

More efficient formulas for efficiency correction of cumulants and effect of using averaged efficiencyToshihiro Nonaka,¹ Masakiyo Kitazawa,^{2,3} and Shinichi Esumi¹¹*Center for Integrated Research in Fundamental Science and Engineering, University of Tsukuba, Tsukuba, Ibaraki 305, Japan*²*Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan*³*J-PARC Branch, KEK Theory Center, Institute of Particle and Nuclear Studies, KEK, 203-1, Shirakata, Tokai, Ibaraki 319-1106, Japan*

(Received 25 February 2017; revised manuscript received 1 May 2017; published 26 June 2017)

We derive formulas for the efficiency correction of cumulants with many efficiency bins. The derivation of the formulas is simpler than the previously suggested method, but the numerical cost is drastically reduced from the naïve method. From analytical and numerical analyses in simple toy models, we show that use of the averaged efficiency in the efficiency correction might cause large deviations in some cases and should not be used, especially for high order cumulants. These analyses show the importance of carrying out the efficiency correction without taking the average.

DOI: [10.1103/PhysRevC.95.064912](https://doi.org/10.1103/PhysRevC.95.064912)**I. INTRODUCTION**

One of the major goals of heavy ion colliding experiments is to reveal the QCD phase structure. Event-by-event fluctuations of conserved quantities, such as net-baryon and net-charge distributions, have been proposed as experimental probes of the signal from the QCD critical point and phase transitions [1–6]; see recent reviews Refs. [7,8]. The STAR experiment has measured the beam energy dependencies of the third and fourth order cumulants of net-proton and net-charge multiplicity distributions [9,10]. In these studies enhancement and suppression of the cumulants in the low energy region are observed, which might be the signal of the critical point. However, there are still large statistical and systematic errors, especially in the low energy region. The Beam Energy Scan Phase II program is planned at the Relativistic Heavy Ion Collider (RHIC) to accumulate a large amount of statistics. In addition, the experimental group is also trying to measure the sixth order cumulant to find the signal of the phase transition [11].

One of the experimental difficulties in the analyses of higher order cumulants is concerned with finite detector efficiencies. We miss particles with some probability called efficiency, and the imperfect efficiency affects the shape of the event-by-event distributions and their cumulants [12,13]. The correction of this effect has been discussed in the literature [12–18]. Moreover, in real detectors the efficiency often becomes nonuniform for detector acceptance due to many reasons, e.g., detector structures, or detector conditions for certain regions. In this case, the nonuniformity of the efficiency should be taken into account in the correction. Although efficiency corrections with nonuniform efficiencies are proposed in Refs. [14–16], these methods are difficult to apply to higher order cumulants. In the method proposed in Refs. [14,15], the numerical cost grows proportional to M^m , where M is the number of efficiency bins and m is the order of the cumulant. Therefore, we cannot increase the numbers of M and m within a realistic CPU time. In Ref. [16], another method which drastically reduces the numerical cost has been proposed on the basis of cumulant expansion. In this method, however, the derivation of the analytic formulas becomes complicated for higher order cumulants and it is difficult to apply the method to sixth order.

Therefore, an alternative efficient method for this problem is called for.

In the present paper, we propose a new method for the efficiency correction. In this method, the analytic procedure is substantially simplified compared to Ref. [16]. The numerical cost, however, is almost the same as that in Ref. [16], and drastically smaller than those in Refs. [14,15].

We also apply the formulas to practical problems in experiments. Generally, acceptance uniformity of detectors can be violated due to various practical problems. In this case, it is desirable to divide the efficiency bin into different acceptance regions and apply the efficiency correction with the increased bins. However, it is practically difficult to implement those corrections because of the large numerical cost, especially for higher order cumulants. Then, one has to use a single averaged efficiency for these acceptance regions. In this paper, we study the effect of using the averaged efficiency in simple toy models. We show that use of the averaged efficiency in the efficiency correction might cause large deviations in some cases and should not be used, especially for high order cumulants.

This paper is organized as follows. In Secs. II and III, we derive formulas for the efficiency correction with many efficiency bins. We first derive the result in a simple case with a single efficiency in Sec. II, and then extend it to the multivariate case in Sec. III. In Sec. IV, analytic calculations are performed to study the effect of using averaged efficiency in a toy model. In Sec. V, we study this effect numerically in toy models assuming net-charge fluctuation and nonuniform acceptance.

II. SINGLE VARIABLE CASE

Although the main goal of this paper is to derive formulas for the efficiency correction with many efficiency bins, in this section we start from the case with a single efficiency bin, because this analysis serves as a simple illustration of the multivariate case, which will be addressed in the next section.

A. Cumulants and factorial cumulants

Let us consider a probability distribution function $P(N)$ for an integer stochastic variable N . The m th order cumulant of

$P(N)$ is defined as [7]

$$\langle N^m \rangle_c = \left. \frac{\partial^m}{\partial \theta^m} K(\theta) \right|_{\theta=0}, \quad (1)$$

with the cumulant generating function

$$K(\theta) = \ln \sum_N e^{N\theta} P(N) = \ln \langle e^{N\theta} \rangle. \quad (2)$$

In this study, we fully make use of another set of quantities called factorial cumulants $\langle N^m \rangle_{fc}$, which are defined as

$$\langle N^m \rangle_{fc} = \left. \frac{\partial^m}{\partial s^m} K_f(s) \right|_{s=1}, \quad (3)$$

with the factorial-cumulant generating function

$$K_f(s) = \ln \langle s^N \rangle. \quad (4)$$

Cumulants can be represented by the sum of factorial cumulants, and vice versa. To obtain these relations, it is convenient to use the fact that the generating functions (2) and (4) are related to each other by the change of variables $s = e^\theta$ or $\theta = \ln s$. The mutual derivatives of s and θ are given by

$$\left. \frac{\partial^m}{\partial \theta^m} s \right|_{s=1} = 1, \quad (5)$$

$$\left. \frac{\partial^m}{\partial s^m} \theta \right|_{\theta=0} = (-1)^{m-1} (m-1)!. \quad (6)$$

Using Eq. (5), the first and second order cumulants are converted into factorial cumulants as

$$\langle N \rangle_c = \frac{\partial}{\partial \theta} K = \frac{\partial s}{\partial \theta} \frac{\partial K_f}{\partial s} = \langle N \rangle_{fc}, \quad (7)$$

$$\begin{aligned} \langle N^2 \rangle_c &= \frac{\partial^2}{\partial \theta^2} K = \frac{\partial}{\partial \theta} \left(\frac{\partial s}{\partial \theta} \frac{\partial K_f}{\partial s} \right) \\ &= \frac{\partial^2 s}{\partial \theta^2} \frac{\partial K_f}{\partial s} + \left(\frac{\partial s}{\partial \theta} \right)^2 \frac{\partial^2 K_f}{\partial s^2} = \langle N^2 \rangle_{fc} + \langle N \rangle_{fc}, \end{aligned} \quad (8)$$

where it is understood that $\theta = 0$ or $s = 1$ is substituted. Repeating the same manipulation, we obtain the relations up to sixth order as

$$\langle N^3 \rangle_c = \langle N^3 \rangle_{fc} + 3\langle N^2 \rangle_{fc} + \langle N \rangle_{fc}, \quad (9)$$

$$\langle N^4 \rangle_c = \langle N^4 \rangle_{fc} + 6\langle N^3 \rangle_{fc} + 7\langle N^2 \rangle_{fc} + \langle N \rangle_{fc}, \quad (10)$$

$$\begin{aligned} \langle N^5 \rangle_c &= \langle N^5 \rangle_{fc} + 10\langle N^4 \rangle_{fc} + 25\langle N^3 \rangle_{fc} \\ &\quad + 15\langle N^2 \rangle_{fc} + \langle N \rangle_{fc}, \end{aligned} \quad (11)$$

$$\begin{aligned} \langle N^6 \rangle_c &= \langle N^6 \rangle_{fc} + 15\langle N^5 \rangle_{fc} + 65\langle N^4 \rangle_{fc} + 90\langle N^3 \rangle_{fc} \\ &\quad + 31\langle N^2 \rangle_{fc} + \langle N \rangle_{fc}. \end{aligned} \quad (12)$$

Using Eq. (6), factorial cumulants can also be expressed in terms of cumulants as

$$\langle N \rangle_{fc} = \langle N \rangle_c, \quad (13)$$

$$\langle N^2 \rangle_{fc} = \langle N^2 \rangle_c - \langle N \rangle_c, \quad (14)$$

$$\langle N^3 \rangle_{fc} = \langle N^3 \rangle_c - 3\langle N^2 \rangle_c + 2\langle N \rangle_c, \quad (15)$$

$$\langle N^4 \rangle_{fc} = \langle N^4 \rangle_c - 6\langle N^3 \rangle_c + 11\langle N^2 \rangle_c - 6\langle N \rangle_c, \quad (16)$$

$$\langle N^5 \rangle_{fc} = \langle N^5 \rangle_c - 10\langle N^4 \rangle_c + 35\langle N^3 \rangle_c - 50\langle N^2 \rangle_c + 24\langle N \rangle_c, \quad (17)$$

$$\begin{aligned} \langle N^6 \rangle_{fc} &= \langle N^6 \rangle_c - 15\langle N^5 \rangle_c + 85\langle N^4 \rangle_c - 225\langle N^3 \rangle_c \\ &\quad + 274\langle N^2 \rangle_c - 120\langle N \rangle_c, \end{aligned} \quad (18)$$

which are summarized in a compact form,

$$\langle N^m \rangle_{fc} = \langle N(N-1)\cdots(N-m+1) \rangle_c. \quad (19)$$

B. Binomial model

Next, we consider the efficiency correction of cumulants in the binomial model [7]. We assume that a multiplicity distribution of a particle number N is given by $P(N)$. We then suppose that individual particles are observed with a probability p , which is independent for different particles. We denote the number of observed particles as n , and the distribution of n as $\tilde{P}(n)$. Then, $\tilde{P}(n)$ is related to $P(N)$ using the binomial distribution function as [7]

$$\tilde{P}(n) = \sum_N P(N) B_{p,N}(n), \quad (20)$$

where the binomial distribution and its factorial-cumulant generating function are given by

$$B_{p,N}(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}, \quad (21)$$

$$\bar{k}_{p,N}(s) = \ln \sum_n s^n B_{p,N}(n) = N \ln(1-p+ps). \quad (22)$$

We call Eq. (20) the binomial model. From Eq. (20), we obtain the factorial-cumulant generating functions of $\tilde{P}(n)$ as

$$\begin{aligned} \tilde{K}_f(s) &= \ln \sum_n \tilde{P}(n) s^n = \ln \sum_N P(N) e^{\bar{k}_{p,N}(s)} \\ &= \ln \sum_N P(N) (1-p+ps)^N = K_f(1+p(s-1)), \end{aligned} \quad (23)$$

where $K_f(s) = \ln \sum_N P(N) s^N$ is the factorial-cumulant generating function of $P(N)$. From Eq. (23), one obtains

$$\left. \frac{\partial^m}{\partial s^m} \tilde{K}_f(s) \right|_{s=1} = p^m \left. \frac{\partial^m}{\partial s^m} K_f(s) \right|_{s=1}. \quad (24)$$

From this result and the definition of the factorial cumulant Eq. (3), we obtain simple relations between the factorial cumulants of $P(N)$ and $\tilde{P}(n)$,

$$\langle n^m \rangle_{fc} = p^m \langle N^m \rangle_{fc}. \quad (25)$$

We note that the same relation holds for factorial moments, which is used in Ref. [13] to derive the formula of efficiency correction.

C. Efficiency correction

In order to perform the efficiency correction, we have to represent $\langle N^m \rangle_c$ in terms of $\langle n^m \rangle_c$. These relations are obtained by the following three steps:

- (1) Convert $\langle N^m \rangle_c$ into $\langle N^m \rangle_{fc}$ using Eqs. (7)–(12).
- (2) Convert $\langle N^m \rangle_{fc}$ into $\langle n^m \rangle_{fc}$ using Eq. (25).
- (3) Convert $\langle n^m \rangle_{fc}$ into $\langle n^m \rangle_c$ using Eqs. (13)–(18).

The specific procedures for the first and second orders are as follows:

$$\langle N \rangle_c = \langle N \rangle_{fc} = \frac{1}{p} \langle n \rangle_{fc} = \frac{1}{p} \langle n \rangle_c, \quad (26)$$

$$\langle N^2 \rangle_c = \langle N^2 \rangle_{fc} + \langle N \rangle_{fc} = \frac{1}{p^2} \langle n^2 \rangle_{fc} + \frac{1}{p} \langle n \rangle_{fc} = \frac{1}{p^2} (\langle n^2 \rangle_c - \langle n \rangle_c) + \frac{1}{p} \langle n \rangle_c = \frac{1}{p^2} \langle n^2 \rangle_c + \left(\frac{1}{p} - \frac{1}{p^2} \right) \langle n \rangle_c. \quad (27)$$

Similar manipulations up to sixth order are obtained as

$$\langle N^3 \rangle_c = \frac{1}{p^3} \langle n^3 \rangle_c + \left(-\frac{3}{p^3} + \frac{3}{p^2} \right) \langle n^2 \rangle_c + \left(\frac{2}{p^3} - \frac{3}{p^2} + \frac{1}{p} \right) \langle n \rangle_c, \quad (28)$$

$$\langle N^4 \rangle_c = \frac{1}{p^4} \langle n^4 \rangle_c + \left(-\frac{6}{p^4} + \frac{6}{p^3} \right) \langle n^3 \rangle_c + \left(\frac{11}{p^4} - \frac{18}{p^3} + \frac{7}{p^2} \right) \langle n^2 \rangle_c + \left(-\frac{6}{p^4} + \frac{12}{p^3} - \frac{7}{p^2} + \frac{1}{p} \right) \langle n \rangle_c, \quad (29)$$

$$\begin{aligned} \langle N^5 \rangle_c &= \frac{1}{p^5} \langle n^5 \rangle_c + \left(-\frac{10}{p^5} + \frac{10}{p^4} \right) \langle n^4 \rangle_c + \left(\frac{35}{p^5} - \frac{60}{p^4} + \frac{25}{p^3} \right) \langle n^3 \rangle_c + \left(-\frac{50}{p^5} + \frac{110}{p^4} - \frac{75}{p^3} + \frac{15}{p^2} \right) \langle n^2 \rangle_c \\ &+ \left(\frac{24}{p^5} - \frac{60}{p^4} + \frac{50}{p^3} - \frac{15}{p^2} + \frac{1}{p} \right) \langle n \rangle_c, \end{aligned} \quad (30)$$

$$\begin{aligned} \langle N^6 \rangle_c &= \frac{1}{p^6} \langle n^6 \rangle_c + \left(-\frac{15}{p^6} + \frac{15}{p^5} \right) \langle n^5 \rangle_c + \left(\frac{85}{p^6} - \frac{15 \times 10}{p^5} + \frac{65}{p^4} \right) \langle n^4 \rangle_c + \left(-\frac{225}{p^6} + \frac{15 \times 35}{p^5} - \frac{65 \times 6}{p^4} + \frac{90}{p^3} \right) \langle n^3 \rangle_c \\ &+ \left(\frac{274}{p^6} - \frac{15 \times 50}{p^5} + \frac{65 \times 11}{p^4} - \frac{90 \times 3}{p^3} + \frac{31}{p^2} \right) \langle n^2 \rangle_c + \left(-\frac{120}{p^6} + \frac{15 \times 24}{p^5} - \frac{65 \times 6}{p^4} + \frac{90 \times 2}{p^3} - \frac{31}{p^2} + \frac{1}{p} \right) \langle n \rangle_c. \end{aligned} \quad (31)$$

We can extend the manipulation to much higher orders straightforwardly. These relations are equivalent to those given in Ref. [7], which are obtained based on the cumulant expansion.

III. MULTIVARIATE CASE

A. Cumulant and factorial cumulant

Next, we extend the analysis in the previous section to the multivariate case. We consider the probability distribution function

$$P(N) = P(N_1, N_2, \dots, N_M), \quad (32)$$

for M stochastic variables N_1, N_2, \dots, N_M . Here, N_i with different i represent, for example, particle numbers entering detectors which cover different acceptances. We then consider the cumulants of charges, which are given by the linear combination of N_i [16],

$$Q_{(a)} = \sum_{i=1}^M a_i N_i. \quad (33)$$

For example, when one considers the net-baryon number, $a_i = 1$ and -1 for baryons and anti-baryons. For net electric charge, a_i represents the electric charge of particle i .

Defining the cumulant generating function of Eq. (32) as

$$K(\boldsymbol{\theta}) = \ln \left[\sum_N e^{\theta_1 N_1 + \dots + \theta_M N_M} P(N) \right], \quad (34)$$

the m th order cumulant of $Q_{(a)}$ is given by

$$\langle Q_{(a)}^m \rangle_c = \partial_{(a)}^m K(\boldsymbol{\theta})|_{\boldsymbol{\theta}=0} \quad (35)$$

with

$$\partial_{(a)} = \sum_{i=1}^M a_i \frac{\partial}{\partial \theta_i}. \quad (36)$$

Similarly, the mixed cumulants are defined by

$$\begin{aligned} \langle Q_{(a)} Q_{(b)} \rangle_c &= \partial_{(a)} \partial_{(b)} K(\boldsymbol{\theta})|_{\boldsymbol{\theta}=0}, \\ \langle Q_{(a)} Q_{(b)} Q_{(c)} \rangle_c &= \partial_{(a)} \partial_{(b)} \partial_{(c)} K(\boldsymbol{\theta})|_{\boldsymbol{\theta}=0}, \end{aligned} \quad (37)$$

and so forth. The factorial cumulants of $P(N)$ are defined from the generating function [19]

$$K_f(\mathbf{s}) = \ln \left[\sum_N P(N) \prod_{i=1}^M s_i^{N_i} \right] \quad (38)$$

as

$$\begin{aligned} \langle Q_{(a)}^m \rangle_{fc} &= \bar{\partial}_{(a)}^m K_f(\mathbf{s})|_{\mathbf{s}=1}, \\ \langle Q_{(a)} Q_{(b)} \rangle_{fc} &= \bar{\partial}_{(a)} \bar{\partial}_{(b)} K_f(\mathbf{s})|_{\mathbf{s}=1}, \end{aligned} \quad (39)$$

and so forth, with

$$\bar{\partial}_{(a)} = \sum_{i=1}^M a_i \frac{\partial}{\partial s_i}. \quad (40)$$

The relations between cumulants and factorial cumulants can be obtained similarly to the previous section. Using the fact that $K(\boldsymbol{\theta})$ and $K_f(\mathbf{s})$ are connected with each other by the

change of variables $s_i = e^{\theta_i}$, we obtain

$$\langle \mathcal{Q}_{(a)} \rangle_c = \partial_{(a)} K = \sum_{i=1}^M a_i \frac{\partial}{\partial \theta_i} K = \sum_{i=1}^M a_i \frac{\partial s_i}{\partial \theta_i} \frac{\partial}{\partial s_i} K_f = \bar{\partial}_{(a)} K_f = \langle \mathcal{Q}_{(a)} \rangle_{fc}, \quad (41)$$

$$\begin{aligned} \langle \mathcal{Q}_{(a)} \mathcal{Q}_{(b)} \rangle_c &= \partial_{(a)} \partial_{(b)} K = \left(\sum_{i=1}^M a_i \frac{\partial}{\partial \theta_i} \right) \left(\sum_{j=1}^M b_j \frac{\partial}{\partial \theta_j} \right) K = \sum_{i=j=1}^M a_i b_j \left(\frac{\partial s_i}{\partial \theta_i} \frac{\partial}{\partial s_i} \right) \left(\frac{\partial s_j}{\partial \theta_j} \frac{\partial}{\partial s_j} \right) K_f \\ &= \left(\sum_{i=1}^M a_i \frac{\partial s_i}{\partial \theta_i} \frac{\partial}{\partial s_i} \right) \left(\sum_{j=1}^M b_j \frac{\partial s_j}{\partial \theta_j} \frac{\partial}{\partial s_j} \right) K_f + \sum_{i=1}^M a_i b_i \frac{\partial^2 s_i}{\partial \theta_i^2} \frac{\partial}{\partial s_i} K_f = (\bar{\partial}_{(a)} \bar{\partial}_{(b)} + \bar{\partial}_{(ab)}) K_f = \langle \mathcal{Q}_{(a)} \mathcal{Q}_{(b)} \rangle_{fc} + \langle \mathcal{Q}_{(ab)} \rangle_{fc}, \end{aligned} \quad (42)$$

where we assumed $\theta = 0$ or $s = 1$, and in the last line we defined

$$\bar{\partial}_{(ab)} = \sum_{i=1}^M a_i b_i \frac{\partial}{\partial s_i}, \quad \mathcal{Q}_{(ab)} = \sum_{i=1}^M a_i b_i N_i. \quad (43)$$

We also define $\partial_{(ab)}$ and the symbols with more than two subscripts, such as $\mathcal{Q}_{(abc)}$, in a similar manner. This manipulation can be extended straightforwardly to arbitrary higher orders. In this analysis, we use the following relations between θ and s derivatives valid for $s = 1$:

$$\partial_{(a)} = \bar{\partial}_{(a)},$$

$$\partial_{(a)} \partial_{(b)} = \bar{\partial}_{(a)} \bar{\partial}_{(b)} + \bar{\partial}_{(ab)}$$

$$\partial_{(a)} \partial_{(b)} \partial_{(c)} = \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} + \bar{\partial}_{(a)} \bar{\partial}_{(bc)} + \bar{\partial}_{(b)} \bar{\partial}_{(ca)} + \bar{\partial}_{(c)} \bar{\partial}_{(ab)} + \bar{\partial}_{(abc)}, \quad (44)$$

$$\begin{aligned} \partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(d)} &= \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cd)} + (6 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(bcd)} + (4 \text{ comb.})] \\ &\quad + [\bar{\partial}_{(ab)} \bar{\partial}_{(cd)} + (3 \text{ comb.})] + \bar{\partial}_{(abcd)}, \end{aligned} \quad (45)$$

$$\begin{aligned} \partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(d)} \partial_{(e)} &= \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} \bar{\partial}_{(e)} + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(de)} + (10 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cde)} + (10 \text{ comb.})] \\ &\quad + [\bar{\partial}_{(a)} \bar{\partial}_{(bc)} \bar{\partial}_{(de)} + (15 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(bcde)} + (5 \text{ comb.})] + [\bar{\partial}_{(ab)} \bar{\partial}_{(cde)} + (10 \text{ comb.})] + \bar{\partial}_{(abcde)}, \end{aligned} \quad (46)$$

$$\begin{aligned} \partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(d)} \partial_{(e)} \partial_{(f)} &= \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} \bar{\partial}_{(e)} \bar{\partial}_{(f)} + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} \bar{\partial}_{(ef)} + (15 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(def)} + (20 \text{ comb.})] \\ &\quad + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cd)} \bar{\partial}_{(ef)} + (45 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cdef)} + (15 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(bc)} \bar{\partial}_{(def)} + (60 \text{ comb.})] \\ &\quad + [\bar{\partial}_{(ab)} \bar{\partial}_{(cd)} \bar{\partial}_{(ef)} + (15 \text{ comb.})] + [\bar{\partial}_{(abc)} \bar{\partial}_{(def)} + (10 \text{ comb.})] \\ &\quad + [\bar{\partial}_{(ab)} \bar{\partial}_{(cdef)} + (15 \text{ comb.})] + [\bar{\partial}_{(a)} \bar{\partial}_{(bcdef)} + (6 \text{ comb.})] + \bar{\partial}_{(abcdef)}. \end{aligned} \quad (47)$$

Here, (comb.) represents terms obtained by all possible combinations of subscripts, for example,

$$\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cd)} + (6 \text{ comb.}) = \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(cd)} + \bar{\partial}_{(a)} \bar{\partial}_{(c)} \bar{\partial}_{(bd)} + \bar{\partial}_{(a)} \bar{\partial}_{(d)} \bar{\partial}_{(bc)} + \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(ad)} + \bar{\partial}_{(b)} \bar{\partial}_{(d)} \bar{\partial}_{(ac)} + \bar{\partial}_{(c)} \bar{\partial}_{(d)} \bar{\partial}_{(ab)}. \quad (48)$$

Note that the number shows the total number of the combinations. The conversions from factorial cumulants to cumulants can be carried out with the following relations valid for $\theta = 0$:

$$\bar{\partial}_{(a)} = \partial_{(a)}, \quad (49)$$

$$\bar{\partial}_{(a)} \bar{\partial}_{(b)} = \partial_{(a)} \partial_{(b)} - \partial_{(ab)}, \quad (50)$$

$$\bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} = \partial_{(a)} \partial_{(b)} \partial_{(c)} - \partial_{(a)} \partial_{(bc)} - \partial_{(b)} \partial_{(ca)} - \partial_{(c)} \partial_{(ab)} + 2\partial_{(abc)}. \quad (51)$$

$$\begin{aligned} \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} &= \partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(d)} - [\partial_{(a)} \partial_{(b)} \partial_{(cd)} + (6 \text{ comb.})] + 2[\partial_{(a)} \partial_{(bcd)} + (4 \text{ comb.})] \\ &\quad + [\partial_{(ab)} \partial_{(cd)} + (3 \text{ comb.})] - 6\partial_{(abcd)}, \end{aligned} \quad (52)$$

$$\begin{aligned} \bar{\partial}_{(a)} \bar{\partial}_{(b)} \bar{\partial}_{(c)} \bar{\partial}_{(d)} \bar{\partial}_{(e)} &= \partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(d)} \partial_{(e)} - [\partial_{(a)} \partial_{(b)} \partial_{(c)} \partial_{(de)} + (10 \text{ comb.})] + 2[\partial_{(a)} \partial_{(b)} \partial_{(cde)} + (10 \text{ comb.})] \\ &\quad + [\partial_{(a)} \partial_{(bc)} \partial_{(de)} + (15 \text{ comb.})] - 6[\partial_{(a)} \partial_{(bcde)} + (5 \text{ comb.})] - 2[\partial_{(ab)} \partial_{(cde)} + (10 \text{ comb.})] + 24\partial_{(abcde)}, \end{aligned} \quad (53)$$

$$\begin{aligned}
\bar{\partial}_{(a)}\bar{\partial}_{(b)}\bar{\partial}_{(c)}\bar{\partial}_{(d)}\bar{\partial}_{(e)}\bar{\partial}_{(f)} &= \partial_{(a)}\partial_{(b)}\partial_{(c)}\partial_{(d)}\partial_{(e)}\partial_{(f)} - [\partial_{(a)}\partial_{(b)}\partial_{(c)}\partial_{(d)}\partial_{(ef)} + (15 \text{ comb.})] + 2[\partial_{(a)}\partial_{(b)}\partial_{(c)}\partial_{(def)} + (20 \text{ comb.})] \\
&+ [\partial_{(a)}\partial_{(b)}\partial_{(cd)}\partial_{(ef)} + (45 \text{ comb.})] - 6[\partial_{(a)}\partial_{(b)}\partial_{(cdef)} + (15 \text{ comb.})] - 2[\partial_{(a)}\partial_{(bc)}\partial_{(def)} + (60 \text{ comb.})] \\
&- [\partial_{(ab)}\partial_{(cd)}\partial_{(ef)} + (15 \text{ comb.})] + 4[\partial_{(abc)}\partial_{(def)} + (10 \text{ comb.})] \\
&+ 6[\partial_{(ab)}\partial_{(cdef)} + (15 \text{ comb.})] + 24[\partial_{(a)}\partial_{(bcdef)} + (6 \text{ comb.})] - 120\partial_{(abcdef)}. \tag{54}
\end{aligned}$$

B. Efficiency correction in binomial model

Next, we extend the binomial model Eq. (20) to the multivariate case. We suppose that a particle labeled by i is observed with efficiency p_i . Assuming the independence of the efficiencies of individual particles, the probability distribution function $\tilde{P}(\mathbf{n})$ of observed particle numbers n_i is related to $P(N)$ as [16]

$$\tilde{P}(\mathbf{n}) = \sum_N P(N) \prod_{i=1}^M B_{p_i, N_i}(n_i). \tag{55}$$

The factorial-cumulant generating function of $\tilde{P}(\mathbf{n})$ is then given by

$$\tilde{K}_f(\mathbf{s}) = \ln \sum_N P(N) \prod_{i=1}^M [1 + p_i(s_i - 1)]^{N_i} = K_f(\mathbf{s}'), \tag{56}$$

with $s'_i = 1 + p_i(s_i - 1)$. From Eq. (56), one finds that $\bar{\partial}_{(a)}\tilde{K}_f = \bar{\partial}_{(ap)}K_f$ and

$$\bar{\partial}_{(a)}K_f = \bar{\partial}_{(a/p)}\tilde{K}_f, \quad \bar{\partial}_{(a)}\bar{\partial}_{(b)}K_f = \bar{\partial}_{(a/p)}\bar{\partial}_{(b/p)}\tilde{K}_f, \tag{57}$$

and so forth, where it is understood that $s = 1$ is substituted and $\bar{\partial}_{(a/p)} = \sum_{i=1}^M (a_i/p_i)(\partial/\partial s_i)$. Equation (57) connects the factorial cumulants of $\tilde{P}(\mathbf{n})$ and $P(N)$.

For the efficiency correction, one must represent the cumulants of $P(N)$ by those of $\tilde{P}(\mathbf{n})$. Similar to the procedure in Sec. II, these relations are obtained by the following steps:

- (1) Convert a cumulant of $P(N)$ into factorial cumulants.
- (2) Convert the factorial cumulants of $P(N)$ into factorial cumulants of $\tilde{P}(\mathbf{n})$.
- (3) Convert the factorial cumulants of $\tilde{P}(\mathbf{n})$ into cumulants.

The explicit manipulation up to the third order is shown as follows:

$$\langle Q_{(a)} \rangle_c = \langle Q_{(a)} \rangle_{fc} = \langle q_{(a/p)} \rangle_{fc} = \langle q_{(a/p)} \rangle_c, \tag{58}$$

$$\langle Q_{(a)}^2 \rangle_c = \langle Q_{(a)}^2 \rangle_{fc} + \langle Q_{(a^2)} \rangle_{fc} = \langle q_{(a/p)}^2 \rangle_{fc} + \langle q_{(a^2/p)} \rangle_{fc} = \langle q_{(a/p)}^2 \rangle_c - \langle q_{(a^2/p^2)} \rangle_c + \langle q_{(a^2/p)} \rangle_c, \tag{59}$$

$$\begin{aligned}
\langle Q_{(a)}^3 \rangle_c &= \langle Q_{(a)}^3 \rangle_{fc} + 3\langle Q_{(a)}Q_{(a^2)} \rangle_{fc} + \langle Q_{(a^3)} \rangle_{fc} = \langle q_{(a/p)}^3 \rangle_{fc} + 3\langle q_{(a/p)}q_{(a^2/p)} \rangle_{fc} + \langle q_{(a^3/p)} \rangle_{fc} \\
&= \langle q_{(a/p)}^3 \rangle_c - 3\langle q_{(a/p)}q_{(a^2/p^2)} \rangle_c + 2\langle q_{(a^3/p^3)} \rangle_c + 3(\langle q_{(a/p)}q_{(a^2/p)} \rangle_c - \langle q_{(a^3/p^2)} \rangle_c) + \langle q_{(a^3/p)} \rangle_c, \tag{60}
\end{aligned}$$

where we defined the linear combination of n_i as

$$q_{(a)} = \sum_{i=1}^M a_i n_i, \quad q_{(ab/p)} \equiv \sum_{i=1}^M (a_i b_i / p_i) n_i, \tag{61}$$

and so forth. The explicit results up to the sixth order are given by

$$\langle Q \rangle_c = \langle q_{(1,1)} \rangle_c, \tag{62}$$

$$\langle Q^2 \rangle_c = \langle q_{(1,1)}^2 \rangle_c + \langle q_{(2,1)} \rangle_c - \langle q_{(2,2)} \rangle_c, \tag{63}$$

$$\langle Q^3 \rangle_c = \langle q_{(1,1)}^3 \rangle_c + 3\langle q_{(1,1)}q_{(2,1)} \rangle_c - 3\langle q_{(1,1)}q_{(2,2)} \rangle_c + \langle q_{(3,1)} \rangle_c - 3\langle q_{(3,2)} \rangle_c + 2\langle q_{(3,3)} \rangle_c, \tag{64}$$

$$\begin{aligned}
\langle Q^4 \rangle_c &= \langle q_{(1,1)}^4 \rangle_c + 6\langle q_{(1,1)}^2 q_{(2,1)} \rangle_c - 6\langle q_{(1,1)}^2 q_{(2,2)} \rangle_c + 4\langle q_{(1,1)}q_{(3,1)} \rangle_c + 3\langle q_{(2,1)}^2 \rangle_c + 3\langle q_{(2,2)}^2 \rangle_c - 12\langle q_{(1,1)}q_{(3,2)} \rangle_c \\
&+ 8\langle q_{(1,1)}q_{(3,3)} \rangle_c - 6\langle q_{(2,1)}q_{(2,2)} \rangle_c + \langle q_{(4,1)} \rangle_c - 7\langle q_{(4,2)} \rangle_c + 12\langle q_{(4,3)} \rangle_c - 6\langle q_{(4,4)} \rangle_c, \tag{65}
\end{aligned}$$

$$\begin{aligned}
\langle Q^5 \rangle_c &= \langle q_{(1,1)}^5 \rangle_c + 10\langle q_{(1,1)}^3 q_{(2,1)} \rangle_c - 10\langle q_{(1,1)}^3 q_{(2,2)} \rangle_c + 10\langle q_{(1,1)}^2 q_{(3,1)} \rangle_c - 30\langle q_{(1,1)}^2 q_{(3,2)} \rangle_c + 20\langle q_{(1,1)}^2 q_{(3,3)} \rangle_c + 15\langle q_{(2,2)}^2 q_{(1,1)} \rangle_c \\
&+ 15\langle q_{(2,1)}^2 q_{(1,1)} \rangle_c - 30\langle q_{(1,1)}q_{(2,1)}q_{(2,2)} \rangle_c + 5\langle q_{(1,1)}q_{(4,1)} \rangle_c - 35\langle q_{(1,1)}q_{(4,2)} \rangle_c + 60\langle q_{(1,1)}q_{(4,3)} \rangle_c - 30\langle q_{(1,1)}q_{(4,4)} \rangle_c \\
&+ 10\langle q_{(2,1)}q_{(3,1)} \rangle_c - 30\langle q_{(2,1)}q_{(3,2)} \rangle_c + 20\langle q_{(2,1)}q_{(3,3)} \rangle_c - 10\langle q_{(2,2)}q_{(3,1)} \rangle_c + 30\langle q_{(2,2)}q_{(3,2)} \rangle_c - 20\langle q_{(2,2)}q_{(3,3)} \rangle_c + \langle q_{(5,1)} \rangle_c \\
&- 15\langle q_{(5,2)} \rangle_c + 50\langle q_{(5,3)} \rangle_c - 60\langle q_{(5,4)} \rangle_c + 24\langle q_{(5,5)} \rangle_c, \tag{66}
\end{aligned}$$

$$\begin{aligned}
\langle Q^6 \rangle_c = & \langle q_{(1,1)}^6 \rangle_c + 15\langle q_{(1,1)}^4 q_{(2,1)} \rangle_c - 15\langle q_{(1,1)}^4 q_{(2,2)} \rangle_c + 20\langle q_{(1,1)}^3 q_{(3,1)} \rangle_c - 60\langle q_{(1,1)}^3 q_{(3,2)} \rangle_c + 40\langle q_{(1,1)}^3 q_{(3,3)} \rangle_c - 90\langle q_{(1,1)}^2 q_{(2,2)} q_{(2,1)} \rangle_c \\
& + 45\langle q_{(1,1)}^2 q_{(2,1)}^2 \rangle_c + 45\langle q_{(1,1)}^2 q_{(2,2)}^2 \rangle_c + 15\langle q_{(2,1)}^3 \rangle_c - 15\langle q_{(2,2)}^3 \rangle_c + 15\langle q_{(1,1)}^2 q_{(4,1)} \rangle_c - 105\langle q_{(1,1)}^2 q_{(4,2)} \rangle_c + 180\langle q_{(1,1)}^2 q_{(4,3)} \rangle_c \\
& - 90\langle q_{(1,1)}^2 q_{(4,4)} \rangle_c - 45\langle q_{(2,1)}^2 q_{(2,2)} \rangle_c + 45\langle q_{(2,2)}^2 q_{(2,1)} \rangle_c + 60\langle q_{(1,1)} q_{(2,1)} q_{(3,1)} \rangle_c - 180\langle q_{(1,1)} q_{(2,1)} q_{(3,2)} \rangle_c \\
& + 120\langle q_{(1,1)} q_{(2,1)} q_{(3,3)} \rangle_c - 60\langle q_{(1,1)} q_{(2,2)} q_{(3,1)} \rangle_c + 180\langle q_{(1,1)} q_{(2,2)} q_{(3,2)} \rangle_c - 120\langle q_{(1,1)} q_{(2,2)} q_{(3,3)} \rangle_c + 6\langle q_{(1,1)} q_{(5,1)} \rangle_c \\
& - 90\langle q_{(1,1)} q_{(5,2)} \rangle_c + 300\langle q_{(1,1)} q_{(5,3)} \rangle_c - 360\langle q_{(1,1)} q_{(5,4)} \rangle_c + 144\langle q_{(1,1)} q_{(5,5)} \rangle_c + 15\langle q_{(2,1)} q_{(4,1)} \rangle_c - 105\langle q_{(2,1)} q_{(4,2)} \rangle_c \\
& + 180\langle q_{(2,1)} q_{(4,3)} \rangle_c - 90\langle q_{(2,1)} q_{(4,4)} \rangle_c - 15\langle q_{(2,2)} q_{(4,1)} \rangle_c + 105\langle q_{(2,2)} q_{(4,2)} \rangle_c - 180\langle q_{(2,2)} q_{(4,3)} \rangle_c + 90\langle q_{(2,2)} q_{(4,4)} \rangle_c \\
& + 10\langle q_{(3,1)}^2 \rangle_c - 60\langle q_{(3,1)} q_{(3,2)} \rangle_c + 40\langle q_{(3,1)} q_{(3,3)} \rangle_c + 90\langle q_{(3,2)}^2 \rangle_c - 120\langle q_{(3,2)} q_{(3,3)} \rangle_c + 40\langle q_{(3,3)}^2 \rangle_c + \langle q_{(6,1)} \rangle_c - 31\langle q_{(6,2)} \rangle_c \\
& + 180\langle q_{(6,3)} \rangle_c - 390\langle q_{(6,4)} \rangle_c + 360\langle q_{(6,5)} \rangle_c - 120\langle q_{(6,6)} \rangle_c, \tag{67}
\end{aligned}$$

where we used

$$q_{(r,s)} = q_{(a^r/p^s)} = \sum_{i=1}^M (a_i^r/p_i^s) n_i. \tag{68}$$

In the Appendix A, we show a specific example of these results for the net-particle number with $M = 2$.

In Eqs. (62)–(67), the cumulants of $P(N)$ are expressed in terms of the (mixed) cumulants of $\tilde{P}(\mathbf{n})$. These formulas thus can be used for the efficiency correction. We note that the number of cumulants does not depend on the number of efficiency bins M . This property is contrasted to the method proposed in Refs. [14,15], in which the number of expectation values to be calculated increases as $\sim M^m$ for an m th order cumulant. The numerical cost for the efficiency correction with Eqs. (62)–(67) thus is drastically reduced compared to the formulas proposed in Refs. [14,15] for large M . In the formulas proposed in Ref. [16], the number of terms is much more reduced compared to Eqs. (62)–(67) and thus the numerical cost is smaller than in our method. However, the derivation in Ref. [16] is complicated and it is quite difficult to extend the analysis in Ref. [16] to sixth and much higher orders. We have numerically verified that our method gives completely the same result as those in Refs. [14,15] and Ref. [16]. In actual analyses, it would be convenient to implement the derivation of Eqs. (62)–(67) as a numerical algorithm.

C. Mixed cumulants

So far, we considered the efficiency correction of the cumulants of a single charge $Q_{(a)}$. Exactly the same discussion can be applied to the efficiency correction of mixed cumulants, which probe correlations between different conserved quantities, e.g. net-baryon, net-strangeness, and net-charge numbers. Below we show the formulas for mixed cumulants $\langle Q_{(x)}^{m_1} Q_{(y)}^{m_2} \rangle_c$ up to fourth order ($m_1 + m_2 \leq 4$):

$$\langle Q_{(x)} Q_{(y)} \rangle_c = \langle q_{(1,0,1)} q_{(0,1,1)} \rangle_c + \langle q_{(1,1,1)} \rangle_c - \langle q_{(1,1,2)} \rangle_c, \tag{69}$$

$$\begin{aligned}
\langle Q_{(x)}^2 Q_{(y)} \rangle_c = & \langle q_{(1,0,1)}^2 q_{(0,1,1)} \rangle_c + 2\langle q_{(1,0,1)} q_{(1,1,1)} \rangle_c - 2\langle q_{(1,0,1)} q_{(1,1,2)} \rangle_c + \langle q_{(0,1,1)} q_{(2,0,1)} \rangle_c - \langle q_{(0,1,1)} q_{(2,0,2)} \rangle_c \\
& + \langle q_{(2,1,1)} \rangle_c - 3\langle q_{(2,1,2)} \rangle_c + \langle q_{(2,1,3)} \rangle_c, \tag{70}
\end{aligned}$$

$$\begin{aligned}
\langle Q_{(x)}^2 Q_{(y)}^2 \rangle_c = & \langle q_{(1,0,1)}^2 q_{(0,1,1)}^2 \rangle_c + \langle q_{(1,0,1)}^2 q_{(0,2,1)} \rangle_c - \langle q_{(1,0,1)}^2 q_{(0,2,2)} \rangle_c + \langle q_{(0,1,1)}^2 q_{(2,0,1)} \rangle_c - \langle q_{(0,1,1)}^2 q_{(2,0,2)} \rangle_c + 4\langle q_{(1,0,1)} q_{(0,1,1)} q_{(1,1,1)} \rangle_c \\
& - 4\langle q_{(1,0,1)} q_{(0,1,1)} q_{(1,1,2)} \rangle_c + 2\langle q_{(1,0,1)} q_{(1,2,1)} \rangle_c - 6\langle q_{(1,0,1)} q_{(1,2,2)} \rangle_c + 4\langle q_{(1,0,1)} q_{(1,2,3)} \rangle_c + 2\langle q_{(0,1,1)} q_{(2,1,1)} \rangle_c \\
& - 6\langle q_{(0,1,1)} q_{(2,1,2)} \rangle_c + 4\langle q_{(0,1,1)} q_{(2,1,3)} \rangle_c - 4\langle q_{(1,1,1)} q_{(1,1,2)} \rangle_c + 2\langle q_{(1,1,1)}^2 \rangle_c + 2\langle q_{(1,1,2)}^2 \rangle_c + \langle q_{(2,0,1)} q_{(0,2,1)} \rangle_c \\
& - \langle q_{(2,0,1)} q_{(0,2,2)} \rangle_c - \langle q_{(2,0,2)} q_{(0,2,1)} \rangle_c + \langle q_{(2,0,2)} q_{(0,2,2)} \rangle_c + \langle q_{(2,2,1)} \rangle_c - 7\langle q_{(2,2,2)} \rangle_c + 12\langle q_{(2,2,3)} \rangle_c - 6\langle q_{(2,2,4)} \rangle_c, \tag{71}
\end{aligned}$$

$$\begin{aligned}
\langle Q_{(x)}^3 Q_{(y)} \rangle_c = & \langle q_{(1,0,1)}^3 q_{(0,1,1)} \rangle_c + 3\langle q_{(1,0,1)}^2 q_{(1,1,1)} \rangle_c - 3\langle q_{(1,0,1)}^2 q_{(1,1,2)} \rangle_c + 3\langle q_{(2,0,1)} q_{(1,0,1)} q_{(0,1,1)} \rangle_c - 3\langle q_{(2,0,2)} q_{(1,0,1)} q_{(0,1,1)} \rangle_c \\
& + 3\langle q_{(1,0,1)} q_{(2,1,1)} \rangle_c - 9\langle q_{(1,0,1)} q_{(2,1,2)} \rangle_c + 6\langle q_{(1,0,1)} q_{(2,1,3)} \rangle_c + 3\langle q_{(2,0,1)} q_{(1,1,1)} \rangle_c - 3\langle q_{(2,0,1)} q_{(1,1,2)} \rangle_c \\
& - 3\langle q_{(2,0,2)} q_{(1,1,1)} \rangle_c + 3\langle q_{(2,0,2)} q_{(1,1,2)} \rangle_c + \langle q_{(3,0,1)} q_{(0,1,1)} \rangle_c - 3\langle q_{(3,0,2)} q_{(0,1,1)} \rangle_c + 2\langle q_{(3,0,3)} q_{(0,1,1)} \rangle_c + \langle q_{(3,1,1)} \rangle_c \\
& - 7\langle q_{(3,1,2)} \rangle_c + 12\langle q_{(3,1,3)} \rangle_c - 6\langle q_{(3,1,4)} \rangle_c, \tag{72}
\end{aligned}$$

where we used the symbol

$$q_{(r,s,t)} = q_{(x^r y^s / p^t)} = \sum_{i=1}^M (x_i^r y_i^s / p_i^t) n_i. \tag{73}$$

TABLE I. Comparison of CPU time to calculate the sixth order cumulant between the conventional and new methods.

M	Factorial moment	New method
4	64.7 s	30.8 s
8	17.3×10^2 s	31.3 s
12	14.1×10^3 s	32.3 s
200		62.7 s

D. Calculation cost

Finally we discuss how efficient Eqs. (62)–(67) are compared to the conventional method based on factorial moments [14]. In the conventional method, a cumulant is decomposed into mixed factorial moments. In this decomposition for an m th order cumulant, all possible combinations of mixed factorial moments between different efficiency bins with order r satisfying $r \leq m$ appear. The number of the r th order factorial moments is given by ${}_{r+M-1}C_r$ with M being different efficiency bins. The total number of the mixed factorial moments satisfying $r \leq m$ is thus given by

$$N_m^{\text{fm}} = \sum_{r=1}^m {}_{r+M-1}C_r = {}_{m+M}C_m - 1. \quad (74)$$

When this method is adopted to the analysis of the net-particle number, the numbers of efficiency bins of particle and antiparticle are given by $M/2$.

Assuming that the numerical cost to calculate one mixed factorial moment is insensitive to the order, the cost in the conventional method is proportional to Eq. (74), which grows as $\sim M^m$ for large M . On the other hand, in the new method with Eqs. (62)–(67), the total number of cumulants,

$$\begin{aligned} N_1^c &= 1, & N_2^c &= 3, & N_3^c &= 6, \\ N_4^c &= 13, & N_5^c &= 24, & N_6^c &= 48, \end{aligned} \quad (75)$$

is independent of M . By comparing Eqs. (74) and (75), it is clear that the new method becomes more advantageous for larger M , especially for higher order cumulants. In the actual numerical analyses, the cost to calculate one factorial moment or cumulant can grow with increasing M depending on the implementation and data structure. As this M dependence can be common for both methods, we neglect this effect here. In Fig. 1 we show the number of terms in both methods, Eqs. (74) and (75), as functions of M for fourth and sixth orders.

In Table I, we show the comparison of the CPU time to calculate the sixth order cumulant in both methods in specific implementations. The codes are executed on the same CPU (3 GHz Intel Core i7) for 1×10^5 events and $M = 4, 8, 12$, and 200. One finds that the CPU time with the conventional method rapidly increases as M becomes larger, while CPU time in the new method is insensitive to M ; even with $M = 200$, the CPU time is only about twice larger than $M = 4$ in the new method. This result is consistent with the cost estimate in Fig. 1. Moreover, the new method is about two orders faster than the conventional one already at $M = 8$. We, however,

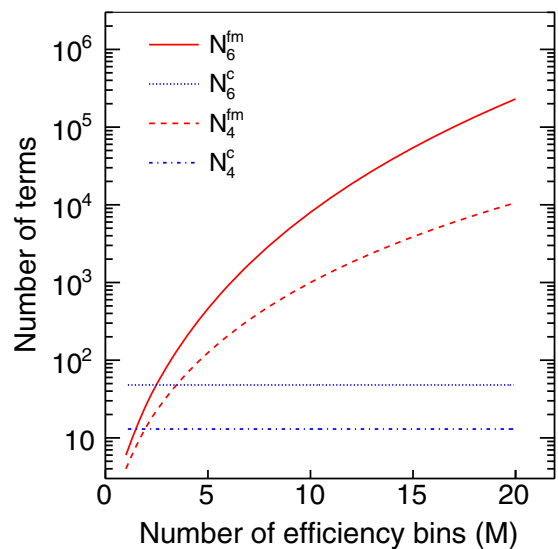


FIG. 1. Number of terms (factorial moments or cumulants) to be calculated in the conventional method based on the factorial moments [14] and the new method in this paper, Eqs. (74) and (75), respectively, for the efficiency correction of the fourth and sixth order cumulants as a function of efficiency bins.

note that the calculation cost, of course, is strongly dependent on the implementation.

It is also notable that the new method is advantageous in simplifying the code and reducing memory requirements.

IV. TWO-DISTRIBUTION MODEL

In the rest of this paper we focus on the effect of using the averaged efficiency for different efficiency bins. In this section, we first consider a simple problem which can be treated analytically.

We consider a measurement of two kinds of particle number distributions $P(N_A)$ and $P(N_B)$ by detectors having different efficiencies ε_A and ε_B , respectively. We assume that the two distributions are equivalent and independent, and their cumulants are given by

$$\langle N_A^m \rangle_c = \langle N_B^m \rangle_c = C_m. \quad (76)$$

We are interested in the cumulants of the total particle number $N = N_A + N_B$. Due to the additive property of cumulants for independent stochastic variables [7], cumulants of N are given by

$$K_m \equiv \langle N^m \rangle_c = \langle N_A^m \rangle_c + \langle N_B^m \rangle_c = 2C_m. \quad (77)$$

Because of the efficiency loss, the observed particle numbers n_A and n_B have different distributions from those of N_A and N_B . The cumulants of n_A and n_B are represented by C_m by the inverse procedure of Eqs. (26) and (27) [7]. For the first and second orders we have

$$\langle n_X \rangle = \varepsilon_X C_1, \quad (78)$$

$$\langle n_X^2 \rangle_c = \varepsilon_X^2 C_2 + \varepsilon_X(1 - \varepsilon_X)C_1, \quad (79)$$

with $X = A$ and B . By substituting Eqs. (78) and (79) into Eqs. (62) and (63) with $M = 2$, the correct value of K_m is recovered.

Now, we consider a case such that the efficiency correction is performed by regarding $n = n_A + n_B$ as a particle number described by a single distribution function measured by an averaged efficiency $\bar{\varepsilon} = (\varepsilon_A + \varepsilon_B)/2$. Then, the efficiency correction would be performed by substituting $n = n_A + n_B$ and $p = \bar{\varepsilon}$ into the result in Sec. II such as Eqs. (26) and (27). For the first order, the result of this efficiency correction is

$$\begin{aligned} K_1^{(\text{ave})} &= \langle N_1 \rangle + \langle N_2 \rangle = \frac{\langle n_1 \rangle}{\bar{\varepsilon}} + \frac{\langle n_2 \rangle}{\bar{\varepsilon}} \\ &= \frac{\varepsilon_1 C_1}{\bar{\varepsilon}} + \frac{\varepsilon_2 C_1}{\bar{\varepsilon}} = 2C_1. \end{aligned} \quad (80)$$

Therefore, the correct cumulant Eq. (77) is recovered to this order. This, however, is not the case for higher order cumulants. By denoting the deviation of the reconstructed cumulant with average efficiency $K_m^{(\text{ave})}$ from the original one as

$$\Delta K_m = K_m - K_m^{(\text{ave})} = 2C_m - K_m^{(\text{ave})}, \quad (81)$$

ΔK_m is calculated to be

$$\Delta K_2 = \frac{1}{2} \left(\frac{\Delta \varepsilon}{\bar{\varepsilon}} \right)^2 (C_2 - C_1), \quad (82)$$

$$\Delta K_3 = \frac{3}{2} \left(\frac{\Delta \varepsilon}{\bar{\varepsilon}} \right)^2 (C_3 - 2C_2 + C_1), \quad (83)$$

$$\begin{aligned} \Delta K_4 &= \frac{1}{2} \left(\frac{\Delta \varepsilon}{\bar{\varepsilon}} \right)^2 (6C_4 - 18C_3 + 19C_2 - 7C_1) \\ &\quad + \frac{1}{8} \left(\frac{\Delta \varepsilon}{\bar{\varepsilon}} \right)^4 (C_4 - 6C_3 + 11C_2 - 6C_1), \end{aligned} \quad (84)$$

with $\Delta \varepsilon = \varepsilon_A - \varepsilon_B$. The nonzero ΔK_m shows that the reconstructed cumulant does not agree with the original one. These results clearly show that the use of the averaged efficiency gives rise to a deviation in the result of the efficiency correction. Only for Poisson distribution ($C_1 = C_2 = \dots = C_m$) does the deviation vanish.

Let us see ΔK_m in specific distributions. We first consider a Gauss distributions with $C_1 = 20$, $C_2 = 25$, and $C_m = 0$ for $m \geq 3$. In the top panel of Fig. 2, ΔK_m with $m = 2, 3, 4$ are plotted as functions of $\Delta \varepsilon$. One finds that ΔK_m becomes large with increasing $\Delta \varepsilon$ and m . Next, we consider a distribution with $C_1 = 20$ and $C_m = 19$ for $m \geq 2$; this distribution is close to Poissonian but cumulants higher than the first order are 5% smaller than Poissonian values. The $\Delta \varepsilon$ dependencies of ΔK_m in this case is shown in the bottom panel of Fig. 2. From the figure, one again obtains the same conclusion that ΔK_m becomes large for higher orders and larger $\Delta \varepsilon$. These results show the importance of the use of the separated efficiencies in the experimental analysis especially for higher orders.

V. NUMERICAL ANALYSIS IN TOY MODELS

In this section, we study the effects of using averaged efficiency numerically in toy models by generating random events.

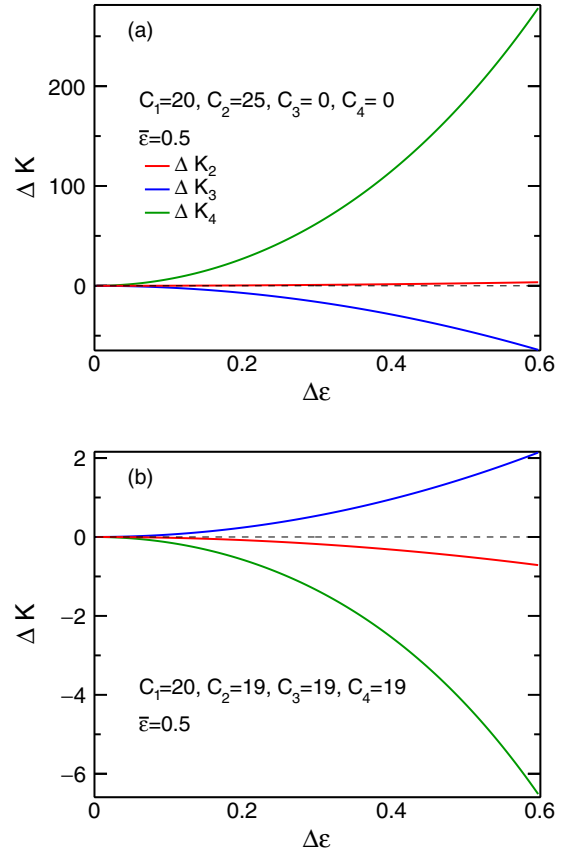


FIG. 2. Deviation of the efficiency corrected values of cumulants using averaged efficiency ΔK_m assuming (a) Gauss distribution and (b) distribution that has 5% smaller cumulants than Poisson distribution.

A. Two-distribution model

First, we analyze the two-distribution problem discussed in the previous section numerically. Two particle numbers N_A and N_B are independently generated according to Gauss distribution, and they are randomly sampled with the efficiencies ε_A and ε_B to obtain the measured particle numbers n_A and n_B . We generated 10^8 events, and this analysis was repeated 30 times independently for the estimate of the statistical error. We perform the efficiency correction by the following two methods:

- (1) Efficiency correction with separated efficiencies for A and B.
- (2) Efficiency correction using the averaged efficiency $\bar{\varepsilon} = (\varepsilon_A + \varepsilon_B)/2$.

We set $C_1 = 20$, $C_2 = 25$, $\varepsilon_A = 0.3$, and $\varepsilon_B = 0.7$.

The results of ΔK_m in these analyses are shown in Fig. 3 as a function of the order of the cumulant m . Blue circles are the results with separated efficiency correction. The figure shows that they reproduce correct input cumulants with $\Delta K_m = 0$ within statistical error. Red squares represent results from the averaged efficiency. The figure shows that these results give wrong values with $\Delta K_m \neq 0$ for $m \geq 2$. These deviations are

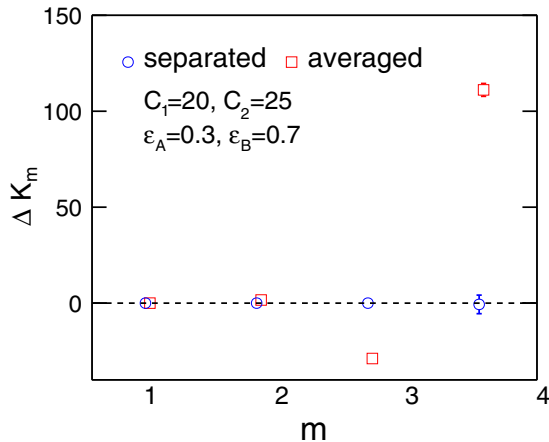


FIG. 3. Deviation of the efficiency corrected m th order cumulant from input value ΔK_m up to fourth order. Blue circles represent the results from separated efficiencies, and red squares are results with the averaged efficiency. Parameters for the distribution are $C_1 = 20$ and $C_2 = 25$. Efficiencies are $\varepsilon_A = 0.3$ and $\varepsilon_B = 0.7$. The result can be directly compared with analytical calculation in Sec. IV.

compared with the analytic results in Eqs. (82)–(84) in Table II. The table shows that they are consistent with each other.

B. Averaged efficiencies for different particle species

In Secs. IV and VA, we discussed the case with a single particle species with a unit charge. Next we extend the discussion to the case of the net-charge fluctuation. In this case, we measure the charged particles without particle identifications, and there seems to be no problem using averaged efficiency of charged particles for the correction. However, when we consider the fact that the charged particles mainly consist of π^\pm , K^\pm and p^\pm , this assumption would be violated, because those particles have different efficiencies experimentally and their net-particle distributions could have different probability distributions. Therefore, we perform a toy model analysis in order to study the effect of using the averaged efficiency assuming the net-charge distribution. At high beam energies, one can expect that the produced pion distribution is closer to Gaussian than the distributions of kaons and protons due to the large production of pions. In this toy model, therefore, we simply set the distribution for π^\pm as a Gauss distribution as an extreme case, while for K^\pm and p^\pm we set them as Poisson distributions. These particles are observed with different efficiencies for different particle species. These different efficiencies are used in the analysis of

TABLE II. Comparison of ΔK_m between the numerical and the analytical calculations in the two-distribution model.

m	Separated	Averaged	Analytical
1	$0 \pm 0.7 \times 10^{-3}$	$0 \pm 0.6 \times 10^{-3}$	0
2	0 ± 0.01	1.63 ± 0.01	1.60
3	0 ± 0.3	-28.9 ± 0.2	-28.8
4	-0.7 ± 4.8	111.1 ± 3.4	115.1

TABLE III. Parameters used in the toy model discussed in Sec. VB.

Particles	$P(N)$	Charge	Mean	Sigma	Efficiency
π^+	Gauss	+1	30	8	0.3
K^+	Poisson	+1	10		0.6
p	Poisson	+1	8		0.9
π^-	Gauss	-1	25	7	0.25
K^-	Poisson	-1	4		0.55
\bar{p}	Poisson	-1	3		0.85

separated efficiency correction. We also perform the efficiency correction with the averaged efficiencies for positively and negatively charged particles:

$$\varepsilon_{\pi K p}^\pm = \frac{\sum_i \varepsilon_i^\pm N_i^\pm}{\sum_i N_i^\pm}, \quad (85)$$

where i denotes particle species (π , K , p) and N is the number of produced particles. Note that the use of the averaged efficiency for positively and negatively charged particles can give rise to other artificial effects discussed in Ref. [18]. Parameters are shown in Table III.

Relative deviations of efficiency corrected m th order cumulant from input value $\Delta K_m/K_m$ are shown in Fig. 4 up to the fourth order. The figure shows that the result with the averaged efficiencies again cannot reproduce the correct value. Thus, we must not use averaged efficiency if there are different physics in different efficiency bins.

C. Two detectors with a common source

In current analysis for net-proton distribution at STAR, an efficiency bin is divided into two p_T regions, $0.4 < p_T < 0.8$ and $0.8 < p_T < 2.0$ GeV/c [20], because the measurements of particles are performed in different ways for these p_T regions: Energy loss measured by the time projection chamber (TPC) is used for proton identification at $0.4 < p_T < 0.8$ GeV/c, while the mass squared measured by the time-of-flight (TOF) detector is also used at $0.8 < p_T < 2.0$ GeV/c. By including the

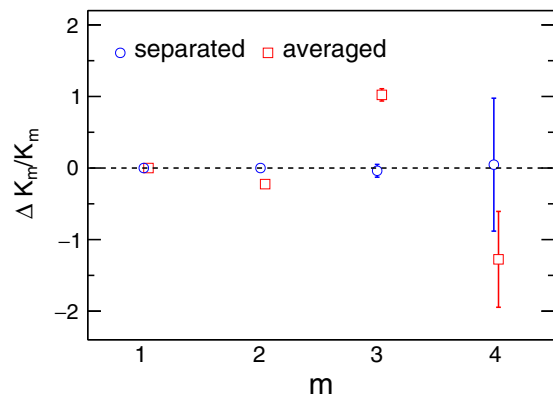


FIG. 4. Relative deviation of efficiency corrected m th order cumulant from input value ($\Delta K_m/K_m$). Blue circles and red squares represent the results from separated and averaged efficiencies, respectively.

TABLE IV. Parameters used in the toy model discussed in Sec. V C.

$P(N)$	Charge	Mean	Sigma	Efficiency
Gauss	+1	20	$\sqrt{32}$	$\varepsilon_A^+ = 0.9, \varepsilon_B^+ = 0.3$
Gauss	-1	8	$\sqrt{8}$	$\varepsilon_A^- = 0.4, \varepsilon_B^- = 0.8$

TOF detector, the efficiency drops at $0.8 < p_T < 2.0$ GeV/ c . This p_T dependent efficiency is implemented by dividing the p_T region at 0.8 GeV/ c . Similarly, efficiencies would depend on ϕ direction. The TPC and TOF detectors cover full azimuthal angle and have excellent particle identification capability. However, some of the TPC sectors are sometimes in a bad condition, which leads to the nonuniform acceptance in the ϕ direction. Let us discuss the effect by using the averaged efficiency in these conditions assuming two detectors, which may not be the case discussed in Sec. V B, because the distribution at each detector would not be determined separately. In other words, even if there are different kinds of particle distributions, we cannot identify those distributions at the detector level.

Setup for the toy model is as follows. Particles are randomly generated according to Gauss distributions $P(N)$, and we let those particles randomly incident on the detector A or B with 50% probability. Then particles are randomly sampled by efficiencies ε_A^\pm and ε_B^\pm . We apply efficiency correction on $P(N_A)$ and $P(N_B)$ with separated efficiencies or with averaged efficiency between two detectors. We consider the net particle number by generating charge ± 1 particles assuming the measurement of net-proton number cumulants. Parameters are shown in Table IV.

The last row in Table IV represents efficiencies that are characterized for each detector and electric charge. Results of ΔK_m are shown in Fig. 5. From the figure, one finds that there is no deviation for all the order of cumulants. Note that the value of the denominator K_m is not common for different m . This leads to the larger error for third order than fourth order in Fig. 5. At first glance this looks strange, but we can provide a simple explanation as follows. When one focuses on a particle in this model, it is measured with a probability $(\varepsilon_A^\pm + \varepsilon_B^\pm)/2$ randomly and independently. Therefore, this is

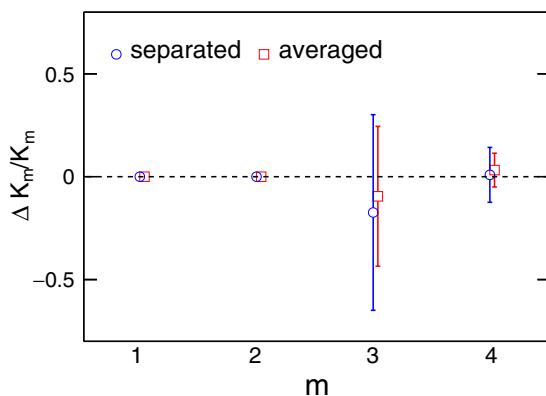


FIG. 5. Relative deviation of efficiency corrected m th order cumulant from input value $\Delta K_m/K_m$.

exactly the case of a single efficiency bin with the averaged efficiency. This result indicates that the efficiency correction with averaged efficiency works well when the underlying physics is identical for different efficiency bins. However, for nonuniform acceptance in a real experiment, one needs to check whether the results obtained from averaged efficiencies are consistent with the separated efficiencies.

VI. SUMMARY

In this paper, we derived formulas for the efficiency correction with many efficiency bins. In our method, the formulas are obtained easily compared to Ref. [16], but the numerical cost is drastically reduced compared to Refs. [14,15] when the number of efficiency bins and order of the cumulant are large. The efficiency correction for higher order cumulants with many bins thus can be carried out effectively in our method. The result is then applied to the efficiency correction in simple models to study the effect of using averaged efficiency in Secs. IV and V. We have shown that the use of the averaged efficiency can lead to wrong corrected values if underlying physics is different in efficiency bins. This result indicates that separated efficiencies have to be used to perform the efficiency correction correctly. For example, it would be important to take account of the nonuniform acceptance along azimuthal angle and the p_T dependencies of efficiency for the accurate efficiency correction.

Final remarks are in order. First, although we used the binomial model throughout this paper, this model is justified only when the efficiencies for individual particles are independent [7]. When the correlations between individual particles are not negligible, these effects have to be considered [17]. Second, experimental analyses usually measure proton number cumulants as proxies of baryon number cumulants. In Refs. [12,21], it is shown that the measurement of protons corresponds to the measurement of baryons with 50% efficiency loss. Therefore, the baryon number cumulants can in principle be constructed from those of protons using efficiency correction. In this case, the use of the binomial model is justified owing to isospin randomization [21].

ACKNOWLEDGMENTS

The authors thank X. Luo and N. Xu for useful discussions. T.N. thanks P. Tribedy for the idea of the toy model discussed in Sec. V B. M.K. is thankful for stimulating discussions in the INT program “Exploring the QCD Phase Diagram through Energy Scans,” Seattle, Sep. 19 – Oct. 14, especially with A. Bzdak and V. Koch. We acknowledge support from MEXT and JSPS KAKENHI Grant No. 25105504 and the Super Global University Program in University of Tsukuba.

APPENDIX: NET-PARTICLE NUMBERS IN A SIMPLE CASE

In the case of net-particle numbers with a single efficiency bin, explicit formulas for efficiency correction can be derived from Eqs. (62)–(68) by substituting $M = 2, a_1 = 1, a_2 = -1,$

and $p_1 = p_2 = p$. By defining $n_{\text{net}} = n_1 - n_2$ and $n_{\text{tot}} = n_1 + n_2$, the formulas up to sixth order are given by

$$\langle Q \rangle_c = \frac{1}{p} \langle n_{\text{net}} \rangle_c, \quad (\text{A1})$$

$$\langle Q^2 \rangle_c = \frac{1}{p^2} \langle n_{\text{net}}^2 \rangle_c + \left(-\frac{1}{p^2} + \frac{1}{p} \right) \langle n_{\text{tot}} \rangle_c, \quad (\text{A2})$$

$$\langle Q^3 \rangle_c = \frac{1}{p^3} \langle n_{\text{net}}^3 \rangle_c + \left(-\frac{3}{p^3} + \frac{3}{p^2} \right) \langle n_{\text{net}} n_{\text{tot}} \rangle_c + \left(\frac{2}{p^3} - \frac{3}{p^2} + \frac{1}{p} \right) \langle n_{\text{net}} \rangle_c, \quad (\text{A3})$$

$$\begin{aligned} \langle Q^4 \rangle_c &= \frac{1}{p^4} \langle n_{\text{net}}^4 \rangle_c + \left(-\frac{6}{p^4} + \frac{6}{p^3} \right) \langle n_{\text{net}}^2 n_{\text{tot}} \rangle_c + \left(\frac{8}{p^4} - \frac{12}{p^3} + \frac{4}{p^2} \right) \langle n_{\text{net}}^2 \rangle_c + \left(\frac{3}{p^4} - \frac{6}{p^3} + \frac{3}{p^2} \right) \langle n_{\text{tot}}^2 \rangle_c \\ &+ \left(-\frac{6}{p^4} + \frac{12}{p^3} - \frac{7}{p^2} + \frac{1}{p} \right) \langle n_{\text{tot}} \rangle_c, \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} \langle Q^5 \rangle_c &= \frac{1}{p^5} \langle n_{\text{net}}^5 \rangle_c + \left(-\frac{10}{p^5} + \frac{10}{p^4} \right) \langle n_{\text{net}}^3 n_{\text{tot}} \rangle_c + \left(\frac{20}{p^5} - \frac{30}{p^4} + \frac{10}{p^3} \right) \langle n_{\text{net}}^3 \rangle_c + \left(\frac{15}{p^5} - \frac{30}{p^4} + \frac{15}{p^3} \right) \langle n_{\text{net}} n_{\text{tot}}^2 \rangle_c \\ &+ \left(-\frac{50}{p^5} + \frac{110}{p^4} - \frac{75}{p^3} + \frac{15}{p^2} \right) \langle n_{\text{net}} n_{\text{tot}} \rangle_c + \left(\frac{24}{p^5} - \frac{60}{p^4} + \frac{50}{p^3} - \frac{15}{p^2} + \frac{1}{p} \right) \langle n_{\text{net}} \rangle_c, \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} \langle Q^6 \rangle_c &= \frac{1}{p^6} \langle n_{\text{net}}^6 \rangle_c + \left(-\frac{15}{p^6} + \frac{15}{p^5} \right) \langle n_{\text{net}}^4 n_{\text{tot}} \rangle_c + \left(\frac{40}{p^6} - \frac{60}{p^5} + \frac{20}{p^4} \right) \langle n_{\text{net}}^4 \rangle_c \\ &+ \left(\frac{45}{p^6} - \frac{90}{p^5} + \frac{45}{p^4} \right) \langle n_{\text{net}}^2 n_{\text{tot}}^2 \rangle_c + \left(-\frac{15}{p^6} + \frac{45}{p^5} - \frac{45}{p^4} + \frac{15}{p^3} \right) \langle n_{\text{tot}}^3 \rangle_c + \left(-\frac{210}{p^6} + \frac{480}{p^5} - \frac{345}{p^4} + \frac{75}{p^3} \right) \langle n_{\text{net}}^2 n_{\text{tot}} \rangle_c \\ &+ \left(\frac{184}{p^6} - \frac{480}{p^5} + \frac{430}{p^4} - \frac{150}{p^3} + \frac{16}{p^2} \right) \langle n_{\text{net}}^2 \rangle_c + \left(\frac{90}{p^6} - \frac{270}{p^5} + \frac{285}{p^4} - \frac{120}{p^3} + \frac{15}{p^2} \right) \langle n_{\text{tot}}^2 \rangle_c \\ &+ \left(-\frac{120}{p^6} + \frac{360}{p^5} - \frac{390}{p^4} + \frac{180}{p^3} - \frac{31}{p^2} + \frac{1}{p} \right) \langle n_{\text{tot}} \rangle_c. \end{aligned} \quad (\text{A6})$$

-
- [1] M. Asakawa, U. W. Heinz, and B. Muller, *Phys. Rev. Lett.* **85**, 2072 (2000).
[2] S. Jeon and V. Koch, *Phys. Rev. Lett.* **85**, 2076 (2000).
[3] S. Ejiri, F. Karsch, and K. Redlich, *Phys. Lett. B* **633**, 275 (2006).
[4] M. A. Stephanov, *Phys. Rev. Lett.* **102**, 032301 (2009).
[5] M. Asakawa, S. Ejiri, and M. Kitazawa, *Phys. Rev. Lett.* **103**, 262301 (2009).
[6] B. Friman, F. Karsch, K. Redlich, and V. Skokov, *Eur. Phys. J. C* **71**, 1694 (2011).
[7] M. Asakawa and M. Kitazawa, *Prog. Part. Nucl. Phys.* **90**, 299 (2016).
[8] X. Luo and N. Xu, [arXiv:1701.02105](https://arxiv.org/abs/1701.02105).
[9] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **112**, 032302 (2014).
[10] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **113**, 092301 (2014).
[11] L. Chen (STAR Collaboration), *Nucl. Phys. A* **904-905**, 471c (2013).
[12] M. Kitazawa and M. Asakawa, *Phys. Rev. C* **86**, 024904 (2012); **86**, 069902(E) (2012).
[13] A. Bzdak and V. Koch, *Phys. Rev. C* **86**, 044904 (2012).
[14] A. Bzdak and V. Koch, *Phys. Rev. C* **91**, 027901 (2015).
[15] X. Luo, *Phys. Rev. C* **91**, 034907 (2015).
[16] M. Kitazawa, *Phys. Rev. C* **93**, 044911 (2016).
[17] A. Bzdak, R. Holzmann, and V. Koch, *Phys. Rev. C* **94**, 064907 (2016).
[18] T. Nonaka, T. Sugiura, S. I. Esumi, H. Masui, and X. Luo, *Phys. Rev. C* **94**, 034909 (2016).
[19] M. Kitazawa and X. Luo, [arXiv:1704.04909](https://arxiv.org/abs/1704.04909).
[20] X. Luo, PoS (CPOD2014) (2014) 019.
[21] M. Kitazawa and M. Asakawa, *Phys. Rev. C* **85**, 021901 (2012).