Impact parameter determination in experimental analysis using a neural network

F. Haddad,[1,2] K. Hagel,[1] J. Li,[1] N. Mdeiwayeh,[1] J. B. Natowitz,[1] R. Wada,[1] B. Xiao,[1] C. David,[2] M. Freslier,[2] and J. Aichelin[2]

[1]*Cyclotron Institute, Texas A&M University, College Station, Texas 77843*

[2]*Subatech, 4 Av A. Kastler, 44070 Nantes cedex 03, France*

(Received 11 July 1996)

A neural network is used to determine the impact parameter in $^{40}Ca+^{40}Ca$ reactions. The effect of the detection efficiency as well as the model dependence of the training procedure has been studied carefully. An overall improvement of the impact parameter determination of 25% is obtained using this technique. The analysis of Amphora $^{40}Ca+^{40}Ca$ data at 35 MeV per nucleon using a neural network shows two well-separated classes of events among the selected ''complete'' events. [S0556-2813(97)05903-7]

PACS number(s): 25.70−z, 07.05.Mh, 24.10.Cn

I. INTRODUCTION

Experimental heavy-ion data become more and more difficult to analyze without the help of theoretical calculations. A key parameter of such calculations is the impact parameter which is strongly correlated with the violence of the reaction. Unfortunately, the impact parameter remains difficult to extract experimentally despite intensive experimental works [1–5].

Numerous attempts have been made to extract this quantity by analyzing the correlation between a single observable and the violence of the collision. For example, the charged particle multiplicity [3], the perpendicular momentum [1], the neutron number [4], etc., have been extensively used. It turns out that all these single observable methods, very efficient for peripheral reactions, failed for central collisions due to the saturation of the considered observables in the energy range under consideration here (between 20 and 100 MeV per nucleon).

With the availability of 4π detectors, it would be of interest to use as much information as possible by combining several different observables. Some works in this direction have been undertaken by Llope *et al.* [5] by studying the autocorrelation between different observables in order to maximize the efficiency of the central collision selection. Recently, another promising step in this direction has been made using neural networks [6–8]. Indeed, using simulated $^{197}Au+^{197}Au$ events, David *et al.* [7] have obtained an improvement by a factor of 4 in the determination of the impact parameter compared to more conventional method in central collisions.

The purpose of this paper is to explore the neural network capabilities to help analyze the Amphora 35 MeV per nucleon $^{40}Ca+^{40}Ca$ data [10]. This paper is organized as follows. We first start with a short review on neural networks and then focus on the training procedure. After an illustration of the performance of the neural network in Sec. II, we present the analysis of $^{40}Ca+^{40}Ca$ data in Sec. III and finally conclude in Sec. IV.

II. NEURAL NETWORKS

A. General description

Let us start with a succinct review of neural networks (NN's). For details and general background, the reader is referred to [11]. By definition a neural network is an ensemble of highly connected cells. A cell is an entity which has one or several weighted (ω_i) inputs I_i and an activation threshold θ, and gives an output according to a certain activation function f. Such a cell is represented in the top of Fig. 1. The output S of the cell is generated according to

$$S=f\left(\sum_i \omega_i I_i+\theta\right).$$

For a multilayer network, the activation function of each layer can be different. For the sake of simplicity, we have

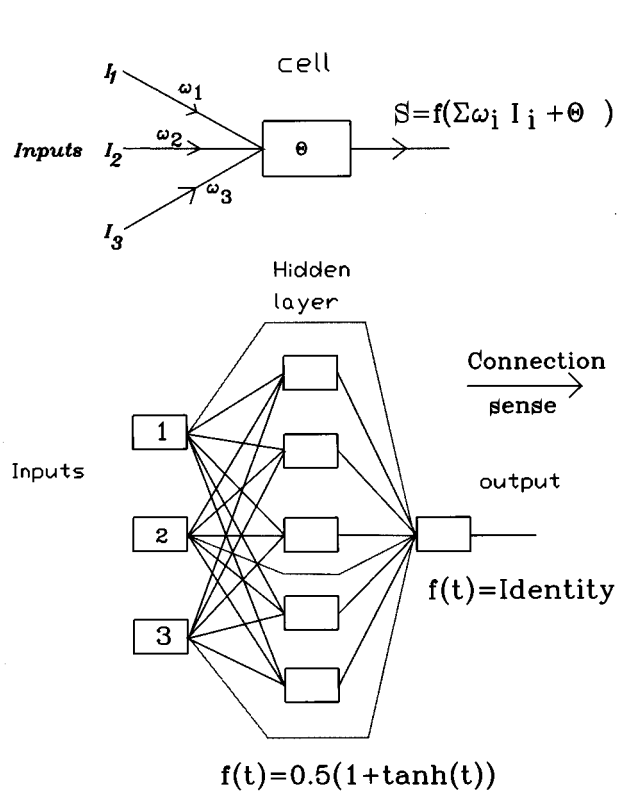

FIG. 1. Schematic view of a cell (upper drawing) and of our network (lower drawing).

restricted ourselves to a three-layer feed-forward network. The first layer, which corresponds to the input layer, is composed of three cells, the median layer is called the hidden layer and contains five cells, and finally the last one is a single-cell output layer. Input cells receive data from the outside (here the value of the physical observables) and the output cell gives the result (here an estimate of the impact parameter). The activation functions used for each layer of our network are displayed in the bottom of Fig. 1.

The use of neural networks is a two-step process: a learning stage followed by an application stage. During the learning phase, the different parameters of the network (θ,ω) are determined. This is done with the help of a sample for which inputs and the expected output are perfectly known (a learning sample).

The parameters are then adjusted in order to minimize, according to the different weights and thresholds, the difference D between the calculated output $\mathrm{out}_{\mathrm{NN}}$ and the known one ξ for the whole training ensemble. The function D is defined in our work as

$$D(\omega_i,\theta_i)=0.5\left(\sum_{\mu=1}^{n}(|\zeta^{\mu}-\mathrm{out}_{\mathrm{NN}}^{\mu}|)^2\right),$$

where n is the total number of elements of the training sample. The details of the whole minimization procedure can be found in Ref. [7].

B. Training procedure

In our case, it is not possible to make a learning sample. A theoretical model has then to be used to generate this sample. In this work, two different models have been used to generate the learning sample: (i) a dynamical transport model QMD [12], coupled with GEMINI [13]. This hybrid model has already been very successful in reproducing many features of the ^{40}Ca+^{40}Ca reaction at 35 MeV per nucleon [10]. (ii) An event generator based on a statistical model EUGENE [14]. This model combines a massive transfer entrance channel with a statistical deexcitation and therefore does not contain deep inelastic processes.

The choice of these two models was motivated by the fact that they are very different in their philosophies and then will give a good estimate of the neural network capabilities. We also required the models to reproduce correctly the Z distribution obtained in the ^{40}Ca+^{40}Ca experiment at 35 MeV per nucleon [10].

The Z distributions obtained for the two models are displayed in Fig. 2. The solid circles correspond to the experimental data, the shaded histogram to the EUGENE simulation, and the other histogram to the QMD results. It has to be noted that as in the experiment, only complete events are taken into account in both calculations. That means that to be selected an event has to have a multiplicity greater than 10 and that 85% of the total charge has to be measured after passing through the filter. According to the geometry of our detector [15], this implies that we are already dealing with central events. As can be seen in Fig. 2 fairly good agreement with the experimental data is obtained in both cases. This can give us some confidence in using these calculations to generate a learning sample for our neural network.

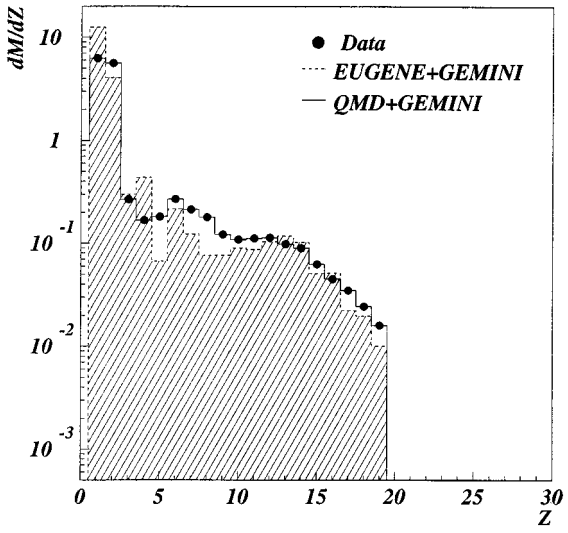

FIG. 2. Z distributions obtained for the complete events in ^{40}Ca+^{40}Ca at 35 MeV per nucleon. Experimental data are represented by solid circles, the EUGENE calculation by the shaded histogram, and the QMD calculations using a solid line. Calculations have been filtered.

We then generate sets of learning samples composed of 1000 events uniformly distributed between 0 fm and 8 fm. The three inputs of our neural networks are the charged particle multiplicity (CP), the perpendicular momentum (P_{perp}), and $E_{\mathrm{rat}}=(\Sigma P_t^2/2m)/(\Sigma P_z^2/2m)$. They correspond to the most efficient combination of the available experimental observables for ^{40}Ca+^{40}Ca at 35 MeV per nucleon [9].

III. BEHAVIOR OF OUR NEURAL NETWORK

To compare the performance of the neural network (NN) with other commonly used methods, an observable is defined:

$$\mathrm{deviation}=\frac{1}{N}\sum_{i=1}^{N}|B_i^{\mathrm{QMD}}-B_i^{\mathrm{var}}|,$$

which gives an estimate of the dispersion over the overall range of impact parameter. B^{var} stands for the impact parameter value determined using one of the observables NN, CP, P_{perp}, and E_{rat}. For methods other than the neural network, let us explain the way the impact parameter is extracted. Using the training sample, the total distribution of a given observable (var) is cut into ten equal-size bins for which the average impact parameter is calculated. Then, using a fit of a polynomial expression of these points, an impact parameter value B^{var} is associated with each value of the observable.

We have reported the results of such a comparison for Ca+Ca reactions as a function of the incident energy in Fig. 3.

For all energies, the neural network gives the lowest deviation. It is the most accurate of the methods used here. It can be seen also that as the incident energy increases, the impact parameter determination becomes better. This is true for all the different methods. Nevertheless, the neural network always allows an improvement of around 25% compared to the others.

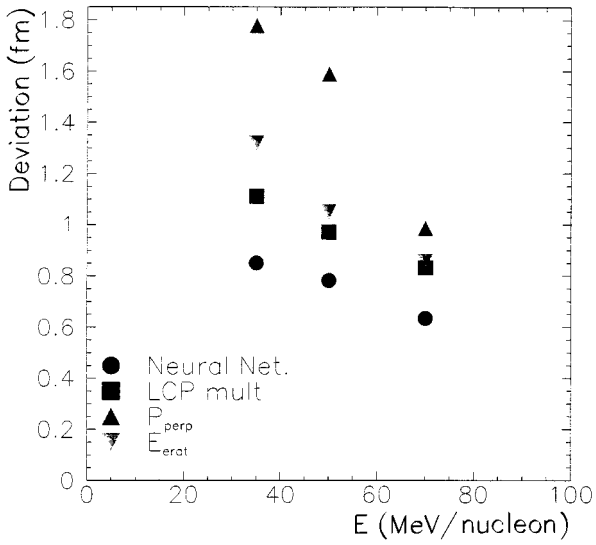

FIG. 3. Energy dependence of the deviation (see definition in the text) for Ca+Ca reactions obtained from QMD+GEMINI.

In Fig. 4, the recognition capability of our neural network is given for a theoretical case (no filter was applied). In this figure, the neural network output value $\mathrm{out}_{\mathrm{NN}}$ is displayed as a function of the corresponding QMD impact parameter B_{QMD} for ^{40}Ca+^{40}Ca at 70 MeV per nucleon. The correlation between these two quantities is very good above 1.5 fm. For very central collisions, $\mathrm{out}_{\mathrm{NN}}$ saturates due to the saturation of input observables. It can also be noticed that the dispersion around the mean value increases with decreasing impact parameter.

In most previous neural network works, model calculations have been used without taking into account any experimental filter. The effect of such a filter is far from negligible and has the tendency to increase the apparent fluctuations. The Amphora detector filter has been applied to study this effect. The result is presented in Fig. 5. As expected, the recognition by the neural network is poorer than without the

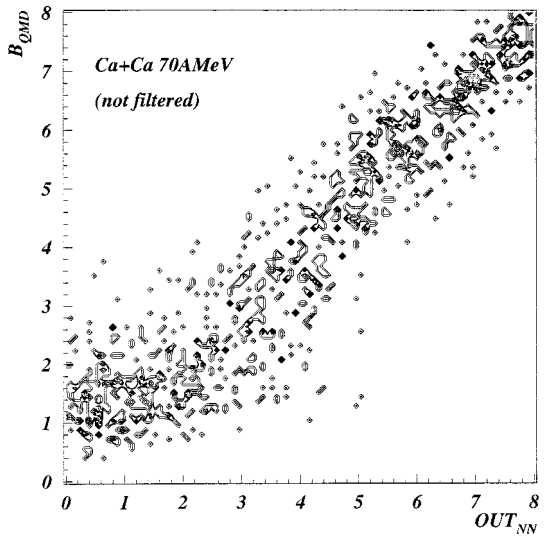

FIG. 4. Neural network output $\mathrm{out}_{\mathrm{NN}}$ as a function of the QMD impact parameter B_{QMD} for the ^{40}Ca+^{40}Ca at 70 MeV per nucleon.

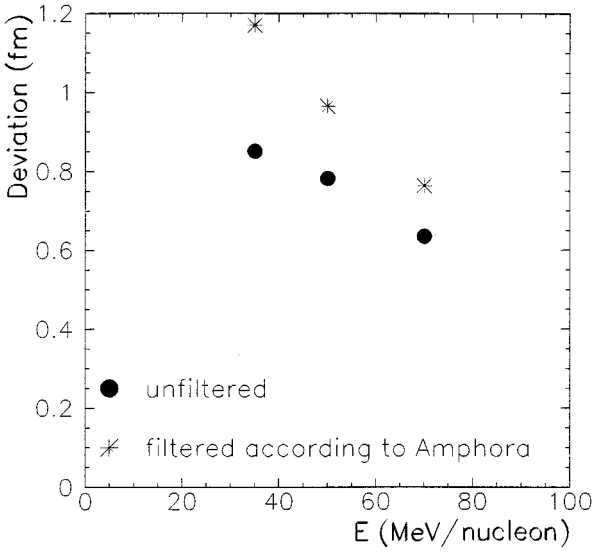

FIG. 5. Effect of a filter in the neural network performances.

filter. This clearly shows the necessity to use a network trained as closely to the experimental condition as possible.

IV. APPLICATION ON REAL DATA

The learning stage finished, it is possible to apply the neural network on real data. In Fig. 6, the neural network output obtained for both trained networks is displayed for each selected experimental event (complete event). The shaded histogram corresponds to the EUGENE-trained network whereas the other histogram corresponds to the QMD-trained network.

The neural network outputs are always below 7 fm. This means that complete events deal mainly with central events as expected. The QMD-trained network gives higher impact parameter value due to the fact that the QMD model contains

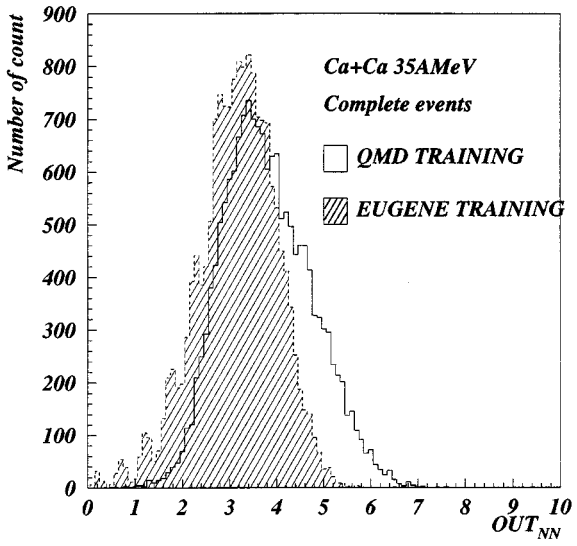

FIG. 6. Neural network output obtained for each selected experimental event. EUGENE-trained network outputs are represented using the shaded histogram whereas QMD-trained network outputs are displayed by a solid line.

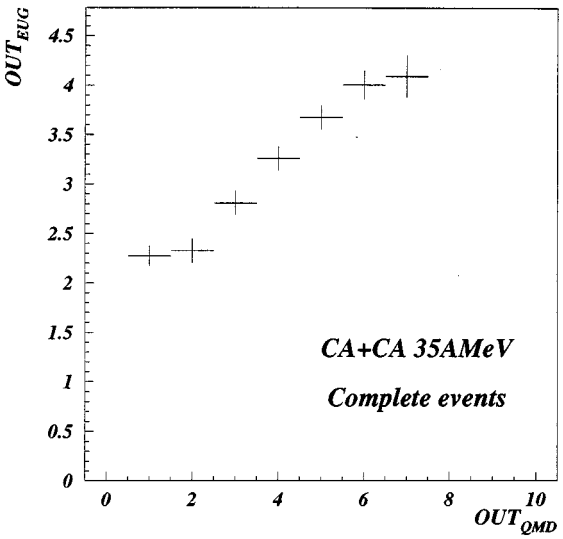

FIG. 7. Correlation between the EUGENE-trained network output and the QMD-trained network output.

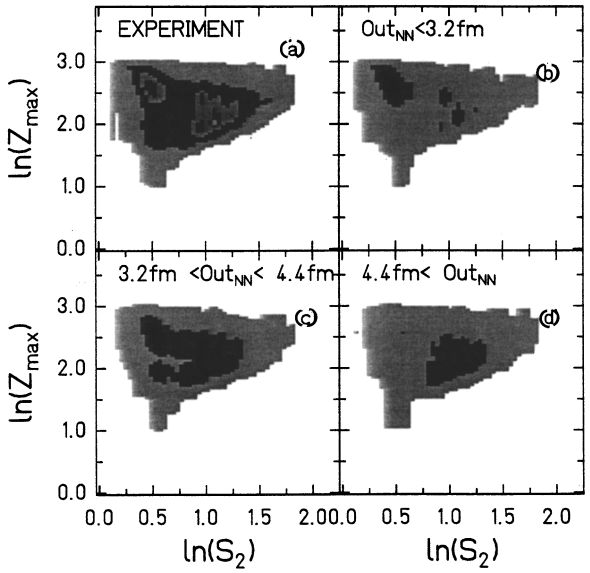

FIG. 8. Logarithmic distribution of $Z_{\max}$ vs $S_2'=\Sigma_{i,Z\neq Z_{\max}}Z_i^2M(Z_i)/\Sigma_{i,Z\neq Z_{\max}}Z_iM(Z_i)$. Each contour represents constant value units of relative $d^2Y/d\ln S_2'Z_{\max}$ where Y is the yield. The outside contour level is at level 10, and each inner contour represents a progression in yield of 150.

deep inelastic processes contrary to the EUGENE model. A depletion occurs for very central impact parameter values. This is a result of two facts: We are working with real data which means that events are weighted according to $2\pi bdb$ where b stands for the impact parameter. This makes central events less abundant than peripheral ones. In addition, as shown in Fig. 4, the neural network recognition failed in very central collisions and gives systematically an overestimated impact parameter value. The combination of these two facts explains the shape of the spectra at low out_{NN}. The maxima of both distributions are located around 3 fm.

It is interesting now to look at the correlation between the two network outputs for our experimental events. In Fig. 7 the averaged EUGENE-trained network output out_{EUG} is plotted as a function of the QMD-trained network output out_{QMD}. A quite good correlation exists between the two outputs. A high out_{QMD} corresponds to a high out_{EUG}. This result and the fact that the two distribution maxima are close are quite encouraging for the use of a neural network in an experimental analysis.

A precise individual impact parameter determination at this energy by our network does not seem reasonable. Nevertheless, we are going to separate the data into three groups according to out_{NN}. The limits of these groups are $\text{out}_{\text{NN}}\leqslant 3.2$ fm, $3.2<\text{out}_{\text{NN}}\leqslant 4.4$ fm, and $\text{out}_{\text{NN}}>4.4$ fm and have been chosen to make three equally populated bins.

For these three different classes, the so-called ''Campi plot'' has been generated. This plot allows an exploration of the moments of the multiplicity distribution and has been suggested as a useful means to identify a possible critical behavior in the deexcitation patterns. In Fig. 8, such plots are displayed for all events as well as for the three out_{NN} classes.

Two peaks occur in the experimental contour plot in panel (a) of Fig. 8. One is located at large values of $\ln Z_{\max}$ and small values of $\ln S_2'$ and the other is located at small values of $\ln Z_{\max}$ and large values of $\ln S_2'$. Plots obtained for the different cuts in out_{NN} show quite distinct behavior. For the lower values of out_{NN}, only the low $\ln S_2'$ peak remains. On the other hand, for higher out_{NN} values, only the high $\ln S_2'$ peak is present. This systematic nice behavior shows that neural networks can be very useful in data analysis by allowing the grouping of events according to the correlation of several observables (here CP, P_{perp}, and E_{rat}).

V. CONCLUSION

In this paper, a neural network has been used for the first time to determine the impact parameter of real experimental events ($^{40}\text{Ca}+^{40}\text{Ca}$ at 35 MeV per nucleon). An improvement of about 25% is obtained compared to commonly used methods. Applied to real data, the $^{40}\text{Ca}+^{40}\text{Ca}$ reaction at 35 MeV per nucleon, the network provides a clear separation of the different peaks obtained in the experimental Campi plots. At the same time, a careful study of the influence of the model used for the training shows that in both cases the neural network outputs are consistent together. This tends to indicate that the neural network can be a valuable tool in data analysis.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Grant No. DE-FG05-86ER40256, the National Science Foundation, The Robert A. Welch Foundation and the French Centre National de la Recherche Scientifique.

[1] J. Péter *et al.*, Nucl. Phys. A (to be published).
[2] L. Phair *et al.*, Nucl. Phys. **A548**, 489 (1992).
[3] M. Betty Tsang *et al.*, Phys. Rev. C **40**, 1685 (1989).
[4] D. Utley, R. Wada, K. Hagel, J. Li, X. Bin, Y. Lou, R. Tezkratt, J. B. Natowitz, and M. Gonin, Phys. Rev. C **49**, R1737 (1994).
[5] X. Llope *et al.*, Phys. Rev. C **51**, 1325 (1995).
[6] S. A. Bass *et al.*, J. Phys. G **20**, L21 (1994).
[7] C. David *et al.*, Phys. Rev. C **51**, 1453 (1995).
[8] S. A. Bass *et al.*, Phys. Rev. C **53**, 2358 (1996).
[9] F. Haddad, K. Hagel, J. Li, N. Mdeiwayeh, J. B. Natowitz, R. Wada, B. Xiao, C. David, M. Freslier, and J. Aichelin, in Proceedings of the 209th ACS National Meeting, Anaheim, 1995.
[10] K. Hagel *et al.*, Phys. Rev. C **50**, 2017 (1994).
[11] B. Müller and J. Reinhardt, *Neural Networks* (Springer, Berlin, 1991).
[12] T. Naruyama *et al.*, Prog. Theor. Phys. **87**, 1367 (1992).
[13] B. Charity *et al.*, Nucl. Phys. **A483**, 371 (1988).
[14] D. Durand, Nucl. Phys. **A541**, 266 (1992).
[15] D. Drain *et al.*, Nucl. Instrum. Methods Phys. Res. A **281**, 528 (1989).