A LOOK AT MULTIPLICITY DISTRIBUTIONS VIA MODIFIED COMBINANTS

The experimentally measured multiplicity distributions exhibit, after a closer inspection, the peculiarly enhanced void probability and the oscillatory behavior of modified combinants. We show that both these features can be used as additional sources of information, not yet fully explored, on the mechanism of multiparticle production. We provide their theoretical understanding within the class of compound distributions.

Keywords: multiplicity distributions, combinants, void probabilities, compound distributions.

1. Introduction
The experimentally measured (non-single diffractive (NSD charged) multiplicity distributions, \(P(N)\) (which are one of the most thoroughly investigated and discussed sources of information on the mechanism of the production process [1]), exhibit, after a closer inspection, the peculiarly enhanced void probability, \(P(0) > P(1)\) [2, 3], and the oscillatory behavior of the so-called modified combinants, \(C_j\), introduced by us in [4, 5] (and thoroughly discussed in [6, 7]; they are closely connected with the combinants \(C^*_N\) introduced in [8] and discussed occasionally for some time [9–14]). Both features were only rarely used as a source of information. We demonstrate that the modified combinants can be extracted experimentally from the measured \(P(N)\) by means of a recurrence relation involving all \(P(N < j)\), and that new information is hidden in their specific distinct oscillatory behavior, which, in most cases, is not observed in the \(C_j\) obtained from the \(P(N)\) commonly used to fit experimental results [4–7]. We discuss the possible sources of such behavior and the connection of \(C_j\) with the enhancement of void probabilities, and their impact on our understanding of the multiparticle production mechanism, with emphasis on understanding both phenomena within the class of compound distributions.

2. Recurrence Relation and Modified Combinants
The dynamics of the multiparticle production process is hidden in the way, in which the consecutive measured multiplicities \(N\) are connected. There are two ways of characterizing the multiplicity distributions: by means of generating functions, \(G(z) = \sum_{N=0}^{\infty} P(N)z^N\), or by some form of a recurrence relation between \(P(N)\). In the first case, one uses the Poisson distribution as a reference and characterizes deviations from it by means of combinants \(C^*_N\) defined as [8]

\[
C^*_j = \frac{1}{j!} \frac{d}{dz} \ln G(z) \Big|_{z=0}.
\]

(1)

or by the expansion

\[
\ln G(z) = \ln P(0) + \sum_{j=1}^{\infty} C^*_j z^j.
\]

(2)

For the Poisson distribution, \(C^*_1 = \langle N \rangle\) and \(C^*_j > 1 = 0\). The combinants were used in the analysis of experimental data in [9–14]. In [10,13], it was demonstrated that they are particularly useful in identifying the nature of the emitting source. It turns out that, in the case of \(S\) sources emitting particles without any restrictions concerning their number, the multiplicity \(P^S(N)\) is a completely symmetric function of degree \(N\) of the probabilities of emission, \(p_i\), the generating
function of which reduces for \( p_i \to 0 \) to the generating function of the Poisson Distribution (PD). For all probabilities remaining the same, \( p_i = p \), it reduces to the generating function of the Negative Binomial Distribution (NBD). In this case, the combinants are given by a power series

\[
C_j^* = \frac{1}{j} \sum_{i=1}^{S} p_i^j,
\]

and are always positive. However, when each of the sources can emit only a given number of particles (let us assume, for definiteness, that at most only one particle), then \( P^S(N) \) is an elementary symmetric function of degree \( N \) in the arguments, and the corresponding combinants are given by

\[
C_j^* = (-1)^{j+1} \frac{1}{j} \sum_{i=1}^{S} \left( \frac{p_i}{1-p_i} \right)^j,
\]

and alternate in sign for different \( j \)'s. For all probabilities remaining the same, \( p_i = p \), a generating function in this scenario reduces to the generating function of the Binomial Distribution (BD) and the combinants oscillate rapidly with period equal to 2.

Note that, in both cases, we were working with probabilities \( p_i \), which were not extracted from experiment, but their values were taken such that the measured multiplicity distributions are reproduced. They are then usually represented by one of the known theoretical formulae for multiplicity distributions, \( P(N) \), which can be defined either by the generating functions mentioned above or by some recurrence relations connecting different \( P(N) \). In the simplest (and most popular) case, one assumes that the multiplicity \( N \) is directly influenced only by its neighboring multiplicities, \((N \pm 1)\), i.e., we have

\[
(N+1)P(N+1) = g(N)P(N), \quad g(N) = \alpha + \beta N. \quad (5)
\]

This recurrence relation yields BD (when \( \alpha = Kp/(1-p) \) and \( \beta = -\alpha/K \), PD (when \( \alpha = \lambda \) and \( \beta = 0 \), and NBD (when \( \alpha = kp \) and \( \beta = \alpha/k \), where \( p \) denotes the particle emission probability). Usually, the first choice of \( P(N) \) in fitting the data is a single NBD [15] or two- [16, 17], three- [18], or multicomponent NBDs [19] (or some other forms of \( P(N) \) [1,15,20]). However, such a procedure only improves the agreement at large \( N \), whereas the ratio \( R = \) data/fit still deviates dramatically from unity at small \( N \) for all fits [4, 5]. This means that the measured \( P(N) \) contains information which is not yet captured by the rather restrictive recurrence relation (5). Therefore, in [4], we proposed to use a more general form of the recurrence relation (used, e.g., in counting statistics when dealing with multiplication effects in point processes [21]):

\[
(N+1)P(N+1) = \langle N \rangle \sum_{j=0}^{N} C_j P(N-j). \quad (6)
\]

This relation connects multiplicities \( N \) by means of some coefficients \( C_j \), which contain the memory of particle \( N+1 \) about all the \( N-j \) previously produced particles. The most important feature of this recurrence relation is that \( C_j \) can be directly calculated from the experimentally measured \( P(N) \) by reversing Eq. (6) [4–7]:

\[
\langle N \rangle C_j = (j+1) \left[ \frac{P(j+1)}{P(0)} \right] - \langle N \rangle \sum_{i=0}^{j-1} C_i \left[ \frac{P(0-i)}{P(0)} \right].
\]

(7)

The modified combinants \( C_j \) defined by the recurrence relation (7) are closely related to the combinants \( C_j^* \) defined by Eq. (1), namely,

\[
C_j = \frac{j+1}{\langle N \rangle} C_j^*.
\]

(8)

Using Leibnitz’s formula for the \( j \)-th derivative of the quotient of two functions \( x = G'(z)/G(z) \),

\[
x^{(j)} = \frac{1}{G} \left( G^{(j)} - j! \sum_{k=1}^{j} \frac{G^{(j-k)}}{(j-k)!} x^{(k-1)} \right),
\]

where \( G'(z)/G(z) = [\ln G(z)]' \) and \( G(z)/N!|_{z=0} = P(N) \), we immediately obtain the recurrence relation (7).

The modified combinants \( C_j \), share with the combinants \( C_j^* \) the apparent ability of identifying the nature of the emitting source mentioned above (with, respectively, Eq. (3) corresponding to the NBD case with no oscillations, and Eq. (4) corresponding to the rapidly oscillating case of BD). This also means that \( C_j \) can be calculated from the generating function \( G(z) \) of \( P(N) \),

\[
\langle N \rangle C_j = \left. \frac{1}{j!} \frac{d^{j+1} \ln G(z)}{dz^{j+1}} \right|_{z=0}.
\]

739
Thus, whereas the recurrence relation, Eq. (7), allows us to obtain the \( C_j \) from the experimental data on \( P(N) \), Eq. (10) allows for their calculation from the same generating function with the same parameters, as used for fitting \( P(N) \).

Note that \( C_j \) provide a similar measure of fluctuations as the set of cumulant factorial moments, \( K_q \), which are very sensitive to the details of the multiplicity distribution and are frequently used in phenomenological analyses of data (cf., [1, 22]),

\[
K_q = F_q = \sum_{i=1}^{q-1} \binom{q-1}{i-1} K_{q-i} F_i,
\]

where \( F_q = \langle N(N-1)(N-2)\ldots(N-q+1) \rangle \) are the factorial moments, and \( K_q \) can be expressed as an infinite series in \( C_j \),

\[
K_q = \sum_{j=q}^{\infty} \frac{(j-1)!}{(j-q)!} \langle N \rangle C_{j-1}.
\]  

However, while the cumulants are best suited to study densely populated regions of the phase space, combinatorial moments are better suited for the study of sparsely populated regions, because, according to Eq. (7), the calculation of \( C_j \) requires only a finite number of probabilities \( P(N < j) \) (which may be advantageous in applications).

The modified combinants share with the cumulants the property of additivity. For a random variable composed of independent random variables, with its generating function given by the product of their generating functions, \( G(x) = \prod G_j(x) \), the corresponding modified combinants are given by the sum of the independent components. To illustrate this property, let us consider the \( e^+e^- \) data and use the generating function \( G(z) \) formally treated as a generating function of the multiplicity distribution \( P(N) \), in which \( N \) consists of both the particles from BD (\( N_{\text{BD}} \)) and from NBD (\( N_{\text{NBD}} \)):

\[
N = N_{\text{BD}} + N_{\text{NBD}}.
\]

In this case, the multiplicity distribution can be written as

\[
P(N) = \sum_{i=0}^{\min\{N,k'\}} P_{\text{BD}}(i) P_{\text{NBD}}(N-i),
\]

and the respective modified combinants as

\[
\langle N \rangle C_j = \langle N_{\text{BD}} \rangle C_j^{\text{BD}} + \langle N_{\text{NBD}} \rangle C_j^{\text{NBD}}.
\]

Figure 1 shows the results of attempts to fit both the experimentally measured [23] multiplicity distributions and the corresponding modified combinants \( C_j \) calculated from these data (cf. [24] for details). The fits shown in Fig. 1 correspond to the parameters: \( k' = 1 \) and \( p' = 0.8725 \) for BD and \( k = 4.2 \) and \( p = 0.75 \) for NBD.

Concerning the void probabilities at all energies of interest, one observes that \( P(0) > P(1) \), a feature which cannot be reproduced by any composition of NBD used to fit the data [7]. To visualize the importance of this result, we note firstly that \( P(0) \) is

\[\text{ISSN} \text{ 2071-0194. Ukr. J. Phys. 2019. Vol. 64, No. 8}\]
strongly connected with the modified combinants \(C_j\), in fact:

\[
P(0) = \exp \left( -\sum_{j=0}^{\infty} \frac{(N)}{j+1} C_j \right).
\]  

(16)

From Eq. (7), one can deduce that the \(P(0) > P(1)\) property is possible only when \((N)C_0 < 1\). For most multiplicity distributions, \(P(2) > P(1)\), which results in an additional condition, \(C_1 > C_0(2-(N)C_0)\); taken together this means that \(C_1 > C_0\). However, because of the normalization condition \(\sum_{j=0}^{\infty} C_j = 1\), such an initial increase of \(C_j\) cannot continue for all ranks \(j\), and we should observe some kind of nonmonotonic behavior of \(C_j\) with rank \(j\) in this case. This means that all multiplicity distributions, for which the modified combinants \(C_j\) decrease monotonically with rank \(j\), do not exhibit the enhanced void probability.

3. Compound Distributions

To continue, we use the idea of compound distributions (CD), which are applicable, when (as in our case) the production process consists of a number \(M\) of some objects (clusters/fireballs/etc.) produced according to a distribution \(f(M)\) (defined by a generating function \(F(z)\)), which subsequently decay independently into a number of secondaries, \(n_i, i = 1,\ldots,M\), following some other (always the same for all \(M\)) distribution, \(g(n)\) (defined by a generating function \(G(z)\)). The resultant multiplicity distribution,

\[
h \left( N = \sum_{i=0}^{M} n_i \right) = f(M) \otimes g(n),
\]  

(17)

is a compound distribution of \(f\) and \(g\) with the generating function

\[
H(z) = F[G(z)].
\]  

(18)

Equation (18) means that, in the case where \(f(M)\) is a Poisson distribution with the generating function

\[
F(z) = \exp[\lambda(z-1)],
\]  

(19)

the combinants for any other distribution \(g(n)\) with a generating function \(G(z)\), which are obtained from the compound distribution \(h(N) = P_{PD} \otimes g(n)\) and calculated with the use of Eq. (10), do not oscillate and are equal to

\[
C_j = \frac{\lambda(j+1)}{(N)} g(j+1).
\]  

(20)

This fact explains why \(C_j\) from NBDs do not oscillate. This is because NBD is a compound distribution of the Poisson and logarithmic distributions. This means that \(g(n) = -p^n/[n \ln(1-p)]\), and \(h(N)\) is NBD with \(k = -\lambda/\ln(1-p)\). In this case, \(C_j\) coincide with those derived before and given by Eq. (3). Actually, this reasoning applies to all more complicated compound distributions, with any distribution itself being a compound Poisson distribution. This property limits the set of distributions \(P(N)\) leading to oscillating \(C_j\), to BD, and to all compound distributions based on it. In this case, the period of oscillations is determined by the number of particles emitted from the source. For the compound distributions based on BD with \(P(n) = \delta_{n,m}\), we have

\[
C_j = (-1)^{j/m+1} \frac{K}{\langle N \rangle} \left( \frac{p}{1-p} \right)^{j/m+1},
\]  

(21)

(for \(j = mk\) and \(C_j = 0\) for \(j \neq mk\), where \(k = 1,2,3,\ldots\)). For broader distributions \(P(n)\), we get a smoother \(C_j\) dependence on rank \(j\). For example, for \(P(n)\) given by the Poisson distribution (with expected value \(\lambda\)), we obtain a Compound Binomial Distribution (CBD) with the generating function

\[
H(z) = \{p \exp[\lambda(z-1)] + 1-p\}^K,
\]  

(22)

and the modified combinants are given by

\[
C_j = \frac{(-1)^{j+1} K e^\lambda \lambda^{j+1} (1-p)}{\langle N \rangle \left( e^{\lambda(p-1)/p} + 1 \right)^{j+1}} A_j \left( \frac{p}{\lambda} - 1 \right),
\]  

(23)

where \(A_j(x)\) are the Eulerian polynomials. As an illustration, we show in Fig. 2 that, by compounding

\[\text{Fig. 2.} \quad C_j \text{ for } BD, \text{BD compounded with } \delta_{n,m} \text{ with } m = 10 \text{ and compounded with the Poisson distribution with } \lambda = 10\]
BD with a Poisson distribution, one gains control over the period of oscillations (now equal to 2\(\lambda\)) and their amplitude. However, it turns out that such a combination does not allow us to fit data.

4. Multicomponent

The situation improves substantially, when one uses a multi-CBD based on Eq. (22). But the agreement is not yet satisfactory. It turns out that the situation improves dramatically, if one replaces the Poisson distribution by NBD, and, additionally, uses a two-component version of such CBD with

\[
P(N) = \sum_{i=1,2} w_i h(N; p_i, K_i, k_i, m_i) \tag{24}
\]

with the generating function of each component equal to

\[
H(z) = \left[ p \left( \frac{1 - p'}{1 - p/z} \right)^K + 1 - p \right]. \tag{25}
\]

In such a case, as can be seen in Fig. 3, one gains a satisfactory control over the periods of oscillations, their amplitudes, and their behavior as a function of the rank \(j\). Moreover, one can nicely fit \(P(N)\) and \(C_j\). Of special importance is the fact that the enhancement \(P(0) > P(1)\) is also reproduced in this approach.

The above result also explains the apparent success in fitting the experimentally observed oscillations of \(C_j\) by using a weighted sum of the three NBD used in [26]. Such a distribution uses freely selected weights and parameters \((p, k)\) of NBDs and, therefore, resembles the compound distribution of BD with NBD. However, we note that the sum of \(M\) variables (with \(M = 0, 1, 2, \ldots\)), each from NBD characterized by parameters \((p, k)\), is described by NBD characterized by \((p, Mk)\). Therefore, as discussed before, it cannot reproduce the void probability \(P(0)\). This can be reproduced only in the case where \(M = 0, 1, \ldots, K\) is distributed according to BD, and we have a \(K\)-component NBD (where the consecutive NBDs have precisely defined parameters \(k\)).

\[
P(N) = \sum_{M=0}^{K} P_{BD}(M) P_{NBD}(N; p, Mk). \tag{26}
\]

In this case, one also has the \(M = 0\) component, which is lacking in the previous multi-NBD case used in [26]. This is the reason for that, whereas the compound (BD\&NBD) distribution reproduces the void probability, \(P(0)\), the single NBD (or any combination of NBDs) do not. This means that the observation of the peculiar behavior of the void probability discussed above signals the necessity of using some compound distribution based on BD to fit data for \(P(N)\) (and the \(C_j\) obtained from it).

5. Summary and Conclusions

Since the time of Ref. [8], one encounters essentially no detailed experimental studies of the combinants and only rather sporadic attempts at their phenomenological use to describe the multiparticle production processes. We demonstrate that the modified combinants \(C_j\) are a valuable tool for the in-

Fig. 3. Multiplicity distributions \(P(N)\) measured in \(p+p\) collisions by ALICE [25] (upper panel) and the corresponding modified combinants \(C_j\) (lower panel). Data are fitted using a two compound distribution (BD+NBD) given by Eqs. (25) and (24) with the parameters: \(K_1 = K_2 = 3, p_1 = 0.9, p_2 = 0.645, k_1 = 2.8, k_2 = 1.34, m_1 = 5.75, m_2 = 23.5, w_1 = 0.24\) and \(w_2 = 0.76\).
vestigations of multiplicity distributions, and $C_j$ deduced from the measured multiplicity distributions, $P(N)$, could provide additional information on the dynamics of the particle production. This, in turn, could allow us to reduce the number of possible interpretations presented so far and, perhaps, answer some of the many still open fundamental questions (that this is possible, despite experimental errors, has been shown in [7, 26]). Finally, let us note that a large number of papers suggest some kind of universality in the mechanisms of hadron production in $e^+e^−$ annihilations and in $pp$ and $pp$ collisions. This arises from observations of the average multiplicities and relative dispersions in both types of processes (cf., e.g., [27, 28]). However, as we have shown here, the modified combinant analysis reveals differences between these processes. Namely, while, in $e^+e^−$ annihilations, we observe oscillations of $C_j$ with period 2, the period of oscillations in $pp$ collisions is $\sim 10$ times longer, and the amplitude of oscillations in both types of processes differs dramatically. At the moment, this problem remains open and awaits a further investigation.

This research was supported in part by the Polish Ministry of Science and Higher Education (Contract No. DIR/WK/2016/2010/17-1), by the National Science Centre (NCN) (Contract No. DEC-2016/23/B/ST2/00176) (G.W.), and by the NCN grant 2016/23/B/ST2/00692 (M.R.). We would like to thank Dr. Nicholas Keeley for reading the manuscript.


Received 08.07.19