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

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

Outline

- Self-similarity and fractality in multiple production at high energies
- Fractals and models of their formation
- Method and Data for Fractal Analysis
- Comparison results of BC, PaC and SePaC methods
- Summary

XXI International Baldin Seminar on High Energy Physics Problems
"Relativistic Nuclear Physics and Quantum Chromodynamics",
JINR, Dubna, Russia, September 10-15, 2012

Self-similarity & z-Scaling

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

High-p_T inclusive particle spectra is 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)

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

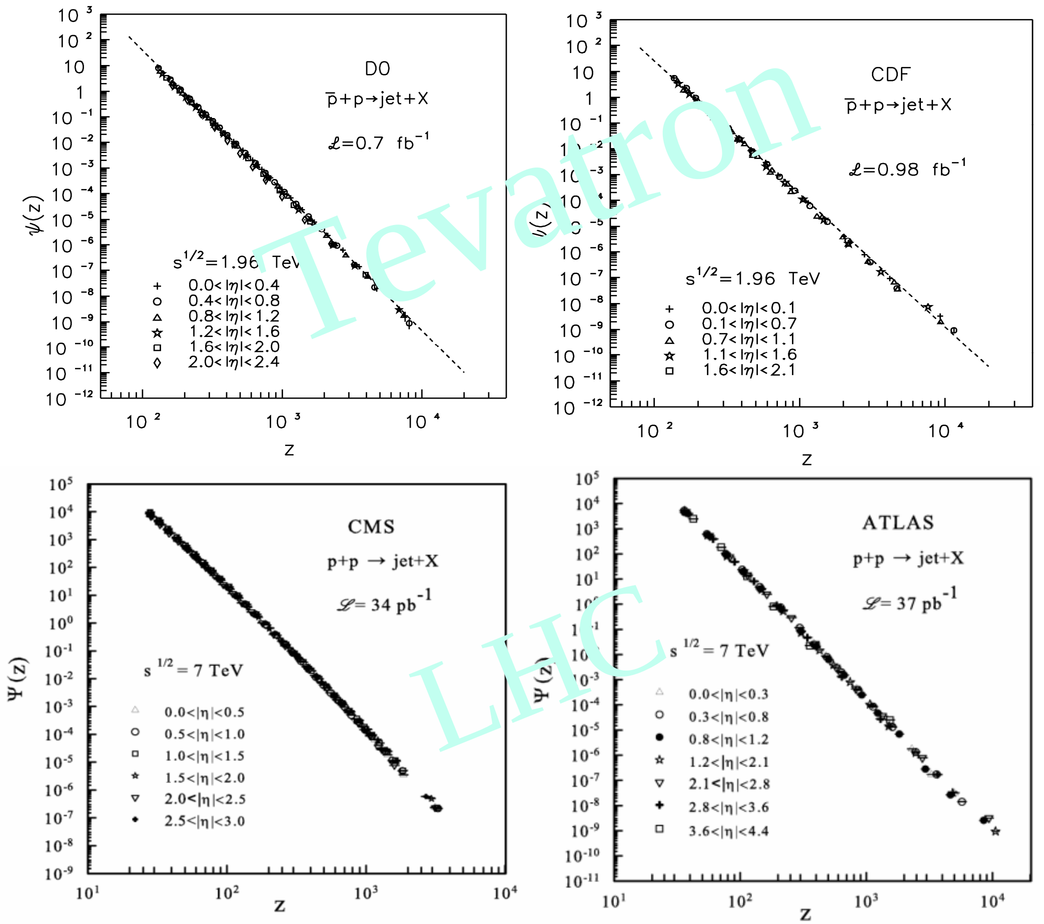
$$z = z_0\,\Omega^{-1}$$

$$\Omega = (1-x_1)^{\delta_1}(1-x_2)^{\delta_2}$$

$$(x_1P_1 + x_2P_2 - p)^2 = M_X^2$$

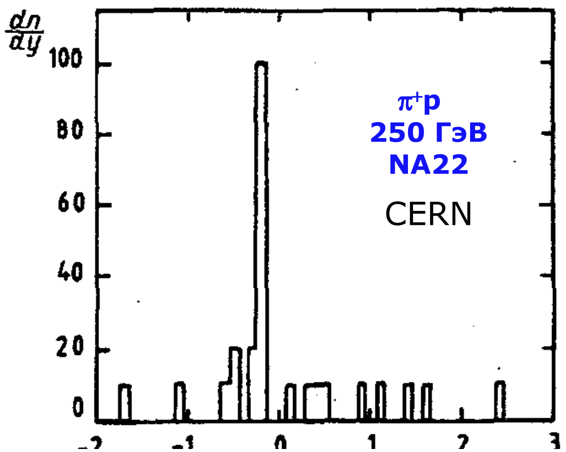
$$M_X = x_1M_1 + x_2M_2 + m_2$$

$\sqrt{s}$ - collision energy
$dN/d\eta$ - multiplicity density
σ_{inel} - total inelastic cross section
$Ed^3\sigma/dp^3$ - the inclusive cross section
J - Jakobian
x_1,x_2 - momentum fractions



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

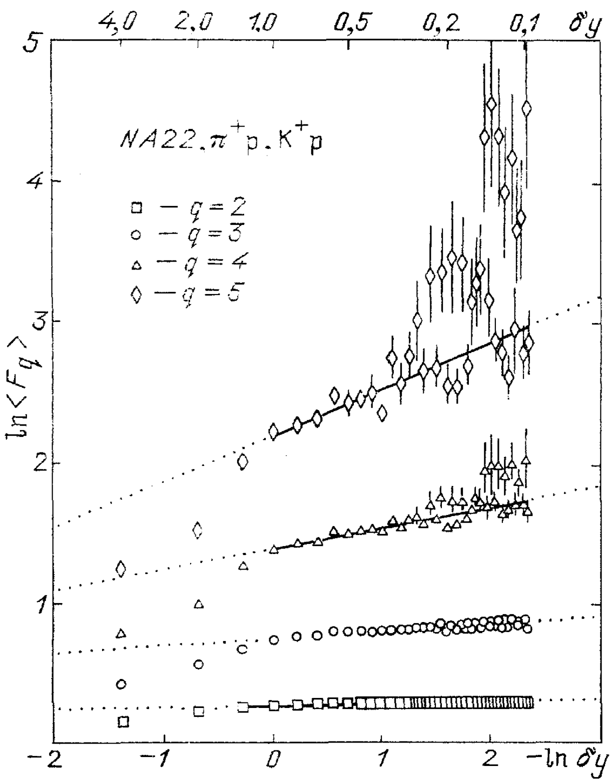
Self-similarity & Fluctuation & Intermittency

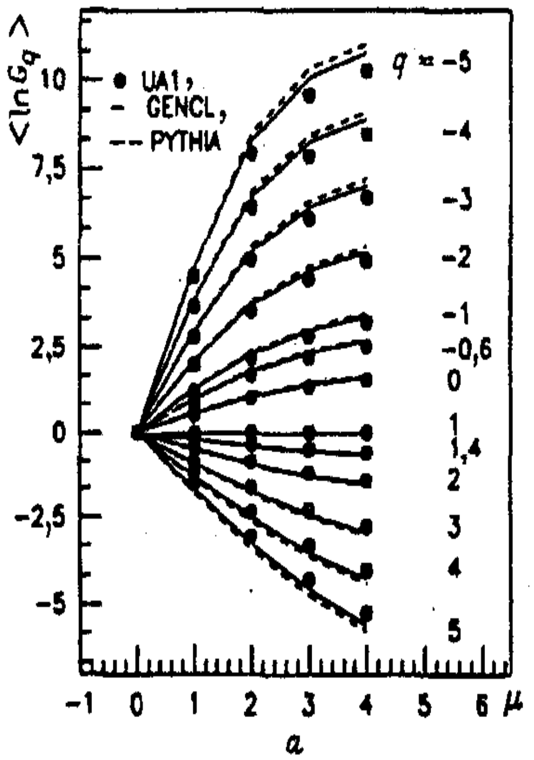


Intermittency: abnormal events with large multiplicity fluctuation were observed in h-h interaction.
NA22: $\delta y=0.1$, $dN/dy=100$, dN/dy / $\langle dN/dy \rangle \approx 60$

Observable fluctuations are dynamical and reflects self-similarity of interaction

Bialas A. // Nucl. Phys. 1986 B273, p.703
Hwa R. // Phys.Rev. 1990. D41, p.1456





Power Law dependence of factorial moments $F_q(\delta y)$, $G_q(\delta y)$ on bin widths δy

$$F_q(\delta y) \sim (\delta y)^{-\phi(q)}$$

$$G_q(\delta y) \sim (\delta y)^{-\tau(q)}$$

$$F_q(\delta y) = M^{q-1} \frac{\left\langle \sum_{k=1}^{M} n_k (n_k - 1) \ldots (n_k - q + 1) \right\rangle}{\langle n \rangle^q}$$

$$G_q = \sum_{m=1}^{M}{}' p_m^q \ , \ p_m = n_m/n, \ n = \sum_{m=1}^{M} n_m$$

$M=2^{\mu}$-number of bins with width δy,
n_k- number of particles in k-bin

Self-similarity & Fractality & Multiple production

$F_q(\delta y) \sim (\delta y)^{-\phi(q)}$, $G_q(\delta y) \sim (\delta y)^{-\tau(q)}$, $\Psi(z) \sim z^{-\beta}$

Power Laws established experimentally, and characterizing self-similarity of particles production on different scales **are typical for fractals**

Fractal is the self-similar object with **nonintegral (fractal) dimension**

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

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

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

Relationship of fractal and multiple production

Power Law exponent τ(q)

(Intermittency: $G_q(\delta y) \sim (\delta y)^{-\tau(q)}$)

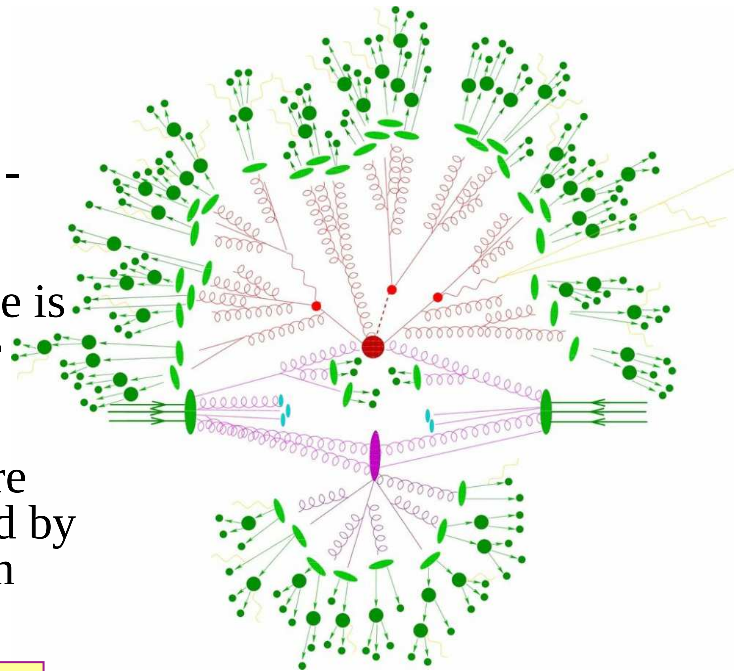
defines

spectrum of fractal dimensions

(generalized fractal dimension)

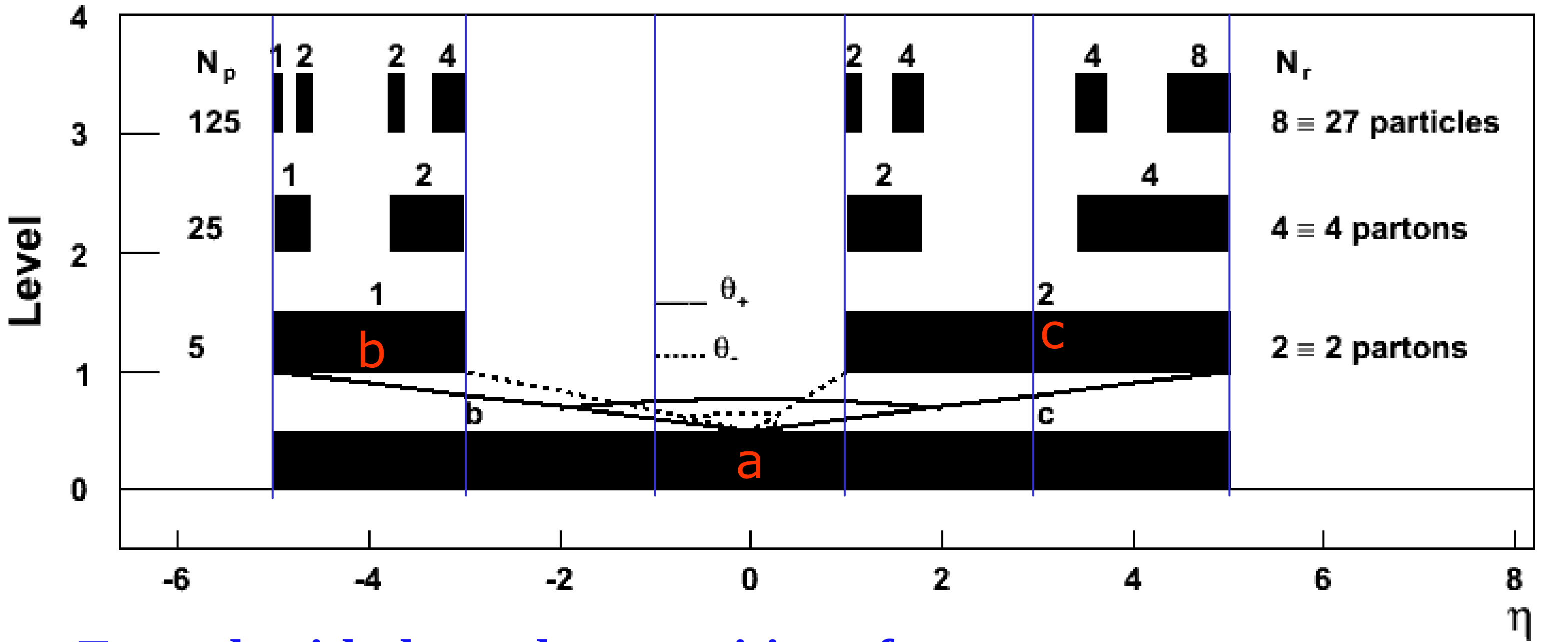
D(q) = t(q)/(q-1)

- Set of hadrons produced in inelastic interaction are 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 of parts

- Outgoing from hard process parton branch a→bc
- θ±- admissible opening angle

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

Black rectangles –permissible ranges

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

Fractal dimension D_F

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

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

$$D_F \approx 0.5639\ldots$$

Box dimension

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

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

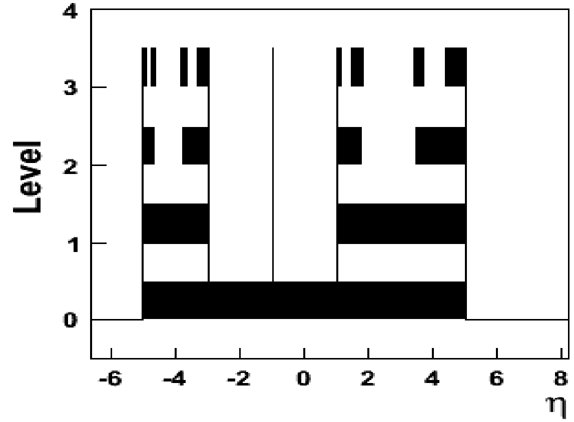
Power Law

$$N_r = N_p^s$$

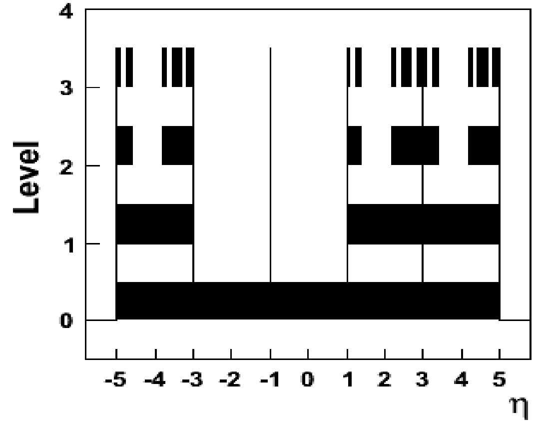
$$s = \ln 2/\ln 5 \approx 0.4307\ldots$$

Models of Fractals Formation

Fractals with dependent partition:
permissible range consisting of several parts are divided as uniform object (dependently)



Fractals with independent partition:
permissible ranges consist of one part and are divided independently



$D_F \neq D_b$ for fractal with dependent partition of parts
$D_F = D_b$ for fractal with independent partition of parts

Goal & Problems of fractal analysis

➢Goal: find the fractal dimension for classification events

➢But: to do it correctly we must know the type of the fractal and number of parts at first level

➢Possible solution:

- try a set of fractal coverage and fractal types and choose the best one
- Introduce the createria: how good the fractal fit of the data

Comparison of existing methods

- Data of analysis
- description of the methods

  BoxCounting

  P-adic Coverages methods

  SePaC method
- Results and Summary

Data & Analysis

- **437 fractals with dependent partition:**
  All possible variants of fractals with number of parts at first level P=4÷8

- **774 Fractals with independent partition:**
  All possible variants of fractals with number of parts at first level P=3÷8

Efficiency of reconstruction of D_F, N_{lev}, P

$$Eff_V = 1 - |Err_V| \quad if\ |Err_V| \le 1$$

$$Eff_V = 0 \quad if\ |Err_V| > 1$$

$$V = D_F, N_{lev}, P \quad Err_V = (V_{test} - V)/V_{test}$$

$$Eff_{Full} = Eff_{D_F} \cdot Eff_{N_{lev}} \cdot Eff_P$$

BC, PaC Methods of Reconstruction of Fractal dimension D_F

Box Counting (BC) & P-adic Coverages (PaC)

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 of set

   are changed as degree of basis P ($M_i=(P)^i$)

   BC: as a rule P=2, PaC: P=2,..P_{Max}
3. Counting 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. Finding the slope parameter D_F and χ^2 of dependence

   of ln N vs. ln M for each P-adic coverage
5. Accuracy condition $\chi^2(P) < \chi^2_{lim}$:

   the set of particles is a fractal (P and $D_F(P)$)

B.B.Mandelbrot
The Fractal Geometry
of Nature

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

BC has one parameter - χ^2_{lim}

PaC has two parameters - P_{Max}, χ^2_{lim}

SePaC Methods of definition of fractal dimension D_F

Systems of the Equations of P-adic Coverage (SePaC)

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

1. Read out data – $\{\eta_i\}$ 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 number of levels N_{lev}= lev
4. Analysis of system of the equations for verification of hypothesis

   (independent/dependent partition):

- **Construction system** of the equations for all levels

  N_{lev} and d_{lev} - number of and long permissible ranges for each level

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

- **Finding** solution D_F^{lev} by a dichotomy method for each level
- **Defining** average value $\langle D_F^{lev} \rangle = \sum_{lev=1}^{N_{lev}} D_F^{lev} / N_{lev}$ and deviation $\Delta D_F^{lev} = | \langle D_F^{lev} \rangle - D_F^{lev} |$
- **Accuracy condition** $\Delta D_F^{lev} < Dev$: set of particles is **a fractal** (P and $D_F = \langle D_F^{lev} \rangle$)

SePaC has two parameters – P_{Max}, Dev

Parameter P_{max} of PaC, SePaC

Search procedure of maximum P-adic Coverage P_{max}

1. Construction of D_F, N_{lev}, P distributions for different P_{Max} at all χ^2_{lim}
2. Calculation of function $\Delta D_V\,(P_{Max})$ of a difference of distributions

   for $V=D_F$, N_{lev}, P

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

   a_i and b_i – bin contents for adjacent distributions ($P_{Max}= P_j$ and P_j+1)

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

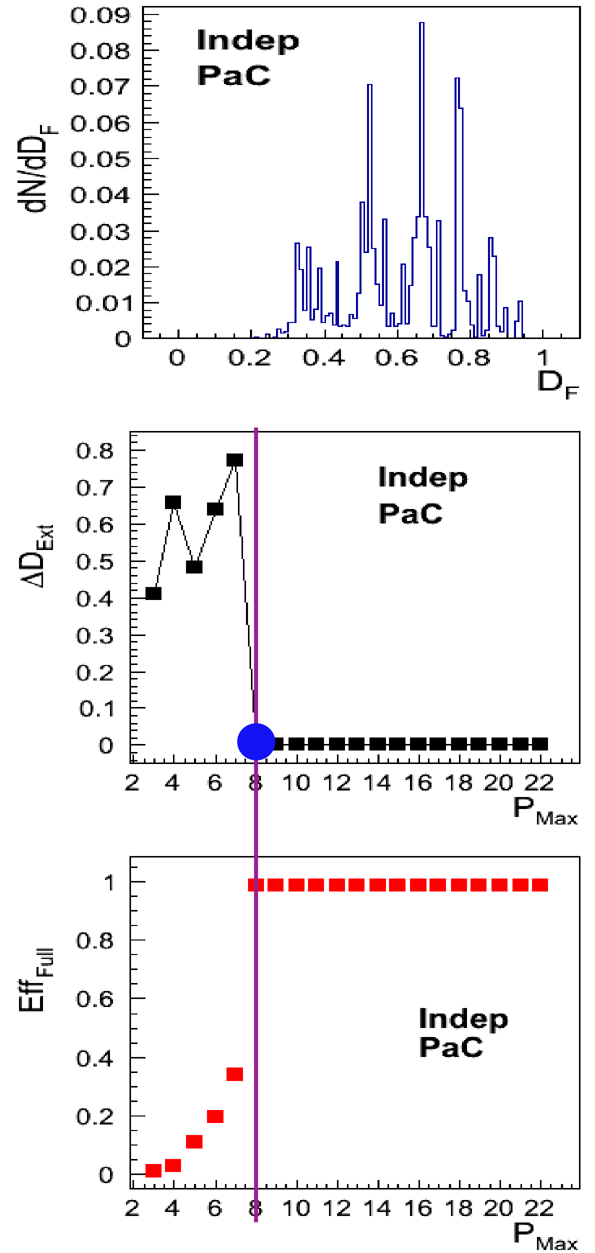
$$\Delta D_{Ext}(P_{Max}) = \Delta D_{D_F}(P_{Max}) + \Delta D_{N_{lev}}(P_{Max}) + \Delta D_P(P_{Max})$$

Preliminary analysis and result

Comparison of dependences $\Delta D_{Ext}(P_{Max})$ and $Eff_{Full}(P_{Max})$

The plateau $\Delta D_{Ext}(P_{Max})$ corresponds to the maximum $Eff_{Full}(P_{Max})$

4. P_{Max} is defined as the minimum value on a plateau of $\Delta D_{Ext}(P_{Max})$



Search procedure of optimal value parameter P_{Max} is developed for PaC and SePaC methods

Parameter χ^2_{lim} of BC, PaC

Search procedure of parameter χ^2_{lim}

1. Construction of D_F, N_{lev}, P distributions for different χ^2_{lim} at optimal P_{Max}
2. Calculation of function $\Delta D_V(\chi^2_{lim})$ of a difference of distributions

   for V=D_F, N_{lev}, P.

$$\Delta D_V(\chi^2_{\lim}) = \sum_{i=1}^{i=N_{bin}} |a_i - b_i|$$

   a_i and b_i – bin content for adjacent distribution ($\chi^2_{lim} = \chi^2_i$ and χ^2_{i+1})

3. Calculation extended function $\Delta D_{Ext}(\chi^2_{lim})$

$$\Delta D_{Ext}(\chi^2_{\lim}) = \Delta D_{D_F}(\chi^2_{\lim}) + \Delta D_{N_{lev}}(\chi^2_{\lim}) + \Delta D_P(\chi^2_{\lim})$$

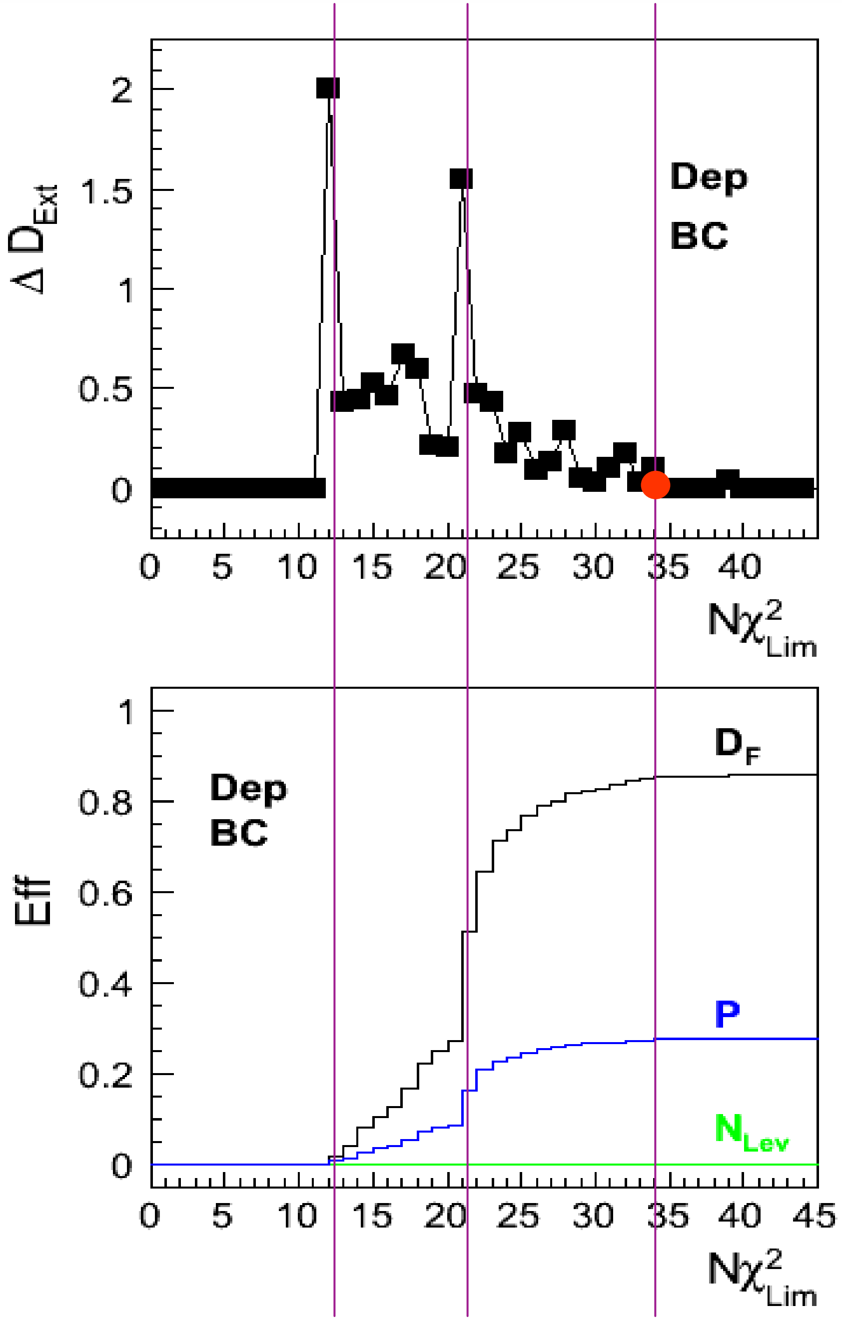
4. Choice of χ^2_{lim} is deremined by preliminary analysis results:

   comparison of $\Delta D_{Ext}(\chi^2_{lim})$ and $Eff_V(\chi^2_{lim})$ dependences

Correspondence of number $N\chi^2_{lim}$ and value χ^2_{lim}

$N\chi^2_{lim}$	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

Choice χ^2_{lim} for BC

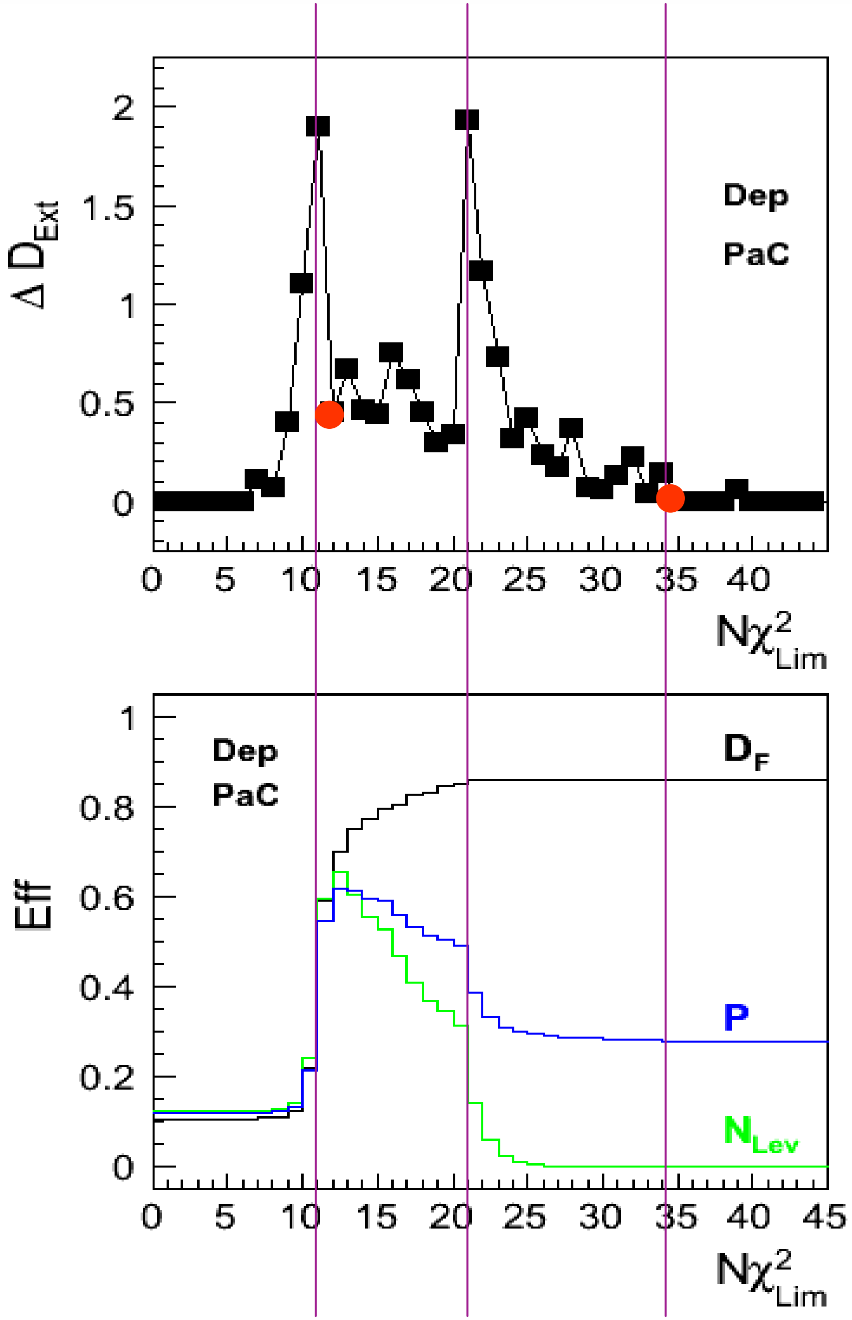


- The first plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ zero value Eff_{DF}, Eff_P and Eff_{Nlev}
- The first peak of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ the first is not zero value Eff_{DF}, Eff_P
- The second peak of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ rapid growth Eff_{DF}, Eff_P
- The second plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ the maximum Eff_{DF}, Eff_P

χ^2_{lim} is defined as the minimum value on a second plateau of $\Delta D_{Ext}(N\chi^2_{lim})$

Search procedure of optimal value χ^2_{lim} parameter is developed for BC methods

Choice χ^2_{lim} for PaC



- The first plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ small value Eff_{DF}, Eff_P and Eff_{Nlev}
- The first peak of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ rapid growth Eff_{DF}, Eff_P and Eff_{Nlev}
- The second peak of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ rapid decrease Eff_P and Eff_{Nlev}
- The range between peaks of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ max Eff_P and Eff_{Nlev}
- The second plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ ~ max Eff_{DF}

χ^2_{lim} is defined as the minimum value between peaks of $\Delta D_{Ext}(N\chi^2_{lim})$ for determination P and N_{lev}

χ^2_{lim} is defined as the minimum value on a second plateau of $\Delta D_{Ext}(N\chi^2_{lim})$ for determination D_F

Search procedure of optimal value χ^2_{lim} parameter is developed for PaC methods

Parameter *Dev* of SePaC

Search procedure of deviation from an average *Dev*

1. Construction D_F, N_{lev}, P distributions for different *Dev* at optimum P_{Max}
2. Calculation function ΔD_V (Dev_{lim}) of a difference of distributions

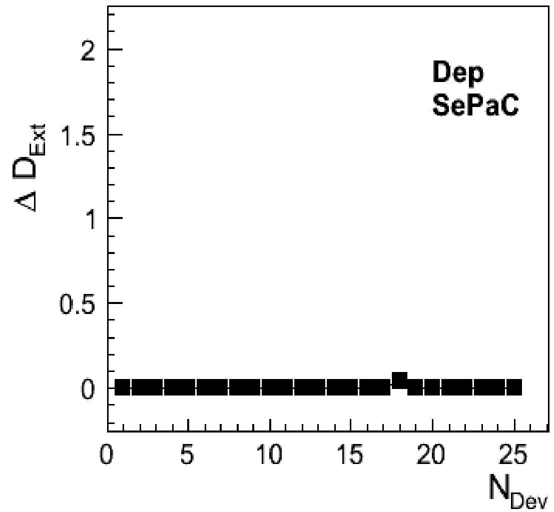
   for each V=D_F, N_{lev}, P

   here a_i and b_i – bin content for adjacent distribution

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

3. Calculation extended function $\Delta D_{Ext}(Dev_{lim})$

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

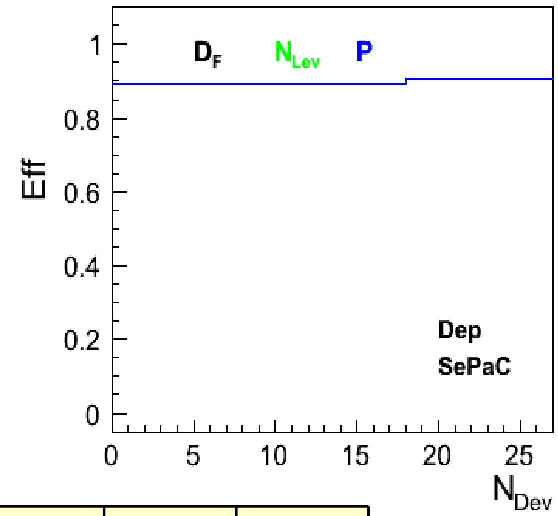


Preliminary analysis and result

Comparition of features of dependences $\Delta D_{Ext}(Dev)$ and $Eff_V(Dev)$:

Negligible change of $\Delta D_{Ext}(Dev)$ ~ negligible change of Eff_V

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



Correspondence of number N_{Dev} and value *Dev*

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

Comparison of BC, PaC, SePaC methods

Fractals with independent partition

Efficiency for optimal values of parameters

Method	Eff_{DF}	Eff_{NLev}	Eff_P	Eff_{Full}	% event Err_{DF}<0.5%	% event Err_{NLev}<0.5%	% event Err_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%

- BC: $Eff_{DF} \approx 90$ %
  D_F is precisely restored for small number of events (16%)
  N_{Lev} and P are not precisely restored
- PaC and SePaC: restoration of D_F, N_{Lev}, P with high efficiency (99-100 %)
  $D_{F,}$ N_{Lev}, P is precisely restored for large number of events (93-100%)

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

Comparison of BC, PaC, SePaC methods

Fractals with dependent partition

Efficiency for optimal values of parameters

Method	Eff_{DF}	Eff_{NLev}	Eff_{P}	Eff_{Full}	% event Err_{DF}<0.5%	% event Err_{NLev}<0.5%	% event Err_{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%

- BC: $Eff_{DF} \approx 85$ %
  D_F is precisely restored for very small number of events (2.5%)
  N_{Lev} and P are not restored
- PaC: $Eff_{DF} \approx 86$ % , $Eff_{NLev} \approx 65$ % and $Eff_{P} \approx 61$%
  D_F is precisely restored for very small number of events (1.5%)
- SePaC: restoration of D_F, N_{Lev}, P with high efficiency 91 %
  D_F, N_{Lev}, P are precisely restored for large number of events (89-90%)

SePaC has advantage in the analysis of fractals with depended partition

Summary

- The analysis of fractal (with independent and dependent partition) restoration by BC, PaC and SePaC methods was performed.

- Search procedure of optimal values of parameters χ^2_{lim} , P_{Max}, *Dev* for determination of fractal dimension D_F, number of levels N_{Lev}, and base P for these methods was developed.

- Comparison of BC, PaC and SePaC methods shown advantages of PaC and SePaC methods for restoration of fractals with independent partition and SePaC method for restoration of fractals with depended partition.

Thank You
for attention

P-adic coverage

The term P-adic coverage is used by analogy with the P-adic numbers

P-adic positive number is written as a series in powers of any number P

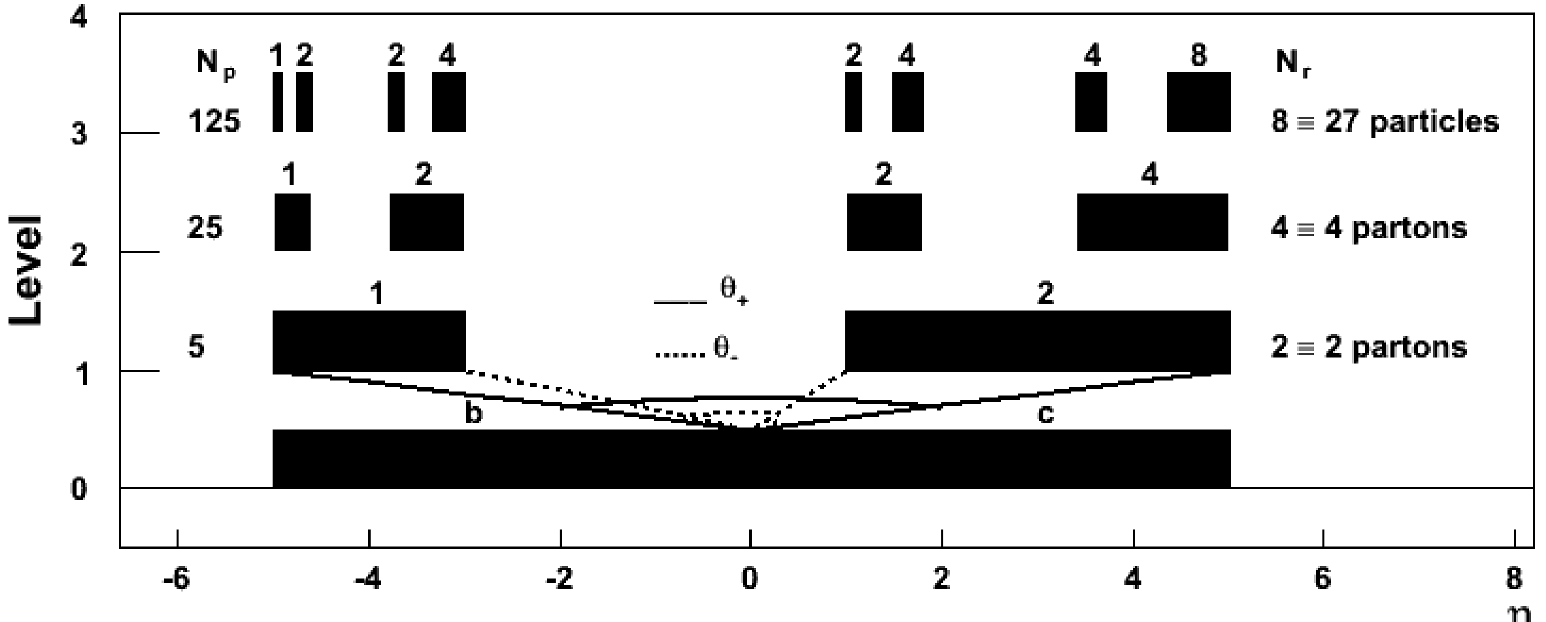
$$x = a_0 P^0 + a_1 P^1 + a_2 P^2 + a_3 P^3 + \ldots$$

P-adic coverage is a set of distributions. The number of bins M in distributions of set are changed as degree of basis P

$$M = \{P^0, P^1, P^2, P^3 + \ldots\}$$

Analysis of the optimum Coverage

regular fractal with dependent partition



Optimum Coverage

consist of a set of probes covering an object which corresponds to fractal formation.

The probes corresponds to permissible range

- System of the equation for every level

  (N-number of probes, d_i – long of the probes)

$$\sum_{i=1}^{N} d_i^{D_F} = 1$$

- The first level $(1/5)^{DF} + (2/5)^{DF} = 1$
- The second level $(1/25)^{DF} + 2*(2/25)^{DF} + (4/25)^{DF} = 1$
- The third level $(1/125)^{DF} + 3*(2/125)^{DF} + 3*(4/125)^{DF} + (8/125)^{DF} = 1$

Numerical decisions of the equations coincide each other

and define fractal dimension $D_F \approx 0.5639...$