Comparison of fractal analysis methods for fractal and random data sets

T.Dedovich & M.Tokarev
JINR, Dubna, Russia

Outline

- Self-similarity and fractality at high energy
- Methods of fractal analysis
- Comparison of methods for fractal and random data sets
- Summary

Self-similarity & z-Scaling

I.Zborovsky
Yu. Panebratsev
M.Tokarev
G.Skoro
Phys.Rev.D54(1996)5548

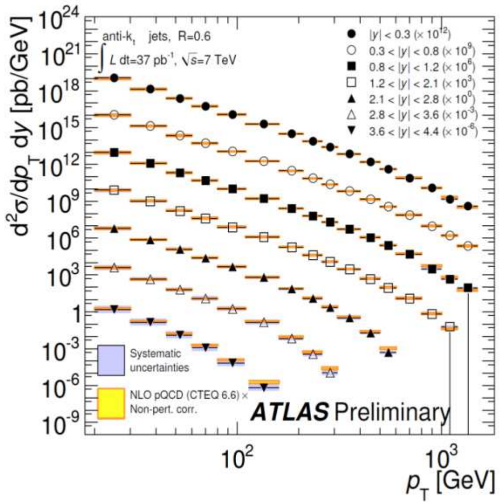
High-p_T inclusive particle spectra are described
by dimensionless function Ψ depending
on single dimensionless variable z

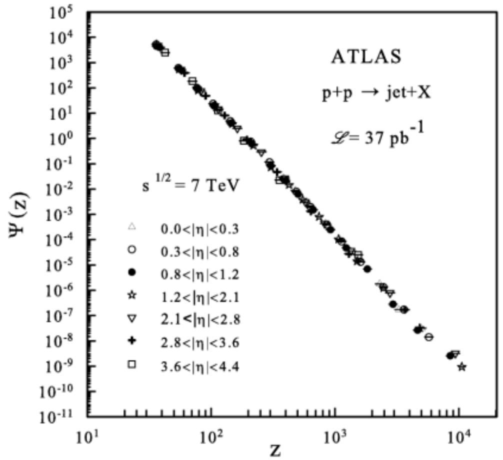
TD, M.Tokarev, I.Zborovsky
Int. Mod. Phys. A 15, 3495 (2000)
Int.Mod. Phys. A27,1250115 (2012)

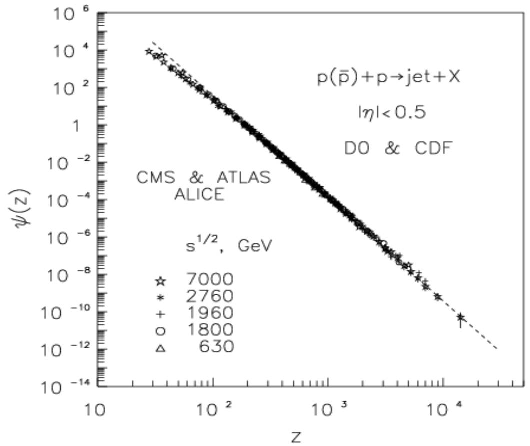
$$\Psi(z) = \frac{\pi s}{(dN/d\eta)\,\sigma_{inel}} J^{-1} E \frac{d^3\sigma}{dp^3}$$

$$z = z_0\, \Omega^{-1} \qquad \Omega = (1-x_1)^{\delta_1}(1-x_2)^{\delta_2}$$

- $\sqrt{s}$ - collision energy
- $dN/d\eta$ - multiplicity density
- σ_{inel} - total inelastic cross section
- $Ed^3\sigma/dp^3$ - inclusive cross section
- J - Jakobian from p_t, p_l to z, η
- $z\,(P_{1,2}, M_{1,2}, m, p, dN/d\eta)$







Energy, angular independence of $\Psi(z)$ and power law $\Psi(z) \sim z^{-\beta}$ over a wide z-range.
It indicates on self-similarity of jet production at various scales.

z-Scaling & Self-similarity and Fractality

z-Scaling - universal description of inclusive cross section over a wide kinematical region based on properties of locality, **self-similar, fractality**

Zborovsky , Panebratsev , Tokarev , Skoro , Phys. Rev. D54, 5548 (1996).

Self-similar: inclusive particles production is described by the set of values which does not contain the characteristic scale independent variables (dimensionless scaling function $\Psi(z)$ depending on dimensionless variable z)

$$\Psi(z) \equiv \frac{1}{N\sigma_{inl}} \frac{d\sigma}{dz}$$

Fractality: Variable z is constructed by analogy of length of the fractal

$$z = z_0 \Omega^{-1}$$

$$\Omega = (1-x_1)^{\delta_1} (1-x_2)^{\delta_2} (1-y_a)^{\varepsilon_a} (1-y_b)^{\varepsilon_b}$$

Here δ_1, δ_2 and ε_a, ε_b are anomalous fractal dimensions of the colliding objects, fragmentation process of inclusive and recoil particles

Motivation

- A quantitative characteristics of fractal structures and processes in z-scaling are dimensions δ_1, δ_2 and ε_a, ε_b .
- Discontinuity of dimension can be regarded as an indication on the change of physical subprocesses underlying inclusive particles/jet production.
- Fractal dimensions δ_1, δ_2 and ε_a, ε_b are defined for one set of experimental data and then verified in analysis of other data sets.
- New procedure: determination of dimensions by using fractal analysis.

- Comparison of existing methods of fractal analysis
- Development of new methods of restoring other characteristics of fractals

Fractals

Fractal is the self-similar object whose $D_F > D_T$

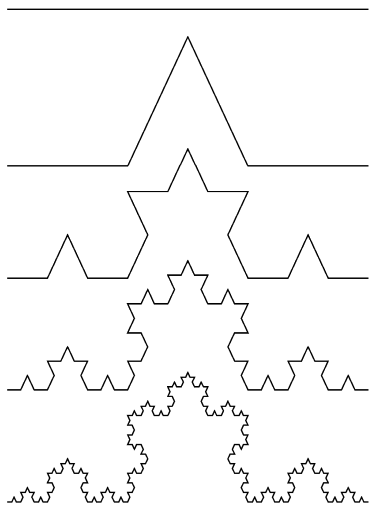
Fractal dimension is the value $\mathbf{D_F}$ which provides the finite limit

N - is number of probes with size $l_i < \delta$ covering the object

$$\lim_{\delta \to 0} \sum_{i=1}^{N} l_i^{D_F} = const$$

For the Euclidean objects measure M(δ) is independent of scale, and has a limit

$$M(\delta) = \lim_{\delta \to 0} N(\delta) \cdot \delta^{D_T} = const$$



For Fractal objects measure tends to infinity.

$$M(\delta) = \lim_{\delta \to 0} N(\delta) \cdot \delta^{D_T} \to \infty$$

The number of N(δ) unlimited increases more, rapidly than decreases δ

In order to evaluate the speed of growth impose Hausdorff measure,

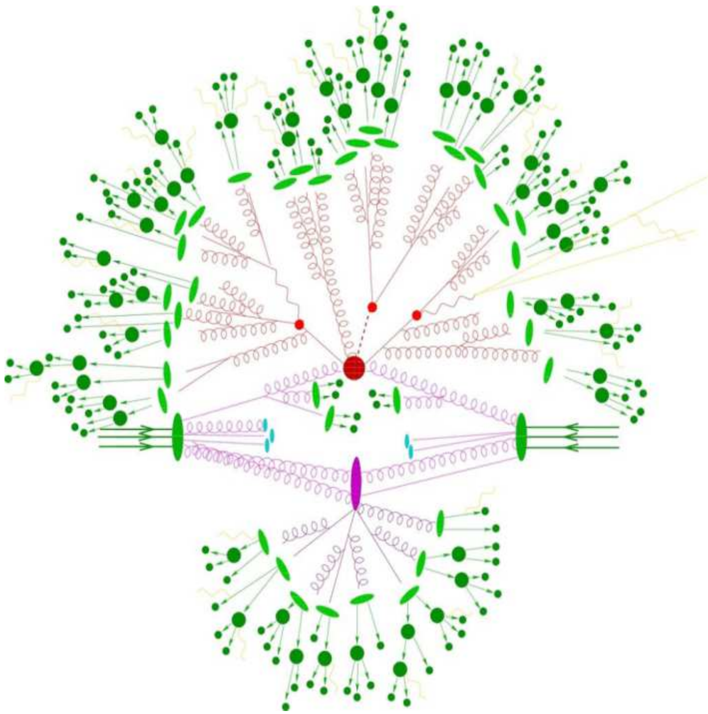
and examine its behavior at $\delta \to 0$

$$M_d = \lim_{\delta \to 0} \sum_{i=1}^{N} l_i^d$$

$$\begin{aligned} M_d &\to 0, & d &> D_F \\ M_d &\to const, & d &= D_F \\ M_d &\to \infty, & d &< D_F \end{aligned}$$

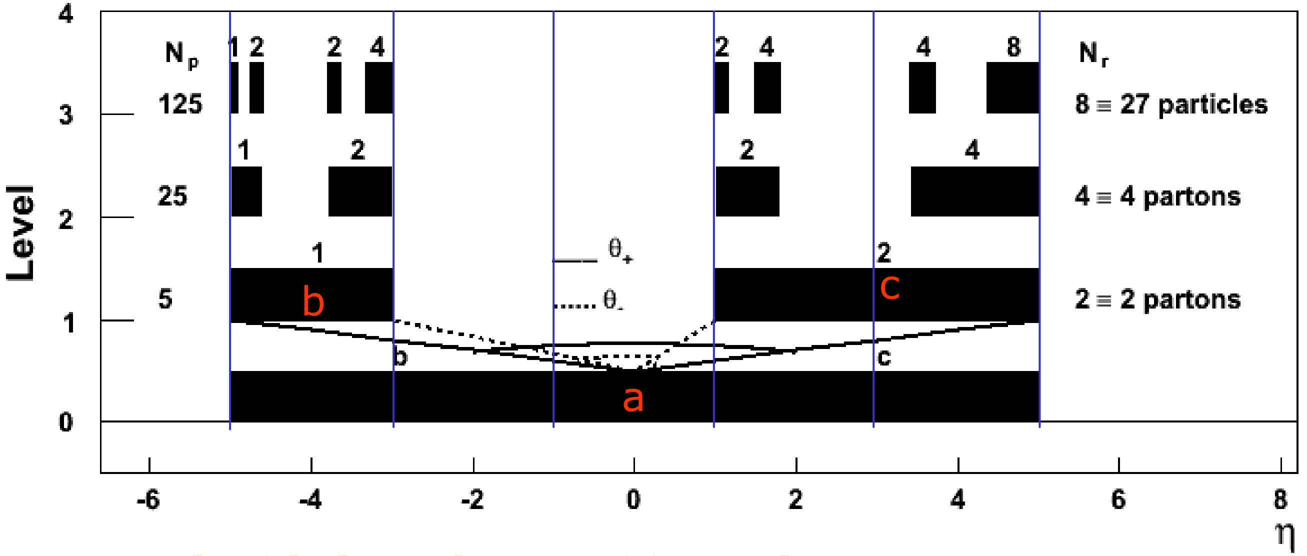
Fractality & Multiple production

- Set of hadrons produced in inelastic interaction is set of points of the three -dimensional phase-space (p_T,y,φ)
- The distribution of points in phase-space is non-uniformly and is determined by the process of particle production
- Set of these points in the phase-space are considered as a fractal and characterized by the fractal dimension, which depends on interaction dynamics



Determination of fractal dimensions is important for reconstruction of interaction dynamics

Scenario of Parton Shower and Hadronization



Fractal with dependent partition on base P=5

- Outgoing from hard process parton branch a→bc
- $\theta_\pm$- admissible opening angle

  Black rectangles –permissible ranges

$$\eta = -\ln(tg(\vartheta/2))$$

- The range (consisting of two parts) is considered as uniform object (dependent parts).
- Further branching and hadronization keeps spatial structure.

Fractal dimension D_F

$$\lim_{\delta\to 0}\sum_{i=1}^{N} l_i^{D_F} = \text{const}$$

$$(1/5)^{D_F} + (2/5)^{D_F} = 1$$

$$D_F \approx 0.5639...$$

Box dimension

$$D_b = -\lim_{\delta\to 0}\frac{\ln N(\delta)}{\ln(\delta)}$$

$$D_b = \ln 3/\ln 5 \approx 0.6826...$$

Power Law

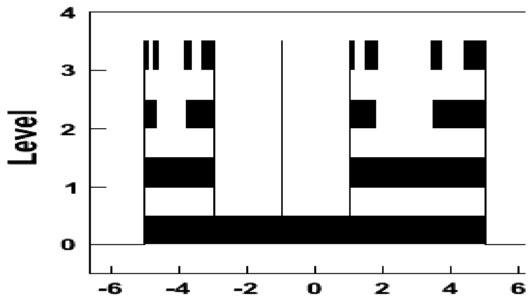
$$N_r = N_p^s$$

$$s = \ln 2/\ln 5 \approx 0.4307...$$

Models of fractals formation

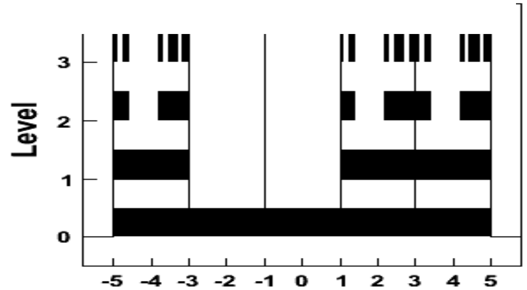
➢ Fractal with dependent partition of parts:

- permissible ranges consisting of arbitrary number of parts and are not in contact
- parts are divided as uniform object (dependently)



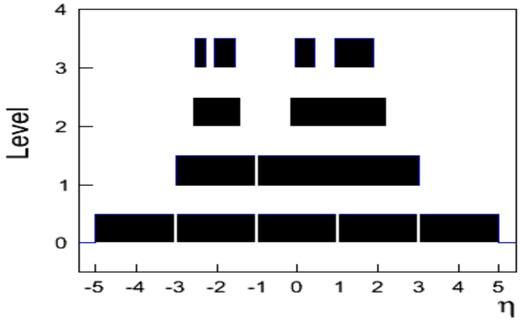
➢ Fractal with independent partition of parts:

- permissible ranges consisting of one part
- permissible parts are divided independently



➢ Fractal with the combined partition of parts:

- the permissible ranges consist of arbitrary number of parts and can come into contact
- parts belonging to the same range are divided dependently, and different ranges - independently



$D_F = D_b$ for fractal with independent partition of parts

$D_F \neq D_b$ for fractal with dependent and combined partition of parts

BC, PaC methods of fractal analysis

1. Read out data – $\{X = \eta , p_T, \ldots\}$ of particles in events
2. Construction of P-adic Coverages:
   Each coverage is a set of distributions of variable X.
   The number of bins M_i in distributions
   are changed as a degree of basis P: $M_i = (P)^i$
   BC: as a rule P = 2, PaC: P = 2,..P_{Max}
3. Count a 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. Base PaC method:
   Processed data for which $N(lev,P) = N(1,P)^{lev}$
5. Finding slope parameter D_F and χ^2
   Dependence of ln N vs. ln M for each P-adic coverage
6. Accuracy condition $\chi^2(P) < \chi^2_{lim}$:
   The set of particles is a fractal (P and $D_F(P)$)

BC, PaC – determine the box dimension
BC has two parameters: Par1=P, Par2=χ^2_{lim}
PaC has two parameters: Par1=P_{Max}, Par2=χ^2_{lim}

PaC
DT, M.Tokarev
Phys.Part.Nucl.Lett.
8 (2011) 521

BC – Box Counting
PaC – P-adic Coverage

SePaC method of fractal analysis

1. Read out data – {η_i } of particles in event
2. Construction of P-adic Coverages: P = 3,..P_{Max}
3. Count a number of non-zero bins N(lev,P):

   saturation condition N(lev,P) = N(lev+1,P) defines number of levels N_{lev}= lev
4. Base SePaC method

   Processed data for which $N_r(lev,P) = N_r(1,P)^{lev}$
5. Analysis of system of equations for verification of hypothesis

   of independent/dependent partition:

- Construction of a system of the equations for all levels
  N_{lev} and d_{lev} are number and length of permissible ranges

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

- Finding solution D_F^{lev} by using a dichotomy method for each level
- Defining average value $< D_F^{lev} >$ and deviation ΔD_F^{lev}
- Accuracy condition $\frac{\Delta D_F^{lev}}{< D_F^{lev} >} < Dev$: set of particles is a fractal

SePaC – determine the fractal dimension
SePaC method has two parameters: Par1=P_{Max}, Par2=Dev

SePaC
DT, M.Tokarev
Phys.Part.Nucl.Lett.
9(2011) 552

SePaC – System
of the Equation
of P-adic Coverage

Results of analysis by BC, PaC and SePaC methods

- Fractals with independent partition: the base PaC and SePaC methods are preferable
- Fractals with dependent partition: the base SePaC and modified PaC methods are preferable
- Fractals with combined partition: the modified SePaC method is preferable

DT, M.Tokarev
Phys.Part.Nucl.Lett.
2013 6(183) p491-500

DT, M.Tokarev
Phys.Part.Nucl.Lett.
2013 6(183) p.791-803

A more flexible procedure for general analysis of a wide class of fractals, taking into account peculiarities of the developed methods, is needed

Two-step procedure of fractal analysis

I step: Analysis of the general data set by the base method

- determination of the optimal values of parameter Par
- analysis of the data using the base method with selected parameters
- determination of characteristics of reconstructed fractals
- selection of unreconstucted fractals

II step: Analysis of unreconstucted fractals by the modified method

Data sample of fractal analysis

1857 fractals

(fractals with independent, dependent and combined partition)

1857 random data sets

(multiplicity distributions for random and fractal data sets are equal)

Reconstruction efficiency and impurities

Efficiency = portion of reconstructed fractals Por_{Frac}

Impurities = portion of reconstructed random sets as a fractal Por_{Rand}

Search Procedure for Method Parameter Par

1. Construction of D_F, N_{lev}, P distributions for different Par

$$\Delta D_V(Par) = \sum_{i=1}^{N_{bin}} | a_i - b_i |$$

2. Calculation of function $\Delta D_V(Par)$ for V=D_F, N_{lev}, P.

   a_i and b_i are bin content for adjacent distributions

3. Calculation of extended function $\Delta D_{Ext}(Par)$

$$\Delta D_{Ext}(Par) = \Delta D_{D_F}(Par) + \Delta D_{N_{lev}}(Par) + \Delta D_P(Par)$$

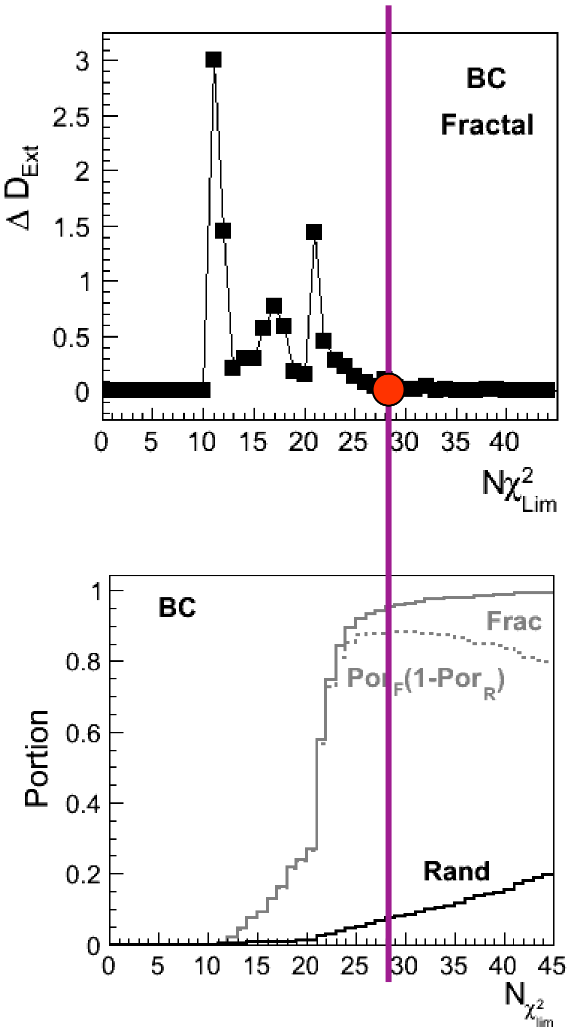
4. Choice of the value Par on the basis of analysis of function $\Delta D_{Ext}(Par)$

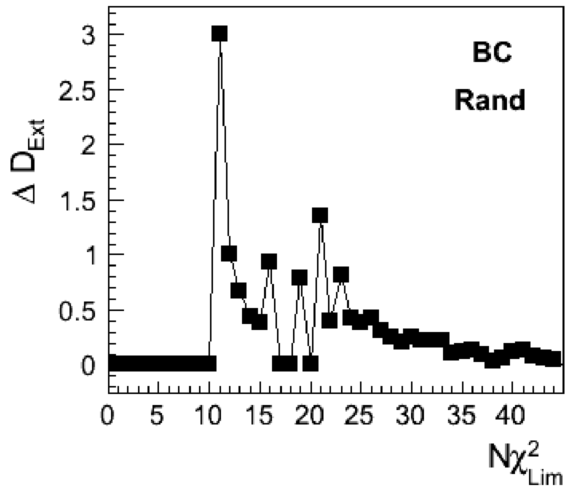
Correspondence of number of and value of Par

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

N_{Dev}	1	2	3	4	5	6	7	8	9	10	11	12	13
Dev	10^{-6}	10^{-5}	10^{-4}	$2\cdot10^{-4}$	$5\cdot10^{-4}$	10^{-3}	$2\cdot10^{-3}$	$5\cdot10^{-3}$	10^{-2}	0.02	0.03	0.04	0.05
N_{Dev}	14	15	16	17	18	19	20	21	22	23	24	25	26
Dev	0.06	0.07	0.08	0.09	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

BC-method

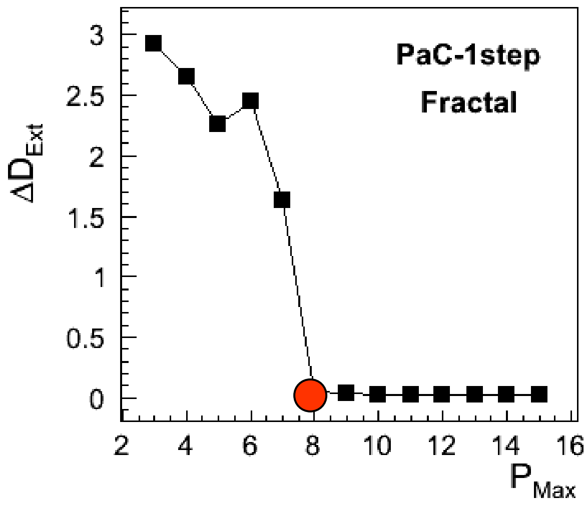


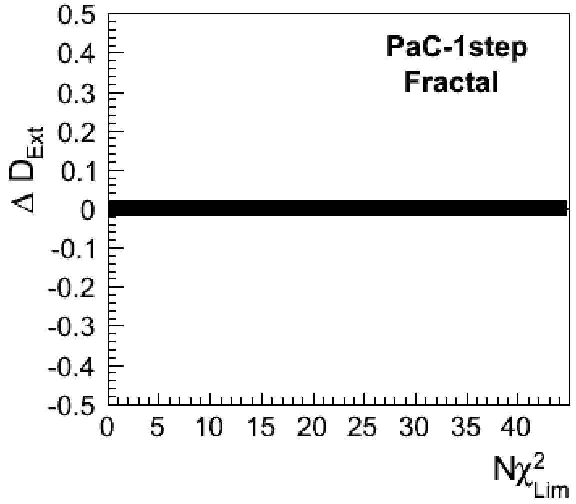


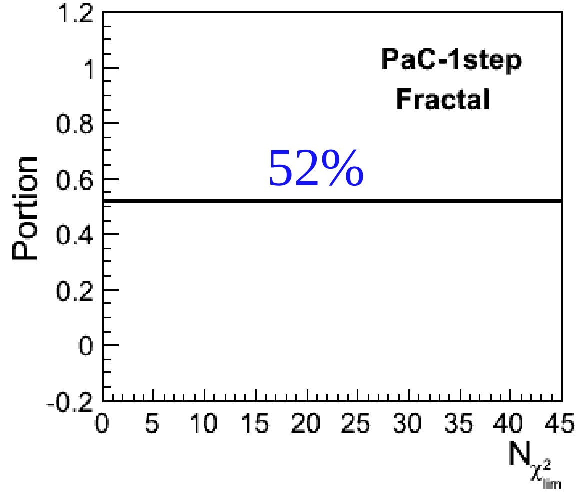
- For fractal: choice χ^2_{lim} as the minimal value on the second plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ corresponds the maximal Por_{Frac}
- Shapes $\Delta D_{Ext}(N\chi^2_{lim})$ for fractal and random set are similar
- Shapes Por_{Frac} and Por_{Rand} are different
- Shapes $Por_{Frac}(1-Por_{Rand})$ allows you to define a valid range χ^2_{lim} at which Por_{Frac} is maximal and Por_{Rand} is minimal
- χ^2_{lim} established for fractal corresponds to this range

T. Dedovich

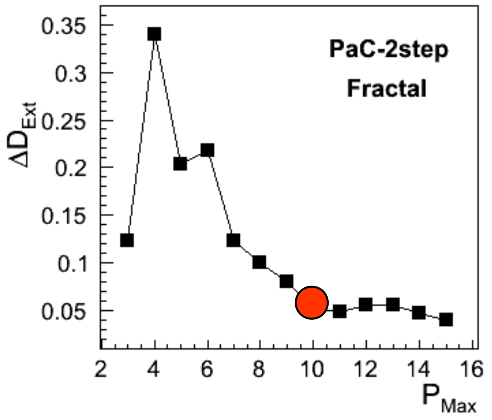
PaC-method (1st-step)



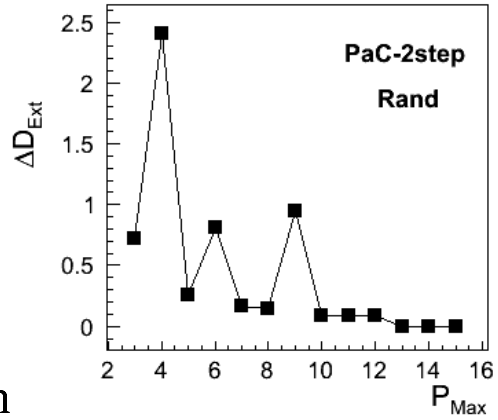




- Choice P_{Max} as the minimal value on the plateau of $\Delta D_{Ext}(P_{Max})$ corresponds to the maximal Por_{Frac}
- Choice χ^2_{lim} is arbitrary (shapes $\Delta D_{Ext}(N\chi^2_{lim})$ and $Por_{Frac}(N\chi^2_{lim})$ does not depend on χ^2_{lim}
- 1st step of PaC-method restore 52% fractal from data analysis
- 1st step of PaC-method does not restore the random set as a fractal (no impurities !)



PaC-method (2nd-step)



- For fractal choice

  P_{Max} as the minimal value on the plateau of $\Delta D_{Ext}(P_{Max})$

  χ^2_{lim} as the value of second peak on $\Delta D_{Ext}(N\chi^2_{lim})$ corresponds the maximal Por_{Frac}

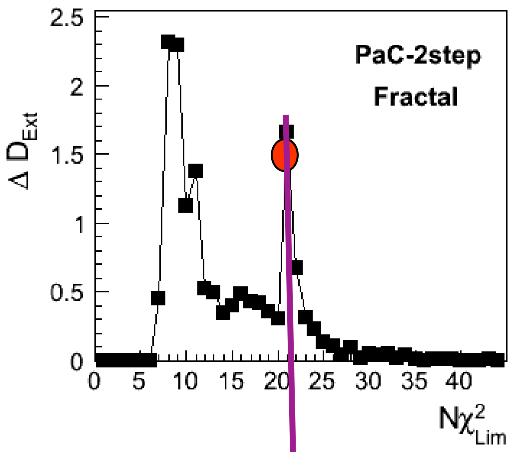
- Shapes of $\Delta D_{Ext}(P_{Max})$ for fractal and random sets have a plateu

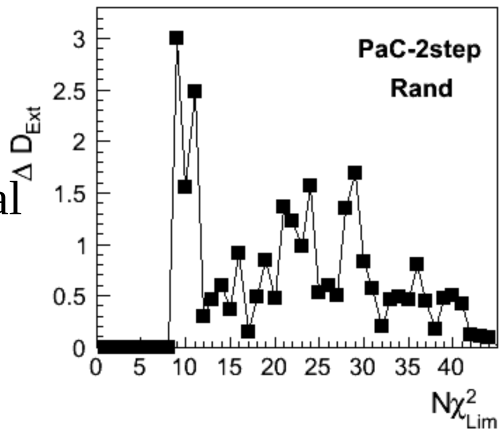
  Shapes of $\Delta D_{Ext}(N\chi^2_{lim})$ for fractal and random sets are different (If $\Delta D_{Ext}(N\chi^2_{lim})$ has two peaks then data set is fractal)

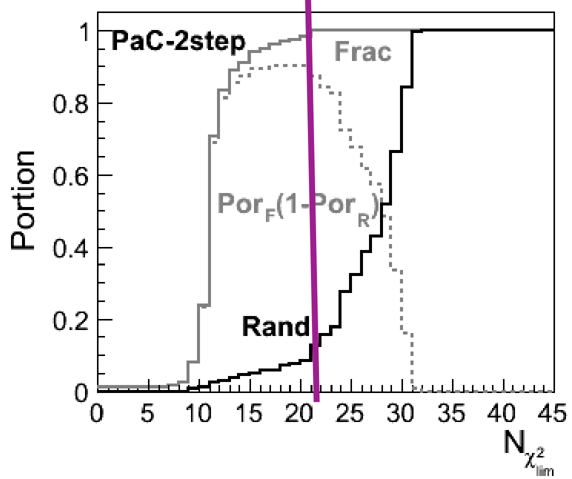
- Shapes Por_{Frac} and Por_{Rand} are different

- Shape $Por_{Frac}(1\text{-}Por_{Rand})$ allows you to define a valid range χ^2_{lim} at which Por_{Frac} is maximal and Por_{Rand} is minimal

- χ^2_{lim} established for fractal corresponds to this range

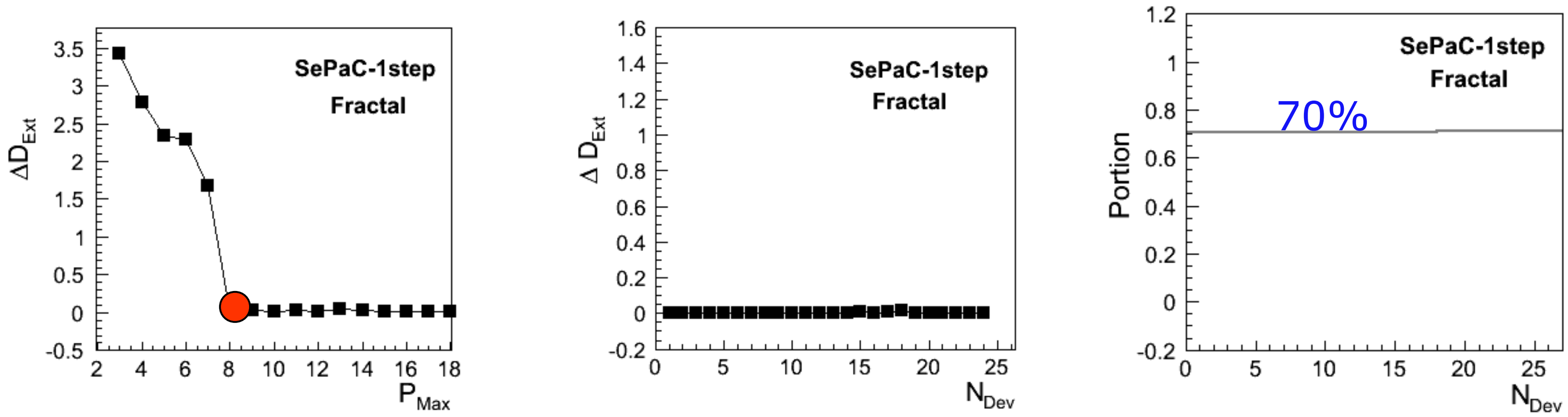






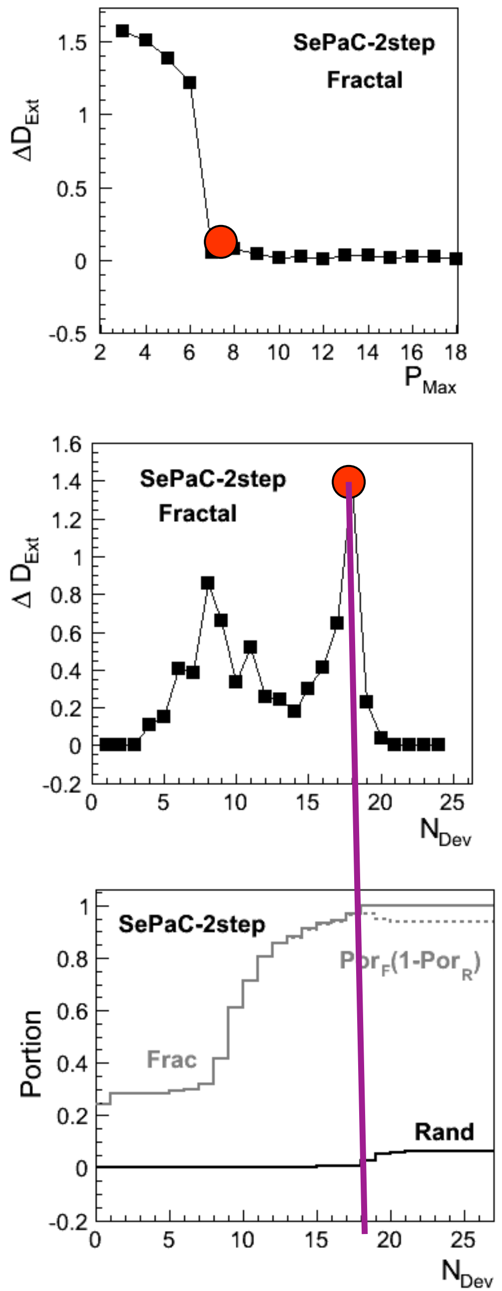
JINR

SePaC-method (1st step)



- Choice P_{Max} as the minimal value on the plateau of $\Delta D_{Ext}(P_{Max})$ corresponds to the maximal Por_{Frac}
- Choice Dev is arbitrary (shapes $\Delta D_{Ext}(N_{Dev})$ and $Por_{Frac}(N_{Dev})$ does not depend on Dev
- 1step of SePaC-method restore 70% fractal from analysis
- 1step of SePaC-method does not restore the random set as a fractal (no impurities !)

SePaC-method (2-step)



- For fractal choice

  P_{Max} as the minimal value on the plateau of $\Delta D_{Ext}(P_{Max})$

  Dev as the value of second peak on $\Delta D_{Ext}(N_{Dev})$ corresponds the maximal Por_{Frac}

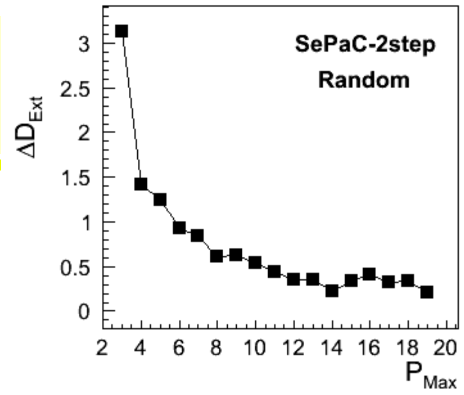
- Shapes of $\Delta D_{Ext}(P_{Max})$ for fractal and random sets are different

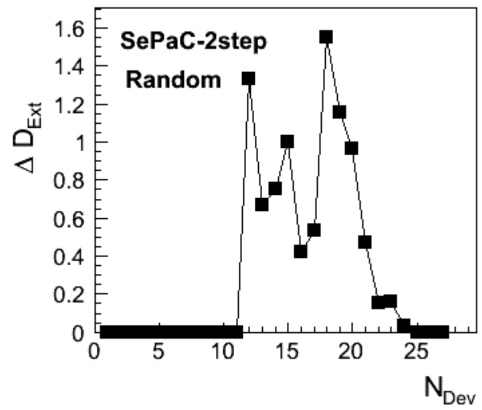
  Shapes of $\Delta D_{Ext}(N_{Dev})$ for fractal and random sets are different (if 1st peak of $\Delta D_{Ext}(N_{Dev})$ is smeared then data set is a fractal)

- Shapes Por_{Frac} and Por_{Rand} are different

- Shape $Por_{Frac}(1-Por_{Rand})$ allows you to define a valid range Dev at which Por_{Frac} is maximal and Por_{Rand} is minimal

- Dev established for fractal corresponds to this range





JINR

Comparison of Methods

Method	Por_{Frac}	Por_{Rand}	$Por_F(1-Por_R)$
BC	0.95	0.08	0.88
2-step PaC	1.0	0.12	0.90
2-step SePaC	1.0	0.03	0.97

Two step procedure of SePaC method

has the advantage for fractal analysis

- fully fractals reconstruction

- lowest impurity

Summary

- Two-step procedure of fractal analysis for PaC и SePaC methods was developed
- The search procedure of optimal values of parameters for BC, PaC and SePaC methods is developed.
- Comparision of the fractal analysis methods for fractals and random data set is carried out.
- The two-step procedure of fractal analysis for SePaC method has advantage before other ones
  - fully fractal reconstruction
  - lowest impurity

Thank You for attention