Proceedings of the 3rd International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/43/20241139

GPU-based computing analysis for charged-particle production in heavy ion collisions

Yuanjia Zhang

Department of Physics, Zhejiang University, Hangzhou, 310058, China

rize_z@zju.edu.cn

Abstract. The innovation of this paper lies in the introduction of GPU acceleration, which has greatly speeded up the project's progress. It begins by offering a brief overview of the research history on quark-gluon plasma (QGP) and the corresponding research tool, heavy ion collisions. The paper mainly focuses on data analysis of a significant process within the hadronic collisions, namely, the charged-particle production. The investigation of this process is an integral part of QGP research, as it occurs promptly after the formation of QGP. By applying the tracklet method to the datasets collected by ALICE detector, an estimate of the charged-particle multiplicity density is presented. The finding then suggests a positive correlation between the number of participating nucleons in the collision and the normalized charged-particle multiplicity density. Notably, this dependence on centrality is similar to the research conducted by the ALICE collaboration, but derived from a distinct dataset. It indicates that there should be an underlying mechanism behind the intriguing phenomenological trend waiting for further exploration. These studies have the potential to enhance the understanding of quantum chromodynamics (QCD) and provide insights into the true nature of the universe.

Keywords: quark-gluon plasma, heavy ion collisions, charged-particle production, tracklet method, GPU acceleration.

1. Introduction

In the 1970s, an ultraviolet behavior of non-Abelian gauge theories known as "asymptotic freedom" was revealed. It indicated a property in QCD that interactions between particles asymptotically weaken as the energy scale increases. Afterward, scientists soon realized the existence of a new state of matter – a plasma of liberated quarks and gluons with extremely high energy [1]. Today, scholars refer to it as quark-gluon plasma, abbreviated as QGP. The experimental efforts have never ceased since the proposal of this concept. In February 2000, CERN (Conseil Européen pour la Recherche Nucléaire, Geneva, Switzerland) publicly declared, for the first time, the existence of experimental evidence indicating the creation of the new state of matter. Five years later, QGP was directly observed at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Lab (BNL, New York, USA) [1, 2]. The nearly 25-year-long "hunt" has come to a conclusion, simultaneously opening the doors to new branches of research.

A majority of nuclear and particle scientists are convinced that research programs focused on QGP and relativistic heavy ion (RHI) collisions are of great significance. The QGP physics is an interdisciplinary field that intersects nuclear and particle physics. It serves as both an extension of nuclear matter research and an inheritance of particle physics theory [3]. The past two decades have seen

DOI: 10.54254/2753-8818/43/20241139

an increasing number of researchers devote themselves to studying this area. Here are some instances of the potential applications of the QGP physics [2, 4].

1) Quarks have a much lighter mass compared to nucleons. Studying the phase transition from hadrons to QGP helps answer the ultimate question: What is mass?

2) The universe was hot in the extremely brief period after the Big Bang. Studying the behavior of the QGP helps in understanding the evolution of the early universe.

3) Liquid QGP is the earliest and, in a sense, the simplest form of complex quantum matter. Research into its emergence helps explore how the common (but more complex) constituents of the world around us come into being.

4) During the collisions, a large number of 2^{nd} and 3^{rd} generation quarks are produced, which are typically rare. Experiments related to QGP can provide substantial experimental material for researching the nature of flavor.

In addition, the concepts and methodologies in this field can also be applied to other areas, including hadron physics and soft condensed matter physics.

Given the current technological capabilities, heavy ion collisions, specifically ultra-relativistic heavy ion collisions, represent the optimal experiments for investigating QGP. According to the predictions from lattice QCD, the critical temperature for hadron melting reaches several hundred MeV [3]. This calls for projectile particles to possess exceedingly high velocities, which justifies the term "ultra-relativistic". At the Large Hadron Collider (LHC), the world's largest and highest-energy experimental setup, collision experiments can achieve $\sqrt{s_{NN}}$ in the TeV range, where $\sqrt{s_{NN}}$ refers to the center-of-mass energy per nucleon pair.

The RHI collisions constitute a unique mechanism of particle production. Insights into this process can be gained through a qualitative analysis rooted in phenomenology, relativity and QCD. The collision starts with two heavy ions. Due to Lorentz contraction, each incident nucleus appears as a disc in the center-of-mass frame. Since electromagnetic forces do not play a dominant role at this point, this experiment will not yield results similar to the Rutherford scattering. Instead, when the two discs collide (or, to put it precisely, when they overlap), they will not experience notable transverse displacement most of the time. In this very brief moment, longitudinal color fields emerge between the discs following some rapid color charge exchange. With the discs receding, the color fields store increasing energy, eventually decaying into a vast number of high-energy quarks, antiquarks and gluons. Thus, QGP is produced. It can be anticipated that its lifetime will not last long. After a period of cooling, these elementary particles will aggregate to form diverse hadrons. The generated hadrons will travel in all directions and the number of them can reach into the thousands. The detailed analysis of this strong interaction process is both intricate and indispensable, as it can contribute to advancing people's understanding of QCD and the nature of QGP [4].

At present, the leading experiment dedicated to the study of QCD matter and QGP is A Large Ion Collider Experiment (ALICE), one of the major research projects conducted at the LHC. The construction of ALICE took nearly 20 years and drew upon the collective efforts of numerous countries and institutes. ALICE is comprised of a central barrel part and a muon spectrometer. The barrel is mainly equipped with two types of detectors: the central-barrel detectors, which consist of multiple cylindrical detection layers and are designed for measuring hadrons, electrons and photons; and a set of smaller forward detectors, serving the purpose of global event characterization, triggering function, and multiplicity research. The innermost component of the central-barrel detectors is the Inner Tracking System (ITS), which includes six tracking layers: two Silicon Pixel Detectors (SPD), two Silicon Drift Detectors (SDD), and two Silicon Strip Detectors (SSD) [5, 6].

The data used in this paper mainly comes from the SPD. When a charged particle hits the detector, it induces the generation of electron-hole pairs, giving rise to a current pulse. This signal enables the determination of the particle's position where it impacts the layer. In the center-of-mass frame, the position can be expressed by a pair of angles – the polar angle θ and the azimuthal angle φ, as illustrated in Figure 1. For the sake of convenience, pseudo-rapidity η ($\eta \equiv -\ln[\tan(\theta/2)]$) is often used as a substitute

DOI: 10.54254/2753-8818/43/20241139

for θ. SPD layer 1 and SPD layer 2 can measure the η range as $|\eta| < 2.0$ and $|\eta| < 1.4$, respectively; while both layers can cover the entire azimuth [5-7].

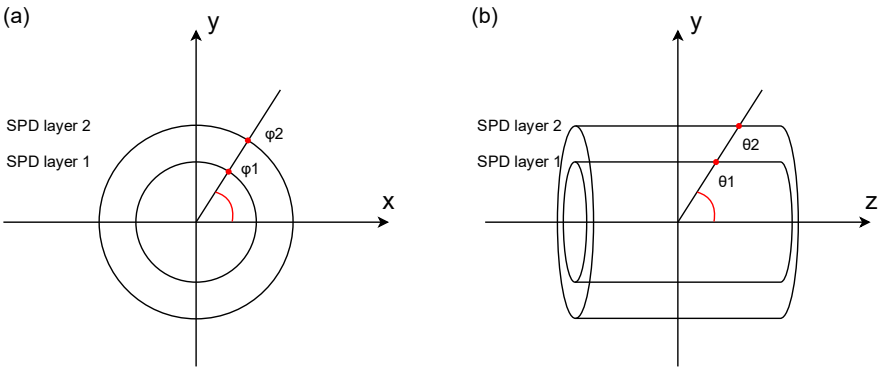

Figure 1. Schematic diagram of the experimental apparatus. (a) cross-sectional view of the detector; (b) side view (z-y plane view) of the detector.

Furthermore, it is important to note that not all observed collision events are head-on collisions. The percentile centrality is a common metric in particle physics to evaluate the geometric properties of collisions. It directly correlates with the impact parameter, or in other words, with the shape and size of the initial overlap region of the two Lorentz-contracted discs. Apparently, different centralities imply distinct initial states. Therefore, categorizing collision events based on this key parameter becomes essential. This paper obtains the centrality information from one of the forward detectors in ALICE, the plastic scintillator detector (VZERO). The VZERO amplitude distribution can be utilized to calculate centrality percentiles of collision events [6, 8].

Another point worth emphasizing is that the nuclear structure is far more complicated than one might imagine. Owing to the influence of quantum fluctuations, a real nucleus is composed of numerous colored quarks, antiquarks and gluons, with a non-uniform density distribution. Therefore, in AA collisions, regardless of whether directly measuring the number of participating nucleons (N_{part}) or deducing it from other observables, both methods face a considerable challenge. Thankfully, a powerful theoretical technique has been developed to estimate this value. It is known as the Glauber model, typically with two mainstream calculation methods: optical and Monte Carlo approaches. Physicists often employ this model to simulate RHI collisions and calculate the centrality dependence of N_{part}. That is to say, for a given collision experiment, the average value $\langle N_{part} \rangle$ can be computed within a specific centrality class [4, 8, 9]. In this paper, the data of $\langle N_{part} \rangle$ is sourced from Ref. [10].

Due to physical characteristics and technological constraints, more information about charged particles is collected in the study of collision products. This article also focuses on the charged particles produced during RHI collisions. Its primary objective is to estimate the charged-particle multiplicity density $dN_{ch}/d\eta$ within different centrality classes and explore the phenomenological connection between them. The paper is structured as follows. Section 2 introduces the core methods adopted in the data analysis, namely, *the tracklet method* and *GPU acceleration*. In Section 3, an elaborate description of the data processing procedure and a presentation of the results are provided. Finally, a conclusion is drawn in Section 4.

2. Methods

2.1. The Tracklet Method

The tracklet method, also known as the tracklet reconstruction algorithm, plays a crucial role in data analysis within the ALICE experiment. It is often used to estimate the number of charged particles produced in high-energy hadronic collisions. By exploiting data from the SPD exclusively, the algorithm can reconstruct the interaction vertex and then charged tracks (tracklets) originating from this position.

DOI: 10.54254/2753-8818/43/20241139

Specifically, it involves the following steps. First, the primary collision location is determined through correlation analysis. Subsequently, transform the original signals into a frame with the reconstructed vertex as the origin, namely, the center-of-mass frame. Tracklets, which represent the charged particles, are then rebuilt by matching those new coordinates between two detector layers [11]. In comparison to other measurement techniques, this method possesses a highly competitive edge. Its minimal detector involvement allows operational simplicity and rapidity, while consistently providing reliable results under a wide range of experimental conditions. Moreover, owing to the mathematical principles behind its algorithm, this approach excels at filtering background noise, including noise caused by detectors and secondary particles [11].

2.2. GPU Acceleration

A GPU (Graphics Processing Unit), initially designed as a microprocessor for accelerating graphics rendering, is now commonly applied to various working fields, such as deep learning and scientific computing. It houses thousands of parallel processing units called shaders or CUDA (Compute Unified Device Architecture) cores, enabling GPUs to execute large-scale computational tasks simultaneously. Thus, GPU acceleration refers to the process of utilizing a GPU to perform parallel computations, ultimately improving program efficiency. When it comes to algorithms suitable for parallel computing, that is, those capable of breaking down complex tasks into smaller and similar ones, GPU acceleration can notably expedite the data processing progress. Today, thanks to the availability of mature platforms, achieving GPU acceleration has become readily accessible. From a logical perspective, kernel functions are written to address divided tasks and are encapsulated within threads. From a physical perspective, CUDA automatically allocates GPU resources to run multiple threads. As is widely known, high energy physics deals with an enormous amount of particle data. Therefore, in addition to amassing computational resources, enhancing algorithm performance is of great importance. Fortunately, the primary algorithm used in the data analysis, the tracklet method, is compatible with GPU acceleration. In this paper, the data processing tools include a PC with an integrated NVIDIA GeForce RTX 2060 and Python with the CUDA GPU acceleration library.

3. Data Analysis and Results

The data sample comes from Pb-Pb collisions at $\sqrt{s}_{NN} = 2.76$TeV within the pseudo-rapidity interval $|\eta| < 1.0$. A total of 20,000 collision events are provided, with each event containing hit information received by the SPD and a VZERO amplitude of the event. It is worth mentioning that these hits have been represented by coordinates (η, φ) in the center-of-mass frame. Therefore, the first step is to establish the fundamental algorithmic framework, that is, to implement tracklet reconstruction through coordinate matching. In an ideal scenario, a charged particle will leave identical signals on both the inner and outer tracking layers. Hence, a pair of coordinates that are approximately equal, one from layer 1 and one from layer 2, are expected to represent a reconstructed tracklet or, in other words, a generated charged particle. In this context, "matching" means searching for such coordinate pairs between the two layers.

A very natural search approach is doing pairwise comparisons, where the simplest form is to calculate differences. Specifically, for any (η_1, φ_1) and (η_2, φ_2) on layer 1 and layer 2, respectively, perform the subtractions $\Delta\eta = \eta_1 - \eta_2$ and $\Delta\varphi = \varphi_1 - \varphi_2$. When creating distributions of $\Delta\eta$ and $\Delta\varphi$, it becomes evident that closely-matching hits cluster around the origin, while non-matching hits are scattered along the x-coordinate axis. The statistical properties of those histograms can then be described. For a single collision event, it is known that the distributions of η_1 (φ_1) and η_2 (φ_2) are roughly the same, following a uniform distribution. The difference between two uniform distributions results in a triangular distribution. Thus, a triangular background is expected after the subtractions. Additionally, η_1 (φ_1) and η_2 (φ_2) are not mutually independent; substantial overlap between them will lead to a peak at the origin. The upcoming step is supposed to select tracklets from this peak for counting the charged-particle multiplicity N_{ch}. However, this step will be postponed to the end of this section, at which point the algorithm will be sufficiently refined.

DOI: 10.54254/2753-8818/43/20241139

The procedure outlined above is primarily tailored for an individual event, yet the computational task requires processing 20,000 events, totaling about 40 million floating-point values. Even the basic subtraction calculations still demand significant computing resources. The traditional method of using loop statements to go through all events would take approximately 5 hours of continuous operation on a standard PC. Accordingly, it is time for GPU acceleration to show up. Taking two arrays as input, the written kernel function accomplishes the following task: it computes pairwise differences between the two arrays, and returns the distribution of the differences in terms of bins and their respective frequencies. Each thread is dedicated to a single event. Operating multiple threads in parallel realizes the concurrent tracklet reconstruction for multiple events. This facilitates plotting the distributions within a set of events, thereby preparing for subsequent steps. In fact, GPU acceleration markedly speeds up the program execution, reducing the required time for traversing each event to less than 2 minutes.

To validate the algorithm's effectiveness, the distributions of Δη and Δφ for all events are plotted, as shown in Figure 2. Note that the shapes of these histograms align with the previous expectation, which, in turn, substantiates the reasonability of the data sample used in this article.

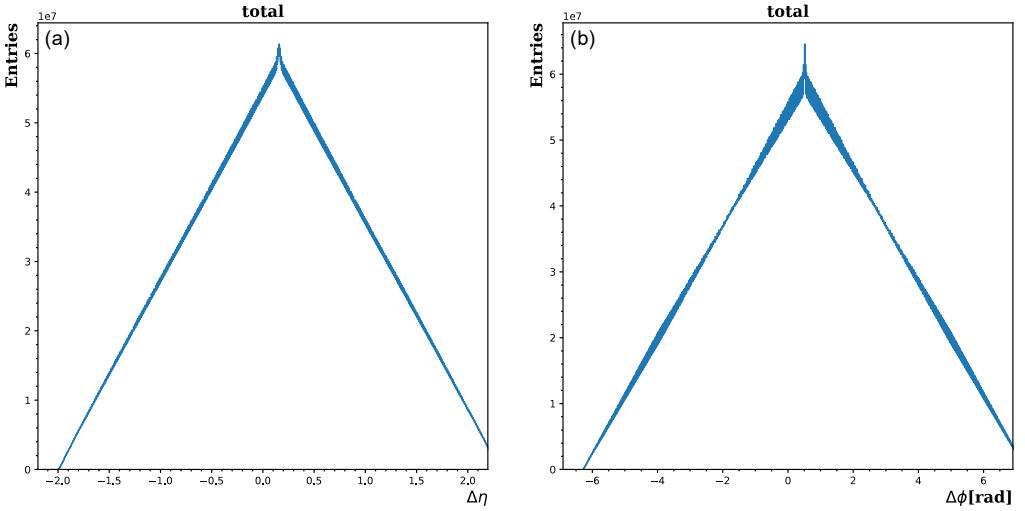

Figure 2. Distributions for all events with bins = 1100. (a) the overall distribution of Δη; (b) the overall distribution of Δφ.

Nevertheless, as mentioned earlier, it is more physically meaningful to investigate an event collection characterized by similar initial states, which necessitates classifying collision events by centrality before the multiplicity estimation. The distribution of the VZERO amplitudes is then depicted in Figure 3, where bins are divided into ten regions based on percentiles, with the same division standards as presented in Ref. [10]. This yields nine distinct centrality classes (excluding the 80-100% part because of the absence of a well-defined primary vertex [10]), each class corresponding to a group of events with similar centralities.

Further, separately input these nine sets of events into the aforementioned algorithm, and distributions of Δη and Δφ within different centrality classes can be obtained. The nine histograms of Δφ are shown in Figure 4 as an example. Upon observing these graphs, it can be found that as the percentage decreases (indicating an increase in centrality), the peak in the distribution becomes less prominent. This suggests that when an event approaches central collisions, the total number of generated particles will increase, leading to a more pronounced dominance of the triangular background in the plot. On the other hand, the Glauber model states that N_{part} increases with the rise in centrality [9]. Given these two points, it can be qualitatively inferred that there is a monotonically increasing relationship between the centrality and the multiplicity.

DOI: 10.54254/2753-8818/43/20241139

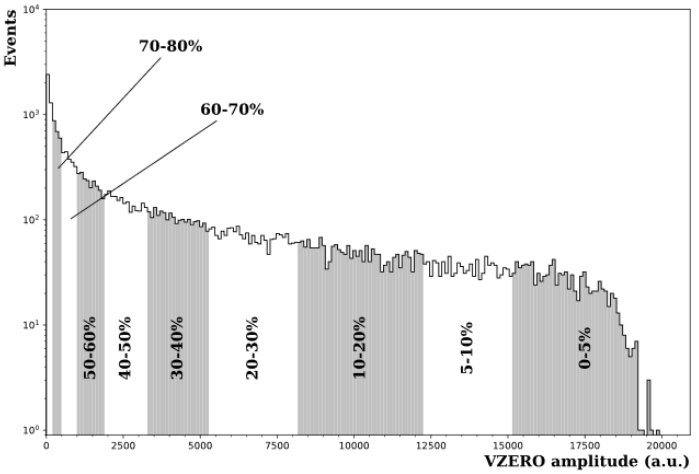

Figure 3. Distribution of the VZERO amplitudes with bins = 200. Classify bins by percentiles 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, and 80% (excluding 80-100%).

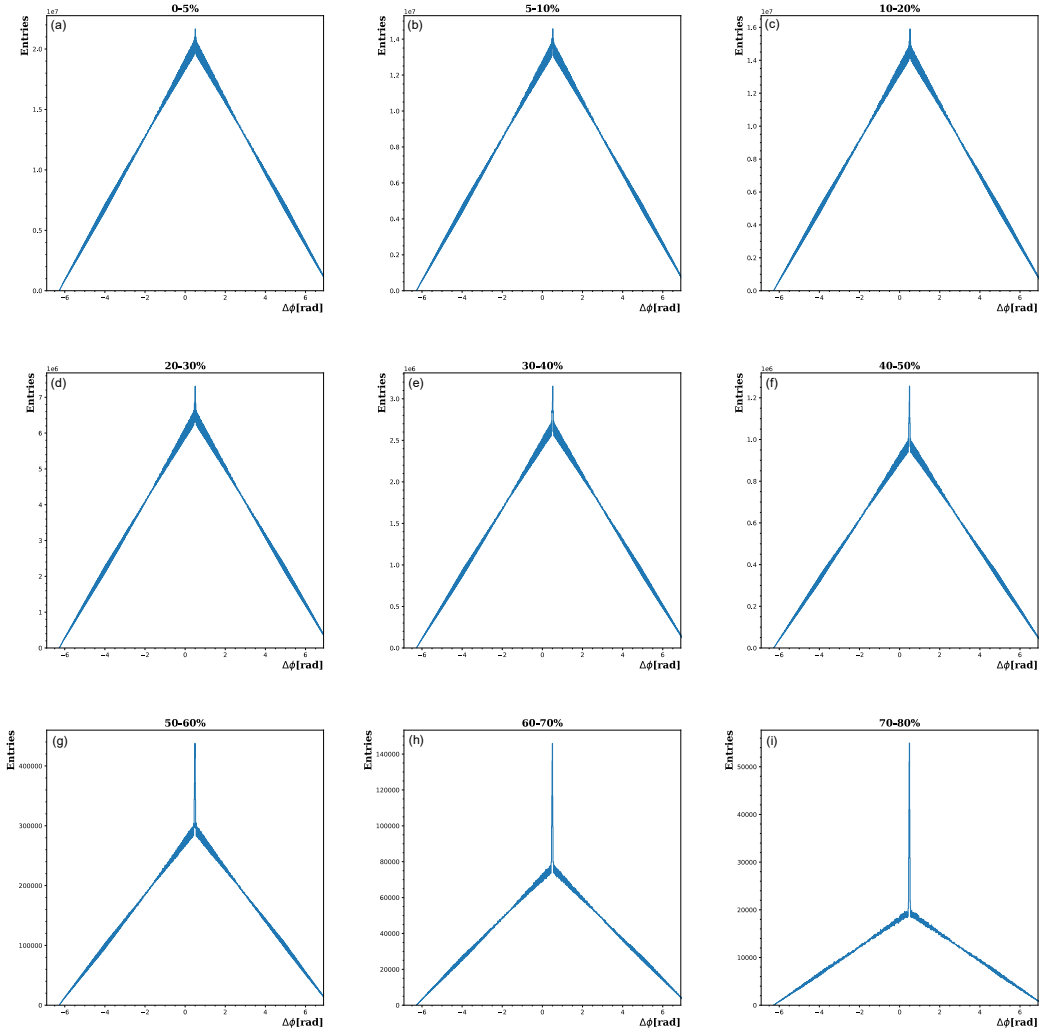

Figure 4. Centrality dependence of the Δφ distribution. (a) – (i) correspond to percentile ranges of 0-5%, 5-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, and 70-80% respectively.

DOI: 10.54254/2753-8818/43/20241139

Moving forward, a more in-depth quantitative analysis will be conducted. Firstly, in order to make the final results more scientific and rigorous, the algorithm needs to be improved by considering η and φ together. Define the distance $\Delta R = \sqrt{(\Delta\eta^2 + \Delta\varphi^2)}$, and incorporate this calculation process into the kernel function. Similarly, the refined algorithm fed by nine event collections will return the distributions of ΔR for different centrality classes. When ΔR is close to zero, the distribution shape is expected to appear as a linear curve, with a bulge emerging near the origin, as shown in Figure 5(a). Clearly, below the straight line is the background noise, and the mountain-like area represents the portion that is anticipated to be retained. Removing the background from the original pattern leaves the selected tracklets, as shown in Figure 5(b). For a given centrality class, the integral of the remaining part corresponds to the total number of charged particles produced in this group of events.

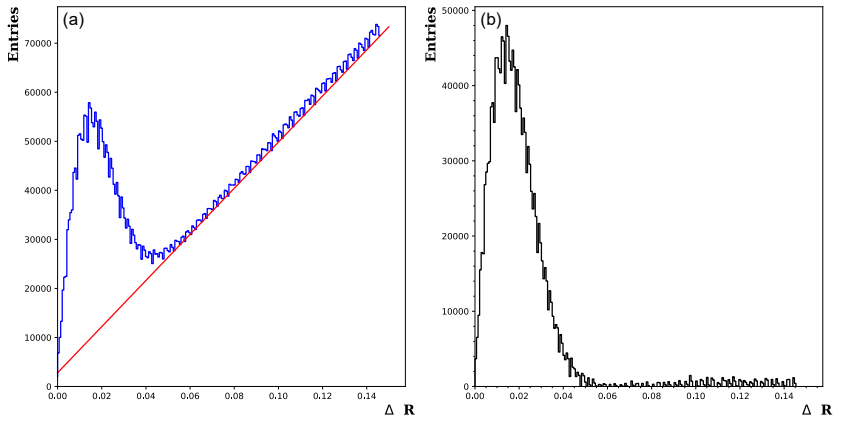

Figure 5. Example for background subtraction. (a) Δ*R* distribution with a fitted curve (red line) for the 30-40% centrality class, $\Delta R \in [0.00, 0.15]$; (b) Δ*R* distribution after background removal.

Considering that the number of events varies across different centrality classes, it is more reasonable to use the charged-particle multiplicity, which represents the number of generated charged particles per event, for subsequent comparative analysis. In the algorithm, the variable N_{ch} is estimated by dividing the total number of particles by the total number of events within one centrality class. Similarly, the measured pseudo-rapidity range may also differ in different experiments. Therefore, in practical analysis, scientists often compute the charged-particle multiplicity density, denoted as $dN_{ch}/d\eta$, and its normalized form $(dN_{ch}/d\eta)/(\langle N_{part}\rangle/2)$. In the algorithm, the variable $dN_{ch}/d\eta$ is estimated by dividing N_{ch} by the range width of η. The calculation results are then displayed in Table 1. In terms of the table, Figure 6 illustrates a relationship diagram between $\langle N_{part}\rangle$ and $(dN_{ch}/d\eta)/(\langle N_{part}\rangle/2)$. A phenomenological trend is revealed: as the centrality increases, the charged-particle multiplicity density per participant-pair also rises.

Table 1. $dN_{ch}/d\eta$ and $(dN_{ch}/d\eta)/(\langle N_{part}\rangle/2)$ within different centrality classes.

Centrality	$\langle N_{part}\rangle$	$N_{ch}/d\eta$	$(dN_{ch}/d\eta)/(\langle N_{part}\rangle/2)$
0-5%	382.8	1261.4	6.6
5-10%	329.7	1072.7	6.5
10-20%	260.5	827.7	6.4
20-30%	186.4	570.2	6.1
30-40%	128.9	376.4	5.8
40-50%	85.0	230.1	5.4
50-60%	52.8	128.4	4.9
60-70%	30.0	67.6	4.5
70-80%	15.8	31.5	4.0

DOI: 10.54254/2753-8818/43/20241139

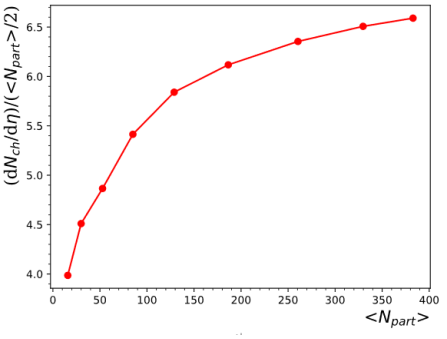

Figure 6. Centrality dependence of the normalized charged-particle multiplicity density.

4. Conclusion

With the adoption of the tracklet method and the introduction of GPU acceleration, this paper successfully estimated the charged-particle multiplicity density and demonstrated its evolution as a function of centrality. The resulting curve can ultimately be utilized for screening and improving existing theoretical models. However, the remaining problem is that the numerical estimation without uncertainty is not yet scientifically rigorous. It is worth highlighting that this article has arrived at experimental conclusions similar to those found in Ref. [10], with the use of distinct datasets. So far, within a broader pseudo-rapidity range, there has been no significant evolution in the shape of the curve above [12]. Human understanding of the mechanism behind such particle production remains unclear. There must be deeper physical connections waiting for researchers to explore, which are essential for advancing the understanding of QCD and QGP.

References

[1] Pasechnik, Roman, and Michal Šumbera. "Phenomenological review on quark–gluon plasma: concepts vs. observations." Universe 11 (2017): 7.

[2] Rafelski, Johann. "Melting Hadrons, Boiling Quarks." arXiv preprint arXiv:1508.03260 (2015).

[3] Csernai, László P. "Introduction to Relativistic heavy ion collisions." Jonh Wiley and Sons Ltd., Chichester, England (2008).

[4] Busza, Wit, Krishna Rajagopal, and Wilke Van Der Schee. "Heavy ion collisions: the big picture and the big questions." Annual Review of Nuclear and Particle Science 68 (2018): 339-376.

[5] Aamodt, Kenneth, et al. "The ALICE experiment at the CERN LHC." Journal of Instrumentation 3.08 (2008): S08002.

[6] Alice Collaboration. "Performance of the ALICE experiment at the CERN LHC." International Journal of Modern Physics A 29.24 (2014): 1430044.

[7] Elia, D., et al. "Performance of ALICE silicon pixel detector prototypes in high energy beams." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 565.1 (2006): 30-35.

[8] Abelev, Betty, et al. "Centrality determination of Pb-Pb collisions at √sNN = 2.76 TeV with ALICE." Physical Review C 88.4 (2013): 044909.

[9] Miller, Michael L., et al. "Glauber modeling in high-energy nuclear collisions." Annu. Rev. Nucl. Part. Sci. 57 (2007): 205-243.

[10] Aamodt, Kenneth, et al. "Centrality dependence of the charged-particle multiplicity density at mid-rapidity in Pb-Pb collisions at √sNN = 2.76 TeV." Physical review letters 106.3 (2011): 032301.

[11] Elia, D., et al. The pixel detector based tracklet reconstruction algorithm in ALICE. No. ALICE-INT-2009-021. 2009.

[12] Adam, Jaroslav, et al. "Centrality evolution of the charged-particle pseudorapidity density over a broad pseudorapidity range in Pb–Pb collisions at √sNN = 2.76 TeV." Physics Letters B 754 (2016): 373-385.