Influence of statistics on the measured moments of conserved quantities in relativistic heavy ion collisions

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 J. Phys. G: Nucl. Part. Phys. 42 065103
(http://iopscience.iop.org/0954-3899/42/6/065103)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 202.114.35.25
This content was downloaded on 08/05/2015 at 03:06

Please note that terms and conditions apply.
Influence of statistics on the measured moments of conserved quantities in relativistic heavy ion collisions

Lizhu Chen$^{1,2}$, Zhiming Li$^2$, Xia Zhong$^1$, Yuncun He$^3$ and Yuanfang Wu$^2$

$^1$ School of Physics and Optoelectronic Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, People’s Republic of China
$^2$ Key Laboratory of Quark and Lepton Physics (MOE) and Institute of Particle Physics, Central China Normal University, Wuhan 430079, People’s Republic of China
$^3$ Faculty of Physics and Electronic Technology, Hubei University, Wuhan 430079, People’s Republic of China

E-mail: wuyf@phy.ccnu.edu.cn

Received 27 January 2015, revised 26 March 2015
Accepted for publication 30 March 2015
Published 7 May 2015

Abstract

We study statistics dependence of the probability distributions and the means of measured moments of conserved quantities, respectively. The required statistics of all interested moments and their products are estimated based on a simple simulation. We also explain why the measured moments are underestimated when the statistics are insufficient. With the statistics at RHIC/BES, the second and third order moments can be reliably obtained based on the method of centrality bin width correction (CBWC), which can not be applied for the fourth order moments at low energy. With planning statistics at RHIC/BES II, and improved CBWC method, $\kappa\sigma^2$ in a finer centrality bin scale should be measurable. This will help us to understand the current observation of energy and centrality dependence of high-order moments.

Keywords: statistics, probability distribution, moments of conserved quantities

(Some figures may appear in colour only in the online journal)

1. Introduction

The main goal of the RHIC Beam Energy Scan (BES) program is to search for the quantum chromodynamics (QCD) critical point and phase boundary [1]. The calculations from lattice QCD and QCD based models have shown that the high-order moments of conserved
quantities, such as net-baryon, net-charge and net-strangeness, are sensitive to the critical fluctuations [2–7]. So the behavior of high-order moments of conserved quantities are greatly interested in heavy ion collisions [8–10].

However, the moments of conserved quantities are also affected by many non-critical backgrounds [11], such as global baryon/charge conservation. The moments are substantially suppressed if global conservation is taken into consideration [12]. In addition, artificial cuts of phase space in experimental analysis are not trivial [13], such as the acceptance in rapidity and transverse momentum space, the selection of centrality of the collisions, and so on.

To reduce those non-critical background contributions, many new techniques are applied in moments analysis at RHIC/STAR [14, 15]. The delta theorem method is used for estimation of statistical uncertainties [16]. The track efficiency correction [17] is firstly considered in an event-by-event fluctuation. The initial size of the fluctuation is one of the important non-critical effects. Although it has not been taken into consideration in most of the theoretical calculations [18–20], it exists in all event variables in experiments. As we know, the initial size is usually quantified by the centrality of the collisions, which is determined by the multiplicity \( N_{ch} \) of the final state. A centrality bin corresponds to a range of multiplicity. For a given centrality bin, i.e., a range of multiplicity, the initial size still fluctuates from event to event [21]. In order to reduce this initial size fluctuation, it is proposed to calculate the moment in each of \( N_{ch} \). The measured moment is averaged over all \( N_{ch} \) in a given centrality, where the average is weighted by the number of events in each of \( N_{ch} \). This is the so-called method of the centrality bin width correction (CBWC) [22].

The CBWC method can only be reliably used if the statistics in each of \( N_{ch} \) are large enough [23]. It is shown that measured moments increase with statistics and finally saturate when statistics are sufficient. If the statistics are insufficient, the measured moments will be underestimated. In fact, it is already found that the fourth order moments would be underestimated at RHIC low energies. In order to increase the statistics, it is suggested to calculate the moment at \( \delta 1\% \) centrality bin width, instead of each of \( N_{ch} \), in the CBWC method, i.e., the so-called improved CBWC method [23]. It can reduce the initial size fluctuations as well.

The published STAR results show \( \kappa \sigma^2 \) of net-proton multiplicity distributions in central collisions (0–5% and 5–10% centrality) is systematically lower than that in the other centralities [14]. The behavior of \( \kappa \sigma^2 \) in a finer centrality bin (such as \( \delta 1\% \) centrality) should be further investigated. At planning RHIC/BES II, the statistics will be increased to a few hundred millions, see table 1 [24]. With these large statistics, and improved CBWC method, it is possible to study the centrality dependence of high-order moments in a finer centrality bin scale, i.e., in each \( \delta 1\% \) centrality bin width for central collisions (0–10% centrality). If this can be done, the centrality dependence of high-order moments can be better understood.

In this paper, we firstly study the statistics dependence of probability distribution of measured moments in section 2. It is shown that when the probability distribution of measured moments can be described by a normal one with goodness of fit \( \chi^2/ndf < 10 \), the corresponding statistics are enough for the moment analysis. The required statistics in each of \( N_{ch} \) or \( \delta 1\% \) centrality bin width for all interested moments and their product is estimated. The statistics dependence of the means of measured moments are studied in section 3. The required statistics are obtained when the mean of the measured moment is saturated to its expectation. By this method, the required statistics are lower than those obtained from probability distribution of measured moments. It also explains why the measured mean of moment is systematically smaller than its expectation when the statistics are short.

According to the statistics dependence of measured moments given in sections 2 and 3, the reliability of measured moments in current RHIC/BES I, and planned RHIC/BES II are discussed in section 4. Finally, the summary is presented in section 5.
2. Statistics dependence of probability distributions of measured moments

It is well known that the probability distribution of the measurement should be a normal one in the case of the sufficient statistics. For the statistics at current heavy ion collisions, usual observables, such as, $K/\pi$ fluctuation [25], and the mean of various multiplicity, whose probability distributions are already normally distributed.

However, for high-order moments, in particular, the fourth, and even the sixth order moments, the required statistics should largely be increased. How much they should be increased is not yet clear. So in this section we firstly demonstrate how they change with statistics from a simple simulation.

Let’s suppose that the particle and anti-particle are produced independently by a Poisson distribution. Then the net-particle follows a Skellam distribution. The parameters of Skellam distributions are the means of particle and anti-particle [26–28]. In our simulation, they are set to $m_1 = 14.5$, $m_2 = 0.6$, reference to the means of proton and anti-proton at low energies of RHIC/STAR [14].

In figure 1, the distribution of mean $\langle N \rangle$, variance $\sigma$, and skewness $S$ of 100 events of Skellam distribution are shown. The distributions are obtained from 5 million (5M) independent samples and presented by black points in figures 1(a), (b) and (c), respectively. The yellow dashed lines are fitted by a normal distribution.

We can see from figure 1(a) that the distribution of measured mean $\langle N \rangle$ is already a normal one. Its goodness of fit $\chi^2/ndf$ is 1.48 as listed in table 2. So the statistics 100 seems enough for first order moment analysis. While, the distribution of the second order moment $\sigma$ in figure 1(b) is slightly deviated from normal distribution, and its $\chi^2/ndf = 46.3$. The distribution of third order moment $S$ in figure 1(c) can not be fitted by a normal one any more with $\chi^2/ndf = 228.9$. So the higher the order of moments, the larger statistics are required.

Increasing the statistics to magnitude 1,000, and 10,000, we re-calculate the second and third order moments. Their probability distributions are presented in the upper and lower panel of figure 2, respectively. The left and right columns correspond to the statistics 1000, and 10000, respectively. From figure 2(a) with statistics 1,000, we can see that the distribution of $\sigma$ can be fitted by a normal one with $\chi^2/ndf = 4.96$. And figure 2(b) with statistics 10000 gives a better normal distribution with $\chi^2/ndf = 1.55$.

Therefore, it seems that we can relax the usual requirement of goodness of fit from $\chi^2/ndf \sim 1$ to $\chi^2/ndf \sim 10$ in our simulation. In this case, if the distribution of measured moments is fitted by a normal one with goodness of fit $\chi^2/ndf < 10$, the corresponding statistics are sufficient for a reliable analysis. By this standard, the statistic 1000 seems enough for the second order moment analysis. It is a magnitude larger than that of the first order moment.

Figure 2(c) shows the probability distribution of the third order moments with statistics 1000. It can still not be well described by a normal one. The goodness of fit is $\chi^2/ndf = 56.9$, which is still larger than 10. When statistics are increased to 10,000, i.e., figure 2(d), it can be fitted by a normal one with $\chi^2/ndf = 7.97$. So the statistics for the third moments are 10,000. It is again a magnitude larger than that of the second moment.

| Energy (GeV) | 7.7 | 9.1 | 11.5 | 13.0 | 14.6 | 19.6 |
|-------------|-----|-----|------|------|------|------|
| No. Events (Million) | 100 | 160 | 230 | 250 | 300 | 400 |

Table 1. The proposed statistics at RHIC/BES II.

J. Phys. G: Nucl. Part. Phys. 42 (2015) 065103 L Chen et al

3
Obviously, the statistics for the fourth order moment, $\kappa$, has to be larger than 10,000. From table 2, we can see that when the statistics increase to 1 million (1 M), its goodness of fit $\chi^2/ndf = 4.00$. So the statistics for the fourth order moment is around 1 M.

The product of the fourth and second order moment, $\kappa \sigma^2$, is another interesting observable [29]. Its statistics should be mainly determined by the fourth order moment $\kappa$. The probability distributions of measured product with statistics 0.01 M, 0.1 M and 1 M are presented in figures 3(a), (b) and (c), respectively. The yellow dashed lines are fitted normal distributions. It shows clearly that figure 3(a) can not be fitted by a normal distribution, and figure 3(b) is still slightly deviated from it. Only figure 3(c) with statistics 1 M can be fitted by a normal one. Consequently, the statistics of $\kappa \sigma^2$ are 1 M, the same magnitude as those of $\kappa$.

3. Statistics dependence of the means of measured moments

It is known when the statistics are large enough, the mean of measured moments should be saturated and independent of the statistics, i.e.,

$$\langle X \rangle_{n_1} = \langle X \rangle_{n_2} = \langle X \rangle_{n_3},$$

where $X$ is the measured moments of $\langle N \rangle$, $\sigma$, $S$ and $\kappa$. The subscripts $n_1$, $n_2$ and $n_3$ denote different statistics.

In a previous study [23], we defined the difference of the mean of measured moments to its expectation as:

$$d = \left| \frac{\langle X \rangle_{\text{exp}} - \langle X \rangle_{\text{th}}}{\langle X \rangle_{\text{th}}} \right| \times 100\%,$$

where $\langle X \rangle_{\text{th}}$ is theoretical expectation and $\langle X \rangle_{\text{exp}}$ is experimentally measured mean of moments with finite statistics. Consider the statistics at RHIC/BES I and the uncertainty of the measurements, the result is acceptable if $d < 5\%$ [23]. The statistics at RHIC/BES II have a few hundred millions, which is 1–2 magnitudes larger than those at RHIC/BES I. So the
statistical uncertainty will be largely reduced. In this case, $d < 1\%$ is more reasonable within smaller uncertainties at planning RHIC/BES II.

The left panel of table 3 lists statistics dependence of $\langle \sigma \rangle$ and corresponding difference $d$, respectively. It shows that the difference $d$ is already less than 1% when statistics is 100. This statistics is less than 1000, which is obtained in the above section by requiring a normal distribution with the goodness of fit $\chi^2/ndf < 10$. So the normal distribution of the measured moment is a more restrictive requirement for the statistics.

Table 2. $\chi^2/ndf$ of normally fit for probability distributions of $\langle N \rangle$, $\sigma$, $S$ and $\kappa$ with statistics 100, 1000, 10000, 100000 and 1 M, respectively.

| Statistics | 100  | 1000 | 10000 | 100000 | 1 M  |
|------------|------|------|-------|--------|------|
| $\chi^2/ndf$ (Mean) | 1.48 | 1.08 | 1.05  | 0.99   | 1.00 |
| $\chi^2/ndf$ ($\sigma$) | 46.3 | 4.96 | 1.55  | 0.97   | 0.95 |
| $\chi^2/ndf$ ($S$) | 228.9| 56.9 | 7.97  | 2.08   | 0.96 |
| $\chi^2/ndf$ ($\kappa$) | 4962 | 1955 | 321   | 37.6   | 4.00 |

Figure 2. Probability distributions of the measured $\sigma$ (upper panel), and $S$ (lower panel) with statistics: 1000 (left column), and 10000 (right column), respectively. The yellow dashed lines are the corresponding normal fittings.
In figures 4(a) and (b), the statistics dependence of \( \kappa \sigma^2 \) for two sets of Skellam parameters: (a) \( m_1 = 14.5 \) and \( m_2 = 0.6 \) and (b) \( m_1 = 79.0 \) and \( m_2 = 65.0 \) (reference to the means of charge and anti-charge within RHIC/STAR acceptance [15]) are presented, respectively. With an increase of statistics, \( \kappa \sigma^2 \) in both figures 4(a) and (b) increases firstly and then converges to their expectation unity. It converges more quickly in figure 4(a) than that in figure 4(b). So the required statistics are related to parent distribution.

The statistics dependence of a few typical \( \kappa \sigma^2 \) in figure 4(a) is given in the right panel of table 3. It shows that the difference \( d \) is less than 1%, when statistics are above 10,000. It is about two magnitudes smaller than that obtained from the above section. So the normal distribution of the measured moment is also more restrictive for high-order moments.

Figure 3. Probability distributions of \( \kappa \sigma^2 \).

Figure 4. Statistics dependence of \( \langle \kappa \sigma^2 \rangle \), which are all averaged from 5 M independent samples randomly generated from the simulations of Skellam distributions with the input parameters: (a) \( m_1 = 14.5 \) and \( m_2 = 0.6 \) and (b) \( m_1 = 79.0 \) and \( m_2 = 65.0 \), respectively.
Figures 4(a) and (b) demonstrate that the mean of measured moments are underestimated, when statistics is insufficient. The poorer the statistics, the smaller the mean. To search for the reasons, the probability distributions of measured $\kappa \sigma^2$ at statistics 100, 500 and 10000 are presented in figures 5(a), (b), and (c), respectively. The black dashed lines shows the peak of the distributions. The blue lines are the mean of their probability distribution, i.e., $\langle \kappa \sigma^2 \rangle$. The red solid lines are theoretical expectation. Figure 5(a) shows that the peak (black dashed line) and mean (blue line) of the distribution are on the left side of the expectation (unity). The probability on the left side of unity is systematically higher than that on the right side. So the measured moment is underestimated when the statistics are poor. This is due to the lost information of the tail of net-particle multiplicity distributions, which can be found from figures 6. Figures 6(a), (b) and (c) show the net-particle multiplicity distributions with statistics 1000, 10000, and 0.1 M. These three samples are randomly selected from the sub-samples used for calculating moments in this paper. It shows that the tail of the net-particle multiplicity distributions is dependent on the statistics. With insufficient statistics, the tail of net-particle multiplicity distributions shown in figure 6(a) and (b) can not be well detected. The smaller the statistics, the poorer information about the tail of the net-particle multiplicity distributions we can detect. With the statistics increased, this information can be better seen, such as shown in figure 6(c). That is why $\langle \kappa \sigma^2 \rangle$ decreases as the statistics decreases in case of insufficient statistics shown in figure 4.

The influence of tail of net-particle multiplicity distributions on measured moments is similar to what had been observed at measurement of sixth order moment, where the tail of the net-proton multiplicity distributions is artificially truncated [30]. It also shows that the more the lost information at the tail of the distribution, the smaller $C_6/C_2$ is measured.

| Statistics 100 | 500 | 1000 | Statistics 1000 | 2000 | 5000 | 10000 |
|---------------|-----|------|---------------|------|------|-------|
| $\langle \sigma \rangle$ | 3.85629 | 3.87996 | 3.88286 | $\langle \kappa \sigma^2 \rangle$ | 0.893 | 0.946 | 0.979 | 0.990 |
| $d$ (%) | 0.76 | 0.15 | 0.077 | $d$ (%) | 10.7 | 5.45 | 2.15 | 0.97 |
If the statistics are sufficient, the peak, mean and expectation, i.e., the three lines in figure 5, should be identical. So we can see from figures 5(a) and (b), and finally to figure 5(c) that they become closer and closer.

In addition, the statistical uncertainty of figure 5, i.e., the widths of probability distributions, looks much larger than the deviation of the average (blue dashed lines in figure 5) from the expectation (red solid lines). However, it is the statistical uncertainty of one sub-sample, similar to the experimental measurement at a fixed multiplicity. An experimental point at a given centrality is averaged over all multiplicities in the centrality, analogous to the case of figure 4. The corresponding statistical uncertainty is usually smaller than those given in figure 5. Moreover, it has been shown in figure 1 in [23] that the fluctuation of many sub-samples are out of $\sigma$ of expectation. So the deviation of average from its expectation due to insufficient statistics can not be ignored.

### 4. Measurement of $\kappa\sigma^2$ at RHIC/BES I and II

Until now, we have only considered the required statistics for a given distribution. In fact, this corresponds to statistics in each of $N_{ch}$ with CBWC method, or in each $\delta 1\%$ centrality bin with improved CBWC method. The statistics of a real sample in minimum bias has to be larger than this. For example, the required statistics for $\{\kappa\sigma^2\}$ of net-charge is about 0.1 M (cf figure 4 and table 3), while the corresponding statistics in whole minimum bias sample is around 200 M, assuming there is 100 multiplicity bin in 0–5% centrality [31]. That is why we propose the improved CBWC method on moment analysis.

As we have mentioned in motivation, another advantage of the improved CBWC method with each $\delta 1\%$ centrality bin is that we can further investigate why $\kappa\sigma^2$ in central collisions is systematically lower than the other centralities. With tens of millions of statistics at RHIC/BES I, the number of events in each $\delta 1\%$ is around 0.1 M, similar statistics as figure 3(b). If we study the behavior of $\kappa\sigma^2$ in each $\delta 1\%$ centrality bin, the statistical uncertainty shown in figure 3(b) is around 0.29. Consequently, the statistical uncertainty is already larger than the measured difference of centrality dependence of $\kappa\sigma^2$ shown in [14].
At RHIC/BES II, the statistics in minimum bias will be increased to 100–400 millions, cf table 1. The statistics in the suggested $\delta 1\%$ centrality bin will be about 1–4 millions. The width in figure 3(c) is about 0.1, which would be slightly larger than statistical uncertainties at RHIC/BES II. So in this case, the statistics should be sufficient for $\kappa \sigma^2$ analysis in each $\delta 1\%$ centrality bin. This will be helpful to clarify why measured $\kappa \sigma^2$ of net-proton multiplicity distributions in central collisions (0–5% and 5–10%) is much lower than those in the other centralities.

5. Summary

In summary, the moments of conserved quantities are important measurements to study QCD phase structure and bulk properties. If the statistics are insufficient, the shape of net-proton/net-charge multiplicity distributions, especially for the tail of the distributions, can not be well detected. Consequently, the measured moments are underestimated and the obtained moments probability distributions would be deviated to normal distribution.

In this paper, statistics dependence of the probability distributions and the means of measured moments and their product $\kappa \sigma^2$ are systematically presented. The behavior of simulation in this paper can be referenced to moments of net-proton multiplicity distributions at RHIC low energies. It is found that when the probability distribution of the measured moment can be fitted by a normal one with goodness of fit $\chi^2/\text{ndf} < 10$, the corresponding statistics are enough for the moment analysis. The required statistics can also be obtained when the mean of the measured moment is saturated to its expectation within statistical error. In this way, the required statistics are 1–2 magnitude less than those obtained from probability distribution of measured moments. So the normal distribution of the measured moment is a more restrictive requirement for the statistics.

Currently, the CBWC method in each of $N_{ch}$ is employed to moments analysis. With the restrictive requirement of the normally probability distribution, it is found that the statistics at RHIC/BES I is enough to extract the second and third order moments. However, at RHIC low energy, the statistics is not enough for $\kappa \sigma^2$ either from the restrictive normally distribution or quite looser requirement obtained from the saturation of the mean of measured moments. The measured moment will be underestimated if statistics are not sufficient. With the improved CBWC method in each $\delta 1\%$ centrality bin, we can obtain much reliable $\kappa \sigma^2$ at RHIC low energies.

The current measured $\kappa \sigma^2$ of net-proton multiplicity distributions shows that $\kappa \sigma^2$ in central collisions is systematically lower than that in the other centralities. With planning statistics for RHIC/BES II, and improved CBWC method, it is possible to measure the high moments of conserved quantities in finer centrality bin scale. This will help us to understand the observation of the centrality and energy dependence of high-order moments.

Acknowledgments

This work is supported in part by the Major State Basic Research Development Program of China under Grant No. 2014CB845402, the NSFC of China under Grants No. 11405088, No. 11221504, No. 11005046, No. 11447190, and the Ministry of Education of China with Project No. 20120144110001.
References

[1] Aggarwal M M et al (STAR Collaboration) arXiv: 1007.2613
[2] Stephanov M A 2009 Phys. Rev. Lett. 102 032301
[3] Cheng M et al 2009 Phys. Rev. D 79 074505
[4] Ejiri S, Karsch F and Redlich K 2006 Phys. Lett. B 633 275
[5] Aggarwal M M et al (STAR Collaboration) 2010 Phys. Rev. Lett. 105 022302
[6] Skokov V, Friman B and Redlich K 2012 Phys. Lett. B 708 179
[7] Friman B et al 2011 Eur. Phys. J. C 71 1694
[8] Bazavov A et al (Hot QCD Collaboration) 2012 Phys. Rev. Lett. 109 192302
[9] Karsch F 2012 Central Eur. J. Phys. 10 1234
[10] Gupta S et al 2011 Science 332 1525
[11] Koch V 2014 Critical point and onset of deconfinement (University of Bielefeld, Germany) (www2.physik.uni-bielefeld.de/cpod2014.html)
[12] Bzdak A, Koch V and Skokov V 2013 Phys. Rev. C 87 044901
[13] Luo X, Xu J, Mohanty B and Xu N 2013 J. Phys. G 40 105104
[14] Adamczyk L et al (STAR Collaboration) 2014 Phys. Rev. Lett. 112 032302
[15] Adamczyk L et al (STAR Collaboration) 2014 Phys. Rev. Lett. 113 092301
[16] Luo X 2012 J. Phys. G. 39 025008
[17] Bzdak A and Koch V 2012 Phys. Rev. C 86 044904
[18] Skokov V, Friman B and Redlich K 2013 Phys. Rev. C 88 034911
[19] Gavai R V and Gupta S 2011 Phys. Lett. B 696 459
[20] Karsch F and Redlich K 2011 Phys. Lett. B 695 136
[21] Heiselberg H 2001 Phys. Rep. 351 161
[22] Luo X (For the STAR Collaboration) 2011 J. Phys. : Conf. Ser. 316 012003
[23] Chen L, Li Z and Wu Y 2014 J. Phys. G 41 105107
[24] Cebra D 2014 Critical point and onset of deconfinement (University of Bielefeld, Germany) (www2.physik.uni-bielefeld.de/cpod2014.html)
[25] Abelev B I et al (STAR Collaboration) 2009 Phys. Rev. C 103 92301
[26] Skellam J G 1946 J. Roy. Stat. Soc. 109 296
[27] Braun-Munzinger P et al 2011 Phys. Rev. C 84 064911
[28] Pan X et al 2014 Phys. Rev. C 89 044904
[29] Stephanov M A 2009 Phys. Rev. Lett. 102 032301
[30] Chen L (For the STAR Collaboration) 2013 Nucl. Phys. A 904 471c
[31] Adamczyk L et al (STAR Collaboration) 2013 Phys. Rev. C 88 014902