LABORATOIRE DE PHYSIQUE NUCLEAIRE

UNITE ASSOCIEE AU CNRS/IN2P3 (UA 57)

2, rue de la Houssinière 44072 NANTES CEDEX 03

UNIVERSITE DE NANTES   I N2 P3

IMPACT PARAMETER DETERMINATION FOR HEAVY ION COLLISIONS BY USE OF A NEURAL NETWORK

C. David, M. Freslier, J. Aichelin

Rapport Interne LPN - 94-05

UNIVERSITE DE NANTES

IMPACT PARAMETER DETERMINATION FOR HEAVY ION COLLISIONS BY USE OF A NEURAL NETWORK

Christophe David*, Marc Freslier and Jörg Aichelin†

Laboratoire de Physique Nucléaire, CNRS/IN2P3 Université de Nantes,

2, Rue de la Houssinière, 44072 Nantes Cedex 03, France

Abstract

To determine the impact parameter of heavy ion collisions from the observables is of crucial importance for comparing the experimental results with theory. For central collisions all methods which rely on a single observable have failed. We propose to determine the impact parameter by a neural network. We find that a combination of three observables allows to determine the impact parameter four times more accurate than a single observable. We investigate in detail which combination of observables provide the best results.

PACS numbers : 25.70.-z 25.70.Np 02.70.+d

LPN 94-04 May 1994

Submitted to Phys. Rev. C

*email : "DAVID@nanhp2.in2p3.fr"

†email : "AICHELIN@nanvs2.in2p3.fr"

The study of relativistic heavy ion collisions is plagued by the complexity of the reaction. For the same combination of projectile and target a multitude of different processes takes place, depending on the impact parameter. In peripheral collisions we observe only a small energy transfer between projectile and target. Both get excited and behave like compound nuclei or undergo fission. More central reactions cause the multifragmentation of the combined system and many intermediate mass fragments can be observed. In central collisions, finally, we observe the formation of a fireball, which disintegrates mainly by emission of nucleons or very light fragments up to α particles. There is, however, no clear cut between these processes and the fluctuation of the observables are large even for a given impact parameter.

To separate these processes is, however, of crucial importance if one would like to investigate the different processes in detail and if one desires to compare it with experiment. This is especially true for the most central collisions. In these reactions the nuclei get compressed and part of the system may reach densities up to three times the normal nuclear matter density ρ_0. If one has no means to select these most central collisions all signals are washed out by the much more numerous more peripheral reactions where such a compression is not present.

Up to now in most experiments the observables have been analyzed as a function of the multiplicity of the observed charged particles. As is well known and as we will see later the binning as a function of the multiplicity can give a coarse classification into peripheral, semi central and central events but it is by no means sufficient to separate the most central reactions. Recently efforts have been made for a better selection of the central events by using the stopping and the directivity for the classification [1,2]. Although this transverse momentum analysis yields somewhat better results it is not sufficient for selecting precisely the desired impact parameters range.

In this article we will address the question of selecting the impact parameter in a more systematic way. Using the new technique of neural networks we investigate in detail which combination of observables yields the most precise determination of the impact parameter.

Up to now there exist only two exploratory studies of the use of neural networks in heavy ion collisions [3,4]. The authors of [4] use three input quantities , the flow angle θ, the directed momentum $p_{x\ dir}$ and zz component (Q_{zz}) of the momentum tensor. These variables can be reduced to the longitudinal momentum, the absolute value of the transverse momentum in forward rapidity and the width of the momentum distribution in z direction. With this three parameters the authors obtained only a marginal improvement of the determination of the impact parameter selection as compared with a determination using (Q_{zz}) only. As we will show this choice of variables is not optimized, because both, θ and $p_{x\ dir}$ approach zero for the central collisions and therefore contribute little to an exact determination.

In our approach we use standard neural network techniques which are sufficient for this purpose. We limit ourselves to "theoretical data", i.e. simulated events where the impact parameter is known and investigate to which degree the impact parameter is reproduced by the network. The application to experiments is straight forward, however it requires a detailed knowledge of the acceptance of the detector which is used. When the training of the network is performed with these filtered "theoretical" data it can be directly applied to experiments.

For our investigation we apply the Quantum Molecular Dynamics Model (QMD) which has been successfully used in the last years to simulate heavy ion reactions [5]. In this model the nucleons are represented by Gaussian wave functions which move under the influence of mutual interactions given by the Brückner G- Matrix. The real part of the G - matrix acts like a density dependent two body force between the nucleons. The imaginary part can be formulated as a cross section. Initially the nucleons of projectile and target are located in a sphere of the radius $r = 1.14A^{1/3}$ and the initial momentum has been randomly selected in between 0 and the local Fermi momentum. For details of this approach we refer to ref.[5]. The analysis is performed with 1037 events of the reaction Au (600 MeV) + Au with an impact parameter randomly chosen between 0 and $b^{max} = 14fm$.

Before we start to discuss the neural network and the results obtained we calculate as a benchmark how well the impact parameter can be determined by use of one observable only. This analysis was done for three observables, the total multiplicity of protons (MULT) , the largest fragment observed in each collision (AMAX) and the energy ratio (ERAT) in the center of mass system defined by

$$ERAT = \frac{\sum p_t^2/2m}{\sum p_z^2/2m}. \tag{1}$$

For this purpose we define the normalized impact parameter

$$b = \frac{b^{true}}{b^{max}} \tag{2}$$

Thus the range of the reduced impact parameter is [0,1]. In bins of 0.05 we calculated the mean value and the variance of the abovementioned observables and approximated the mean value by a spline fit. The result is presented in fig. 1. We see that MULT as well as AMAX is constant for impact parameters smaller than 3.5 fm, thus they cannot be used to select the most central collisions. Only ERAT shows an appreciable dependence on the impact parameter in that impact parameter domain. In peripheral reactions AMAX approaches a constant value and the other observables show a dependence on b.

To estimate the precision of the determination of the impact parameter we invert the fit function in order to obtain the reduced impact parameter b^{fit} as a function of the observables. Then we calculate the standard deviation

$$C = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(b_i^{QMD} - b_i^{fit})^2} \quad (3)$$

for central, semicentral and peripheral events. The results are presented in table 1.

$C=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(b_i^{QMD} - b_i^{fit})^2}$	$0 \leq$ b $\leq$ 0.15	0.3 $\leq$ b $\leq$ 0.7	0.85 $\leq$ b $\leq$ 1
AMAX	0.0791	0.0673	0.0449
MULT	0.1046	0.0426	0.0532
ERAT	0.1286	0.0531	0.0813

Table 1: Uncertainty of the impact parameter determination using one observable only.

We see, as already expected from the discussion above that the results obtained for AMAX and MULT are quite satisfactory for semicentral and central collisions, giving an uncertainty of the impact parameter determination of about $C \cdot b^{max} = .7fm$. The methods fails , however, for central events. The seemingly reasonable values for the variable AMAX and MULT are of no use as discussed above. The small standard deviation is a consequence of the fact that in central collisions ($b < 3fm$) rarely a cluster with mass number larger than 5 is produced. Therefore, the variance of AMAX is small and the same is true - due to particle number conservation - for MULT. The only candidate ERAT fails badly giving an uncertainty of $C \cdot b^{max} = 1.8fm$. In fig. 2 we investigate in detail the difference between b^{fit} and b^{QMD} obtained for this variable. For this purpose we ordered the events corresponding to their impact parameter and plotted for each simulation the difference between b^{fit} and b^{QMD}. We see that the huge fluctuations in central collisions are the cause for this failure.

Although one may think of other observables there is the general tendency that with one observable only it seems to be impossible to select central collisions.

Next we will show that an artificial neural network which makes use of several variables improves the situation. We start with a short description of the neural network we use. For details and the general background we refer to ref [6]. In our calculation we have used a feed-forward network with three layers of neurons. The three layers consist of an input layer of three cells, a hidden layer of five cells and a output layer of one cell (fig.3). The input layer receives data from outside (i.e. the values for the observables) and the output layer gives the result (i.e. the impact parameter) whose difference to the known impact parameter has to be minimized.

For the learning phase we use the method of error backpropagation for the change of the synaptic connections in order to obtain agreement between impact parameter calculated by the network and the known impact parameter of the QMD simulations.

For the detailed description we use the following notations :

k ——————- Input

$\bar{w}_{jk}$,$\bar{\theta}_j$,$\bar{h}_j$ $\Downarrow$

j ——————- Hidden (state: s_j)

w_{ij},θ_i,h_i $\Downarrow$

i ——————- Output (state: S_i)

with i = 1 and

$\bar{w}_{jk}$ synaptic connections between input cells and hidden cells
$\bar{\theta}_j$ threshold of hidden cells
w_{ij} synaptic connections between hidden cells and output cell (i=1)
θ_i threshold of output cell
h_i output i-cell activation
$\bar{h}_j$ hidden j-cell activation
f activation function
s_j output function of hidden j-neuron
S_i output function of output i=1 - neuron

The equations governing the state of the network are

$$S = f(h_1) \quad , \quad h_1 = \sum_j \omega_{1j} s_j - \theta_1 \tag{4}$$

$$s_j = f(\bar{h}_j) \quad , \quad \bar{h}_j = \sum_k \bar{\omega}_{jk} \sigma_k - \bar{\theta}_j \tag{5}$$

The activation function of the hidden neurons is given by f(t)=$\frac{1}{2}$[1+tanh(t)].

The aim of the training phase of the network is to minimize the output function D, i.e. the difference between the known impact parameter and that determined by the network,

$$D[\omega_{ij}, \theta_i, \bar{\omega}_{jk}, \bar{\theta}_j] = \frac{1}{2} \sum_\mu \sum_i [\xi_i^\mu - f(h_i^\mu)]^2 \tag{6}$$

by varying the activations and the synaptic connections. For this purpose we have to evaluate the gradient of D with respect to every variable (synaptic connections and thresholds) to minimize D. For finding the minimum of D we applied the conjugate gradient method as well as a subroutine of the NAGLIB library. The results obtained are identical.

The procedure to find the minimum consists of two steps:

In the first step we vary only synaptic connections of the output neurons ($\omega_{ij} \rightarrow \omega_{ij} + \delta\omega_{ij}$):

$$\delta\omega_{ij} = -\varepsilon \frac{\partial D}{\partial \omega_{ij}} = \varepsilon \sum_\mu [\xi_i^\mu - f(h_i^\mu)] f'(h_i^\mu) \frac{\partial h_i^\mu}{\partial \omega_{ij}} = \varepsilon \sum_\mu \Delta_i^\mu s_j^\mu \tag{7}$$

$$\delta\theta_i = -\varepsilon \frac{\partial D}{\partial \theta_i} = \varepsilon \sum_\mu [\xi_i^\mu - f(h_i^\mu)] f'(h_i^\mu) \frac{\partial h_i^\mu}{\partial \theta_i} = -\varepsilon \sum_\mu \Delta_i^\mu \tag{8}$$

with the abbreviation

$$\Delta_i^\mu = [\xi_i^\mu - f(h_i^\mu)] f'(h_i^\mu) \tag{9}$$

The parameter ε is the intensity of the learning phase. It must be chosen to ensure optimal convergence when eqs. 7-15 are applied in a repetitive way. We determine the value of ε in each iteration step as that value which minimize D with respect to ε, i.e. $\frac{\delta D}{\delta \varepsilon} = 0$. This garanties a fast convergence to the stable solution.

In the second step we vary the variables associated with the connection between the input and hidden layer ($\delta\bar{\omega}_{jk}$).

$$\delta\bar{\omega}_{jk} = -\varepsilon \frac{\partial D}{\partial \bar{\omega}_{jk}} = \varepsilon \sum_{\mu,i} [\xi_i^\mu - f(h_i^\mu)] f'(h_i^\mu) \frac{\partial h_i^\mu}{\partial s_j} \frac{\partial s_j}{\partial \bar{\omega}_{jk}} \tag{10}$$

$$\delta\bar{\omega}_{jk} = \varepsilon \sum_{\mu,i} \Delta_i^\mu \omega_{ij} f'(\bar{h}_j^\mu) \frac{\partial \bar{h}_j}{\partial \bar{\omega}_{jk}} \tag{11}$$

$$\delta\bar{\omega}_{jk} = \varepsilon \sum_{\mu} \bar{\Delta}_j^\mu \sigma_k^\mu \tag{12}$$

$$\delta\bar{\theta}_j = -\varepsilon \frac{\partial D}{\partial \bar{\theta}_j} = \varepsilon \sum_{\mu,i} [\xi_i^\mu - f(h_i^\mu)] f'(h_i^\mu) \frac{\partial h_i^\mu}{\partial s_j} \frac{\partial s_j}{\partial \bar{\theta}_j} \tag{13}$$

$$\delta\bar{\theta}_j = \varepsilon \sum_{\mu,i} \Delta_i^\mu \omega_{ij} f'(\bar{h}_j^\mu) \frac{\partial \bar{h}_j}{\partial \bar{\theta}_j} \tag{14}$$

$$\delta\bar{\theta}_j = -\varepsilon \sum_{\mu} \bar{\Delta}_j^\mu \tag{15}$$

with the new abbreviation

$$\bar{\Delta}_i^\mu = (\sum_i \Delta_i^\mu \omega_{ij}) f'(\bar{h}_j^\mu) \tag{16}$$

The training proceeds in iterations.

The three input observables are chosen from the following set of observables

The mass of the largest fragment (AMAX)

The multiplicity of intermediate fragment (IMF)

The multiplicity of protons (MULT)

The flow of particles $= \frac{1}{N} \sum sign(y_{cm}) \cdot p_x$, x being the direction of the impact parameter (FLOW)

The directivity (DIR) $= \frac{|\sum \vec{p}_t|}{\sum |\vec{p}_t|}$

The energy ratio in the center of mass system (eq.1) (ERAT)

FSD defined as

$$\frac{N_c}{N_c - 1} \cdot \frac{(\sum \vec{p}_t)^2 - \sum p_t^2}{(\frac{p_{proj}}{a_{proj}} \cdot \sum A)^2} \tag{17}$$

We try several combinations of three observables among the seven above and present each event 10000 time at the network. The quality of the network response is quantified by the standard deviation

$$C = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (b_i^{QMD} - b_i^{network})^2} \tag{18}$$

with

b_i^{QMD} QMD impact parameter

$b_i^{network}$ impact parameter estimated by the network

N number of events

The time evolution of the standard deviation is displayed in fig. 4 for a typical example. We see that after 5000 iteration the network has practically reached its asymptotic value.

To see how well the different impact parameter ranges are reproduced we calculated the standard deviation for three intervals $0 \leq b^{QMD} \leq 0.15$ $0.3 \leq b^{QMD} \leq 0.7$ $0.85 \leq b^{QMD} \leq 1$ separately.

We come now to the results. Table 2 gives a survey of all calculations we performed selected for central, semicentral and peripheral events.

First of all, we observe that especially for central collisions the quality of the impact parameter selection depends strongly on the observables which are used. The combination AMAX, IMF and FLOW leads to a standard deviation for central collisions of 0.8 fm , whereas the combination IMF, ERAT and FLOW allows to determine the impact parameter quite well (standard deviation = .31 fm). In any case the values are at least a factor of three better than if one uses ERAT only, as discussed above. Thus a neural network is indeed a proper tool to select the observables which contains the most information.

Let us investigate in detail how the network performs. In fig. 5 we display the impact parameter ordered deviation between $b^{network}$ and b^{QMD} for 4 combinations. In the top row we display the two combinations which give the best result. We see that up to b = .8 there are practically no systematic structures. Thus the standard deviation is caused by the fluctuations of the observables in simulations with almost the same impact parameters. Above b = 0.8 (= 11.2 fm) we observe systematic deviations. All the observables, MULT, FLOW, ERAT and IMF tend to zero for large impact parameters. Therefore their value for a determination of the impact parameter of peripheral reactions is rather limited. Most probably observables like the scattering angle of the large residue are more appropriate. The bottom shows the combinations MULT DIR ERAT and AMAX IMF FLOW. the first comes closest to the analysis with which the experimental groups [1,2] tried to select central events. We see that the network produces some systematic deviations close to the most central collisions. AMAX IMF FLOW is the worst case. We observe systematic structures in $b^{network} - b^{QMD}$. Since all of the three values tend towards zero for central collision the network has large problems to perform well in this region.

It is worthwhile to compare the results with those obtained in ref. [4]. The quantity Δb which defines there the deviation between the known impact parameter and that produced by the network is defined as

$$\Delta b = \frac{1}{N} \sum |b_i^{network} - b_i^{QMD}| \tag{19}$$

For the most central collisions ($b < 2.1 fm$) we obtain values of $\Delta b \cdot b^{max}$ of .62 ,.26 and .23, for the combination AMAX IMF FLOW, MULT FLOW ERAT, and IMF FLOW ERAT, respectively. Thus for the optimal combination of observables we reduce the error of the impact parameter determination by about 30% as compared to [4].

To verify whether an increase of the number of observables improves the situation we also performed a calculation with 4 input cells using the combination IMF, MULT, FLOW and ERAT, which represent the observables which give the best prediction using three entry cells. We do not observe an improvement of the prediction. If we change the number of hidden cells for the selection of three observables we do also not obtain a better prediction.

In summary, we have presented a systematic study of a new method, the application of a neural network, to determine the impact parameter of heavy ion reactions. The neural network allows to lower the standard deviation between the known impact parameter of "theoretical data" and the impact parameter derived from the observables by a factor of four as compared to the use of one observable only. Thus the most interesting central events can now be selected with a precision of about .3 fm. Using instead of the "theoretical" data those which have been filtered by the acceptance of the experiment, the network - after being trained with this data set - can be directly used to select the most central experimental reactions by using the measured values of the observables as the input variables. One could even imagine that this method can be used online to select already during the experiment the impact parameter range desired for the later analysis.

References

[1] J.P. Alard et al, Phys. Rev. Lett. 69 (1992) 889

[2] T. Wienold, GSI Report 93-28

[3] M. Freslier, DEA Thesis, Nantes 1993, unpublished

[4] S.A. Bass et al., J. Phys. G20 (1994) L21

[5] J. Aichelin and H. Stöcker, Phys. Lett. B176 (1986) 14
J. Aichelin, Phys. Rep. 202 (1991) 233,
C. Hartnack et al. submitted to Nucl. Phys.

[6] Bernt Müller, J. Reinhardt, Neural Networks, Springer, Berlin 1991

FIGURE CAPTIONS

Fig.1: The multiplicity of charged particles MULT, the largest fragment AMAX and the energy ratio ERAT as a function of the reduced impact parameter b. We display the fit function as well.

Fig.2: The difference between b^{QMD} and b^{fit} for the observable ERAT for the impact parameter ordered events.

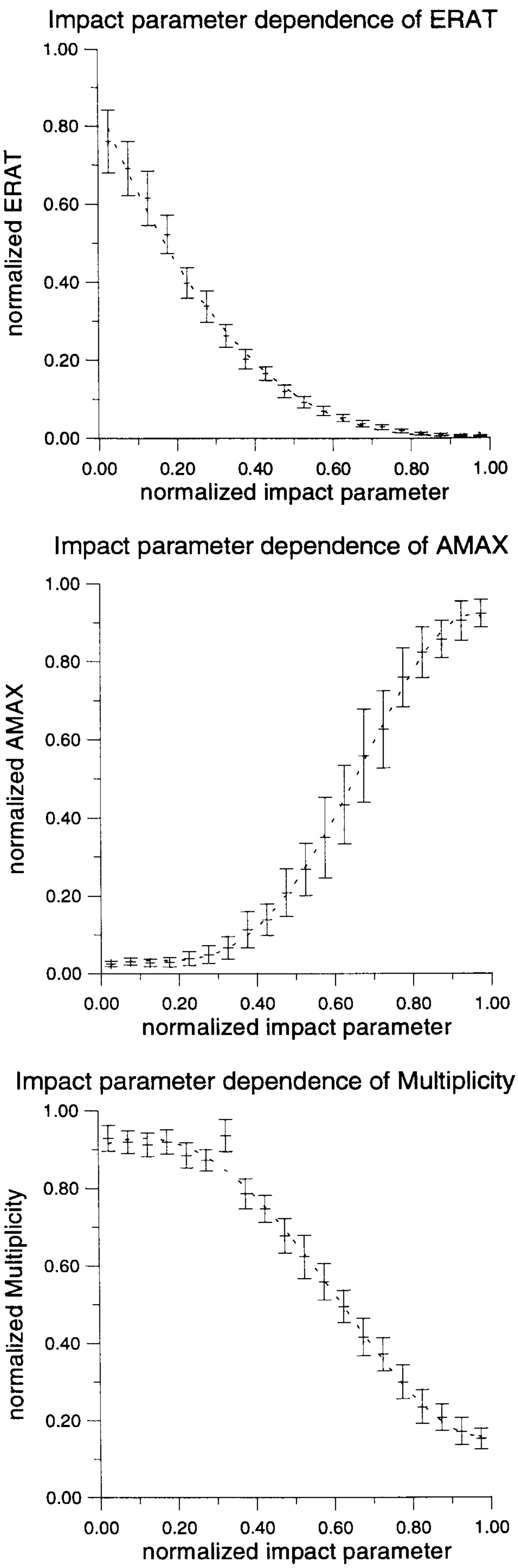
Fig.3: The forward feed network we use for our investigation.

Fig.4: The standard deviation between b^{QMD} and $b^{network}$ as a function of the number of iterations.

Fig.5: The difference between b^{QMD} and $b^{network}$ for the impact parameter ordered events for 4 combinations of observables.

$C=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(b_i^{QMD}-b_i^{net})^2}$	$0 \leq b \leq 0.15$	$0.3 \leq b \leq 0.7$	$0.85 \leq b \leq 1$
AMAX IMF FLOW	0.0569	0.049	0.0481
AMAX IMF ERAT	0.0331	0.0277	0.0441
AMAX IMF FSD	0.0455	0.0433	0.0469
AMAX MULT FLOW	0.0521	0.0489	0.0442
AMAX MULT ERAT	0.0356	0.0281	0.0422
AMAX MULT FSD	0.0475	0.0457	0.0415
AMAX FLOW ERAT	0.04	0.032	0.043
IMF MULT FSD	0.0546	0.0414	0.041
IMF MULT ERAT	0.0335	0.025	0.0406
IMF MULT FLOW	0.0543	0.0455	0.0422
AMAX DIR ERAT	0.0327	0.0357	0.0427
AMAX ERAT FSD	0.0277	0.029	0.0469
IMF DIR ERAT	0.027	0.03	0.0564
IMF ERAT FSD	0.0242	0.027	0.0492
MULT FLOW DIR	0.0354	0.0256	0.0375
MULT DIR ERAT	0.0234	0.0232	0.0387
MULT ERAT FSD	0.0251	0.023	0.0395
MULT FLOW ERAT	0.0234	0.0217	0.0353
IMF FLOW ERAT	0.0224	0.0224	0.0357
IMF MULT FLOW ERAT	0.0251	0.0232	0.0368

Table 2: Uncertainty of the impact parameter determination using three (four) observable as the input of the network



Impact parameter dependence of ERAT
normalized ERAT
0.00
0.20
0.40
0.60
0.80
1.00
normalized impact parameter
Impact parameter dependence of AMAX
normalized AMAX
normalized impact parameter
Impact parameter dependence of Multiplicity
normalized Multiplicity
normalized impact parameter

fig. 1

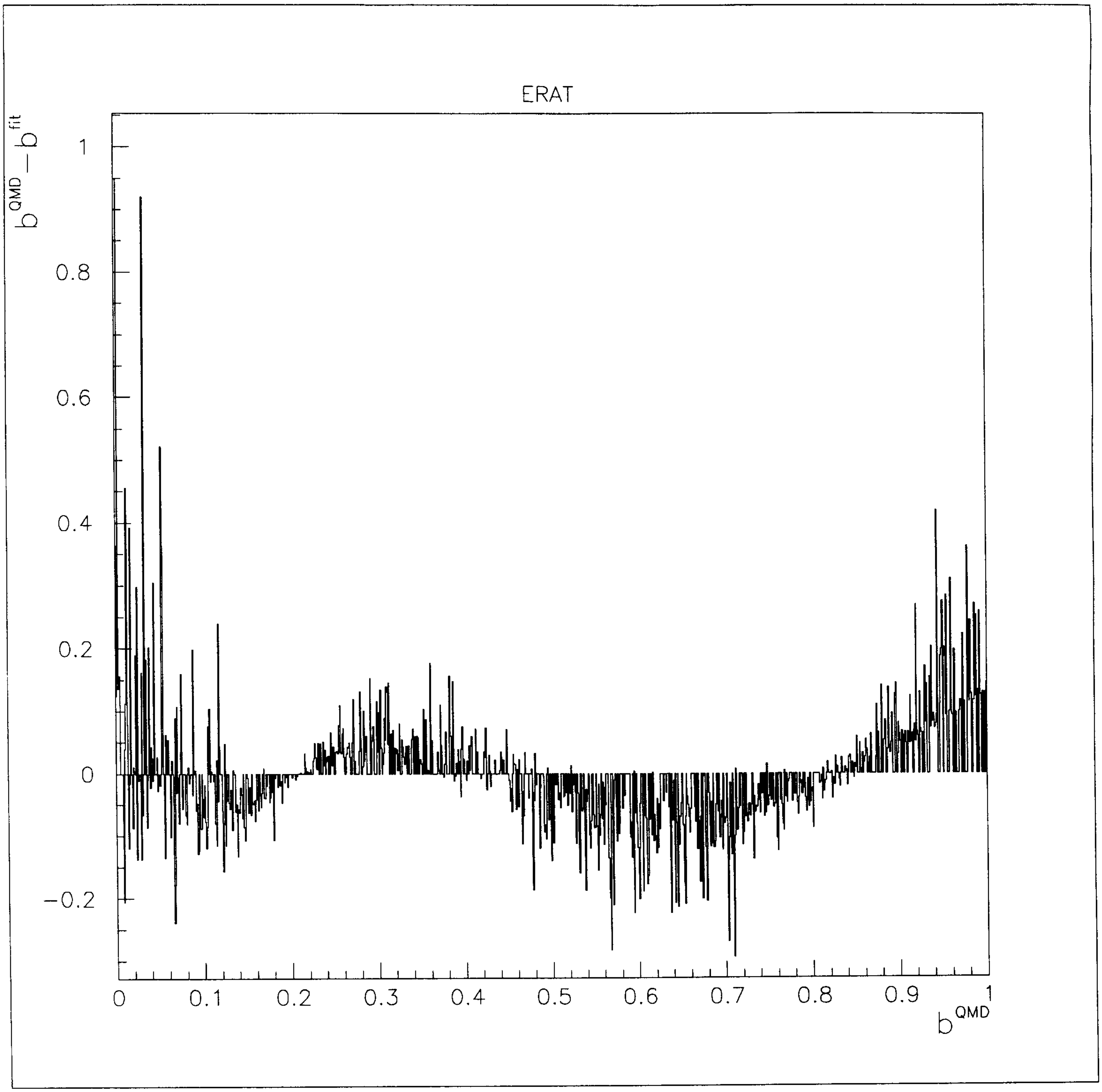


Fig. 2

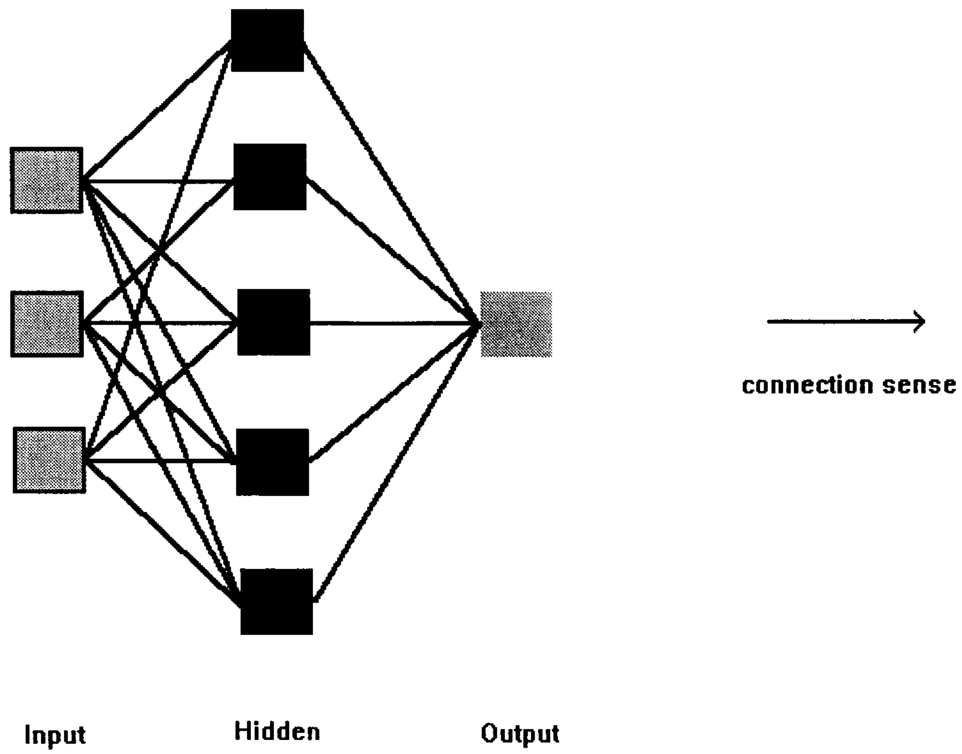


Fig. 3

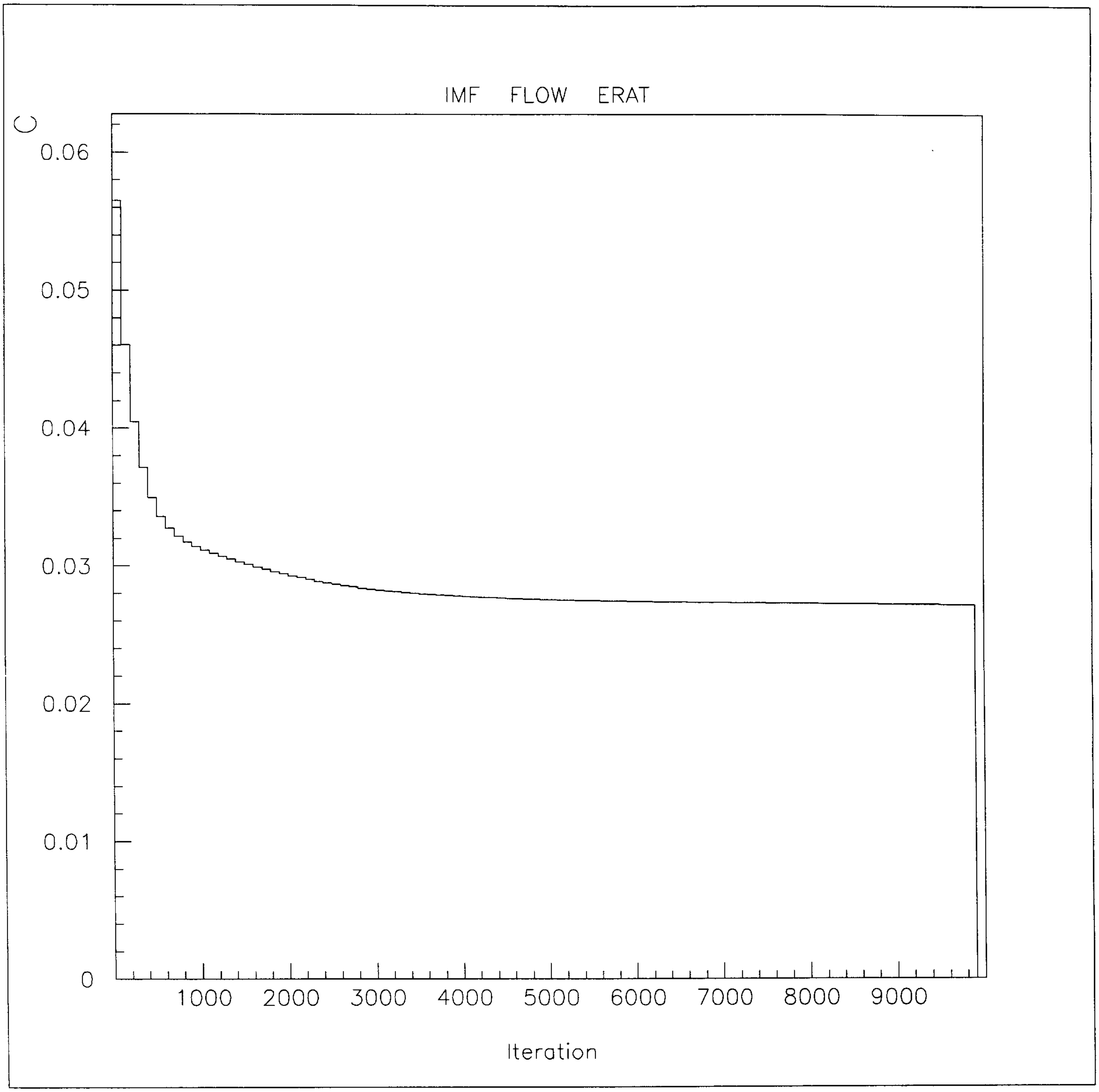

IMF FLOW ERAT
C
0.06
0.05
0.04
0.03
0.02
0.01
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Iteration

Fig. 4

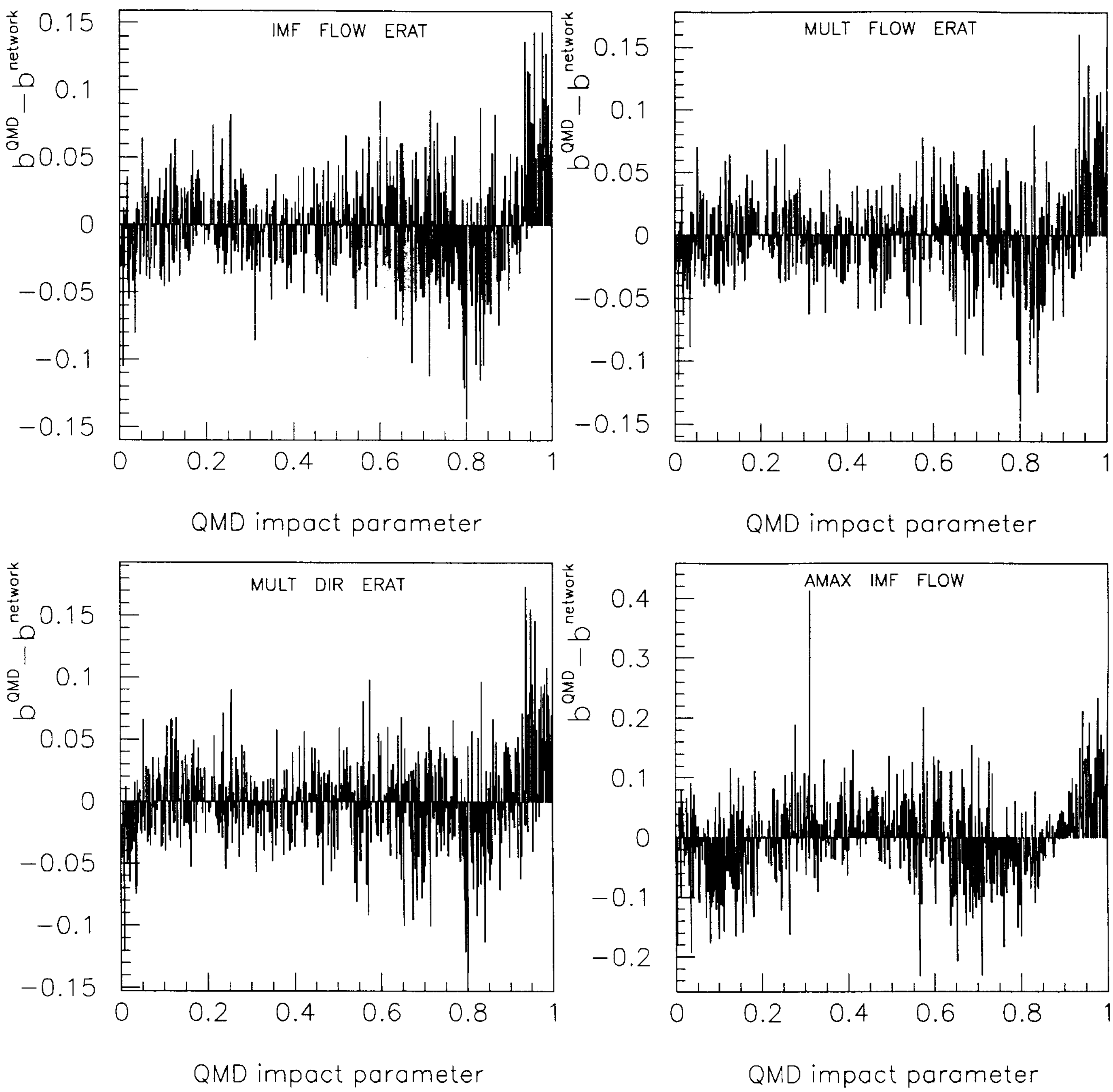

IMF FLOW ERAT
MULT FLOW ERAT
MULT DIR ERAT
AMAX IMF FLOW
$b^{QMD} - b^{network}$
QMD impact parameter
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
0.4
0.3
0.2
−0.2
0.6
0.8
1

Fig. 5