Effects of the detection efficiency on multiplicity distributions

A. H. Tang[1] and G. Wang[2]
[1]*Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*University of California, Los Angeles, California 90095, USA*
(Received 11 May 2013; published 12 August 2013)

In this paper we investigate how a finite detection efficiency affects three popular multiplicity distributions, namely, the Poisson, the binomial, and the negative binomial distributions. We found that a multiplicity-independent detection efficiency does not change the characteristic of a distribution, while a multiplicity-dependent detection efficiency does. We layout a procedure to study the deviation of moments and their derivative quantities from the baseline distribution due to a multiplicity-dependent detection efficiency.

DOI: 10.1103/PhysRevC.88.024905 PACS number(s): 25.75.Ld

I. INTRODUCTION

One of the main purposes of relativistic heavy ion collision experiments is to explore the QCD phase boundary [1], in particular to look for signatures of a first order phase transition [2,3] and a critical end point [4,5]. Moments of the distributions of conserved quantities, such as net-baryon number, net-charge, and net-strangeness, have been argued to be sensitive to the phase transition and the critical end point and are drawing increased attention from both experimentalists [6–8] and theorists [9–12]. In the study of higher order moments and their derivative quantities, an abnormal deviation from the baseline distribution is usually interpreted as an interesting physics signal. In practice, such a deviation is complicated by experimental effects, such as a finite detection efficiency. In this paper, we address how a finite efficiency would change three widely used multiplicity distributions, namely, the Poisson, the binomial, and the negative binomial distributions. We will discuss the case of a multiplicity-independent efficiency, followed by the case of a multiplicity-dependent efficiency, where we layout a procedure to investigate how the efficiency affects the three multiplicity distributions. The procedure also applies to the difference distribution of two multiplicity distributions.

II. MULTIPLICITY DISTRIBUTIONS WITH A MULTIPLICITY-INDEPENDENT EFFICIENCY

A. Poisson distribution

The probability mass function for the Poisson distribution is given by

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{1}$$

where k is a nonnegative integer (the same for the other two distributions discussed below), and λ is both the mean and the variance of the distribution. The probability-generating function for the Poisson distribution is given by

$$G(z) = e^{-\lambda(1-z)}, \tag{2}$$

where z is a complex number with $|z| \leqslant 1$.

We treat observing and not-observing a particle as "decay" modes of a particle, and apply the cluster decay theorem [13] by replacing z with the generating function

$$g(y) = (1-\epsilon) + \epsilon y, \tag{3}$$

where ϵ is the probability of seeing a particle, in practice less than unity due to the finite acceptance and detection efficiency. Without losing generality, below we refer to ϵ as the detection efficiency inclusive of both sources.

Then Eq. (2) becomes

$$G(z) = G(g(y)) = e^{-\lambda(1-[(1-\epsilon)+\epsilon y])} = e^{-\lambda\epsilon(1-y)}. \tag{4}$$

One immediately identifies that the new generating function, for an experimental observable with a finite detection efficiency, still maintains the form of a Poisson distribution, with the mean of the distribution reduced to $\lambda\epsilon$.

Note that Eq. (3) is simply the generating function for a binomial process with $n = 1$ (see below). With Eq. (3) convoluted into Eq. (2), the fluctuation of event-by-event efficiency has been taken into account, similar to the procedure proposed in [14].

B. Binomial distribution

The probability mass function for the binomial distribution is given by

$$f(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}, \tag{5}$$

where $p \in [0,1]$, and the nonnegative integer $n \geqslant k$. The mean and the variance of the distribution are given by np and $np(1-p)$, respectively. The corresponding probability-generating function is given by

$$G(z) = (1-p+pz)^n. \tag{6}$$

Similarly with a finite detection efficiency,

$$\begin{aligned} G(z) = G(g(y)) &= [1-p+p(1-\epsilon+\epsilon y)]^n \\ &= [1-p\epsilon+p\epsilon y]^n. \end{aligned} \tag{7}$$

We have recovered the probability-generating function for the binomial distribution with the replacement of $p \to p'(= p\epsilon)$. The mean of the new distribution is given by $\mu' = \mu\epsilon$. The calculation of other quantities under the influence of a finite detection efficiency is thus straightforward. For example,

$$\kappa\sigma^2 = \frac{C_4}{C_2} = 1 - 6p + 6p^2, \quad (8)$$

where κ is the kurtosis and C_i is the ith order cumulant. When taking the detection efficiency into account, one simply replaces every p with $p\epsilon$,

$$\kappa\sigma^2 = \frac{C_4}{C_2} = 1 - 6p\epsilon + 6p^2\epsilon^2. \quad (9)$$

Such knowledge is useful for quantifying the deviation of the observable of interest from the original distribution due to the finite detection efficiency.

C. Negative binomial distribution

The probability mass function for the negative binomial distribution is given by

$$f(k; r, p) = \binom{k + r - 1}{k}(1 - p)^k p^r, \quad (10)$$

where $p \in [0, 1]$, and the real number $r > 0$. It has identities of $p = \frac{\mu}{\sigma^2}$ and $r = \frac{\mu p}{1-p}$, where μ and σ^2 are the mean and the variance, respectively. Its probability-generating function has the form of

$$G(z) = \left(\frac{\frac{r}{\mu}}{1 + \frac{r}{\mu} - z}\right)^r = \left(\frac{p}{1 - (1 - p)z}\right)^r, \quad (11)$$

where $p = \frac{\mu}{\sigma^2} = \frac{r}{\mu+r}$.

Likewise, in the case of a finite detection efficiency, we have

$$G(z) = G(g(y)) = \left(\frac{p}{1 - (1 - p)(1 - \epsilon + \epsilon y)}\right)^r = \left(\frac{p'}{1 - (1 - p')y}\right)^r, \quad (12)$$

where $p' = \frac{p}{\epsilon+p-p\epsilon}$, and r is unchanged. The form of the probability-generating function for the negative binomial distribution is recovered, with $p \to p'$ and $\mu \to \mu'(= \mu\epsilon)$. Again, other quantities with a finite detection efficiency can be evaluated with the two simple replacements. For example, replacing p with $\frac{p}{\epsilon+p-p\epsilon}$ everywhere in

$$\kappa\sigma^2 = \frac{C_4}{C_2} = \frac{6 - 6p + p^2}{p^2} \quad (13)$$

gives the $\kappa\sigma^2$ for the case with a finite detection efficiency.

III. MULTIPLICITY DISTRIBUTIONS WITH A MULTIPLICITY-DEPENDENT EFFICIENCY

Usually the detection efficiency decreases with increased multiplicity, as the reconstruction of a particle becomes more difficult in a crowded environment. In this case, the detection efficiency is expressed as a function of k, $\epsilon(k)$. Now for all the three distributions, the probability-generating function can no longer be written in a concise form. Instead, we take the general definition

$$G(y) = \sum_{k=0}^{\infty} f(k)z^k = \sum_{k=0}^{\infty} f(k)[1 - \epsilon(k) + \epsilon(k)y]^k. \quad (14)$$

Generally one cannot recover the generating function of the same type. That means, a multiplicity-dependent efficiency will distort the original distribution, unlike the case of a multiplicity-independent efficiency, where the detector effect will change the mean and width of the distribution, but keep the characteristic shape (as the same type). Nevertheless, with $\epsilon(k)$ as input, one can still calculate the mean μ' and the variance σ'^2:

$$\mu' = \langle M\rangle = F_1, \quad (15)$$

$$\sigma'^2 = \langle M^2\rangle - \langle M\rangle^2 = \langle M(M - 1)\rangle + \langle M\rangle - \langle M\rangle^2 = F_2 + F_1 - F_1^2, \quad (16)$$

where F_i is the factorial moment $\langle M(M - 1)\cdots(M - i + 1)\rangle$, given by $F_i \equiv \frac{\partial^i G(y)}{\partial y^i}|_{y=1}$.

For the Poisson distribution

$$F_i = e^{-\lambda}\sum_{k=i}^{\infty}\frac{\lambda^k}{(k - i)!}\epsilon(k)^i, \quad (17)$$

for the binomial distribution

$$F_i = \sum_{k=i}^{\infty}\frac{n!}{(k - i)!(n - k)!}p^k(1 - p)^{n-k}\epsilon(k)^i, \quad (18)$$

and for the negative binomial distribution

$$F_i = \sum_{k=i}^{\infty}\frac{(k + r - 1)!}{(k - i)!(r - 1)!}(1 - p)^k p^r \epsilon(k)^i. \quad (19)$$

With Eqs. (17), (18), and (19), F_i can be numerically calculated with known $\epsilon(k)$, and the calculation is no more complicated than that for the corresponding distributions with the perfect detection. Note that in practice one only needs to perform the summation over k to a value that is large enough, say, a few σ above the mean value, so that F_i has little change with further increase of k.[1]

The third and the fourth central moments are given by

$$\langle(M - \langle M\rangle)^3\rangle = F_1 + 2F_1^3 + 3F_2 - 3F_1(F_1 + F_2) + F_3, \quad (20)$$

and

$$\langle(M - \langle M\rangle)^4\rangle = F_1 - 3F_1^4 + 7F_2 + 6F_1^2(F_1 + F_2) + 6F_3 - 4F_1(F_1 + 3F_2 + F_3) + F_4. \quad (21)$$

With the mean, the variance, and the third and the fourth central moments, the first few cumulants can be calculated as usual:

$$C_1 = \langle(\delta M)\rangle = 0 \quad C_2 = \langle(\delta M)^2\rangle$$
$$C_3 = \langle(\delta M)^3\rangle \quad C_4 = \langle(\delta M)^4\rangle - 3\langle(\delta M)^2\rangle^2, \quad (22)$$

[1] To avoid the problem of the overflowing with large numerical numbers, one may also calculate, for example, $e^{-\lambda}\sum_{k=i}^{\infty}\frac{\lambda^k}{(k-i)!}\epsilon(k)^i$ as $\exp\{\ln[\sum_{k=i}^{\infty}\frac{\lambda^k}{(k-i)!}\epsilon(k)^i] - \lambda\}$.

where $\delta M = M - \langle M \rangle$. One can further calculate skewness and kurtosis based on cumulants, which is straightforward and thus is not repeated here.

Note that although we addressed three specific multiplicity distributions, the procedure discussed in this section can be extended to other multiplicity distributions, as long as the factorial moments can be conveniently calculated.

IV. DIFFERENCE DISTRIBUTION OF TWO MULTIPLICITY DISTRIBUTIONS

The difference between two independent variables is useful for studying the fluctuation of conserved quantities, e.g., the net-charge and the net-baryon numbers. The difference between two variables, each following the Poisson distribution, is called the Skellam distribution, and its probability-generating function is given by

$$G(z; \mu_1, \mu_2) = e^{-(\mu_1+\mu_2)+\mu_1 z+\mu_2/z}. \tag{23}$$

It follows from one of the properties of the probability-generating function: for the difference of two independent random variables $S = X_1 - X_2$, the generating function is given by $G_S(z) = G_{X_1}(z)G_{X_2}(z^{-1})$. The generating function for the difference between two binomial variables is

$$G(z; n_1, p_1, n_2, p_2) = (1 - p_1 + p_1 z)^{n_1}(1 - p_2 + p_2/z)^{n_2}, \tag{24}$$

and the generating function for the difference between two negative binomial variables is

$$G(z; r_1, p_1, r_2, p_2) = \left(\frac{p_1}{1-(1-p_1)z}\right)^{r_1}\left(\frac{p_2}{1-(1-p_2)/z}\right)^{r_2}. \tag{25}$$

When we take into account the finite detection efficiency, none of the three generating functions above can recover the form of the same type. Fortunately, they describe the *difference between two quantities*, to both of which the argument on the detection efficiency still applies. This facilitates the calculation of cumulants of the three difference-distributions with the finite detection efficiency under consideration. For example, for the net-charge distribution, the additivity of cumulants directly gives $C_{\Delta\text{charge}} = C_+ - C_-$, where C_+ and C_- are cumulants for positively and negatively charged particles, respectively. The $C_{\Delta\text{charge}}$ with a finite detection efficiency can be calculated this way as long as the distributions of separate charges are independent of each other. Here we assume that the two underlying distributions are completely independent of each other, to solely investigate how a nonphysics effect (finite detection efficiency) disturbs the baseline distribution, when studying cumulants of the difference of two variables. This treatment is different from that in [14] where the derivation starts from the cumulants of the difference distribution, with the correlation between the two variables already taken into account.

V. CONCLUSION

We have shown that for the Poisson, the binomial, and the negative binomial distributions, a multiplicity-independent efficiency will modify the mean and the width of the original distribution, but it does *not* change the distribution type. With a known multiplicity-independent efficiency, the original distribution can be completely reconstructed from the measured one, and vice versa. However, a multiplicity-dependent efficiency will distort the original distribution. In this case it is difficult to recover the original distribution. Nevertheless, one can still study how a finite, multiplicity-dependent detection efficiency changes the original distribution. The procedure applies also to the difference distribution of two independent distributions. With a known form of $\epsilon(k)$, the deviation of moments and their derivative quantities from the baseline distributions can be estimated following the procedure presented in this paper. Knowledge obtained in this work will help avoid the misinterpretation of certain observables as signals of the possible phase transition and/or the critical end point.

ACKNOWLEDGMENTS

The work for the case of a multiplicity-dependent detection efficiency was stimulated by a discussion with J. Dunlop, P. Sorensen, and H. Wang. We thank A. Bzdak and V. Koch for a fruitful discussion. A.T. was supported by the US Department of Energy under Grants No. DE-AC02-98CH10886 and No. DE-FG02-89ER40531. G.W. was supported by the US Department of Energy under Grant No. DE-FG02-88ER40424.

[1] I. Arsene *et al.*, Nucl. Phys. A **757**, 1 (2005); B. B. Back *et al.*, *ibid.* **757**, 28 (2005); J. Adams *et al.*, *ibid.* **757**, 102 (2005); K. Adcox *et al.*, *ibid.* **757**, 184 (2005); S. Chatrchyan *et al.*, J. High Energy Phys. 05 (2012) 063; B. Abelev *et al.*, Phys. Rev. Lett. **109**, 072301 (2012).
[2] S. Ejiri, Phys. Rev. D **78**, 074507 (2008).
[3] E. S. Bowman and J. I. Kapusta, Phys. Rev. C **79**, 015202 (2009).
[4] M. A. Stephanov, Prog. Theor. Phys. Suppl. **153**, 139 (2004).
[5] Z. Fodor and S. D. Katz, J. High Energy Phys. 04 (2004) 050.
[6] M. M. Aggarwal *et al.*, Phys. Rev. Lett. **105**, 022302 (2010).
[7] STAR Collaboration, D. McDonald, Nucl. Phys. A **904-905**, 907c (2013).
[8] PHENIX Collaboration, J. T. Mitchell, Nucl. Phys. A **904-905**, 903c (2013).
[9] M. A. Stephanov, Phys. Rev. Lett. **102**, 032301 (2009).
[10] M. Asakawa, S. Ejiri, and M. Kitazawa, Phys. Rev. Lett. **103**, 262301 (2009).
[11] C. Athanasiou, K. Rajagopal, and M. Stephanov, Phys. Rev. D **82**, 074008 (2010).
[12] M. A. Stephanov, Phys. Rev. Lett. **107**, 052301 (2011).
[13] J. Pumplin, Phys. Rev. D **50**, 6811 (1994).
[14] A. Bzdak and Volker Koch, Phys. Rev. C **86**, 044904 (2012).