Measuring pseudorapidity distributions in Au+Au collisions at RHIC with the STAR EPD

Molnár Mátyás

Physics

Master's Thesis

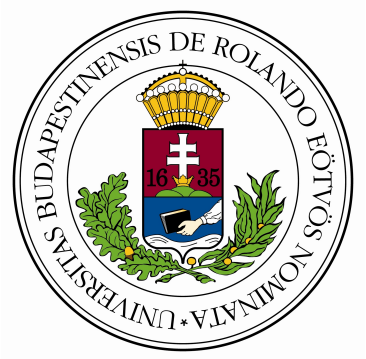
Supervisor:

Csanád Máté

Eötvös Loránd University, Faculty of Science

Department of Atomic Physics

2024



Abstract

Understanding the properties of the quark-gluon plasma (QGP) gives us a glimpse back to the early Universe. This very fact shows the importance of experiments to create and study it. Such matter, however, can exist at temperatures of – at least – about 150–200 MeV (∼2–109 K), so it can only be created in high-energy particle accelerators.
The QGP "freezes out" almost immediately, so its properties can only be inferred from the created particles, which can be measured by complex detector systems. In heavy-ion physics, it is essential to determine the dependence of this particle yield on pseudorapidity, as this is a key point for comparison with theoretical models.

In 2018, in preparation for the Beam Energy Scan phase II, the STAR detector at RHIC (Relativistic Heavy Ion Collider) was upgraded with the Event Plane Detector (EPD). This instrument, as its name suggests, enhanced STAR's capabilities in event plane resolution for flow measurements, in triggering overall, and also centrality determination for fluctuation measurements. Due to its fine radial granularity, it can also be utilized to indirectly measure pseudorapidity distributions of the produced particles.
The response of the detector to the primary particles created in the collision has to be understood well: using simulated measurement data, I used an iterative unfolding procedure to determine the relationship of EPD data to input distributions, and measured pseudorapidity distributions. I also did several systematic error checks of the method, with the motivation to develop this technique to be used at other beam energies of the RHIC. Experiments at several energies should give a more complete picture of the process at which QGP is transformed into hadronic matter, thus "scanning" the phase diagram of QCD (quantum chromodynamics) describing the strong interaction.
In my thesis, I concentrated on refining my results at $\sqrt{s_{NN}} = 19.6$ and 27.0 GeV with further systematic checks and comparing them via the total particle yield, furthermore, obtaining results at 14.5 GeV. I did some comparisons of the 19.6 GeV results with several other results at this beam energy, such as the MUFFIN-SMASH hydrodynamic model's prediction and PHOBOS experiment's data.

Contents

---

1 Introduction

1.1 High-energy heavy-ion physics

1.1.1 Nuclear physics and the Big Bang

Up until the first few microseconds after the Big Bang, an unimaginably dense and hot matter was present in the Universe, known as quark-gluon plasma (QGP), as illustrated in Fig. 1 [1]. In that hot "soup-like" matter, the quarks and gluons could exist in their free form; according to our current knowledge, this kind of matter is formed in high-energy particle collisions (with at least 150–200 MeV collision energy). However, the high energy density blows up the QGP almost instantly [2], and it freezes out to different particles, as quarks are turned into hadrons due to *colour confinement* [3]: with the growing distance between the quarks, the increasing field strength creates gluons and quark-antiquark pairs until, all particles are colourless, i.e. have zero colour charge.

Using the known conservation laws, these frozen out particles can be identified; their properties can be used to infer the various characteristics of the QGP and hence of the near-Big Bang state.

1.1.2 Pseudorapidity measurements in heavy-ion physics

From the vast amount of experimental data obtained in high-energy particle collisions, the first aim is to extract the basic physical parameters of the particles flying out from the collision centre, in order to obtain the geometrical properties of the hadronic matter after the QGP freezeout.

One of such quantities is the angle θ between the three-momentum $\mathbf{p}$ of the flying out particle and the collider beam. In experimental particle physics, however, a more convenient quantity, the pseudorapidity η is used, which is by definition:

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2}\ln\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right), \tag{1}$$

where p_z is the z component of the momentum, and the z direction is chosen to coincide with the direction of the beam [4]. This quantity is practical mostly because of two of its properties:

- in the ultrarelativistic limit ($m << |\mathbf{p}| \quad \Longrightarrow \quad E \approx |\mathbf{p}|$, where E is the energy of

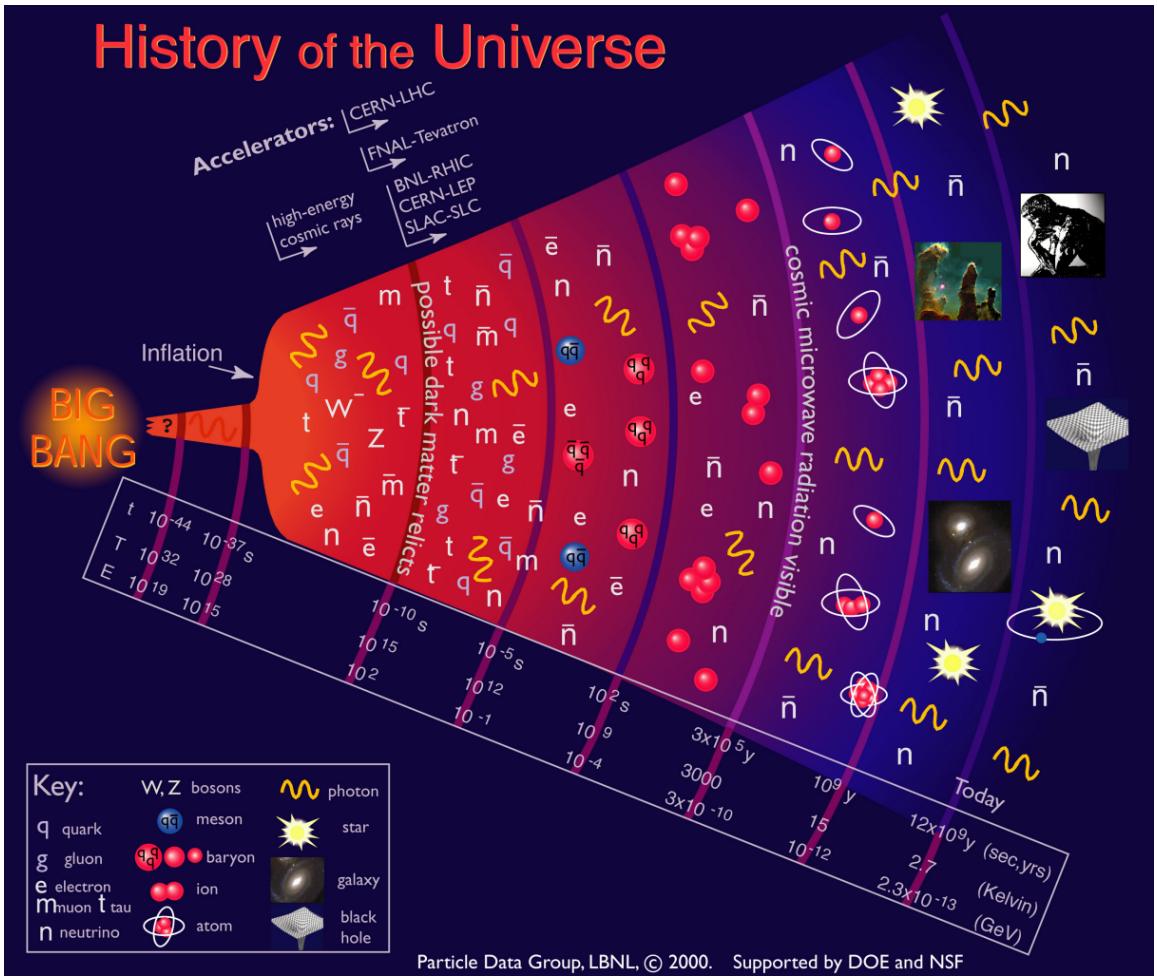


Figure 1: The story of the Universe after the Big Bang. Quarks and gluons, before being confined into bound states of hadrons after the first milliseconds, existed in the form of quark-gluon plasma. The study of this primordial matter is currently possible using methods of high-energy heavy ion physics. [1]

the particle in $c = 1$ unit system, where c is denotes the speed of light) it approaches to rapidity

$$\eta \approx y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2}$$

known from special relativity and

- in the midrapidity region (around $\eta = 0$) it gives a rather uniform particle yield as a function of rapidity in contrary to θ – as seen e.g. later in Sec. 3.4 as well in PHOBOS' results. This measure remains invariant under Lorentz transformations along the beam direction [4].

In general, measuring pseudorapidity is essential for analysing the angular distribution

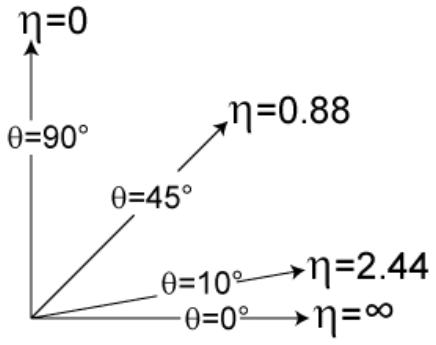


Figure 2: Illustration of pseudorapidity η and its relation to the angle θ between the three-momentum of the particle and the direction of the beam (usually chosen as axis z). Source of the picture: [5].

of particles produced in heavy-ion collisions, as it allows us to study particle production mechanisms and energy distributions in a manner that is consistent across different reference frames – particularly in high-energy environments where relativistic effects are significant.

1.1.3 Relevance to high-energy nuclear physics in general

Measuring pseudorapidity distributions is mostly relevant in the exploration of the QGP. These measurements provide indirect evidence for the existence of the QGP via revealing changes in particle production patterns (and thus, in energy distributions); furthermore, the phase transition between normal nuclear matter and QGP as predicted by Quantum Chromodynamics (QCD) [1].

Pseudorapidity distributions offer detailed information about the dynamics of heavy-ion collisions themselves as well, including the aforementioned energy deposition along the beam axis and the symmetry of the collision. Via this information, accurate models of the collision process can be constructed, and we can get closer to understanding the interplay between different physical phenomena, such as thermalization and collective flow [6].

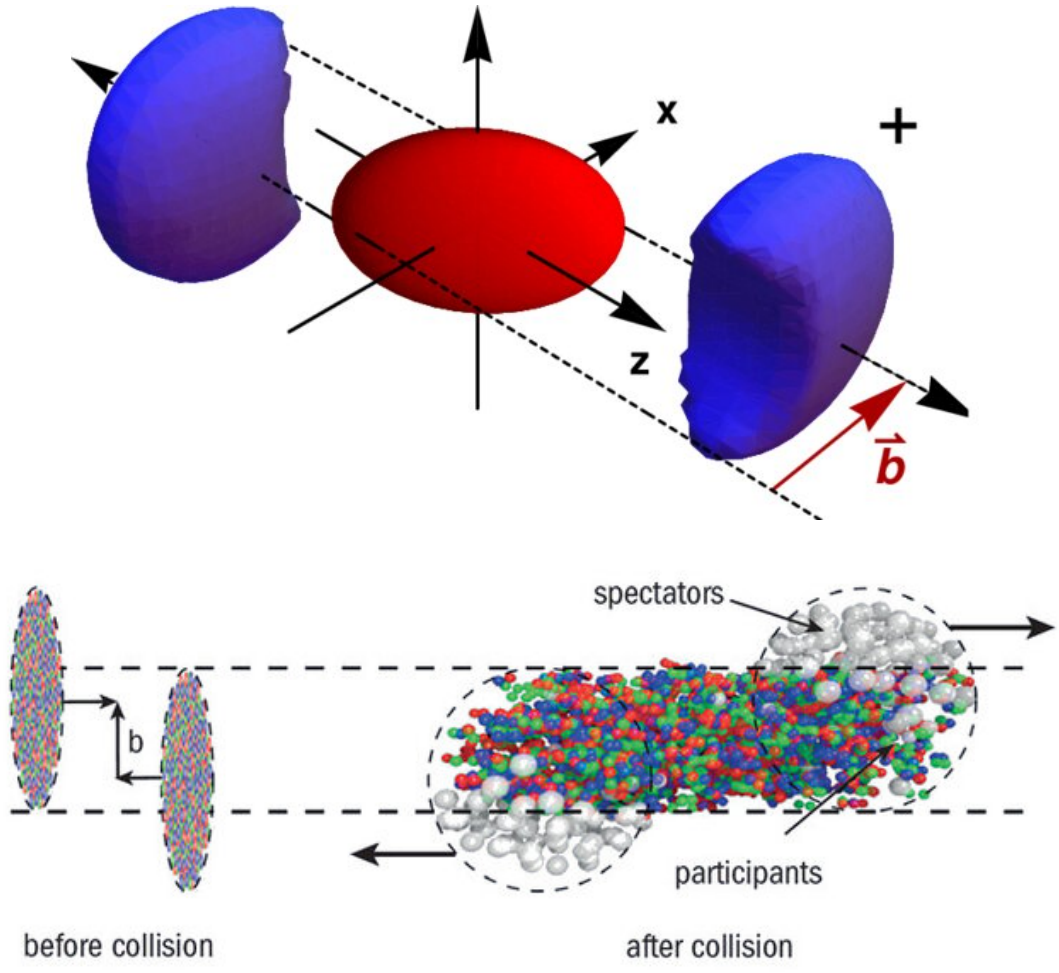


Figure 3: Simulated geometry of the collision of two symmetric nuclei. Centrality, which is the measure of how much the nuclei overlap, is specified by the impact vector $\vec{b}$. Sources: [7], [8].

1.1.4 Relevance to other observables

Centrality refers to the overlap region (or impact parameter, see Fig. 3) of two colliding heavy ions, quantifying how central or peripheral the collision is. The particles in the nuclei participating in the collision are called then *participants*
Centrality is often expressed as a percentile of the hadronic cross section corresponding to a certain energy deposited.

Particle multiplicity refers to the total number of produced particles (charged had-

[1]Quantum Chromodynamics is a theory describing the strong interaction, among other aspects, the colour confinement as well. The phase diagram describing transitions between normal baryonic (nuclear) matter and QGP (additionally, the hypothetical colour superconducting phase), referred to as the QCD phase diagram.

rons, photons, etc.) in a collision event. High multiplicity events correspond to more central collisions and vice versa low multiplicity events to more peripheral collisions, thus, it can be used to determine centrality.
In experiments like ALICE at the LHC, or in this analysis, STAR at RHIC, centrality classes are defined based on the simulated multiplicity distribution [9].

The *Glauber model* is widely used to describe the dependence of the number of participating nucleons (N_{part}) and the number of binary collisions (N_{coll}) on the impact parameter ($\mathbf{b}$) in heavy-ion collisions [10]. By coupling a Monte Carlo implementation of the Glauber model to a model of particle production (often based on a negative binomial distribution), one can predict the particle multiplicity distribution for different centrality classes [9].

Although in this case the particle multiplicity (`RefMult`) and consequently the centrality classes were already determined for the collision events in order to perform the pseudorapidity measurements, in the reverse case, pseudorapidity distributions can provide insights into particle production rates and multiplicity distributions. Such experimental results are crucial for testing theoretical models of particle production and, in overall, understanding the dynamics of the collision process.

1.1.5 Experimental techniques

Experimental setups for measuring pseudorapidity typically may involve detectors such as Time Projection Chambers (TPCs) and calorimeters. These kinds of detector systems capture the trajectories and energies of produced particles.

However, in this analysis, a single-layer scintillation detector was utilized – as for angle distribution measurement no particle trajectories nor energies are needed. This experimental setup, along with its "relative simplicity", needs special data analytical tools to achieve high precision and accuracy.

1.1.6 (Pseudo)rapidity distributions and relativistic hydrodynamics

Pseudorapidity distributions of particles produced in high-energy heavy-ion collision can be inferred from relativistic hydrodynamic models, and vice versa, measured pseudorapidity distributions can be fitted to a hydrodynamic model.

Here I will use the $c = 1$ unit system, with c being the speed of light.
Equations of relativistic hydrodynamics can be described by the continuity of conserved charges:

$$\partial_\nu (n u^\nu) = 0 \tag{3}$$

(where n is the conserved charge, u is the velocity field), and by the conservation of the energy-momentum tensor T:

$$\partial_\nu T^{\mu\nu} = 0. \tag{4}$$

Let us assume that the modelled QGP behaves as an ideal fluid (which turns out to be a very good approximation, acc. to [11]), then the energy-momentum tensor can be expanded as:

$$T^{\mu\nu} = (\epsilon + p) u^\mu u^\nu - p g^{\mu\nu}, \tag{5}$$

where ϵ is the energy density, p is the pressure, $g^{\mu\nu}$ is the metric tensor. That means, in the locally comoving frame:

$$u^\mu = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad T = \begin{pmatrix} \epsilon & & & \\ & -p & & \\ & & -p & \\ & & & -p \end{pmatrix}.$$

A possible equation of state can be

$$\epsilon = \kappa p \tag{6}$$

with a temperature-independent $\kappa = \frac{1}{c_s^2}$ constant, c_s being the speed of sound in the (near-)perfect-fluid medium.

To complete the formalism, let us use Rindler-coordinates where

- $\tau = \sqrt{t^2 - \mathbf{r}^2}$ is a coordinate proper-time with t and $\mathbf{r}$ being the components of the space-time coordinates $x^\mu = (t, \mathbf{r}) = \begin{pmatrix} t \\ r_x \\ r_y \\ r_z \end{pmatrix}$,
- $\eta_s = \frac{1}{2} \sqrt{\frac{t+|\mathbf{r}|}{t-|\mathbf{r}|}}$ is the space-time rapidity.

As cited in Reference [12], the solutions to these equations are

$$u^{\mu} = \begin{pmatrix} \cosh \lambda \eta_s \\ \sinh \lambda \eta_s \end{pmatrix} \tag{7}$$

$$n = n_f \frac{\tau_f^{\lambda}}{\tau^{\lambda}} \tag{8}$$

$$T = T_f \left(\frac{\tau_f}{\tau} \right)^{\frac{\lambda}{\kappa}}, \tag{9}$$

where λ is a parameter controlling the acceleration (if $\lambda = 1$, there is no acceleration) and f denotes quantities right at the freeze-out. If $\lambda = 2$, then κ becomes the number of dimensions. T here denotes the temperature.

The next step is the application via calculating the hadron momentum distribution. Firstly, as freeze-out conditions

- let T_f be defined as where $\eta_s = 0$,
- let the freeze-out hypersurface be pseudo-orthogonal to the velocity field (let the vector-measure of the freeze-out hypersurface $\mathrm{d}\Sigma_{\mu}(x) \,||\, u^{\mu}(x)$).

As cited in Ref. [12], in a 1+1 dimensional solution, the equation of the hypersurface will be

$$\left(\frac{\tau_f}{\tau} \right)^{1-\lambda} = \cosh\big((\lambda - 1)\eta_s\big). \tag{10}$$

The rapidity distribution will be

$$\frac{\mathrm{d}N}{\mathrm{d}y} \approx N_0 \cosh^{\frac{\alpha}{2}-1} \left(\frac{y}{\alpha} \right) e^{-\frac{m}{T_f} \cosh^{\alpha}\left(\frac{y}{\alpha}\right)}, \tag{11}$$

where N is the total number of hadrons, y is the rapidity (see Eq. 2); furthermore, $\alpha = \frac{2\lambda-1}{\lambda-1}$, m is the average mass of the hadrons, N_0 is a normalization parameter, all to be fitted with the experimental data.

The relation between rapidity and pseudorapidity distributions via a double differential and using an average transverse momentum p_t:

$$\frac{E}{p}\frac{1}{p_t}\frac{\mathrm{d}N}{\mathrm{d}p_t \mathrm{d}\eta} = \frac{1}{p_t}\frac{\mathrm{d}N}{\mathrm{d}p_t \mathrm{d}y} \tag{12}$$

$$\frac{\mathrm{d}N}{\mathrm{d}\eta} = \frac{\overline{p}_T \cosh \eta}{\sqrt{m^2 + \overline{p}_T^2 \cosh^2 \eta}} \frac{\mathrm{d}N}{\mathrm{d}y}. \tag{13}$$

Here, $\overline{p}_T$ indicates the rapidity dependency of an average transverse momentum of the particles. This can be estimated from the effective temperature (T_{eff}) of hadron spectra

using the Buda–Lund model, as cited in Ref. [12], this gives

$$\overline{p}_T = \frac{T_{\text{eff}}}{1+\sigma^2 y^2} \tag{14}$$

with

$$\sigma^2 = \frac{T_f T_{\text{eff}}}{m^2\left(\Delta y^2 + \frac{T_f}{m}\right)} \tag{15}$$

and

$$T_{\text{eff}} = T_f + \frac{m\langle u_t\rangle^2}{1+\frac{m}{T_f}}. \tag{16}$$

The model parameters at the freezeout are

- T_f describing the central temperature [2],
- $\langle u\rangle$ the average transverse flow [3]
- Δy the rapidity distribution width,

all to be determined by the fits. This approach leads to a reasonable fit, as can be seen in Fig. 4 [12].

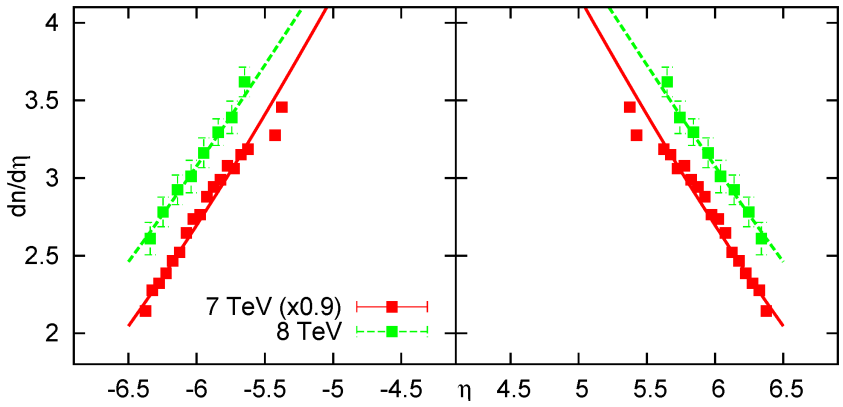


Figure 4: Charged particle pseudorapidity distributions (denoted in this figure as $\mathrm{d}n/\mathrm{d}\eta$) from the LHC TOTEM experiment, fitted with the result of the relativistic hydrodynamical model's solution. [12]

To sum up, there is plenty of motivation towards obtaining pseudorapidity distributions of the created particles in high-energy heavy-ion collisions with high precision.

[2] For a simpler fit, it could be assumed to be 170 MeV, then only deal with the rest of the model parameters.

[3] Transverse flow is the collective motion of nucleons within the "fireball" created during a high-energy heavy-ion collision. It is related to the average transverse momentum $\langle p_T\rangle = \frac{1}{N}\sum_{i=1}^{N} p_{T,i}$ [13].

1.2 Event Plane Detector at STAR

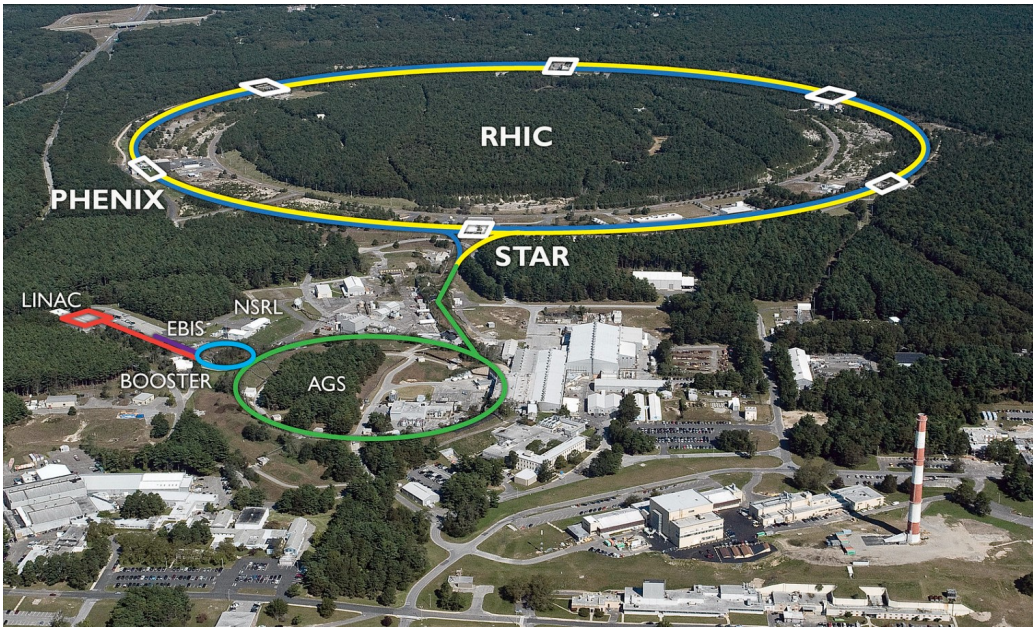


Figure 5: The RHIC facility, including the STAR experiment. Source: [14]

STAR (Solenoidal tracker, see Fig. 6) experiment's detector system is one of the facilities of the Relativistic Heavy Ion Collider (RHIC, see Fig. 5). Its primary purpose is to study the formation and several properties of the QGP.

Due to the complexity of the high-energy collisions, several simultaneous experiments are needed to draw strong conclusions about QGP. Therefore, STAR is also composed of several detectors specialised for different particles, or overall, characterising their motion.

A quantity that will be important for the measurement of pseudorapidity later on is the index of the EPD *ring*: this corresponds to the number of tiles at the same distance from the beam axis. For the axial angle θ measurement that is directly related to the pseudorapidity, this will be the important value to extract from raw EPD data.
For a ring on one side (`Side=-1`, east side), I denoted from 0 to 15 from the centre outwards, while on the other side (`Side=1`, west side) it is denoted from 16 to 31 from the outside inwards. Thus, the pseudorapidity as a function of the ring index number is as "continuous" as possible as a function of pseudorapidity (of course, the ring numbers are discrete values, and also there is a jump between index 15 and 16, where the pseudorapidity effectively changes sign).

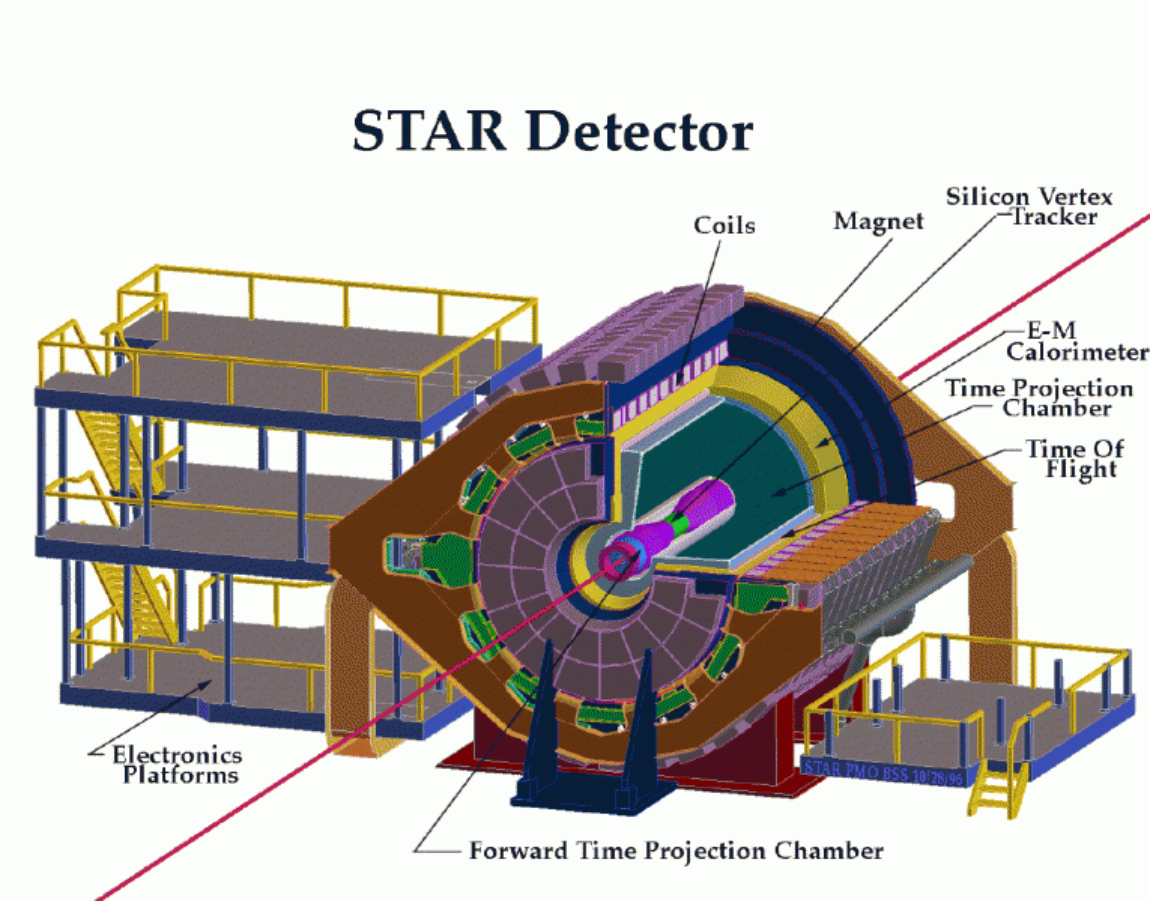


Figure 6: An earlier schematic of the STAR experiment; the EPD detector is not yet drawn. Image source: [15]

The Event Plane Detector (EPD) is one of the latest upgrades as a newly installed detector of the STAR experiment. It was installed in 2018 as part of the preparations for the BES-II program. The motivations for building it included enhancing the event plane resolution for flow measurements, providing independent centrality determination for fluctuation measurements, and serving as a trigger in the high-luminosity environment during the BES-II program.

In simple terms, its purpose is to measure the pattern of forward scattering charged particles emitted during high-energy collisions.
Determining the initial geometry of each collision is a critical step in understanding the complex physics of the system. Experimentally, this geometry is quantified in the form of *event planes*, for which well-established methods exist. In this thesis, I will not discuss this quantity in detail.

1.2.1 Structure of the EPD

The detector consists of two highly segmented discs ("wheels") with 1.2 cm thick scintillator plates connected via wavelength-shifting optical fibres to the silicon photoelectron multipliers.

The schematic structure of a wheel is shown in Fig. 7.

Each wheel consists of 12 "supersectors" covering an azimuthal angle of 30°, divided into 31 additional tiles. All the supersectors deliver light to a silicon photoelectron multiplier (SiPM) via a bundle of 31 optical fibres; these signals are amplified and sent to the STAR digitizing and data acquisition system (STAR DAQ). [16]

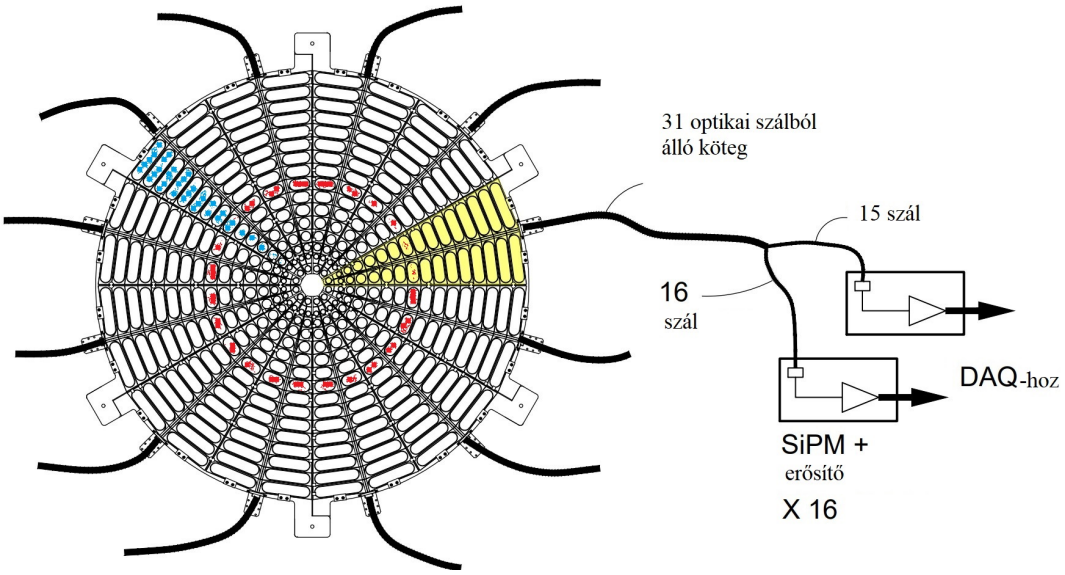


Figure 7: Schematic of the EPD structure. One of the two EPD "wheels" is shown. The 31 tiles from each of the 12 supersectors are connected via bundles of optical fibres to the silicon photoelectron multipliers and amplifier electronics. A single *EPD ring* is highlighted in red, a single *EPD supersector* in yellow. Source: [16]

The determination of the types of particles (*particle identification*, PID) and whether these particles originate from the original collision nodes (vertex) or are possibly daughter particles born from their decay is irrelevant for the definition of the event plane. Although the trajectory of charged particles is deflected in the 0.5 T magnetic field utilized at the STAR experiment, simulations suggest that this does not significantly affect the determination of the event plane.

As a consequence, the EPD can neither determine the type of particles nor track their

trajectories. [16] This property of the detector system will be relevant in how it can be used to measure the pseudorapidity distribution of charged particles originating in the collision vertex.

1.2.2 Geometry of the EPD

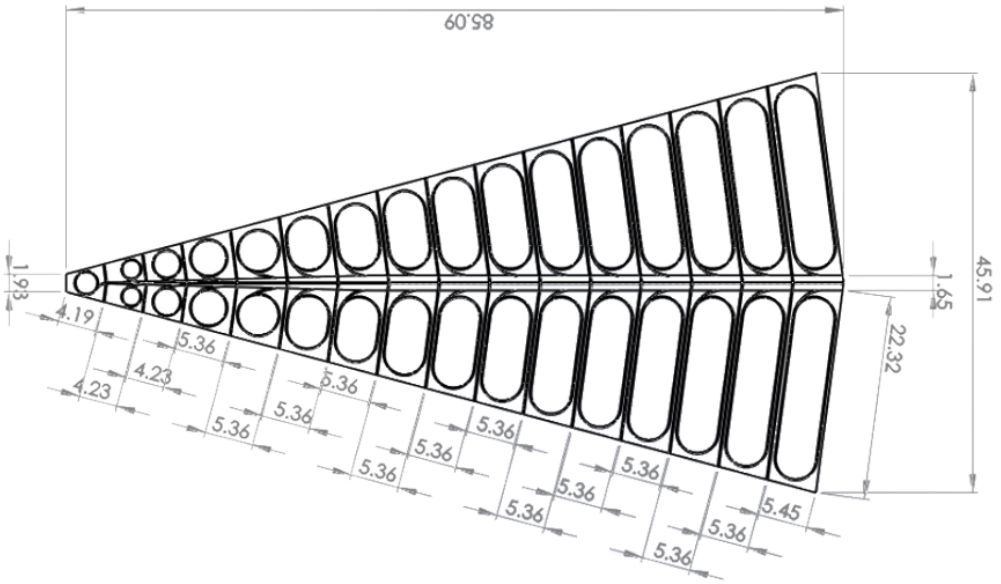


Figure 8: Detailed diagram of a supersector of the EPD; dimensions in centimeters.

Since the event planes are determined by the Fourier components of the azimuthal dependence of the particle yield [17], the EPD also shows azimuthal symmetry. Due to the placement of the rest of the detectors of the STAR experiment, the EPD is mounted at $z = \pm 3.75$m position.

The EPD is capable of detecting forward and backward scattered particles, in a range of

$$0.7^\circ < \theta < 13.5^\circ$$

angle between the direction of the impacting particle and the collider beam. In pseudorapidity terms, it means a

$$2.14 < |\eta| < 5.09$$

acceptance.

The EPD also has its predecessor, Beam–Beam Counter (BBC) with much less fine granularity than the EPD: only 36 tiles, with the 18 inner smaller tiles used for the same

purpose – compared to the 372 tiles of the EPD [16]. It also has smaller acceptance of $3.3 < |\eta| < 5.0$ in pseudorapidity [18]. Thus, the EPD significantly improved the event plane resolution – for example, by about a factor of 2 in Au + Au collisions at $\sqrt{s_{NN}} = 19.6$ GeV [4].

When designing the detector segmentation, several requirements in terms of measurement performance had to be taken into account, and the tile sizes were optimised accordingly.
In the end, based on previous measurements [16], double impacts occur with a probability of about 10% on tiles at any radial size, measured in central Au + Au collisions at energy $\sqrt{s_{NN}} = 19.6$ GeV.
The impact of these effects should be determined from simulations for the online trigger and, similarly, in offline analyses to be taken into account accordingly.

1.2.3 Calibrated ADC of the EPD electronics

At the end of the STAR DAQ, the raw data is simply the digitized electronic signal of scintillations, therefore firstly one has to ensure its relation to the real particle hits.

Each EPD tile detects hits, primarily from Minimum Ionizing Particles (MIPs). Assuming that the probability distribution of a single hit's measured signal follows a Landau distribution, multiple hits will produce a convolution of several Landau distributions.

The observed Analog Digital Count (ADC) distributions were fitted using a so-called multi-MIP Landau function, as illustrated in Fig. 9. The properly chosen Landau distributions, representing the ADC contributions from n number of MIPs, were convolved with different weights (n-MIP weights).

The conclusion was that using fewer than 5 n-MIP weights is sufficient for a good fit, as the contribution from the 5-MIP weight was already zero within uncertainties. This was based on the assumption, validated during data analysis, that the MIP weights followed a Poisson distribution.

Given this finding, the systematic uncertainty from fitting only up to 5 n-MIP weights is considered negligible [19]. Further systematic uncertainties will be discussed in Sec. 3.2.

[4] $\sqrt{s_{NN}}$: centre-of-mass energy per nucleon in a nucleon-nucleon collision.

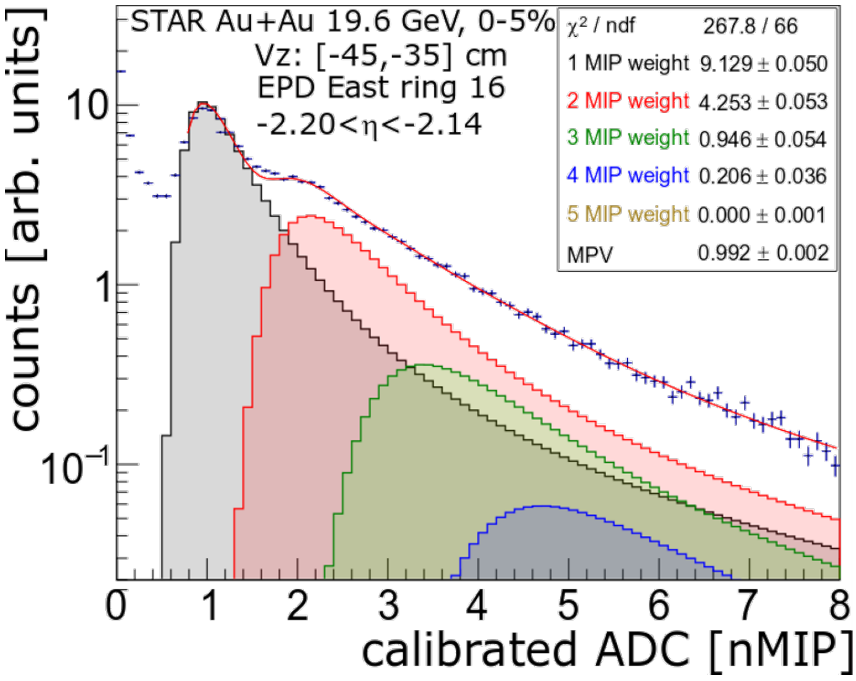


Figure 9: An example of a multi-MIP Landau fit for the ADC count distribution in ring #16, with ADC counts given in arbitrary units. The blue points with error bars depict the data, while the red continuous line represents the fitted function. [19]

2 Analysis

During the whole analysis progress, including the Monte Carlo simulation, the unfolding procedure and the plotting, I used the `ROOT 5.34/38 [20]` *open-source data analysis framework. Although a slightly older version at the time of writing the thesis, this was the version readily available on STAR's servers.*

2.1 Raw EPD data to charged particle pseudorapidity measurements

2.1.1 Extracting EPD Ring hits

The first step is to extract the N_{hits} [5] per EPD tile in each (triggered and thus recorded) collision event.

The raw data at STAR is stored PicoDst datafiles. This format is designed for general physics analysis, similar to a ROOT [20] tree, derived from the STAR standard

[5] Number of hits

MuDst [21], [22] data. Then, as mentioned in Sec. 1.2.3, the next step is to fit the data with convoluted Landau distributions.

In order to continue in obtaining the pseudorapidity distributions, several other parameters are needed as well, mostly different for the specific particle accelerator runs (as they depend on accelerator energy, the current experiment setup's fine-tuned parameters etc.). For the sake of illustration, a few examples of these are:

- relevant (minimum bias) triggers,
- `RefMult` cuts (see Sec. 1.1.4),
- centrality definition of the collision events.

As these are mostly very technical parameters, in this thesis I will not get into details.

The conversion from EPD tile index to Ring number is very straightforward; for illustration, in the case of the Monte Carlo simulation, a C++ function like this handles the conversion of a `StMuEpdHit` class' member to an integer EPD Ring number:

```
int get_ring(StMuEpdHit* epdhit){
   int ring_no = epdhit->tile()/2;
   if(epdhit->side()>0)
   {
     ring_no = 31-ring_no;
   }
   cerr << "Current tile number: "<< epdhit->tile() << endl;
   cerr << "Current ring number: "<< ring_no << endl;
   return ring_no;
}
```

A 2D heatmap visualisation of such a datafile with $N(i_{\text{Ring}})$ distribution ($\mathrm{d}N/\mathrm{d}i_{\text{Ring}}$) can be seen in Fig. 10.

2.1.2 Detector distortions

As stated before, in addition to event plane determination, the EPD's fine radial granularity theoretically makes pseudorapidity measurements possible. That means, the raw N_{hits} numbers from the EPD can be used to calculate the pseudorapidity distribution of

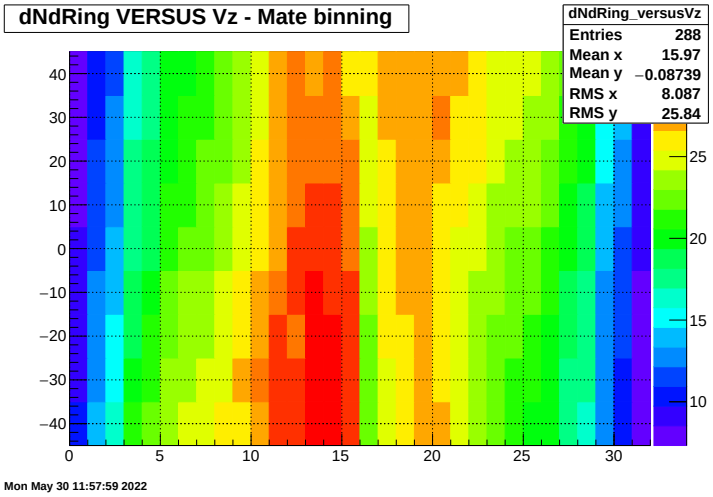


Figure 10: Visualisation of an EPD datafile with $N(i_{\text{Ring}})$ distribution.

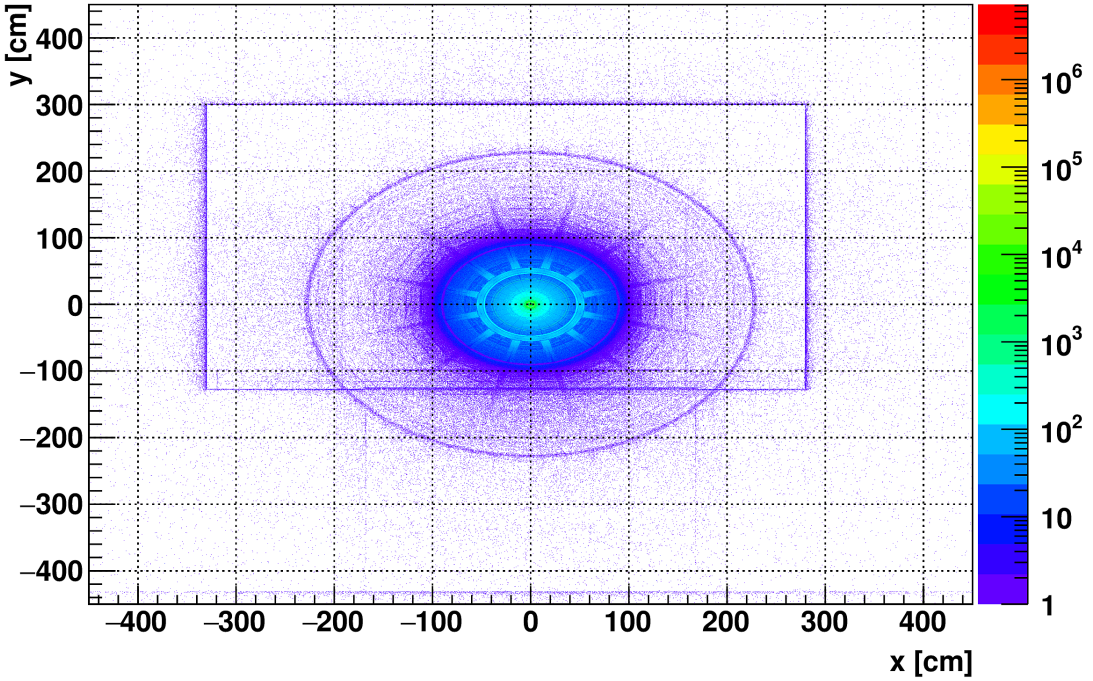


Figure 11: Vertices of particles detected by the EPD, based on a `HIJING` [23] + `Geant4` [24] Monte Carlo detector simulation. The plots display the vertex distribution in the x–y plane, integrated along the z axis, revealing the detector structure and different materials surrounding it.

charged particles ($derN_{ch}/\mathrm{d}\eta$) by using the η value corresponding to each ring.

However, this kind of simple measurement also includes secondary particles that do not originate from the primary vertex. As the EPD is located behind the rest of the

detector system and far from the interaction point, various factors distort the measured distribution.

The main factors causing significant distortion are:

- charged primary particles scatter in the detector material (or occasionally even on each other), producing secondary particles that add a significant contribution to $derN_{ch}/\mathrm{d}\eta$ calculated from a naive measurement. This is demonstrated in Fig. 11, where the vertices (origins) of particles causing hits in the EPD are depicted. Of course, such data comes from a Monte Carlo (MC) simulation of the detector material.
- Neutral primary particles also may contribute through decays (e.g., a neutral Λ baryon decaying into a proton and a pion). Data shown in Fig. 12 demonstrates that this contribution is considerable, based on the same simulation with PID analysis.

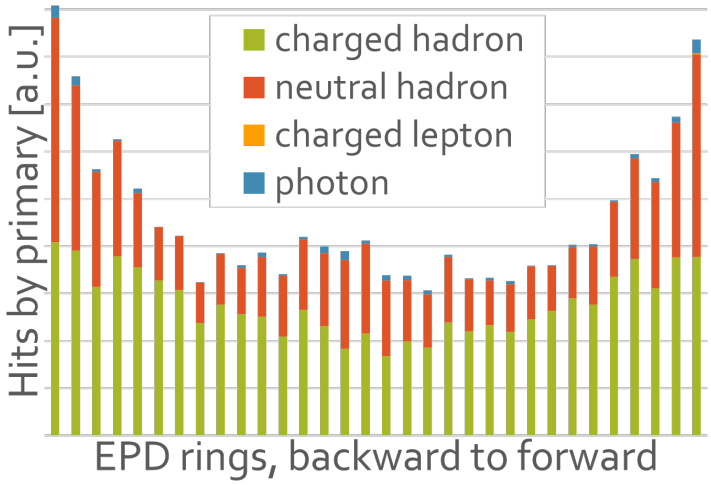


Figure 12: Distribution of various simulated primary particles hitting the EPD, shown ring-by-ring. The rings in the backward direction are on the left side of the panel, while the rings in the forward direction are on the right side – ordered by the apparent spatial rapidity of each ring. [19]

Due to these distortions, the analysis process of the experimental data gets more complicated even in theory.

Let us denote the vector storing EPD hits of given events $\mathbf{N}$. Its elements: $N(i_{\mathrm{ring}})_i$ let us index with Ring numbers $i \in \{0, 1, ..., 31\}$.

Let us Formally denote the number of primary particles created during the collision,

separated into discrete values, falling within a given pseudorapidity range $\left(\frac{\mathrm{d}N}{\mathrm{d}\eta}\right)_j$. Assuming a linear dependence:

$$N(i_{\mathrm{ring}})_i = \sum_j R_{ij}(\eta_j, i_{\mathrm{ring}}) \left(\frac{\mathrm{d}N}{\mathrm{d}\eta}\right)_j \tag{17}$$

This R_{ij} matrix is not an identity matrix but a so-called *response matrix* [6]. In this analysis, it contains the number of hits in the given ring number distribution's bin, originating from a particle at the given η pseudorapidity distribution's bin. In a continuous case [7], it formally works as a convolution:

$$\mathrm{d}N(i_{\mathrm{ring}}) = \int R(\eta, i_{\mathrm{ring}}) \left(\frac{\mathrm{d}N}{\mathrm{d}\eta}\right) \mathrm{d}\eta. \tag{18}$$

The **R** matrix needs to be determined via Monte Carlo simulations, then the $\frac{\mathrm{d}N}{\mathrm{d}\eta}$ distribution can be unfolded using the measured EDP ring hits $N(i_{\mathrm{Ring}})$ distribution [8].

2.1.3 Response matrix

For this analysis, events were generated using the STAR `HIJING` Monte Carlo event generator, combined with `Geant4` to accurately simulate the geometry of the EPD.

Note that no (light) ion fragments can be simulated in `HIJING`. These kinds of flying out particles are, in reality, inevitable with heavy-ion collisions. However, this shortfall should not change the results significantly. According to PHOBOS results [25], the contribution from light ion fragments causes at least an order of magnitude smaller contribution to measured pseudorapidity distributions than the resulting ones in this analysis, without light ion fragment simulation (see Sec. 3.2).

Such a response matrix can be seen in Fig. 13 obtained for all particles (both charged and neutral). The visualisation was done using `HresponseNoOverflow()` procedure of the `TH2` histogram function class of `ROOT`.

[6] The response matrix encodes the response of the detector, i.e., connects a detector-level distribution with the true distribution to be measured [19].

[7] The pseudorapidity distribution is a continuous quantity anyway, but in numerical calculations we work with binned distributions.

[8] In order to obtain a physically meaningful result, naturally, the Ring hit distribution also needs to be normalized with the number of events processed

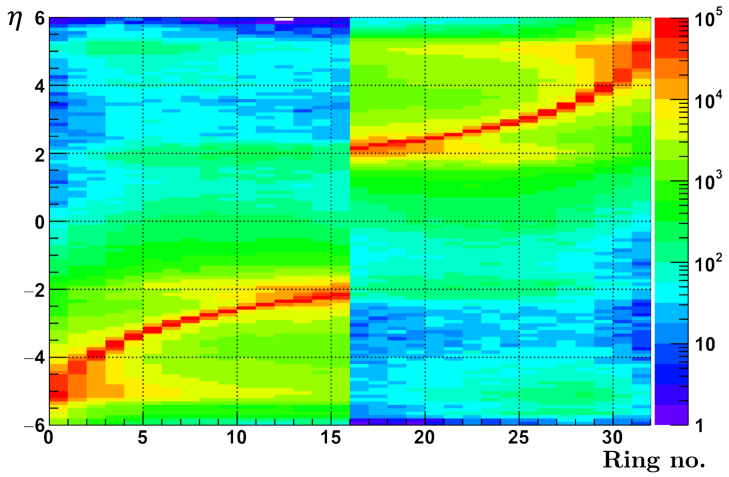


Figure 13: Heatmap of the R response matrix, connecting bins containing numbers of EPD ring hits (caused by either primary or secondary particles) with bins corresponding to primary particles at given η pseudorapidity. The left side (0–15 or 1–16 depending on the indexing of the bin borders) corresponds to East EPD wheel, the right side (16–31 or 17–32) to West EPD wheel. It is worth noting that many primaries create hits even in the opposite side EPD via secondary particles, as seen in upper left and bottom right quarters.

2.1.4 Iterative Bayesian unfolding [26], [27]

As a matter of fact, in experiments, the distribution of measured, observable quantities is different from the corresponding "real" physical quantities, due to physical effects and imperfections in the detectors.

For a single physical variable, it is enough to use *bin-to-bin* correction, via a *general efficiency* [9] determined from Monte Carlo simulation. This kind of correction does not take into account correlation between bins, thus, it is only valid if the migration between bins is negligible and the smearing is less than the width of the bin.

Another option is using a response matrix already discussed in Sec.2.1.2.
However, no probabilistic consideration guarantees this matrix to be invertible; therefore, a simple (or might be that not even a regularized) matrix inversion to extract the real physical quantities (such as pseudorapidity distributions as formally described in Eq. 18) might not be an option even if the exact form of R would be known.

[9] Here we calculate the ratio of the number of events in a given bin for the reconstructed variable versus the number of events in the same bin for the real variable.

In this analysis, a method called Bayesian iterative unfolding was used. It is described in detail in Ref. [26] and Ref. [27] on which this section is based, here I list the main points of its theoretical background.

Let C_i $(i = 1, 2, ..., n_C)$ be independent causes causing effect E. Let $P(C_i)$ be the initial probability of the causes, $P(E|C_i)$ the conditional probability of i-th to cause the effect.
Then, according to the Bayes theorem:

$$P(C_i|E) = \frac{P(E|C_i) \cdot P(C_i)}{\sum_{l=1}^{n_C} P(E|C_l) \cdot P(C_l)}. \tag{19}$$

$P(C_i|E)$ depends on the initial probability of the causes, as the final distribution also depends on $P(E|C_i)$. These probabilities need to be determined via Monte Carlo simulations.

If there is an effect E and we are observing $n(E)$ events, the expectation value of the number of events attributable to each cause:

$$\langle n(C_i) \rangle = n(E) \cdot P(C_i|E) \tag{20}$$

In the case of all the possible E_j effect $(j = 1, 2, ..., n_E)$ of a C_i cause the Bayes formula needs to be evaluated [10]:

$$P(C_i|E_j) = \frac{P(E_j|C_i) \cdot P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l) \cdot P_0(C_l)}, \tag{21}$$

where $P_0(C_i)$ is the initial probability of the different cases. Normalization is also needed:

$$\sum_{i=1}^{n_C} P_0(C_i) = 1 \quad \text{és} \quad \sum_{i=1}^{n_C} P(C_i|E_j) = 1 \tag{22}$$

Also, the following condition must be fulfilled:

$$0 \leq \varepsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|E_i) \leq 1 \tag{23}$$

$P(E_j|C_i)$ probabilities are formally the $\mathbf{R}$ *response matrix*.

After N_{obs} experimental observations the obtained distribution comes as $\mathbf{n}(E) = \{n(E_1), n(E_2), ..., n(E_{n_E})\}$. These can be calculated via an assignment as follows:

$$\langle n(C_i)|_{obs} \rangle = \sum_{j=1}^{n_E} n(E_j) \cdot P(C_i|E_j). \tag{24}$$

[10] This is also called the creation of a $\mathbf{S}$ smearing matrix.

Taking ε_i efficiency into account, the expectation value of the real event number is estimated as:

$$\langle n(C_i)\rangle = \frac{1}{\varepsilon_i}\sum_{j=1}^{n_E} n(E_j)\cdot P(C_i|E_j) \quad \wedge \quad \varepsilon_i \neq 0 \tag{25}$$

The true total number of events can be approximated from these unfolded events:

$$\langle N_{true}\rangle = \sum_{i=1}^{n_C}\langle n(C_i)\rangle.$$

The final probability of the causes:

$$\langle P(C_i)\rangle \equiv P(C_i|n(E)) = \frac{\langle n(C_i)\rangle}{\langle N_{true}\rangle}.$$

The final efficiency:

$$\langle \varepsilon\rangle = \frac{N_{obs}}{\langle N_{true}\rangle},$$

which may differ from the efficiency of the generated and reconstructed Monte Carlo events ε_0:

$$\varepsilon_0 = \frac{N_{\text{reconstructed}}}{\langle N_{\text{generated}}\rangle} = \frac{\sum_{i=1}^{n_C}\varepsilon_i P_0(C_i)}{\sum_{i=1}^{n_C} P_0(C_i)}.$$

Therefore, the unfolding is done iteratively:

1. choose an initial $\mathbf{P_0}(C)$ with the closest possible resemblance of the studied process;
2. compute $\langle \mathbf{n}(C)\rangle$ and $\langle \mathbf{P}(C)\rangle$;
3. perform χ^2 test between $\langle \mathbf{n}(C)\rangle$ and $\langle \mathbf{n_0}(C)\rangle$;
4. substitute $\mathbf{P_0}(C)$ with $\langle \mathbf{P}(C)\rangle$, $\mathbf{n_0}(C)$ with $\langle \mathbf{n}(C)\rangle$;
5. iterate again from point 2 until χ^2 gets "low enoug".

In this analysis, 4 iterations [11] were proven to be enough for all unfolding procedures performed.

Introducing the unfolding matrix

$$M_{ij} = \frac{P(E_j|C_i)\cdot P_0(C_i)}{[\sum_{l=1}^{n_E} P(E_l|C_i)]\cdot[\sum_{l=1}^{n_C} P(E_j|C_l)\cdot P_0(C_l)]}$$

[11] These were considered optimal according to the package's manual I performed the unfolding with. Too large iteration number may cause unreliability and fluctuations, and has to be addressed e.g. via linear regularization.

Eq. 25 can be rewritten as:

$$\langle n(C_i)\rangle = \sum_{j=1}^{n_E} M_{ij} \cdot n(E_j). \tag{26}$$

The $\mathbf{V}$ *covariance matrix* is composed of these two linear components:

$$V_{kl} = V_{kl}(\mathbf{n}(E)) + V_{kl}(\mathbf{M}), \tag{27}$$

where

$$V_{kl}(\mathbf{n}(E)) = \sum_{j=1}^{n_E} M_{kj}M_{lj}n(E_j)\left(1 - \frac{n(E_j)}{\langle N_{true}\rangle}\right) - \sum_{i,j=1;\, i\neq j}^{n_E} M_{ki}M_{lj}\frac{n(E_i)n(E_j)}{\langle N_{true}\rangle}; \tag{28}$$

and

$$V_{kl}(\mathbf{M}) = \sum_{i,j=1}^{n_E} n(E_i)n(E_j) \cdot \mathrm{Cov}(M_{ki}, M_{lj}), \tag{29}$$

in which

$$\mathrm{Cov}(M_{ki}, M_{lj}) = \sum_{\{ru\}\{su\}} \frac{\partial M_{ki}}{\partial P(E_r|C_u)} \frac{\partial M_{lj}}{\partial P(E_s|C_u)} \cdot \mathrm{Cov}[P(E_r|C_u),\, P(E_s|C_u)];$$

$$\frac{\partial M_{ki}}{\partial P(E_r|C_u)} = M_{ki}\left[\frac{\delta_{ku}\delta_{ri}}{P(E_r|C_u)} - \frac{\delta_{ku}}{\varepsilon_u} - \frac{\delta_{ri}M_{ui}\varepsilon_u}{P(E_i|C_u)}\right];$$

$$\mathrm{Cov}[P(E_r|C_u),\, P(E_s|C_u)] = \begin{cases} \frac{1}{n_u}P(E_r|C_u)[1 - P(E_s|C_u)] & \text{if } (r = s) \\ -\frac{1}{n_u}P(E_r|C_u)P(E_s|C_u) & \text{if } (r \neq s) \end{cases}$$

where δ_{ij} is the Dirac delta, n_u is the number of generated events in C_u cause cell.

The **background** can be then taken into account via introducing additional cause cells C_{n_c+1} and initial probabilities $P(C_{n_c+1})$.

Unfolding the errors is also possible via this method.

A simple regularization procedure requires inverting the covariance matrix to obtain the correct χ^2:

$$\chi^2 = (\mathbf{x}_{\mathrm{meas}} - \mathbf{x}_{\mathrm{true}})^T \mathbf{V}^{-1} (\mathbf{x}_{\mathrm{meas}} - \mathbf{x}_{\mathrm{true}}). \tag{30}$$

However, $\mathbf{V}$ is mostly not well-conditioned. With bayesian unfolding, this does not cause a problem: even though $\mathbf{V}$ is calculated via error propagation from $n(E_j)$-ből, M_{ij} is assumed to be independent. This is true for the first iteration, after that $n_0(C_i)$ is substituted with $\langle n(C_i)\rangle$ from previous iteration. The next $\langle n(C_i)\rangle$ is already dependent on $n(E_j)$. To take this into account, the error propagation matrix is utilized:

$$\frac{\partial\langle n(C_i)\rangle}{\partial n(E_j)} = M_{ij} + \sum_{k=1}^{n_E} M_{ik}n(E_k)\left(\frac{1}{n_0(C_i)}\frac{\partial n_0(C_i)}{\partial n(E_j)} - \sum_{l=1}^{n_C}\frac{\varepsilon_l}{n_0(C_l)}\frac{\partial n_0(C_l)}{\partial n(E_j)}M_{lk}\right) \tag{31}$$

The $\frac{\partial n_0(C_i)}{\partial n(E_j)}$ matrix is in fact $\frac{\partial \langle n(C_i)\rangle}{\partial n(E_j)}$ from previous iteration, thus, after the first iteration, the second term vanishes:

$$\frac{\partial n_0(C_i)}{\partial n(E_j)} = 0 \quad \text{és így} \quad \frac{\partial \langle n(C_i)\rangle}{\partial n(E_j)} = M_{ij}.$$

The error propagation matrix can be used to obtain the covariance matrix of the non-unfolded distribution $V(n(E_i), n(E_j))$ from the measurement:

$$V(\langle n(C_k)\rangle, \langle n(C_l)\rangle) = \sum_{i,j=1}^{n_E} \frac{\partial \langle n(C_k)\rangle}{\partial n(E_i)} V(n(E_i), n(E_j)) \frac{\partial \langle n(C_l)\rangle}{\partial n(E_j)} \tag{32}$$

To sum up, the Bayesian unfolding is a powerful method both for unfolding and uncertainty estimation.

2.1.5 RooUnfold

The software used to perform the unfolding procedure detailed in Sec. 2.1.4 is `RooUnfold` [28].

The software package is implemented in `C++`, building on `ROOT` classes and running within the `ROOT` environment. It defines classes for the different unfolding algorithms and a class for the response matrix.

The response matrix can be filled using `RooUnfold`'s built-in classes:

- `Fill(`$x_{\texttt{meas}}$`,` $x_{\texttt{true}}$`)`: fill a real measured value and the corresponding output value;
- `Miss(`$x_{\texttt{true}}$`)`: can be interpreted as an additional row in R, a non-real output value;
- `Fake(`$x_{\texttt{meas}}$`)`: can be interpreted as an additional column in R, a "fake" input value.

2.2 Extracting pseudorapidity distributions

I obtained my results from $\mathrm{d}N/\mathrm{d}i_{\mathrm{Ring}}$ EPD distributions via bayesian iterative unfolding, inferring the statistical uncertainty of the results via propagating the covariance matrix.

I published the results for analysis performed at energies 19.6 and 27 GeV already in 2023 (see Reference [19]), therefore the results detailed here will be in strong agreement with the results of the article.

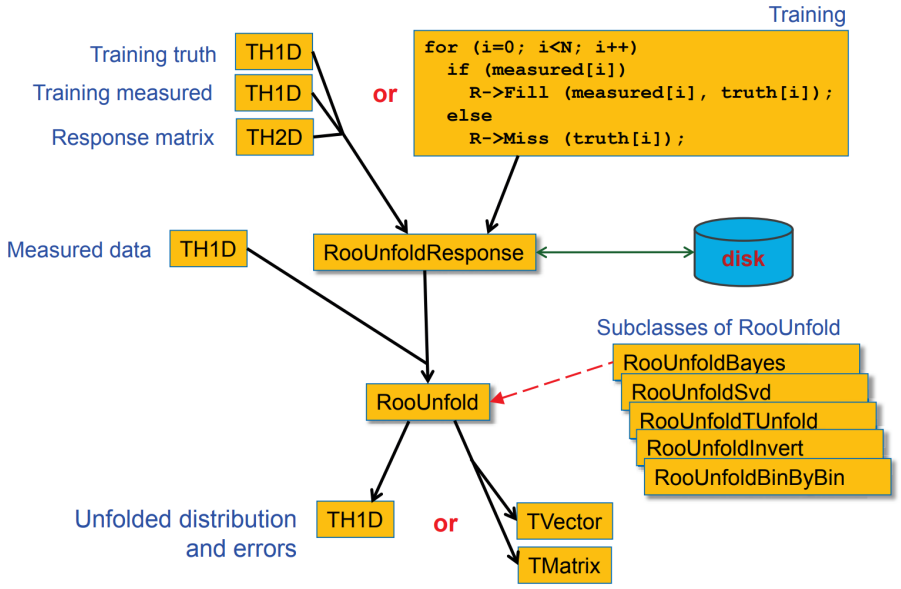


Figure 14: RooUnfold classes. The training truth, training measured, measured and unfolded distributions can be TH1D, TH2D or TH3D histogram inputs. Source: [29], Fig.1. In this thesis, `RooUnfoldBayes` and `RooUnfoldResponse` classes were utilized.

2.2.1 Efficiency: multiple count correction

The unfolding procedure results in one unfolded track for each individual EPD hit. However, as detailed in the previous sections, one primary track can cause multiple hits. Therefore, this effect needs to be corrected for – either via a bin-by-bin correction calculated from Monte Carlo simulation data (via a function of psuedorapidity, a `Number of hits from 1 primary`(η) distribution), or by adding weights to the values that are going to be filled in the response matrix, such that it could compensate for the multiple counts during the unfolding. Both should lead to the same results; in this analysis, the first method was used.

In Fig. 15, the bin-by-bin efficiency correction is visualised for both all particles and only the charged particles as well.

2.2.2 Methods on extracting $\mathrm{d}N_{ch}/\mathrm{d}\eta$

In this analysis, I concentrated on obtaining charged particle pseudorapidity distributions ($\mathrm{d}N_{ch}/\mathrm{d}\eta$). The reason behind only dealing with charged particles is that it was assumed that EPD can only provide reliable results on charged particles.

It is also important to note that the same is assumed for the reliability outside EPD's

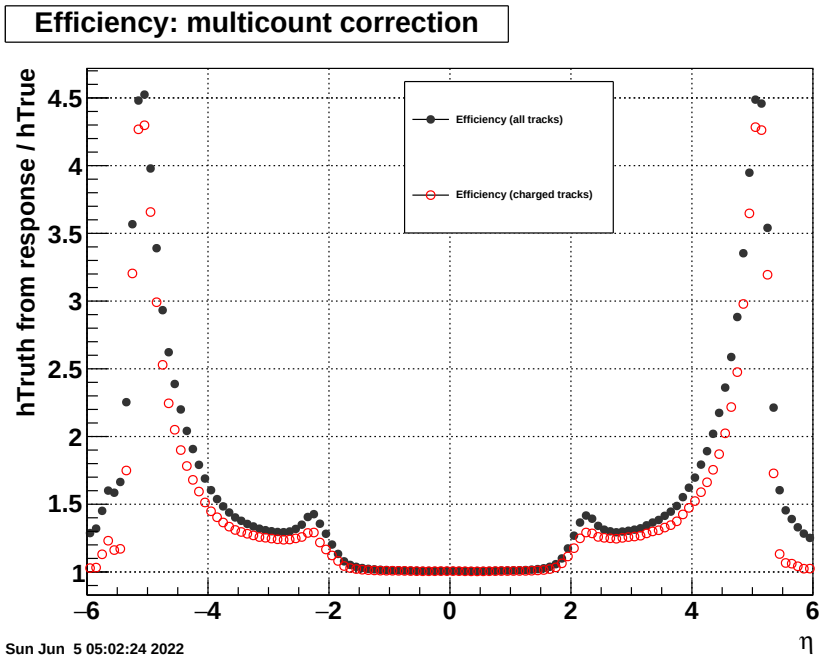


Figure 15: Efficiency correction visualised for all particles (black) and only the charged particles (red).)

pseudorapidity range. Even though some of the figures have values plotted outside the $-5.09 \geq \eta \geq -2.14$ and $2.14 \geq \eta \geq 5.09$ range, these should NOT be interpreted as realistic results.

In order to obtain the charged particle distribution ($\mathrm{d}N_{ch}/\mathrm{d}\eta$) from $\mathrm{d}N/\mathrm{d}\eta$, the fraction of all the particles that are charged has to be taken into account. This can either be done using different bin-by-bin corrections, or neutral particles can be marked as background ("fake").
In Fig. 16, I demonstrated how efficiency and then charged factor correction is applied to the raw unfolded η distribution.

In this analysis, I used the following methods as the charged factor correction:

1. Bin-by-bin correction of the already unfolded $\mathrm{d}N/\mathrm{d}\eta$ using the charged particle fraction $N_{charged}(\eta)/N_{all}(\eta)$ from Monte Carlo simulation (MC) data (such a correction can be seen in Fig. 17);

2. Bin-by-bin correction of the raw EPD data via $N_{charged}(i_{\mathrm{Ring}})/N_{all}(i_{\mathrm{Ring}})$ from MC data; then unfolding only the EPD charged particle distribution (see in Fig. 18).

3. Mark neutral particles as background and fill the response matrix as in the second

method, except that the hits from neutral primaries are considered as "fake".

In the latter two cases, another type of response matrix has to be used that was filled only with the charged particles' data. The visualisation of both of these matrices is the same, one of such can be seen in Fig. 19.

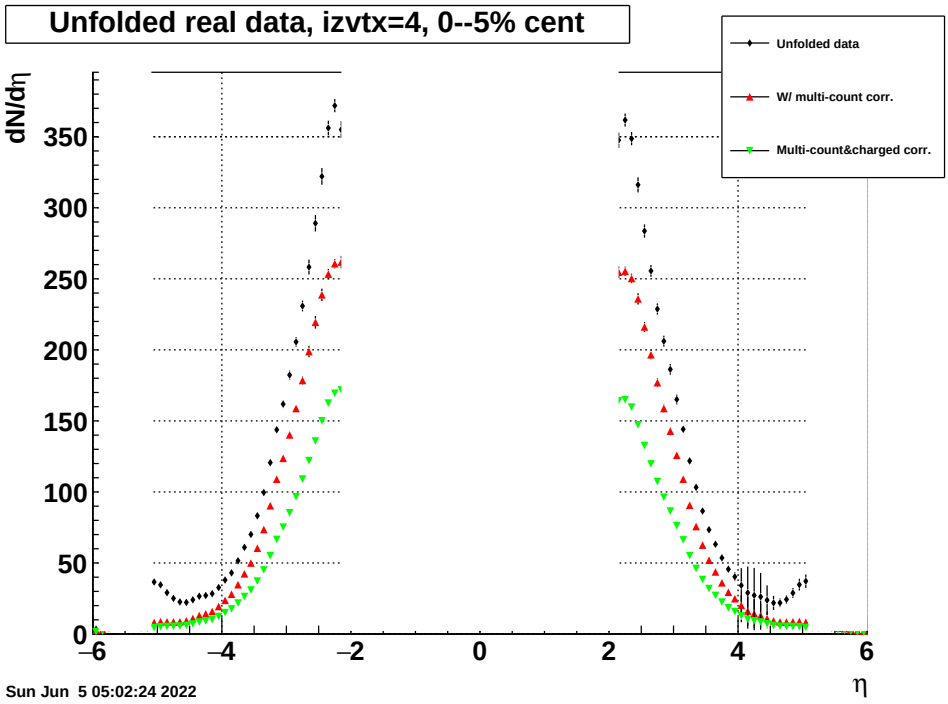


Figure 16: Caption

The three different methods can later be used to estimate the systematic uncertainty of the unfolding procedure itself. An example comparison of the three methods can be seen in Fig. 20 and 21. In the EPD range, all three methods return roughly the same $\mathrm{d}N_{ch}/\mathrm{d}\eta$ within error.

2.2.3 Consistency check

Before unfolding the real EPD data, a closure test was done to check whether the unfolding method can recover the "true" training data itself (MC "truth").

I found that unfolding done on the input training MC sample reproduces well the input η distribution. Even with some added noise ($\pm$1-–10%) to the training sample, the resulting unfolded distribution was in agreement with the input distribution within $< 3\%$ error. Such a test can be seen in Fig. 22. All in all, the unfolding itself was found to work well.

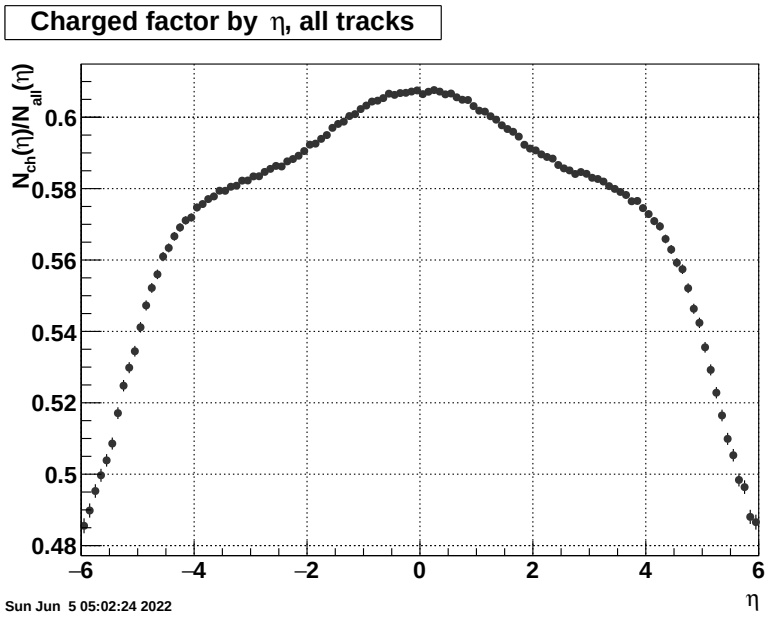


Figure 17: Bin-by-bin correction of the $\mathrm{d}N/\mathrm{d}\eta$ using the charged particle fraction $N_{charged}(\eta)/N_{all}(\eta)$ obtained from MC data.

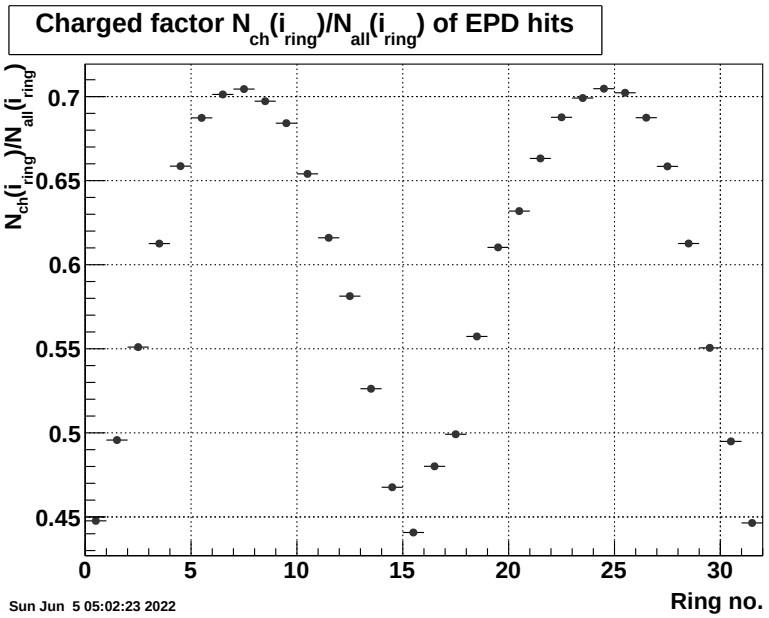


Figure 18: Bin-by-bin correction of the raw EPD data via $N_{charged}(i_{\mathrm{Ring}})/N_{all}(i_{\mathrm{Ring}})$ from MC data

Furthermore, after applying the multiple counting correction and the three different methods of charged factor correction on the unfolded distribution [12], the resulting distri-

[12]Note that the mentioned unfolding procedure was at this stage still done on the MC EPD ring distribution, thus on the training sample. Not to be confused with the unfolding done on the real, experimental EPD data.

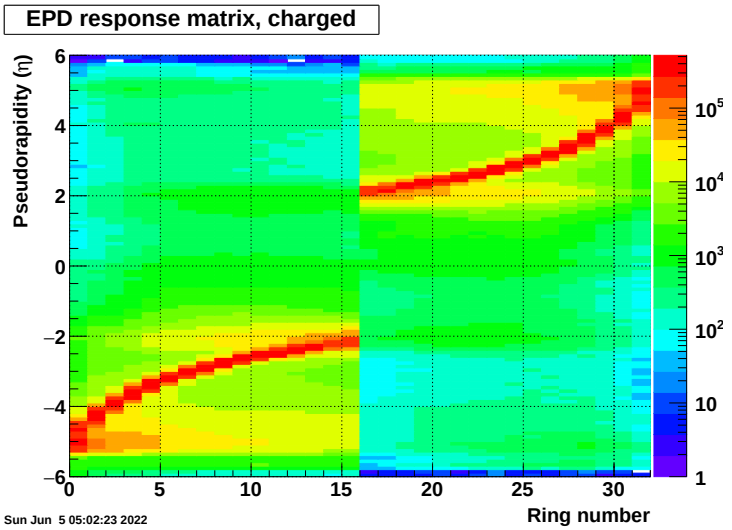


Figure 19: Visualisation of a response matrix containing the charged particles' data. The plottable response matrix used with `Fakes` method look the same.

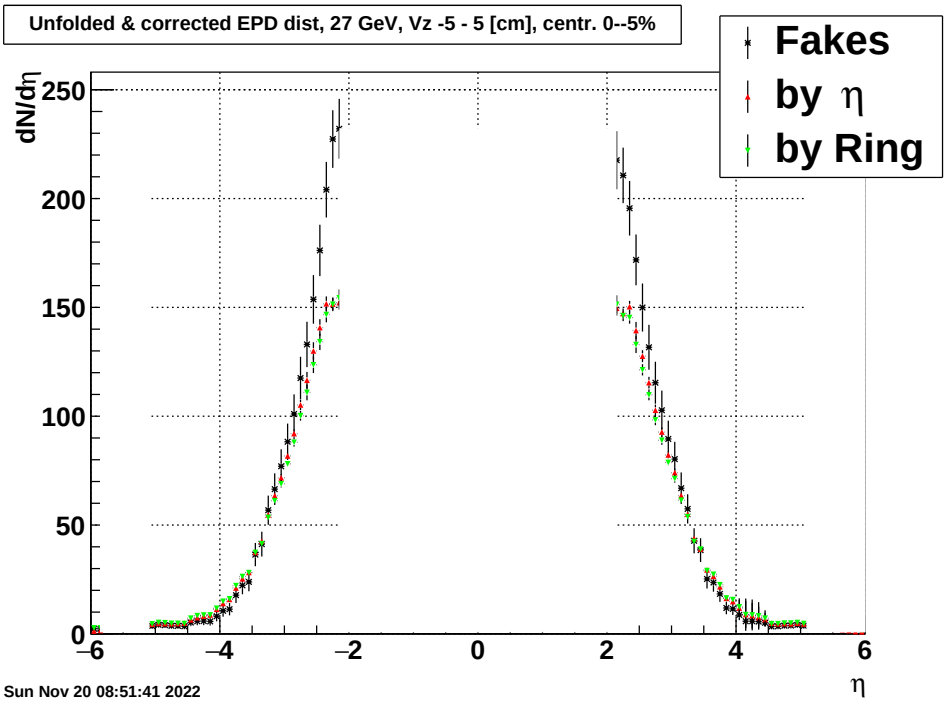


Figure 20: Comparison of the three methods used to obtain the charged particle distribution ($\mathrm{d}N_{ch}/\mathrm{d}\eta$) from the pseudorapidity distribution of all particles ($\mathrm{d}N/\mathrm{d}\eta$). The unfolding was done on real EPD data at $\sqrt{s_{NN}} = 27$ GeV.

butions were compared to the original MC dataset's $\mathrm{d}N_{ch}/\mathrm{d}\eta$.

In Fig. 23 I visualised the unfolding and corrections done with all three methods. As it can be seen, the points overlap, it is then reasonable to plot the relative difference

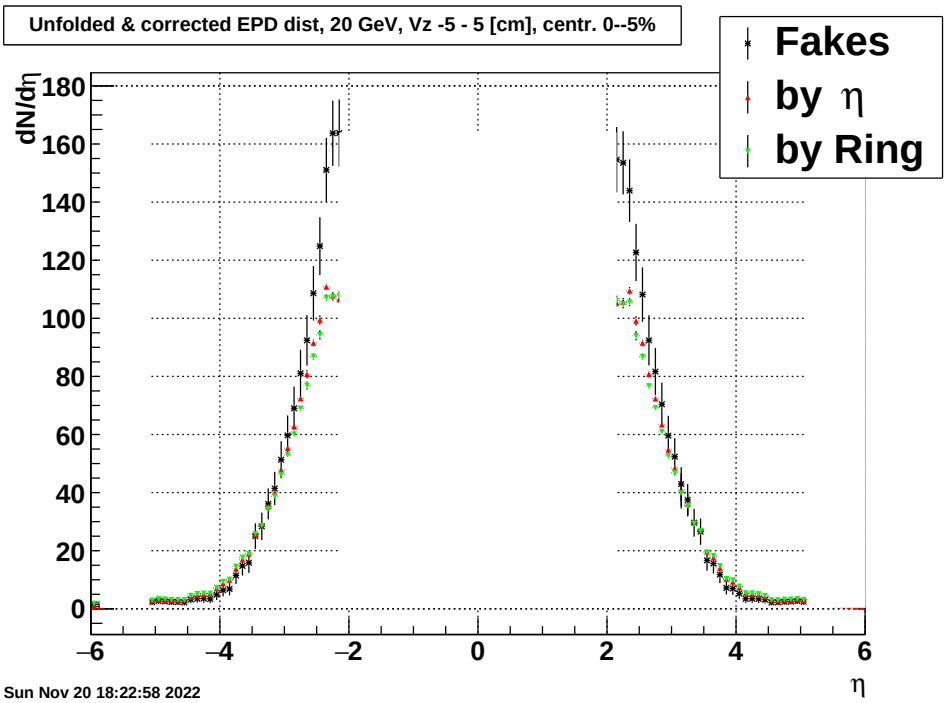


Figure 21: Comparison of the three methods used to obtain the charged particle distribution ($\mathrm{d}N_{ch}/\mathrm{d}\eta$) from the pseudorapidity distribution of all particles ($\mathrm{d}N/\mathrm{d}\eta$). The unfolding was done on real EPD data at $\sqrt{s_{NN}} = 19.6$ GeV.

compared to the MC dataset as well.

As it is visible in Fig. 24, the maximal relative deviation is only

- up to 2% in certain bins for the first method;
- less than 0.1% for the other two methods.

It's important to highlight that, despite the third method (marking neutral particles as fakes) yielding the most precise results in this consistency check, further systematic checks revealed it to be the least reliable, as it heavily relies on the Monte Carlo input used for the response matrix.

Given the result of the closure test, the unfolding and correction methods were considered adequately self-consistent.

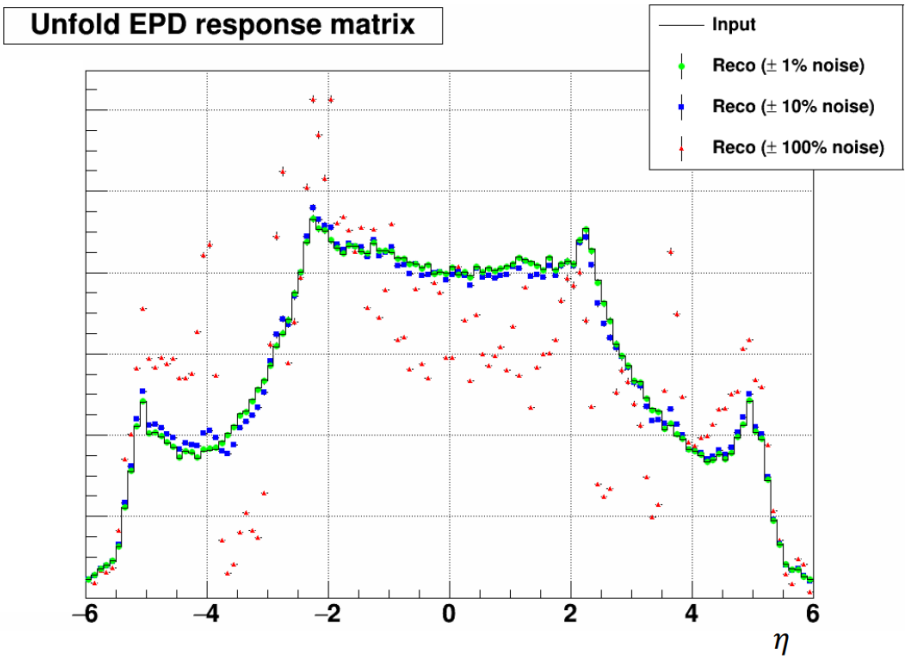


Figure 22: Unfolding the MC EPD distribution with added noise compared to the input MC data. The unfolding seems to be robust to noise.

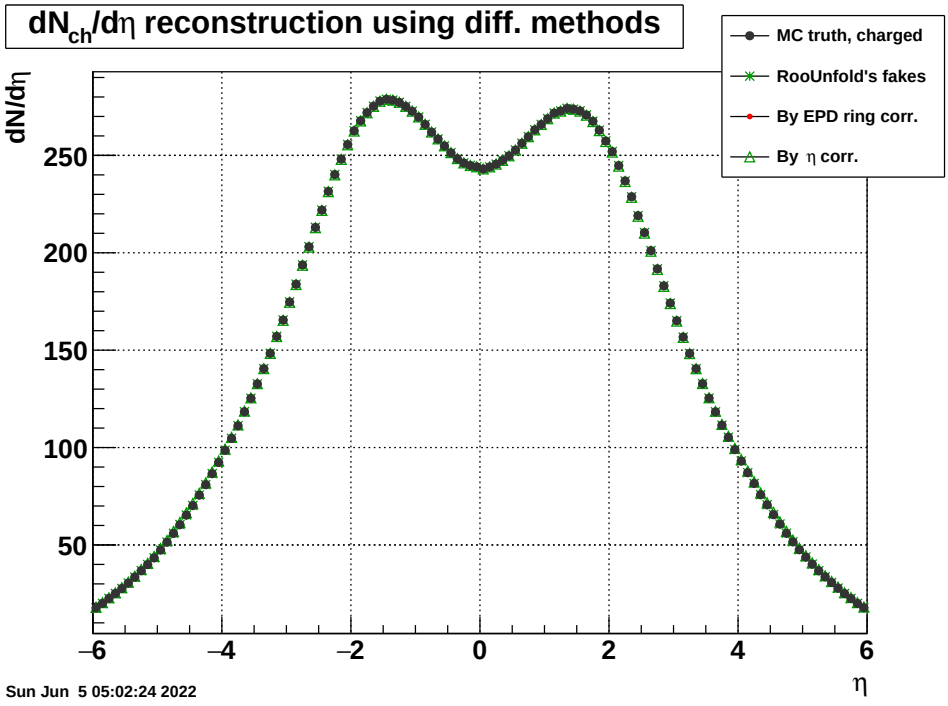


Figure 23: Consistency check: comparison of the three methods used to obtain the charged particle distribution ($\mathrm{d}N_{ch}/\mathrm{d}\eta$) applying to the Monte Carlo simulation data itself.

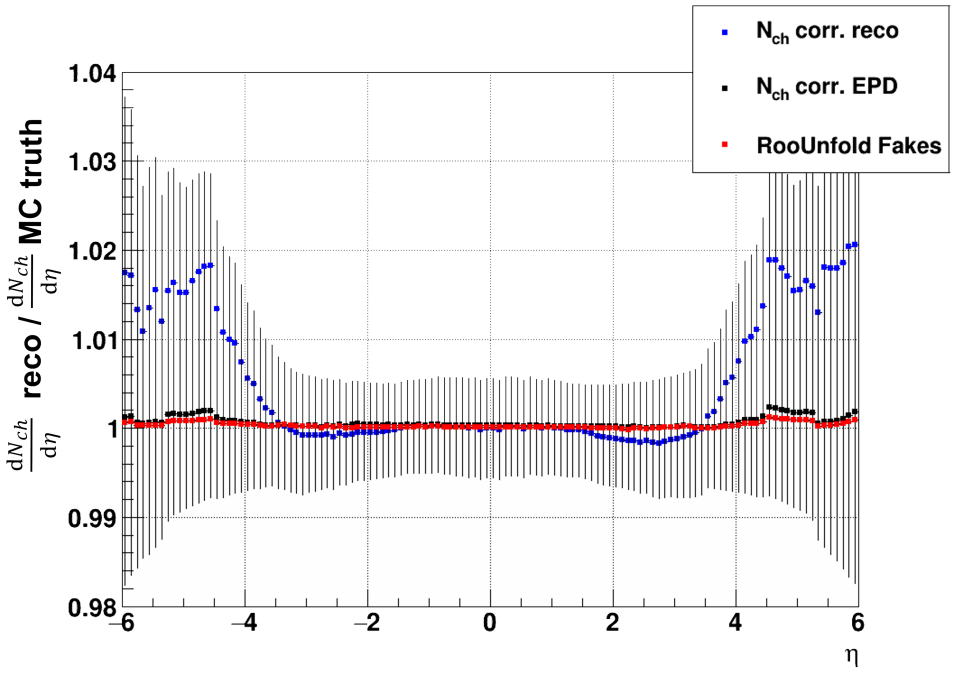


Figure 24: Consistency check of the three different methods to get $\mathrm{d}N_{ch}/\mathrm{d}\eta$ from Monte Carlo simulation-generated EPD ring distribution. The difference is shown as unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ over Monte C "truth", the distributions divided bin-by-bin. Blue marker represents the first method (η-dependent charged factor correction), black shows the second method (EPD ring number dependent charged factor correction), and red represents the third method (marking neutral particles), relative to MC truth's $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the errorbars are only plotted for informative purposes: they were calculated using the `ROOT`'s `TH1` class' default square root of sum of squares of weights (`Sumw2()`).

3 Results

3.1 Results at various energies

As the consistency checks proved the usability of the unfolding method, the next step was to apply the procedure to real EPD data.

In Fig. 25 and 26, the results from 19.6 GeV and 27 GeV measurements can be seen, with reasonable statistical error.

In addition to published results, I did the analysis on $\sqrt{s_{NN}} = 14.6$ GeV data as well. In Fig. 27 and 28, rather consistent results were obtained. However, the statistical uncertainties represented by the errorbars have no physical significance, as the uncertainties

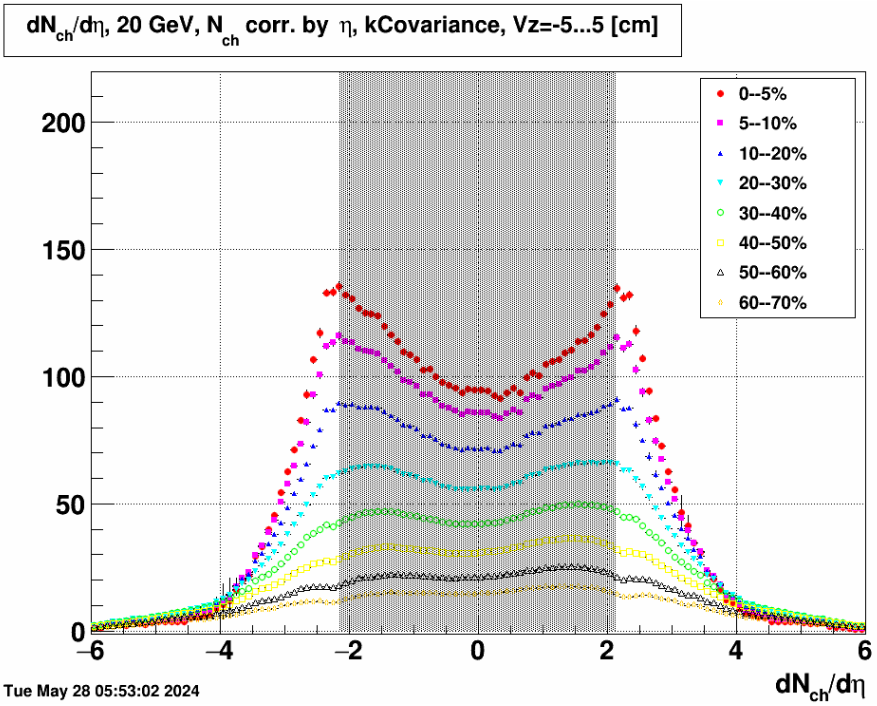


Figure 25: Example of unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at RHIC energy $\sqrt{s_{NN}} = 19.6$ GeV, in z-vertex range of -5 to $+5$ cm. The errorbars represent the statistical error propagated via the covariance matrix. The horizontal axis represent η, the vertical $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the results are only valid in the $2.14 < |\eta| < 5.09$ range. 1st method used.

during the extraction of the raw EPD data were ill-defined.

In Fig. 29, the `Fakes` method was used as charged factor correction; in this case, the error propagation did not work, and the same is true for all the unfolded distributions done with this method. However, these data still do have significance, as they enabled to check the systematic uncertainty of the different unfolding methods.

3.2 Systematic error checks [19]

In the following section, the examined systematic uncertainty sources and their contribution to the results are discussed.

3.2.1 Dependence on input MC distribution

The bayesian iterative unfolding process, via its iterative nature, should mostly overcome differences in response matrix from real response that are not related to distortion effects, such as detector geometry [26]. However, as the exact response matrix cannot be de-

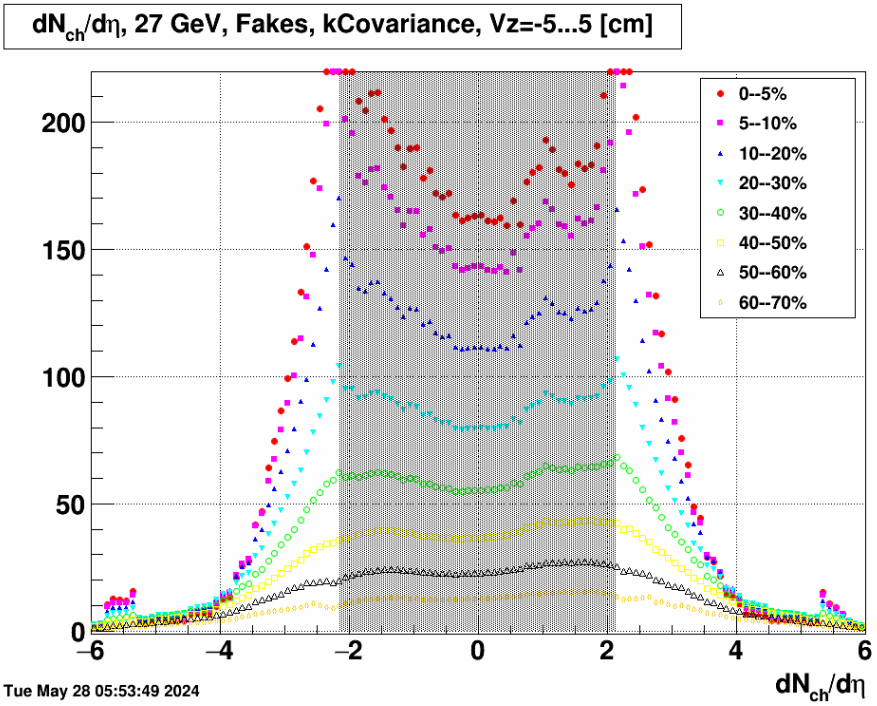


Figure 26: Example of unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at RHIC energy $\sqrt{s_{NN}} = 27$ GeV, in z-vertex range of -5 to $+5$ cm. The errorbars represent the statistical error propagated via the covariance matrix. The horizontal axis represent η, the vertical $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the results are only valid in the $2.14 < |\eta| < 5.09$ range. 1st method used.

termined even with precise MC simulations and the unfolding process itself is not perfect, some dependencies on the various parameters in the MC simulations can occur. Those are considered as systematic uncertainties of the measurement.

3.2.2 Tightening and shifting the input MC $\mathrm{d}N/\mathrm{d}\eta$

Firstly, the simulated sample's $\mathrm{d}N_{ch}/\mathrm{d}\eta$ was modified ("suppressed") using a Gaussian shape with width σ and mean η_0. These suppression factors can be seen in Fig. 30a. This was done via a random selection based on Gaussian distribution while filling the response matrices.

Using this approach, all combinations could be analysed, that is, unfolding the i-th MC sample's EPD ring hit distribution via response from j-th MC sample. In case of $i = j$, the unfolding was as close to perfect as expected.

Unfolding results with the Gaussian width of $\sigma \lessapprox 1$ were not considered here, as in this case there are almost no particles in the EPD range. Otherwise, there was less than

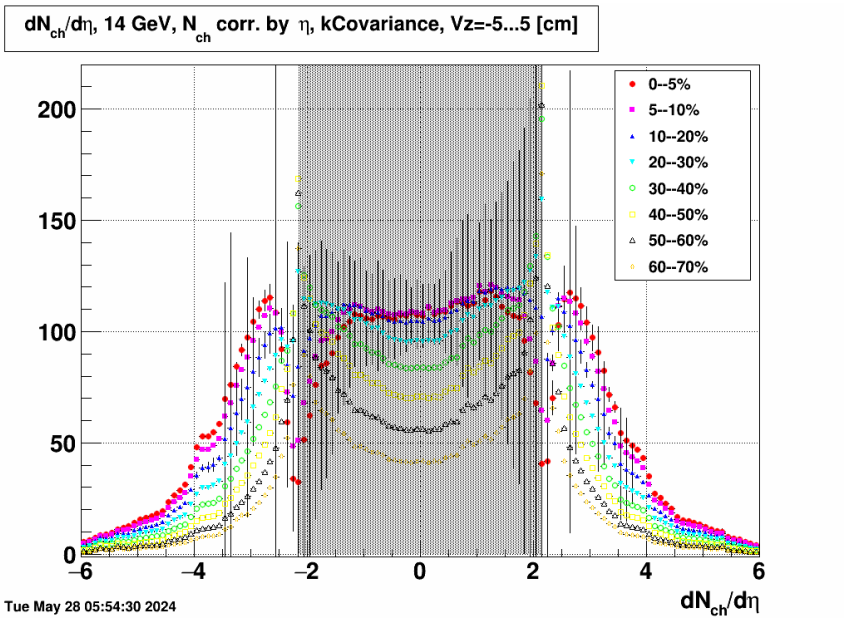


Figure 27: Example of unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at RHIC energy $\sqrt{s_{NN}} = 14.6$ GeV, in z-vertex range of -5 to $+5$ cm. The errorbars represent the statistical error propagated via the covariance matrix, however, in this case they have no physical significance. The horizontal axis represent η, the vertical $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the results are only valid in the $2.14 < |\eta| < 5.09$ range. 1st method used.

a few percent variations in the EPD's η region.

Overall, in the analysis the effect of tightening the $\mathrm{d}N_{ch}/\mathrm{d}\eta$ of the training sample to $\sigma = 2$ and shifting it by ± 3 units of pseudorapidity was investigated.

Broadening the input MC $\mathrm{d}N/\mathrm{d}\eta$

Similar to modification done in Sec. 3.2.2., here the tracks were modified with a factor of

$$\exp\left(\frac{\eta^2 - \eta_{\max}^2}{2\sigma_{\text{broad}}}\right). \tag{33}$$

There was no suppression utilized for $|\eta| > \eta_{\max}$, with $\eta_{\max} = 6$. The resulting shape of the distributions can be seen in Fig. 31.

While unfolding the data with these input MC distributions, a significant decrease at midrapidity values was observed. However, this occurred mostly outside the EPD's η region; the unfolding was considered acceptable down to $\sigma_{\text{broad}} \approx 3$.

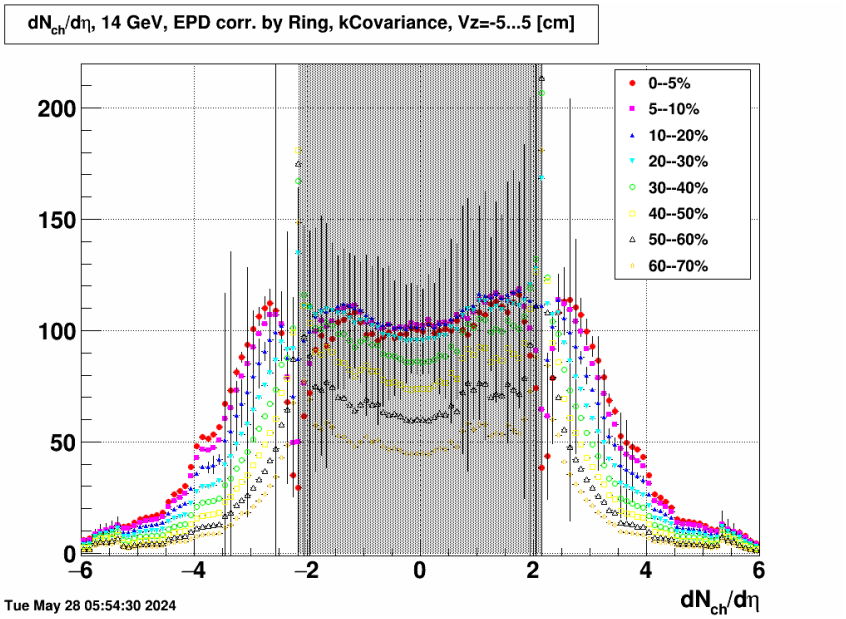


Figure 28: Example of unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at RHIC energy $\sqrt{s_{NN}} = 14.6$ GeV, in z-vertex range of -5 to $+5$ cm. The errorbars represent the statistical error propagated via the covariance matrix, however, in this case they have no physical significance. The horizontal axis represent η, the vertical $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the results are only valid in the $2.14 < |\eta| < 5.09$ range. 2nd method used.

3.2.3 Changing the charged fraction in the MC training dataset

The fraction of the charged particles in the MC input data was changed by ±15%. This was achieved by randomly rejecting either the neutral or the charged particles.

3.2.4 Changing the p_{T} slope of the MC training dataset

The transverse momentum (p_{T}) distribution slope of the MC input data was changed by ±10% via randomly rejecting particles of small or large p_{T}.

3.2.5 Centrality and z-vertex selection

It was investigated, by how much the unfolded distribution would change if either the z-vertex or the centrality selection are modified. For the former investigation, a ±5 cm calibration uncertainty in the z-vertex measurement of the real EPD data was employed; for the second one, ±5% calibration uncertainty was assumed in centrality determination of the real EPD data.

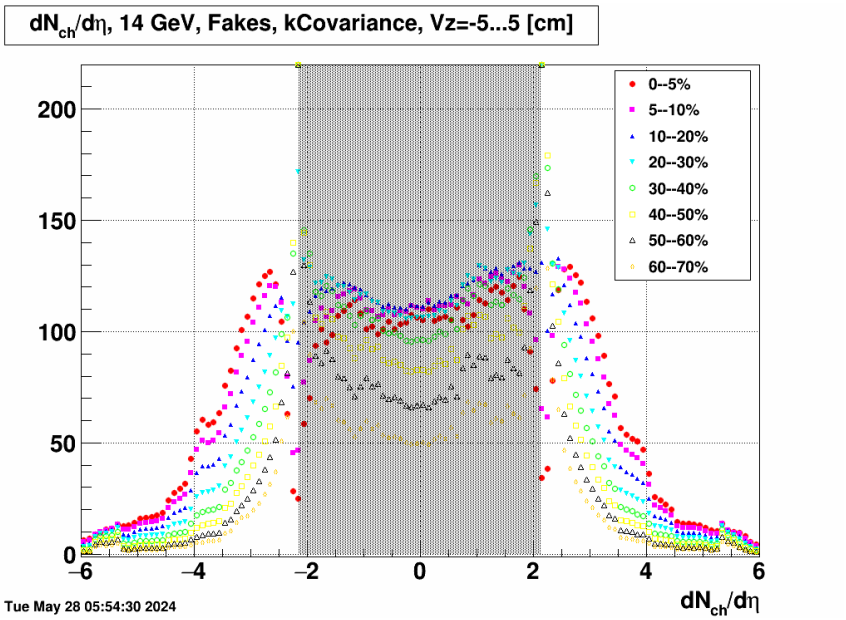


Figure 29: Example of unfolded $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at RHIC energy $\sqrt{s_{NN}} = 14.6$ GeV, in z-vertex range of -5 to $+5$ cm. The horizontal axis represent η, the vertical $\mathrm{d}N_{ch}/\mathrm{d}\eta$. Note that the results are only valid in the $2.14 < |\eta| < 5.09$ range. 3rd method used; in this case, the error propagation did not work at all.

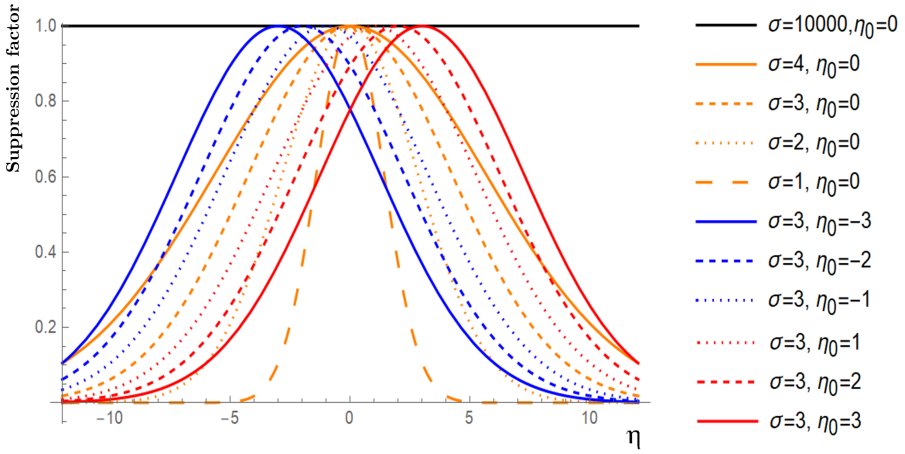
3.2.6 z-vertex choice

Due to the detector geometry, it is important to also take into account the interaction point's z-vertex position in the calculations, as the resulting pseudorapidity distribution should not depend on it.

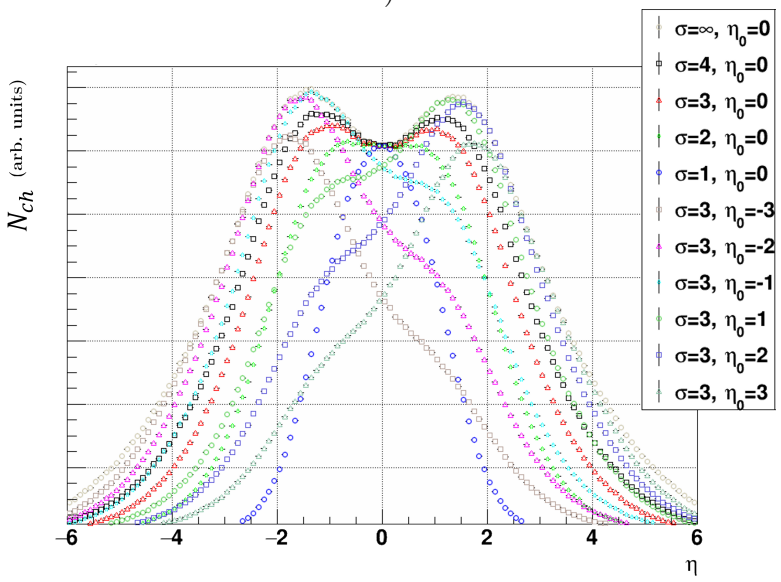
The EPD data, as well as the responses, were collected in nine different z-vertex classes, equally distributed from -45 to $+45$ cm. Depending on which range was unfolded, the resulting distribution still may differ and has to be taken into account as systematic uncertainty.

3.2.7 Unfolding method choice

The most significant systematic uncertainty contribution was caused by the difference between the results achieved using different unfolding and correction methods (as listed in Sec. 2.2.2.). The first method was used as a benchmark, from which the differences were calculated.



a)



b)

Figure 30: Tightening and shifting the MC input distribution using random selection based on Gaussian distribution of σ width and η_0 curve peak position. (**a**) Demonstration of the Gaussian suppression factors used. (**b**) The $\mathrm{d}N_{ch}/\mathrm{d}\eta$ of the distorted MC input samples.

3.2.8 EPD related uncertainties

As previously stated, the EPD electronics were considered fully efficient (except some "dead areas" in the detector from e.g. glue and gaps, but these were assumed to be correctly handled in the simulation). The uncertainty from multi-MIP Landau fit was considered negligible compared to other systematic sources.

In conclusion, the systematic uncertainties coming from the detector system itself were

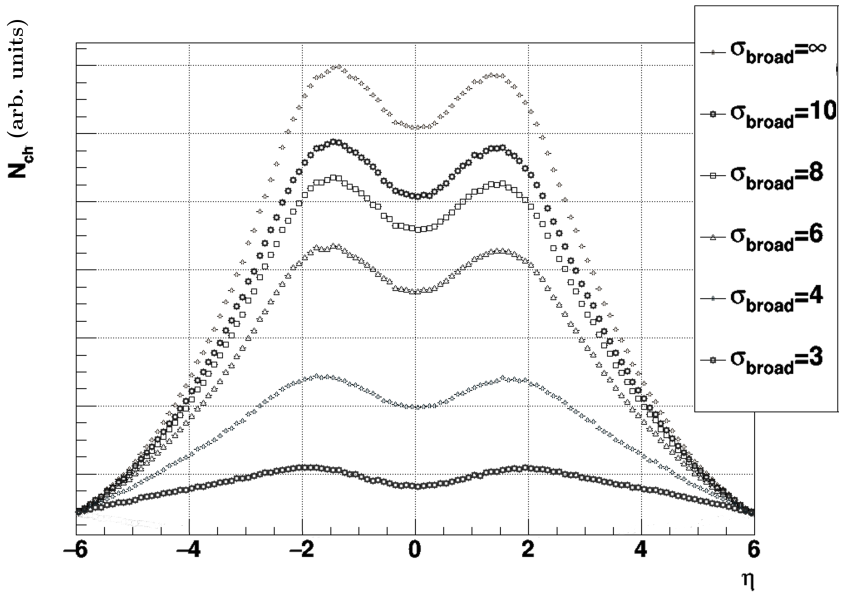


Figure 31: Broadening the MC input distribution using random selection based on Gaussian distribution of σ_{broad} width.

considered negligible.

Table 1: Summary of systematic uncertainty sources and their contribution.

Source	**Systematic uncertainty**
MC input $\mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ tightening, shifting	6%
MC input $\mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ broadening	4%
Charged fraction in MC	6%
p_{T} slope change in MC	1%
Centrality selection	2%
z-vertex selection	negligible
z-vertex choice	1%
Unfolding method choice	8%
EPD related uncertainties, electronics, efficiency	negligible

The different systematic uncertainty sources and their contribution with informative percentage values can be seen in Table 1.

3.3 Results summary

In this thesis, charged particle pseudorapidity distributions with systematic uncertainties listed in Sec. 3.2 were obtained at two RHIC energies, in the EPD pseudorapidity range. The results at $\sqrt{s_{NN}} = 19.6$, 27.0 and 14.6 GeV can be seen in Fig. 32, 33 and 34, respectively. The caption #MIP ≤ 5 written on the plot indicates the number of convolution members in the multi-MIP Landau fit.

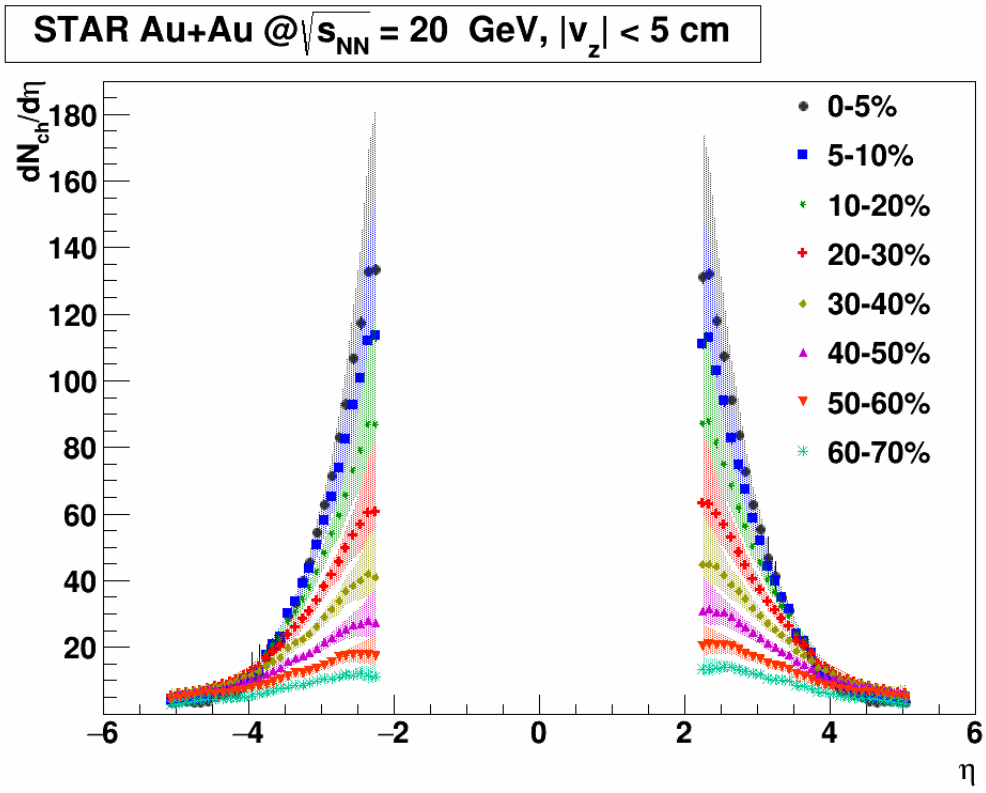


Figure 32: Charged particle pseudorapidity distributions measured with STAR EPD on RHIC energy $\sqrt{s_{NN}} = 19.6$ GeV. The data was processed in eight centrality classes, presented with the different markers. The statistical uncertainties, marked by errorbars, are not visible on this plot, as the markers themselves are larger. The coloured area indicates the systematic uncertainties of the measurement.

3.4 Comparison with PHOBOS results

Another experiment of the RHIC complex was the PHOBOS experiment, which completed data taking in 2006. The PHOBOS was a large acceptance silicon detector, covering almost 2π in azimuth and $|\eta| < 5.4$ in pseudorapidity [25]. Compared to STAR's EPD, there are differences in both detector topology and granularity: the silicone pad detectors

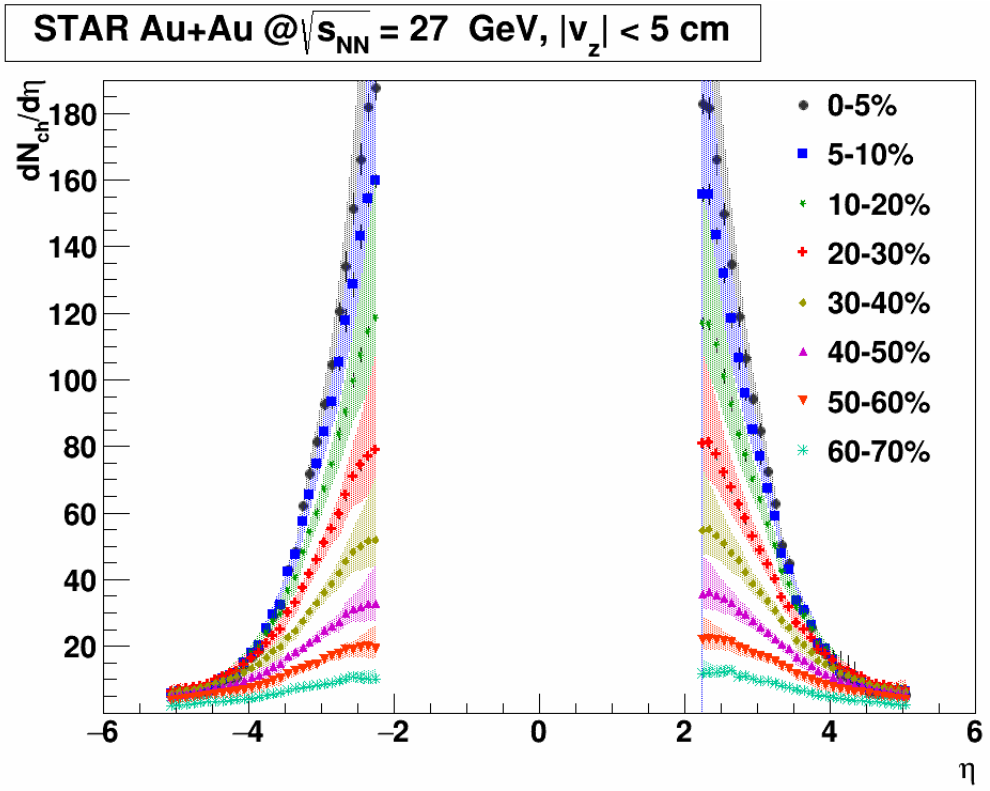


Figure 33: Charged particle pseudorapidity distributions measured with STAR EPD on RHIC energy $\sqrt{s_{NN}} = 27.0$ GeV. The data were processed in eight centrality classes, presented with the different markers. The errorbars represent the statistical uncertainty, the coloured area indicates the systematic uncertainties of the measurement.

measure the total number of charged particles emitted in the collision, with modules mounted onto a centrally located octagonal frame (Octagon) covering $|eta| \leq 3.2$, as well as three annular frames (Rings) on either side of the collision vertex, extending the coverage out to $|\eta| \leq 5.4$ [30].

The PHOBOS also measured $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ at 19.6, 62.4, 130, 200 GeV energies [31]. Although in that paper a slightly different centrality binning was used (0–3%, 3–6% and 6–10% instead of 0–5% and 5–10%; the other centrality classes were the same), at 19.6 GeV the results can be compared.

In Figure 35, it is apparent that the two measurements show sizeable differences, depending on η: around up to a factor of two, increasing from small $|\eta|$ towards forward/-backward rapidities.

The exact reasons behind this discrepancy are not yet known but the difference cannot

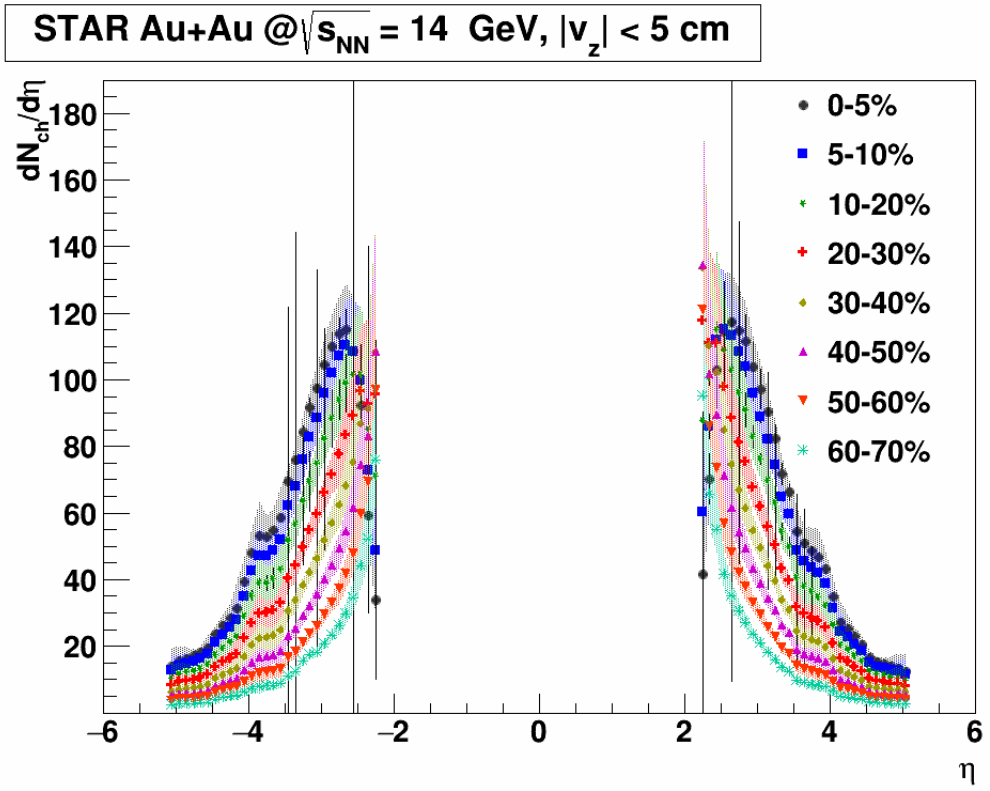


Figure 34: Charged particle pseudorapidity distributions measured with STAR EPD on RHIC energy $\sqrt{s_{NN}} = 14.6$ GeV. The data were processed in eight centrality classes, presented with the different markers. The errorbars are not of physical significance, the coloured area indicates the systematic uncertainties of the measurement.

be explained by the systematic uncertainties described in Sec. 3.2.

3.5 Next steps

3.5.1 Comparison with MUFFIN-SMASH simulation

MUFFIN-SMASH is a next-generation hybrid event-by-event three-fluid hydrodynamic model, suitable for simulations of heavy-ion collisions in the energy range from few up to tens of GeV per colliding nucleon-nucleon pair. [32]

As seen in Fig. 36, this project also has results in charged particle pseudorapidity distributions. It would be worth comparing my results with this model's.

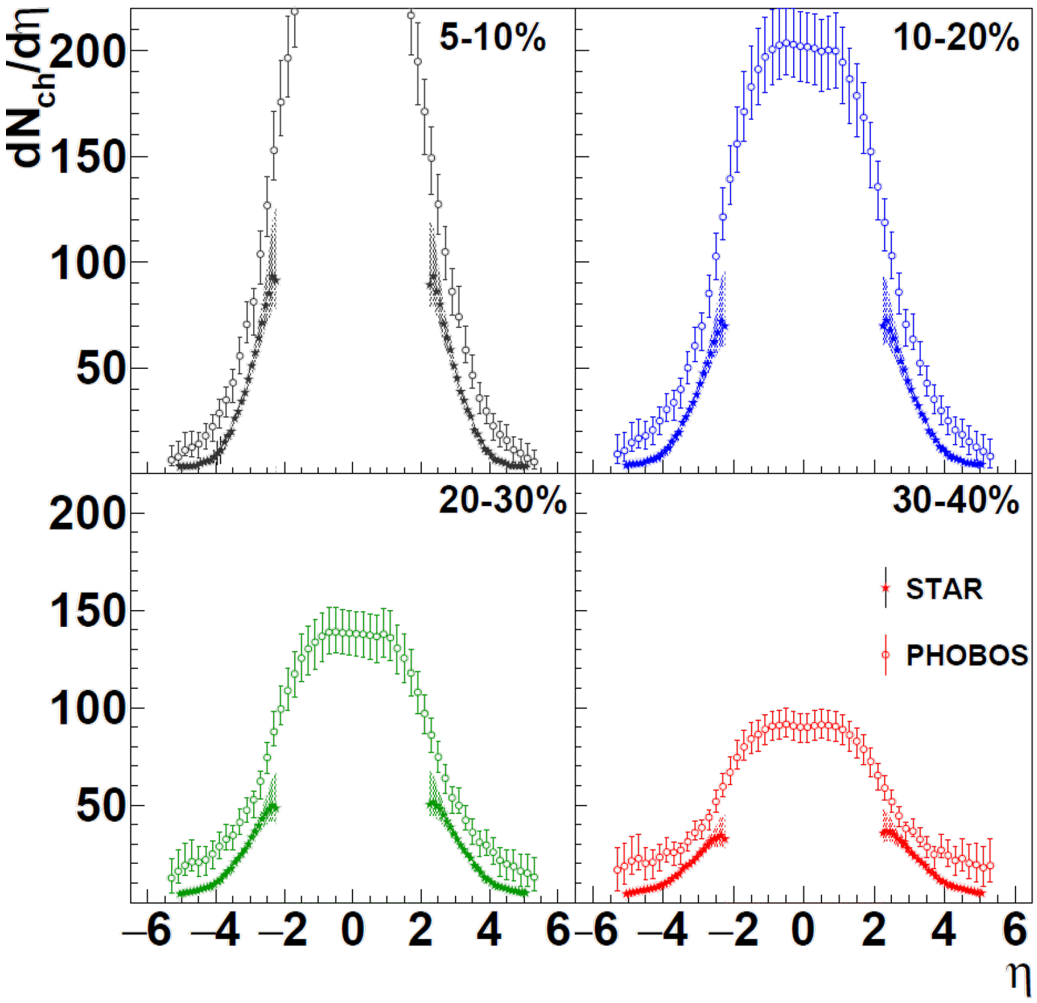


Figure 35: Charged particle pseudorapidity distributions measured in PHOBOS (hollow circles) and STAR (star markers). Note that on the upper left graph the centrality class of the PHOBOS experiment's result is actually 6–10%.

3.6 $N_{\mathbf{part}}$ corrected comparison of pseudorapidity distributions

In theory, the pseudorapidity distributions corrected with the number of particles created, should all overlap. This has to be done.

4 Summary

Based on EPD ring-by-ring distributions, charged particle pseudorapidity measurements at $\sqrt{s_{\mathrm{NN}}} = 19.6$, 27.0 and 14.6 GeV were performed with detailed systematic investigations regarding simulation data, calibration data, and unfolding methods.

The results at $\sqrt{s_{NN}} = 19.6$ GeV show significant difference compared to the results

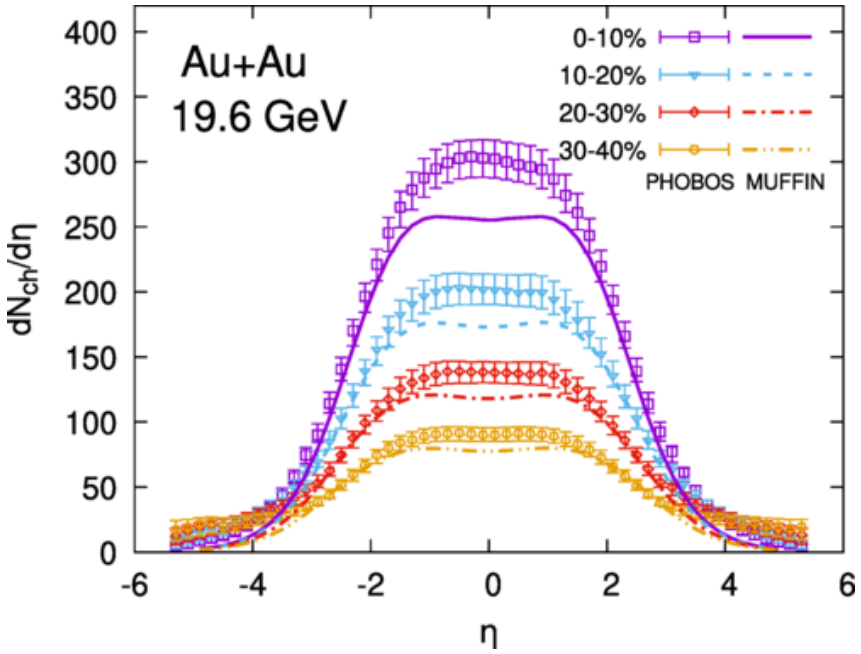


Figure 36: MUFFIN-SMASH results at 19.6 GeV. [32]

from PHOBOS. There are four components in this comparison: EPD spectrum measurement, `Geant4` simulation, unfolding procedure from the STAR part, and the PHOBOS data itself.

The results at 14.6 GeV need to be refined, with proper statistical uncertainty analysis.

The method presented in this manuscript is to be extended to other $\sqrt{s_{NN}}$ values (as part of the BES-II program) and to fixed target data – mainly at energies where the QCD critical point is expected [33]. Refining this measurement method is also important for the search of the QCD critical point, in order to fine-tune the models used in these analyses.

List of Figures

List of Tables

Köszönetnyilvánítás

Köszönetemet szeretném kifejezni témavezetőmnek, Csanád Mátének, a STAR kutatócsoportjának és azon belül is Mike Lisa-nak a rendelkezésre juttatott EPD adatokért, Xiaouyu Liu-nak a segítségért a 14.6 GeV-es adatok kifejtéséhez. Ezen kívül húgaimnak, hogy a legnagyobb krízis idején egyszerűen csak őszintén mellettem álltak.

References

[1] S. Weinberg), 'The first three minutes,' in Basic Books, 1977.

[2] M. Csanád, 'Nagyenergiás atommagütközések téridőbeli szerkezete,' *MTA doktori értekezés*, 2019.

[3] D. Horváth, 'A részecskefizika anyagelmélete: A standard modell,' *Fizikai Szemle*, vol. 7-8. 2008.

[4] U. Cheuk-Yin Wong (Oak Ridge National Laboratory, 'Introduction to high-energy heavy-ion collisions,' in World Scientific, 1994, p. 24. DOI: 10.1142/1128.

[5] JabberWok. 'Wikipedia-en, wikimedia commons.' (2007), [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/3/30/Pseudorapidity2.png.

[6] U. W. Heinz and P. F. Kolb, 'Early thermalization at rhic,' *arXiv preprint hep-ph/0111075*, 2001.

[7] R. Ryblewski, *Collective phenomena in the early stages of relativistic heavy-ion collisions*, 2012. arXiv: 1207.0629 [nucl-th].

[8] M. E. Tejeda-Yeomans, 'Heavy-ion physics: Freedom to do hot, dense, exciting qcd,' *arXiv preprint arXiv:2004.13812*, 2020.

[9] A. Collaboration. 'Centrality determination in heavy ion collisions.' (), [Online]. Available: https://inspirehep.net/files/7616388e2d08c6f210a6b20a0367e21a.

[10] M. L. Miller, K. Reygers, S. J. Sanders and P. Steinberg, 'Glauber modeling in high-energy nuclear collisions,' *Annual Review of Nuclear and Particle Science*, vol. 57, no. Volume 57, 2007, pp. 205–243, 2007, ISSN: 1545-4134. DOI: https://doi.org/10.1146/annurev.nucl.57.090506.123020. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev.nucl.57.090506.123020.

[11] E. Shuryak, 'Why does the quark–gluon plasma at rhic behave as a nearly ideal fluid?' *Progress in Particle and Nuclear Physics*, vol. 53, no. 1, pp. 273–303, 2004, Heavy Ion Reaction from Nuclear to Quark Matter, ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2004.02.025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0146641004000444.

[12] M. Csanád, T. Csörgő, Z.-F. Jiang and C.-B. Yang, 'Initial energy density of sqrt(s) = 7 and 8 tev p–p collisions at the lhc,' *Universe*, vol. 3, no. 1, 2017, ISSN: 2218-1997. DOI: 10.3390/universe3010009. [Online]. Available: https://www.mdpi.com/2218-1997/3/1/9.

[13] P. Bozek, W. Broniowski and S. Chatterjee, 'Transverse momentum fluctuations and correlations,' *arXiv preprint arXiv:1707.04420*, 2017.

[14] STAR. 'Star images, https://drupal.star.bnl.gov/star/public/img/.' (), [Online]. Available: https://www.flickr.com/photos/brookhavenlab/7979381212/in/album-72157613690851651/.

[15] 'The star experiment.' (), [Online]. Available: https://www.star.bnl.gov/central/experiment/.

[16] J. Adams, A. Ewigleben, S. Garrett *et al.*, 'The star event plane detector,' *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 968, p. 163 970, 2020, ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.163970. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900220304344.

[17] A. M. Poskanzer and S. A. Voloshin, 'Methods for analyzing anisotropic flow in relativistic nuclear collisions,' *Phys. Rev. C58*, pp. 1671–1678. 1998. DOI: 10.1103/PhysRevC.58.1671..

[18] J. Whitten C. A. and S. Collaboration, 'The Beam-Beam Counter: A Local Polarimeter at STAR,' *AIP Conference Proceedings*, vol. 980, no. 1, pp. 390–396, Feb. 2008, ISSN: 0094-243X. DOI: 10.1063/1.2888113. eprint: https://pubs.aip.org/aip/acp/article-pdf/980/1/390/11985098/390_1_online.pdf. [Online]. Available: https://doi.org/10.1063/1.2888113.

[19] M. Molnár, 'Charged particle pseudorapidity distributions measured with the star epd,' *Universe*, vol. 9, no. 7, 2023, ISSN: 2218-1997. DOI: 10.3390/universe9070335. [Online]. Available: https://www.mdpi.com/2218-1997/9/7/335.

[20] I. Antcheva, M. Ballintijn, B. Bellenot *et al.*, 'Root — a c++ framework for petabyte data storage, statistical analysis and visualization,' *Computer Physics Communications*, vol. 182, no. 6, pp. 1384–1385, 2011, ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2011.02.008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010465511000701.

[21] T. S. Experiment. 'Stmudstmaker class reference.' (), [Online]. Available: https://www.star.bnl.gov/webdata/dox/html/classStMuDstMaker.html.

[22] S. Hall, *A comparison of mudst and picodst data analysis methods*, Bachelor's Thesis, 2022.

[23] X.-N. Wang and M. Gyulassy, 'Hijing 1.0: A monte carlo program for parton and particle production in high energy hadronic and nuclear collisions,' *Comp. Phys. Comm.*, vol. 83, pp. 307–331, 1995. DOI: https://doi.org/10.48550/arXiv.nucl-th/9502021.

[24] S. Agostinelli, J. Allison, K. Amako *et al.*, 'Geant4—a simulation toolkit,' *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003, ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[25] B. Alver, B. B. Back, M. D. Baker *et al.*, 'Participant and spectator scaling of spectator fragments in au + au and cu + cu collisions at $\sqrt{s_{NN}}$ = 19.6 and 22.4 gev,' *Phys. Rev. C*, vol. 94, p. 024 903, 2 2016. DOI: 10.1103/PhysRevC.94.024903. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.94.024903.

[26] G. D'Agostini, 'A multidimensional unfolding method based on bayes' theorem,' *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 362, pp. 487–498, 1995. DOI: 10.1016/0168-9002(95)00274-X.

[27] G. D'Agostini, *Improved iterative bayesian unfolding*, 2010. arXiv: 1010.0632 [physics.data-an]

[28] T. Adye. 'Roounfold programcsomag és dokumentáció.' (), [Online]. Available: http://hepunx.rl.ac.uk/~adye/software/unfold/RooUnfold.html.

[29] T. Adye, 'Unfolding algorithms and tests using roounfold,' *arXiv*, 2011. DOI: 10.48550/arXiv.1105.1160.

[30] B. Back, M. Baker, M Ballintijn *et al.*, 'The phobos perspective on discoveries at rhic,' *Nuclear Physics A*, vol. 757, no. 1-2, pp. 28–101, 2005.

[31] B. Alver, B. B. Back, M. D. Baker *et al.*, 'Charged-particle multiplicity and pseudorapidity distributions measured with the phobos detector in Au + Au, Cu + Cu, **d** + Au, and **p** + **p** collisions at ultrarelativistic energies,' *Phys. Rev. C*, vol. 83, p. 024 913, 2 2011. DOI: 10.1103/PhysRevC.83.024913. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.83.024913.

[32] J. Cimerman, I. Karpenko, B. Tomášik and P. Huovinen, 'Next-generation multifluid hydrodynamic model for nuclear collisions at s n n from a few gev to a hundred gev,' *Physical Review C*, vol. 107, no. 4, p. 044 902, 2023.

[33] G. Odyniec, 'Rhic beam energy scan program—experimental approach to the qcd phase diagram,' *Journal of Physics G: Nuclear and Particle Physics*, vol. 37, 2010. DOI: doi:10.1088/0954-3899/37/9/094028.

NYILATKOZAT

Név: Molnár Mátyás

ELTE Természettudományi Kar, szak: Fizikus MSc

NEPTUN azonosító: FKLESX

Szakdolgozat címe:

Pszeudorapiditás-eloszlások mérése Au+Au ütközésekben a Relativisztikus Nehézion-ütkeztető STAR Eseménysík-detektora segítségével

Measuring pseudorapidity distributions in Au+Au collisions at RHIC with the STAR EPD

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló szellemi alkotásom, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2024.05.30.

a hallgató aláírása