Structure of correlation functions

P. Carruthers
Department of Physics, University of Arizona, Tucson, Arizona 85721
(Received 9 October 1990)

Experimental determinations of correlation functions typically involve events of differing particle number. In order to interpret such data, it is important to relate these measurements to fixed-number correlations. We exhibit the total correlation function in terms of fixed n correlations and fluctuations with respect to the average. We note that moments, constructed as integrals of the appropriate correlation functions, can be dominated in the case of broad distributions by cumulant moments constructed from single-particle density fluctuations.

I. INTRODUCTION

Consider an ensemble composed of subsystems of variable particle number n. We imagine the sample to be sufficiently large that one can define the probabilities q_n to find an n-particle subsystem within which the joint probability $Q_n(x_1,x_2,\ldots,x_n)$ is defined in some space. We assume normalizations

$$\int Q_n(x_1,\ldots,x_n)dx_1\cdots dx_n=1\ , \quad \sum_n q_n=1\ . \tag{1.1}$$

Hence the joint probability of finding a system of n particles which are located at the indicated points is

$$p_n(x_1,\ldots,x_n)=q_nQ_n(x_1,\ldots,x_n)\ . \tag{1.2}$$

In Eq. (1.1) the integrations of all the x variables are taken over the same "volume" Ω.

Examples include multiparticle production in nuclear and high-energy physics, where $P_n=\sigma_n/\sigma_{\rm in}$ is typically the probability of charged particle production constructed from the "n-prong" cross section σ_n and $\sigma_{\rm in}$ the inelastic charged cross section; galaxy counts, and of course statistical-mechanical systems where the probabilities follow from the Gibbs ensemble. It should be noticed that our physical format is a little different from the grand ensemble of statistical mechanics, from a mathematical point of view.

It might seem that everything should already be known about the description of such systems. However, the detailed analysis of particular physical problems produces specific question, points of view, and conjectures needing a fresh analysis. A case in point concerns the construction of higher-order cumulant correlations from two-particle correlations. A specific formulation apparently describes galaxy-galaxy correlations,[1,2] cluster-cluster correlations,[3] as well as multihadron correlations[4,5] observed in high-energy collisions. It is not yet known what the origin might be for these structures: occurrence in such different physical systems points to a statistical origin of such behavior.

The present paper analyzes the effect of admixing different particle numbers in the construction of correlation functions. It can happen that more or less trivial contributions overwhelm "true" correlations existing in a system with fixed particle number. In particle physics the following well-known example occurs.[6] The variable x becomes either momentum, or some other measure such as longitudinal rapidity $y=\frac{1}{2}\ln(E+p_z)/(E-p_z)$. Let $\rho_1^{(n)}(x)$, $\rho_2^{(n)}(x,x')$ be the one- and two-particle density correlations for fixed n (defined in Sec. II). In terms of the overall densities $\rho_1(x)=\sum_n q_n\rho_1^{(n)}(x)$, $\rho_2(x,x')=\sum_n q_n\rho_2^{(n)}(x,x')$ the cumulant correlation functions

$$C_2(x,x')\equiv\rho_2(x,x')-\rho_1(x)\rho_1(x')\ , \tag{1.3}$$

$$C_2^{(n)}(x,x')\equiv\rho_2^{(n)}(x,x')-\rho_1^{(n)}(x)\rho_1^{(n)}(x')\ , \tag{1.4}$$

are connected as follows:

$$\begin{aligned} C_2(x,x')&=\sum_n q_n\rho_2^{(n)}(x,x')-\sum_n q_n\rho_1^{(n)}(x)\sum_m q_m\rho_1^{(m)}(x') \\ &=\sum_n q_n[\rho_2^{(n)}(x,x')-\rho_1^{(n)}(x)\rho_1^{(n)}(x')] \\ &\quad+\sum_n q_n[\rho_1^{(n)}(x)-\rho_1(x)] \\ &\qquad\times[\rho_1^{(n)}(x')-\rho_1(x')]\ ; \end{aligned} \tag{1.5}$$

$$C_2(x,x')=\langle C_2^{(n)}(x,x')\rangle+\langle\Delta\rho_n(x)\Delta\rho_n(x')\rangle \tag{1.6}$$

where $\Delta\rho_n=\rho_1^{(n)}-\rho_1$ is the single-particle density fluctuation. Here the bracket notation clearly denotes $\sum_n q_n\theta_n=\langle\theta_n\rangle$.

This example shows how to decompose the "total" correlation function C_2, defined in the obvious way, in terms of the averages over the fixed multiplicity $C_2^{(n)}$ and the single-particle density fluctuation. The main goal of this paper is to generalize Eq. (1.6) to higher orders. We quote the following results:

$$C_3(x,x',x'')=\langle C_3^{(n)}(x,x'x'')\rangle+\sum_{(3)}\langle\Delta C_2^{(n)}(x,x')\Delta\rho_n(x'')\rangle+\langle\Delta\rho_n(x)\Delta\rho_n(x')\Delta\rho_n(x'')\rangle\ , \tag{1.7}$$

$$C_4(x,x',x'',x''')=\langle C_4^{(n)}(x,x',x'',x''')\rangle+\sum_{(4)}\langle \Delta C_3^{(n)}(x,x',x'')\Delta\rho_n(x''')\rangle$$
$$+\sum_{(3)}\langle \Delta C_2^{(n)}(x,x')\Delta C_2^{(n)}(x'',x''')\rangle+\sum_{(6)}\langle C_2^{(n)}(x,x')\Delta\rho_n(x'')\Delta\rho_n(x''')\rangle$$
$$+\langle \Delta\rho_n(x)\Delta\rho_n(x')\Delta\rho_n(x'')\Delta\rho_n(x''')\rangle-\sum_{(3)}\langle\Delta\rho_n(x)\Delta\rho_n(x')\rangle\langle\Delta\rho_n(x'')\Delta\rho_n(x''')\rangle \ . \qquad (1.8)$$

In Eqs. (1.7) and (1.8), the fluctuations $\Delta C_p^{(n)}$ are referred to the mean:

$$\Delta C_2^{(n)}=C_2^{(n)}-\langle C_2^{(n)}\rangle \ ,$$
$$\Delta C_3^{(n)}=C_3^{(n)}-\langle C_3^{(n)}\rangle \qquad (1.9)$$
$$\vdots$$

The final terms of Eqs. (1.8) will be recognized as the cumulant moments of the variable $\Delta\rho_n$.

II. DEFINITIONS OF CORRELATION FUNCTIONS

It is convenient to define a set of δ-function density operators, $\hat\rho_1(x)$, $\hat\rho_2(x,x')$, $\hat\rho_3(x,x',x'')$, ... following, for example, Klimontovich.[7] We shall consider *all particles to belong to one species, e.g.,* the species of charged particles (without regard to whether they are plus or minus), the species of any galaxy (without regard to whether they are spiral, spherical, etc.). Note that we avoided the word "identical," which carries quantum-mechanical overtones, unnecessary for the present job of connecting correlation integrals to the counting of particles. Reference 7 deals with the complications of populations composed of distinct species (for example, if we do distinguish plus from minus in a charged population). For clarity we defer these issues to another paper in preparation. Some applications to charged particle correlations can be found in Ref. 8.

For notational purposes, let x_a denote points of observation and y_i $(i=1,2,\ldots,n)$ the positions of the particles in the n-particle system. The density operator $\hat\rho_1$ is

$$\hat\rho_1^{(n)}(x,y_1,\ldots,y_n)=\sum_{i=1}^{n}\delta(x-y_i) \ . \qquad (2.1)$$

Equation (2.1) provides an easy way to count particles inside a volume Ω in x space:

$$\int_\Omega dx\,\hat\rho_1^{(n)}(x,\mathbf{y})=n(\Omega) \ . \qquad (2.2)$$

Equation (2.1), of course, corresponds to a specific realization of the n-particle system. By using the probability $Q_n(\mathbf{y})$ we get the average single-particle density

$$\rho_1^{(n)}(x)=\int d\mathbf{y}\,Q_n(\mathbf{y})\hat\rho_1^{(n)}(x,\mathbf{y}) \ . \qquad (2.3)$$

Finally, the total single-particle density is

$$\rho_1(x)=\sum_n q_n\rho_1^{(n)}(x) \qquad (2.4)$$

so that

$$\int_\Omega dx\,\rho_1(x)=\sum_n nq_n=\langle n\rangle \ . \qquad (2.5)$$

Proceeding to higher-order density correlation operators we write[7]

$$\hat\rho_2^{(n)}(x_1,x_2;\mathbf{y})=\sum_{i,j=1}^{n}{}'\,\delta(x_1-y_i)\delta(x_2-y_j) \ , \qquad (2.6)$$

$$\hat\rho_3^{(n)}(x_1,x_2,x_3;\mathbf{y})=\sum_{i,j,k=1}^{n}{}'\,\delta(x_1-y_i)\delta(x_2-y_j)\times\delta(x_3-y_k) \ , \qquad (2.7)$$

etc., where the prime indicates that $i\neq j$, $i\neq j\neq k\neq i$, etc. The necessity of this exclusion in constructing sensible distribution functions for a single species is explained in detail in Ref. 7.

If we integrate $\hat\rho_2^{(n)},\hat\rho_3^{(n)},\ldots,\hat\rho_p^{(n)}$ over identical ranges of the x_a variables, we find

$$\int_\Omega dx_2\int_\Omega dx_2\hat\rho_2^{(n)}(x_1,x_2;\mathbf{y})=n(\Omega)[n(\Omega)-1] \ ,$$
$$\int_\Omega dx_1\int_\Omega dx_2\int_\Omega dx_3\hat\rho_3^{(n)}(x_1,x_2,x_3;\mathbf{y})=n(\Omega)[n(\Omega)-1][n(\Omega)-2] \ , \qquad (2.8)$$
$$\int_\Omega\prod dx_j\hat\rho_p^{(n)}(x_1,\ldots,x_p;\mathbf{y})=n(\Omega)[n(\Omega)-1]\cdots[n(\Omega)-p+1] \ .$$

The correlation *functions* for fixed n and $\rho_m^{(n)}(x_1,\ldots,x_m)$ are

$$\rho_2^{(n)}(x_1,x_2)\equiv\int d\mathbf{y}\,Q_n(\mathbf{y})\hat\rho_2^{(n)}(x_1,x_2;\mathbf{y}) \ ,$$
$$\rho_3^{(n)}(x_1,x_2,x_3)\equiv\int d\mathbf{y}\,Q_n(\mathbf{y})\hat\rho_3^{(n)}(x_1,x_2,x_3;\mathbf{y}) \qquad (2.9)$$
$$\vdots$$

and the total correlation functions are

$$\rho_2(x_1,x_2)=\sum_n q_n \rho_2^{(n)}(x_1,x_2)\ ,$$
$$\rho_3(x_1,x_2,x_3)=\sum_n q_n \rho_3^{(n)}(x_1,x_2,x_3) \tag{2.10}$$
$$\vdots$$

Their integrals [see Eq. (2.8)] yield factorial moments

$$\int_\Omega dx_1 \int_\Omega dx_2 \rho_2(x_1,x_2)=\langle n(n-1)\rangle_\Omega\ ,$$
$$\int_\Omega \prod_{i=1}^{p} dx_i \rho_p(x_1,\ldots,x_p)=\langle n(n-1)\cdots(n-p+1)\rangle_\Omega\ . \tag{2.11}$$

As shown in Eqs. (1.3) and (1.4), one needs to subtract away uncorrelated background densities to exhibit physically interesting correlations, be they dynamical (due to forces) or statistical (e.g., due to Bose-Einstein correlations). If $\rho_2(x_1,x_2)$ factorizes into the product $\rho_1(x_1)\rho_1(x_2)$ we say that x_1 and x_2 are statistically independent, and C_2 vanishes.

The general procedure for removing lower-order correlations is to construct cumulant correlations (as we shall see, they should be called factorial cumulant correlations in our case) by the well-known construction[9,10]

$$C_2(x_1,x_2)=\rho_2(x_1,x_2)-\rho_1(x_1)\rho_1(x_2)\ ,$$
$$C_3(x_1,x_2,x_3)=\rho_3(x_1,x_2,x_3)-\sum_{(3)}\rho_2(x_1,x_2)\rho_1(x_3)+2\rho_1(x_1)\rho_1(x_2)\rho_1(x_3)\ , \tag{2.12}$$
$$C_4(x_1,x_2,x_3,x_4)=\rho_4(x_1,x_2,x_3,x_4)-\sum_{(4)}\rho_3(x_1,x_2,x_3)\rho_1(x_4)-\sum_{(3)}\rho_2(x_1,x_2)\rho_2(x_3,x_4)$$
$$+2\sum_{(6)}\rho_2(x_1,x_2)\rho_1(x_3)\rho_1(x_4)-6\rho_1(x_1)\rho_1(x_2)\rho_1(x_3)\rho_1(x_4)\ .$$

In the case of fixed n each symbol is given a superscript (n). In Sec. III we will verify the result that the C_p vanish if any variable becomes statistically independent of the others. Hence nonvanishing cumulants imply true statisitical dependence. Inverting Eqs. (2.12) shows how ρ_p is composed of lower-order cumulant correlations

$$\rho_2(x_1,x_2)=\rho_1(x_1)\rho_1(x_2)+C_2(x_1,x_2)\ ,$$
$$\rho_3(x_1,x_2,x_3)=\rho_1(x_1)\rho_1(x_2)\rho_1(x_3)+\sum_{(3)}C_2(x_1,x_2)\rho_1(x_3)+C_3(x_1,x_2,x_3)\ ,$$
$$\rho_4(x_1,x_2,x_3,x_4)=\rho_1(x_1)\rho_1(x_2)\rho_1(x_3)\rho_1(x_4)+\sum_{(4)}C_3(x_1,x_2,x_3)\rho_1(x_4) \tag{2.13}$$
$$+\sum_{(3)}C_2(x_1,x_2)C_2(x_3,x_4)+\sum_{(6)}C_2(x_1,x_2)\rho_1(x_3)\rho_1(x_4)+C_4(x_1,x_2,x_3,x_4)\ .$$

In Sec. III we will construct the specific generating functionals which automatically produce these structures.

We see [Eq. (2.11)] that the integrated density correlations produce factorial moments. Correspondingly integration of C_p gives factorial cumulants

$$\int_\Omega dx_1 \int_\Omega dx_2 C_2(x_1,x_2)=\langle n(n-1)\rangle_\Omega-\langle n\rangle_\Omega^2\ ,$$
$$\int_\Omega dx_1 \int_\Omega dx_2 \int_\Omega dx_3 C_3(x_1,x_2,x_3)=\langle n(n-1)(n-2)\rangle-3\langle n(n-1)\rangle_\Omega\langle n\rangle_\Omega+2\langle n\rangle_\Omega^3 \tag{2.14}$$
$$\vdots$$

Denoting the factorial moments (cumulants) by $\xi_r(f_r)$ recall that the generating functions are[11]

$$Q(\lambda)=\sum_r (1-\lambda)^r P_r=\sum_r(-\lambda)^r \xi_r/r!\ ,$$
$$\ln Q(\lambda)=\sum_r(-\lambda)^r f_r/r!\ . \tag{2.15}$$

Note that a vanishing cumulant C_r (i.e., statistical independence) implies the vanishing of the corresponding factorial cumulant f_r. Also note that for Poissonian statistics all f_r $(r\geq 2)$ vanish, since $Q(\lambda)=e^{-\lambda\langle n\rangle}$. This result establishes a link (for our definitions appropriate for one species of particle) between statistical independence of the correlation functions and Poissonian counting statistics. This should be contrasted with the usual cumulants for Gaussian random variables: in this case all cumulants beyond the second one vanish.

III. DECOMPOSITION OF CUMULANT MOMENTS: GENERATING FUNCTIONAL TECHNIQUES

A straightforward application of generating functional methods allows the derivation of Eqs. (1.7)–(1.9). Define $Z_n[\lambda]$ by

$$Z_n[\lambda] \equiv \int d\mathbf{y}\, Q_n(\mathbf{y}) \exp\left[i \int dx\, \lambda(x) \hat{\rho}_1^{(n)}(x,\mathbf{y}) \right] . \tag{3.1}$$

Change of notation. To save space we no longer exhibit the coordinates $x_1, x_2, \ldots$ but simply write $1,2,\ldots$. For example, $\lambda(x_1)\to\lambda_1$, $C_2(x_1,x_2)\to C_2(1,2)$, etc.

The density correlations follow from

$$\left.\frac{\delta^p Z_n[\lambda]}{\delta\lambda_1,\ldots,\delta\lambda_p}\right|_{\lambda=0} = i^p \langle \hat{\rho}^{(n)}(x_1) \times \hat{\rho}^{(n)}(x_2),\ldots,\hat{\rho}^{(n)}(x_p)\rangle \equiv i^p \rho_p^{(n)}(x_1,\ldots,x_p) . \tag{3.2}$$

Likewise $\ln Z_n$ generates the cumulant moments:

$$\left.\frac{\delta^p \ln Z_n[\lambda]}{\delta\lambda_1,\ldots,\delta\lambda_p}\right|_{\lambda=0} = i^p C_p^{(n)}(x_1,x_2,\ldots,x_p) . \tag{3.3}$$

If we now average over n, we find

$$Z[\lambda] = \sum_n q_n Z_n[\lambda] \tag{3.4}$$

which generates density correlations $\rho_p(x_1,\ldots,x_p)$, and

$$\ln Z[\lambda] \tag{3.5}$$

which generates cumulants C_p in terms of the *total* ρ_p. The reader should work out a few examples to confirm this.

Since $\ln Z$ generates complete cumulants and $\ln Z_n$ fixed n cumulants, we can find the relation between the two sets beginning with

$$\frac{\delta \ln Z}{\delta\lambda_1} = \sum_n q_n \frac{1}{Z}\frac{\delta Z_n}{\delta\lambda_1} \equiv \sum_n q_n \xi_n \frac{\delta \ln Z_n}{\delta\lambda_1} , \tag{3.6}$$

$$\xi_n \equiv Z_n/Z . \tag{3.7}$$

As it turns out, $\ln\xi_n$ is an important generating functional in its own right. Note that

$$\frac{\delta \ln\xi_n}{\delta\lambda_1} = \frac{\delta \ln Z_n}{\delta\lambda_1} - \frac{\delta \ln Z}{\delta\lambda_1} \underset{\lambda\to 0}{\to} i[\rho_1^{(n)}(1) - \rho_1(1)] \equiv i\Delta\rho_n(1) , \tag{3.8}$$

$$\frac{\delta^2 \ln\xi_n}{\delta\lambda_1\delta\lambda_2} = \frac{\delta^2 \ln Z_n}{\delta\lambda_1\delta\lambda_2} - \frac{\delta^2 \ln Z}{\delta\lambda_1\delta\lambda_2} \underset{\lambda\to 0}{\to} i^2[C_2^{(n)}(1,2) - C_2(1,2)] , \tag{3.9}$$

$$\frac{\delta^m \ln\xi_n}{\delta\lambda_1\cdots\delta\lambda_m} \to C_m^{(n)}(1,2,\ldots,m) - C_m(1,2,\ldots,m) .$$

From Eq. (3.6) we now derive (writing $\delta\xi_n/\delta\lambda = \xi_n \delta\ln\xi_n/\delta\lambda$)

$$\frac{\delta^2 \ln Z}{\delta\lambda_1\delta\lambda_2} = \sum_n q_n \left[\xi_n \frac{\delta^2 \ln Z_n}{\delta\lambda_1\delta\lambda_2} + \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_2}\frac{\delta \ln Z_n}{\delta\lambda_1} \right] . \tag{3.10}$$

Note that we can replace Z_n by $\xi_n = Z_n/Z$ in the second term of Eq. (3.10) since $\delta/\delta\lambda_2 \sum_n q_n \xi_n) = \delta/\delta\lambda_2(1) = 0$.

Change of notation. In order to save writing we denote the average over q_n by brackets $\langle\ \rangle$. Hence Eq. (3.10) becomes

$$\frac{\delta^2 \ln Z}{\delta\lambda_1\delta\lambda_2} = \left\langle \xi_n \frac{\delta^2 \ln Z_n}{\delta\lambda_1\delta\lambda_2} \right\rangle + \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta \ln\xi_n}{\delta\lambda_2} \right\rangle . \tag{3.11}$$

From Eq. (3.11) we now derive Eq. (1.7) as $\lambda_i \to 0$:

$$C_2(1,2) = \left\langle C_2^{(n)}(1,2) \right\rangle + \left\langle \Delta\rho_n(1)\Delta\rho_n(2) \right\rangle . \tag{3.12}$$

It is now straightforward to evaluate higher-order terms; in third order we find [see (3.11)]

$$\frac{\delta^3 \ln Z}{\delta\lambda_1\delta\lambda_2\delta\lambda_3} = \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta \ln\xi_n}{\delta\lambda_2}\frac{\delta \ln\xi_m}{\delta\lambda_3} \right\rangle + \sum_{(3)} \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta^2 \ln\xi_n}{\delta\lambda_2\delta\lambda_3} \right\rangle + \left\langle \xi_n \frac{\delta^3 \ln Z_n}{\delta\lambda_1\delta\lambda_2\delta\lambda_3} \right\rangle \tag{3.13}$$

where the (3) denotes three terms in the summation over permutated labels. Setting $\lambda_i = 0$ in Eq. (3.13), we derive the result

$$C_3(1,2,3) = \langle C_3^{(n)}(1,2,3) \rangle + \sum_{(3)} \langle \Delta C_2^{(n)}(1,2)\Delta\rho_n(3) \rangle + \langle \Delta\rho_n(1)\Delta\rho_n(2)\Delta\rho_n(3) \rangle ; \tag{3.14}$$

$$\Delta C_2^{(n)}(1,2) = C_2^{(n)}(1,2) - \langle C_2^{(n)}(1,2) \rangle . \tag{3.15}$$

The emerging pattern by which the full cumulant C_p is expressed as fixed n quantities and deviations from the averages of lower densities and correlations becomes still more evident in fourth order. From the relation

$$\frac{\delta^4 \ln Z}{\delta\lambda_1\delta\lambda_2\delta\lambda_3\delta\lambda_4} = \left\langle \xi_n \frac{\delta^4 \ln Z_n}{\delta\lambda_1\delta\lambda_2\delta\lambda_3\delta\lambda_4} \right\rangle + \sum_{(4)} \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta^3 \ln\xi_n}{\delta\lambda_2\delta\lambda_3\delta\lambda_4} \right\rangle + \sum_{(3)} \left\langle \xi_n \frac{\delta^2 \ln\xi_n}{\delta\lambda_1\delta\lambda_2}\frac{\delta^2 \ln\xi_n}{\delta\lambda_3\delta\lambda_4} \right\rangle + \sum_{(6)} \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta \ln\xi_n}{\delta\lambda_2}\frac{\delta^2 \ln\xi_n}{\delta\lambda_3\delta\lambda_4} \right\rangle + \left\langle \xi_n \frac{\delta \ln\xi_n}{\delta\lambda_1}\frac{\delta \ln\xi_n}{\delta\lambda_2}\frac{\delta \ln\xi_n}{\delta\lambda_3}\frac{\delta \ln\xi_n}{\delta\lambda_4} \right\rangle , \tag{3.16}$$

taking $\lambda_i \to 0$ and using Eq. (3.9), we find the expression

$$C_4(1,2,3,4)=\langle C_4^{(n)}(1,2,3,4)\rangle+\sum_{(4)}\langle \Delta\rho_n(1)[C_3^{(n)}(2,3,4)-C_3(2,3,4)\rangle$$
$$+\sum_{(3)}\langle [C_2^{(n)}(1,2)-C_2(1,2)][C_2^{(n)}(3,4)-C_2(3,4)]\rangle$$
$$+\sum_{(6)}\langle \Delta\rho_n(1)\Delta\rho_n(2)[C_2^{(n)}(3,4)-C_2(3,4)]\rangle+\langle \Delta\rho_n(1)\Delta\rho_n(2)\Delta\rho_n(3)\Delta\rho_n(4)\rangle \;. \tag{3.17}$$

Now using Eqs. (3.12), (3.14), (3.15), and (1.9) to eliminate C_2 and C_3 from Eq. (3.17) we find

$$C_4(1,2,3,4)=\langle C_4^{(n)}(1,2,3,4)+\sum_{(4)}\langle \Delta C_3^{(n)}(1,2,3)\Delta\rho_n(4)\rangle+\sum_{(3)}\langle \Delta C_2^{(n)}(1,2)\Delta C_2^{(n)}(3,4)\rangle$$
$$+\sum_{(6)}\langle \Delta C_2^{(n)}(1,2)\Delta\rho_n(3)\Delta\rho_n(4)\rangle+\langle \Delta\rho_n(1)\Delta\rho_n(2)\Delta\rho_n(3)\Delta\rho_n(4)\rangle$$
$$-\sum_{(3)}\langle \Delta\rho(1)\Delta\rho(2)\rangle\langle \Delta\rho_n(3)\Delta\rho_n(4)\rangle \;. \tag{3.18}$$

Now one can write down higher-order expansions by inspection. Note that the last two terms are just the cumulants formed from the variable $\Delta\rho_n$. To sharpen the meaning of these expansions, suppose that the first q density moments were to factorize (statistical independence in the n-particle sector)

$$\begin{aligned}
&\rho_2^{(n)}(1,2)=\rho_1^{(n)}(1)\rho_1^{(n)}(2)\;,\\
&\rho_3^{(n)}(1,2,3)=\rho_1^{(n)}(1)\rho_1^{(n)}(2)\rho_1^{(n)}(3)\;,\\
&\vdots\\
&\rho_q^{(n)}(1,2,3,\ldots,q)=\rho_1^{(n)}(1)\rho_1^{(n)}(2)\cdots\rho_q^{(n)}(q)\;.
\end{aligned} \tag{3.19}$$

In this case the generating function factorizes and $C_p^{(n)}=0$ for $p=2,3,\ldots,q$, i.e., there are no cross terms. The surviving terms in the $\Delta\rho$ cumulants are [$\mathcal{P}$ indicates the distinct permutations of the pairs (12) (34)]

$$\begin{aligned}
&\tilde{C}_2(\Delta\rho)=\langle \Delta\rho_n(1)\Delta\rho_n(2)\rangle=\langle \rho_1^{(n)}(1)\rho_1^{(n)}(2)\rangle-\rho_1(1)\rho_1(2)\;,\\
&\tilde{C}_3(\Delta\rho)=\langle \Delta\rho_n(1)\Delta\rho_n(2)\Delta\rho_n(3)\rangle\;,\\
&\tilde{C}^4(\Delta\rho)=\langle \Delta\rho_n(1)\Delta\rho_n(2)\Delta\rho_n(3)\Delta\rho_n(4)\rangle-\sum_{\mathcal{P}}\langle \Delta\rho_n(1)\Delta\rho_n(2)\rangle\langle \Delta\rho_n(3)\Delta\rho_n(4)\rangle\;.
\end{aligned} \tag{3.20}$$

The generating functional for the $\tilde{C}_4$ is $\ln\sum_n q_n \exp[i\int dx\,\lambda(x)\Delta\rho_n(x)]$.

The lesson of Eqs. (3.19) and (3.20) is that *even in the absence of true correlations in the n-particle sector, the mixing of different n inherent in the definition of the C_q induces a nonzero, often large component.* In such cases, the physical significance of the correlations is nearly trivial, following from the single-particle density fluctuation averaged over the number probability q_n.

IV. INTEGRAL PROPERTIES; RELATION TO MOMENTS

In Sec. II we observed the relation between factorial moments and certain integrals over the variables appearing in the correlation functions. (By choosing more general integration domains, quite a variety of correlations can be obtained.[5]) Here we show that knowledge of the moments puts strong constraints on the relative magnitude of the contributions of the individual terms on the right-hand sides of Eqs. (1.6)–(1.8). In fact, the latter will be seen to be, in essence, a local version of an expansion of factorial cumulant moments in terms of cumulant moments.[10]

Our exposition depends on the identities

$$\int_\Omega \prod_{i=1}^{p} dx_i C_p^{(n)}(x_1,\ldots,x_p)=(-1)^{p-1}(p-1)!n\;, \tag{4.1}$$

$$\int \prod_{i=1}^{p} dx_i \tilde{C}_p(x_1,\ldots,x_p)=K_p \tag{4.2}$$

where K_p is the ordinary cumulant moment of order p [see Eq.(3.20)]. The ordinary moments μ_p and cumulant moments K_p are related by the generating functions

$$\begin{aligned}
&M(\lambda)\equiv\sum_n e^{-\lambda n}P_n=\sum_{n=1}^{\infty}(-\lambda)^n\mu_n/n!\;,\\
&\ln M(\lambda)\equiv\sum_{n=1}^{\infty}(-\lambda)^n K_n/n!\;.
\end{aligned} \tag{4.3}$$

These definitions are to be constrasted with their factorial counterparts defined in Eq. (2.15). If we take $P_n=\delta_{nn_0}$ the averaged factorial cumulant moment reduces to Eq. (3.1); we write

$$\begin{aligned}
\ln\sum_n(1-\lambda)^n P_n&=\ln(1-\lambda)^{n_0}\\
&=n_0\ln(1-\lambda)=-n_0\sum_{p=1}^{\infty}\frac{\lambda^p}{p}\\
&\equiv\sum_{p=1}^{\infty}(-\lambda)^p f_p/p!\;.
\end{aligned} \tag{4.4}$$

These relations immediately lead to Eq. (4.1). A corollary of Eq. (4.1) follows:

$$\int \prod_{i=1}^{p} dx_i \Delta C_p^{(n)}(x_1,\ldots,x_p) = (-1)^{p-1}(p-1)!(n-\langle n\rangle) . \quad (4.5)$$

Integration of Eqs. (1.6)–(1.8) leads to [using Eqs. (4.1) and (4.2)]

$$\begin{aligned} f_2 &= \langle n(n-1)\rangle - \langle n\rangle^2 = -\langle n\rangle + K_2 ,\\ f_3 &= \langle n(n-1)(n-2)\rangle - 3\langle n(n-1)\rangle\langle n\rangle + 2\langle n\rangle^3 \\ &= 2\langle n\rangle - 3K_2 + K_3 ,\\ f_4 &= \langle n(n-1)(n-2)(n-3)\rangle \\ &\quad -4\langle n(n-1)(n-2)\rangle\langle n\rangle - 3\langle (n-\langle n\rangle)^2\rangle \\ &\quad +2\langle (n-\langle n\rangle)^2\rangle\langle n\rangle^2 - 6\langle n^4\rangle \\ &= -6\langle n\rangle + 4\times 2K_2 - K_3 + K_4 . \end{aligned} \quad (4.6)$$

Here the cumulant moments are

$$\begin{aligned} K_2 &= \langle (n-\langle n\rangle)^2\rangle = \langle n^2\rangle - \langle n\rangle^2 ,\\ K_3 &= \langle (n-\langle n\rangle)^3\rangle = \langle n^3\rangle - 3\langle n^2\rangle\langle n\rangle + 2\langle n\rangle^2 , \\ K_4 &= \langle (n-\langle n\rangle)^4\rangle - 3\langle (n-\langle n\rangle)^2\rangle^2 . \end{aligned} \quad (4.7)$$

Although the correlation functions have dynamical significance, they are subject to the normalization conditions, Eqs. (4.1) and (4.2).

Before analyzing specific examples we reexpress Eqs. (4.5) in terms of "reduced" moments $f_p/\langle n\rangle^p$ and $\gamma_p = K_p/\langle n\rangle^p$:

$$\begin{aligned} \frac{f_2}{\langle n\rangle^2} &= -\frac{1}{\langle n\rangle} + \gamma_2 ,\\ \frac{f_3}{\langle n\rangle^3} &= \frac{2}{\langle n\rangle^2} - \frac{3}{\langle n\rangle}\gamma_2 + \gamma_3 , \\ \frac{f_4}{\langle n\rangle^4} &= -\frac{6}{\langle n\rangle^3} + \frac{11}{\langle n\rangle^2}\gamma_2 - \frac{6}{\langle n\rangle}\gamma_3 + \gamma_4 . \end{aligned} \quad (4.8)$$

For convenience we have defined the reduced cumulants $K_p/\langle n\rangle^p$ as γ_p:

$$\begin{aligned} \gamma_1 &= 1 ,\\ \gamma_2 &= \frac{\langle n^2\rangle - \langle n\rangle^2}{\langle n\rangle^2} ,\\ \gamma_3 &= \frac{\langle n^3\rangle - 3\langle n^2\rangle\langle n\rangle + 2\langle n\rangle^3}{\langle n\rangle^3} , \\ \gamma_4 &= \frac{\langle n^4\rangle - 4\langle n\rangle^3\langle n\rangle - 3\langle n^2\rangle^2 + 12\langle n^2\rangle^2 - 6\langle n\rangle^4}{\langle n\rangle^4} . \end{aligned} \quad (4.9)$$

Now we see that Eqs. (1.6)–(1.8) lead *directly* to the expression of the normalized factorial cumulants $f_p/\langle n\rangle^p$ as an expansion in normalized cumulant moments γ_p.

The manner in which Eqs. (4.8) are satisfied depends completely on the nature of the underlying probability distribution P_n. If P_n is Poisson, $M(\lambda)=e^{-\lambda\langle n\rangle}$ and all f_p in Eq. (4.8) vanish. Hence the only role of the right-hand side is to cancel. On the other hand, if the γ_p are of order unity, as for a broad distribution, then there is an ordering of numerical significance expressed in Eqs. (4.8).

V. APPLICATIONS

We have seen that the expansion of the correlation function in terms of $C_p^{(n)}$ and the lower-order fluctuations $\Delta C_q^{(n)}$ has a very simple structure. The corresponding moment expressions relate the factorial cumulants $\gamma_p(\Omega)$. In fact, each term [see, for example, (3.18)] integrates to a particular cumulant moment.

The sum rule (4.1) provides a quantitative constraint on the relative magnitude of the various terms in the expansion in terms of averages over q_n. In many cases the results are nonintuitive. One might expect that the pth-order correlation would be dominated by $C_p^{(n)}$. For "broad" distributions, the opposite is the case, as we shall prove in a specific example, the negative binominal distribution. For such *broad distributions, in the limit of large* $\langle n\rangle$, *the dominating terms in the cumulant are the single-particle* $\Delta\rho$ *cumulants (3.20). Hence knowledge of the global* $P_n(\Omega)$ *and* $\Delta\rho_n=\rho_n-\rho_1$ *determines the basic form of* C_p *in suitable limits.*

Let us analyze the case of negative binomial distribution, which plays a useful role in describing final state hadronic multiplicity data.[11–13] For this system the usual variable upon which the correlation function depends is the longitudinal rapidity variable $y=\frac{1}{2}\ln(E+p_2)/(E-p_z)$, with $E=(\mathbf{p}^2+m^2)^{1/2}$ and p_z the longitudinal (i.e., parallel to the c.m. collision axis) momentum. y is additive under a change of (longitudinal) reference frames; the phase space interval $d^3p=dyd^2\mathbf{p}_\perp$ for each particle, $\mathbf{p}_\perp$ being the transverse momentum, which at high energies is typically (except for scarce jets) confined to a few hundred MeV/c in magnitude.

It is well established that for a wide range of rapidity acceptance ΔY ($\rightarrow\Omega$ of the general discussion of this paper) and reaction types ($e^+-e^-,\mu p, pp, \pi p, \cdots$) the probability P_n of finding n charges in ΔY agrees (except for the 900-eV UA5 Collaboration data[14]) with the negative binomial distribution (NBD)

$$P_n^k = \frac{\Gamma(n+k)}{\Gamma(n)\Gamma(k)}\frac{(\bar{n}/k)^n}{(1+\bar{n}/k)^{n+k}} . \quad (5.1)$$

The cell parameter k can be any real number, depending on the particular dynamical context in which it arises.[11] The earliest physics derivation seems to be that of Planck,[15] who composed k Bose-Einstein distributions of equal average occupancy $\bar{n}/k$. [For $k=1$, (5.1) reduces to the geometric distribution $\bar{n}^n(1+\bar{n})^{n+1}$, which coincides with a thermalized Bose-Einstein distribution when $1/\bar{n}=\exp(E/kT)-1$.] Equation (5.1) has a nice scaling form for large n, $\langle n\rangle$. If we set $x=n/\langle n\rangle$ and let $n\rightarrow\infty$, k and x fixed, we get

$$\langle n\rangle P_n^k \rightarrow \frac{k^k}{(k-1)!}x^{k-1}e^{-kx} . \quad (5.2)$$

Although the differences between (5.1) and (5.2) can be seen in careful data analysis, the form (5.2) is very useful. We note that these results are already known from semiclassical photocount theory of Gaussian random fields from Mandel's 1959 paper.[16] In the photon case k would be the number of independent polarizations, or laser modes.

For hadronic multiparticle phenomenology the physical meaning of k is not understood. The UA5 Collaboration (Ref. 14) has found that k varies with the squared c.m. energy $s=W^2$ as $k^{-1}=A+B\ln s$ for given ΔY. Thus the scaling law (5.2) is violated at sufficiently high energy. In addition, k decreases. Typical values are $k\approx 3$, $\langle n\rangle\approx 30$ for c.m. energy of 540 GeV, $\Delta Y\approx 10$. For small ΔY, k decreases, e.g., for $\Delta Y\approx 1$, $k\approx 1.7$. At lower energies, characteristic of earlier Fermilab and CERN Intersecting Storage Ring experiments, $k\approx\langle n\rangle\approx 10$.

The factorial moment generating function [see Eq. (2.15)] for (5.1) is

$$Q(\lambda)=\left[1+\frac{\lambda\bar{n}}{k}\right]^{-k} \tag{5.3}$$

which is just the product of k independent Bose-Einstein sources with average population $\bar{n}/k$. The factorial cumulants f_p defined by

$$\ln Q=\frac{\sum_p(-\lambda)^p}{p!}f_p \tag{5.4}$$

lead to the NBD value

$$\frac{f_p}{\langle n\rangle^p}=\frac{(p-1)!}{k^{p-1}} . \tag{5.5}$$

In the limit $k\to\infty$, $Q(\lambda)\to e^{-\lambda\langle n\rangle}$ approaches the Poisson limit, and $f_p\to 0$, $p>1$, as noticed earlier.

The first several ordinary reduced cumulant moments are

$$\begin{aligned}\gamma_2&=\frac{K_2}{\langle n\rangle^2}=\frac{1}{k}+\frac{1}{\langle n\rangle} ,\\ \gamma_3&=\frac{K_3}{\langle n\rangle^3}=\frac{2}{k^2}+\frac{3}{k\langle n\rangle}+\frac{1}{\langle n\rangle^2} ,\\ \gamma_4&=\frac{K_4}{\langle n\rangle^4}=\frac{6}{k^3}+\frac{7}{k^2\langle n\rangle}+\frac{12}{k\langle n\rangle^2}+\frac{1}{\langle n\rangle^3} .\end{aligned} \tag{5.6}$$

Clearly, when $\langle n\rangle/k\gg 1$ (5.6) reduces to the factorial cumulant values (5.5). From (4.8) we see that this implies that the "true" n-particle correlations contribute negligibly to the integrated cumulant.

The identities of Eq. (4.8) are, for the NBD,

$$\begin{aligned}\frac{1}{k}&=-\frac{1}{\langle n\rangle}+\gamma_2 ,\\ \frac{2}{k^2}&=\frac{2}{\langle n\rangle^2}-\frac{3}{\langle n\rangle}\gamma_2+\gamma_3 ,\\ \frac{6}{k^3}&=-\frac{6}{\langle n\rangle^3}+\frac{11}{\langle n\rangle^2}\gamma_2-\frac{6}{\langle n\rangle}\gamma_3+\gamma_4 .\end{aligned} \tag{5.7}$$

For $\langle n\rangle\gg k$, $f_p/\langle n\rangle^p\simeq\gamma_p$ and the difference terms are of $O(1/\langle n\rangle)$. Hence in this domain little can be learned about correlations from mixed-number data, which are dominated by fluctuations in the single-particle density. In particular, we note Eqs. (3.19) and (3.20) give a substantial total correlation even when fixed n correlations vanish. We also note the relation of k to the second-order cumulant

$$\frac{f_2}{\langle n\rangle^2}=\int_{\Delta_Y}dy_1dy_2\frac{C_2(y_1,y_2)}{\langle n\rangle^2}=\frac{1}{k} . \tag{5.8}$$

To summarize, for high-energy and nuclear applications, high-energy, large multiplicity, large ΔY acceptance tells us little about true fixed n correlations. At lower energy, all terms contribute and a careful analysis is required to disentangle the components.

The distribution of galaxy counts also has a broad distribution. Hubble[17] found a log-normal distribution for counts in $6^0\times 6^0$ photographic plates. We found[18] that conditional counts (probabilities in Zwicky clusters) obeyed a negative binomial or the gamma distribution of Eq. (5.2) with $\langle n\rangle\approx 100$, $k\approx 6$. Saslaw and Hamilton have recommended[19] a different but similar distribution for galaxy counts.

The main results of this paper are based on combinatoric identities. While trivial in this physical sense, the identities have powerful implications. Clearly, to see nontrivial effects the observation volume must be chosen small enough that the average particle number is not too large compared to some measure of the width of the distribution (k for the NBD).

There exist many aspects of higher-order correlations that invite further investigation. For example, classical systems have phase space distribution functions coupled by the Bogoliubov-Born-Green-Kirkwood-Yvon hierarchy.[9,20] What does the present analysis imply for this problem? Secondly, in the standard statistical mechanics we can construct the cluster expansions of Ursell, Mayer, and others. For both short- and long-range forces[21] there is a systematic procedure to expand all physical quantities in terms of two-particle potentials, correlation functions, etc. Presently such simplicity does not exist for the example systems mentioned above, hadrons and galaxies.

Nevertheless, an appealing conjecture[1-4] about the form of the higher cumulant correlations has been rather successful both for galaxy distributions and multihadron distributions. In astronomy one speaks of the "hierarchical model" and in hadronics[4] of the linked pair approximation (LPA). The idea is simply to compose the pth cumulant as a (symmetrized) product of $p-1$ linked (no closed loops) two-particle cumulants. Although it was our interest in the LPA that drew our attention to the analysis of the present paper, we have so far not succeeded in connecting the two approaches. We hope to pursue such connections elsewhere.

The integral (moment) relations discussed in this paper were restricted to the specially simple case that each variable was integrated over the *same* range Ω. There are many interesting generalizations of this technique. For example, correlations between different parts of the phase space are related to suitably chosen integration domains. If we partition Ω into subspaces (e.g., M identical ones, with increasing M) it is possible to test for scaling, possibly fractal behavior from moment data rather than direct correlation functions. We refer the reader to Ref. 5 for current examples of these ideas.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. Department of Energy, Division of High Energy Physics and Division of Nuclear Physics, and the Alexander von Humboldt Foundation.

[1] P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton Unversity Press, Princeton, 1980).

[2] A. J. S. Hamilton, Astrophys. J. **332**, 67 (1988).

[3] G. Toth, J. Jollos, and A. S. Szalay, Astrophys. J. **344**, 75 (1989).

[4] P. Carruthers and I. Sarcevic, Phys. Rev. Lett. **63**, 1562 (1989).

[5] P. Carruthers, H. C. Eggers, Q. Gao, and I. Sarcevic, Int. J. Mod. Phys. A (to be published).

[6] L. Foa, Phys. Rep. C **22**, 1 (1975).

[7] Yu L. Klimontovich, *The Statistical Theory of Non-Equilibrium Processes in a Plasma* (Pergamon, Oxford, 1967).

[8] P. Carruthers, Int. J. Mod. Phys. A **4**, 5587 (1989).

[9] R. L. Liboff, *Introduction to the Theory of Kinetic Equations* (Krieger, Huntington, NY, 1979).

[10] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics,* 5th ed. (Oxford University Press, New York, 1987).

[11] P. Carruthers and C. C. Shih, Int. J. Mod. Phys. A **2**, 1447 (1987).

[12] A. Giovannini, Nuovo Cimento A **10**, 1713 (1972); **15**, 542 (1973); **24**, 421 (1974); **34**, 647 (1976).

[13] P. Carruthers and C. C. Shih, Phys. Lett. **127B**, 242 (1983).

[14] UA5 Collaboration (G. J. Alner *et al.*), Phys. Rep. **154**, 247 (1987).

[15] M. Planck, Sitzungsber. Deutsch. Akad. Wiss. Berlin Kl. Math. Phys. Tech. **33**, 335 (1932).

[16] L. Mandel, Proc. Phys. Soc. London **74**, 233 (1959).

[17] E. Hubble, Astrophys. J. **79**, 8 (1934).

[18] P. Carruthers and Minh Duong-Van, Phys. Lett. **131B**, 116 (1983).

[19] W. C. Saslaw and A. J. S. Hamilton, Astrophys. J. **276**, 13 (1984).

[20] N. Bogoliubov, in *Studies in Statistical Mechanics,* edited by J. De Boer and G. E. Uhlenbeck (North-Holland, Amsterdam, 1962), Vol. 1.

[21] R. Brout and P. Carruthers, *Lectures on the Many-Electron Problem* (Wiley-Interscience, New York, 1963).