IMPLEMENTATION OF THE RECOVERING CORRECTIONS INTO THE INTERMITTENT DATA ANALYSIS

BEATA ZIAJA

Department of Theoretical Physics
Institute of Nuclear Physics
Radzikowskiego 152, 31-342 Cracow, Poland
e-mail: beataz@qcd.ifj.edu.pl

(Received July 29, 1999)

The improved method of intermittent data analysis is proposed. It exploits, in addition to the standard density moments, the information on the bin–bin correlations, observed in the data and expressed in terms of the density correlators. The improving recovering corrections are implemented into the data analysis in the form of the recursive algorithm, and tested in the framework of multiplicative cascading models.

PACS numbers: 12.38.Mh

The first signals on possible intermittent behaviour in high-energy multiparticle production [1] were found in the data of the single event recorded by the JACEE collaboration [2]. The presence of large dynamical fluctuations manifesting a scaling behaviour was registered also afterwards in other accelerator experiments [3]. Many different models [4] have been proposed since to explain the power-law rise of the multiparticle moments, described by the coefficients called the intermittency exponents. Some of the models suggested that the observed scaling may be the result of final state multiparticle cascading [4], and the intermittent data represent the last stage of the cascade. In this approach the main problem lies in the extraction of the information on the previous cascade stages which are in some way encoded in the last stage data. The standard method of recovering the history of the cascade was proposed and applied originally to the JACEE event data. Since that time it has become a standard tool of multiparticle data analysis [4], especially in the event-by-event analysis [5,6]. However, the theoretical problem what is the interplay between the cascade recovered in the standard analysis and the true cascade which generated the data has been studied only recently [7,8]. It ended up with the proposal of introducing the new recovering corrections to the data analysis (see also [9]).

In this paper we would like to summarize the results on improvement of the standard data analysis achieved by including the recovering corrections. Most of these results were derived in Ref. [8]. In what follows we will concentrate on the technique of implementing the recovering corrections into the data analysis in the way which may be useful for experiment.

The **recovering corrections** aim to improve the method of recovering the history of the particle cascade, also called the rebinning [4]. The standard method reconstructs the cascade from the last step data, represented by a sample of M numbers: $x_i^{(n)}$ $(i=1,\ldots,M)$. They describe *e.g.* the distribution of particle density into bins. For simplicity assume $M=2^n$, where n denotes the number of cascade steps. The recovered particle density $y_i^{(n-k)}$ in the ith bin at the $(n-k)$th cascade step $(k=0,\ldots,n-1)$ takes then the form:

$$y_i^{(n-k)} = \frac{1}{2^k}\sum_{j=0}^{2^k-1} x^{(n)}_{2^k\times i+j}\,. \tag{1}$$

The intermittency exponents are extracted from unnormalized **reconstructed density moments** $z^{(k)}_{q;\,\mathrm{rec.}}$ [1]:

$$z^{(k)}_{q;\,\mathrm{rec.}} = \frac{1}{2^k}\sum_{j=0}^{2^k-1}\left(y_j^{(k)}\right)^q\,, \tag{2}$$

assuming that $z^{(k)}_{q;\,\mathrm{rec.}}$ manifests a power law behaviour:

$$z^{(k)}_{q;\,\mathrm{rec.}} \sim 2^{k\cdot\phi'_q}\,, \tag{3}$$

and the normalized intermittency exponent $\phi_{q;\,\mathrm{norm.}}=\phi'_q$,
where $\phi_{q;\,\mathrm{norm.}} := \phi_q - q\phi_1$. The exponent ϕ'_q is estimated as a slope of the linear χ^2-fit applied to the points $(k,\log z^{(k)}_{q;\,\mathrm{rec.}})$ $(\log x \equiv \log_2 x)$.

However, it was found and proved in [7,8] that there exists a difference between the true density moments:

$$z_q^{(k)} = \frac{1}{2^k}\sum_{i=0}^{2^k-1}\left(x_i^{(k)}\right)^q \tag{4}$$

[1] Actually, the factorial moments are normally used (to reduce the statistical noise) but this does not concern us here.

and the reconstructed ones (3) which influences the estimation of intermittency exponents. This difference may be expressed in the form of correcting factor $p_q(k)$:

$$z_{q;\,\mathrm{rec.}}^{(n-k)} = z_q^{(n-k)} \cdot p_q(k)\,, \tag{5}$$

which was called the **recovering correction** [8].

The recovering corrections contain information on the specific process which generated the true cascade. In what follows we restrict ourselves to the corrections considered in the framework of multiplicative random cascading models [10–12] which are nowadays widely recognized in multiparticle data analysis [4]. Similarly as in [8], we consider the multiplicative models with possible neighbour-node memory which generate a uniform distribution of particle density into bins. In the multiplicative cascade the root of the cascade is set equal 1: $x_0^{(0)} = 1$. One generates the next stages of the cascade recursively, following the scheme:

$$\begin{aligned} x_{2i}^{(k+1)} &:= W_1 \cdot x_i^{(k)}\,, \\ x_{2i+1}^{(k+1)} &:= W_2 \cdot x_i^{(k)}\,, \end{aligned} \tag{6}$$

where W_1 and W_2 are random variables of m model parameters a_j, $j = 1,\ldots,m$:

$$\begin{aligned} W_1 &= a_j \text{ with probability } p_{a_j}\,, \\ W_2 &= a_j \text{ with probability } p_{a_j}\,, \end{aligned} \tag{7}$$

with normalized probability weights $\sum_{j=1}^{m} p_{a_j} = 1$. The distribution of particle density will be uniform if the following condition is fulfilled:

$$p(W_1 = a_i, W_2 = a_j) = p(W_1 = a_j, W_2 = a_i)\,, \tag{8}$$

where $p(W_1 = a_i, W_2 = a_j)$ denotes probability of choosing in (6) $W_1 = a_i$ and $W_2 = a_j$ $(i,j = 1,\ldots,m)$. Then unnormalized density moments $z_q^{(k)}$ fulfil the scaling relation:

$$z_q^{(k)} \sim 2^{k\cdot\phi_q}\,, \tag{9}$$

and intermittency exponents ϕ_q read:

$$\phi_q = \log(a_1^q p_{a_1} + \ldots + a_m^q p_{a_m})\,. \tag{10}$$

The popular models: $\alpha-$, $p-$models [10, 12] and the $(p+\alpha)-$model introduced in [8] are special cases of multiplicative rule (6).

It was proved in [8] that for any multiplicative process which obeys rules (6), (8) recovering corrections $p_q(k)$ fulfil the recurrence equation[2]:

$$p_q(k) = \frac{1}{2^q} \sum_{j=0}^{q} \left(\begin{array}{c} q \\ j \end{array} \right) p_j(k-1) p_{q-j}(k-1) \langle W_1^j W_2^{q-j} \rangle \tag{11}$$

with the initial conditions:

$$\begin{aligned} p_q(0) &= 1, \\ p_0(k) &= 1\,. \end{aligned} \tag{12}$$

We introduce a notation:

$$\langle W_1^j W_2^l \rangle \equiv k_{j,l}\,. \tag{13}$$

Formula (11) implies that coefficients $k_{j,l}$ are the only parameters of the multiplicative model needed for calculating the value of $p_q(k)$. Furthermore, it is not difficult to establish the values of $k_{j,l}$ from the model. For either $j = 0$ or $l = 0$ they equal:

$$k_{j,0} = k_{0,j} = 2^{\phi_j}\,, \tag{14}$$

where ϕ_j's are ordinary intermittency exponents (9). To find the value of $k_{j,l}$ for both $j, l \neq 0$ we use the **unnormalized density correlators** $c_{j,l}^{(k)}$ [1,4,13]:

$$c_{j,l}^{(k)} = \frac{1}{2^{k-1}} \sum_{i=0}^{2^{k-1}-1} \left(x_{2i}^{(k)} \right)^j \left(x_{2i+1}^{(k)} \right)^l\,. \tag{15}$$

The correlators and the density moments fulfil the relation (see [8]):

$$c_{j,l}^{(k)} = z_{j+l}^{(k-1)} \cdot k_{j,l}\,, \tag{16}$$

which can be also rewritten as:

$$\log c_{j,l}^{(k)} = (k-1)\phi_{j+l} + \log k_{j,l}\,. \tag{17}$$

Both relations (16), (17) imply that we may derive $k_{j,l}$ in a straightforward way by calculating correlators and density moments from data, and applying to them the standard χ^2−fit.

Applying the standard method to the correlators at the previous cascade stages, we would expect to find the similar difference between the reconstructed correlators and the true ones, as it was observed for the density

[2] A similar recurrence relation has been obtained in a different context in [9].

moments. It was proved in [8] that this difference may be expressed in terms of the same recovering correction $p_q(k)$ (11) as for the density moments:

$$c^{(n-k)}_{j,l;\,\mathrm{rec.}} = c^{(n-k)}_{j,l} p_{j+l}(k)\,. \tag{18}$$

Now we have all tools needed for implementation of recovering corrections into the multiplicative data analysis. To illustrate the problem we describe the improved estimation of intermittency exponents of the second rank. The recovering correction of the second rank derived from (11) reads:

$$p_2(k) = \tfrac{1}{4}(\,p_2(k-1)k_{2,0} + p_2(k-1)k_{0,2} + 2p_1^2(k-1)k_{1,1}\,) \tag{19}$$

with the initial condition $p_2(0) = 1$. It was proved in [8] that:

$$p_1(k) = 2^{k\,\phi_1} = \left(z_1^{(n)}\right)^{\frac{k}{n}}\,. \tag{20}$$

The parameters needed to calculate $p_2(k)$ are following: $z_1^{(n)}$, ϕ_2 (*cf.* (14)) and $k_{1,1}$. Derivation of $z_1^{(n)}$ is straightforward, and to estimate the values of ϕ_2 and $k_{1,1}$ we propose the following recursive procedure. The primary values of ϕ_2 and $k_{1,1}$ may be obtained in the standard way from (3),(16). We substitute them to formula (19) to derive the approximate form of correction $p_2(k)$. Now, applying again the approximate form of $p_2(k)$ to equations (3), (16), one derives adjusted parameters ϕ_2 and $k_{1,1}$ and compares them with the primary values. If the relative difference is large, one repeats the recursive adjusting till the parameters do not change within a given accuracy.

One may generalize the above scheme for the intermittency exponents of any rank. Following [8], below we present the implementation algorithm which recursively adjusts the primary parameters ϕ_q, $k_{j,l}$ ($j+l=q$, $jl>0$) obtained after applying the standard method to the data:

(**INPUT**) parameters $\boldsymbol{\phi}_1,\ldots,\boldsymbol{\phi}_{q-1}$, $\boldsymbol{k}_{j,l}$ ($j+l=1,\ldots,q-1$) obtained after applying the implementation algorithm for $q=1,2,\ldots,q-1$ step-by-step:

(**1**) derive $\boldsymbol{\phi}'_q$, $\boldsymbol{k}'_{j,q-j}$ ($j=1,\ldots,q-1$) from data, using the standard method *i.e.* reconstruct the cascade using (1) and derive the parameters from relations:

$$\log z^{(k)}_{q;\,\mathrm{rec.}} \;=\; k\cdot\phi'_q + b\,, \tag{21}$$

$$c^{(k)}_{j,l;\,\mathrm{rec.}} \;=\; z^{(k-1)}_{j+l;\,\mathrm{rec.}}\cdot k'_{j,l}\,, \tag{22}$$

where $k=1,\ldots,n$ (*cf.* (3), (16)),

(**2**) derive $\boldsymbol{\phi}_{q;\,\mathrm{corr.}}$, $\boldsymbol{k}_{j,q-j;\,\mathrm{corr.}}$ $(j=1,\ldots,q-1)$ in the following substeps:

(**2.0**) calculate $p_q(k)$ from relation (*cf.* (11)):

$$p_q(k) = \frac{1}{2^q}\sum_{j=0}^{q}\left(\begin{array}{c} q \\ j \end{array}\right) p_j(k-1)p_{q-j}(k-1)k_{j,q-j}\,, \tag{23}$$

using $\boldsymbol{\phi}'_q$, $\boldsymbol{k}'_{j,q-j}$ derived in step (**1**), and estimate $\boldsymbol{\phi}_{q;\,\mathrm{corr.}}$ from:

$$\log z^{(n-k)}_{q;\,\mathrm{rec.}} - \log(p_q(k)) = (n-k)\cdot\phi_{q;\,\mathrm{corr.}} + b\,, \tag{24}$$

(**2.1**) calculate $p_q(k)$ from (23) using $\phi_{q;\,\mathrm{corr.}}$ (other parameters as after step (1)), and estimate $\boldsymbol{k}_{1,q-1;\,\mathrm{corr.}}$ from relation (*cf.* (17), (23)):

$$\log c^{(n-k)}_{j,l;\,\mathrm{rec.}} - \log(p_{j+l}(k)) = (n-k-1)\phi_{j+l} + \log k_{j,l;\,\mathrm{corr.}}, \tag{25}$$

$\ldots,$

(**2.q–1**) calculate $p_q(k)$ from (23), using all previously derived parameters $\boldsymbol{\phi}_{q;\,\mathrm{corr.}}$, $\boldsymbol{k}_{j,q-j;\,\mathrm{corr.}}$, and estimate $\boldsymbol{k}_{q-1,1;\,\mathrm{corr.}}$ from (25),

(**3**) compare the values of $\boldsymbol{\phi}'_q$, $\boldsymbol{k}'_{j,q-j}$ and $\boldsymbol{\phi}_{q;\,\mathrm{corr.}}$, $\boldsymbol{k}_{j,q-j;\,\mathrm{corr.}}$ $(j=1,\ldots,$ $q-1)$. If the relative difference is large, assume:

$$\begin{aligned} \phi'_q &:= \phi_{q;\,\mathrm{corr.}}\,, \\ k'_{j,q-j} &:= k_{j,q-j;\,\mathrm{corr.}} \end{aligned}$$

and repeat steps (2),(3) recursively until the relative difference between parameters before and after step (2) is small enough. Then go to the output, assuming $\boldsymbol{\phi}_q := \boldsymbol{\phi}'_q$, $\boldsymbol{k}_{j,q-j} := \boldsymbol{k}'_{j,q-j}$.

(**OUTPUT**) parameters $\boldsymbol{\phi}_1,\ldots,\boldsymbol{\phi}_q$; $\boldsymbol{k}_{j,l}$ $(j+l=1,\ldots,q)$.

Technical details and problems which appear when applying the algorithm to data were discussed in detail in [8].

We have performed numerical simulations of the $\alpha-$, $p-$ and $(p+\alpha)-$models [8] in order to test how the implementation algorithm works in practice. We generated 10000 cascades of the 10 step length for the $\alpha-$ and $(p+\alpha)-$models, and one cascade of the 10 step length for the $p-$model[3] for two different parameter sets separately.

[3] It can be proved that for a given parameter set the $p-$model generates always the same values of the correlators and density moments.

Implementation algorithm analized the data of the last cascade step. For each event it estimated the value of normalized intermittency exponents $\phi_{2;\,\mathrm{norm.}}$, $\phi_{3;\,\mathrm{norm.}}$ ($\phi_{i;\,\mathrm{norm.}} := \phi_i - i \cdot \phi_1$), using the standard method (step 1) with recovering corrections included (steps 2,3). The selected results (for one set of parameters) are presented in Fig. 1 and in Tabs. I, II.

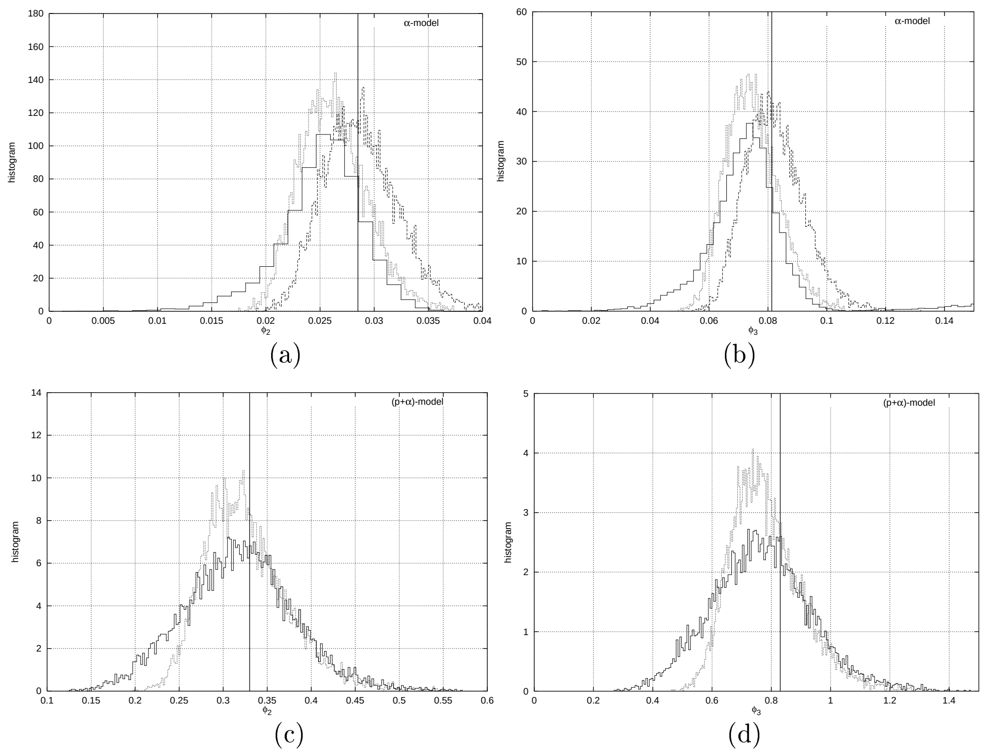

Fig. 1. (a), (b) — Estimation of normalized intermittency exponents $\phi_{2;\,\mathrm{norm.}}$ and $\phi_{3;\,\mathrm{norm.}}$ for $\alpha-$model, using the standard method (dotted line), the improved method with the implementation algorithm (thin solid line), and dedicated α-corrections [8] (dashed line) compared with the theoretical values (solid line), performed for one set of $\alpha-$model parameters: $a_1 = 0.8$, $a_2 = 1.1$, $p_1 = 1/3$; (c), (d) — Estimation of normalized intermittency exponents $\phi_{2;\,\mathrm{norm.}}$ and $\phi_{3;\,\mathrm{norm.}}$ for $(p+\alpha)-$model, using the standard method (dotted line), the improved method with the implementation algorithm (thin solid line), compared with the theoretical values (solid line), performed for one set of $(p+\alpha)-$model parameters: $a_{2i} = 1 - a_{2i-1}$, $p_{2i} = p_{2i-1}$ for $i = 1, \ldots, 10$,$a_1 = 0.2$, $a_3 = 0.5$, $a_5 = 0.6$, $a_7 = 0.3$, $a_9 = 0.45$,$a_{11} = 0.25$, $a_{13} = 0.1$, $a_{15} = 0.15$, $a_{17} = 0.87$, $a_{19} = 0.66$,$2p_1 = 0.05$, $2p_3 = 0.15$, $2p_5 = 0.25$, $2p_7 = 0.40$, $2p_9 = 0.05$,$2p_{11} = 0.05$, $2p_{13} = 0.02$, $2p_{15} = 0.02$, $2p_{17} = 0.005$, $2p_{19} = 0.005$.

TABLE I

Estimation of normalized intermittency exponents $\phi_{2;\,\mathrm{norm.}}$ and $\phi_{3;\,\mathrm{norm.}}$ and their dispersions for the α-model, using the standard method (second column), the improved method with the implementation algorithm (third column), and dedicated α- corrections [8] (fourth column), compared with the theoretical values (first column), performed for one set of α-model parameters (*cf.* Figs. 1(a), 1(b)).

	theor.	standard	algorithm	α-corr.
$\phi_{2;\,\mathrm{norm.}}$	0.0285	0.0251 ± 0.004	0.0246 ± 0.0033	0.0288 ± 0.004
$\phi_{3;\,\mathrm{norm.}}$	0.0813	0.0757 ± 0.010	0.0727 ± 0.009	0.0798 ± 0.0111

TABLE II

Estimation of normalized intermittency exponents $\phi_{2;\,\mathrm{norm.}}$ and $\phi_{3;\,\mathrm{norm.}}$ and their dispersions for the $(p+\alpha)$-model, using the standard method (second column), the improved method with the implementation algorithm (third column), compared with the theoretical values (first column), performed for one set of $(p+\alpha)$-model parameters (*cf.* Figs. 1(c), 1(d)).

	theor.	standard	algorithm
$\phi_{2;\,\mathrm{norm.}}$	0.177	0.170 ± 0.023	0.173 ± 0.029
$\phi_{3;\,\mathrm{norm.}}$	0.478	0.438 ± 0.069	0.470 ± 0.092

For the α-model the histograms of $\phi_{2;\mathrm{norm.}}$, $\phi_{3;\mathrm{norm.}}$ obtained in the standard method and the histograms with recovering corrections included are almost identical (see Figs. 1(a), 1(b) and Table I). In this case the recovering corrections can be implemented better when one applies directly dedicated α-model recovering correction [8] *i.e.* if one substitutes coefficient $k_{j,q-j}$ in (11) by the product: $k_{j,q-j} = 2^{\phi_j} \cdot 2^{\phi_{q-j}}$.

On the contrary, the implementation algorithm works well for the $(p+\alpha)$-model (see Figs. 1(c), 1(d) and Table II). For the $(p+\alpha)$-model the histogram with the recovering corrections included approximates well the theoretical value of normalized intermittency exponent. The histogram obtained by using the standard method is moved slightly to the left in comparison to the histogram with recovering corrections included.

We have checked that for the p-model the theoretical values of normalized intermittency exponents are estimated perfectly by both standard method and implementation algorithm [8].

It should be also mentioned that the histograms generated by the implementation algorithm (recovering corrections) are symmetric, in contrast to the standard ones, and their dispersions are of the same order as those derived for the standard method (*cf.* Tabs. I, II).

To sum up we analyzed the estimation of intermittency exponents from the data which were generated by a multiplicative random cascading process. The following methods were applied: the standard method of cascade recovering (1) and the improved method which included recursively the recovering corrections. The improved method was applied in the form of the implementation algorithm. Numerical simulations have been performed to check how both methods work in practice. The conclusions may be summarized as follows:

(a) standard method of estimation of intermittency exponents does not apply for the whole class of multiplicative models: its accuracy depends on the specific properties of the model and its parameters. The method does not detect a conservation law if present in the model;

(b) we propose an improved method of estimation of intermittency exponents. It exploits, in addition to the standard density (factorial) moments, the information on the bin–bin correlations, observed in the data and expressed in terms of the density correlators;

(c) the method is formulated in the form of the recursive algorithm which, starting from the parameters obtained from the density moments and correlators, allows successive improvements of the result;

(d) the method was tested in MC simulations which show that it is workable and indeed brings the experimental estimates closer to their theoretical values. Moreover, the improved distributions are symmetric with approximately the same dispersions as the uncorrected ones.

I would like to thank Prof. A. Białas for reading the manuscript and many suggestions and comments and Dr. R. Janik for discussions. This work was supported in part by Polish Government grant Project (KBN) 2P03B04214.

REFERENCES

[1] A. Bialas, R. Peschanski, *Nucl. Phys.* **B273**, 703 (1986); A. Bialas, R. Peschanski, *Nucl. Phys.* **B207**, 59 (1988).

[2] JACEE coll. T.H. Burnett *et al.*, *Phys. Rev. Lett.* **50**, 2062 (1983).

[3] P. Carlson (UA5), 4th Topical Workshop on p–p Collider Physics, Bern, March 1983; G. J. Alner *et al.*, *Phys. Rep.* **154**, 247 (1987).

[4] for recent reviews see E.A. de Wolf, I.M. Dremin, W. Kittel, *Phys. Rep.* **270**, 1 (1996), P. Bozek, M. Ploszajczak, R. Botet, *Phys. Rep.* **252**, 101 (1995).

[5] A. Bialas, R. Hwa, *Phys. Lett.* **B253**, 436 (1991); L. Lian-shou, F. Jinghua, W. Yuanfang, hep-ph/9903277.

[6] A. Bialas, B. Ziaja, *Phys. Lett.* **B378**, 319 (1996).

[7] R. Janik, B. Ziaja, *Acta Phys. Pol.* **B30**, 259 (1999).

[8] B. Ziaja, *Eur. Phys. J.* **C9**, 633 (1999).

[9] P. Desvallees, R. Ouziel, R. Peschanski, *Phys. Lett.* **B235**, 317 (1990); Y. Gabellini, J.-L. Meunier, R. Peschanski, *Z. Phys.* **C55**, 455 (1992).

[10] R. Peschanski, *Nucl. Phys.* **B327**, 144 (1989); P. Brax, R. Peschanski, *Nucl. Phys.* **B346**, 65 (1990); A. Bialas, R. Peschanski, *Phys. Lett.* **B207**, 59 (1988); H.C. Eggers, M. Greiner, P. Lipa, *Phys. Rev. Lett.* **80**, 1553 (1998); H.C. Eggers, M. Greiner, P. Lipa, preprint hep-ph 9811204.

[11] A. Bialas, A. Szczerba, K. Zalewski, *Z. Phys.* **C46**, 163 (1990).

[12] C. Meneveau, K. R. Sreenivasan, *Phys. Rev. Lett.* **59**, 1424 (1987).

[13] R. Peschanski, J. Seixas, preprint CERN-TH-5903-90.