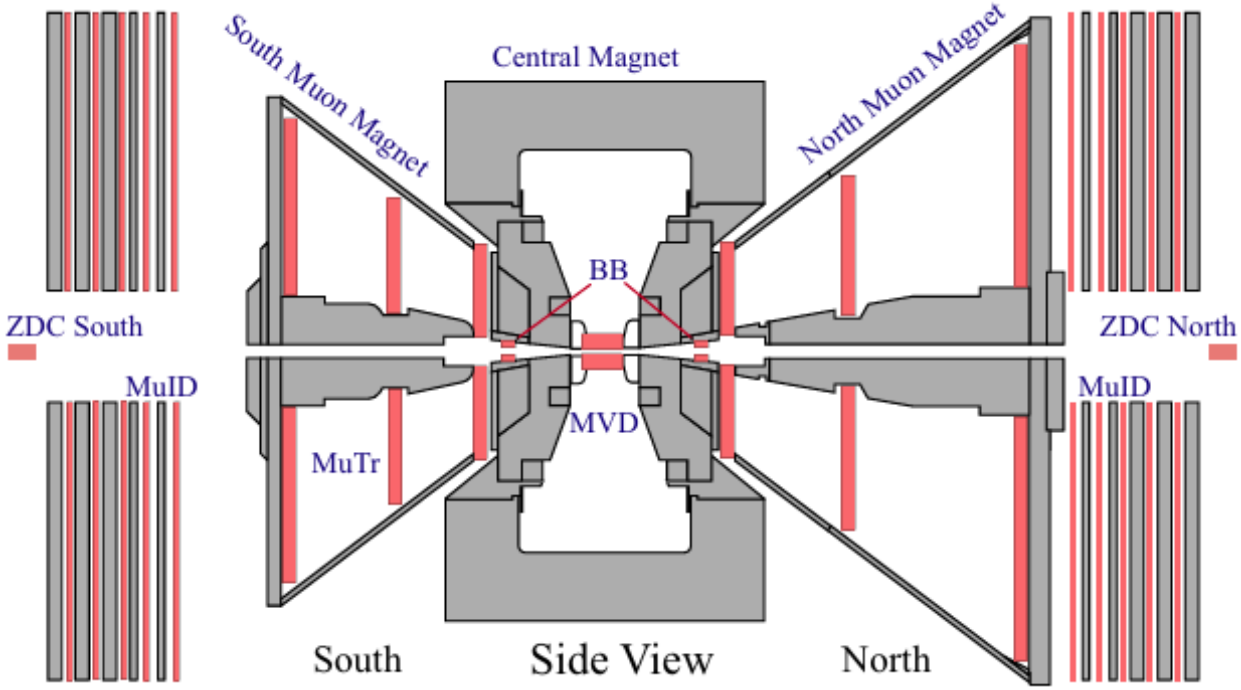


Figure 1.10: Inner detectors of the Runs 2 - 5 PHENIX detector, relevant to I731133 and I778168

Chapter 2

Methodology

The systematic approach to producing general analyses was described in Section 1.4. This chapter aims to describe improvements to the process by which we write an analysis in this thesis. It will follow a chronological path of the work done in this thesis. The utilized operating software is described. The process of setting up the operating environment using Docker is included. The necessary files for each analysis in the GitHub Repo are discussed. The functionality issues due to the update from Rivet 3 to Rivet 4 are explained. To ensure the same methodology is used in the Rivet analysis as in the data, a bin shift is applied to each analysis. Divide by zero (DB0) protections are implemented for all analyses. The final section will cover specific fixes to the existing analyses that got them to the point of being ready for approval by the collaboration. [1]

2.1 Setting Up Docker

Docker Desktop allows for the usage of closed environments called containers [14]. These containers can interact with local files on the host device while operating with an independent operating system. It is used to store an up-to-date version of Rivet and PYTHIA8-Angantyr. The local files are attached to the official GitHub repository. All of the experimental data, required scripts, and calibration data are stored in that repository.

[1] Please note that the analyses in this work were started by other students, but left unfinished and unverified at higher statistics. Those who came before more will be cited as such.

A complete set of steps for setting up Docker, the container, and gaining access to the GitHub repository will be provided in the appendix section B.

2.2 Required Files

Each complete analysis contains a set of files required for proper function. They will be named and detailed here. The naming convention follows 'Collaboration_Year_Inspire ID' for several of the files. Figure 2.1 shows the files for the first analysis in this work inside the repository.

The '.cc' file contains all of the information discussed in sections 1.2 and 1.4. The physics we want to replicate and capture are defined here.

The '.info' file defines all of the information associated with the experiment, including publication details, analysis authors, status of analysis validation, and the RunInfo. The RunInfo contains general details about the experiment and is checked before the analysis is run by Rivet to ensure the correct collision system is enabled in both the '.cc' and '.info' files.

The '.plot' file sets up the axis labels, applied scaling, and titles of the plotted histograms. The simulation and reference data for each measurement are plotted together.

The '.yoda' and '.yoda.gz' are the same file, where the latter is a compressed version. They both contain the experimental data from the experiment as stored on HEPData.

The 'README.md' is a living document that briefly tells the user what data the analysis handles, the status of the analysis, and the author history.

The 'RunAnalysis.sh' and 'RunFIFOAuAu130.sh' are the two executable scripts used for testing and running the analysis, respectively:

- 'RunAnalysis.sh' purely runs Rivet over a set of dummy data to ensure that the '.cc' will compile. No simulation occurs here. It only compares the dummy data to experimental data. After running, it will produce a log of errors in the terminal (if applicable), a '.so' file, and a Rivet.yoda file. The '.so' file is a compiled machine code from our '.cc' file. The 'Rivet.yoda' file is the output that contains dummy data filled as defined in the '.cc' file.

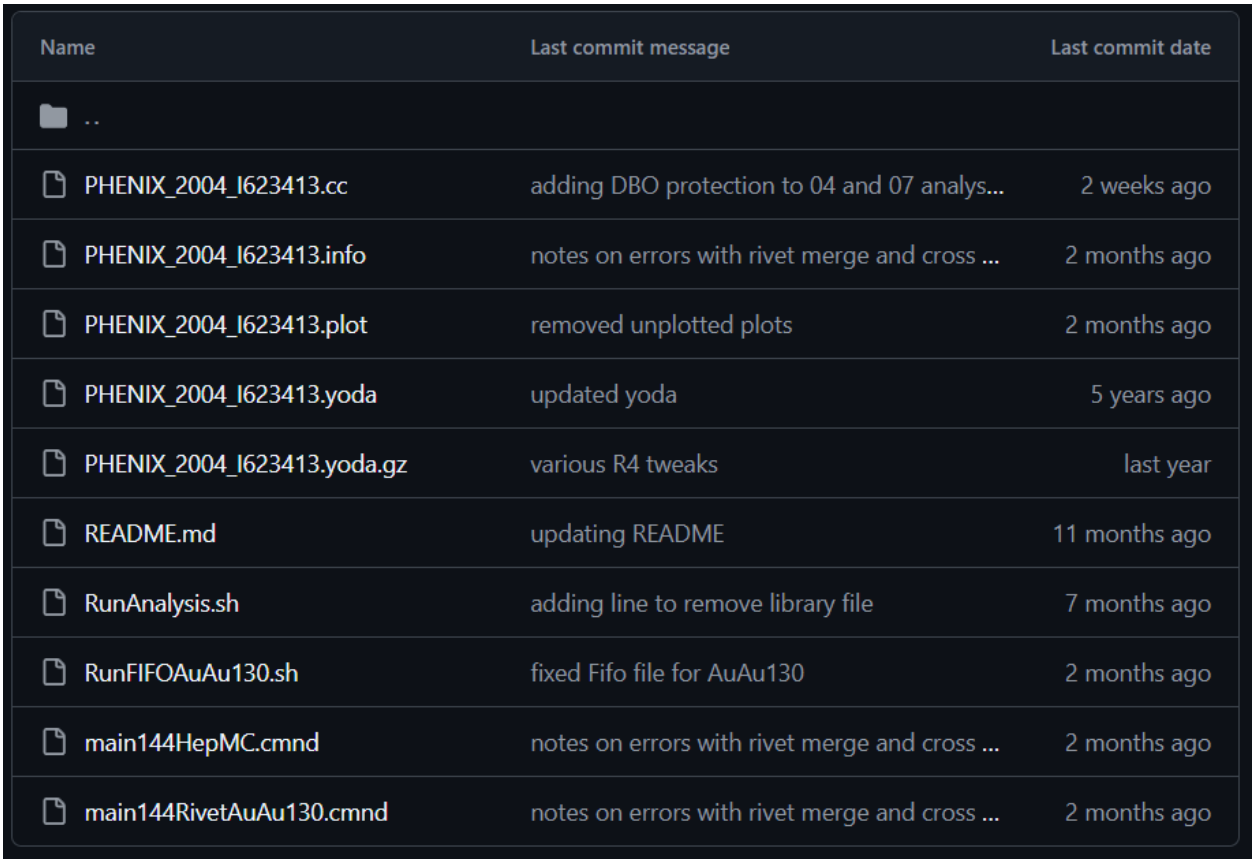
- 'RunFIFOAuAu130.sh' runs Pythia simulations to test the model. It runs with a 'first in, first out' logic, such that the generator (for the models) immediately sends the test data to Rivet for analysis to avoid bloated data files. After running, the compiling errors are displayed in the terminal, and the same '.so' is produced. The '.HEPMC' file produced details of the raw data from the generator. An 'AuAu130GeV.yoda (or similar for different systems and energies) file is produced instead of the 'Rivet.yoda' that contains the same collision information from the generator, but the data follows the binning as defined in the '.cc.'

The 'main144HepMC.cmnd' and 'main144RivetAuAu130.cmnd' are the command files associated with the analysis. These configuration files are used by PYTHIA and RIVET to determine things such as which Rivet analyses are run, how many events are simulated, and what collision system and energy are simulated. The former is only responsible for setting the number of collisions to run when using FIFO mode. The latter defines which analyses to run and which file to use for centrality calibration.

2.3 Primary Analyses Issues

Data from a high-energy experiment is only posted to HEPData after paper publication. The Rivet analysis is then written that experiment. This is normally done consecutively. The PHENIX experiments finished in 2016, and the HEPData entries for PHENIX did not begin until 2017. Review and approval of entries only started in 2020. As a result, the Rivet analyses and HEPData entries were not written simultaneously. To get the HEPData entries approved, they were often changed such that the entry notation did not always match the notation used in the analysis before its (HEPData's) approval. This required updates to incorrect notation in each analysis and sometimes to errors in the HEPData entry. This was done in the init() section of the analysis referred to as 'booking.'

The upgrade from Rivet3 to Rivet4 implemented changes that 'broke' existing analyses that were not approved yet [3]. With each update, any analysis that is approved by the collaboration and added to the official Rivet repo (not the repo referenced in Appendix B) is updated by the Rivet developers. The update saw the change from Yoda1 to Yoda2 that



Name	Last commit message	Last commit date
..		
PHENIX_2004_I623413.cc	adding DBO protection to 04 and 07 analys...	2 weeks ago
PHENIX_2004_I623413.info	notes on errors with rivet merge and cross ...	2 months ago
PHENIX_2004_I623413.plot	removed unplotted plots	2 months ago
PHENIX_2004_I623413.yoda	updated yoda	5 years ago
PHENIX_2004_I623413.yoda.gz	various R4 tweaks	last year
README.md	updating README	11 months ago
RunAnalysis.sh	adding line to remove library file	7 months ago
RunFIFOAuAu130.sh	fixed Fifo file for AuAu130	2 months ago
main144HepMC.cmnd	notes on errors with rivet merge and cross ...	2 months ago
main144RivetAuAu130.cmnd	notes on errors with rivet merge and cross ...	2 months ago

Figure 2.1: Complete set of files in the RIVETAnalyses repository

changed how the data with histograms was stored and required updates to the reference files for all analyses [2]. Several other naming conventions were changed regarding how histogram types were produced and stored data[3].

A collision system check was implemented to ensure that the end user does not attempt to run a beam type not present in the analysis. If the beam type is not specified, the user may manually input the collision system (AuAu130, pp200, eg.) and the collision centrality via the command line. This is not required for approval by the collaboration, as required by RIVET's formulation (or due to the change to Rivet4), but is done as an additional feature in the analysis for future use.

[2]More information on the changes from Yoda1 to Yoda2 can be found here: https://arxiv.org/abs/2312.15070

[3]Many of these updates were handled in these analyses before my interaction with them.

2.4 Bin Shift Correction

The PHENIX collaboration chose to report the value of the spectrum at the center of the bin. We therefore have to correct the weighted average across the bin, where the spectrum is roughly exponentially falling with p_T. We do this using a local approximation of the spectrum as an exponential[4].

Once the Monte Carlo generator has processed all events and filled the corrected bins of the associated histogram, the values are stored as the average density of data points within that bin. This is stored at the center of the bin. However, if the data used contained more points at the low end of the bin, the value placed in the center of the bin will be higher than the actual value if the data points were evenly distributed. For a given bin, i, the stored bin average density is:

$$\overline{f_i} = \frac{1}{\Delta p} \int_{p_{low}}^{p_{high}} f(p)\ dp. \tag{2.1}$$

The variables p_{high} and p_{low} refer to the upper and lower edges of a given bin, respectively. Δp is the difference between those two points. We then assume that the local distribution of the spectrum is exponential, such that $f(p) = Ae^{-bp}$ where p refers to a specific value on the x-axis. Equation 2.1 then becomes:

$$\overline{f_i} = \frac{A}{b\Delta p}(e^{-b\cdot p_{low}} - e^{-b\cdot p_{high}}). \tag{2.2}$$

We want to consider the point at the center of each bin, which we will define as:

$$p_c = \frac{p_{low} - p_{high}}{2}, \tag{2.3}$$

where we will replace p with p_c where $f(p) \rightarrow f(p_c)$. We want to define a correction factor that shifts this value at the center of the bin *vertically*, that also follows the assumed local distribution, such that $f(p_c) = f_{corr} \cdot \overline{f_i}$. Our bin shift correction factor then becomes:

$$f_{corr} = \frac{f(p_c)}{\overline{f_i}} = \frac{-b \cdot (p_{high} - p_{low}) \cdot e^{-b\cdot(p_{high}+p_{low})/2}}{e^{(-b\cdot p_{high})} - e^{(-b\cdot p_{low})}}. \tag{2.4}$$

[4]The collaboration did this because they were aiming to make the results clear and comparable to theories.

To estimate b, we continue to assume our local function, $f(p)$. Thus the bin contents scale as $f(p_i) \propto e^{-b \cdot p_i}$. The ratio between the neighboring bin values yields:

$$\frac{f_{i-1}}{f_{i+1}} = e^{-b(p_{i-1}-p_{i+1})}. \tag{2.5}$$

Taking the natural log of both sides and solving for b gives:

$$b = \frac{1}{(p_{high} - p_{low})} \cdot ln\left(\frac{f_{i-1}}{f_{i+1}}\right). \tag{2.6}$$

The above derivation assumes that the values of $f_{i\pm1}$ are taken at the center of their respective bins. This at least more or less reproduces the same bin shift as used by PHENIX, especially as the corrections are typically less than 1%.

To avoid producing undefined values such as 'not a number' (NAN) or 'infinity' (INF) in the YODA file, three primary conditions are considered before applying the bin shift:

1. Ensuring that we are not including the overflow bins at the edges of the histogram in the .yoda file. These are edge bins that contain overflow data and are not relevant to the analysis.

2. The histogram must have at least three physical bins to perform the bin shift.

3. Either the bin immediately to the left or right of the bin being operated on is non-zero.

For the outermost physical edge bins, it is impossible to consider a bin to the left or right of the left-most or right-most bin, respectively. If it did, it would use the overflow edge bins that do not represent physical measurements. In this case, the correction factor is calculated using $f_{i,i-1}$ or $f_{i+1,i}$, as opposed to $f_{i+1,i-1}$ seen in equation 2.6.

2.5 Divide By Zero Protection

In finalize(), normalization, scaling, and histogram division occur. Each histogram is scaled by the 'sum of weight' value associated with the counter of its event centrality. For example, the histogram covers all centrality ranges for a given collision system, and it is scaled by $1/(Number\ of\ events)$.

Similarly, the measurements of R_{dA} and R_{AA} also experience this scaling, for both the p+p and A+A/d+A data. For R_{dA} and R_{AA}, the histograms of A+A/d+A yields are divided by the p+p yields. Rivet was designed to run one collision type per simulation. Thus, the Nuclear Modification Factor requires the running of Rivet twice. Because finalize() is fixed for a given analysis, when the two .yoda files are merged, finalize() is re-run over the raw data. In the output .yoda file, either collision type yields NAN values for those sections.

To prevent this, divide-by-zero protections are implemented for normalizations, histograms requiring division, and bin shifts. This is done by implementing an if statement that requires the counter for a given measure to be greater than zero. The .yoda file indicates the counter value as the sum of weights (sumW). If the if statement fails, the scaling or division does not take place. The ignored division operation leaves zeros in the .yoda file.

2.6 Updates by Analysis

The issues specific to each analysis are listed below. Analyses are tested at high statistics before review. I623413 is to be submitted to the collaboration for review. Some minor, fixable issues persist for I731133 and I778168.

2.6.1 PHENIX 2004 I623413

The analysis of 'PHENIX 2004 I623413' contains Au-Au collisions at 130GeV. The yields of π^+, π^-, K^+, K^-, p, and $\bar{p}$ as a function of momentum, p_T, are recorded. Each yield histogram has a separate output for the following centrality ranges: 0-5%, 5-15%, 15-30%, 30-60%, 60-92%, and min-bias, which contains 0-92%. The rapidity range is restricted to $|y| < 0.5$.

Several datasets were removed from the analysis. Table 2 of [9] contains the dN/dy at midrapidity for hadrons produced at midrapidity in each centrality class. N_{part} and N_{coll} are parameters that do not exist in the data, as they are estimated using a model.

When booking the hadrons invariant yields, they were previously booked as profiles. In this scenario, clunking code was used to normalize results. These were changed to histograms to simplify code functionality.

Table 4 of [9] contains the resulting inverse slopes in MeV after fitting an exponential to the spectra in each event centrality class. These are also removed because direct comparisons between the data and the model are more robust than comparisons of fits to data and models.

The data in Figure 11a-1, 11a-2, and 11a-3 are a pure ratio of central to peripheral measurements of yields for each particle. The HEPData entry for these figures was incorrectly input. To prevent incorrect tuning and comparison, these plots are omitted from the final analysis. While the correct data can be tracked down at a later date and the histograms added back, it is omitted from this work.

The integrated means from Figure 13 (of the HEPData) or Table 3 of [9] were changed from histograms to profiles to use the internal function of Rivet to calculate the integrated means (or the average value). This required adding an integer counter for each relevant particle that incremented by 1 for each particle produced.

2.6.2 PHENIX 2007 I731133

From 'PHENIX 2007 I731133' we present measurements for η and π^0 in p-p, d-Au, and Au-Au collisions at 200 GeV. For η, the inelastic cross section, invarient yields, R_{dA}, and R_{AA} are measured. The ratio of η to π^0 is measured. All are plotted for set centrality ranges.

The particle η was listed as a charged particle in the analysis, which caused it to be missed completely when looping over all particles. It was updated to be treated as a neutral particle.

The analysis incorrectly labeled and defined the projections for Primary & Unstable particles in init(). This analysis does not consider any primary particles and thus was removed from init(). The definition used for unstable particles was updated to:

Listing 2.1: unstable particles

```
1    const UnstableParticles np((Cuts::abspid == 111 ||
2    Cuts::abspid == 221) && Cuts::absrap <
3    0.5 && Cuts::pT > 1*GeV);
4          declare(np, "np");
```

This analysis still requires work on the R_{dA} and R_{AA} for proper comparisons. During merging, the cross sections are not kept separate, but are added together instead. This results in invalid scaling of all histograms. This is a common issue with analyses, as merging is still a developmental feature in Rivet.

2.6.3 PHENIX 2008 I778168

The analysis of 'PHENIX 2008 I77816' contains p-p and Au-Au collisions at 200GeV. The invariant yield and R_{AA} [5] for π^0 is measured and recorded for select centralities ranging from 0-92%. The corrections to this analysis are as follows.

In analyze(), the nesting of several if() statements for filling histograms was corrected. Previously, for the p+p system, π^0 were filling the incorrect histogram. The Au+Au system was nested inside the p+p if statement, causing π^0 to not populate any of the histograms.

This measurement also compared R_{AA} as a function of centrality, which still requires implementation due to the way in which the output HEPmc file records data.

[5] This analysis suffers the same issue as 'PHENIX 2007 I731133' with respect to R_{AA} measurements.

Chapter 3

Results

This chapter will cover the physics implications of the PYTHIA8-Angantyr model in reference to the PHENIX data. Yields, yield ratios, cross sections, and the nuclear modification factor will be discussed. Each analysis will be discussed separately, using the selected plots produced by the analysis code finished in this thesis to make the comparisons. All other plots can be found in Appendix D, E, & F.

3.1 Results by Analysis

3.1.1 PHENIX 2004 I623413

The model was run over 10^5 events[1]. We will cover the invariant yields, integrated yields (dN/dy), and $< p_T >$ (average p_T) of π^+, π^-, K^+, K^-, p, and $\bar{p}$.

In the invariant yields of the charged pions, we see, regardless of centrality, an underestimation of particles at low p_T while the model begins to overestimate at high p_T. Angantyr does not account for radial particle flow associated with hydrodynamic evolution or a QGP phase model. This results in no rescattering and reduces the amount of interactions that we should see at low p_T where particles spend more time in the hot, dense medium. Above .75 Gev/c, the missing radial flow will change the shape of the spectrum, but not the integral. Flow moves particles around and tends to move particles closer to $< p_T >$.

[1]Due to the linear scaling of the simulation, running 10^6 events would have taken 30 hours.

The integrated yields and $< p_T >$ (average p_T) of the model match the simulation well. Figure 3.1 shows the plots of the most central collisions for each particle class.

The model's charged kaon spectrum of invariant yields is underestimated for $K^{+/-}$ at low p_T but reaches agreement with experiment at higher p_T. Due to the Lund String model used, $s\bar{s}$ production is suppressed relative to $u\bar{u}$ and $d\bar{d}$. As with the charged pion production, there are no contributions from QGP formulation or hydrodynamic evolution, leaving only string fragmentation and other decays to produce the necessary s and $\bar{s}$. Due to the low total statistics, some of the less central collision systems experience few entries and some overestimation. The model's integrated yields match well with experiment. The integrated yields suffer from the same underestimation issue as above.

The proton & anti-proton invariant yields primarily underestimate for all p_T bins and centrality ranges. Under Pythia's Lund fragmentation model, during and after the initial collision, recombination of the protons is suppressed. The average transverse momentum is underestimated consistently, regardless of centrality. The underestimation of the integrated yields is once again due to the lack of hydrodynamics and QGP considerations.

The invariant yield patterns in the most central collisions for each particle are shown in Figure 3.1. A full list of results can be found in the Appendix Section D.

3.1.2 PHENIX 2007 I731133

For this analysis, Rivet was run 3 separate times for each collision system. Output Yoda files were merged to rerun the correct final calculations from the analysis. The p+p simulation was ran for 10^7 events. The d+Au system has 10^6 collisions and the Au+Au system has 10^5. We report on the cross section for the p+p and d+Au collision systems and the ratio of η/π for all collision systems. Only the Au+Au and d+Au collision systems report the invariant yields of η.

The p+p cross section holds well even to high p_T. The more central collisions of the d+Au system underestimate the cross section, while the more peripheral collisions agree with the data.

In the p+p collision system, the ratio of π to η measurement from the model begins to break down around 5 GeV/c. The d+Au system underestimates the ratio for most centrality

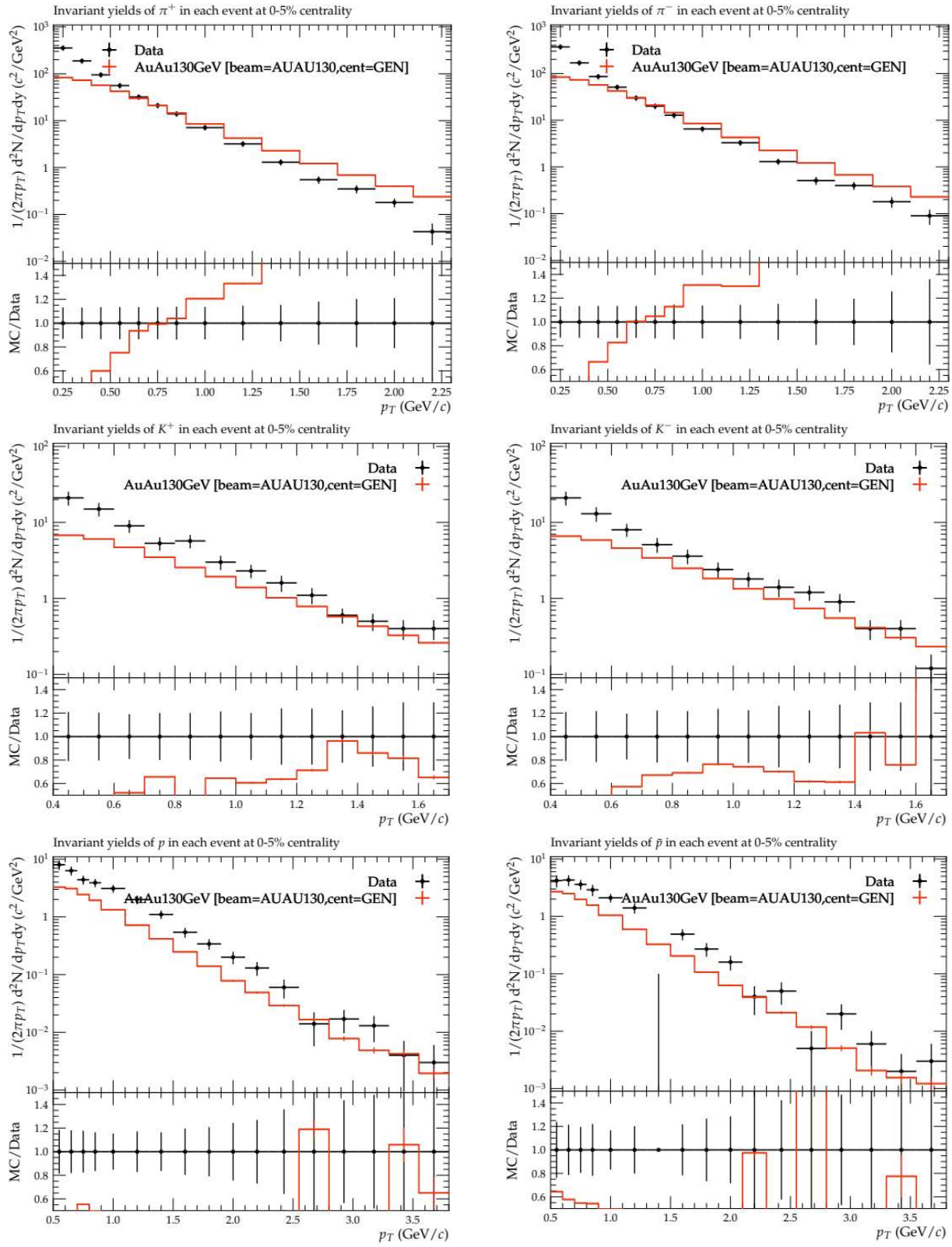


Figure 3.1: Invariant yields of the top 0-5% most central collision of (top to bottom, left to right) $\pi^{+/-}$, $K^{+/-}$, and p & $\bar{p}$ from PHENIX 2004 I623413.

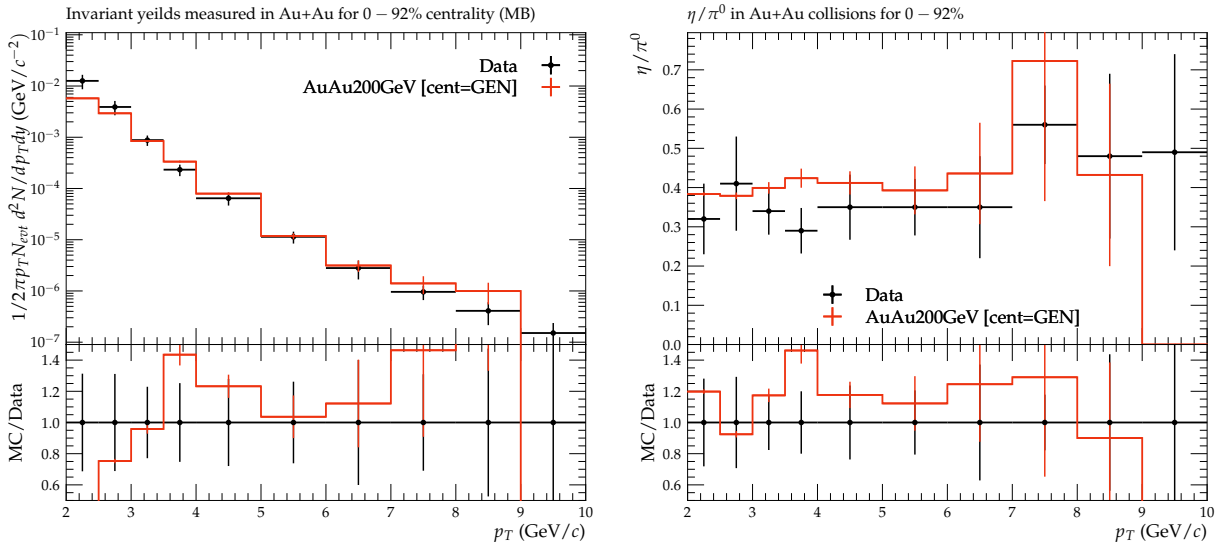


Figure 3.2: Invarient Yields of π^0 (left) and the ratio of η to π (right) for the Au+Au collision system at min bias ranges of PHENIX 2007 I731133.

ranges, despite several outliers present with lower statistics. The model estimates are within the uncertainty given by the experiment for the Au+Au system, with larger error bars at higher p_T in the simulation.

The d+Au collision system underestimates the η yield at low p_T with statistical fluctuations at high p_T. The Au+Au invariant yield of the model fluctuates above and below the data and error bars. A full list of results can be found in the Appendix Section E, while Figure 3.2 displays the min bias invariant yields and η to π ratio of Au+Au.

3.1.3 PHENIX 2008 I778168

This Rivet analysis reports on the Au+Au system that ran for 10^6 events. We report on the comparison of the invariant yields of π^0 to data.

The Yields of π^0 match experiment consistently for all centrality ranges within most error bars. Due to low event values, statistical uncertainty begins to appear near the tail end of each simulation plot. The full list of plots can be found in Appendix F.

Table 3.1: Particle Masses

Particle	Mass (GeV/c^2)
π^0	0.134977
$\pi^{+/-}$	0.13957
$K^{+/-}$	0.493677
K^0	0.497614
η	0.547853
p & $\bar{p}$	0.938272

3.2 Discussion

We note several expectations from the experiments to apply to each simulation. Yields increase strongly with collision centrality because more participant nucleons contribute to particle production [9], [10], [6]. For particle production, we expect lower mass particles to be produced most abundantly. From Table 3.1, simulation and experiment should see yields such that $\pi^x > K^x > p$ $(or$ $\bar{p})$ [16]. Due to this, running at low statistics will introduce statistical fluctuations where particle production is already low. QGP effects are expected to be strongest in the most central Au+Au collisions due to the larger overlap region and greater number of participating nucleons [5].

Chapter 4

Conclusions

We have completed three Rivet Analyses comparing data from the PHENIX collaboration to the model Pythia8-Angantyr. These are to be submitted for approval by the collaboration. This consisted of debugging existing code and comparing the simulation against known results and expectations from the model.

Once approved, these analyses will be added to the official Rivet Repository, where they will be usable by anyone seeking to use them. They will be used as tuning parameters in ongoing Bayesian Parameter Estimations with AMPT.

The utility of Rivet as a data analysis tool has been demonstrated while breaking down its fundamental parts. Thus, this thesis will also act as a roadmap for future students to develop an understanding of how this code operates to further complete the repository of Rivet analyses for PHENIX.

Bibliography

[1] Concept and History — phenix.bnl.gov. https://www.phenix.bnl.gov/experiment/concept.html. [Accessed 31-03-2026]. 1, 12

[2] Zero Degree Calorimeters and Shower Max Detectors — phenix.bnl.gov. https://www.phenix.bnl.gov/detectors/zdc.html. [Accessed 31-03-2026]. viii, 15

[3] (2024). doc/tutorials/mig3to4.md · release-4-1-x · Cedar / Rivet · GitLab — gitlab.com. https://gitlab.com/hepcedar/rivet/-/blob/release-4-1-x/doc/tutorials/mig3to4.md?ref_type=heads. [Accessed 18-06-2026]. 23

[4] Abelev, B. I. et al. (2009). Systematic Measurements of Identified Particle Spectra in pp, d^+ Au and Au+Au Collisions from STAR. *Phys. Rev. C*, 79:034909. 6

[5] Adams, J. et al. (2005). Experimental and theoretical challenges in the search for the quark–gluon plasma: The star collaboration's critical assessment of the evidence from rhic collisions. *Nuclear Physics A*, 757(1):102–183. First Three Years of Operation of RHIC. 34

[6] Adare, A. et al. (2008). Suppression pattern of neutral pions at high transverse momentum in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV and constraints on medium transport coefficients. *Phys. Rev. Lett.*, 101:232301. 17, 34

[7] Adare, A. et al. (2011). Identified charged hadron production in $p + p$ collisions at $\sqrt{s} = 200$ and 62.4 gev. *Phys. Rev. C*, 83:064903. 6

[8] Adcox, K. et al. (2003). Phenix detector overview. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 499(2):469–479. The Relativistic Heavy Ion Collider Project: RHIC and its Detectors. 17

[9] Adcox, K. et al. (2004). Single identified hadron spectra from s(NN)**(1/2) = 130-GeV Au+Au collisions. *Phys. Rev. C*, 69:024904. 14, 27, 28, 34

[10] Adler, S. S. et al. (2007). High transverse momentum η meson production in p^+p, d^+ Au and Au+Au collisions at $S(NN)^{(1/2)}$ = 200-GeV. *Phys. Rev. C*, 75:024909. 14, 34

[11] Bierlich, C., Chakraborty, S., Desai, N., Gellersen, L., Helenius, I., Ilten, P., Lönnblad, L., Mrenna, S., Prestel, S., Preuss, C. T., Sjöstrand, T., Skands, P., Utheim, M., and Verheyen, R. (2022). A comprehensive guide to the physics and usage of pythia 8.3. 7

[12] Bierlich, C., Gustafson, G., Lönnblad, L., and Shah, H. (2018). The Angantyr model for Heavy-Ion Collisions in PYTHIA8. *JHEP*, 10:134. 7

[13] Collaboration, A. (2017). The ALICE definition of primary particles. 6

[14] Docker (2026). Docker Desktop: The #1 Containerization Tool for Developers — Docker — docker.com. https://www.docker.com/products/docker-desktop/. [Accessed 08-06-2026]. 21

[15] Gohd, C. (2026). What is Dark Energy? Inside Our Accelerating, Expanding Universe — science.nasa.gov. https://science.nasa.gov/dark-energy/. [Accessed 07-04-2026]. viii, 3

[16] Group, P. D. (2026). PDG Particle List. https://www.phy.bnl.gov/twister/bee/particles/. [Accessed 22-06-2026]. 34

[17] Karsch, F. (2002). *Lattice QCD at High Temperature and Density*, page 237. Springer Berlin Heidelberg, Berlin, Heidelberg. 2

[18] Nagamiya, S., Aronson, S. H., Young, G. R., and Paffrath, L. (1993). Phenix conceptual design report. an experiment to be performed at the brookhaven national laboratory relativistic heavy ion collider. 12

[19] RIVET (2025). Rivet – the particle-physics MC analysis toolkit — rivet.hepforge.org. https://rivet.hepforge.org/. [Accessed 03-04-2026]. 5

[20] Shen, C. and Heinz, U. (2015). The road to precision: Extraction of the specific shear viscosity of the quark-gluon plasma. *Nuclear Physics News*, 25(2):6–11. viii, 4

[21] Shuryak, E. V. (1980). Quantum chromodynamics and the theory of superdense matter. *Physics Reports*, 61(2):71–158. 2

[22] Watt, G. and Byers, J. (2026). About HEPData — hepdata.net. https://www.hepdata.net/about. [Accessed 08-06-2026]. 5

[23] Zajc, W. (2003). Phenix experiment at rhic: Decadal plan 2004-2013. 12

Appendix

A Summary of Equations & Definitions

Below contains equations & definitions of variables that, while important, are often common knowledge and are given for those unfamiliar with them.

A.1 Pseudorapidity and Rapidity

Rapidity and Pseudorapidity qualitatively describe how far you are from the beam pipe.

Pseudorapidity, η, is a coordinate representing the angle of a particle relative to the beam axis:

$$\eta = -ln\left[tan\frac{\theta}{2}\right]. \tag{4.1}$$

Rapidity measures longitudinal velocity as:

$$y = \frac{1}{2}ln\left[\frac{E + p_z c}{E - p_z c}\right] \tag{4.2}$$

Mid-Rapidity refers to $y \approx 0$, where the produced particle has no component of velocity in the direction of the beam. Forward and backward rapidity y in high-energy collisions refer to particle production along the beam axis, with positive (forward) typically denoting the direction of the proton/deuteron beam and negative (backward) the direction of the heavier nucleus, used to probe different kinematic regions and nuclear effects.

B Docker Start-Up

Below is the process of setting up the Docker Container and GitHub Repo on your Windows Machine. File pathways using (\) are for Windows convention when needed and / is for the Linux convention. Scripts that include “” are kept. Scripts including ‘’ are not kept and only used to identify scripts from text.

1. Start with Downloading Docker

https://docs.docker.com/desktop/setup/install/windows-install/

You don't need to use an email.

If asked to update the Linux subsystem using 'wsl –update' that's completely fine.

If on a (custom) PC, may have to enable virtualization in the BIOS. Once the program is installed and running, you must keep the application open to run the containers(s).

2. Establish a local directory in which Docker. Create the **BOLD** folder

   Location: C:\Users\UserName**rivet_research**

   OR use the terminal command: mkdir C:\Users**UserName****rivet_research** (replace **UserName** with your account user name)

3. Pull the image needed for docker:

   A Docker image is a read-only template that contains all the necessary instructions, application code, libraries, dependencies, and configurations needed to create a fully functional, isolated Docker container.

   Run this in the command Prompt: docker pull hepstore/rivet-pythia

4. Mount the directory to use:

   Add the file directory in the text box. Go through the setting in Docker Via:

   Settings → Resources → File Sharing.

   I left my pathway as: C:\Users**UserName**

5. Attach the local directory for docker to work in conjunction with:

   docker run -it –name rivet_pythia -v "C:\Users**UserName**\rivet_research:/rivet_work" hepstore/rivet-pythia bash

6. This should produce a working environment in the docker container. It should have this file pathway (or something similar): root@26c3413e0d0e:/work

7. This is the temporary work station where nothing is saved. You need to move into the environment where actions sync to the folder we made earlier in step 2. Use 'cd /rivet_work' (without the quotations) to move into the correct environment.

   /work → default working directory (temporary)

   /rivet_work → your mounted Windows folder (persistent)

8. Use ls to check where you need to go and you can start running Rivet!

   In the terminal, the follow commands will open and close the docker container:

   Start: docker start -ai rivet_pythia

   You would then use step 7 to move into the live directory.

   Close: exit

9. To clone the repo into the **rivet_research**, run the following command:

   git clone https://github.com/cnattras/RIVETAnalyses.git

10. At this point, we do not have write permissions to the repo. The following steps explain how to gain access. From your working directory: /rivet_work/RIVETAnalyses

    Run:

    ssh-keygen -t ed25519 -C "your_email@example.com"

    You will then select the defaults for location by hitting enter. Once the key fingerprint and randomart image are produced, run:

    cat /.ssh/id_ed25519.pub

    This will produce the code you need. Copy 'ssh-ed25519.....your_email@example.com'. Navigate to your setting in GitHub. (You github account should match the email used above). Under settings, select 'SSH and GPG Keys' and select 'Add New Key'. The naming convention is up to you. The key type is 'Authentication Key.' You then past the code from earlier and add the SSH key.

11. You now can submit changes to the repo as you are working and have full functionality of the Docker Container. You may repeat steps 5-10 to produce more containers. This

will be useful when running one analysis and wanting to work on another in the second container, just don't bog down your computer.

C Code

This section contains snippets of the code added into the analysis that are described in Chapter 2.

C.1 Bin Shift

Listing 4.1: Bin-shifting algorithm

```
void binShift(YODA::Histo1D& histogram) {

    const auto& binlist = histogram.bins();

    for (size_t n = 1; n < binlist.size()-1; ++n) {

        const auto& bins = binlist[n];

        double p_high = bins.xMax();
        double p_low  = bins.xMin();

        double left;
        double right;

        // Define neighboring bins
        if (n == 1) {
            left  = binlist[1].sumW();
            right = binlist[2].sumW();

        } else if (n == binlist.size()-2) {
            left  = binlist[binlist.size()-3].sumW();
```

```
            right = binlist[binlist.size()-2].sumW();

        } else {
            left  = binlist[n-1].sumW();
            right = binlist[n+1].sumW();
        }

        // Protect logarithm
        if (left <= 0 || right <= 0) {
            continue;
        }

        double b = 1.0 / (p_high - p_low) * log(left / right);

        double denom =
            pow(M_E, -b * p_high) -
            pow(M_E, -b * p_low);

        // Protect division by zero
        if (std::abs(denom) < 1e-12) {
            continue;
        }

        double f_corr =
            -b * (p_high - p_low)
            * pow(M_E, -b * (p_high+p_low) / 2)
            / denom;

        histogram.bin(n).scaleW(f_corr);
    }
}
```

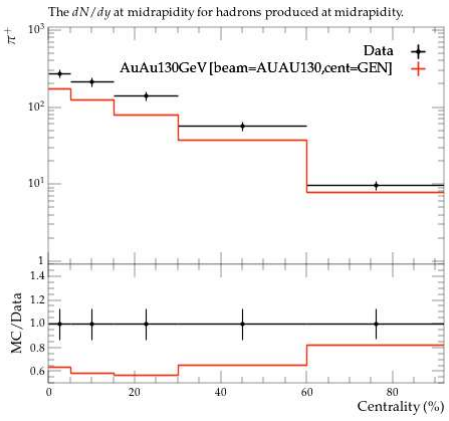
D PHENIX_2004_I623413

From running (rivet-mkhtml –pwd AuAu130GeV.yoda) inside the container, the following plots are produced with the comparison to simulation and data:

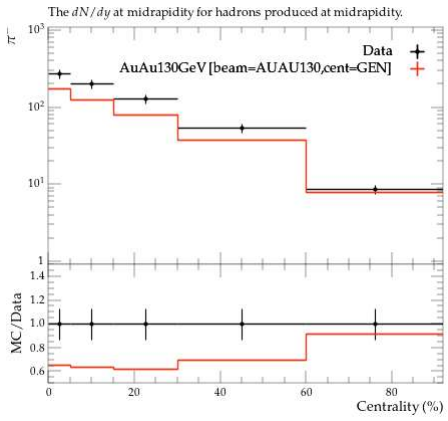
PHENIX_2004_I623413

Transverse momentum spectra and yields of hadrons are measured by the PHENIX collaboration in Au + Au collisions at $\sqrt{s_{NN}}$ = 130 GeV at the Relativistic Heavy Ion Collider (RHIC). The time-of-flight resolution allows identification of pions to transverse momenta of 2 GeV/c and protons and antiprotons to 4 GeV/c. The yield of pions rises approximately linearly with the number of nucleons participating in the collision, while the number of kaons, protons, and antiprotons increases more rapidly. The shape of the momentum distribution changes between peripheral and central collisions. Simultaneous analysis of all the p_T spectra indicates radial collective expansion, consistent with predictions of hydrodynamic models. Hydrodynamic analysis of the spectra shows that the expansion velocity increases with collision centrality and collision energy. This expansion boosts the particle momenta, causing the yield from soft processes to exceed that for hard to large transverse momentum, perhaps as large as 3 GeV/c.

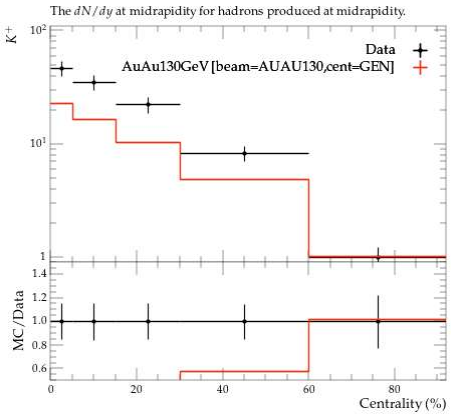
d02-x01-y03:



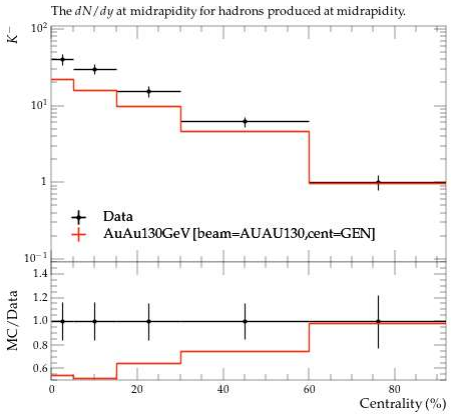
d02-x01-y04:



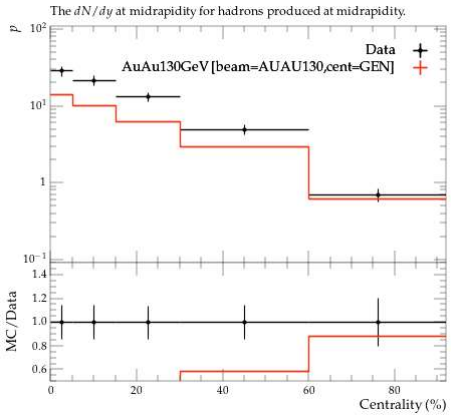
d02-x01-y05:



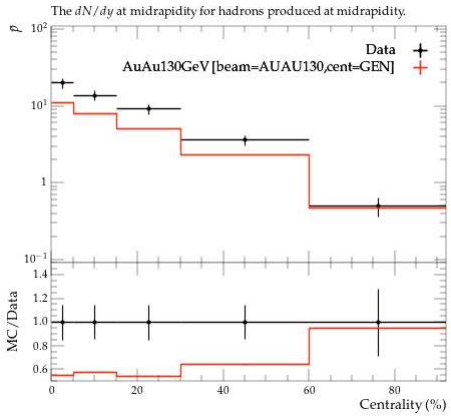
d02-x01-y06:



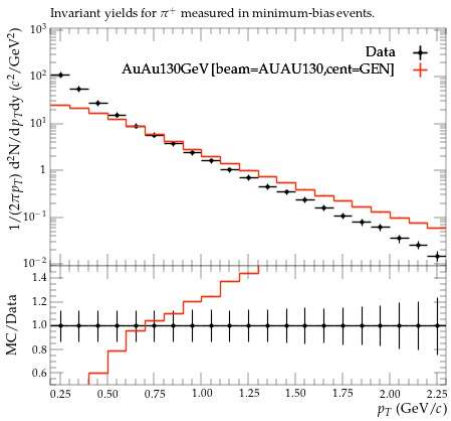
d02-x01-y07:



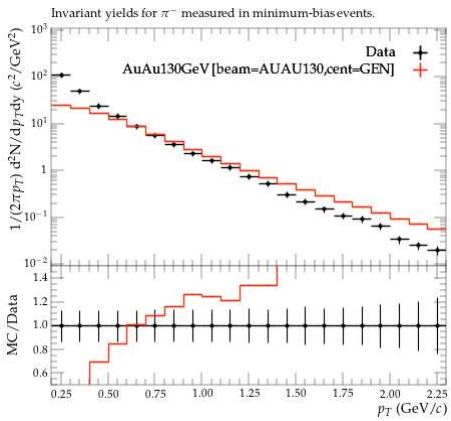
d02-x01-y08:



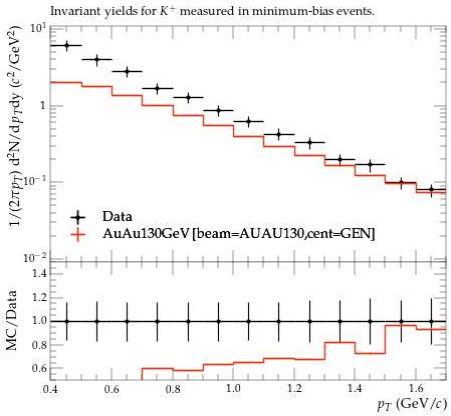
d07-x01-y01:



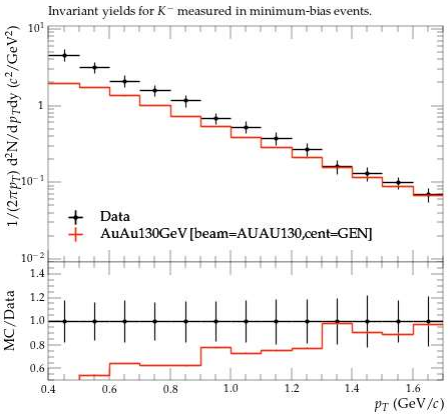
d07-x01-y02:



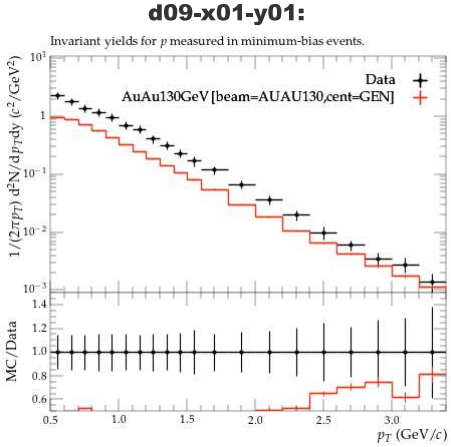
d08-x01-y01:



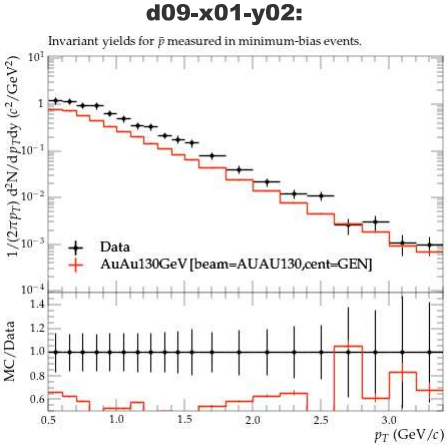
d08-x01-y02:



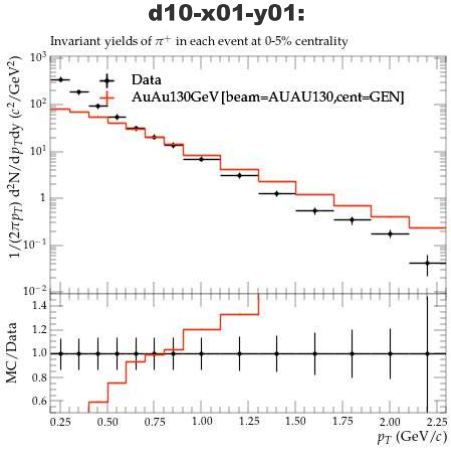
d09-x01-y01:



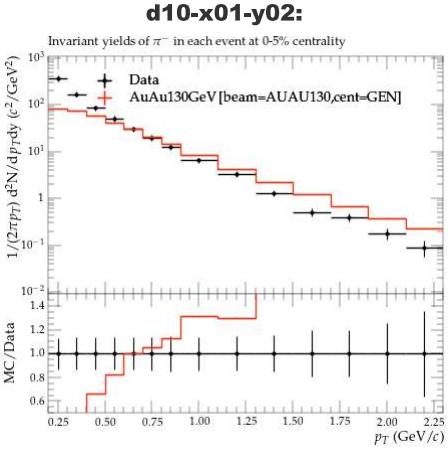
d09-x01-y02:



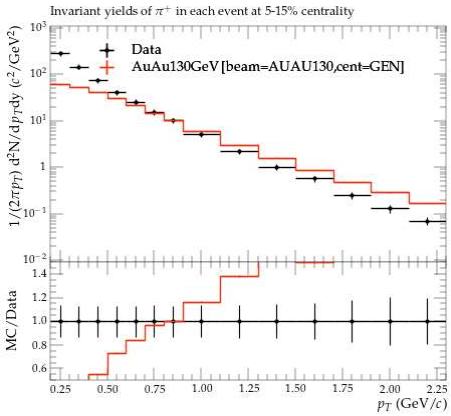
d10-x01-y01:



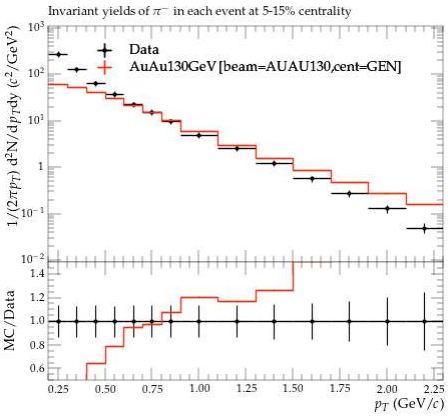
d10-x01-y02:



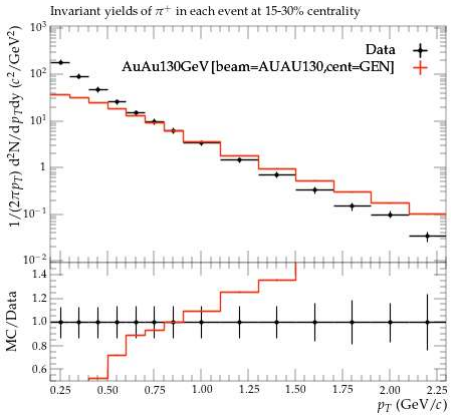
d10-x01-y03:



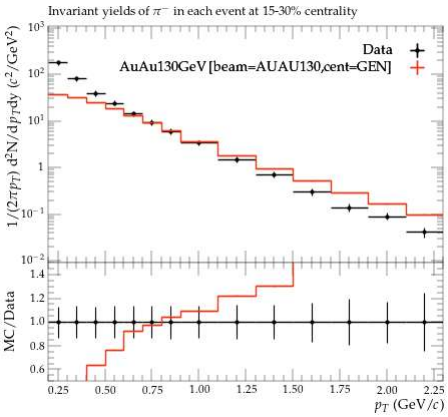
d10-x01-y04:



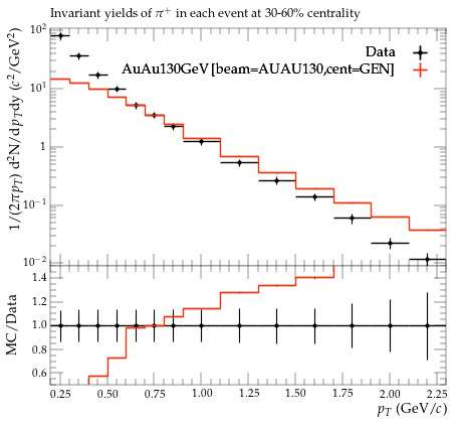
d10-x01-y05:



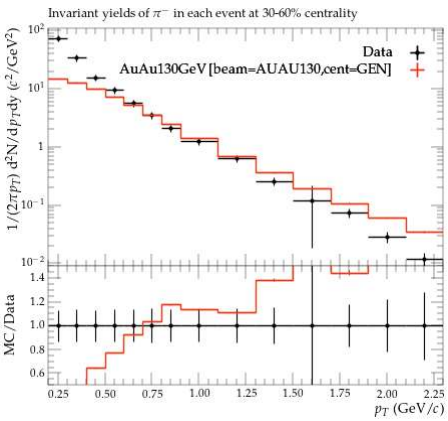
d10-x01-y06:



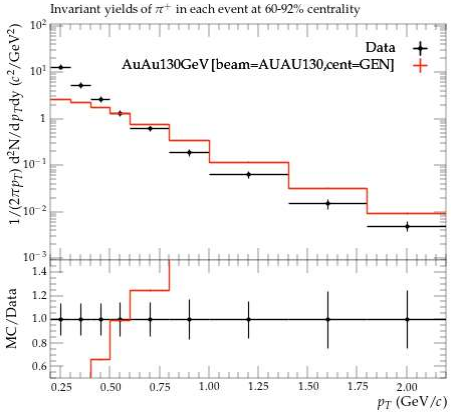
d10-x01-y07:



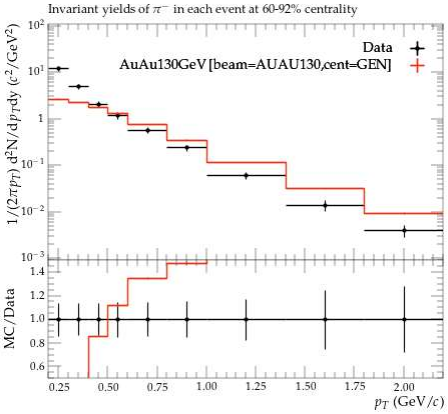
d10-x01-y08:



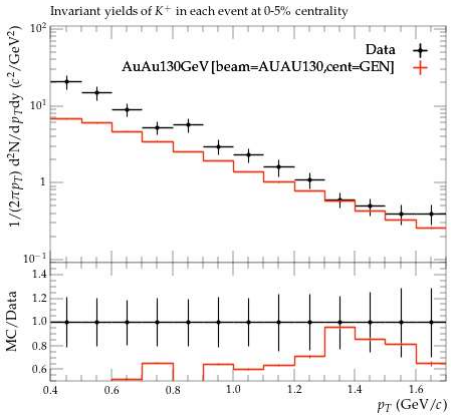
d11-x01-y01:



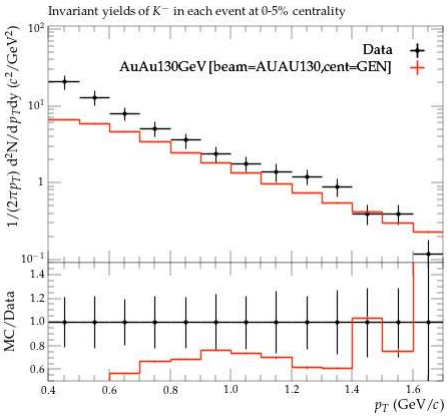
d11-x01-y02:



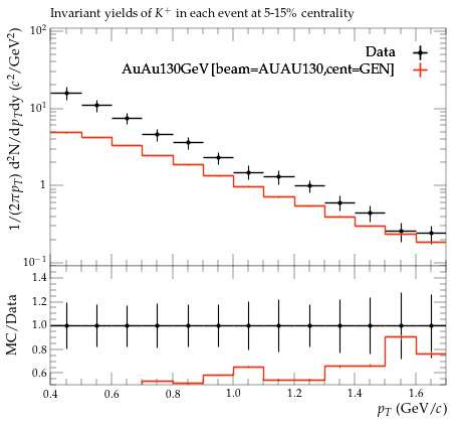
d12-x01-y01:



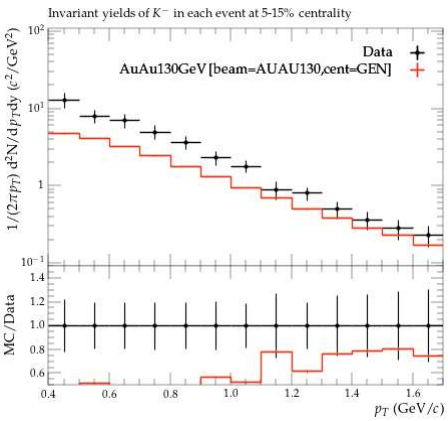
d12-x01-y02:



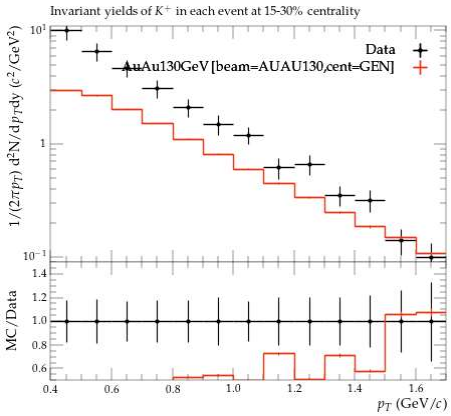
d12-x01-y03:



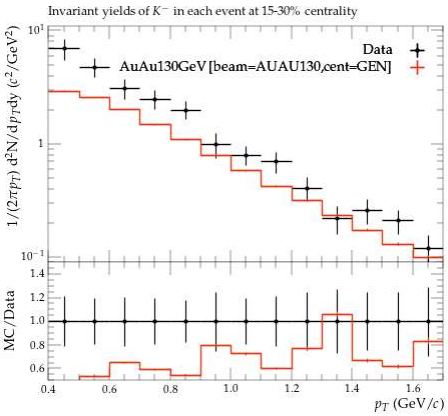
d12-x01-y04:



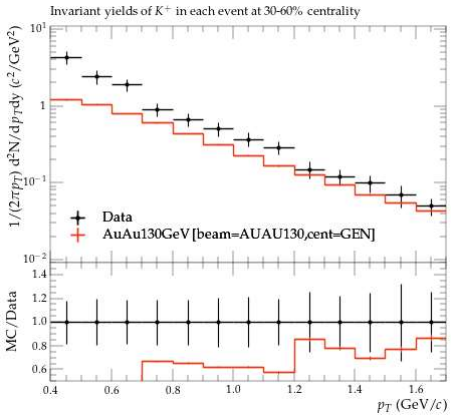
d12-x01-y05:



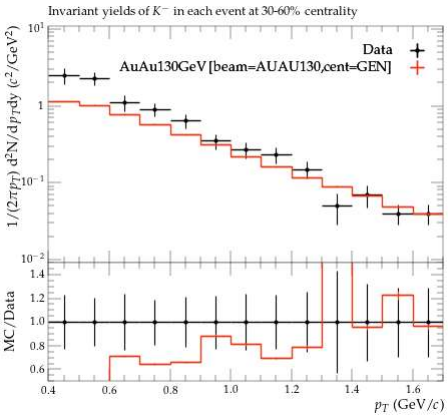
d12-x01-y06:



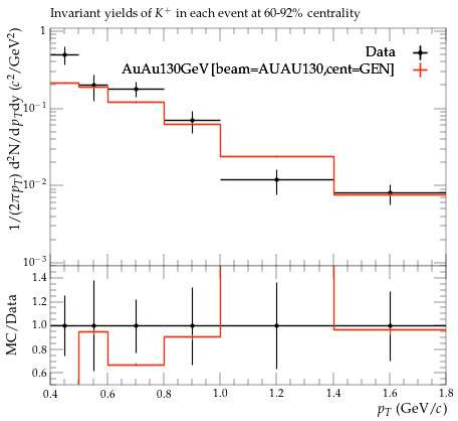
d12-x01-y07:



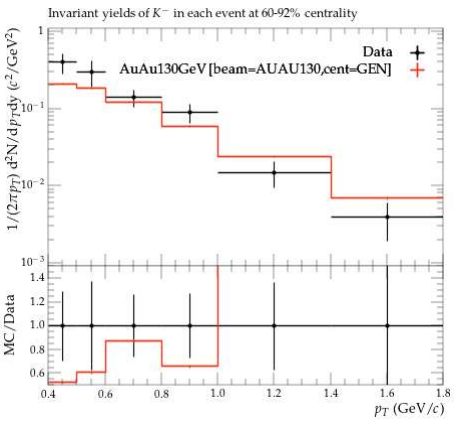
d12-x01-y08:



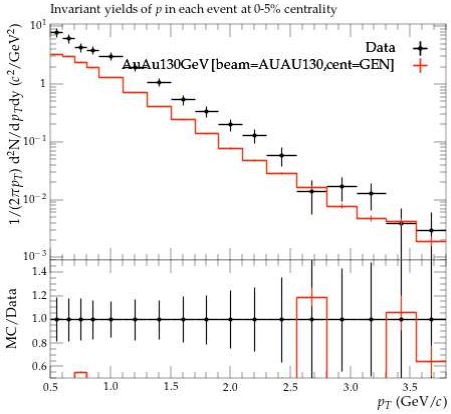
d13-x01-y01:



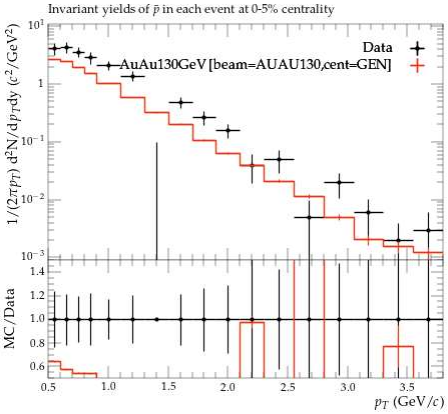
d13-x01-y02:



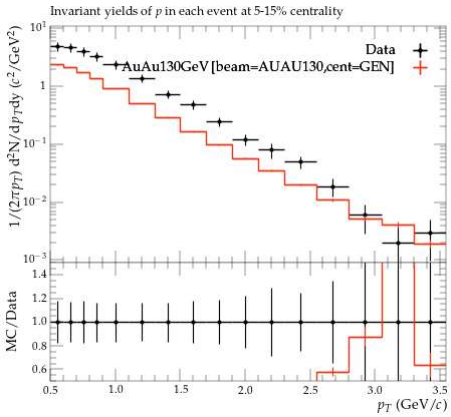
d14-x01-y01:



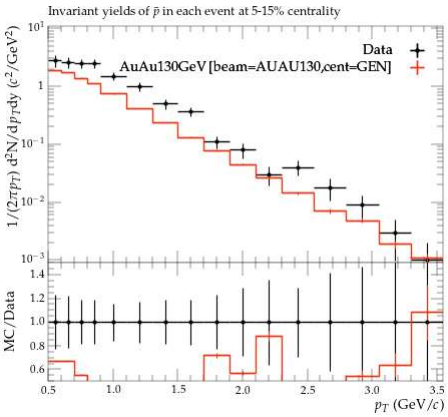
d14-x01-y02:



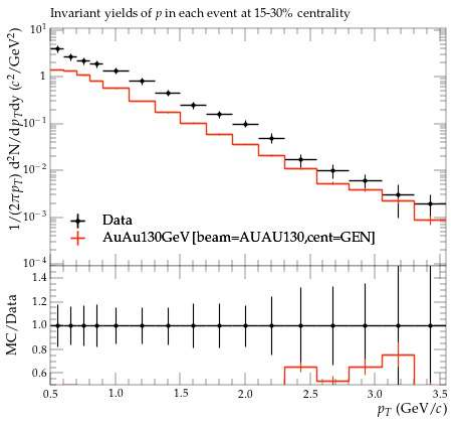
d15-x01-y01:



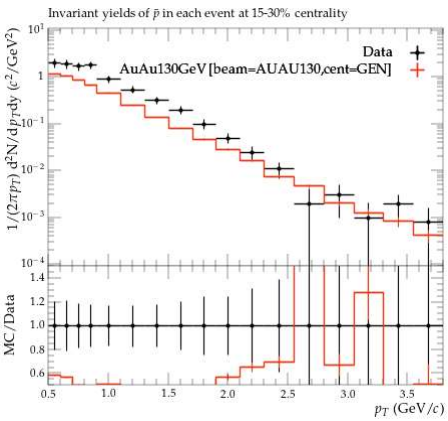
d15-x01-y02:



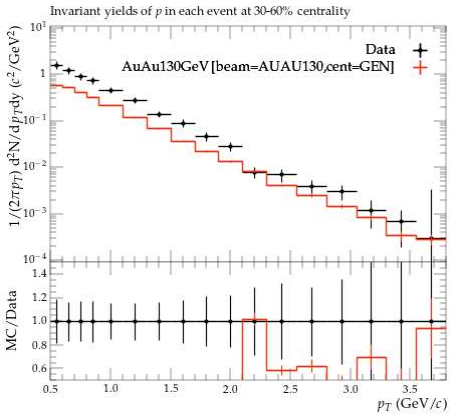
d16-x01-y01:



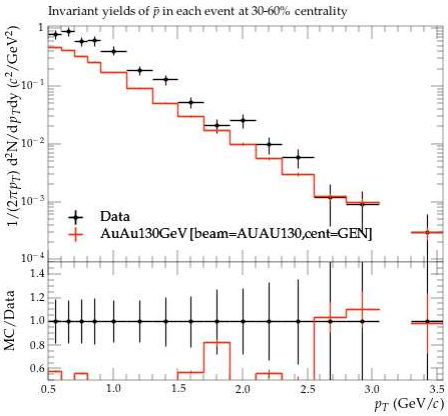
d17-x01-y01:



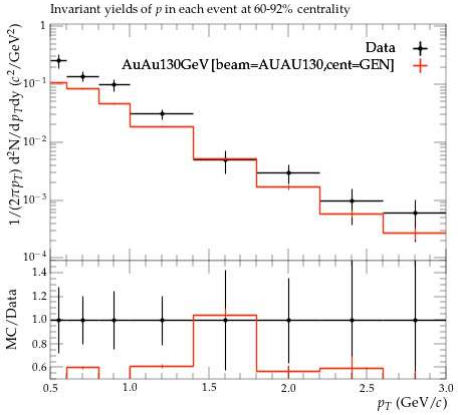
d18-x01-y01:



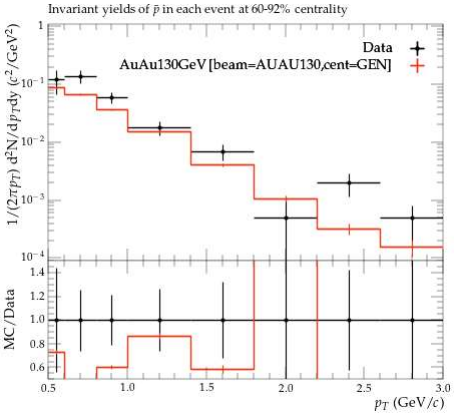
d19-x01-y01:



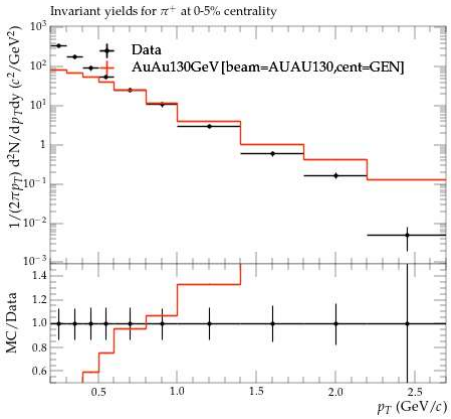
d20-x01-y01:



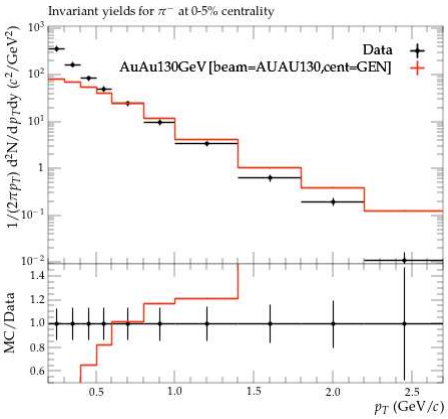
d20-x01-y02:



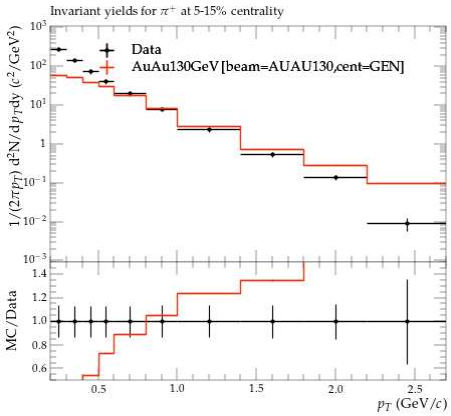
d24-x01-y01:



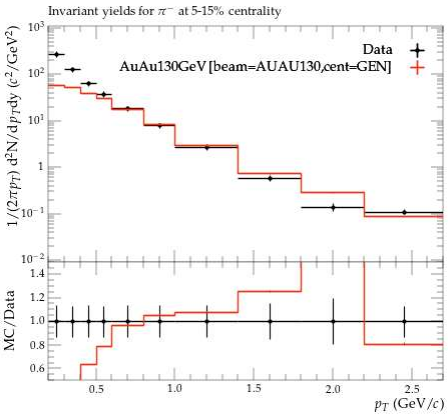
d24-x01-y02:



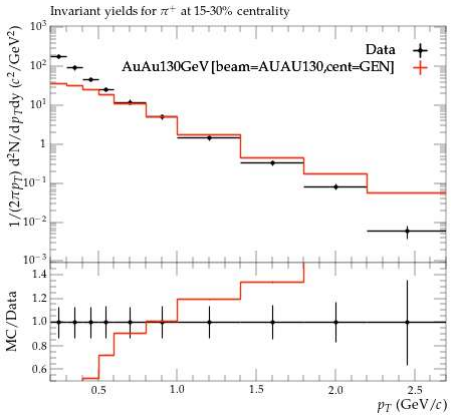
d24-x01-y03:



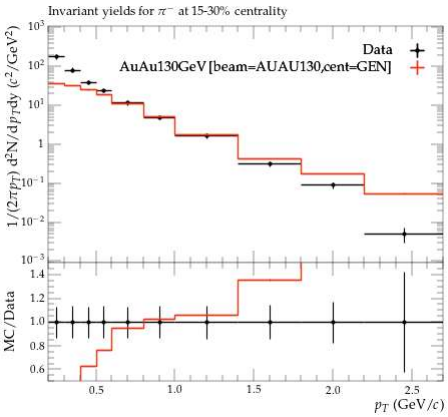
d24-x01-y04:



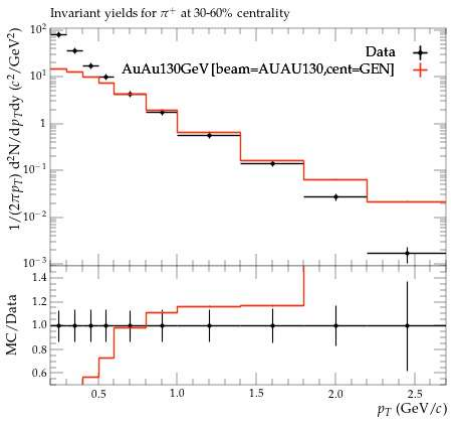
d24-x01-y05:



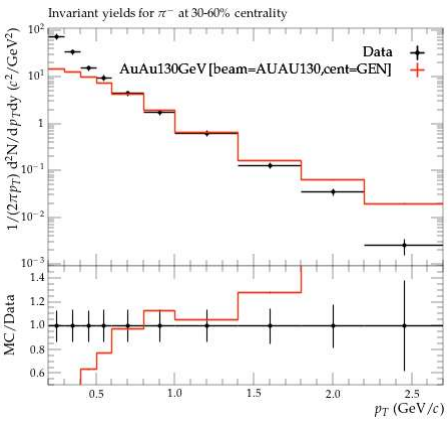
d24-x01-y06:



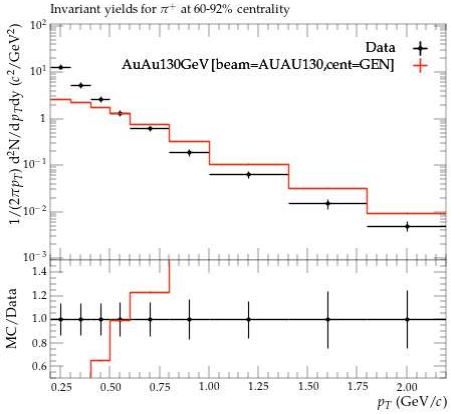
d24-x01-y07:



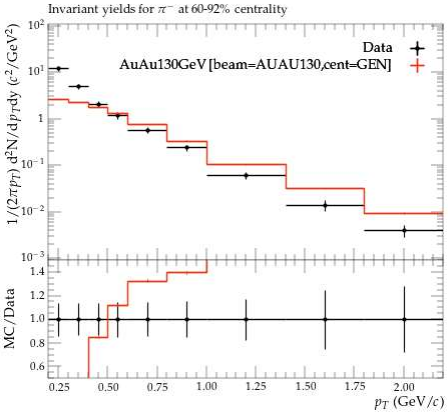
d24-x01-y08:



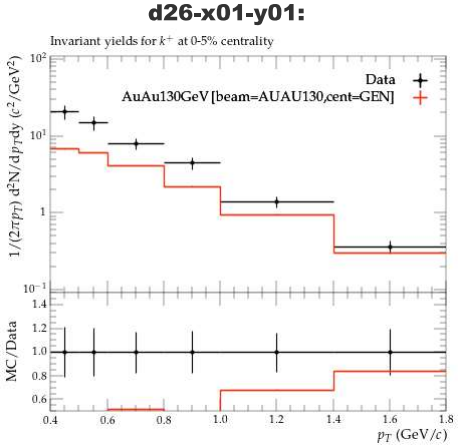
d25-x01-y01:



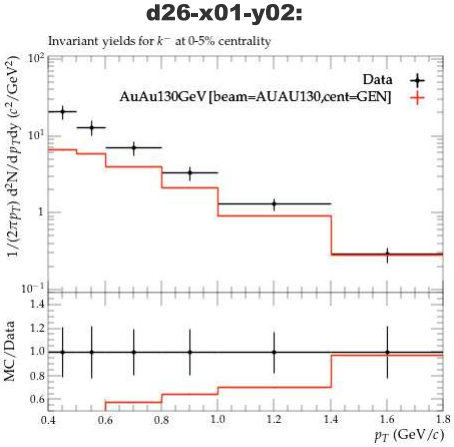
d25-x01-y02:



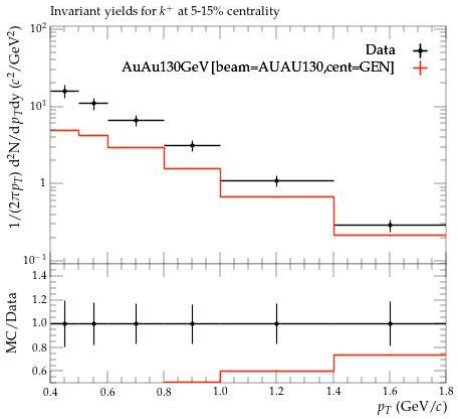
d26-x01-y01:



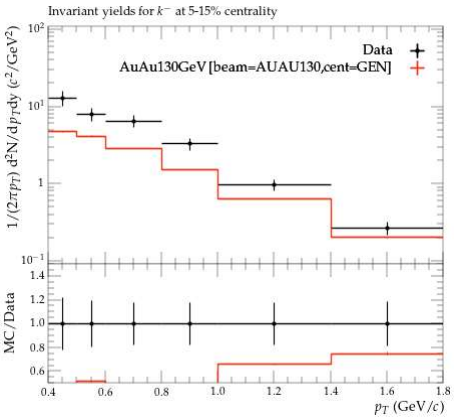
d26-x01-y02:



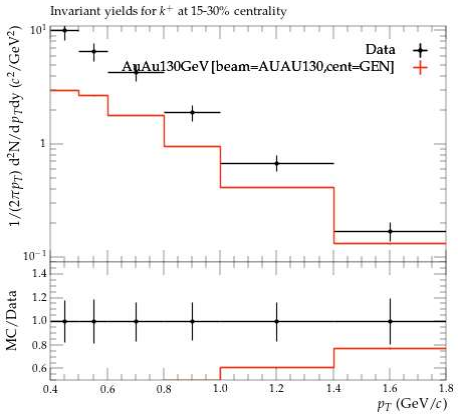
d26-x01-y03:



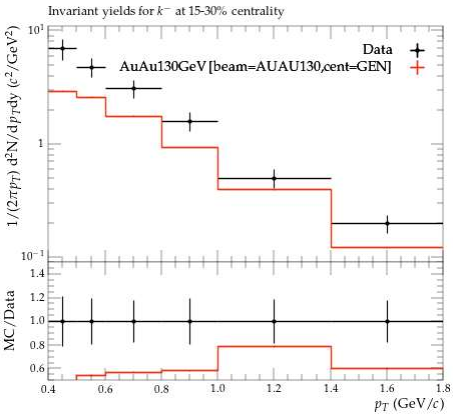
d26-x01-y04:



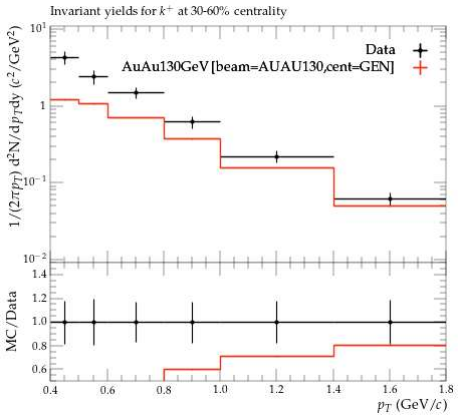
d26-x01-y05:



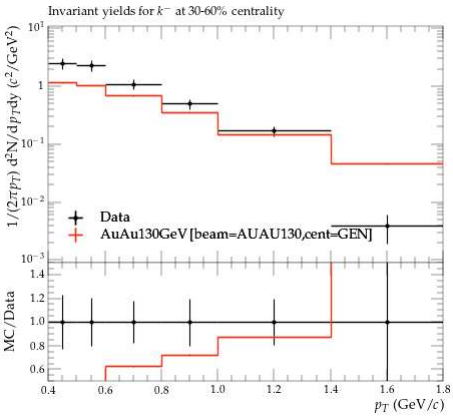
d26-x01-y06:



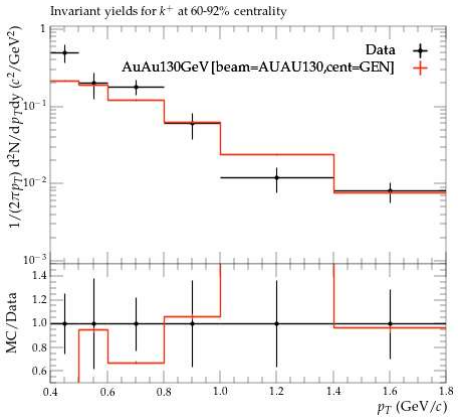
d26-x01-y07:



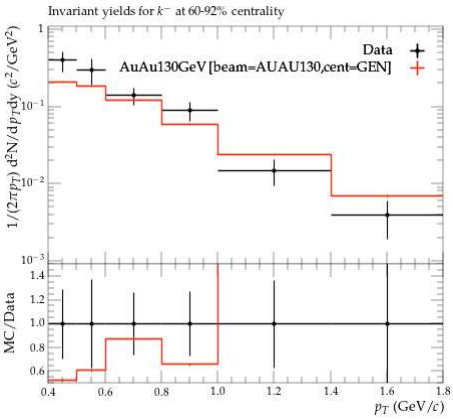
d26-x01-y08:



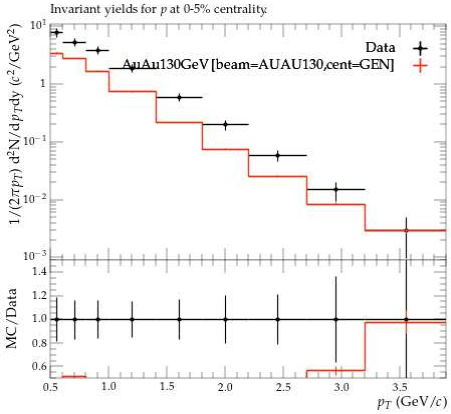
d26-x01-y09:



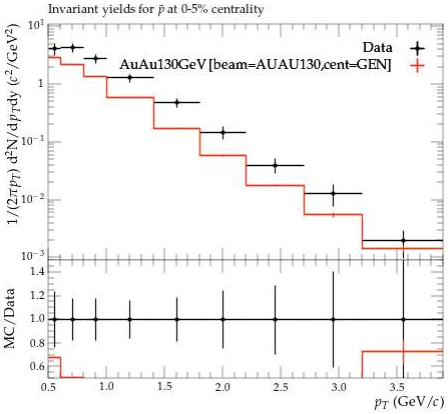
d26-x01-y10:



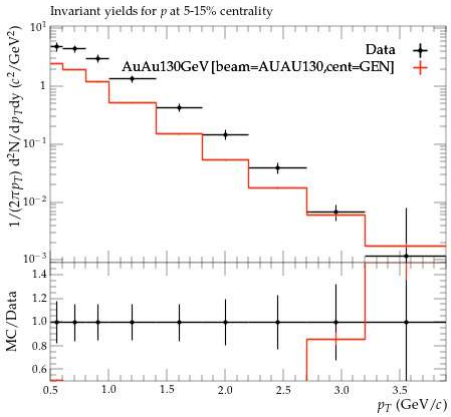
d27-x01-y01:



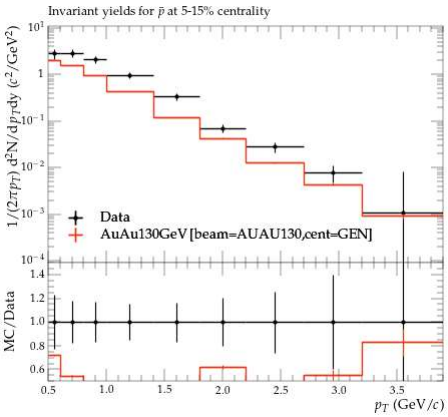
d27-x01-y02:



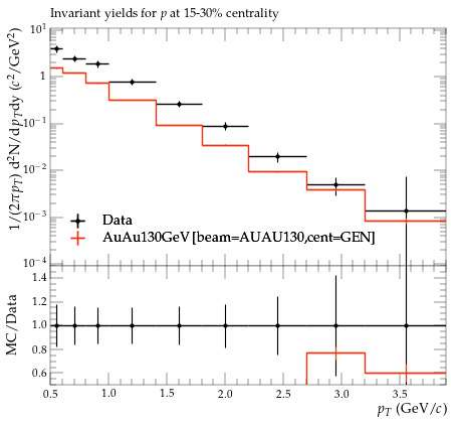
d27-x01-y03:



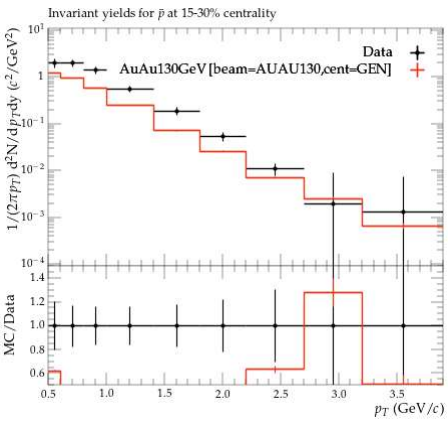
d27-x01-y04:



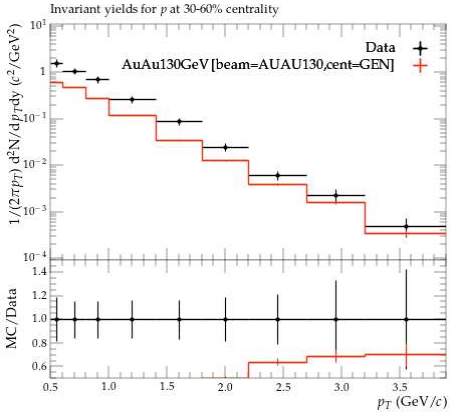
d27-x01-y05:



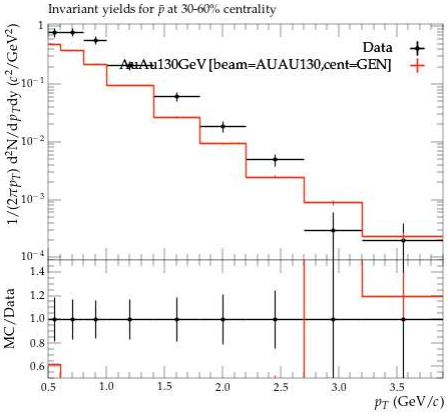
d27-x01-y06:



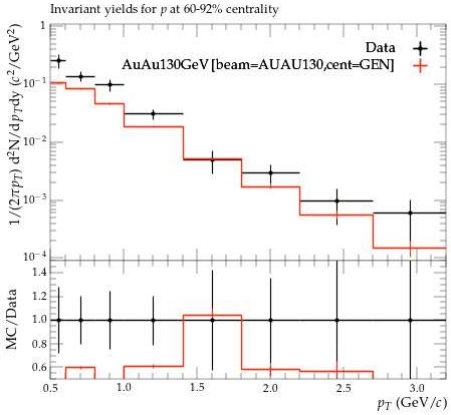
d27-x01-y07:



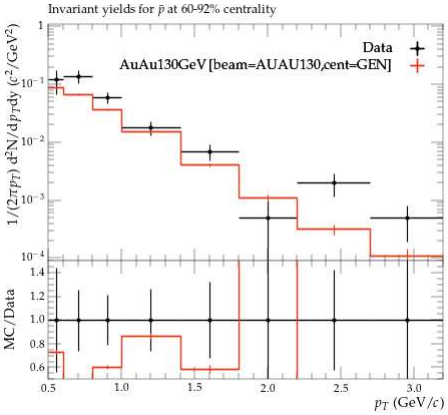
d27-x01-y08:



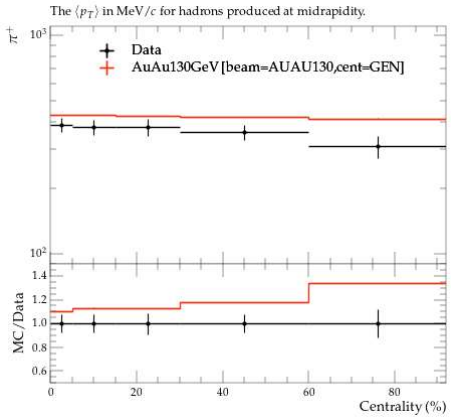
d28-x01-y01:



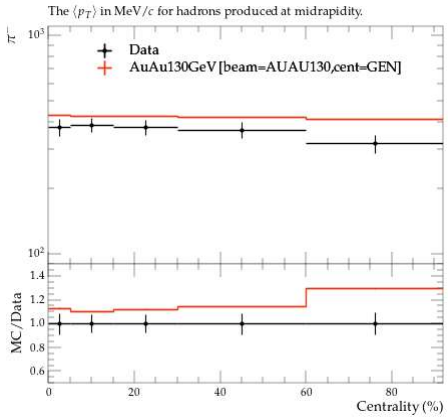
d28-x01-y02:



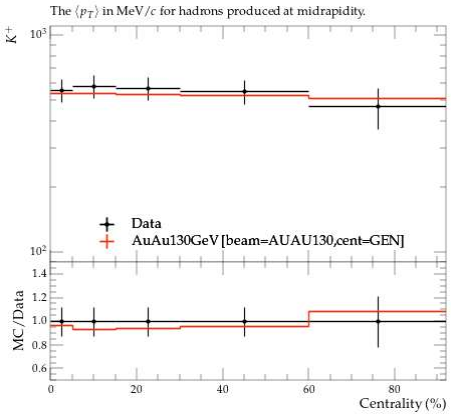
d29-x01-y01:



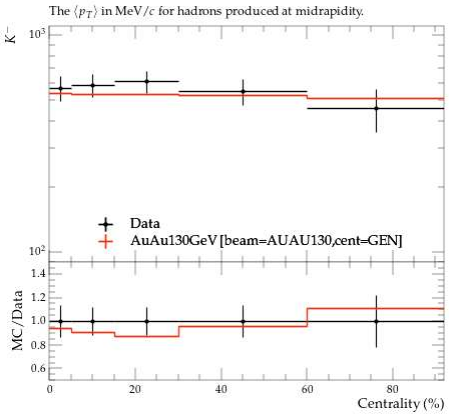
d29-x01-y02:



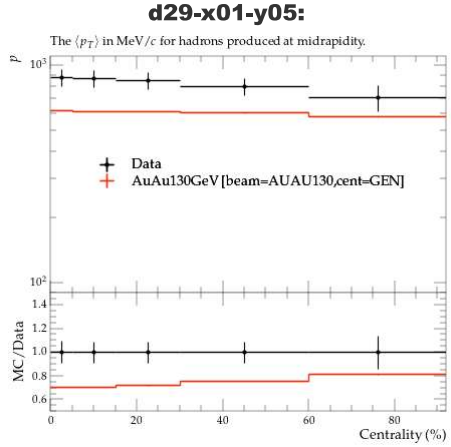
d29-x01-y03:



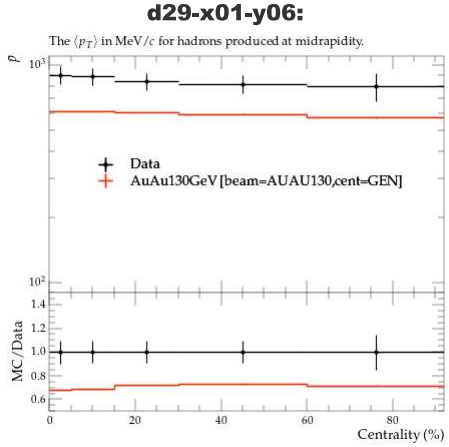
d29-x01-y04:



d29-x01-y05:



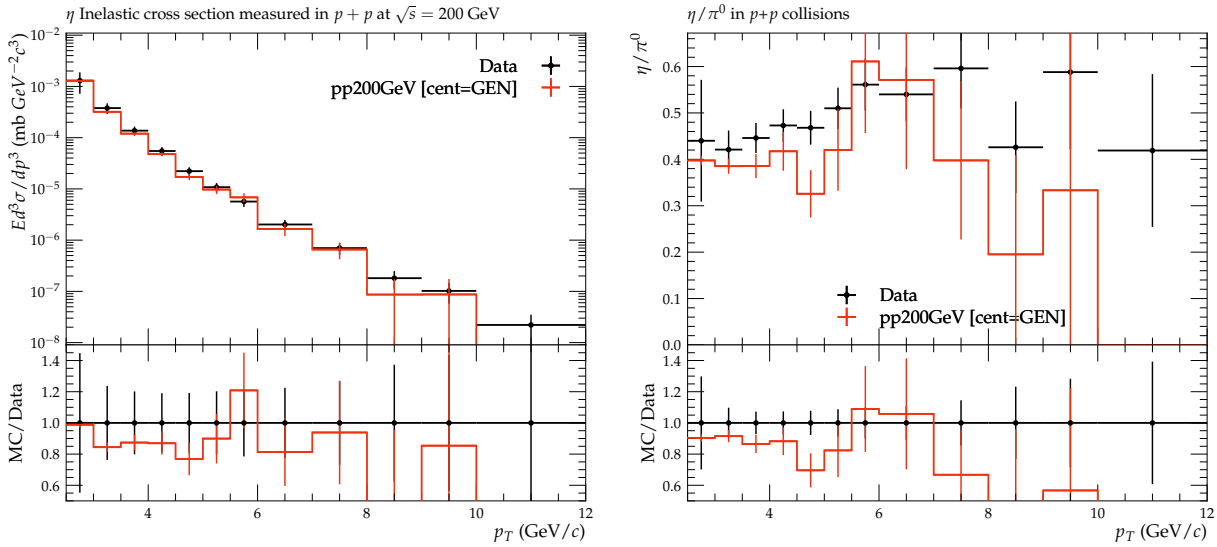
d29-x01-y06:



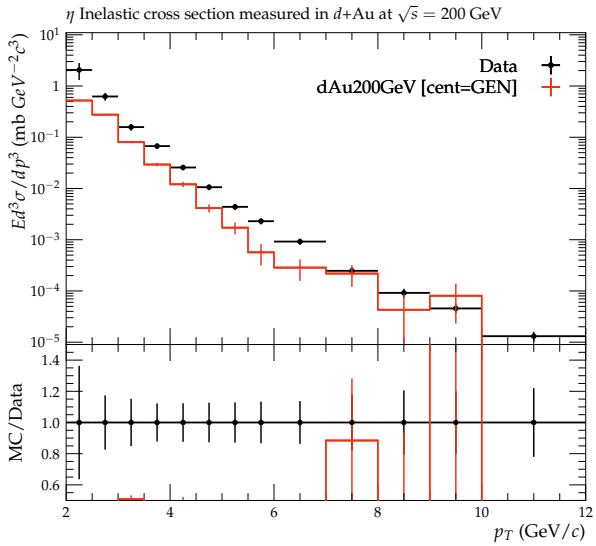
Generated at Monday, 22. June 2026 01:53PM

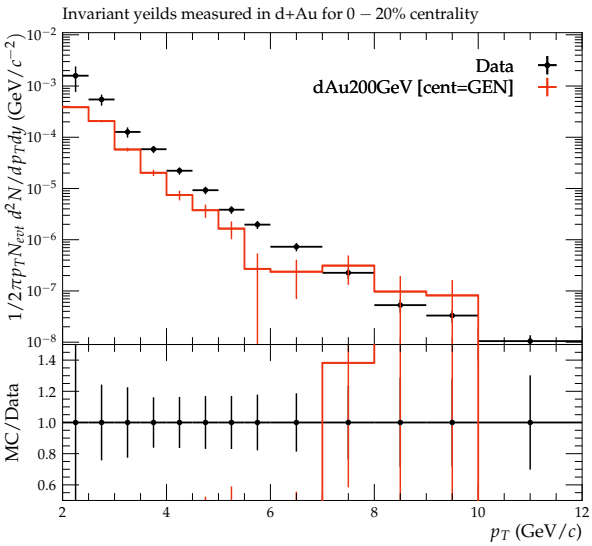
E PHENIX_2007_I731133

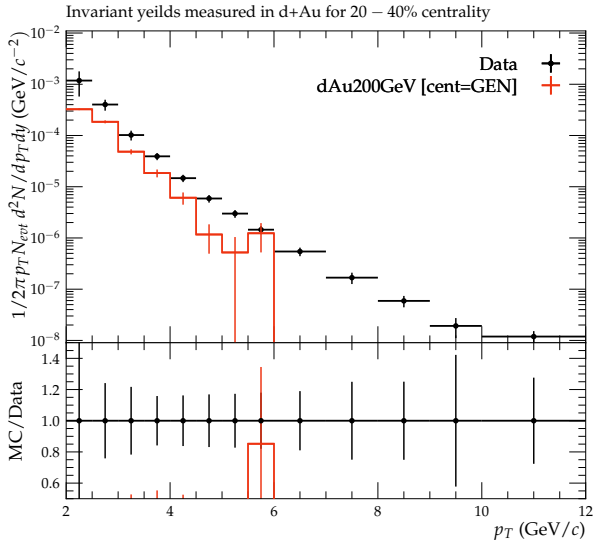
From running (rivet-mkhtml –pwd pp200geV.yoda) inside the container, the following PDF pages are produced with the comparison to simulation and data:

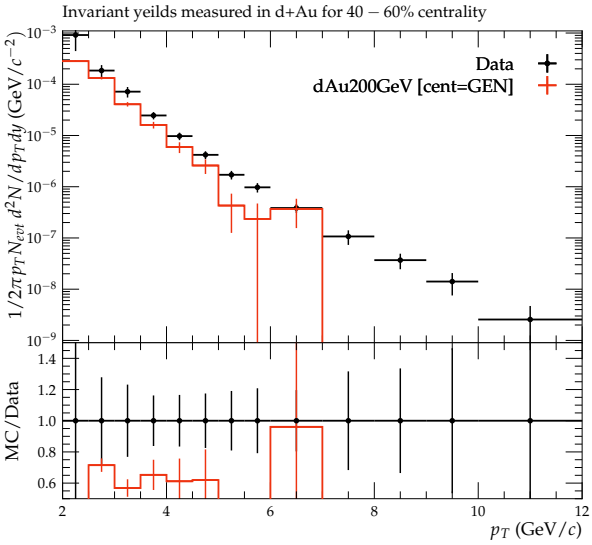


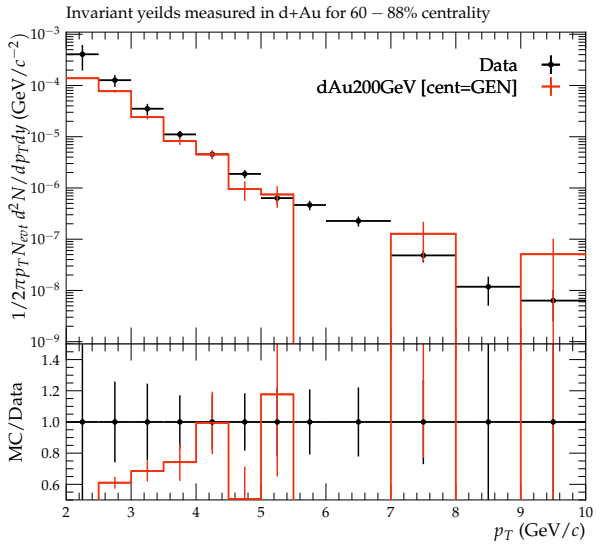
From running (rivet-mkhtml –pwd dAu200geV.yoda) inside the container, the following plots are produced with the comparison to simulation and data:

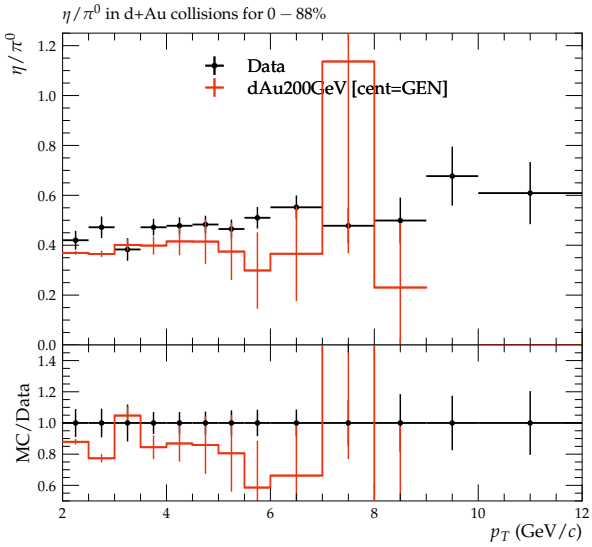

η Inelastic cross section measured in d+Au at $\sqrt{s} = 200$ GeV
Data
dAu200GeV [cent=GEN]
$Ed^3\sigma/dp^3$ (mb GeV$^{-2}c^3$)
MC/Data
p_T (GeV/c)

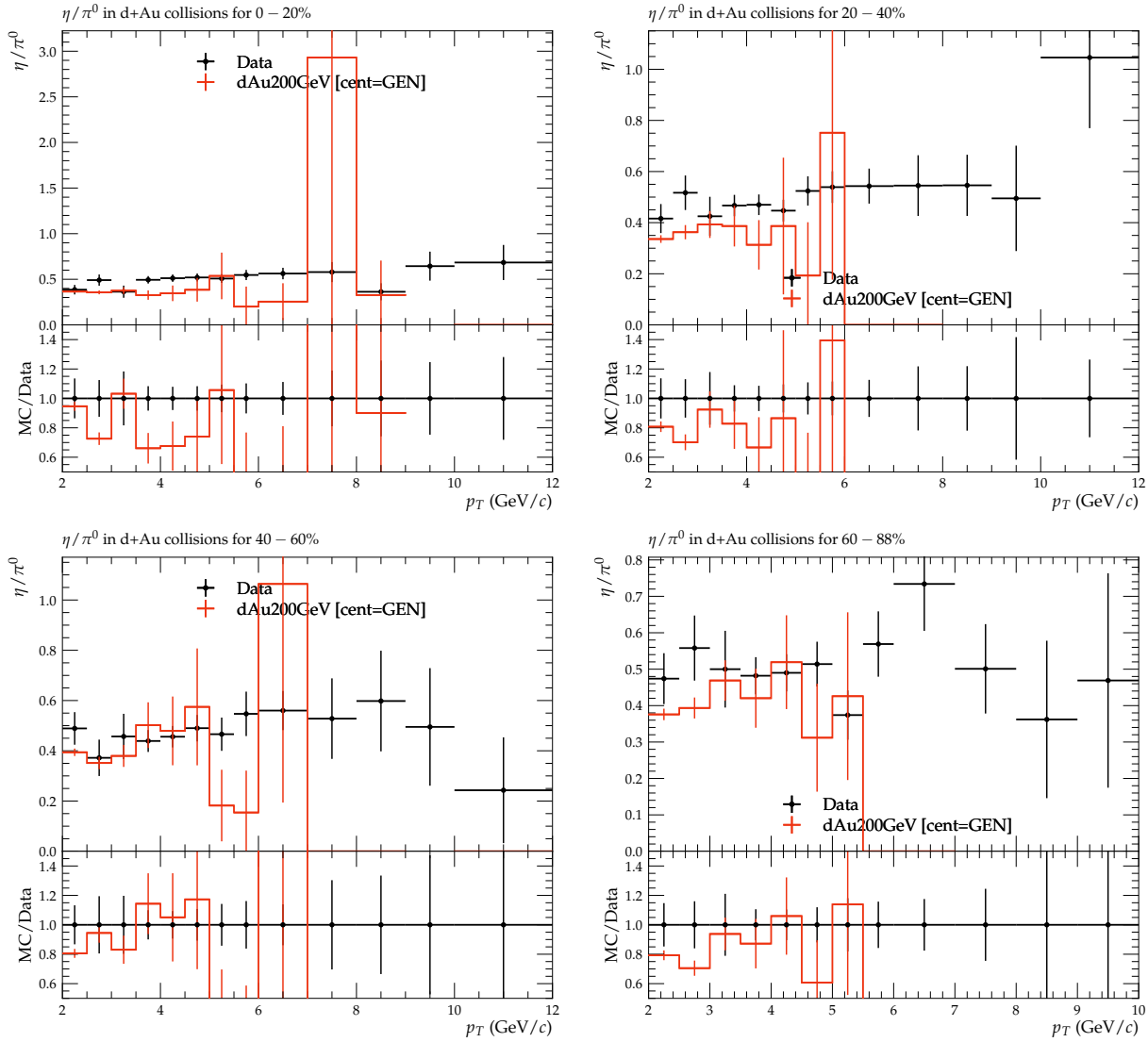

Invariant yeilds measured in d+Au for $0-20\%$ centrality
Data
dAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy$ (GeV/c^{-2})
MC/Data
p_T (GeV/c)


Invariant yeilds measured in d+Au for $20-40\%$ centrality
Data
dAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy$ (GeV/c^{-2})
MC/Data
p_T (GeV/c)

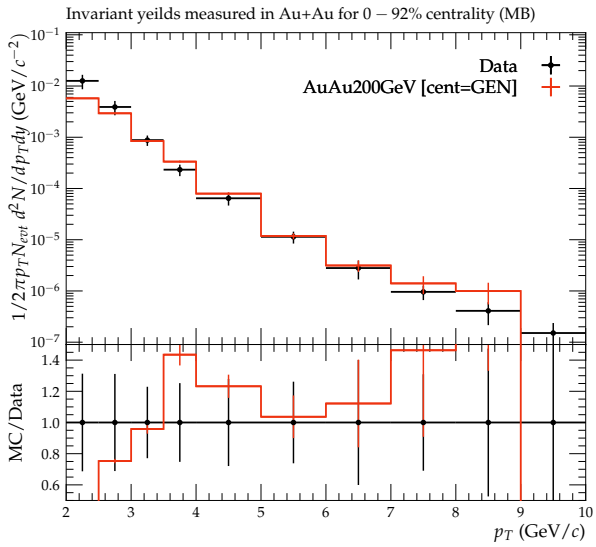

Invariant yeilds measured in d+Au for $40-60\%$ centrality
Data
dAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy$ (GeV/c^{-2})
MC/Data
p_T (GeV/c)

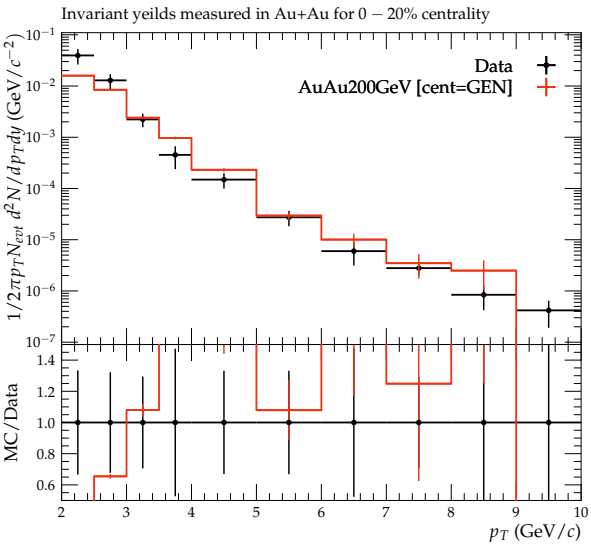

Invariant yeilds measured in d+Au for $60-88\%$ centrality
Data
dAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy$ (GeV/c^{-2})
MC/Data
p_T (GeV/c)

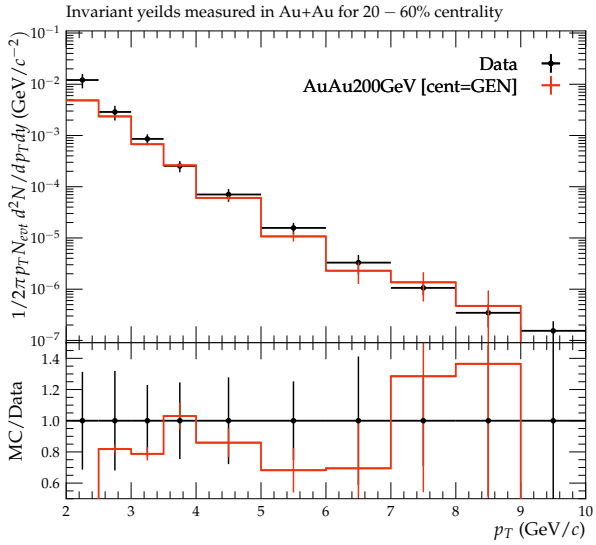

η/π^0 in d+Au collisions for $0-88\%$
Data
dAu200GeV [cent=GEN]
η/π^0
MC/Data
p_T (GeV/c)

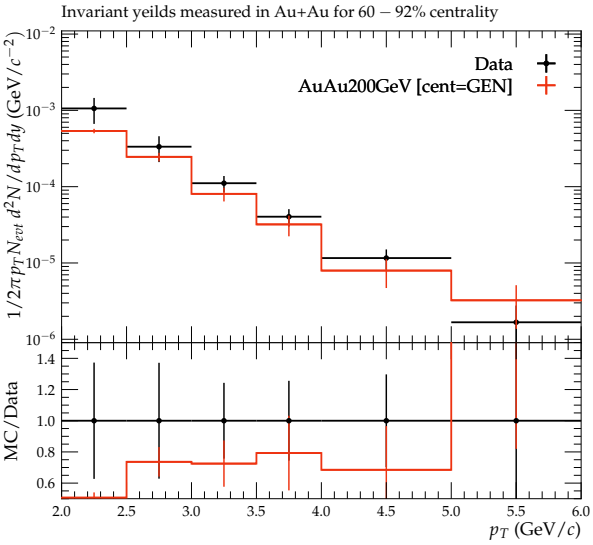


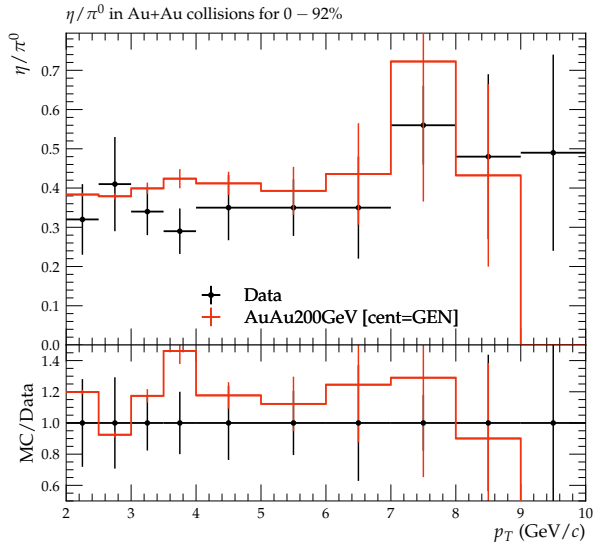
From running (rivet-mkhtml –pwd AuAu200geV.yoda) inside the container, the following plots are produced with the comparison to simulation and data:

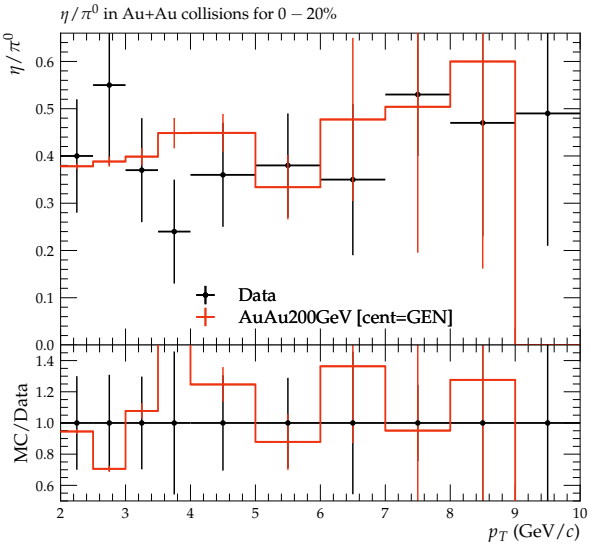

Invariant yeilds measured in Au+Au for 0 − 92% centrality (MB)
Data
AuAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy\ (\mathrm{GeV}/c^{-2})$
MC/Data
p_T (GeV/c)

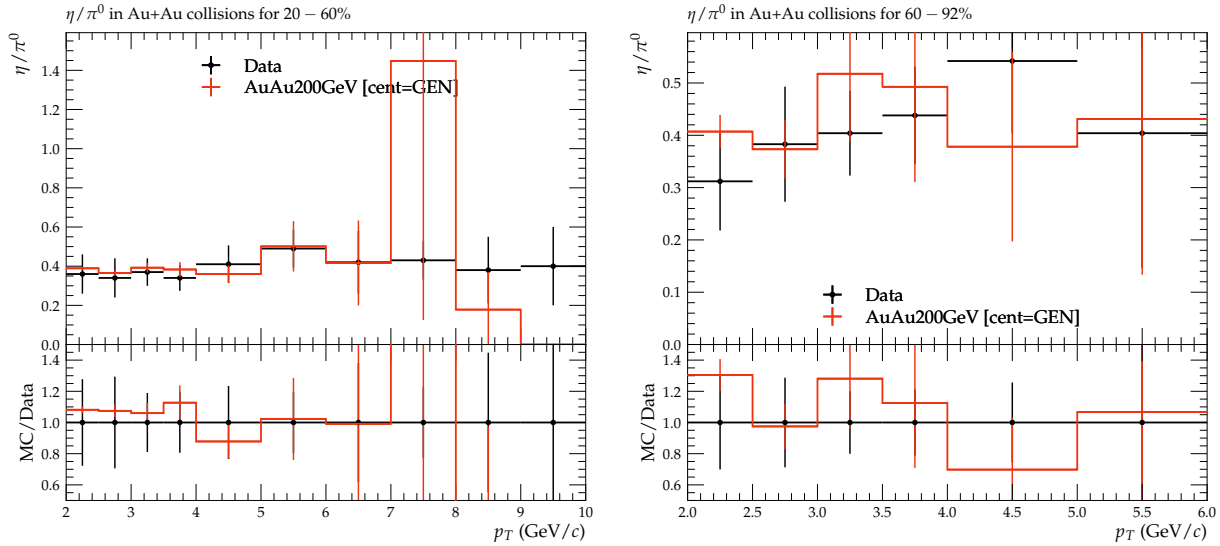

Invariant yeilds measured in Au+Au for 0 − 20% centrality
Data
AuAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy\ (\mathrm{GeV}/c^{-2})$
MC/Data
p_T (GeV/c)


Invariant yeilds measured in Au+Au for 20 − 60% centrality
Data
AuAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy\ (\mathrm{GeV}/c^{-2})$
MC/Data
p_T (GeV/c)


Invariant yeilds measured in Au+Au for 60 − 92% centrality
Data
AuAu200GeV [cent=GEN]
$1/2\pi p_T N_{evt}\, d^2N/dp_T dy\ (\mathrm{GeV}/c^{-2})$
MC/Data
p_T (GeV/c)

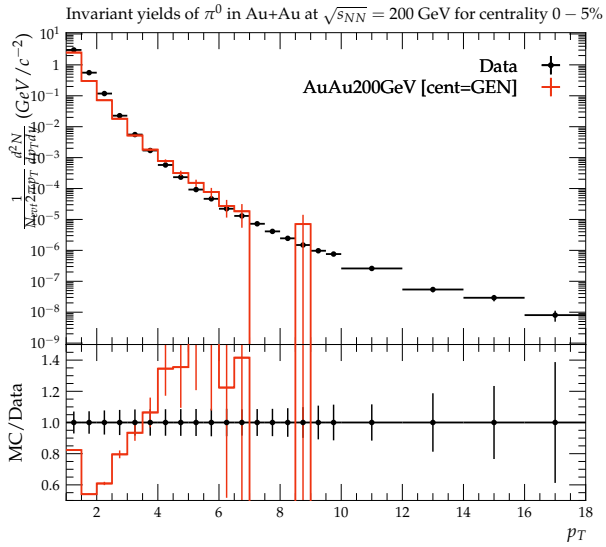

η/π^0 in Au+Au collisions for 0 − 92%
Data
AuAu200GeV [cent=GEN]
η/π^0
MC/Data
p_T (GeV/c)

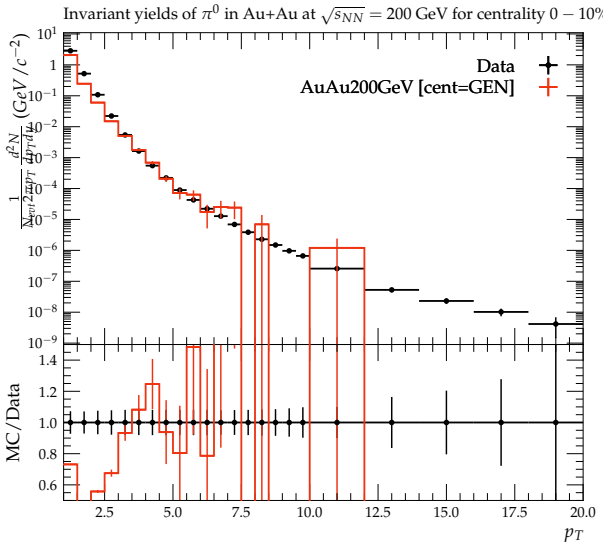

η/π^0 in Au+Au collisions for 0 − 20%
Data
AuAu200GeV [cent=GEN]
η/π^0
MC/Data
p_T (GeV/c)

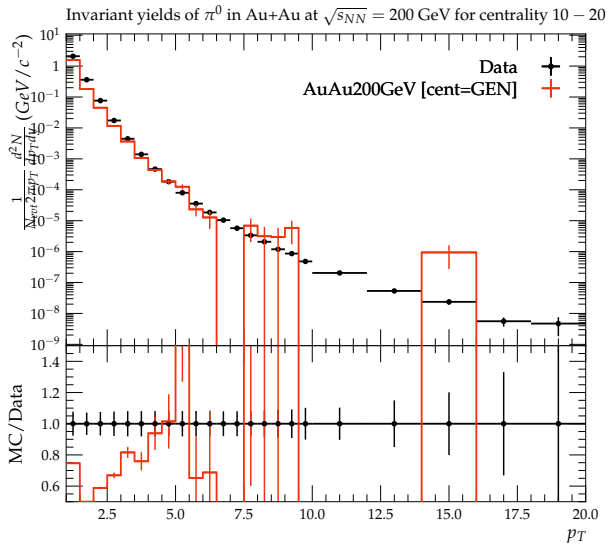


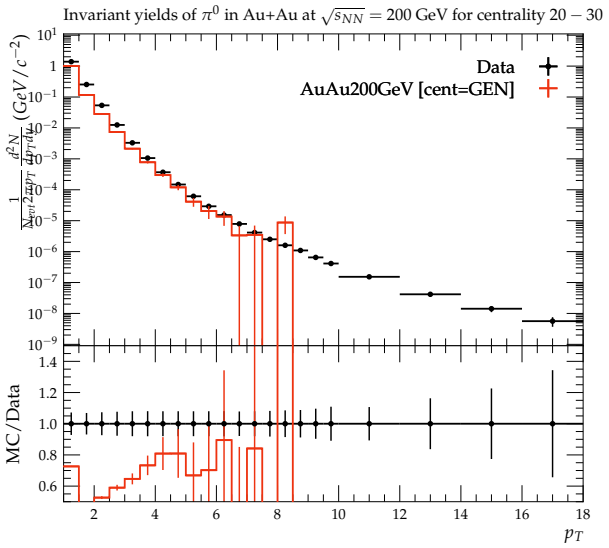
F PHENIX_2008_I778168

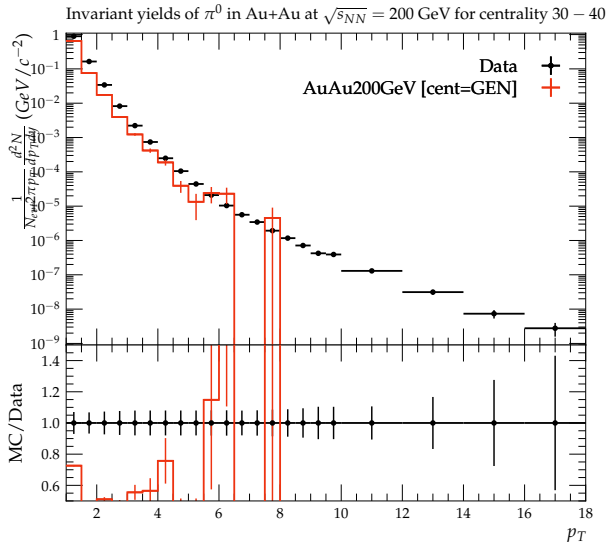
From running (rivet-mkhtml –pwd AuAu200GeV.yoda) inside the container, the following plots are produced with the comparison to simulation and data:

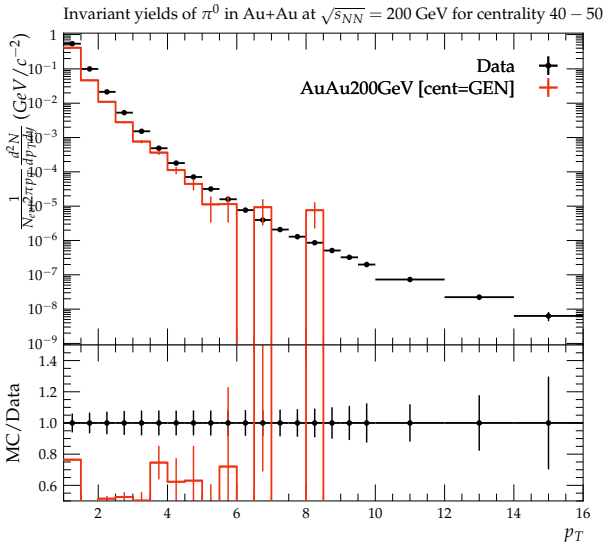

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $0 - 5\%$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

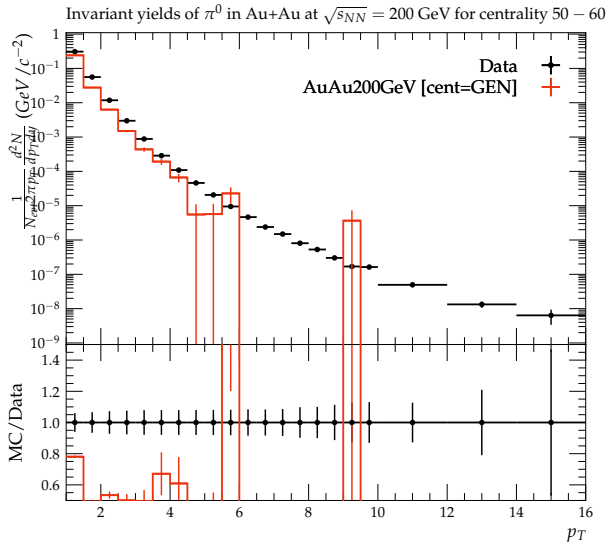

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $0 - 10\%$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

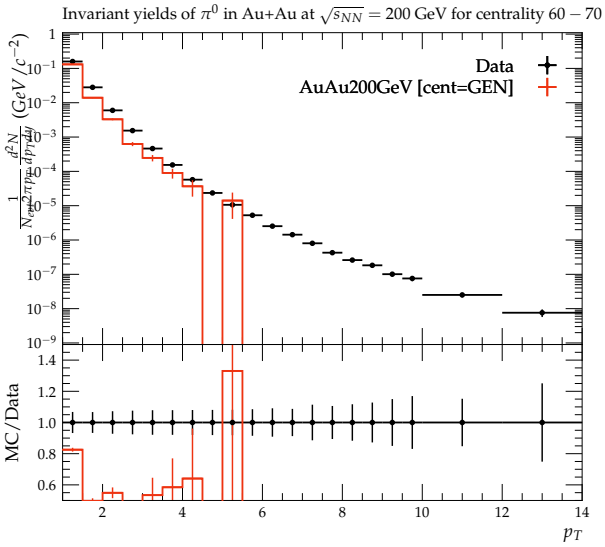

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $10 - 20$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

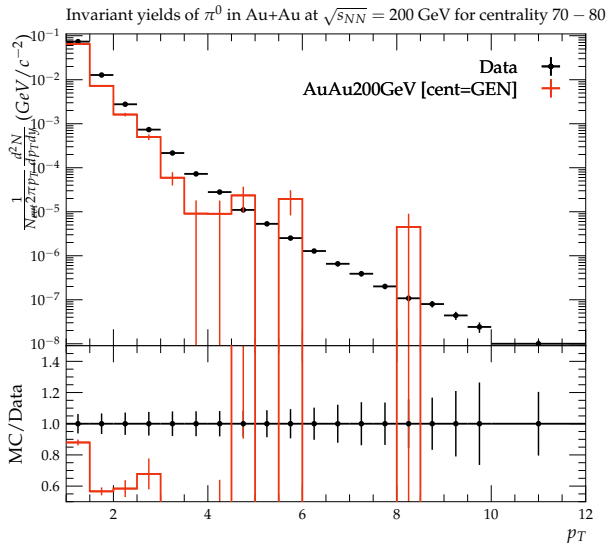

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $20 - 30$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

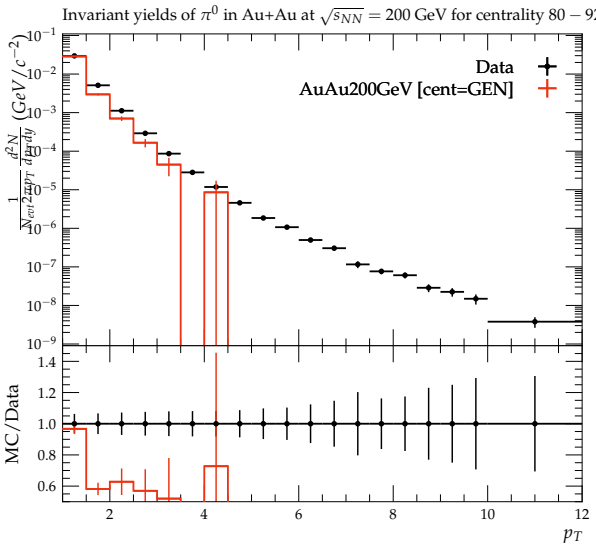

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $30 - 40$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

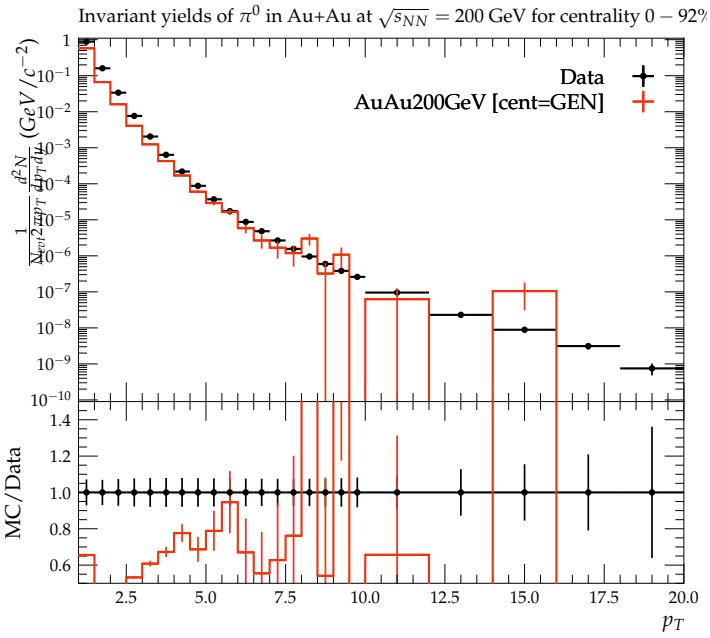

Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $40 - 50$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T


Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $50 - 60$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T


Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $60 - 70$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T


Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $70 - 80$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T


Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $80 - 92$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T


Invariant yields of π^0 in Au+Au at $\sqrt{s_{NN}} = 200$ GeV for centrality $0 - 92\%$
Data
AuAu200GeV [cent=GEN]
$\frac{1}{N_{evt}} \frac{d^2N}{2\pi p_T dp_T dy}$ (GeV/c^{-2})
MC/Data
p_T

Vita

Arthur is a 2nd-year graduate student at the University of Tennessee, Knoxville, where he received his Master's of Science in Physics in 2026. He holds a Bachelor of Science in Physics and a Minor in Math from Georgia College & State University in 2024. He also holds an Associate of Applied Science with a Diploma in Welding and Joining Technology from 2019.

He has served as a Graduate Teaching Assistant for the last two years and as a Graduate Research Assistant, working primarily on data analysis and mechanical systems. In undergrad, he fabricated parts for GCSU's linear accelerator, the Ga PEAcH. He served as the lead tutor for Math/Physic in the college's tutoring center and held a position as a machinist in the physics department's machine shop. During and before undergrad, he held several positions in the automotive and welding field, working as a mechanic for race teams and as a fabricator for an industrial site.

In November 2023, he presented a poster at the 90th Annual SESAPS meeting, presenting the updates he contributed to the Ga PEAcH. In the summer of 2022, he attended an REU where he worked with Oakland University in Michigan, where he assisted in the study of Lithium-Sulfur test cells where he presented his work at MSU's summer conference.

At the end of his undergrad, he was named 'Outstanding Physics Major' of the graduating physics majors in 2024. He was also inducted into Phi Kappa Phi in the Spring of 2024. He was named on the President's List Fall 2020 - Spring 2021, Spring 2022, and Fall 2023. Further, he was named on the Dean's List Spring 2023 and Spring 2024.