

# Factorial Moments in Complex Systems

Laurent Schoeffel

CEA Saclay, Irfu/SPP, 91191 Gif/Yvette Cedex,  
France

## Abstract

Factorial moments are convenient tools in particle physics to characterize the multiplicity distributions when phase-space resolution ( $\Delta$ ) becomes small. They include all correlations within the system of particles and represent integral characteristics of any correlation between these particles. In this letter, we show a direct comparison between high energy physics and quantitative finance results. Both for physics and finance, we illustrate that correlations between particles lead to a broadening of the multiplicity distribution and to dynamical fluctuations when the resolution becomes small enough. From the generating function of factorial moments, we make a prediction on the gap probability for sequences of returns of positive or negative signs. The gap is defined as the number of consecutive positive returns after a negative return, thus this is a gap in negative return. Inversely for a gap in positive return. Then, the gap probability is shown to be exponentially suppressed within the gap size. We confirm this prediction with data.

# 1 Introduction

Multiplicity distributions and correlations between final-state particles in nuclear interactions are an important testing ground for analytic perturbative theory, as well as for Monte-Carlo (MC) models describing the hadronic final state [1]. Two-particle angular correlations have been extensively studied experimentally [2]. Specific statistical tools, namely the normalized factorial moments, have emerged in order to analyze in much details multiplicity distributions measured in restricted phase-space regions. The normalized factorial moments are defined as

$$F_q(\Delta) = \langle n(n-1) \dots (n-q+1) \rangle / \langle n \rangle^q, \quad q = 2, 3, \dots,$$

for a specified phase-space region of size  $\Delta$ . The number,  $n$ , of particles is measured inside  $\Delta$  and angled brackets  $\langle \dots \rangle$  denote the average over all events. The factorial moments, along with cumulants [3] and bunching parameters [4], are convenient tools to characterize the multiplicity distributions when  $\Delta$  becomes small. Indeed, for uncorrelated particle production within  $\Delta$ , Poisson statistics holds and  $F_q = 1$  for all  $q$ . Correlations between particles lead to a broadening of the multiplicity distribution and to dynamical fluctuations. In this case, the normalized factorial moments increase with decreasing  $\Delta$ . This effect is frequently called intermittency [5]. As a matter of fact, it has been noticed in [5] that the use of factorial moments allows to extract the dynamical signal from the Poisson noise in the analysis of the multiplicity signal in high energy reactions.

In addition, it has been shown that it is possible to define and compute a multi-fractal dimension,  $\mathcal{D}_q$ , for the theory of strong interactions [6, 7]

$$F_q(\Delta) = \langle n(n-1) \dots (n-q+1) \rangle / \langle n \rangle^q \propto \Delta^{(q-1)(1-\mathcal{D}_q/d)} \quad (1)$$

where  $d$  is the dimension of the phase space under consideration ( $d = 2$  for the whole angular phase space, and  $d = 1$  if one has integrated over, say the azimuthal angle). In the constant coupling case  $\mathcal{D}_q$  is well defined and reads

$$\mathcal{D}_q = \gamma_0 \frac{q+1}{q} \quad (2)$$

where  $\gamma_0^2 = 4C_A\alpha_S/2\pi$ ,  $\alpha_S$  is the strong interaction coupling constant,  $C_A$  is the gluon color factor [6, 7]. The choice of the factorial moments as a specific tool in order to study the scaling behavior of the high energy multiplicity distributions is then useful to analyze the underlying dynamics of the processes. In principle, we can extend this last idea to other fields where factorial moments can be defined.

## 2 Generating Function for Factorial Moments

The multiplicity distribution is defined as  $P_n = \sigma_n / \sum_{n=0}^{\infty} \sigma_n$ , where  $\sigma_n$  is the cross section of an  $n$ -particle production process (the so-called topological cross section) and the sum is over all possible values of  $n$  so that

$$\sum_{n=0}^{\infty} P_n = 1. \quad (3)$$

The generating function can be defined as

$$G(z) = \sum_{n=0}^{\infty} P_n (1+z)^n, \quad (4)$$

which substitutes an analytic function of  $z$  in place of the set of numbers  $P_n$ . Then, we obtain the factorial moment or order  $q$  as

$$F_q = \frac{1}{\langle n \rangle^q} \cdot \left. \frac{d^q G(z)}{dz^q} \right|_{z=0}, \quad (5)$$

and the corresponding definition for cumulants

$$K_q = \frac{1}{\langle n \rangle^q} \cdot \left. \frac{d^q \ln G(z)}{dz^q} \right|_{z=0}, \quad (6)$$

The expression for  $G(z)$  can then be re-written as

$$G(z) = \sum_{q=0}^{\infty} \frac{z^q}{q!} \langle n \rangle^q F_q \quad (F_0 = F_1 = 1), \quad (7)$$

$$\ln G(z) = \sum_{q=1}^{\infty} \frac{z^q}{q!} \langle n \rangle^q K_q \quad (K_1 = 1). \quad (8)$$

The physical meaning of these moments has been discussed in the previous section. Another interpretation can be seen from the above definitions if they are presented in the form of integrals of correlation functions. Let the single symbol  $y$  represent all kinematic variables needed to specify the position of each particle in the phase space volume  $\Delta$  [8]. A sequence of inclusive  $q$ -particle differential cross sections  $d^q \sigma / dy_1 \dots dy_q$  defines the factorial moments as

$$F_q \sim \frac{1}{\langle n \rangle^q} \int_{\Delta} dy_1 \dots \int_{\Delta} dy_q \frac{d^q \sigma}{dy_1 \dots dy_q}. \quad (9)$$

Therefore, factorial moments include all correlations within the system of particles under consideration. They represent integral characteristics of any correlation between the particles whereas cumulants of rank  $q$  represent genuine  $q$ -particle correlations not reducible to the product of lower order correlations.

### 3 Experimental Applications

In Ref. [9], it has been shown that factorial moments can be applied conveniently to quantitative finance. Namely, if we divide price series  $y(t)$  in consecutive time windows of lengths  $\Delta$ , we can define a set of events. In each window, we have a certain number of positive returns  $n_+$ , where  $y(t) - y(t-1) > 0$ , and similarly of negative returns  $n_-$ . If the sequence of returns is purely uncorrelated, following a Gaussian statistics at all scales, we expect  $F_q = 1$  for all  $q$ .

As explained in previous sections, correlations between returns may lead to a broadening of the multiplicity distributions ( $n_+$  or  $n_-$  or even a combination of both) and to dynamical fluctuations. In this case, the factorial moments may increase with decreasing  $\Delta$ , or increasing the number of bins that divide  $\Delta$ . In Ref. [9], we have used the factorial moment of second order  $F_2$  for like-sign returns

$$F_2^{++} = \frac{1}{N_{events}} \sum_{events} \frac{\sum_{k=1}^{n_{bins}} \{n_k^+(n_k^+ - 1)\} / n_{bins}}{(\langle n^+ \rangle / n_{bins})^2} \quad (10)$$

where  $\langle n \rangle$  is the average number of positive returns in the full time window ( $\Delta$ ),  $n_{bins}$  denotes the number of bins in this window and  $n_k^+$  is the number of positive returns in  $k$ -th bin. Similarly, we have defined the unlike-sign returns for  $F_2$

$$F_2^{+-} = \frac{1}{N_{events}} \sum_{events} \frac{\sum_{k=1}^{n_{bins}} \{n_k^+ n_k^-\} / n_{bins}}{\langle n^+ \rangle \langle n^- \rangle / (n_{bins})^2} \quad (11)$$

with similar notations. Considering different price series, we have shown that for a small resolution in time window [9], below 4 to 8 hours, a deviation with respect to pure Gaussian fluctuations is observed from the shape of second order factorial moments (10) and (11) [9]. Results are reminded in Fig. 1.

In order to compare the strength of the intermittency obtained in financial data, it is interesting to provide a comparison with standard high energy phenomenon in particle physics [8, 10]. In what follows we discuss the results obtained with two of the most widely used  $e^+e^-$  MC event generators, JETSET [11] and HERWIG [12]. As it is done in the finance case, we define moments for like-charge and unlike-charge combinations of particles separately. Fig. 2 shows the behavior of  $F_2^{++,+-}(n_{bins})$  at a centre-of-mass energy 91.5 GeV [10]. Without entering into details, we observe the rise of both observables with  $n_{bins} \equiv b$  as in Fig. 1.

By comparing Fig. 1 and Fig. 2, we notice also that the intermittency is weak in the financial data w.r.t. high energy data, which is a reasonable observation.

## 4 Gap Probability

We can use the formalism introduced in section 2 in order to derive some further statements. From Eq. (7), we get

$$G(-1) = p_0$$

which corresponds to the probability to have zero particle in a phase space  $\Delta$  or to have zero positive (resp. negative) returns in a given time window (finance case). This defines a gap, either in rapidity for particles or in duration for positive (resp. negative) returns. Let us use Eq. (6) to express  $G(-1)$  in another way using cumulants  $K_q$

$$G(-1) = \exp(-K_1 + K_2/2! + \dots)$$

When we can neglect correlations within a large time window [9], we have  $K_2 = 0$  and we conclude

$$p_0 \simeq \exp(-K_1) = \exp(-\langle n_{+,-} \rangle) = \exp(-\rho\Delta) \quad (12)$$

where  $\Delta$  represents either the phase space or the time window. This last expression is very simple and instructive. It states that the gap probability is exponentially suppressed either in rapidity or in time. This is illustrated in Fig. 3 for the finance case. We observe the distribution of events (probability distribution) as a function of the size of the gap. The gap is defined as the number of consecutive positive returns after a negative return, thus this is a gap in negative return. Inversely for a gap in positive return. This gap is given in number of time units for the financial time series considered. In Fig. 3, we display results for the Euro future contract series over 10 years, sampled in two different time units. In both cases we observe effectively an exponential fall of the probability distribution as a function of the gap size. As illustrated also in Fig. 3, this exponential fall does not depend on the sampling. This confirms the prediction of Eq. (12). The case of particle physics is more complex, see Ref. [13].

## 5 Conclusion

In this letter, we have discussed a direct comparison between high energy physics and quantitative finance results on factorial moments analysis. Both for physics and finance, we have illustrated that correlations between particles lead to a broadening of the multiplicity distribution and to dynamical fluctuations when the resolution becomes small enough. From the generating function of factorial moments, we have shown that  $p_0 \simeq \exp(-\rho\Delta)$ . This expression states that the gap probability is exponentially suppressed within the gap size. The gap is defined as the number of consecutive positive returns after a negative return, thus this is a gap in negative return. Inversely for a gap in positive return. We have confirmed this prediction with data.

## References

- [1] E.A. De Wolf, I.M. Dremin and W. Kittel, *Phys. Rep.* 270 (1996) 1.
- [2] ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J C* 12 (2000) 53.
- [3] A.H. Mueller, *Phys. Rev. D* 4 (1971) 150.
- [4] S.V. Chekanov and V.I. Kuvshinov, *Acta Phys. Pol. B* 25 (1994) 1189; S.V. Chekanov, W. Kittel and V.I. Kuvshinov, *Z. Phys. C* 74 (1997) 517.
- [5] A. Białas and R. Peschanski, *Nucl. Phys. B* 273 (1986) 703; *Nucl. Phys. B* 308 (1988) 857.
- [6] Y.L. Dokshitzer and I.M. Dremin *Nuclear Physics B* **402** (1993) 139.
- [7] W. Ochs and J. Wosiek, *Phys. Lett. B* **289** (1992) 159, and *Phys. Lett. B* **305** (1993) 144.
- [8] I. M. Dremin, V. Arena, G. Boca, G. Gianini, S. Malvezzi, M. Merlo, S. P. Ratti, C. Riccardi *et al.*, *Phys. Lett. B* **336** (1994) 119-124; I. M. Dremin, R. C. Hwa, *Phys. Rev. D* **49** (1994) 5805-5811.

- [9] L. Schoeffel, arXiv:1108.5596.
- [10] J. Rames, arXiv:hep-ph/9411349.
- [11] T. Sjöstrand, CERN-TH.7112/93; H. U. Bengtsson and T. Sjöstrand, *Comp. Phys. Comm.* **46** (1987) 43.
- [12] G. Marchesini et al., *Comp. Phys. Comm.* **67** (1992) 465.
- [13] L. Schoeffel, *Prog. Part. Nucl. Phys.* **65** (2010) 9; *Prog. Theor. Phys. Suppl.* **187** (2011) 179.

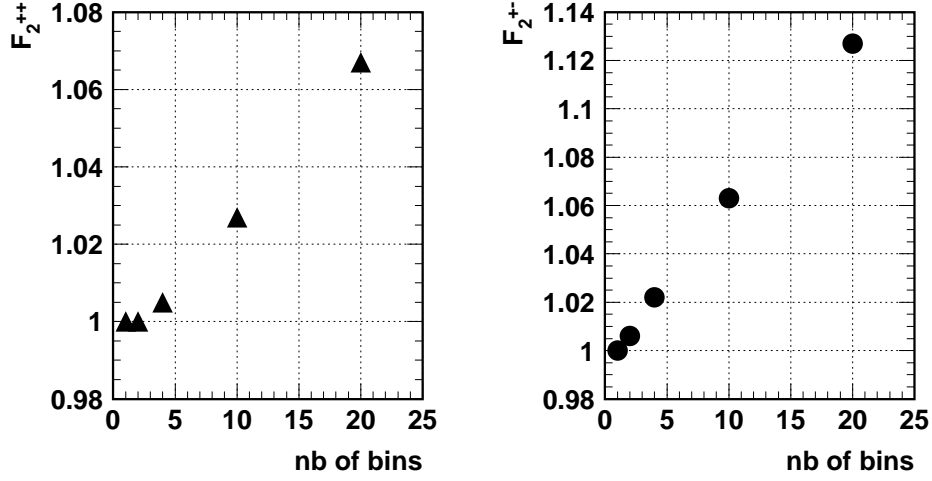


Figure 1: Euro future contract. Like-sign and unlike-sign factorial moments  $F_2^{++}$  and  $F_2^{+-}$  are displayed for 1, 2, 4, 10 or 20 bins. The global time window of 16.5 hours ( $n_{bins} = 1$ ). This provides a time resolution ranging from 1.6 hours for 20 bins till 16.5 hours for 1 bin. We observe that for 1 and 2 bins segmentation,  $F_2^{++}$  is found equal to 1 and  $F_2^{++}$  is increasing above 1 when the number of bins gets larger than 4. This confirms that non-Gaussian fluctuations in the sequence of returns start to play a role when the resolution is below 4 hours.

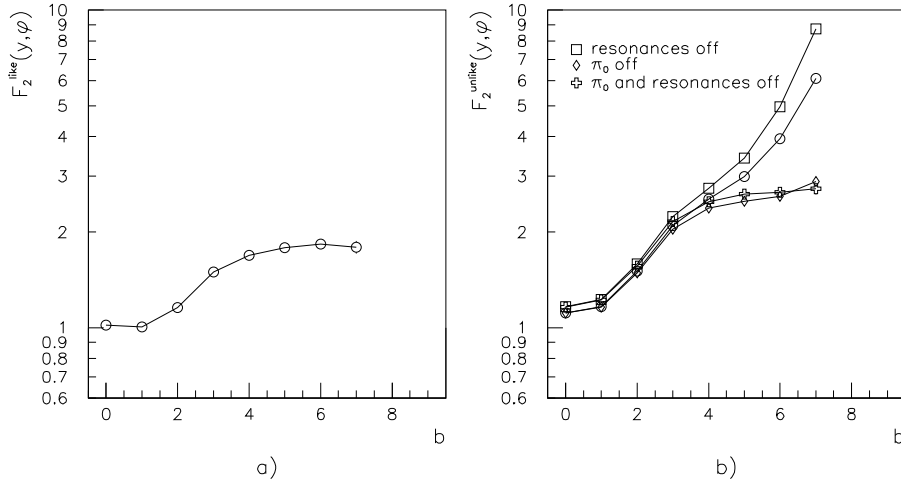


Figure 2:  $e^+e^-$  collisions at high energy. Factorial moments  $F_2(y, \phi)$  for like-charge (a) and unlike-charge (b) factorial moments  $F_2^{++}$  and  $F_2^{+-}$ , calculated using JETSET 7.4 for a center of mass energy of 91.5 GeV. See text for details.

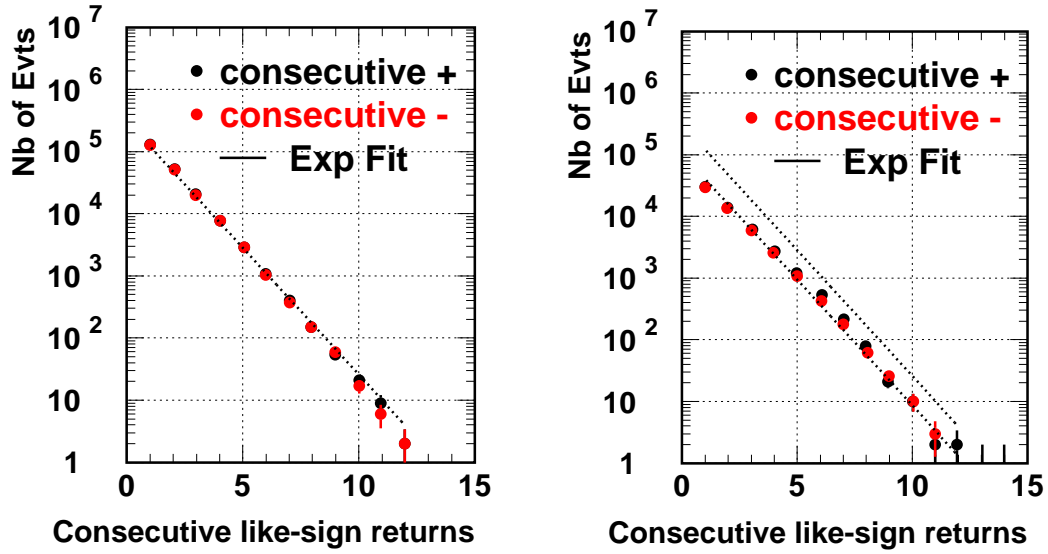


Figure 3: Left: Euro future contract sampled in 5min time units. Right: Euro future contract sampled in 30min time units. The distribution of events (probability distribution) is displayed as a function of the size of the gap. The gap is defined as the number of consecutive positive returns after a negative return, thus this is a gap in negative return. Inversely for a gap in positive return. The Euro future contract series over 10 years is sampled in two different time units (5min-Left and 30min-Right). In both cases, we observe effectively an exponential fall of the probability distribution as a function of the gap size. This confirms the relation derived in this letter  $p_0 \simeq \exp(-K_1) = \exp(-\langle n_{+,-} \rangle) = \exp(-\rho\Delta)$ . We have presented an exponential fit on top of each plot. On the Right plot (30 minutes sampling), we super-impose also the fit obtained on the 5 minutes sampling, in order to show that the exponential slope is identical in both cases.