ISTITUTO NAZIONALE DI FISICA NUCLEARE

Sezione di Roma II

M. Casolino, P. Picozza:

A CELLULAR AUTOMATON TO FILTER EVENTS IN A HIGH ENERGY PHYSICS DISCRETE CALORIMETER

A CELLULAR AUTOMATON TO FILTER EVENTS IN A HIGH ENERGY PHYSICS DISCRETE CALORIMETER

M. Casolino, P. Picozza
Dipartimento di Fisica dell'Università di Roma II "Tor Vergata", Italia
INFN, Isituto Nazionale di Fisica Nucleare, Sezione di Roma II, Italia

Abstract

In this paper the noise-filtering capabilities of a cellular automaton are studied. The tests have been performed on 4 GeV pion experimental events taken with a Silicon Tungsten tracking calorimeter at CERN PS. Particle interaction with the material of the calorimeter can produce straight tracks or hadronic showers. The evolution rules devised for the CA have therefore to reckon with these different topologies in order to remove noise and restore interrupted tracks. At the same time, strips hit by particles have to be affected as little as possible. The distributions of some discriminating parameters are compared with Montecarlo data before and after filtering by the automaton and the agreement is shown to improve. The Cellular Automaton constitutes a simple (Boolean logic) and fast (with possible on-line applications) filter. Its applications range from particle discrimination to track restoration in time-worn detectors.

1 Introduction

Cellular automata (CA) were introduced in 1970 by M. Gardner who presented Conway's game "Life" [1]. They consist of a set of update rules which define a local dynamical system on a discrete space. Once these rules are fixed and an initial configuration is specified, the system is free to evolve without external influence. "Life" was originally conceived as a game and indeed it is interesting just to watch its evolution, an example of how cellular automata produce very complex and interesting behaviors even with simple rules. In the last years the growth of computational power available has increased the studies and applications of CA. Dedicated hardware has often been developed to speed up calculations[2]. Today cellular automata are currently used in modeling physical systems such as reaction-diffusion models[3] and fluid dynamics[4, 5]. Also the mathematical and statistical properties of CA have been studied [6, 7, 8]. There have been applications in high energy particle physics to filter noise from tracks obtained in multiwire detectors[9].

In this paper a cellular automaton is instead used to filter noise from the experimental data of 4 GeV pions not only giving straight tracks but also producing hadronic showers. Therefore the CA has to provide track reconstruction and spurious hits elimination for these different topologies.

2 Experimental data

The data used in these tests have been taken at CERN PS beam T7N on a prototype of the Silicon calorimeter now currently used, full sized, with other detectors in balloon flights to measure the cosmic ray antimatter (i.e. $\bar{p}$, e^+) component[10].

The prototype consists of 20 silicon strip planes, interleaved by 0.175 (0.5 X_0) tungsten slabs to

provide showering material for the incoming particles. A single plane is composed of 2 (for X and Y views) $6 \times 6\ cm^2$ wafers, each divided in 16 strips 3.6 mm wide, ortogonally placed. Each relativistic charged particle crossing the depleted area of a silicon strip is minimum ionizing and releases a roughly fixed amount of energy. This value is taken equal to 1 in *minimum ionizing particle* (mip) units. The amount of energy deposited in each strip is therefore proportional to the number of particles crossing it: this signal is detected and stored by an electronic read-out chain. The performance of this calorimeter prototype has been reported in ref.[11].

For each event there are two 20×16 matrixes (X and Y views) containing topological and energetic information of the interaction of the particle with the detector.

The response of each strip of the calorimeter has been studied in the calibration phase of the instrument. The distance between the peak of the electronic noise pedestal and the peak of the Landau energy loss distribution is normalized to 1. For each strip this normalization factor is determined in the calibration phase. To eliminate noise from the events, a threshold T is placed and all values below this treshold are set to 0. The choice of T is critical and determines the amount of noise removal and the track detetion efficiency of the calorimeter. If a value too low of T is chosen, the efficiency is high but the number of noisy strips increases. If T is too high, only a little noise is taken but also a significant part of data can be lost. As will be shown, a CA can operate in a large range of values of T without great differences in the output This may be particularly useful in larger detectors, currently flying on a balloon and in the future on a satellite, where the work conditions may often change. Cellular automata can cope with these variations, restoring the missing signal and removing noise.

3   Cellular automata rules

A cellular automaton (CA) is defined in a two dimensional square lattice[1] by a set of *local* updating rules which determine the evolution of the system at discrete time steps. Each point in the lattice (a *cell*) can have only two states: 0 (*dead*) or 1 (*alive*). At every iteration, the state of each cell may be modified according to the disposition of its 8 neighboring cells[2]. The process is iterated until convergence (no more deaths or births) of the system is achieved.

The state of a cell may therefore remain unchanged, may pass from 0 to 1 (*birth*) or from 1 to 0 (*death*). The CA is defined by the rules that govern the evolution of the cells; consequently, in physical process modeling, all the approximations and assumptions have to be made at this stage. In our case the 16 × 20 grid of each view of the calorimeter is the space where cells (the Si strips) evolve. The initial configuration is given by the unfiltered event. Even if the silicon calorimeter provides energetic and topological information, only this latter is used: strips with an energy above 0 are set to 1, all others are set to 0.

The birth rules are defined to restore straight tracks with interruptions due to a faulty (or broken) strip which has failed to reveal the particle passing through it. These rules are shown in Figure 1a, 1b, and 1c for a dead cell having (step t) two nearby living cells, one on its left and another on its right. Figure 1a shows a particle crossing orthogonally (from left to right) the Si planes in straight line, while Figures 1b and 1c consider particles undergoing scattering in the detector. At the next step ($t+1$), the central cell will be "on". To preserve the energetic information of the data

---

[1] Also hexagonal lattices may be used.

[2] Since the status of each cell depends only from its nearest neighbors, in case of very large systems, the update process may be easily parallelized to speed up calculations.

for a subsequent physics analysis, the average value of the energy among the two neighboring strips is taken for the born cell.

The "death rules" determine the kind and the amount of noise the CA has to remove. We consider two different cases. In the first a cell having no alive companions either on its left side, its top or its bottom, is killed (Figure 1d). The presence of living cells on the right side is not important since the particle beam hit the calorimeter from left to right.

The simplest example is a signal from an isolated strip (Figure 1e). Since this cell has no neighboring cells alive, its signal is generally due to noise and is eliminated in the update phase.

This rule reduces also noise coming from clusters due to background particles hitting sideways the calorimeter thus producing signals in addition to the track of the particle which had given the trigger.

A second rule, shown in Figures 1f and 1g allows the removal of a single living cell which is next to a straight track.

The update of the grid is two-step synchronous. All birth conditions are considered and cells are born before death evaluation is performed, to avoid an effect of "signal propagation" along interrupted tracks. This happens when a cell having two nearby alive cells, one in the left and one in the right positions is born, but at the same time the cell on its right, finding no alive cells on its left, is killed. The total effect - which is removed if all births occur before deaths - is to move the missing signal one plane on the right every update cycle.

Since the status of each cell may be represented by a bit, the speed of the evolutionary process may be increased using simple and fast Boolean operations. No update is performed on the first plane of the calorimeter, where, however, only one hit strip per view is required.

The effect of the CA approach on a simple π straight track is shown in Figure 2a. The left image shows the unfiltered event having one interruption and three noisy strips. The right image shows the same event after filtering by the CA. The track has been restored and the noisy strips removed.

In this example only one update step is required to achieve stability of this dynamical system: more complex cases require more iterations.

In Figures 2b, 2c, and 2d the effect of the automaton on four hadronically interacting pions is shown. However, to judge in this case the overall performance of the CA it is necessary to consider the discriminating parameter distributions.

4   Results

A sample of 2000 4 GeV pion events has been chosen and some discriminating parameter distributions before and after the application of the CA have been compared.

Let us first consider the number of hits released in a view of the calorimeter. A pion crossing the silicon wafers without hadronic interactions intersects a strip per plane, so the average number of hits h expected for straight track pions should be $h = 20$. In Figure 3 are shown the distributions of the hit numbers before and after the filtering by the CA. The enhancement of a sharp peak at $h = 20$ is evident and the agreement with Montecarlo simulated data (using Geant 3.15 code) is remarkably improved. Events having $h < 20$ - due to hadronic interactions occurring early in the calorimeter with a small number of secondaries - are in quantitative agreement down to values of $h \simeq 15$, where a small excess of filtered data is found. Filtered data show also a second peak at $h = 22$. This is due to the fact that 2 adjacent strips on the same plane constitute a stable config-uration. On the other hand, the contribution at $h = 21$ due to a straight track and only one strip

hit far from the track is cancelled by the CA, so that the occurrence of 21-hits events is reduced. This and other similar noise effects give to straight track events values of $h = 20 - 23$. The number and position of hadronic showers with a large number of secondaries ($h > 25$) match the simulated data. On the overall, the distribution of the filtered events is in much better agreement with Montecarlo calculations than the unfiltered data.

In Figure 4 distribution of the "Shower aggregation" before and after CA filter is shown for pions. The shower aggregation s is defined as follows: for each living cell we count the number of its neighboring state-1 cells; this value is then averaged over all the living cells of the calorimeter to give the shower aggregation. For pions leaving straight tracks, the most likely value is $s = 2$ (one cell on the left and one on the right, as in Figure 2a), while for interacting pions this value is obviously greater. It is possible to note how the number of events having $s < 2$ (that is, with missing strips) is greatly reduced after filtering, and that the tail of data with $s > 2$ (mainly hadronic showers) remains unaltered. This means that the topology of the events is essentially unaltered by the CA. The CA works without using energetic information of the event. As far as the energy is concerned, an average value between the two nearby cells is given to born cells; on the other hand the cells killed have mostly a low energetic value due to noise. Therefore, the energetic response of the calorimeter for pions is not appreciably modified by the filter, as shown in Figure 5. The peak at $E = 20 \, mips$ is due to straight tracks.

Since for weight and power consumption reasons this calorimeter is not a full containment detector already at GeV energies and is used for particle discrimination purposes, the use of topological discriminating parameters in conjunction with energetic variables is fundamental in particle classification[12, 13]. This discrimination may be greatly improved with the CA filtering.

The amount and type of noise filtered is shown in Figure 6, where the distribution of the difference $D = h_{raw} - h_{filtered}$ between the numbers of strips "on" of the raw and the filtered data is given. It is possible to see how only in about 5% of the events there is predominance of track restoration over noise filtering ($D < 0$). This happens for straight tracks having some interrupted strips and a very low noise content. However, in the vast majority of the cases $D > 0$, that is noise filtering (cell death) is much more common than track reconstruction (cell birth). This clearly appears in Figure 7a, where the number of cells killed vs. the iteration steps is shown. Deaths are present in up to 20 steps, while births take place (Figure 7b) only in the first 5 steps and in a much smaller number. As a consequence of the cell deaths, the convergence of the CA may take up to 21 steps (Figure 7c), since one more evolution is required to check that no more births or deaths are present.

If the CA is to be used as a on-line filter, the procedure may be stopped after a given number of iterations according to trigger rate and processing time requirements. Indeed, more than 50% of the events converge after only 2 steps.

To measure the independency degree of the CA-filtered data from the amount of noise present, and therefore from the value of the treshold T chosen in the calibration phase (see section 2), tests have been done using $T = 0.5 \, mip$. The results are compared with those previously considered, obtained with $T = 0.7 \, mip$. A value of $T = 0.5 \, mip$ means a lower cut and a higher amount of noise present. As can be seen in Figure 8, where the distributions of the number of hits for these two cases before and after the application of the CA are shown, the differences have been reduced by the CA. It is however possible to see how the number of 22-hit events (due to two nearby cells) increases with noise and also a small 24-hit peak due to two double-cell clusters appears.

5 Conclusions

A cellular automaton has been tested to filter noise and restore broken tracks in different conditions: straight track and hadronic interaction. The distributions of selected discriminating parameters have been led back to the expected values after such filtering and the agreement with Montecarlo simulated data has been shown to increase. This is done almost independently from the amount of the noise present. This algorithm may be used as the first stage of a particle identification process, or as a noise-removal filter (to save memory occupation, for example in flight experiments.) The simplicity of the CA makes it suitable for on-line applications.

References

[1] M. Gardner, Scientific American, **223** (4), (1970) 120.

[2] T. Toffoli, N. Magolus, *Cellular Automata Machines*, Mit Press, Cambridge (1987).

[3] B. Chopard, Proceedings of PC'94, Lugano, CH, R. Gruber and M. Tomassini eds., 515, EPS, Geneva, CH (1994).

[4] S. Di Gregorio et al., Proceedings of PC'94, Lugano, CH, R. Gruber and M. Tomassini eds., 523, EPS, Geneva, CH (1994).

[5] A. Lejeune et al., Proceedings of PC'94, Lugano, CH, R. Gruber and M. Tomassini eds., 527 , EPS, Geneva, CH (1994).

[6] S. Wolfram (ed.), *Theory and application of Cellular Automata*, World Scientific, Singapore (1988).

[7] B. Chopard, M. Droz, Europhys. Lett., **15**, 459 (1991).

[8] J. Hemmingsson, H.J. Herrmann, Europhys. Lett., **23** (1), 15 (1993).

[9] A. Glazov, I. Kisel, E. Konotopskaya, G. Ososkov, NIM, **A329**, 262 (1993).

[10] G.Barbiellini, M. Casolino, P. Picozza et al, *Nucl. Phys.*, **B32**, 77, (1993).

[11] M.Bocciolini, P. Picozza et al., *NIM*, **A333**, 560, (1993).

[12] M. Casolino, A. Morselli, E. Pasqualucci, P. Picozza, Proceedings of PC'94, Lugano, CH, R. Gruber and M. Tomassini eds. 719, Geneva, CH (1994).

[13] M. Casolino, P. Picozza, et al., Proc. of Frontier Detectors for Frontier Physics, 6th Pisa meeting on advanced detectors, La Biodola, Italy, 22-28/5/94, (1994) in press.
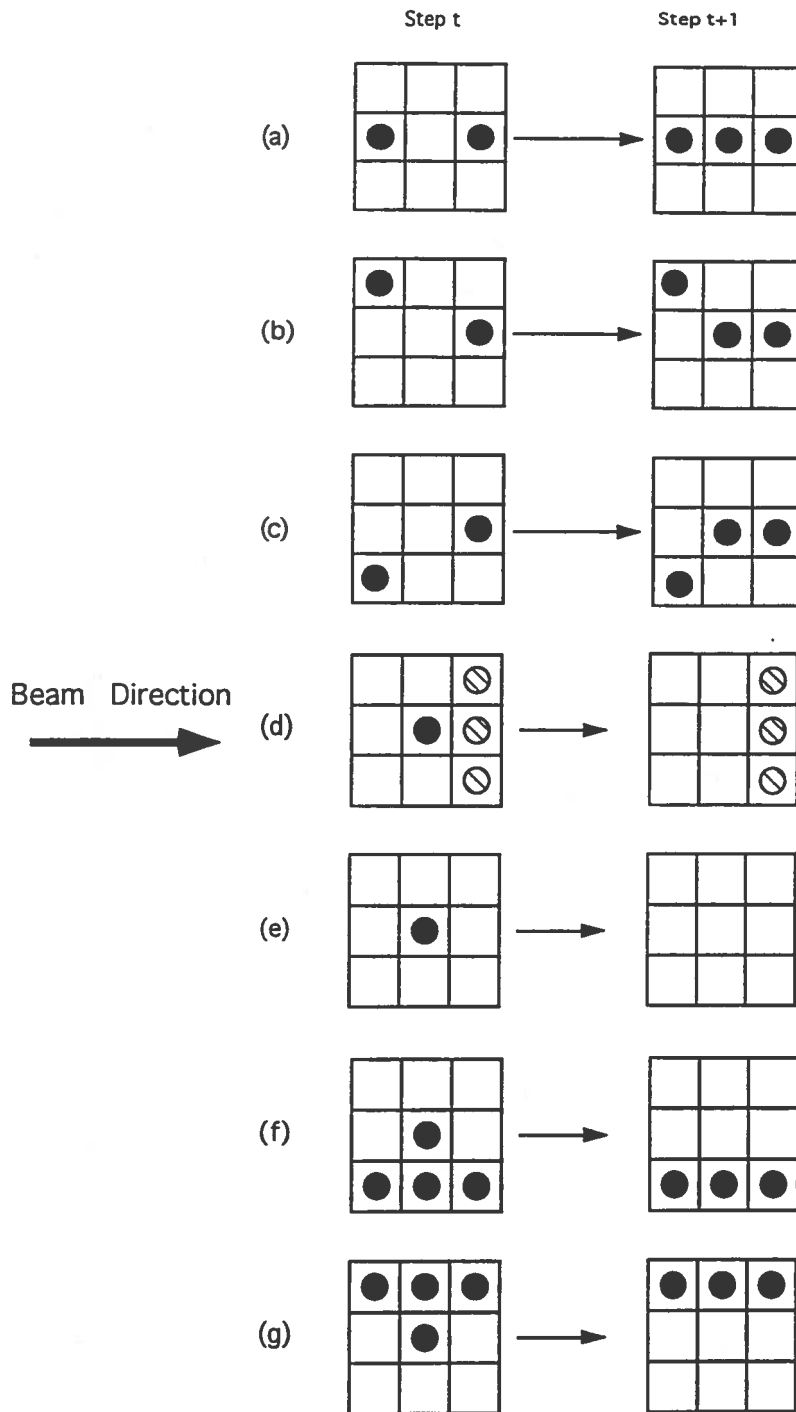
Figure 1: Cellular automata birth (a,b,c) and death (d,e,f,g) evolution rules. *Left*: before update; *right*: after update. The particle traveling direction is from left to right (This explains the asymmetry of the update rules). The status of the banded cells in case (d) is not important for the evolution of the central cell.
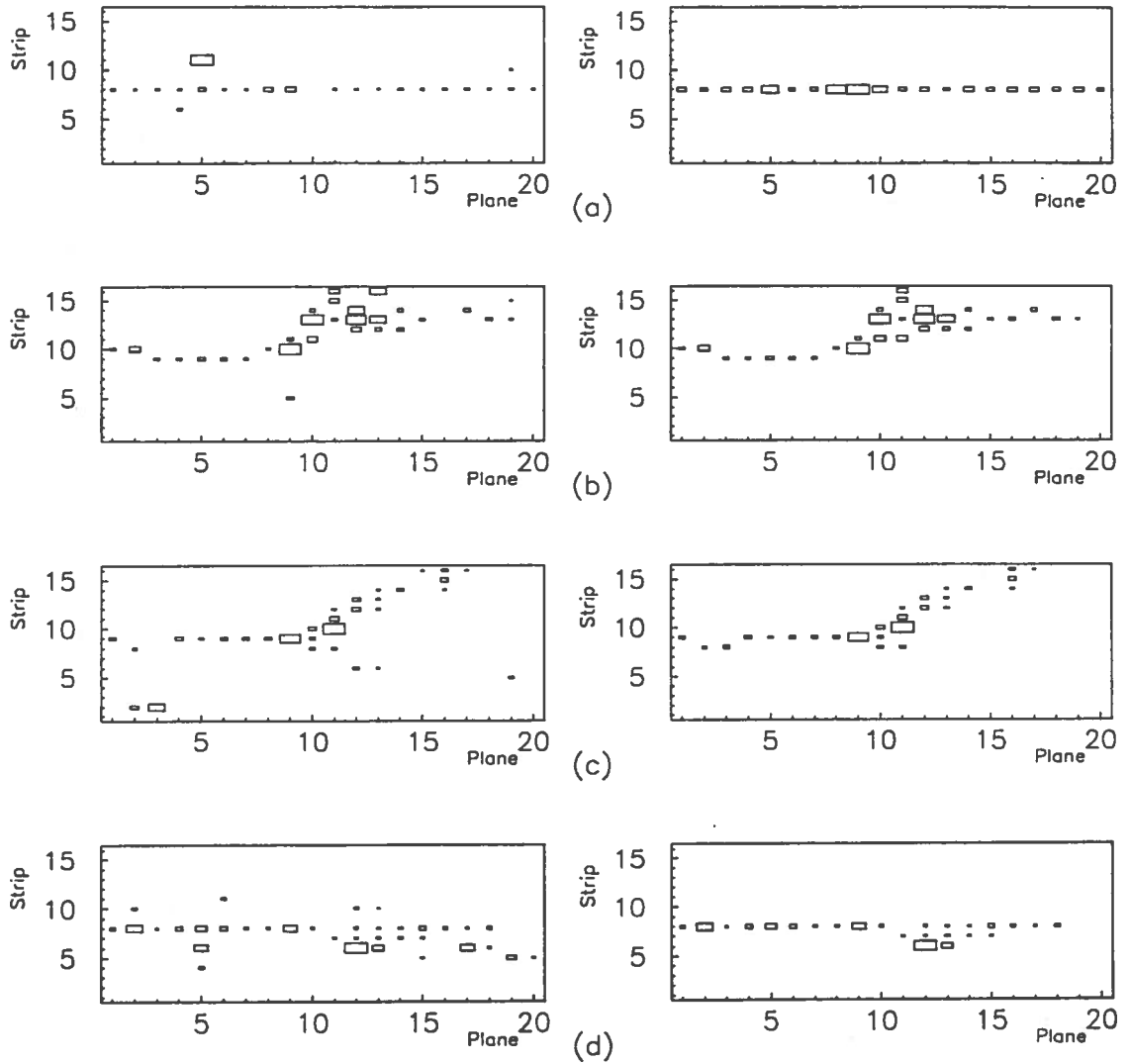
Figure 2: A straight track of a pion *(a)* and three hadronically interacting pions *(b,c,d)* before *(left)* and after *(right)* filtering by the CA. The size of each box is proportional to the energy released in each strip. On the X axis the plane number (1-20) is shown; on the Y axis the strip number (1-16). Only one view (X view) is given. The pion beam hits the calorimeter from left to right.
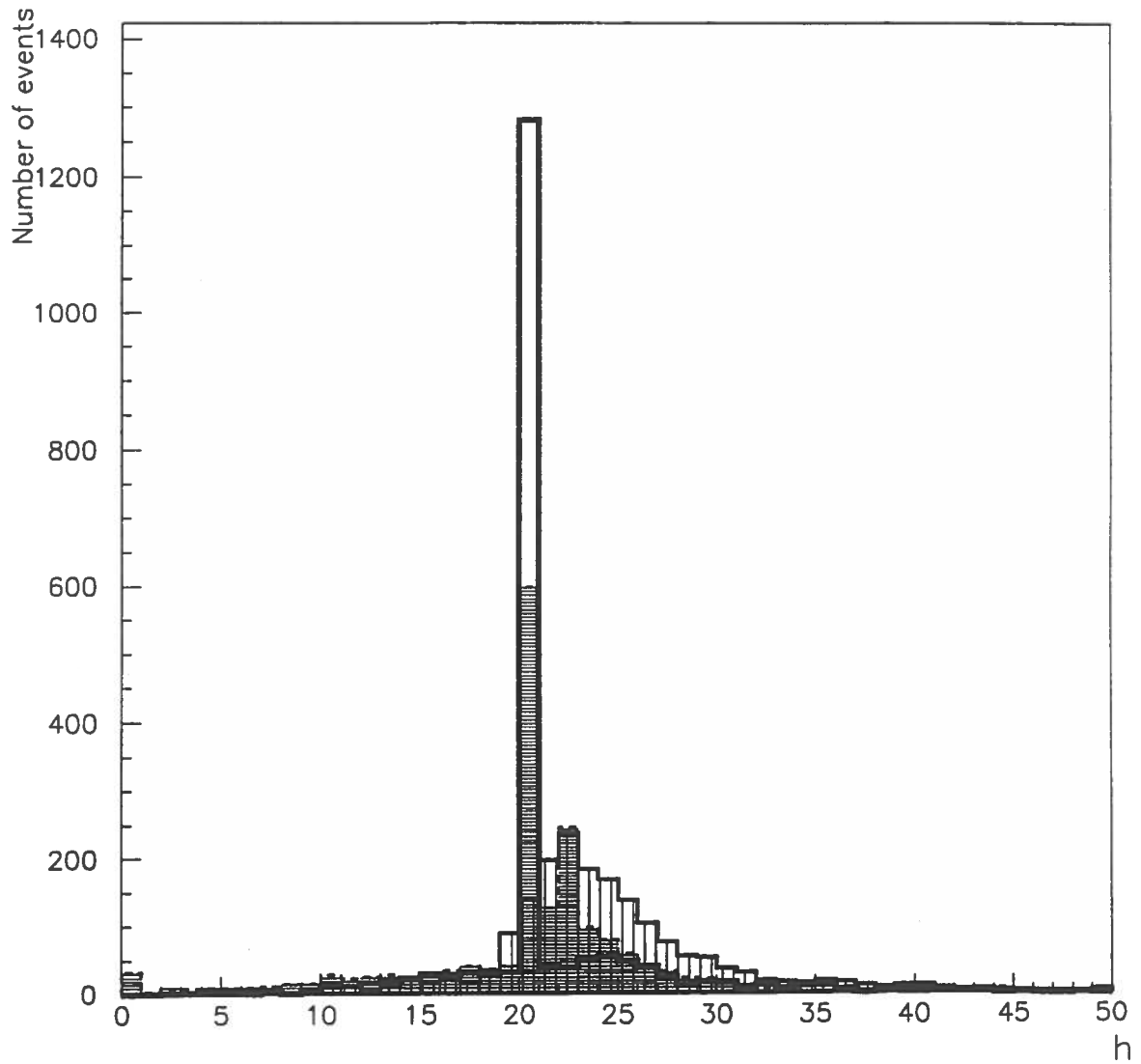
Figure 3: Distribution of the number of events vs. the number of hits h released in the calorimeter by 4 GeV pions. *Thick line:* Montecarlo simulations; *Vertical bands:* before filter; *Horizontal bands:* after filter.
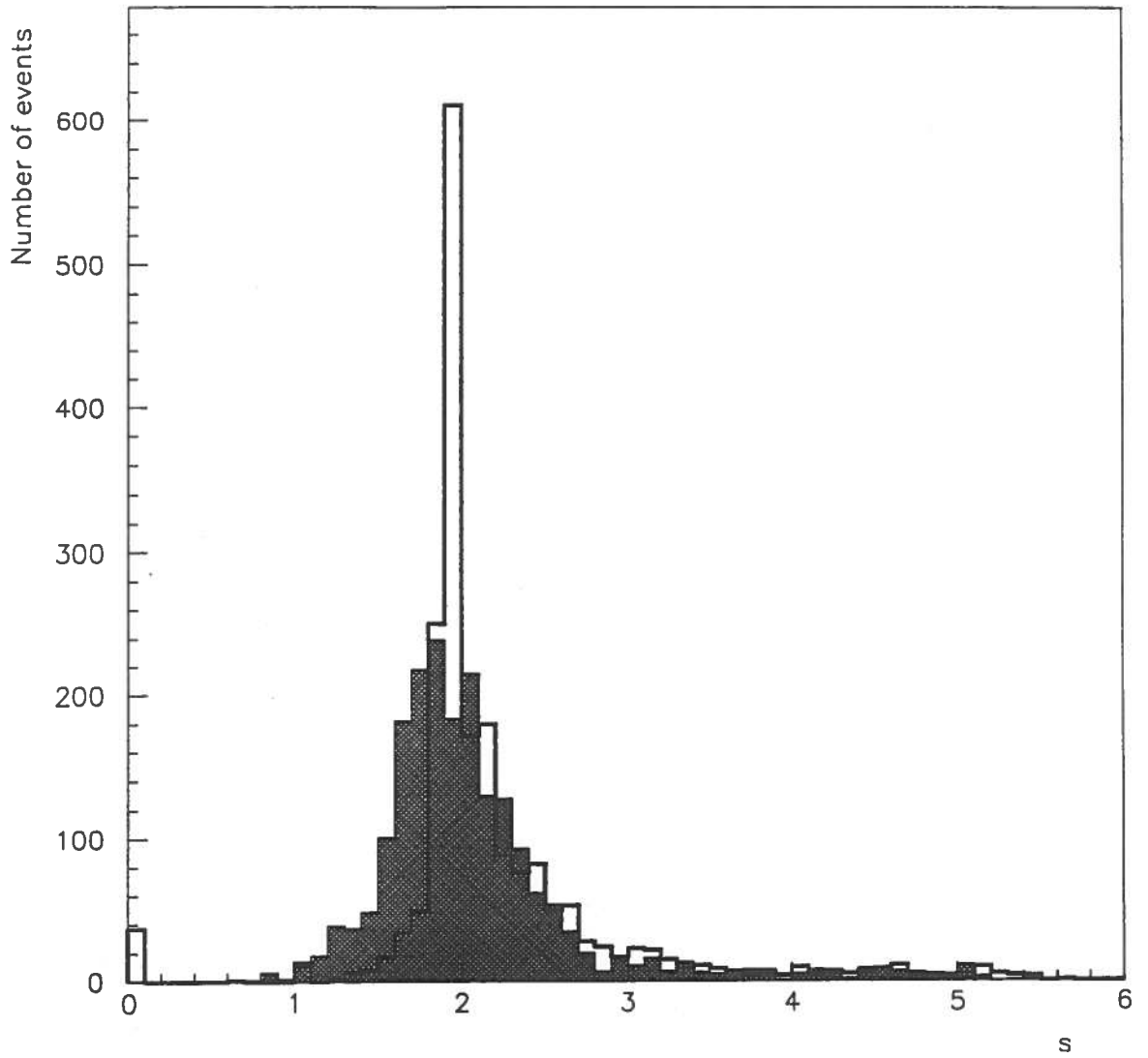
Figure 4: Distribution of the Shower aggregation *s* (see text) of 4 GeV pions before (*shaded area*) and after the CA application (*Thick line*).
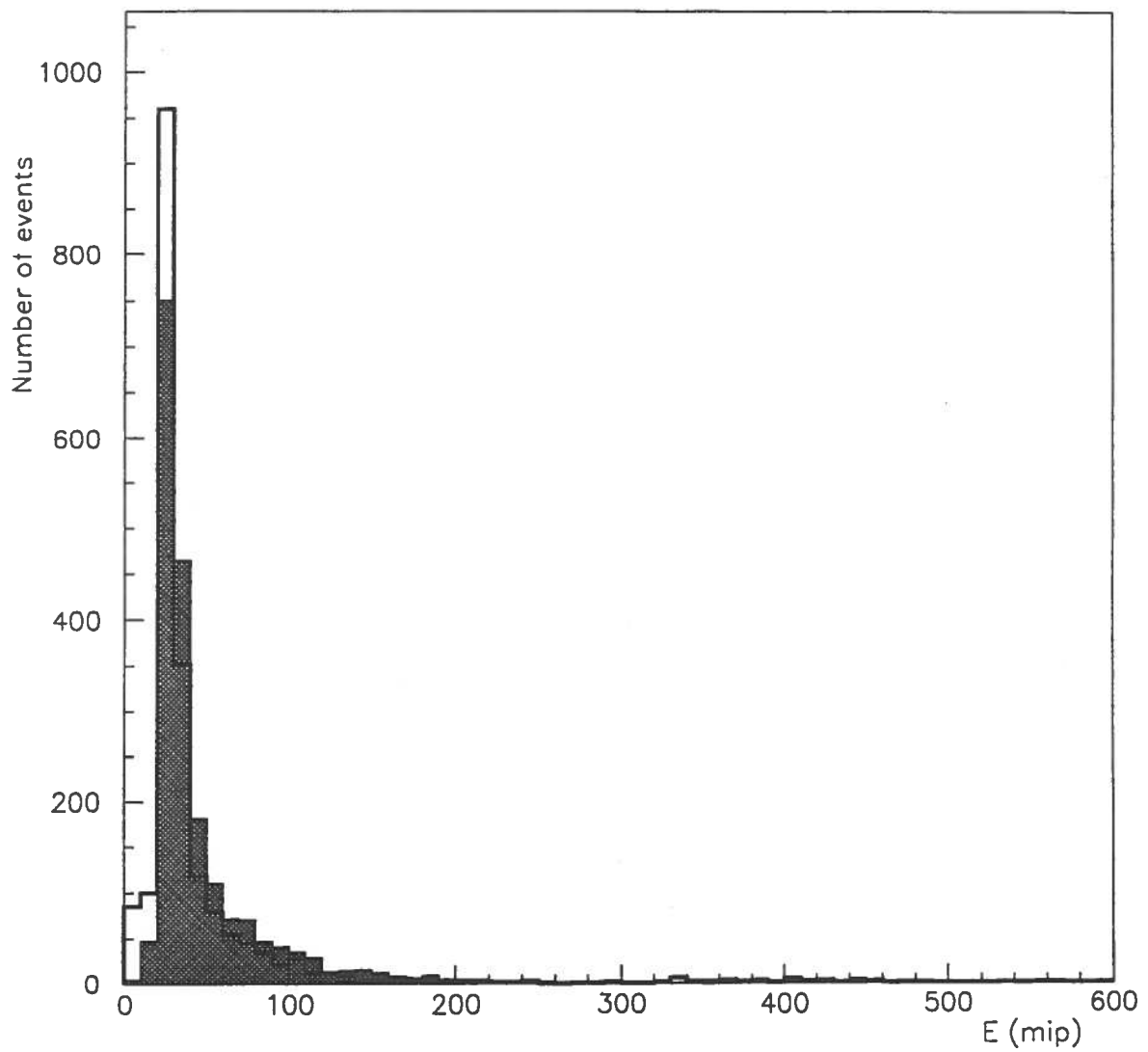
Figure 5: Distribution of the Total energy E (*mip*) released in the calorimeter before (*shaded area*) and after the CA filter (*thick line*) from 4 GeV pions.
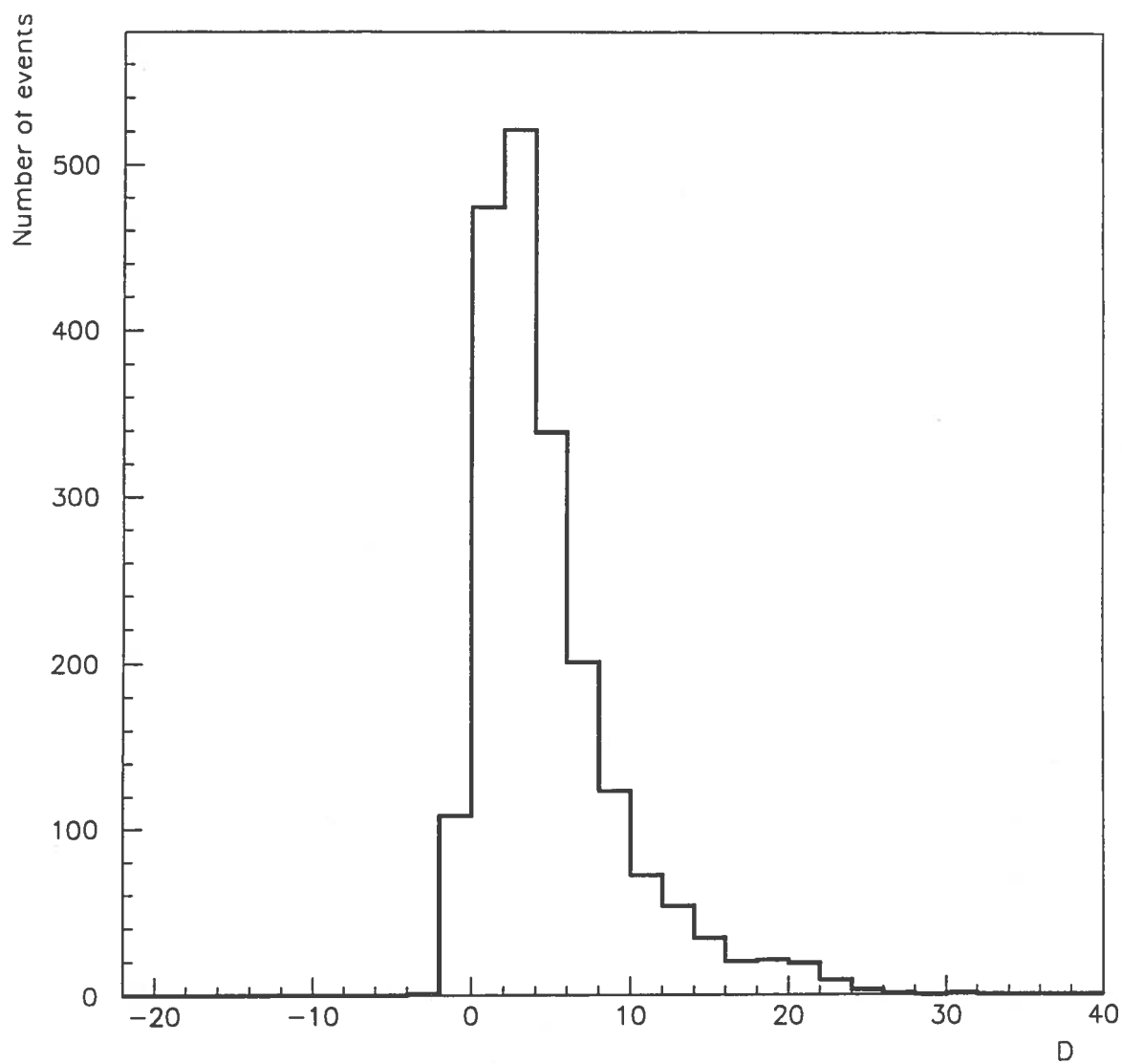
Figure 6: Distribution of the difference D between the number of hit strips before and after the use of CA.
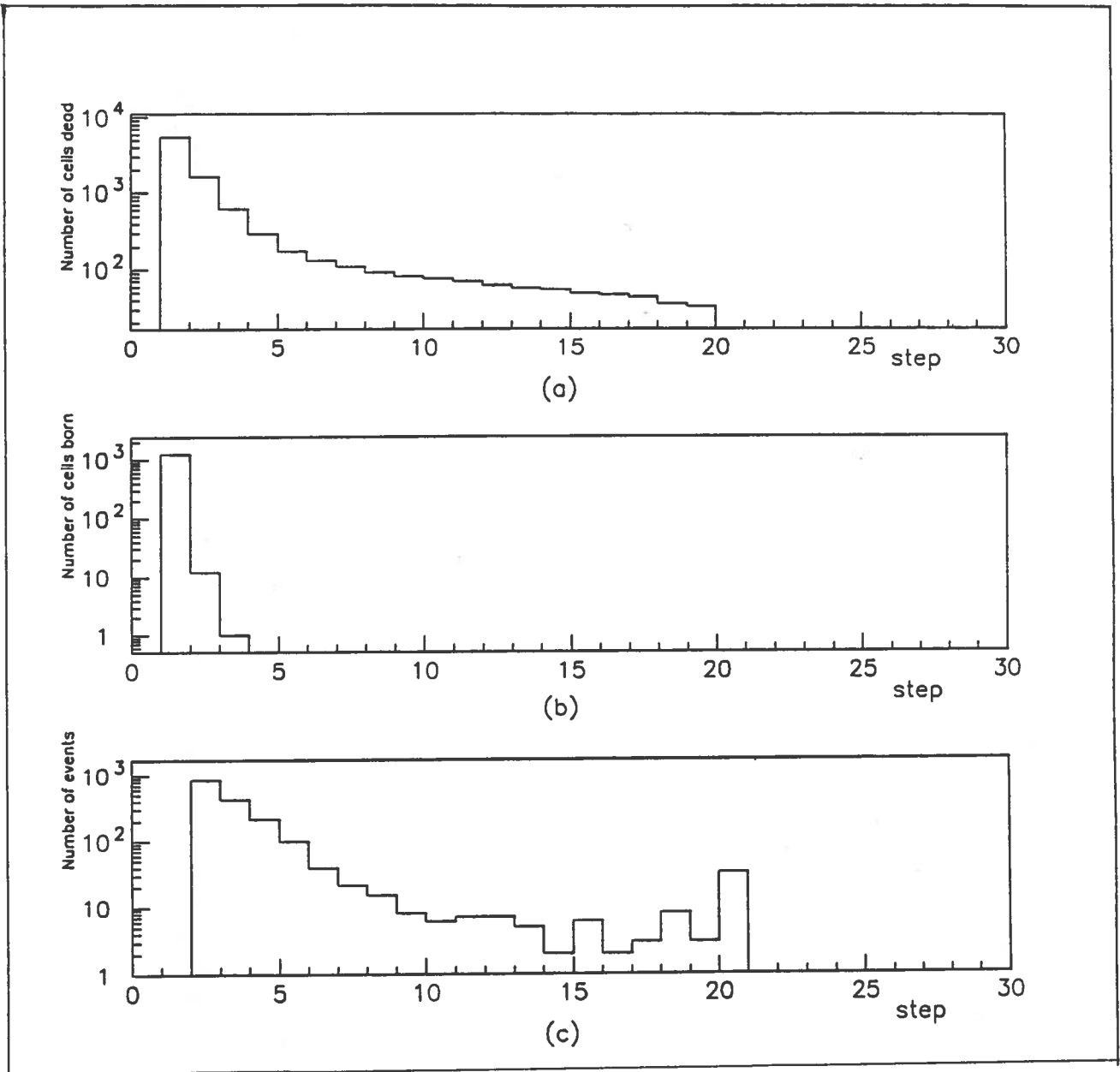
Figure 7: *(a)* Number of deaths vs. update cycles; *(b)* Number of births vs. update cycles; *(c)* Number of events vs. number of cycles required for convergence.
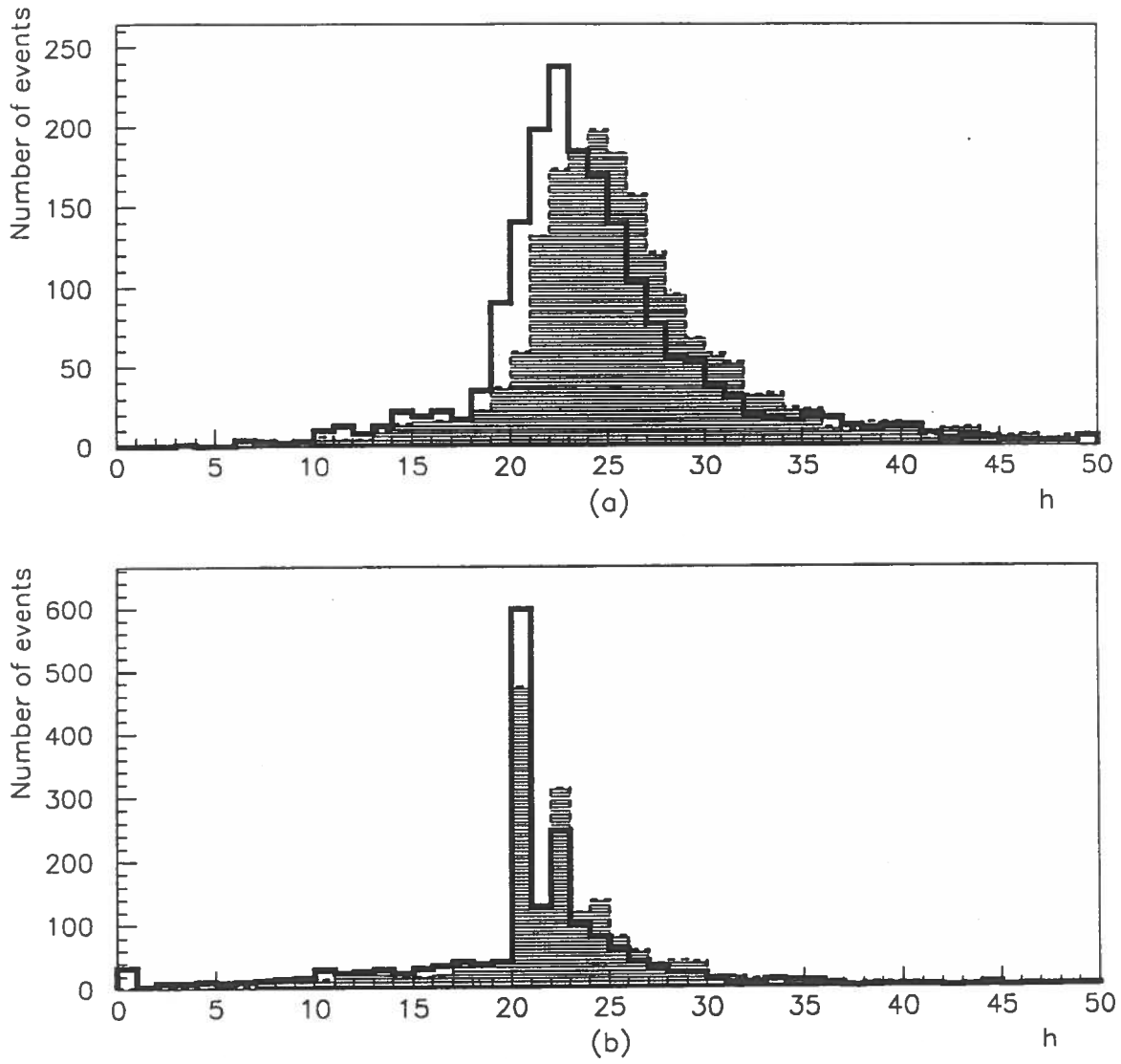
Figure 8: Distribution of the number of hit strips h in presence of low noise ($T = 0.7$ *mip* - *thick line*) and high noise ($T = 0.5$ *mip* - *banded*) before (a) and after (b) the use of the CA.