CRN

STRASBOURG

CRN 93-32

Correction of measured multiplicity distributions by the simulated annealing method

Mohamed HAFIDOUNI

CENTRE DE RECHERCHES NUCLEAIRES

STRASBOURG

IN2P3
CNRS

UNIVERSITE
LOUIS PASTEUR

Correction of measured multiplicity distributions by the simulated annealing method

M. Hafidouni*

Centre de Recherches Nucléaires
B.P. 20, 67037 Strasbourg, France

Abstract. Simulated annealing is a method used to solve combinatorial optimization problems. It is used here for the correction of the observed multiplicity distribution from S-Pb collisions at 200 GeV/c per nucleon.

Introduction.

Because of limited geometrical acceptance and efficiency of detectors, observed distributions, in general, differ from the real distributions. The goal of data analysis in an experiment is to recover the true (or real) distributions starting from the observed (or measured) ones.

The usual procedure of correcting observed distributions starts from simulating detector geometry and electronic response. A generated distribution is passed through the simulation codes yielding a 'measured' distribution. The generated distribution is usually not of major importance. In the following multiplicity distributions are discussed.

Dividing the observed distribution O into M bins and the true one, T, into N bins, one has the following relation [1,2]

$$O_m = \sum_{n=1}^{N} P_{mn} T_n \qquad (1)$$

where O_m ($m = 1, ..., M$) is the fraction of detected events with observed multiplicity in the bin m and T_n is the fraction of events with true multiplicity in the bin n; P_{mn} is the probability that an event

* crnvax::hafidouni or hafidouni@crnvax.in2p3.fr

with a true multiplicity in the bin n be observed as an event with observed multiplicity in the bin m. Consequently, the P_{mn}'s satisfy the following relation:

$$\sum_{m=1}^{M} P_{mn} = 1. \tag{2}$$

The matrix P is determined by simulation, where one knows O and T.

The underconstrained system of equations (1) cannot be solved in a straightforward way; it may lead, e.g., to negative values of T_n. One must rather look for a solution $(T_1, ..., T_M)$ that describes the data well in a statistical sense which means that the differences $\left|O_m - \sum_{n=1}^{N} P_{mn} T_n\right|$ are of the order of the corresponding statistical errors σ_m.

A method to choose one probability distribution from a set of distributions that are compatible with the data consists in maximizing the Shannon entropy [3]

$$S = -\sum_{n=1}^{N} t_n \; Log \; t_n, \qquad t_n = \frac{T_n}{\sum_{i=1}^{N} T_i} \tag{3}$$

under constraints imposed by the data. This is known as the principle of the maximum entropy [4]. It has been shown that the maximum entropy method (MEM) is the only consistent method of inference for underconstrained problems [5].

With the method [6] as it is applied here, one tries to maximize the entropy and at the same time to minimize χ^2 given by

$$\chi^2 = \sum_{m=1}^{M} \left(\frac{O_m - \sum_{n=1}^{N} P_{mn} T_n}{\sigma_m} \right)^2 \tag{4}$$

where σ_m is the statistical error. This is equivalent to minimizing the expression

$$F = -S + \lambda \, \chi^2 \tag{5}$$

where S is given by (3). The factor λ is a Lagrange multiplier which determines the relative strength between these two tendencies. One notes that F looks like the free energy (of a statistical system) divided by temperature with λ and χ^2 assuming, respectively, the role of an inverse temperature and internal energy.

Simulated annealing method.

To minimize the function F, the simulated annealing method [7] is now used. This method was developped to deal with combinatorial optimization problems of very large scale and is based on the analogy with cooling and annealing of metals.

To find the true ground state of a material, the annealing technique is of great importance. It consists in lowering slowly the temperature of a material and spending a long time at temperatures in

the vicinity of the freezing point. When cooling is completed, the material is in an ordered low energy (single crystal) state, provided that the starting temperature is high enough and the cooling process sufficiently slow. If the material is cooled quickly, it does not reach this state and the resulting crystal will have many defects.

Slow cooling means that the system is constantly in thermal equilibrium during the annealing process. For a system in thermal equilibrium at temperatute T, the probability of occurence of a state with energy E is given by the Boltzmann distribution

$$P(E) \propto exp(-E/kT) \tag{6}$$

where k is the Boltzmann constant. $P(E)$ depends on the temperature and energy of the state. At temperatures $kT \gg E$, all states are equally probable. As temperature decreases, $P(E)$ is large for the low energy states, and finally when the temperature approaches zero, only the minimum energy states have a non-zero probability.

The Metropolis algorithm [8] provides a framework for the simulation of systems in thermal equilibrium at a temperature T. The sequences of states generated are distributed according to (6). The principle is the following: the current configuration of the system is slightly disturbed and the resulting change in energy ΔE is computed. If $\Delta E < 0$ the displacement is accepted. The case $\Delta E > 0$ is treated probabistically: a random number, selected in the interval between 0 and 1 with uniform probability, is generated and compared to $P(\Delta E) = exp(-\Delta E/kT)$. If it is less than $P(\Delta E)$, the displacement is accepted otherwise the latter is rejected. The possibility of generating configurations with $\Delta E > 0$ makes the Metropolis algorithm suitable for minimization problems. These configurations help the system to jump out of local minima of the energy in favor of finding a better, more global, minimum.

To make use of the Metropolis algorithm for other than thermodynamic systems, for example combinatorial optimization problems, one must define configurations by a set of parameters, a cost function C analogous to the energy and a control parameter T analogous to temperature.

The simulated annealing method is applied to optimization problems by selecting a starting configuration at a high temperature T_0. To estimate the value of the starting temperature T_0, random configurations are generated and the corresponding values of the cost function are computed. T_0 is then chosen larger than the largest cost difference between two configurations such that all displacements are accepted. A possible choice is given by [9]

$$T_0 = \rho \times max\,|\Delta C|, \qquad \rho \sim 5 - 10 \tag{7}$$

where ΔC is the difference between two given configuration costs.

At each temperature, configurations are generated by changing one or few variables. The generated configuration is accepted as the next state of the system if it fulfils the Metropolis criterion.

The next step consists in lowering the temperature. This is done according to a cooling schedule. The most popular one is the exponential schedule where a new temperature is obtained from the preceding one by multiplication by a factor α: $0 < \alpha < 1$. The choice of the cooling schedule may have some effects on the results obtained [9].

When temperature is very low, the number of accepted configurations becomes very small. The system does not change significantly from one temperature to another: it is frozen to some configuration and no further changes occur. In practice, one considers the system as frozen when the number of accepted configurations is below some (arbitrary) threshold. The final frozen configuration is taken as the solution of the problem at hand. The simulated annealing method was applied to many problems, among which the design of complex integrated circuits [7] and image restoration [9].

Application to the multiplicity distribution.

The simulated annealing method is applied here to the correction of observed multiplicity distribution of negative particles produced in S-Pb interactions at 200 GeV/c per nucleon (see the ref. [10] for the datails of the analysis). The observed negative mutltiplicity distribution to be corrected in the following is shown in figure 1.

To apply the simulated annealing method, a configuration was defined by the (normalized) fractions of events in the true multiplicity bins $t_1, t_2, ..., t_N$ and the cost function is F given by (5). A new configuration is chosen by generating a random t_i in the interval between 0 and 1 and then normalizing all the t_n's to match the condition (3). This configuration is then accepted or rejected according to Metropolis criterion. Next, this step is repeated with t_{i+1} and so on. At each temperature, 9750 configurations are tried; 75 configurations per each bin. The number of temperatures used was 350; the temperature was lowered by exponential cooling. At the frozen configuration the $\chi^2 \approx M$.

Figure 2 shows the corrected multiplicity distribution obtained after 3.4125×10^6 iterations. The error bars are purely statistical.

The mean multiplicity, dispersion $D_- = \sqrt{< n_-^2 > - < n_- >^2}$ and their ratio are

$$< n_- >= 57.19 \pm 0.39 \tag{8}$$

$$D_- = 44.23 \pm 0.83 \tag{9}$$

$$\frac{< n_- >}{D_-} = 1.293 \pm 0.026 \tag{10}$$

The corrected multiplicity distribution have the same shape as the one obtained by other minimization procedures [11]. The smoothness of the corrected distribution may be improved using better Monte-Carlo statistics to determine the matrix P and by smoothing the P_{mn}, for fixed n (see the ref.

[2]). One may also improve the smoothness by increasing the number of temperatures or the number of configurations at each temperature. Here, the number of configurations/temperature was fixed to 75 because of the limitations in the computing time; the simulated annealing method is very time consuming.

Another algorithm was also tried. It consists in changing a bin until a a proposed change is not accepted and then moving to the next bin. This algorithm gives the same results than the previous one and doesn't reduce the CPU time.

Conclusion.

Correction of the observed multiplicity in heavy ion collisions by the simulated annealing method is presented. The corrected distribution is is comparable to the distributions obtained by other methods.

Acknowledgements.

Thanks to W. Geist for reading the manuscript.

References.

[1] R.E. Ansorge et al., Z. Phys. C 43 (1989) 357.

[2] D. Decamp et al., Phys. Lett. B 273 (1991) 181.

[3] C.E. Shannon, Bell Syst. Tech. J. 27 (1948) 379. This paper is reproduced in:
D. Slepian (Editor), Key Papers in The Development of Information Theory (IEEE Press, New-York, 1974).
C.E. Shannon and W. Weaver, The Mathematical Theory of Communication (University of Illinois Press, Urbana, 1949).

[4] E.T. Jaynes, Phys. Rev. 106 (1957) 620; 108 (1957) 171.

[5] J.E. Shore and R.W. Johnson, IEEE Trans. Inf. Th. IT-26 (1980) 26; IT-29 (1983) 942.
Y. Tikochinsky, N.Z. Tishby and R.D. Levine, Phys. Rev. Lett. 52 (1984) 1357.

[6] M. Schmelling, Ph.D. thesis, Dortmund University, Germany (1987); CERN preprint CERN-CN/93-03 (1993).

[7] S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi, Science 220 (1983) 671.
S. Kirkpatrick, J. Stat. Phys. 34 (1984) 975.

[8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, J. Chem. Phys. 21 (1953) 1087.
G. Bhanot, Rep. Prog. Phys. 51 (1988) 429.

[9] H. Raittinen and K. Kaski, Phys. Scripta T 33 (1990) 126.

[10] M. Hafidouni, Doctorat de l'Universite Louis Pasteur de Strasbourg, CRN/HE 92-35 (1992).

[11] M. Hafidouni, Strasbourg preprint CRN 93-28 (1993).

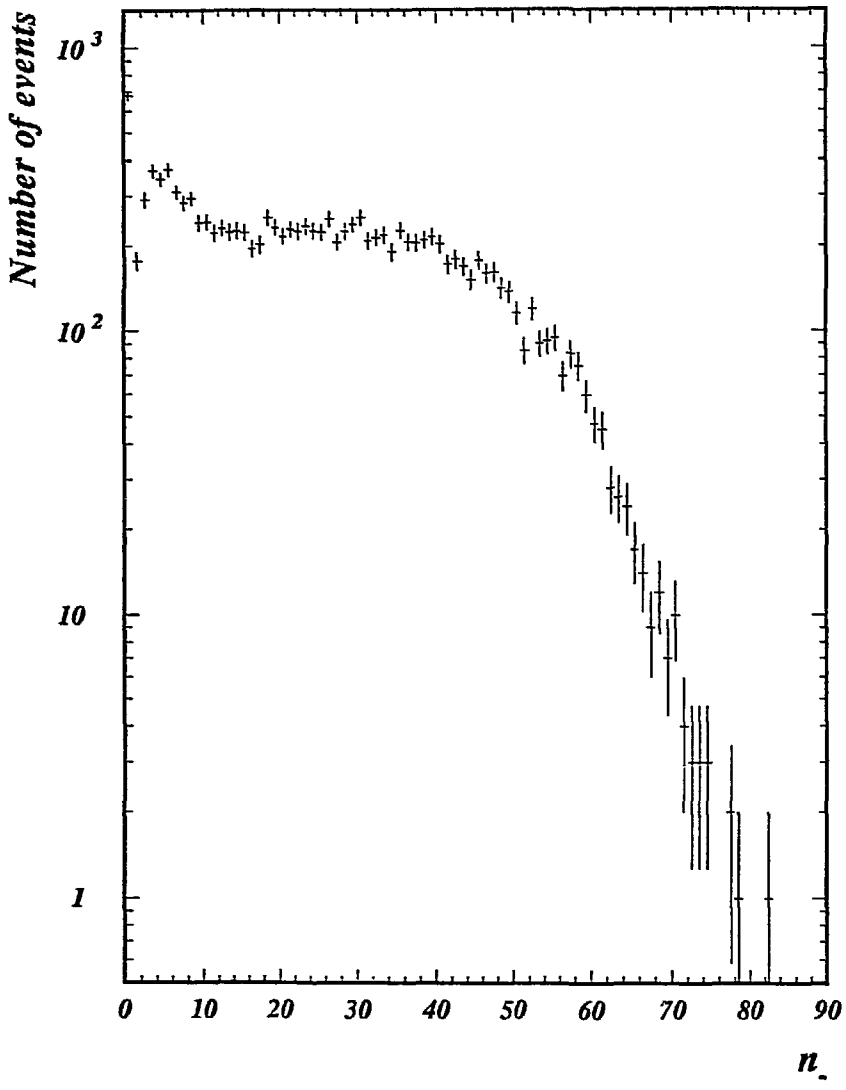

Fig. 1. Observed negative multiplicity distribution.

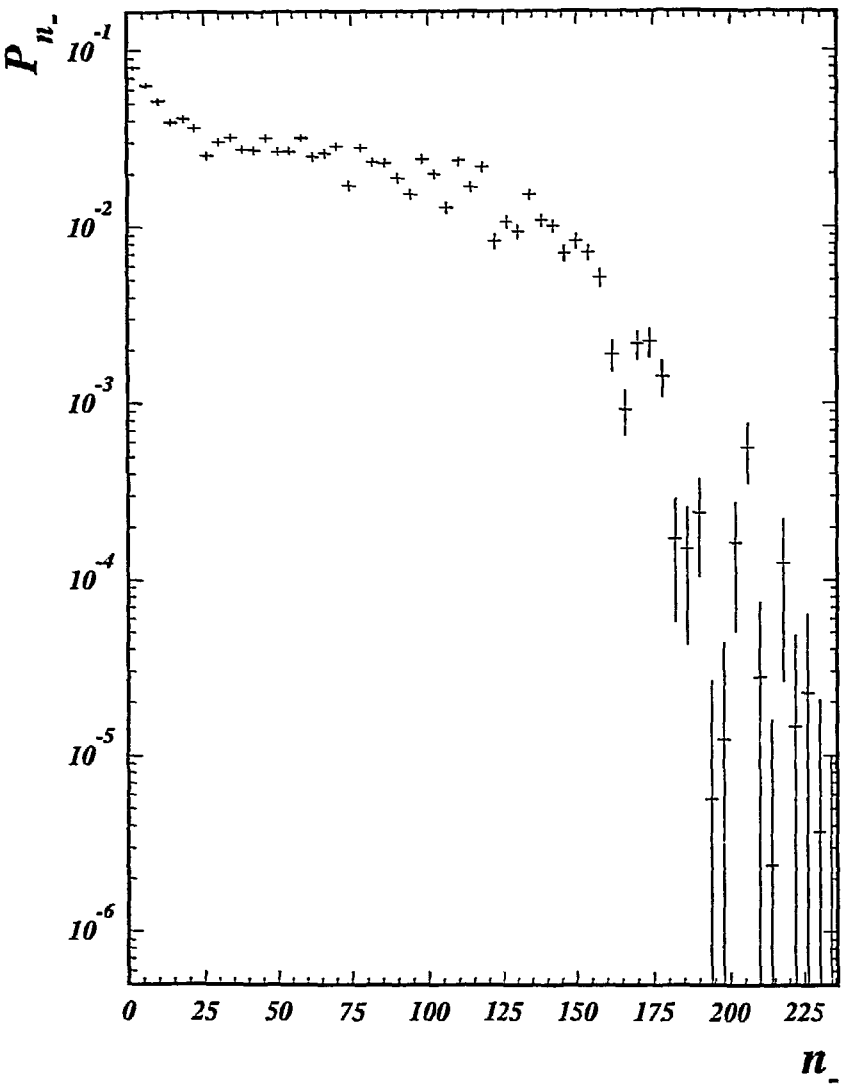

Fig. 2. Corrected negative multiplicity distribution.