WHAT CAN WE LEARN FROM CLAN STRUCTURE ANALYSIS ON RAPIDITY GAP PROBABILITY

A. GIOVANNINI, S. LUPIA and R. UGOCCIONI

Dipartimento di Fisica Teorica, Università di Torino and INFN, Sezione di Torino, via Giuria 1, 10125 Torino, Italy

ABSTRACT

Clan concept is generalized to the whole class of Compound Poisson Distributions which are the natural framework for studying multiplicity distributions of a dynamical two steps production process. It is shown that the probability to detect no particle in a given rapidity interval has an intriguing connection with the clan concept as well as with n-order normalized factorial cumulants. The need for analyzing data on multiplicity distributions in rapidity intervals both at HERA and TEVATRON is pointed out.


E-mail addresses: giovannini@to.infn.it, lupia@to.infn.it, ugoccioni@to.infn.it

Work supported in part by M.U.R.S.T. (Italy) under Grant 1994
WHAT CAN WE LEARN FROM CLAN STRUCTURE ANALYSIS ON RAPIDITY GAP PROBABILITY

A. GIOVANNINI*, S. LUPIA and R. UGOCCIONI
Dipartimento di Fisica Teorica, Università di Torino and INFN, Sezione di Torino, via Giuria 1, 10125 Torino, Italy

ABSTRACT

Clan concept is generalized to the whole class of Compound Poisson Distributions which are the natural framework for studying multiplicity distributions of a dynamical two steps production process. It is shown that the probability to detect no particle in a given rapidity interval has an intriguing connection with the clan concept as well as with n-order normalized factorial cumulants. The need for analyzing data on multiplicity distributions in rapidity intervals both at HERA and TEVATRON is pointed out.

1. Clan structure and rapidity gap probability

One of the most stable result in multiparticle dynamics is the two steps nature of the particles production process: the initial independent emission of certain objects called in the literature clans, clusters, strings, multiperipheral diagrams... is followed in the second step by their decay according to a QCD-inspired self similar Markov branching process dominated by gluon self-interaction.

This talk is an attempt to answer the following questions.
I. Which is the most general framework for describing final particles multiplicity distributions (MD's) in full phase space and in rapidity intervals according to the just mentioned two steps process?
II. How does the resulting general structure of final particles MD's influence related properties of n-particle correlation functions?
III. What are the implications of answers to question I. and II. on multiparticle phenomenology and in particular on the study of rapidity gap probability?

Before answering to the above three questions some remarks are needed. It should be reminded that the two steps nature of the production process has been anticipated both at hadron and parton level by clan structure analysis1.

At hadron level multiparticle production is described indeed by a two steps process controlled by the average number of clans, $N$, [in the first step], and by the average number of particles contained in an average clan, $\bar{n}_e$ [in the second step].

* Talk presented by A. Giovannini
How sharp is this new parametrization is witnessed by the fact that it provides an interpretation of the wide occurrence of approximate NB regularity and reveals new striking regularities in the behavior of the average number of produced clans in a fixed rapidity interval (see Ref. 2). Notice that these regularities are considered still puzzling in the sense that no convincing theoretical explanation of their occurrence has been given so far.

At parton level, the two steps nature of the production process has been anticipated by the discovery that parton shower evolution in virtuality \( y = \frac{A}{N_{c} - 2N_{f}} \log \frac{1}{\epsilon} \) can be described in terms of a QCD Markov branching process controlled by gluon bremsstrahlung by an initial quark, \( A = N_{c}^{A} - 1 \), by gluon self-interaction, \( A = N_{c}^{A} \) (the dominant mechanism), and gluon fission into quarks \( B = N_{f}^{A} \) (to be neglected in the first approximate description of the process) \((\epsilon \ll 1\) fixed infrared cutoff, \(N_{c}\) and \(N_{f}\) the number of colors and flavors of the theory respectively). In this framework, which is limited to the evolution of the parton shower in virtuality only, clans can be identified with gluon jets independently emitted via bremsstrahlung by an initial quark. The two steps nature of the production process is revealed by the peculiar dependence of clan structure parameters \( \bar{N} \) and \( \bar{n}_{c} \) on the dominant QCD vertices \( A \) and \( A \). \( \bar{N} \) depends in fact on \( A \) only and \( \bar{n}_{c} \) on \( A \) only:

\[
\bar{N} = \bar{A} y \\
\bar{n}_{c} = \frac{1}{\bar{A} y} (e^{\bar{A} y} - 1)
\]

(1)

The two steps occurring in clans emission and decay separate the dependence on QCD vertices \( A \) and \( A \) which are on the contrary mixed up in the standard parametrization of the MD:

\[
\bar{n} = \frac{\bar{A}}{\bar{A}} (e^{\bar{A} y} - 1) \\
k = \frac{\bar{A}}{\bar{A}}
\]

(2)

where \( \bar{n} \) is the average multiplicity and \( k \) is related to the dispersion by the relation \( D^{2} = \bar{n} + \bar{n}^{2}/k \).

In view of these remarks clan structure analysis is expected to play an important role in the generalized scheme of multiparticle production which we are looking for.

**Answer to question I.** It turns out that any distribution belonging to the class of Compound Poisson Distributions (CPD) or discrete Infinitely Divisible Distributions can be used to describe a two steps process both in full phase space and in rapidity intervals, at hadron as well as at parton level. In addition, clan structure analysis can be generalized to the whole class of CPD's to which Negative Binomial (NB) MD is belonging. **Generalized** clan structure analysis (GCSA) can indeed be easily defined in terms of the average number of generalized clans, \( \bar{N}_{g\text{-clan}} \), and the average number of particles contained in an average \( g \)-clan, \( \bar{n}_{c\text{-}g\text{-clan}} \). In fact, a CPD can be represented in full phase space by the equation

\[
I_{CPD}(z) = e^{\bar{N}_{g\text{-clan}}[g(z) - 1]}
\]

(3)
Table 1. Generating function \( f_{CPD}(z) \) of final particles MD, average number of \( g \)-clans, \( \bar{N}_g \)-clan, and generating function of particles MD inside an average \( g \)-clan, \( g(z) \), for NB, Thomas, Pólya-Aeppli distributions as a function of the average number of particles \( \bar{n} \) and the second-order normalized factorial cumulant \( \kappa_2 \) of each distribution. Generating function of the MD, average number of \( g \)-clans and generating function of the MD inside a \( g \)-clan for the PCLD are also shown in terms of its three parameters \( A, B, C \).

with

\[
g(z)\Big|_{z=0} = 0 \quad (4)
\]

where \( \bar{N}_g \)-clan is the average number of the above mentioned independent intermediate objects generated according to a Poisson MD and \( g(z) \) is the generating function of the MD of an average \( g \)-clan which contains (by definition) at least one particle. Some interesting examples of CPD’s are shown in Table 1.

In the rapidity interval \( \Delta y \), Eqs. (3) and (4) become

\[
f_{CPD}(z; \Delta y) = e^{\bar{N}_g \text{-clan}[g(z;\Delta y) - 1]}, \quad (5)
\]

with

\[
g(z; \Delta y)\Big|_{z=0} = q_0(\Delta y) \neq 0 \quad (6)
\]

where \( q_0(\Delta y) \) is the probability that a \( g \)-clan does not produce any particle within the interval \( \Delta y \) (this probability is not in general equal to zero).

Eqs. (5) and (6) can of course be rescaled in order to maintain the validity of Eq. (4) in rapidity space. One obtains:

\[
f_{CPD}(z; \Delta y) = e^{\bar{N}_g \text{-clan}(\Delta y)\tilde{g}(z;\Delta y) - 1}, \quad (7)
\]

with

\[
\tilde{g}(z; \Delta y)\Big|_{z=0} = 0 \quad (8)
\]
where \( \tilde{g}(z; \Delta y) = \frac{g(z; \Delta y) - \lambda_0(\Delta y)}{1 - \lambda_0(\Delta y)} \)

Notice that Eqs. (5) and (7) are linked at g-clan level by binomial convolution and one can express \( \tilde{N}_{g\text{-clan}}(\Delta y) \) in terms of \( N_{g\text{-clan}} \) in full phase space according to the equation:

\[
\tilde{N}_{g\text{-clan}}(\Delta y) = \tilde{N}_{g\text{-clan}}[1 - \lambda_0(\Delta y)] .
\]

It is to pointed out also that the probability to detect no particle in a given rapidity interval \( \Delta y \), \( P_0(\Delta y) \), is by definition \( f_{CPD}(z; \Delta y) \big|_{z=0} \). This fact in view of Eq. (7) establishes an intriguing link between \( P_0(\Delta y) \), n-order normalized factorial cumulants \( \kappa_n(\Delta y) \) and the average number of g-clans present in a limited region of phase space, \( \tilde{N}_{g\text{-clan}}(\Delta y) \). We propose to explore this connection in the following. Results of our search\(^5\) are summarized in three theorems.

• **Theorem 1.** The probability to detect no particle in a given rapidity interval \( \Delta y \), \( P_0(\Delta y) \), is determined\(^5\) by the average number of g-clans of the corresponding MD belonging to the class of CPD’s in the same interval, according to the equation:

\[
P_0(\Delta y) = e^{-\tilde{N}_{g\text{-clan}}(\Delta y)}
\]

Example: for the NB MD

\[
P_0(\Delta y) = e^{-k_\Delta \log(1 + \frac{\tilde{N}_{g\text{-clan}}(\Delta y)}{\tilde{N}_{g\text{-clan}}})}
\]

• **Theorem 2.** The probability to detect no particle in a given rapidity interval \( \Delta y \), \( P_0(\Delta y) \), is related\(^6\) to the n-order normalized factorial cumulants of the corresponding CPD in the same interval \( \Delta y \), \( \kappa_n(\Delta y) \), by the equation:

\[
P_0(\Delta y) = \exp \left[ -\sum_{n=1}^{\infty} \frac{\tilde{n}(\Delta y)^n}{n!} \kappa_n(\Delta y) \right]
\]

Example: for the NB MD

\[
P_0(\Delta y) = e^{-k_\Delta \log(1 + \frac{\tilde{N}_{g\text{-clan}}(\Delta y)}{\tilde{N}_{g\text{-clan}}})} = \exp \left[ -\sum_{n=1}^{\infty} \frac{\tilde{n}(\Delta y)^n}{n!} \frac{(n - 1)!}{\kappa^{n-1}(\Delta y)} \right]
\]

Theorems 1. and 2. lead just by inspection to the following Theorem 3., which says:

• **Theorem 3.** The average number of g-clans of a CPD in a given rapidity interval \( \Delta y \), \( \tilde{N}_{g\text{-clan}}(\Delta y) \), is directly connected to the n-order normalized
factorial cumulants of the same distribution in the interval $\Delta y$, $\kappa_n(\Delta y)$, by the equation:

$$\bar{N}_{g\text{-clan}}(\Delta y) = -\sum_{n=1}^{\infty} \frac{[-\bar{n}(\Delta y)]^n}{n!} \kappa_n(\Delta y)$$

(14)

Example: for the NB MD

$$\bar{N}_{g\text{-clan}}(\Delta y) = k(\Delta y) \log \left( 1 + \frac{\bar{n}(\Delta y)}{k(\Delta y)} \right)$$

$$= -\left[ \sum_{n=1}^{\infty} \frac{[-\bar{n}(\Delta y)]^n}{n!} \frac{(n-1)!}{k^{n-1}(\Delta y)} \right]$$

(15)

One comment is needed. Since it has been shown in Ref. 6 that n-particle probability in a rapidity interval $\Delta y$, $P_n(\Delta y)$, is related to zero-particle probability, $P_0(\Delta y)$, by the equation

$$P_n(\Delta y) = \frac{(\bar{n}(\Delta y))^n}{n!} \frac{\partial^n P_0(\Delta y)}{\partial n^0(\Delta y)^n}$$

(16)

the knowledge of $P_0(\Delta y)$ or, for a CPD, of the corresponding $\bar{N}_{g\text{-clan}}(\Delta y)$ determines $P_n(\Delta y)$. Consequently what we believe one of the main goals of Multiparticle Dynamics, i.e., an integrated description in rapidity space of MD's and n-particle correlation functions becomes now possible for the class of CPD in terms of $P_0(\Delta y)$ or $\bar{N}_{g\text{-clan}}(\Delta y)$. This fact is striking per se.

Answer to question II. A partial answer to this question has been anticipated in the final part of the previous answer. Here the main attention is on the second parameter of the GCSA, i.e., on the link between CPD's and hierarchical structure of the corresponding n-particle correlation functions. It is remarkable that this link can be formulated in terms of the average number of particles contained in an average g-clan, $\bar{n}_{c,g\text{-clan}}(\Delta y)$. It should be reminded (see Ref. 7,8) that n-particle correlation structure can be studied in terms of a “void scaling function” $V(\Delta y)$. It scales for hierarchical systems with energy and rapidity as a function of the product of the second order normalized factorial cumulants, $\kappa_2(\Delta y)$, times the average number of particles $\bar{n}(\Delta y)$. It is amazing that $V(\Delta y)$ turns out to be just the inverse of $\bar{n}_{c,g\text{-clan}}(\Delta y)$:

$$V(\Delta y) = -\log \frac{P_0(\Delta y)}{\bar{n}(\Delta y)} = \frac{1}{\bar{n}_{c,g\text{-clan}}(\Delta y)}$$

(17)

Example: for the NB MD

$$V(\Delta y) = \frac{1}{\bar{n}(\Delta y)\kappa_2(\Delta y)} \log (1 + \bar{n}(\Delta y)\kappa_2(\Delta y))$$

(18)
Figure 1. Rapidity gap probability $P_0(\Delta y)$ as a function of the rapidity width $\Delta y$ obtained from NB fits in $hh$ collisions at different c.m. energies $\sqrt{s} = 22$ GeV, $\sqrt{s} = 200$ GeV and $\sqrt{s} = 546$ GeV.

with $\kappa_2(\Delta y) = 1/k(\Delta y)$.

In conclusion, it can be stated that for a CPD the first parameter of the related GCSA, $N_{\text{clan}}(\Delta y)$, determines the rapidity gap probability of the corresponding MD and the related $n$-order normalized factorial cumulants. Eventual hierarchical properties of the $n$-order factorial cumulants are revealed by the second parameter of GCSA, $\bar{n}_{\text{clan}}(\Delta y)$. Clan structure analysis, firstly introduced in order to interpret the wide occurrence of NB regularity in multiparticle production and considered for a long time just a new set of parameters to be added to the standard ones, can be generalized to the whole class of CPD’s and controls the voids distribution in full phase space, in limited regions of it and the related $n$-particle correlation properties.

Answer to Question III. When one tries to study the impact of previous considerations in multiparticle phenomenology, one should of course deal with NB MD. Accordingly, one will find that for minimum bias events

- $P_0(\Delta y)$ decreases almost linearly with $\Delta y$ for small intervals (for given reaction, fixed energy);
- $P_0(\Delta y)$ then levels toward a constant value as $\Delta y$ approaches full phase space due to kinematical boundaries;
6. \( P_0(\Delta y) \) is approximately energy independent in the region of linearity (at fixed \( \Delta y \));

6. \( P_0 \) in full phase space decreases with energy;

6. \( P_0(\Delta y) \) is steeper in e\(^+\)e\(^-\) than in hh or lh.

These results are a summary of the information contained in Figures 1, 2 and 3 for hadron-hadron\(^9\), deep inelastic scattering\(^10\) and e\(^+\)e\(^-\) annihilation\(^11,12\) respectively.

Comments:

- expected behavior of \( P_0(\Delta y) \) is linear in \( \Delta y \) for the soft gluon radiation\(^13\). It is found in experimental data that the behavior is linear just in small rapidity intervals, then it levels. A similar trend is seen by D0 Collaboration\(^14\) and CDF Collaboration\(^15\) for dijet events with transverse energy of each jet larger than 30 GeV (i.e. a typical hard process\(^13\)). Clan structure analysis could be useful at Tevatron for a deeper understanding of D0 and CDF results: it could estimate the soft background contribution (see also Ref.\(^{16}\)).

- Assuming that the resulting MD for dijet events is also of CPD type, one can determine rapidity gap probability \( P_0(\Delta y) \) in different rapidity intervals by applying GCSA to the MD. This analysis should be contrasted with the standard one based on the search of regions of phase space without any
Figure 3. Rapidity gap probability $P_0(\Delta y)$ as a function of the rapidity width $\Delta y$ obtained from NB fits to the sample of two-jet events in $e^+e^-$ annihilation at c.m. energies $\sqrt{s} = 29$ GeV and $\sqrt{s} = 91$ GeV.

particle. Notice that statistical errors are larger in this second case than in the first one.

- In deep inelastic scattering, the average number of clans has the same behavior as in $hh$ collisions, but clan size is much smaller than in $hh^{10}$. This property could be interesting for interpreting the excess of events with a large rapidity gap and small multiplicity seen at HERA$^{17,18}$.

- In $e^+e^-$ annihilation, the average number of clans when plotted as a function of the rapidity variable has a higher slope than in $hh$, i.e., a steeper slope for the rapidity gap probability. The analysis is here limited to the sample of two-jet events$^{11,12}$. In this case we do not expect rapidity gaps to be filled by the fragmentation of the initial state remnants. This reaction is therefore a "clean" environment to study the physics of rapidity gaps and could be used to investigate the two-jet events contribution to the background in the search proposed in Ref. $^{19}$.

Therefore, in all these cases the study of MD's in rapidity intervals could provide interesting information on the general structure of the reaction. One sees stronger correlation in $hh$ reactions than in $e^+e^-$ annihilation, since $e^+e^-$ annihilation data on rapidity gap probability are near to Poissonian limit, whereas $hh$ data show large deviations from Poisson.
| Rapidity interval: $|y| \leq 1.0$ | Rapidity interval: $|y| \leq 1.5$ |
|-----------------------------|-----------------------------|
| **Reaction**                | **$1/k$**                   | **Reaction**                | **$1/k$**                   |
| $\bar{p}p$ 900 GeV          | 0.59±0.02                   | $\bar{p}p$ 900 GeV          | 0.54±0.02                   |
| $e^+e^-$ 2 jet 29 GeV       | 0.095±0.004                 | $e^+e^-$ 2 jet 29 GeV       | 0.063±0.004                 |
| $e^+e^-$ 2 jet 91 GeV       | 0.130±0.002                 | $e^+e^-$ 2 jet 91 GeV       | 0.127±0.003                 |

Table 2. Comparison in two different rapidity intervals of the aggregation parameter $1/k$ obtained from NB fits for the reactions indicated in the Table.

An interpretation of the above results can be given in terms of MD's by recalling that $1/k$ can be identified with an aggregation parameter:

$$\frac{1}{k} = \frac{P(n=2, N=1)}{P(n=2, N=2)}$$

where $P(n=2, N=1)$ and $P(n=2, N=2)$ are the probabilities for a produced particle to join its partner in the same clan or to form a second clan respectively; $k$ is the standard NB parameter. See Table 2 where a comparison is presented in two different rapidity intervals of the values of the aggregation parameter $1/k$ obtained from NB fits in $\bar{p}p$ and $e^+e^-$ annihilation reactions. It is shown that the aggregation is larger in $\bar{p}p$ than in $e^+e^-$ annihilation and that it grows with energy in $e^+e^-$ annihilation. These facts agree with the observed behavior for the rapidity gap probabilities in the two cases.

2. Conclusions

Clan structure analysis is common to the whole class of CPD's. It characterizes all two steps processes both at hadron and at parton level. The rapidity gap probability allows an integrated description of n-particle correlation functions and MD's in rapidity intervals. The rapidity gap probability for the whole class of CPD's is fully determined by the average number of generalized clans. The study of $P_0(\Delta y)$ in terms of $N_{g-class}(\Delta y)$ starting from MD's in rapidity intervals can provide useful information on minimum bias events and as background of two-jet hard scattering events at Tevatron, as background in $e^+e^-$ annihilation and at HERA. These considerations suggest to study MD's in rapidity intervals, in particular at
TEVATRON and HERA; this analysis will allow in fact to determine clan structure parameters and consequently rapidity gap probability and $n$-order factorial cumulants properties in an elegant way.

3. References