PROCEEDINGS OF SCIENCE

BC, *PaC* and *SePaC* Methods for Fractal Analysis of Events

T. Dedovich*
Joint Institute for Nuclear Research, Dubna, Russia
E-mail: tdedovich@yandex.ru

M. Tokarev
Joint Institute for Nuclear research, Dubna, Russia
E-mail: tokarev@jinr.ru

Results of analysis of a wide class of fractals using the Box Counting (*BC*), *P*-adic Coverage (*PaC*) and System of the Equations of *P*-adic Coverage (*SePaC*) methods are presented. Methods are described and comparison underlying hypotheses is given. The parameters of algorithms - χ^2_{lim} for *BC*, P_{Max}, χ^2_{lim} for *PaC*, and P_{Max}, *Dev* for *SePaC* methods, are discussed. The efficiency of reconstruction of fractal dimension D_F, number of levels N_{lev} and *P*-adic coverage *P* is studied. Procedures to search for optimal values of parameters χ^2_{lim} , P_{Max}, and *Dev* to determine D_F, N_{lev}, and *P* by using these methods, are developed. The features of *BC*, *PaC* and *SePaC* methods for analysis of fractals with independent and dependent partition are noted.

XXI International Baldin Seminar on High Energy Physics Problems,
September 10-15, 2012
JINR, Dubna, Russia

*Speaker.

1. Introduction

The study of experimental data which demonstrates a property of self-similarity of hadron interactions at high energy is traditionally of considerable interest. One of such methods is known as z-scaling the approach. It was suggested in [1] and developed in [2, 3, 4, 5]. The method was applied for analysis of numerous experimental data on inclusive spectra of hadrons, direct photons and jets produced in $pp, p\bar{p}, pA$ and AA collisions over a wide kinematic range. Authors have concluded that the structure of hadrons, mechanism of constituent interactions and hadronization process reveal self-similar properties over a wide scale range. In this approach an inclusive particle spectrum is described in terms of dimensionless function ψ depending on single variable z. Dependence of scaling function ψ on variable z was found to be independent of the collision energy, and angle of produced particle. It is also described by power law $\psi(z) \sim z^{-\beta}$ at high z. The value of slope parameter β is independent of kinematic variables over a wide z-range. It indicates self-similarity of hadron production at various scales.

Anomalous events with high fluctuations of multiplicity were observed in pion-proton collisions at incident pion energy of 250 GeV [6]. These anomalous fluctuations reached a value of about 100 higher than the average of 60. This phenomenon was called intermittency. A.Bjalas [7] and R.Hwa [8] have shown that these fluctuations are dynamical ones. They reflect self-similarity of interactions. It has been found that the factorial moments F_q and G_q are described by the power law $F_q(\delta y) \sim (\delta y)^{-\phi(q)}$ and $G_q(\delta y) \sim (\delta y)^{-\tau(q)}$, where δy is the bin width of the rapidity distribution of particles, $\phi(q)$ and $\tau(q)$ are fractal exponents, q is the order of the moment.

These power laws characterizing self-similarity of particle production at different scales are typical for fractals. A fractal is usually determined as an object having the self-similar structure at different scales. It is characterized by the non-integral fractal dimension.

The fractal dimension is the value of D_F which provides the following finite limit:

$$\lim_{\delta \to \infty} N(\delta) \cdot \delta^{D_F} = const. \tag{1.1}$$

Here $N(\delta)$ is the number of probes having size δ which covers the object.

Exponent $\tau(q)$ characterizing intermittency defines the spectrum of fractal dimensions $D(q) = \tau(q)/(q-1)$. It reflects relationship of fractal and multiple particle production. It is assumed that a set of the hadrons produced in inelastic interactions is a set of points of the three-dimensional phase-space (p_T, η, ϕ). Here p_T is a transverse momentum, η, ϕ are the pseudorapidity and azimuthal angle, respectively. The distribution of points in the phase-space is non-uniformed and determined by the process of particle production. The set of these points is assumed to be a fractal and characterized by the fractal dimension, which depends on interaction dynamics.

Thus, we conclude that determination of fractal dimensions is an important task to understand interaction dynamics.

2. The parton shower and hardronization as fractal

The process of the particle production can lead to formation of the fractal. Let us consider a final-state parton shower in Monte Carlo code PYTHIA [9].

The kinematics of a shower is described by energy fraction z. The range of admissible values $z_- < z < z_+$ is defined by the effective mass of the partons. The range of an opening angle: $\Theta_-(z_-) < \Theta < \Theta_+(z_+)$ is restricted by values z_+, z_-. The opening angles of emitted partons are ordered: $\Theta_b, \Theta_c < \Theta_a$ i.e. the opening angle of a daughter parton can not be more than of the parent's.

Below we describe the scenario of the process which does not contradict the conditions of parton showers and leads to formation of a fractal. Figure 1(a) shows a schematic diagram of the process. The parton a outgoing from a hard process branches into partons b and c. The admissible opening angle Θ_+, Θ_- is shown by the solid and dashed lines and permissible ranges of the pseudorapidity space are shown by black rectangles. The pseudorapidity η of a parton is defined by $\eta = -0.5ln(tg(\Theta/2))$. The number of parts ($N_p$) in the general partition of space is shown on the left side. For example, the space at the first level is divided into 5 parts. The number of the permissible ranges (N_r) and their conformity to the number of partons or particles is shown on the right side. We see that two ranges at the first level are permissible. The first and second ranges consist of one and two parts, respectively. It is assumed that the partons get to each of these ranges. The second range is considered as uniformed object (dependent parts). We assume that branching and hardronization keep a specified spatial structure at different levels. The described fractal is a fractal with a dependent partition of its parts.

Next, we introduce the term P-adic partition(formation) of the fractal in analogy with the P-adic numbers. The P-adic positive number x is written as a series in powers of any prime number P: $x = a_0P^0 + a_1P^1 + a_2P^2 + a_3P^3 + \dots$. P-adic partition of the fractal is partition of the space during the formation of a fractal. The space of the fractal is divided into parts $M = \{P^0, P^1, P^2, P^3 + \dots\}$. The number of parts at each level changes as a degree of base P.

The fractal dimension is defined as a value D_F which provides the finite limit (1.1). The equation for the fractal described above is written in the form of $1/5^{D_F} + 2/5^{D_F} = 1$. The solution of the equation is equal to $D_F \approx 0.5639\dots$ The box dimension D_b is determined as a solution of the following equation:

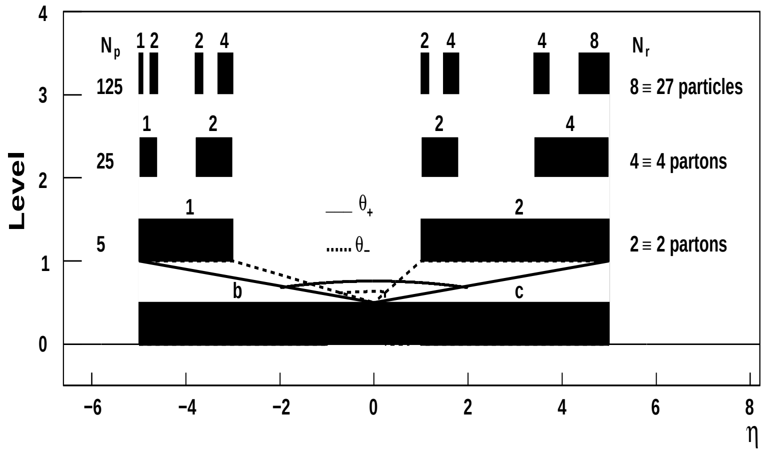
$$D_b = \lim_{\delta \to \infty} \frac{lnN(\delta)}{ln(\delta)}. \tag{2.1}$$

It is equal to $D_b = ln3/ln5 \approx 0.6828\dots$ Note that in the general case the fractal D_F and box D_b dimensions are not equal to each other.

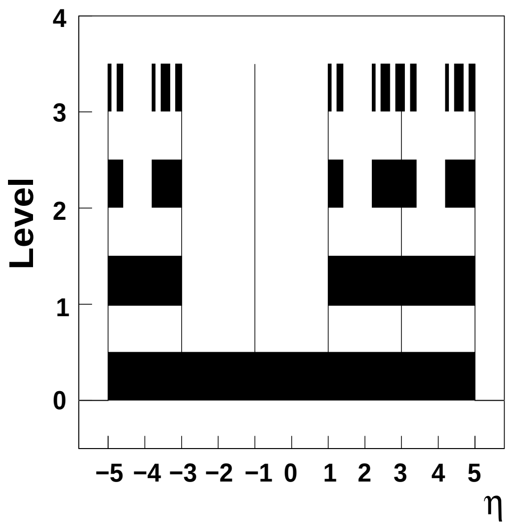
3. Types of fractals

In the present analysis we use two types of fractals. The first one is a set of fractals with dependent partition. The permissible range consists of several parts. It is divided as a uniformed object. The second type is a set of fractals with independent partition (see Fig. 1b). In latter case the permissible ranges consist of one part. They are divided independently. We have found that fractal dimension D_F is equal to box dimension D_b for fractals with independent partition. Dimensions for the fractal with dependent and independent partition are different.

Determination of the fractal dimension D_F is necessary for classification and analysis of events. Information about the type of the fractal and the value of base P in P-adic presentation



(a)



(b)

Figure 1: The parton shower and hardronization as fractal processes. Dependence of the permissible ranges of η-space at the fractal level. Fractals with dependent (a) and independent (b) partition of parts.

of the fractal is necessary to determine dimension D_F. The solution of the task is to choose the most appropriate fractal types and value of base P.

4. Comparison of *BC*, *PaC* and *SePaC* Methods

Here we compare the efficiency of *BC*, *PaC* and *SePaC* methods to determine the fractal dimension D_F, number of levels N_{lev} and base P. The Box counting (*BC*) method is a more popular one [10]. The P-adic Coverages (*PaC*) method [11] and System of the Equations of P-adic Coverages (*SePaC*) [12] are new methods proposed recently.

4.1 Data for Analysis

We analyse 437 fractals with dependent and 774 fractals with independent partition. For the first set of fractals with P-adic partition on base $P = 4 \div 8$ and for the second one - with P-adic partition on base $P = 3 \div 8$ all possible types of fractals have been considered. The fractal dimension D_F, number of levels N_{lev} and base P are recovered. For the recovered values $V = D_F, N_{lev}, P$ we determine errors $Er_V = (V_{test} - V)/V_{test}$ and efficiency

$$Ef_V = 1 - |Er_V| \quad \text{if } |Er_V| \leq 1 \,, \qquad\qquad Ef_V = 0 \quad \text{if } |Er_V| > 1. \tag{4.1}$$

The total efficiency is determined as $Ef_{tot} = Ef_{D_F} \cdot Ef_{Nlev} \cdot Ef_P$.

4.2 Box Counting (*BC*) and P-adic Coverages (*PaC*) Methods

The *BC* and *PaC* methods are based on the definition of the box dimension D_b. The description of the methods contains some common steps. Bellow the general features and differences of *BC* and *PaC* methods are described.

Both methods include the following steps:

1. Read out data ($\{X = \eta, ...\}$ of particles in event).
2. Construction of P-adic coverages: Each coverage is a set of distributions of variable X. The number of bins M_i in distributions of the set is changed as a degree of base P $(M_i = (P)^i)$.
BC: $P = 2$ (as usually), *PaC*: $P = 2, ..., P_{Max}$.
3. The counting number of non-zero bins $N(lev, P)$: saturation condition $N(lev, P) = N(lev + 1, P)$ defines the number of levels $N_{lev} = lev$.
4. Finding the parameter D_F of the power function $N \sim M^{D_F}$ and corresponding χ^2 of linear approximation of this function in a double-log scale for each P-adic coverage.
5. Accuracy condition: if $\chi^2 < \chi^2_{lim}$ then the set of particles is considered to be a fractal (P and $D_F(P)$).

The Box Counting (*BC*) method is based on the definition of the box dimension D_b. It uses one coverage and one parameter χ^2_{lim}. The P-adic Coverage (*PaC*) method is also based on the definition of the box dimension D_b. It includes various P-adic coverages with base $P = 2, ..., P_{max}$ and has two parameters - maximum P-adic coverage P_{Max} and χ^2_{lim}.

4.3 System of Equations of P-adic Coverages (*SePaC*) Method

Here we give brief description and characteristics of *SePaC* method. The method consists of the following steps:

1. Read out data ($\{X = \eta, ...\}$ of particles in event).
2. Construction of P-adic Coverages: $P = 3, ..P_{Max}$.
3. Counting number of non-zero bins $N(lev, P)$: saturation condition $N(lev, P) = N(lev + 1, P)$ defines the number of levels $N_{lev} = lev$.
4. Analysis of the system of the equations to verify the hypothesis (independent/dependent partition):

a) Construction of the system of the equations for all levels:

$$\sum_{i=1}^{N_{lev}} (d_{lev})^{D_F^{lev}} = 1. \tag{4.2}$$

N_{lev} is the number of levels, and d_{lev} is the length of permissible ranges for each level.
b) Finding of the solution D_F^{lev} by dichotomy method for each level.
c) Determination of the averaged value $< D_F^{lev} >$ and deviation ΔD_F^{lev}

$$< D_F^{lev} >= \sum_{lev=1}^{N_{lev}} (D_F^{lev})/N_{lev} \qquad \Delta D_F^{lev} = | < D_F^{lev} > - D_F^{lev}|. \tag{4.3}$$

d) Accuracy condition: if $\Delta D_F^{lev} < Dev$, then the set of particles is considered to be a fractal (P and $D_F = D_F^{lev}$).

The system of the equation of P-adic Coverages (*SePaC*) method is based on the definition of the fractal dimension D_F. The method considers various coverages with base $P = 2, ..., P_{max}$ and has two parameters - maximum P-adic coverage P_{Max} and deviation Dev, that defines the best coverage and whether the set of points is a fractal or not.

4.4 Search Procedure for Parameters of the Methods

BC, *PaC* and *SePaC* methods have parameters whose values should be determined. We have developed the search procedure of optimal values of parameters χ^2_{lim}, P_{Max}, Dev to determine fractal dimension D_F, number of levels N_{lev}, and base P.

4.4.1 Search Procedure of P_{Max} Parameter for *PaC* and *SePaC* Methods

The search procedure of maximum P-adic coverage P_{Max} for *PaC* and *SePaC* methods consists of the following steps:
1. Construction D_F, N_{lev} and base P distributions for different P_{Max} for all values of χ^2_{lim}.
2. Calculation of the difference of distributions ΔD_V for the considered variables for various values of P_{Max}, which are determined by the following formula:

$$\Delta D_V = \sum_{i=1}^{N_{bin}} |a_i - b_i|. \tag{4.4}$$

Here a_i and b_i are bin content for the adjacent distribution ($P_{Max} = P_j$ and P_{j+1}).
3. Calculation of the extended function ΔD_{Ext} for various values of P_{Max}

$$\Delta D_{Ext} = \Delta D_{D_F} + \Delta D_{N_{lev}} + \Delta D_P. \tag{4.5}$$

At this stage a preliminary analysis has been performed. We have compared the features of the dependences of extended function ΔD_{Ext} and total efficiency Ef_{tot} on maximum P-adic coverage P_{Max} (Fig. 2a,c) for fractals with independent partition for *PaC* method.

The plateau of the extended function corresponds to the maximum of the total efficiency Ef_{tot}. The same feature has been found for fractals with dependent and independent partition and for *PaC* and *SePaC* methods. Therefore, we have concluded that
4. The parameter P_{Max} is defined as the minimum value on a plateau for the dependence of extended function ΔD_{Ext} on maximum P-adic coverage P_{Max}. Thus, the search procedure of the optimal value of P_{Max} parameter has been developed for *PaC*, *SePaC* methods.

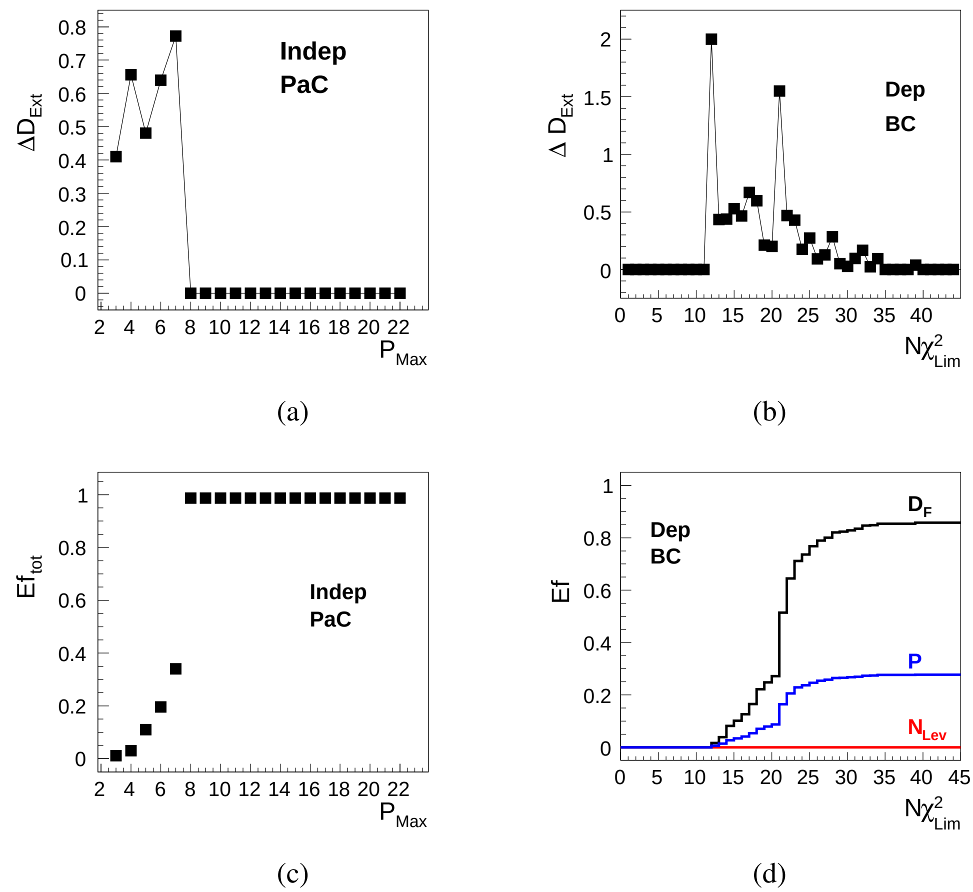


Figure 2: Dependence of extended function ΔD_{Ext} (a) and total efficiency Ef_{tot} (c) on maximum P-adic coverage P_{Max} for fractals with independent partition for *PaC* method. Dependences of extended function ΔD_{Ext} (b) and efficiency Ef (d) on χ^2_{lim} for fractals with dependent partition for *BC* method.

4.4.2 Search Procedure of χ^2_{lim} Parameter for *BC* and *PaC* Methods

The search procedure of χ^2_{lim} for *BC* and *PaC* methods includes the following steps:
1. Construction of D_F, N_{lev} and base P distributions for different χ^2_{lim} at the optimum value of P_{Max}.
2. Calculation of the difference of distributions ΔD_V (4.4) for the considered variables for various values of χ^2_{lim} .
3. Calculation of extended function ΔD_{Ext} (4.5) for various values of χ^2_{lim}.
4. Choice of χ^2_{lim} is determined by preliminary analysis results (comparison of $\Delta D_{Ext}(\chi^2_{lim})$ and $Ef_V(\chi^2_{lim})$ dependences).

We have used values of χ^2_{lim} over the range of $10^{-13} \div 2.5$. The Table (1) shows the correspondence of the number of $N_{\chi^2_{lim}}$ and values of χ^2_{lim}.

Figure 2(c,d) shows extended function ΔD_{Ext} and efficiency Ef_V as a function of $N_{\chi^2_{lim}}$ for fractals with dependent partition for *BC* method. The first plateau of the extended function corresponds to zero value of all efficiencies. The first peak of ΔD_{Ext} corresponds to the first values of the efficiencies Ef_{D_F}, Ef_P which are not zero. The second peak of the extended function corresponds to the rapid growth of Ef_{D_F}, Ef_P. The second plateau of ΔD_{Ext} corresponds to the maximum values of the efficiency of Ef_{D_F}, Ef_P. The same feature is seen for fractals with independent and dependent partition analysed by *BC* method.

So, for *BC* method the value of χ^2_{lim} parameter is defined as the minimum value on the second plateau of the dependence of extended function ΔD_{Ext} on χ^2_{lim}.

Table 1: Correspondence of $N_{\chi^2_{lim}}$ and χ^2_{lim}.

$N_{\chi^2_{lim}}$	1	2	3	4	5	6	7	8	9
χ^2_{lim}	10^{-13}	10^{-12}	10^{-11}	10^{-10}	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}
$N_{\chi^2_{lim}}$	10	11	12	13	14	15	16	17	18
χ^2_{lim}	10^{-4}	10^{-3}	0.01	0.02	0.03	0.04	0.05	0.06	0.07
$N_{\chi^2_{lim}}$	19	20	21	22	23	24	25	26	27
χ^2_{lim}	0.08	0.09	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$N_{\chi^2_{lim}}$	28	29	30	31	32	33	34	35	36
χ^2_{lim}	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
$N_{\chi^2_{lim}}$	37	38	39	40	41	42	43	44	45
χ^2_{lim}	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5

Figure 3(a,c) shows extended function ΔD_{Ext} and efficiency Ef_V as a function of $N_{\chi^2_{lim}}$ for fractals with dependent partition for *PaC* method. The first plateau of the extended function ΔD_{Ext} corresponds to a small value for all efficiencies. The first peak of ΔD_{Ext} corresponds to rapid growth of all the efficiencies. The second peak of ΔD_{Ext} corresponds to rapid decrease of $Ef_{N_{lev}}$, Ef_P. The range between peaks of extended function has the maximum efficiency $Ef_{N_{lev}}$, Ef_P. The second plateau of ΔD_{Ext} has the maximum value of the efficiency of Ef_{D_F}. So, for *PaC* method to determine base P and the number of levels N_{lev}, parameter χ^2_{lim} is defined as the minimum value between these peaks. To determine fractal dimension D_F, parameter χ^2_{lim} is defined as the minimum value on the second plateau of the same dependence.

Thus, the search procedure of the optimal value of χ^2_{lim} parameter has been developed for *BC* and *PaC* methods.

4.4.3 Search Procedure of *Dev* Parameter for *SePaC* Method

The search procedure of *Dev* parameter for *SePaC* method includes the following steps:
1. Construction of D_F, N_{lev} and base P distributions for different values of *Dev* at the optimum value of P_{Max}.
2. Calculation of the difference of distributions ΔD_V (4.4) for the considered variables for various values of *Dev* .
3. Calculation of extended function ΔD_{Ext} (4.5) for various values of *Dev*.
Comparison of features of dependences $\Delta D_{Ext}(Dev)$ and $Ef_V(Dev)$ (Fig. 3(c,d)) has shown that the negligible change of the extended function results in negligible change of efficiency Ef_V. Table

(2) shows the correspondence of N_{Dev} and Dev values.

4. Parameter Dev is defined as any value over the range of $10^{-6} \div 0.9$.

Table 2: Correspondence of $N_{\chi^2_{lim}}$ and χ^2_{lim} values.

N_{Dev}	1	2	3	4	5	6	7	8
Dev	10^{-6}	10^{-5}	10^{-4}	$2 \cdot 10^{-4}$	$5 \cdot 10^{-4}$	10^{-3}	$2 \cdot 10^{-3}$	$5 \cdot 10^{-3}$
N_{Dev}	9	10	11	12	13	14	15	16
Dev	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
N_{Dev}	17	18	19	20	21	22	23	24
Dev	0.09	0.1	0.2	0.3	0.4	0.5	0.6	0.7

Thus, we have concluded that the search procedure of the optimal value of *Dev* parameter has been developed for *SePaC* method.

4.5 Comparison of *BC*, *PaC* and *SePaC* Methods

Here we compare the efficiency of *BC*, *PaC* and *SePaC* methods for recovery of the fractal dimension D_F, number of levels N_{lev} and base P.

4.5.1 Fractals with independent partition

Comparison of these methods was carried out for the fractals with independent partition for optimal values of their parameters. The results of analysis are shown in Table 3. Efficiency of recovery of fractal dimension Ef_{D_F} by using *BC* method has been found to be about 90%. The fractal dimension is precisely restored for a small number of events (16%). The number of levels N_{lev} and base P are not precisely restored. *PaC* and *SePaC* methods restore D_F, N_{lev} and base P with high efficiency ($99-100\%$). The fractal dimension, the number of levels and the base P are precisely restored for a large number of events ($93-100\%$).

PaC and *SePaC* methods have an advantage in the analysis of fractals with independent partition.

Table 3: Results of analysis of fractals with independent partition by *BC*, *PaC* and *SePaC* methods.

$Method$	Ef_{D_F}	$Ef_{N_{lev}}$	Ef_P	Ef_{tot}	% event $Er_{D_F} < 0.5\%$	% event $Er_{N_{lev}} < 0.5\%$	% event $Er_P < 0.5\%$
BC	90.1%	0%	28.1%	0.3%	16%	0%	0%
PaC	99.4%	100%	99.2%	98.6%	93%	100%	97%
SePaC	99.4%	100%	99.2%	98.6%	93%	100%	97%

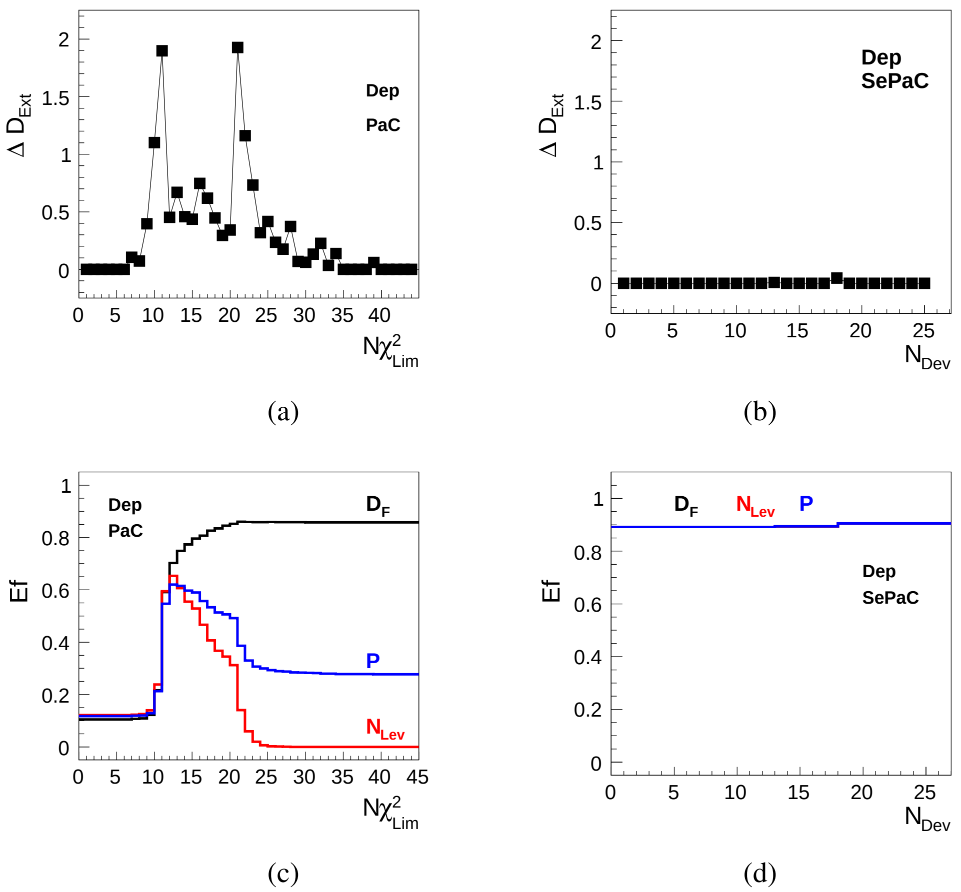


Figure 3: Dependences of extended function ΔD_{Ext} (a) and efficiency Ef (c) on χ^2_{lim} for fractals with dependent partition for *PaC* method. Dependences of extended function ΔD_{Ext} (b) and efficiency Ef (d) on *Dev* for fractals with dependent partition for *SePaC* method.

4.5.2 Fractals with Dependent Partition

Here we compare some results of analysis of fractals with dependent partition performed by *BC*, *PaC* and *SePaC* methods for the optimal values of their parameters. The results are presented in Table 4.

The efficiency of *BC* method of recovery of fractal dimension Ef_{D_F} has been found to be about 85%. For a very small number of events (2.5%) the fractal dimension D_F is precisely restored. The number of levels N_{lev} and base P are not restored.

PaC method has efficiency to recover fractal dimension about 86%, the number of levels $\approx$ 65% and base $\approx$ 61%. The fractal dimension D_F is precisely restored for a very small number of events (1.5%).

SePaC method restores of D_F, N_{lev} and base P with high efficiency (91%). Fractal dimension D_F, the number of levels N_{lev} and base P are precisely restored for a large number of events (89 – 90%).

Thus, based on the obtained results we have concluded that the *SePaC* has an advantage in the

Table 4: Results of analysis of fractals with dependent partition by *BC*, *PaC* and *SePaC* methods.

$Method$	Ef_{D_F}	$Ef_{N_{lev}}$	Ef_P	Ef_{tot}	$\%$ $event$ $Er_{D_F} < 0.5\%$	$\%$ $event$ $Er_{N_{lev}} < 0.5\%$	$\%$ $event$ $Er_P < 0.5\%$
BC	85%	0%	0%	0%	2.5%	0%	0%
PaC	86%	65%	61%	29%	1.5%	43%	26%
SePaC	91%	91%	91%	74%	89%	90%	90%

analysis of fractals with dependent partition.

5. Conclusions

The analysis of fractal (with independent and dependent partition) restoration by *BC*, *PaC* and *SePaC* methods has been performed. The search procedure of the optimal values of parameters χ^2_{lim}, P_{Max}, Dev to determine fractal dimension D_F, the number of levels N_{lev}, and base P for these methods has been developed. Comparison of *BC*, *PaC* and *SePaC* methods has shown advantages of *PaC* and *SePaC* methods to restore fractals with independent partition and *SePaC* method for restoration of fractals with dependent partition.

References

[1] I. Zborovský, Yu.A. Panebratsev, M.V. Tokarev, G. Škoro, *Phys. Rev.* **D 54**, 5548 (1996).

[2] M.V. Tokarev, T.G. Dedovich, *Int. J. Mod. Phys.* **A 15**, 3495 (2000).

[3] M.V. Tokarev, T.G. Dedovich, *Int. J. Mod. Phys. A* **15**, 3495 (2000).

[4] M.V. Tokarev, O.V. Rogachevski, T.G. Dedovich, *Nucl. Part. Phys.* **26**, 1671 (2000).

[5] M.V. Tokarev, T.G. Dedovich, and I.Zborovský, *Int. J. Mod. Phys. A* **27**, 1250115 (2012).

[6] M. Adamus et. al., *Phys. Lett.* **B 185**, 200 (1987).

[7] A. Bjalas, *Nucl. Phys.* **B 273**, 703 (1986).

[8] R. Hwa, *Phys. Rev.* **D 41**, 1456 (1990).

[9] T. Sjostrand et al., *Computer Physics Commun.* **135**, 238 (2001).

[10] B.B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).

[11] T.G. Dedovich, M.V. Tokarev, *Phys. Part. Nucl. Lett.* **8**, 521 (2011).

[12] T.G. Dedovich, M.V. Tokarev, *Phys. Part. Nucl. Lett.* **9**, 552 (2011).