HAL
open science

Easy all-order cumulants

Philippe Di Francesco, Maxime Guilbaud, Matthew Luzum, Jean-Yves Ollitrault

To cite this version:

Philippe Di Francesco, Maxime Guilbaud, Matthew Luzum, Jean-Yves Ollitrault. Easy all-order cumulants. 2016. cea-01472222

HAL Id: cea-01472222

https://cea.hal.science/cea-01472222v1

Preprint submitted on 20 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Easy all-order cumulants

Philippe Di Francesco,[1,2] Maxime Guilbaud,[3] Matthew Luzum,[4] and Jean-Yves Ollitrault[1]

[1] *Institut de physique théorique, Université Paris Saclay, CNRS, CEA, F-91191 Gif-sur-Yvette, France*
[2] *Department of Mathematics, University of Illinois MC-382, Urbana, IL 61821, U.S.A.*
[3] *Rice University, Houston, USA*
[4] *Instituto de Física, Universidade de São Paulo, C.P. 66318, 05315-970 São Paulo, SP, Brazil*

(Dated: December 20, 2016)

We present a systematic procedure for analyzing cumulants to arbitrary order in the context of heavy-ion collisions. It generalizes and improves existing procedures in many respects. In particular, particles which are correlated are allowed to belong to different phase-space windows, which may overlap. It also allows for the analysis of cumulants at any order, using a simple algorithm rather than complicated expressions to be derived and coded by hand. In the case of azimuthal correlations, it automatically corrects for detector non-uniformity, and it is useful for numerous other applications as well. We discuss several of these applications: anisotropic flow, event-plane correlations, symmetric cumulants, net baryon and net charge fluctuations.

arXiv:1612.05634v1 [nucl-th] 16 Dec 2016

I. INTRODUCTION

A nucleus-nucleus collision at ultrarelativistic energies typically emits thousands of particles [1]. The large multiplicity enables one to accurately measure various high-order correlations and cumulants. Cumulants are *connected* correlations, and high-order cumulants are typically used to probe the emergence of collective effects. More specifically, cumulants of azimuthal correlations have been measured [2–6] in order to study the collective expansion of the quark-gluon plasma in the direction of the impact parameter, or elliptic flow [7, 8]. Cumulants of the net proton [9] and net charge [10, 11] distribution are analyzed in order to search for the critical fluctuations associated with the QCD phase transition at finite baryon density [12, 13]. Recent years have witnessed the emergence of a great variety of new correlation observables: Event-plane correlations [14], correlations between transverse momentum and anisotropic flow [15], or between two different Fourier harmonics of anisotropic flow [16, 17], which all involve cumulants of multiparticle correlations.

We propose a new framework for the cumulant analysis, which is more flexible than existing frameworks and allows to fully exploit the potential of multiparticle correlation analyses in high-energy collisions. Our understanding of the collision dynamics has considerably evolved since cumulants were introduced in this context [18]. The importance of event-to-event fluctuations in the flow pattern [19, 20] was only recognized later [21, 22]. More specifically, it was recently shown that the rapidity dependence of anisotropic flow fluctuates event to event [23, 24]. This longitudinal decorrelation [25] induces a sizable variation of azimuthal correlations with the relative rapidity [26]. Existing analysis frameworks [27, 28], where all particles are typically taken from the same rapidity window, do not allow to study this effect.

Precision studies of high-order cumulants are also needed. They have been argued to be a crucial probe of collective behavior in proton-nucleus collisions [29] and proton-proton collisions [30] and first analyses of order 6 and 8 cumulants are promising [31]. High-order cumulants also provide insight into non-Gaussian fluctuations in nucleus-nucleus collisions [32–34], where they have also been measured up to order 8 [6, 35]. The maximum value of 8 is dictacted by existing analysis frameworks, but higher orders are feasible experimentally.

The oldest framework [36] extracts cumulants by numerically tabulating the generating function and using a finite-difference approximation to compute its successive derivatives. The numerical errors resulting from this procedure are hard to evaluate. A new framework by Bilandzic *et al.* [28] uses explicit expressions of cumulants in terms of moments, which is a more robust approach. However, only a finite set of cumulants are provided, and azimuthal symmetry is assumed in order to simplify the algebraic complexity of these expressions. Azimuthal asymmetries in the detector acceptance and efficiency must therefore be corrected beforehand. On the other hand, the cumulant analysis automatically corrects for such asymmetries [18], so it is tempting to use cumulants for this purpose as well.

Our new framework generalizes the approach of Bilandzic [28] in several respects:

- It applies to arbitrary observables, not only to Fourier coefficients of the azimuthal distribution.
- One can correlate particles from different bins in rapidity and transverse momentum.
- No assumption is made regarding the detector acceptance and efficiency.
- Cumulants can be evaluated to arbitrarily high order.

In Sec. II, we illustrate the procedure on the simple example of azimuthal pair correlations. We then generalize results to higher-order correlations in Sec. III, where we show how to remove self-correlations to all orders, and

in Sec. IV, where we give the inversion formulas for cumulants as a function of moments to all orders. Specific implementations are discussed in Sec. V.

II. EXAMPLE: PAIR CORRELATION

We illustrate the steps of the calculation on the simple example of azimuthal pair correlations [22]. One typically wants to compute an average value of $\cos n\Delta\Phi$, where n is the harmonic order and $\Delta\Phi$ is the relative azimuthal angle between two particles in the same event [37]. Specifically, one takes the first particle from a region of momentum space A and the second from region B, and one evaluates

$$\langle \cos n\Delta\Phi \rangle \equiv \frac{\left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle}{\left\langle \sum_{\text{pairs}} \quad 1 \quad \right\rangle}, \tag{1}$$

where ϕ_j and ϕ_k are the azimuthal angles of particles belonging to the same event, "pairs" is a shorthand notation for "$j \in A$, $k \in B$ and $j \neq k$", and angular brackets in the right-hand side denote an average over events.[1]

One could evaluate the sums over pairs as nested loops over j and k, but it is more efficient to instead factorize the sums [28]. For example, if A and B are disjoint:

$$\sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} = \sum_{j\in A} e^{in\phi_j} \sum_{k\in B} e^{-in\phi_k}. \tag{2}$$

In the case where regions A and B overlap, such that they share some of the same particles, one must exclude the extra terms with $j = k$, corresponding to a trivial correlation of a particle with itself (self-correlation):

$$\sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} = \sum_{j\in A} e^{in\phi_j} \sum_{k\in B} e^{-in\phi_k} - \sum_{j\in A\cap B} 1, \tag{3}$$

where the final sum is over the intersection of sets A and B, $A \cap B$, and each term has unit contribution since $e^{in(\phi_j-\phi_j)} = 1$

The next step is to average over a large number of collision events. We first assume that particles are independent, in the sense that the number of particles in two disjoint momentum bins $\mathbf{p}_1$ and $\mathbf{p}_2$ are independent variables. Then, the average number of pairs factorizes as a product of single averages, and

$$\left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle = \left\langle \sum_{j\in A} e^{in\phi_j} \right\rangle \left\langle \sum_{k\in B} e^{-in\phi_k} \right\rangle. \tag{4}$$

For an ideal (isotropic) detector and azimuthal symmetric regions A and B, the right-hand side vanishes identically, since every collision event has an arbitrary azimuthal orientation. In a more realistic experimental situation, the detector efficiency has azimuthal asymmetries and the right-hand side is non-zero. However, Eq. (4) is still valid when particles are independent.

In the more general case where particles are not independent, we define the connected correlation as the difference between the two sides of this equation. It thus isolates the physical correlation, and naturally corrects for asymmetries in the detector:

$$\left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle_c \equiv \left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle - \left\langle \sum_{j\in A} e^{in\phi_j} \right\rangle \left\langle \sum_{k\in B} e^{-in\phi_k} \right\rangle, \tag{5}$$

where the subscript c in the left-hand side denotes the connected part of the correlation, i.e., the cumulant [18].

Note that our definition of independent particles is not exactly equivalent to assuming that for a given pair of particles, azimuthal angles ϕ_j and ϕ_k are independent [28]. With this alternative definition, one can write an equation similar to Eq. (4), where the left-hand side is divided by the average number of pairs, and the single-particle averages in the right-hand side are divided by the average number of particles in A and B. The definition of the cumulant is then modified accordingly:

$$\langle \cos n\Delta\Phi \rangle \equiv \frac{\left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle}{\left\langle \sum_{\text{pairs}} \quad 1 \quad \right\rangle} - \frac{\left\langle \sum_{j\in A} e^{in\phi_j} \right\rangle}{\left\langle \sum_{j\in A} \quad 1 \quad \right\rangle} \frac{\left\langle \sum_{k\in B} e^{-in\phi_k} \right\rangle}{\left\langle \sum_{k\in B} \quad 1 \quad \right\rangle}. \tag{6}$$

The method in this paper can be applied with either choice, as explained at the end of Sec. IV. The advantage of our definition is that it simpler, and allows to treat multiplicity fluctuations and correlations on the same footing as anisotropic flow. Its apparent drawback is that a correlation can be induced by a large fluctuation of the global multiplicity, which is of little physical interest. However, one easily gets rid of this effect by using a fine centrality binning, which is recommended anyway for correlation analyses [38].[2]

Instead of explicitly considering the connected correlation Eq. (5) (or its variant Eq. (6)) to correct for detector

[1] The imaginary part of the right-hand side vanishes if parity is conserved, up to statistical fluctuations and asymmetries in the detector efficiency.

[2] Note that early cumulant analyses, which used wide centrality bins due to limited statistics, used a fixed subset of the multiplicity [3] in order to avoid the effects of multiplicity fluctuations in conjunction with detector asymmetries.

anisotropy, one can do the correction in other ways. Bilandzic *et al.* [28] weight each particle with $1/p$, where p is the efficiency (probability of detection) at the point where the particle hits the detector. After the correction, azimuthal asymmetry can be assumed, and the first term alone in the right-hand side of Eq. (5) is equivalent to the connected part. This "inverse weighting" method can still be used here, but is no longer necessary. In particular, note that inverse weighting cannot be applied when there are holes in the detector, in which case the efficiency p is 0.

Eliminating the nested sums gives the final expression for the numerator and denominator of the measurement

$$\begin{aligned}\left\langle \sum_{\text{pairs}} e^{in(\phi_j-\phi_k)} \right\rangle_c \equiv & \left\langle \sum_{j\in A} e^{in\phi_j} \sum_{k\in B} e^{-in\phi_k} \right\rangle \\ & - \left\langle \sum_{j\in A\cap B} 1 \right\rangle \\ & - \left\langle \sum_{j\in A} e^{in\phi_j} \right\rangle \left\langle \sum_{k\in B} e^{-in\phi_k} \right\rangle .\end{aligned} \tag{7}$$

and

$$\left\langle \sum_{\text{pairs}} 1 \right\rangle = \left\langle \sum_{j\in A} 1 \sum_{k\in B} 1 \right\rangle - \left\langle \sum_{j\in A\cap B} 1 \right\rangle . \tag{8}$$

In the following Sections, we generalize the above discussion to higher-order correlations and arbitrary observables.

III. SUBTRACTING SELF-CORRELATIONS

Cumulants can be constructed from moments, which are correlations that count n-tuples in a collision event, where n now denotes the order of the correlation. Depending on the observable, one may weight every particle differently depending on its momentum, as in Eq. (1). In general, one evaluates in each event multiple sums of the type

$$Q(A_1,\ldots,A_n) \equiv \sum_{j_1\in A_1,\ldots,j_n\in A_n} q_1(j_1)\ldots q_n(j_n), \tag{9}$$

where $j_1,\cdots,j_n$ label particles chosen from n sets labelled $A_1,\cdots,A_n$ (representing, e.g., specific regions in p_T and η, or particular particle species), all indices in the sum are different, and $q_i(j)$ are functions of the particle momentum. In the numerator of Eq. (1), for example, we have $q_1(j)=e^{in\phi_j}$ and $q_2(j)=e^{-in\phi_j}$.

The sum runs over all possible n-tuples in the same event. If one uses nested loops, the time needed to evaluate such sums grows with the multiplicity M like M^n, which can be computationally prohibitive.

In this section, we explain how to express them as a function of simple sums such as:

$$Q(A_i) \equiv \sum_{j\in A_i} q_i(j). \tag{10}$$

This reduces the number of operations from M^n down to M for any order n. In the case of the analysis of anisotropic flow, $Q(A_i)$ is the usual flow vector for particle set A_i [39].

The idea is to factorize the sum, as in Eq. (2). In the case when there is some overlap in the sets of particles A_i, however, one must subtract terms in the sum where the same particle appears more than once, as in Eq. (3). In the case of pair correlations, the notation (10) allows to recast this subtraction in compact form:

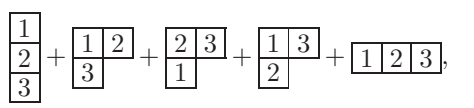
$$Q(A_1,A_2) = Q(A_1)Q(A_2) - Q(A_1\cap A_2). \tag{11}$$

We have introduced the auxiliary notation

$$Q(A_1\cap A_2) \equiv \sum_{j\in A_1\cap A_2} q_1(j)q_2(j), \tag{12}$$

where the sum runs over all particles belonging to both A_1 and A_2.

Consider now a 3-particle correlation. The product $Q(A_1)Q(A_2)Q(A_3)$ contains all possible triplets, plus the self-correlations which must be removed. One separates the sum into different contributions, depending on which particles are identical: (1) j_1, j_2, j_3 all different, (2) $j_1 = j_2 \neq j_3$, (3) $j_2 = j_3 \neq j_1$, (4) $j_1 = j_3 \neq j_2$, (5) $j_1 = j_2 = j_3$. This decomposition can be represented with Young diagrams:

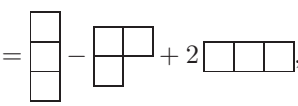


(13)

where each box is associated with a set A_i. In each diagram, different rows correspond to different values of the indices, and the values of the indices are identical for boxes belonging to the same horizontal row. The first term in Eq. (13) is the term we want to isolate and others must be subtracted. The subtraction, which generalizes Eq. (3) to third order, is derived in Appendix B. One obtains:

$$\begin{aligned}Q(A_1,A_2,A_3) = & \; Q(A_1)Q(A_2)Q(A_3) \\ & -Q(A_1\cap A_2)Q(A_3) \\ & -Q(A_2\cap A_3)Q(A_1) \\ & -Q(A_1\cap A_3)Q(A_2) \\ & +2Q(A_1\cap A_2\cap A_3),\end{aligned} \tag{14}$$

which we represent diagrammatically as

$Q(A_1,A_2,A_3) =$  (15)

where we have omitted the labels since the weights are identical for all permutations of the indices. The right-hand side of Eq. (14) is a sum over all partitions of the

set $\{A_1, A_2, A_3\}$. This can be generalized to arbitrary n, as shown in Appendix A. The contribution of a partition is the product of contributions of its subsets, called blocks. Each row in the Young diagram corresponds to a block of the partition. The contribution of a block of k elements $\{A_{i_1}, \cdots, A_{i_k}\}$ is the product of the integer weight $(-1)^{k-1}(k-1)!$ and $Q(A_{i_1} \cap \cdots \cap A_{i_k})$, which is defined by a straightforward generalization of Eq. (12):

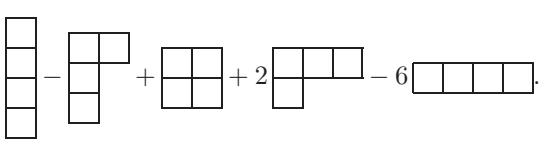
$$Q(A_{i_1} \cap \cdots \cap A_{i_k}) \equiv \sum_{j \in A_{i_1} \cap \cdots \cap A_{i_k}} q_{i_1}(j) \cdots q_{i_k}(j). \quad (16)$$

Blocks of $k = 1$, 2, 3 elements get respective weights of 1, -1, 2, which explains the factors in front of each term in Eq. (14).

We write explicitly, for sake of illustration, the corresponding formula for the 4-particle correlation:

$$\begin{aligned} Q(A_1, A_2, A_3, A_4) = {} & Q(A_1)Q(A_2)Q(A_3)Q(A_4) \\ & -Q(A_1 \cap A_2)Q(A_3)Q(A_4) \\ & -Q(A_1 \cap A_3)Q(A_2)Q(A_4) \\ & -Q(A_2 \cap A_3)Q(A_1)Q(A_4) \\ & -Q(A_1 \cap A_4)Q(A_2)Q(A_3) \\ & -Q(A_2 \cap A_4)Q(A_1)Q(A_3) \\ & -Q(A_3 \cap A_4)Q(A_1)Q(A_2) \\ & +Q(A_1 \cap A_2)Q(A_3 \cap A_4) \\ & +Q(A_1 \cap A_3)Q(A_2 \cap A_4) \\ & +Q(A_1 \cap A_4)Q(A_2 \cap A_3) \\ & +2Q(A_1 \cap A_2 \cap A_3)Q(A_4) \\ & +2Q(A_2 \cap A_3 \cap A_4)Q(A_1) \\ & +2Q(A_1 \cap A_3 \cap A_4)Q(A_2) \\ & +2Q(A_1 \cap A_2 \cap A_4)Q(A_3) \\ & -6Q(A_1 \cap A_2 \cap A_3 \cap A_4), \quad (17) \end{aligned}$$

which we represent diagrammatically as:



(18)

The weight of a given partition can be read directly from the Young diagram. It is the product over all rows (all blocks) of $(-1)^{k-1}(k-1)!$, where k is the number of boxes in the row (number of elements in block).

In order to implement the subtraction in the most general case, one must generate all partitions of the set $\{A_1, \cdots, A_n\}$. An efficient algorithm has been described by Orlov, with a public domain C++ implementation available [40]. A sum can then be taken, with coefficient for each term given by the above formula.

IV. FROM MOMENTS TO CUMULANTS

We now define the cumulants of n-particle correlations. Let $f(p_1, \cdots, p_n)dp_1 \cdots dp_n$ denote the probability of finding a n-tuple in $dp_1 \cdots dp_n$. We call $f(p_1, \cdots, p_n)$ the n-point function. It is normalized to the average number of n-tuples:

$$\int_p f(p_1, \cdots, p_n) = \langle M(M-1) \cdots (M-n+1) \rangle, \quad (19)$$

where $\int_p$ is a shorthand notation for $\int dp_1 \cdots dp_n$, and M denotes the total multiplicity of an event.

The connected n-point function $f_c(p_1, \cdots, p_n)$, or cumulant, is the contribution of the n-particle correlation. For any order n, it is defined recursively by isolating n-particle clusters through an order-by-order decomposition of $f(p_1, \cdots, p_n)$ [41]. The 1-point functions are equal by definition:

$$f(p_1) = f_c(p_1). \quad (20)$$

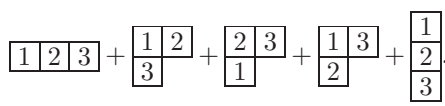
The two-point function is the sum of an uncorrelated part and a correlated part $f_c(p_1, p_2)$:

$$f(p_1, p_2) = f_c(p_1, p_2) + f_c(p_1)f_c(p_2). \quad (21)$$

Similarly, one defines $f_c(p_1, p_2, p_3)$ as the part of $f(p_1, p_2, p_3)$ which is not due to lower-order correlations:

$$\begin{aligned} f(p_1, p_2, p_3) = {} & f_c(p_1, p_2, p_3) \\ & +f_c(p_1, p_2)f_c(p_3) \\ & +f_c(p_2, p_3)f_c(p_1) \\ & +f_c(p_1, p_3)f_c(p_2) \\ & +f_c(p_1)f_c(p_2)f_c(p_3). \quad (22) \end{aligned}$$

The right-hand side of this equation is again a sum over partitions of the set $\{p_1, p_2, p_3\}$, where each cluster corresponds to a block of the partition. It can be represented diagrammatically as:



(23)

Generalization to arbitrary order n is straightforward.

The average of $Q(A_1, \cdots, A_n)$ over many collision events is a weighted integral of $f(p_1, \cdots, p_n)$:

$$\langle Q(A_1, \cdots, A_n) \rangle = \int_p q_1(p_1) \cdots q_n(p_n) f(p_1, \cdots, p_n), \quad (24)$$

We refer to such averages as *moments*. The cumulant decomposition applies to moments after multiplying equations (21), (22) by $q_i(p_i)$ and integrating over p_i. The cumulant of order 2 is thus given by the inversion formula

$$\langle Q(A_1, A_2) \rangle_c \equiv \langle Q(A_1, A_2) \rangle - \langle Q(A_1) \rangle \langle Q(A_2) \rangle, \quad (25)$$

which generalizes Eq. (5), and which we rewrite synthetically as

$$\langle Q_1 Q_2 \rangle_c = \langle Q_1 Q_2 \rangle - \langle Q_1 \rangle \langle Q_2 \rangle. \quad (26)$$

Note that the cumulant is unchanged if one shifts Q_i by a constant value. This property of translational invariance [42], which is true to all orders, explains why cumulants are remarkably stable with respect to detector imperfections.

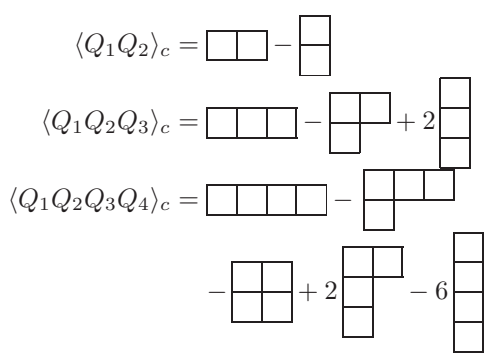
We write for sake of illustration the inversion formulas giving cumulants of order 3 and 4 as a function of the corresponding moments:

$$\begin{aligned}\langle Q_1Q_2Q_3\rangle_c = \langle Q_1Q_2Q_3\rangle \\ -\langle Q_1Q_2\rangle\langle Q_3\rangle \\ -\langle Q_1Q_3\rangle\langle Q_2\rangle \\ -\langle Q_2Q_3\rangle\langle Q_1\rangle \\ +2\langle Q_1\rangle\langle Q_2\rangle\langle Q_3\rangle,\end{aligned} \tag{27}$$

and

$$\begin{aligned}\langle Q_1Q_2Q_3Q_4\rangle_c = \langle Q_1Q_2Q_3Q_4\rangle \\ -\langle Q_1Q_2Q_3\rangle\langle Q_4\rangle \\ -\langle Q_2Q_3Q_4\rangle\langle Q_1\rangle \\ -\langle Q_1Q_3Q_4\rangle\langle Q_2\rangle \\ -\langle Q_1Q_2Q_4\rangle\langle Q_3\rangle \\ -\langle Q_1Q_2\rangle\langle Q_3Q_4\rangle \\ -\langle Q_1Q_3\rangle\langle Q_2Q_4\rangle \\ -\langle Q_1Q_4\rangle\langle Q_2Q_3\rangle \\ +2\langle Q_1Q_2\rangle\langle Q_3\rangle\langle Q_4\rangle \\ +2\langle Q_1Q_3\rangle\langle Q_2\rangle\langle Q_4\rangle \\ +2\langle Q_2Q_3\rangle\langle Q_1\rangle\langle Q_4\rangle \\ +2\langle Q_1Q_4\rangle\langle Q_2\rangle\langle Q_3\rangle \\ +2\langle Q_2Q_4\rangle\langle Q_1\rangle\langle Q_3\rangle \\ +2\langle Q_3Q_4\rangle\langle Q_1\rangle\langle Q_2\rangle \\ -6\langle Q_1\rangle\langle Q_2\rangle\langle Q_3\rangle\langle Q_4\rangle.\end{aligned} \tag{28}$$

The right-hand sides of these equations are again sums over partitions of the sets $\{A_1, A_2, \cdots, A_n\}$ for $n = 2, 3, 4$. We represent them diagrammatically as



(29)

The weight of each term is given by a classic Möbius inversion formula (see Appendix A). It is equal to $(-1)^{n-1}(n-1)!$, where n is the number of blocks of the partition, i.e., the number of rows of the diagram.

Note the striking formal analogy between the subtraction of self-correlations and the cumulant expansion. Both involve a summation over set partitions (so both can be generated by the same algorithm [40]). The only difference is that the weight associated with each Young diagram involves the number of boxes in each row for self-correlations, and the number of rows for the cumulant expansion.

Finally, we point out that the corrections for self-correlations cancel to some extent in the cumulant. Take as an example the order 3 cumulant, given by Eq. (27). Moments of order 2 and 3 in the right-hand side must be corrected for self-correlations using Eqs. (11) and (14), and averaged over events. After summing all terms, the correction to the cumulant from self-correlation can be written as

$$\begin{aligned}&\langle Q(A_1\cap A_2)\rangle\langle Q(A_3)\rangle - \langle Q(A_1\cap A_2)Q(A_3)\rangle \\ &+\langle Q(A_1\cap A_3)\rangle\langle Q(A_2)\rangle - \langle Q(A_1\cap A_3)Q(A_2)\rangle \\ &+\langle Q(A_2\cap A_3)\rangle\langle Q(A_1)\rangle - \langle Q(A_2\cap A_3)Q(A_1)\rangle \\ &+2\langle Q(A_1\cap A_2\cap A_3)\rangle.\end{aligned} \tag{30}$$

The first line is, up to a sign, the connected correlation between $Q(A_1 \cap A_2)$ and $Q(A_3)$, which is usually much smaller than their respective magnitudes. This is true to all orders [18], and the contribution to cumulants from self-correlations becomes smaller and smaller as the order increases.[3] It has been checked experimentally [5, 44] that one obtains compatible results with and without self-correlations in the limit of large cumulant order, through an analysis of Lee-Yang zeros [45, 46].

Self-correlations must be subtracted on an event-by-event basis, and the number of terms increases strongly with the order n (it is the Bell number), therefore they are a limiting factor for the calculation. However, it is expected that they are negligible beyond a certain order (which should explicitly be checked by doing the calculation with and without self correlations). The cumulant expansion can then be pushed to arbitrarily high order, limited only by statistical uncertainty.

Our definition of cumulants in this section follows from our choice of random variables, which are the numbers of n-tuples in a momentum bin $dp_1 \cdots dp_n$. As discussed in Sec. II, an alternative choice is to choose as random variables the momenta $p_1, \cdots, p_n$ themselves [28]. Our method also accommodates this definition, at the expense of a slight complication: One must normalize each moment by the average number of n-tuples before taking the cumulant (as in Eqs. (7) and (8) of Ref. [28]). This number is obtained by repeating the calculations of Sec. III with $q_k(j_k) = 1$ in Eq. (9). However, this complication does not appear to offer any advantage. Our formulation has the advantage that it provides a unified framework for all analyses of correlations and fluctuations, as we now explain.

V. APPLICATIONS

The interest of cumulants is that they subtract the effect of lower correlations, and isolate n-particle correlations. If a nucleus-nucleus collision is a superposition

[3] In the context of the analysis of anisotropic flow, this is consistent with the expectation that higher-order cumulants are less sensitive to nonflow effects, since nonflow effects are of the same order as self-correlations [43].

of a fixed number N of independent nucleon-nucleon collisions, a moment of order n scales like N^n, while the corresponding cumulant is proportional to N. If, on the other hand, there are collective effects in the system, cumulants are typically of the same order as moments, so that large cumulants are a clear signature of collective effects.

By collective effect, one typically means an all-particle correlation. Collective effects often arise from global fluctuations, which affect the whole system. The fluctuation of the total multiplicity, already mentioned in Sec. II, is a mundane collective effect which can be eliminated by a fine centrality binning [38]. On the other hand, fixing the total multiplicity M in each event also generates cumulants to all orders, but they are proportional to M at any order (see Appendix C). As a result, working at a fixed multiplicity doesn't generate fake collective effects.

We now describe how usual cumulant analyses can be implemented within our framework.

A. Net charge fluctuations and related studies

Our framework, where the random variables are numbers of particles in momentum bins (as opposed to momenta of given particles [28]), naturally encompasses correlation studies involving these numbers themselves, such as net charge and net baryon fluctuations.

We first illustrate the formulas derived in previous sections by discussing the simplest case, where one takes all particles from the same set, $A_1 = \cdots = A_n$, and all the functions q_i in Eq. (9) are equal to 1. In this case, the summation in this equation just counts the number of n-tuples of M particles (cf. Eq. (19)):

$$Q(A_1, \cdots, A_n) = M(M-1)\cdots(M-n+1). \tag{31}$$

This result can be used to check the validity of Eqs. (11), (14) and (17). Each factor $Q(...)$ in these equations is equal to M, therefore, they reduce to:

$$\begin{aligned} Q(A_1, A_2) &= M^2 - M \\ Q(A_1, A_2, A_3) &= M^3 - 3M^2 + 2M \\ Q(A_1, A_2, A_3, A_4) &= M^4 - 6M^3 + 11M^2 - 6M, \end{aligned} \tag{32}$$

which agree with the general result (31) after expanding in powers of M.

Averaging $Q(A_1, \cdots, A_n)$ over events, one obtains the (unnormalized) factorial moment F_n, which counts the average number of n-tuples:

$$F_n \equiv \langle M(M-1)\cdots(M-n+1)\rangle \tag{33}$$

If M is distributed according to a Poisson distribution, $F_n = \langle M\rangle^n$.

The cumulants as defined in Sec. IV are the *factorial cumulants* K_n [47] of the distribution of M; unlike traditional multiplicity cumulants, they vanish for $n \geq 2$ for a Poisson distribution, and are therefore automatically corrected for trivial statistical fluctuations.

In the same way, one can study correlations between multiplicities in two different rapidity windows, such as forward-backward correlations [48, 49] which have been proposed as a probe of longitudinal fluctuations [50]. Analyses could easily be extended to higher-order cumulants [51].

Consider next the case where q_i in Eq. (9) is the baryon number or the electric charge. Assuming that $q_i = \pm 1$ for all particles, the moment of order 2, as defined by Eq. (11), is $Q(A_1, A_2) = \Delta N_{\rm ch}^2 - M$ where $= \Delta N_{\rm ch} = \sum_i q_i$ is the net charge and $M = \sum_i (q_i)^2$ the charged multiplicity. Thus trivial fluctuations are again subtracted, and such observables give more direct access to interesting quantities than the raw distribution of $\Delta N_{\rm ch}$ [11] or traditional cumulants [52].

B. Anisotropic flow

The analysis of anisotropic flow is one the most important practical applications of cumulants in high-energy physics. We recall the definition of the relevant observables, and how they are obtained in our framework.

The flow paradigm [22, 53] states that the bulk of particle production is well described by independent-particle emission from an underlying probability distribution. The classic picture is that of a relativistic fluid — near freeze-out, the system is a fluid, composed of well-defined particles that are uncorrelated with each other (e.g., as in a Boltzmann description). In other words, if every event had the same hydrodynamic initial conditions (and a fixed orientation), the connected correlations defined by Eqs. (21), (22) would vanish.

In a single event, then, all information is contained in the single-particle distribution $F(p)$. Note that this single particle distribution is different from $f(p)$ in Eq. (20), which is $F(p)$ averaged over events (and azimuthal orientation). Anisotropic flow is the particular observation that the single-particle distribution depends on azimuthal angle ϕ:

$$F(p) = \frac{N}{2\pi} \sum_{n=-\infty}^{\infty} V_n e^{in\phi}, \tag{34}$$

with $V_0 = 1$, $V_{-n} = V_n^*$, and N is the average number of particles in the event. One denotes by $v_n \equiv |V_n|$ the anisotropic flow coefficient in harmonic n [56] in a particular event.

One can instead absorb the factor of N into the coefficients [57]

$$F(p) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \mathcal{V}_n e^{in\phi}, \tag{35}$$

where now $\mathcal{V}_n = NV_n$. While less standard, these are convenient quantities since they are additive with respect to rebinning in momentum space, and are more natural quantities with respect to the cumulant analysis.

The single-event distribution $F(p)$ fluctuates from one event to the next. Upon averaging over events (and fluctuations), this generates correlations to all orders. Moments and cumulants of multiparticle correlations are moments and cumulants of $F(p)$. [4]

Even if the bulk of particle production can be described according to this single-particle distribution, it is expected for there to be small correlations that cannot be captured by $F(p)$. Such effects are typically referred to as "non-flow", and can arise from sources such as Bose-Einstein correlations, resonance decays, unthermalized jets, momentum conservation, etc. Observables involving higher-order cumulants are expected to be less sensitive to such nonflow effects [18].

The measure of v_n from the cumulant of order $2k$ is usually noted $v_n\{2k\}$, and it can easily be obtained with the procedure outlined above. This generalizes the discussion of pair correlations outlined in Sec. II. One now takes $2k$-tuples, again taken from the same region of phase space ("integrated" flow). The factor q_i in Eq. (9) is equal to $e^{in\phi}$ for the first k particles and to $e^{-in\phi}$ for the next k particles. The subtraction of self correlations carried out in Sec. III generalizes that of Ref. [28]. For instance, one easily checks that the order 4 result Eq. (17) reproduces the numerator of Eq. (18) of Ref. [28].[5] We denote by $\langle |Q_n|^{2k}\rangle_c$, the cumulant obtained after subtracting self-correlations, averaging over events and subtracting lower-order moments (Sec. IV). As explained above, it is a cumulant of the distribution of $\mathcal{V}_n$. We thus define:

$$\mathcal{V}_n\{2k\}^{2k} \equiv \frac{1}{a_{2k}} \langle |Q_n|^{2k}\rangle_c, \tag{36}$$

where a_{2k} is the ratio of the coefficients of the expansions of $\ln I_0(x)$ (cumulants) and $I_0(x)$ (moments) to order x^{2k} [18]:

$$\ln I_0(x) = \sum_{k=1}^{\infty} \frac{a_{2k} x^{2k}}{2^{2k}(k!)^2}, \tag{37}$$

which gives $a_2 = 1$, $a_4 = -1$, $a_6 = 4$, $a_8 = -33$, $a_{10} = 456$, $a_{12} = -9460$, etc. This normalization ensures that $\mathcal{V}_n\{2k\} = |\mathcal{V}_n|$ for all k if $|\mathcal{V}_n|$ were the same for all events.

In order to match the usual normalization convention of the flow coefficients v_n, one normalizes this cumulant by the average number of $2k$-tuples. The standard azimuthal cumulants are thus given by

$$v_n\{2k\}^{2k} \equiv \frac{1}{a_{2k}} \frac{\langle |Q_n|^{2k}\rangle_c}{F_{2k}}. \tag{38}$$

This normalization ensures that again $v_n\{2k\} = v_n$ when there are no fluctuations.

For the traditional measurement of "differential" flow [18], the only difference with the previous case is that the first particle is taken from a different (restricted) phase space window B, while the $2k-1$ remaining particles are taken from the same set of particles A. One then scales the resulting cumulant with the corresponding number of $2k$-tuples (with 1 particle in B and $2k-1$ particles in A). The flow in B, traditionally denoted by $v'_n\{2k\}$, is again given by an equation similar to (38), where one replaces in the left-hand side $v_n\{2k\}^{2k}$ with $v'_n\{2k\} v_n\{2k\}^{2k-1}$. Note that the coefficients a_k are the same as for integrated flow.

In principle, each particle can be taken from an arbitrary region in momentum space, and many other differential analyses are possible [58]. With our normalization conventions, all differential cumulants are additively related to integrated measurements.

C. Symmetric cumulants, event-plane correlations

The normalized symmetric cumulant $\mathrm{NSC}(m,n)$ with $n \neq m$ can be defined by [17][6]

$$\mathrm{NSC}(m,n) = \frac{\langle Q_n Q_m Q_{-n} Q_{-m}\rangle_c}{\langle |Q_n|^2\rangle_c \langle |Q_m|^2\rangle_c} \tag{39}$$

$$\overset{\text{(flow)}}{=} \frac{\langle |\mathcal{V}_m|^2 |\mathcal{V}_n|^2\rangle - \langle |\mathcal{V}_m|^2\rangle\langle |\mathcal{V}_n|^2\rangle}{\langle |\mathcal{V}_m|^2\rangle\langle |\mathcal{V}_n|^2\rangle}. \tag{40}$$

The numerator is a 4-particle cumulant [16] with $q_1 = e^{im\phi}$, $q_2 = e^{in\phi}$, $q_3 = e^{-im\phi}$, $q_4 = e^{-in\phi}$. The fact that it is measured through a 4-particle cumulant guarantees that nonflow effects are small. Equation (40) is the value assuming the flow paradigm of independent particles in each event.

The mean square values $\langle |\mathcal{V}_n|^2\rangle = F_2 v_n\{2\}^2$ in the denominator are 2-particle cumulants obtained as described in Sec. V B, which may be biased by short-range correlations (especially nonflow effects) unless a rapidity gap is applied. The existing analysis [17] implements a gap

[4] There are non-trivial mathematical consequences to having uncorrelated particles in each event, e.g., for pair correlations [54, 55].

[5] The weight of each event is the same as in Ref. [28], in the sense that all n-tuples are treated on an equal footing (and as a result, each *event* can have significantly different contribution, if multiplicity fluctuates). A variety of prescriptions are found in the literature. The scalar-product analysis [2] and the Lee-Yang zero analysis [43] use a prescription equivalent to ours. The cumulant analysis of Ref. [36] weights each event by $1/M$, and the analysis of event-plane correlations [14] implements a similar weight $1/E_T$, where E_T is the energy deposited in the calorimeter. Finally, factors of $1/\sqrt{M}$ were used in the cumulant analysis of Ref. [18] and in studies of the distribution of the flow vector [2]).

[6] As with previous cumulants, the ALICE Collaboration chose to normalize each individual moment by F_2 or F_4. See the discussion at the end of Sec. IV. We choose to have no normalization factors in the normalized measurement.

in the denominator, but not in the numerator, hence neglecting the effect of longitudinal fluctuations. it would be interesting to redo the analysis by implementing the same gap in the numerator and the denominator., and studing how NSC(m,n) varies with the gap.

Event-plane correlations [14] are Pearson correlation coefficients between different complex flow hamonics [59, 60]. For instance, the two-plane correlation between $\mathcal{V}_2$ and $\mathcal{V}_4$ is defined as (we use the notation of ATLAS):

$$\langle\cos(4(\Phi_2-\Phi_4))\rangle_w \equiv \frac{\langle Q_4 Q_{-2} Q_{-2}\rangle_c}{\sqrt{\langle|Q_4|^2\rangle_c \langle|Q_2|^4\rangle}} \tag{41}$$

$$\overset{\text{(flow)}}{=} \frac{\langle \mathcal{V}_4(\mathcal{V}_2^*)^2\rangle}{\sqrt{\langle|\mathcal{V}_4|^2\rangle\langle|\mathcal{V}_2|^4\rangle}}. \tag{42}$$

The numerator is a 3-particle cumulant obtained with $q_1 = e^{4i\phi}$, $q_2 = q_3 = e^{-2i\phi}$. The ATLAS analysis uses two sets of particles separated with a rapidity gap, where particles 2 and 3 belong to the same bin. Thus self-correlations between particles 2 and 3 are not removed. It has been argued that they are small [61], but it will be interesting to check experimentally. The denominator of Eq. (41) involves moments which, as in the case of symmetric cumulants, may be biased by nonflow effect unless a rapidity gap is applied (note that the 4-particle v_2 factor is a moment, not a cumulant).

It has been pointed out [62] that the value of NSC$(4,2)$ measured by ALICE seems large compared to what one would expect based on the corresponding event-plane correlation measured by ATLAS [14], where a large rapidity gap is implemented. It would be interesting to measure both the symmetric cumulant and the event-plane correlation with the same kinematic cuts.

D. Correlation between transverse momentum and anisotropic flow

The correlation between transverse momentum and anisotropic flow recently proposed by Bozek as a further test of hydrodynamic behavior [15] is a straightforward application of our formalism. It is a 3-particle cumulant with $q_1 = p_t$, $q_2 = e^{in\phi}$, $q_3 = e^{-in\phi}$. Bozek recommends to use particles from three different rapidity intervals A, B, C separated with gaps in order to avoid nonflow correlations but predicts that results should be identical if $A = B = C$ provided that self-correlations are subtracted.

VI. CONCLUSION

We have proposed a new, unified framework for cumulant analyses which is more systematic and flexible than existing frameworks, and discussed its practical implementation for the analysis of factorial cumulants and anisotropic flow. A major improvement is that one can correlate particles in arbitrary regions of phase space. Application to proton-nucleus and nucleus-nucleus at RHIC and the LHC should shed light on longitudinal fluctuations. The new framework also extends beyond the analysis of anisotropic flow, and we anticipate a rich program of generalized cumulant analyses on this basis in the near future.

ACKNOWLEDGEMENTS

This work is funded under the USP-COFECUB project Uc Ph 160-16. PDF is supported by the NSF Grant DMS-1301636 and the Morris and Gertrude Fine endowment. MG is supported under DOE Contract No. de-sc0012185 and the Welch Foundation (Grant No. C-1845), and thanks Wei Li and Hubert Hansen for discussions. JYO thanks the Department of Theoretical Physics, Mumbai, where this paper was written, for hospitality and financial support. We thank Giuliano Giacalone for careful reading of the manuscript.

Appendix A: Möbius inversion

In this Appendix, we recall some known facts [63] on Möbius inversion applied to functions of set partitions. We show in particular how it implies the moment/cumulant relations of Sec. IV, as well as the self-correlation subtraction of Section III.

1. Set partitions and Möbius inversion

Let $[n] \equiv \{1, 2, ..., n\}$ and $\mathcal{P}(n)$ the set of its partitions. We denote by $\mathbf{I} = \{I_1, ..., I_k\}$ a partition, where I_j are subsets of $[n]$, called blocks of the partition. The number of blocks k is called the length of the partition $\mathbf{I}$ and denoted by $|\mathbf{I}|$. There is a natural partial order relation on partitions which we denote by $\leq$ and is defined as follows. If $\mathbf{I}$ and $\mathbf{J}$ are two partitions of $[n]$, $\mathbf{I} \leq \mathbf{J}$ if the partition $\mathbf{I}$ is a refinement of the partition $\mathbf{J}$, that is, if each block of $\mathbf{I}$ is included in a block of $\mathbf{J}$. For instance we have $\{\{1,2\},\{3\},\{4\}\} \leq \{\{1,2\},\{3,4\}\}$ but the order relation does not relate $\{\{1,2\},\{3,4\}\}$ to $\{\{1,3\},\{2,4\}\}$. The finest partition is denoted by $\hat{\mathbf{0}} \equiv \{\{1\},\{2\},...,\{n\}\}$ and the coarsest partition by $\hat{\mathbf{1}} \equiv [n]$. For all $\mathbf{I} \in \mathcal{P}(n)$, $\hat{\mathbf{0}} \leq \mathbf{I} \leq \hat{\mathbf{1}}$.

The Möbius inversion formula goes as follows [64]. Assume we have two real functions f, g defined on $\mathcal{P}(n)$, such that for all $\mathbf{I} \in \mathcal{P}(n)$:

$$f(\mathbf{J}) = \sum_{\hat{\mathbf{0}} \leq \mathbf{I} \leq \mathbf{J}} g(\mathbf{I}). \tag{A1}$$

then we have the inverse relation:

$$g(\mathbf{J}) = \sum_{\hat{\mathbf{0}} \leq \mathbf{I} \leq \mathbf{J}} \mu(\mathbf{I}, \mathbf{J})\, f(\mathbf{I}), \tag{A2}$$

where the Möbius function $\mu(\mathbf{I},\mathbf{J})$ is given by

$$\mu(\mathbf{I},\mathbf{J}) = (-1)^{|\mathbf{I}|-|\mathbf{J}|}\prod_{i=1}^{n}((i-1)!)^{r_i(\mathbf{I},\mathbf{J})}, \tag{A3}$$

where $r_i(\mathbf{I},\mathbf{J})$ denotes the number of blocks of $\mathbf{J}$ containing exactly i blocks of $\mathbf{I}$.

Likewise, there is a *dual* Möbius inversion formula: assuming f, g are related via

$$f(\mathbf{I}) = \sum_{\mathbf{I}\leq\mathbf{J}\leq\hat{\mathbf{1}}} g(\mathbf{J}) \tag{A4}$$

for all $\mathbf{I}\in\mathcal{P}(n)$ then we have the inverse relations:

$$g(\mathbf{I}) = \sum_{\mathbf{I}\leq\mathbf{J}\leq\hat{\mathbf{1}}} \mu(\mathbf{I},\mathbf{J})\, f(\mathbf{J}) \tag{A5}$$

2. Applications

An example of the decomposition (A1) is the decomposition of moments into cumulants (Section IV). One defines $f(\mathbf{I})$ as the product of moments over all blocks, and $g(\mathbf{I})$ as the product of cumulants over all blocks, that is;

$$\begin{aligned} f(\mathbf{I}) &\equiv \prod_{j=1}^{|\mathbf{I}|}\left\langle \prod_{l\in I_j} Q_l \right\rangle \\ g(\mathbf{I}) &\equiv \prod_{j=1}^{|\mathbf{I}|}\left\langle \prod_{l\in I_j} Q_l \right\rangle_c . \end{aligned} \tag{A6}$$

Then the Möbius inversion formula (A2) specialized to $\mathbf{J} = \hat{\mathbf{1}}$ gives the connected correlation function $\langle Q_1\cdots Q_n\rangle_c$ in terms of moments as a sum over partitions, with coefficients

$$\mu(\mathbf{I},\hat{\mathbf{1}}) = (-1)^{|\mathbf{I}|-1}\,(|\mathbf{I}|-1)! \tag{A7}$$

Formulas (26), (27), (28) follow as well as the general case.

Similarly, the dual Möbius inversion formula applies directly to the problem of subtracting self-correlations (Section III). We now define $f([n]) \equiv Q(A_1)...Q(A_n)$ (sum over all indices, including self correlations) $g([n]) \equiv Q(A_1,\cdots,A_n)$ (sum over different indices), and more generally:

$$\begin{aligned} f(\mathbf{I}) &\equiv Q(\cap_{i\in I_1}A_i)Q(\cap_{i\in I_2}A_i)...Q(\cap_{i\in I_n}A_i) \\ g(\mathbf{I}) &\equiv Q(\cap_{i\in I_1}A_i, \cap_{i\in I_2}A_i, ..., \cap_{i\in I_n}A_i). \end{aligned} \tag{A8}$$

The relations (A4) are clearly satisfied, as shown by splitting uncontrained sums into sums of constrained ones, for instance:

$$\begin{aligned} Q(A_1)Q(A_2) &= \sum_{i,j} q_1(i)q_2(j) \\ &= \sum_{i\neq j} q_1(i)q_2(j) + \sum_{i=j} q_1(i)q_2(i) \\ &= Q(A_1,A_2) + Q(A_1\cap A_2). \end{aligned} \tag{A9}$$

The dual Möbius inversion formula (A5) for $\mathbf{I} = \hat{\mathbf{0}}$ gives the expression for $g(\hat{\mathbf{0}}) = Q(A_1, A_2, ..., A_n)$ in terms of the $f(\mathbf{J})$, with coefficients

$$\mu(\hat{\mathbf{0}},\mathbf{J}) = \prod_{i=1}^{|\mathbf{J}|}(-1)^{|J_i|-1}\,(|J_i|-1)! \tag{A10}$$

where $|J_i|$ is the cardinality of the block J_i, equal to the length of the i-th row of the Young diagram of $\mathbf{J}$. Formulas (11), (14), (17) and their generalization follow.

Appendix B: Recursion relation

We derive a relation for generating the self-correlation corrections order by order. It is simple to implement and easy to understand, but less efficient numerically than the general method exposed in Sec. III. We want to evaluate sums of the type

$$Q(A_1,\ldots,A_n) \equiv \sum_{j_1\in A_1,\ldots,j_n\in A_n} q_1(j_1)\ldots q_n(j_n), \tag{B1}$$

where all indices in the sum are different. Self correlations can be subtracted order by order. Once one has a formula that works for $n-1$, then to order n one has

$$\begin{aligned} Q(A_1,\ldots,A_n) &= Q(A_1,\ldots,A_{n-1})Q(A_n) \\ &\quad - \sum_{k=1}^{n-1} Q(A_1,...,A_k\cap A_n,...,A_{n-1}). \end{aligned} \tag{B2}$$

The first term in the right-hand side takes into account the conditions that the indices i_1 to i_{n-1} are all different, and the last term subtracts the contributions from $i_k = i_n$. It is understood that the product q_iq_j should be used together with the intersection $A_i\cap A_j$. For $n=2$, Eq. (B2) gives Eq. (3). For $n=3$, it gives

$$\begin{aligned} Q(A_1,A_2,A_3) = Q(A_1,A_2)Q(A_3) - Q(A_1\cap A_3, A_2) \\ -Q(A_1, A_2\cap A_3). \end{aligned} \tag{B3}$$

Substituting Eq. (3) into Eq. (B3), one obtains Eq. (14). For $n=4$, Eq. (B2) gives

$$\begin{aligned} Q(A_1,A_2,A_3,A_4) = Q(A_1,A_2,A_3)Q(A_4) \\ -Q(A_1\cap A_4, A_2, A_3) \\ -Q(A_1, A_2\cap A_4, A_3) \\ -Q(A_1, A_2, A_3\cap A_4) \end{aligned} \tag{B4}$$

Substituting Eq. (B3) into Eq. (B4), one obtains Eq. (17). Generating the subtraction to order n with this method requires $n!$ operations so that it is less efficient in practice than generating partitions, since the number of partitions is the Bell number which grows more slowly with n than $n!$.

Appendix C: Generating function

Moments and cumulants to all orders are conveniently expressed in terms of generating functions. The moment defined by Eq. (24) can be obtained by expanding the generating function:

$$G(z_1, \cdots, z_n) \equiv \left\langle \prod_{j=1}^{M} \left(1 + z_1 q_1(j) + \cdots + z_n q_n(j)\right) \right\rangle, \tag{C1}$$

where the product runs over all particles in the event. If one expands the product, the moment defined by (the average over events of) Eq. (9) is the coefficient in front of $z_1 \cdots z_n$. The corresponding cumulant is given by the expansion of $\ln G$ to the same order:

$$\langle Q_1 \cdots Q_n \rangle_c = \frac{\partial^n}{\partial z_1 \cdots \partial z_n} \ln G(z_1, \cdots, z_n) \bigg|_{z_1 = \cdots = z_n = 0}. \tag{C2}$$

Writing G as a product over all particles [36] guarantees that self-correlations do not appear at any order in the expansion. If one does not subtract self correlations, then the generating function takes the usual exponential form [18]:

$$G(z_1, \cdots, z_n) \equiv \langle \exp\left(z_1 Q(A_1) + \cdots + z_n Q(A_n)\right) \rangle, \tag{C3}$$

with $Q(A_i) = \sum_j q_i(j)$. Note that Eq. (C1) can also be written in an exponential form analogous to (C3) by introducing Grassmann variables [65]. This formal analogy shows that the algebraic relations linking moments to cumulants, derived in Sec. IV, are identical irrespective of whether or not self-correlations are subtracted. Both forms of the generating function have been used [5, 44] in the context of the analysis of elliptic flow with Lee-Yang zeros [43].

As an application of the formalism, we evaluate the generating function (C1) in the simple case of independent particles, where all terms in the product are independent. First, consider the case where the multiplicity M is fixed. Then, Eq. (C1) gives

$$G(z_1, \cdots, z_n) = \left(1 + z_1 \langle q_1 \rangle + \cdots + z_n \langle q_n \rangle\right)^M. \tag{C4}$$

Therefore, $\ln G$ is proportional to M and cumulants of arbitrary order scale like M, as stated in the first paragraph of Sec. V.

If, on the other hand, all the connected n-point functions $f(p_1, \cdots, p_n)$ vanish for $n \geq 2$, then, the multiplicity M follows a Poisson distribution:

$$p_M = \frac{\langle M \rangle^M}{M!} e^{-\langle M \rangle}. \tag{C5}$$

Inserting into Eq. (C1) and summing the series, one obtains

$$G(z_1, \cdots, z_n) = \exp\left(\langle M \rangle (z_1 \langle q_1 \rangle + \cdots + z_n \langle q_n \rangle)\right). \tag{C6}$$

Therefore, $\ln G(z_1, \cdots, z_n)$ is linear in all the variables z_1 and cumulants of order ≥ 2 vanish identically, as expected.

[1] E. Abbas *et al.* [ALICE Collaboration], Phys. Lett. B **726**, 610 (2013) doi:10.1016/j.physletb.2013.09.022 [arXiv:1304.0347 [nucl-ex]].

[2] C. Adler *et al.* [STAR Collaboration], Phys. Rev. C **66**, 034904 (2002) doi:10.1103/PhysRevC.66.034904 [nucl-ex/0206001].

[3] C. Alt *et al.* [NA49 Collaboration], Phys. Rev. C **68**, 034903 (2003) doi:10.1103/PhysRevC.68.034903 [nucl-ex/0303001].

[4] K. Aamodt *et al.* [ALICE Collaboration], Phys. Rev. Lett. **105**, 252302 (2010) doi:10.1103/PhysRevLett.105.252302 [arXiv:1011.3914 [nucl-ex]].

[5] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Rev. C **87**, no. 1, 014902 (2013) doi:10.1103/PhysRevC.87.014902 [arXiv:1204.1409 [nucl-ex]].

[6] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **74**, no. 11, 3157 (2014) doi:10.1140/epjc/s10052-014-3157-z [arXiv:1408.4342 [hep-ex]].

[7] J. Y. Ollitrault, Phys. Rev. D **46**, 229 (1992). doi:10.1103/PhysRevD.46.229

[8] U. Heinz and R. Snellings, Ann. Rev. Nucl. Part. Sci. **63**, 123 (2013) doi:10.1146/annurev-nucl-102212-170540 [arXiv:1301.2826 [nucl-th]].

[9] M. M. Aggarwal *et al.* [STAR Collaboration], Phys. Rev. Lett. **105**, 022302 (2010) doi:10.1103/PhysRevLett.105.022302 [arXiv:1004.4959 [nucl-ex]].

[10] L. Adamczyk *et al.* [STAR Collaboration], Phys. Rev. Lett. **113**, 092301 (2014) doi:10.1103/PhysRevLett.113.092301 [arXiv:1402.1558 [nucl-ex]].

[11] A. Adare *et al.* [PHENIX Collaboration], Phys. Rev. C **93**, no. 1, 011901 (2016) doi:10.1103/PhysRevC.93.011901 [arXiv:1506.07834 [nucl-ex]].

[12] M. A. Stephanov, Phys. Rev. Lett. **102**, 032301 (2009) doi:10.1103/PhysRevLett.102.032301 [arXiv:0809.3450 [hep-ph]].

[13] S. Gupta, X. Luo, B. Mohanty, H. G. Ritter and N. Xu, Science **332**, 1525 (2011) doi:10.1126/science.1204621 [arXiv:1105.3934 [hep-ph]].

[14] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. C **90**, no. 2, 024905 (2014) doi:10.1103/PhysRevC.90.024905 [arXiv:1403.0489 [hep-ex]].

[15] P. Bozek, Phys. Rev. C **93**, no. 4, 044908 (2016) doi:10.1103/PhysRevC.93.044908 [arXiv:1601.04513

[nucl-th]].
[16] A. Bilandzic, C. H. Christensen, K. Gulbrandsen, A. Hansen and Y. Zhou, Phys. Rev. C **89**, no. 6, 064904 (2014) doi:10.1103/PhysRevC.89.064904 [arXiv:1312.3572 [nucl-ex]].
[17] J. Adam *et al.* [ALICE Collaboration], Phys. Rev. Lett. **117**, 182301 (2016) doi:10.1103/PhysRevLett.117.182301 [arXiv:1604.07663 [nucl-ex]].
[18] N. Borghini, P. M. Dinh and J. Y. Ollitrault, Phys. Rev. C **63**, 054906 (2001) doi:10.1103/PhysRevC.63.054906 [nucl-th/0007063].
[19] C. E. Aguiar, Y. Hama, T. Kodama and T. Osada, Nucl. Phys. A **698**, 639 (2002) doi:10.1016/S0375-9474(01)01447-6 [hep-ph/0106266].
[20] M. Miller and R. Snellings, nucl-ex/0312008.
[21] B. Alver *et al.* [PHOBOS Collaboration], Phys. Rev. Lett. **98**, 242302 (2007) doi:10.1103/PhysRevLett.98.242302 [nucl-ex/0610037].
[22] B. Alver and G. Roland, Phys. Rev. C **81**, 054905 (2010) Erratum: [Phys. Rev. C **82**, 039903 (2010)] doi:10.1103/PhysRevC.82.039903, 10.1103/PhysRevC.81.054905 [arXiv:1003.0194 [nucl-th]].
[23] P. Bozek, W. Broniowski and J. Moreira, Phys. Rev. C **83**, 034911 (2011) doi:10.1103/PhysRevC.83.034911 [arXiv:1011.3354 [nucl-th]].
[24] V. Khachatryan *et al.* [CMS Collaboration], Phys. Rev. C **92**, no. 3, 034911 (2015) doi:10.1103/PhysRevC.92.034911 [arXiv:1503.01692 [nucl-ex]].
[25] L. G. Pang, G. Y. Qin, V. Roy, X. N. Wang and G. L. Ma, Phys. Rev. C **91**, no. 4, 044904 (2015) doi:10.1103/PhysRevC.91.044904 [arXiv:1410.8690 [nucl-th]].
[26] L. Adamczyk *et al.* [STAR Collaboration], Phys. Rev. C **88**, no. 1, 014904 (2013) doi:10.1103/PhysRevC.88.014904 [arXiv:1301.2187 [nucl-ex]].
[27] N. Borghini, P. M. Dinh and J. Y. Ollitrault, nucl-ex/0110016.
[28] A. Bilandzic, R. Snellings and S. Voloshin, Phys. Rev. C **83**, 044913 (2011) doi:10.1103/PhysRevC.83.044913 [arXiv:1010.0233 [nucl-ex]].
[29] L. Yan and J. Y. Ollitrault, Phys. Rev. Lett. **112**, 082301 (2014) doi:10.1103/PhysRevLett.112.082301 [arXiv:1312.6555 [nucl-th]].
[30] V. Khachatryan *et al.* [CMS Collaboration], [arXiv:1606.06198 [nucl-ex]].
[31] V. Khachatryan *et al.* [CMS Collaboration], Phys. Rev. Lett. **115**, no. 1, 012301 (2015) doi:10.1103/PhysRevLett.115.012301 [arXiv:1502.05382 [nucl-ex]].
[32] R. S. Bhalerao, M. Luzum and J. Y. Ollitrault, Phys. Rev. C **84**, 034910 (2011) doi:10.1103/PhysRevC.84.034910 [arXiv:1104.4740 [nucl-th]].
[33] H. Grönqvist, J. P. Blaizot and J. Y. Ollitrault, Phys. Rev. C **94**, no. 3, 034905 (2016) doi:10.1103/PhysRevC.94.034905 [arXiv:1604.07230 [nucl-th]].
[34] G. Giacalone, L. Yan, J. Noronha-Hostler and J. Y. Ollitrault, arXiv:1608.01823 [nucl-th].
[35] B. B. Abelev *et al.* [ALICE Collaboration], Phys. Rev. C **90**, no. 5, 054901 (2014) doi:10.1103/PhysRevC.90.054901 [arXiv:1406.2474 [nucl-ex]].
[36] N. Borghini, P. M. Dinh and J. Y. Ollitrault, Phys. Rev. C **64**, 054901 (2001) doi:10.1103/PhysRevC.64.054901 [nucl-th/0105040].
[37] K. Aamodt *et al.* [ALICE Collaboration], Phys. Lett. B **708**, 249 (2012) doi:10.1016/j.physletb.2012.01.060 [arXiv:1109.2501 [nucl-ex]].
[38] F. G. Gardim, F. Grassi, M. Luzum and J. Noronha-Hostler, arXiv:1608.02982 [nucl-th].
[39] A. M. Poskanzer and S. A. Voloshin, Phys. Rev. C **58**, 1671 (1998) doi:10.1103/PhysRevC.58.1671 [nucl-ex/9805001].
[40] M. Orlov, "Efficient generation of set partitions" (2002), Obtained from `https://www.cs.bgu.ac.il/~orlovm/papers/partitions.pdf`; Source code available at `https://www.cs.bgu.ac.il/~orlovm/papers/partitions-source.t`
[41] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).
[42] D. Teaney and L. Yan, Phys. Rev. C **83**, 064904 (2011) doi:10.1103/PhysRevC.83.064904 [arXiv:1010.1876 [nucl-th]].
[43] R. S. Bhalerao, N. Borghini and J. Y. Ollitrault, Nucl. Phys. A **727**, 373 (2003) doi:10.1016/j.nuclphysa.2003.08.007 [nucl-th/0310016].
[44] B. I. Abelev *et al.* [STAR Collaboration], Phys. Rev. C **77**, 054901 (2008) doi:10.1103/PhysRevC.77.054901 [arXiv:0801.3466 [nucl-ex]].
[45] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952). doi:10.1103/PhysRev.87.404
[46] R. S. Bhalerao, N. Borghini and J. Y. Ollitrault, Phys. Lett. B **580**, 157 (2004) doi:10.1016/j.physletb.2003.11.056 [nucl-th/0307018].
[47] I. M. Dremin, Phys. Lett. B **313**, 209 (1993). doi:10.1016/0370-2693(93)91214-8
[48] B. B. Back *et al.* [PHOBOS Collaboration], Phys. Rev. C **74**, 011901 (2006) doi:10.1103/PhysRevC.74.011901 [nucl-ex/0603026].
[49] B. I. Abelev *et al.* [STAR Collaboration], Phys. Rev. Lett. **103**, 172301 (2009) doi:10.1103/PhysRevLett.103.172301 [arXiv:0905.0237 [nucl-ex]].
[50] A. Bzdak and D. Teaney, Phys. Rev. C **87**, no. 2, 024906 (2013) doi:10.1103/PhysRevC.87.024906 [arXiv:1210.1965 [nucl-th]].
[51] A. Bialas and K. Zalewski, Nucl. Phys. A **860**, 56 (2011) doi:10.1016/j.nuclphysa.2011.05.006 [arXiv:1101.1907 [hep-ph]].
[52] P. Braun-Munzinger, A. Rustamov and J. Stachel, arXiv:1612.00702 [nucl-th].
[53] M. Luzum, J. Phys. G **38**, 124026 (2011) doi:10.1088/0954-3899/38/12/124026 [arXiv:1107.0592 [nucl-th]].
[54] F. G. Gardim, F. Grassi, M. Luzum and J. Y. Ollitrault, Phys. Rev. C **87**, no. 3, 031901 (2013) doi:10.1103/PhysRevC.87.031901 [arXiv:1211.0989 [nucl-th]].
[55] R. S. Bhalerao, J. Y. Ollitrault, S. Pal and D. Teaney, Phys. Rev. Lett. **114**, no. 15, 152301 (2015) doi:10.1103/PhysRevLett.114.152301 [arXiv:1410.7739 [nucl-th]].
[56] S. Voloshin and Y. Zhang, Z. Phys. C **70**, 665 (1996) doi:10.1007/s002880050141 [hep-ph/9407282].
[57] A. Mazeliauskas and D. Teaney, Phys. Rev. C **91**, no. 4, 044902 (2015) doi:10.1103/PhysRevC.91.044902 [arXiv:1501.03138 [nucl-th]].

[58] F. Gardim, F. Grassi, M. Guilbaud, M. Luzum and J.-Y. Ollitrault, in preparation.

[59] M. Luzum and J. Y. Ollitrault, Phys. Rev. C **87**, no. 4, 044907 (2013) doi:10.1103/PhysRevC.87.044907 [arXiv:1209.2323 [nucl-ex]].

[60] L. Yan and J. Y. Ollitrault, Phys. Lett. B **744**, 82 (2015) doi:10.1016/j.physletb.2015.03.040 [arXiv:1502.02502 [nucl-th]].

[61] R. S. Bhalerao, J. Y. Ollitrault and S. Pal, Phys. Rev. C **88**, 024909 (2013) doi:10.1103/PhysRevC.88.024909 [arXiv:1307.0980 [nucl-th]].

[62] G. Giacalone, L. Yan, J. Noronha-Hostler and J. Y. Ollitrault, Phys. Rev. C **94**, no. 1, 014906 (2016) doi:10.1103/PhysRevC.94.014906 [arXiv:1605.08303 [nucl-th]].

[63] Stanley, Richard P. "Enumerative Combinatorics. Vol. 1, vol. 49 of Cambridge Studies in Advanced Mathematics." (1997).

[64] Giovanni Peccati, Murad S. Taqqu, preprint `arXiv:0811.1726 [math.CO]` (chapters 1-3).

[65] M. Creutz, in Advanced Series on Directions in High Energy Physics: Volume 11 "Quantum Fields on the Computer", Edited by Michael Creutz (1992)