Neural Networks in HEP triggers

Ricardo Nóbrega [a], João Varela [b],
[a] *LIP, Av. Elias Garcia, 14 P-1000 Lisboa*
[b] *CERN, CH-1211 Geneva*

Neural Networks (NN) have been widely used as classifiers in physics analysis, but comparatively few applications in on-line triggers for high energy physics experiments exist. This paper describes NN based 1st and 2nd level e/γ calorimetry triggers at LHC.

Key words: NN trigger; e/γ triggers at LHC.

1 e/γ triggering in CMS

The 1st level e/γ trigger algorithm[1] in CMS[2] is depicted in figure 1. It combines shower containment criteria, longitudinally (observing the HCAL/ECAL ratio for the central 6×6 crystals trigger cell) and laterally (in both HCAL and ECAL), with the analysis of the *fine grain profile* of the shower, along η (perpendicular to the magnetic field).

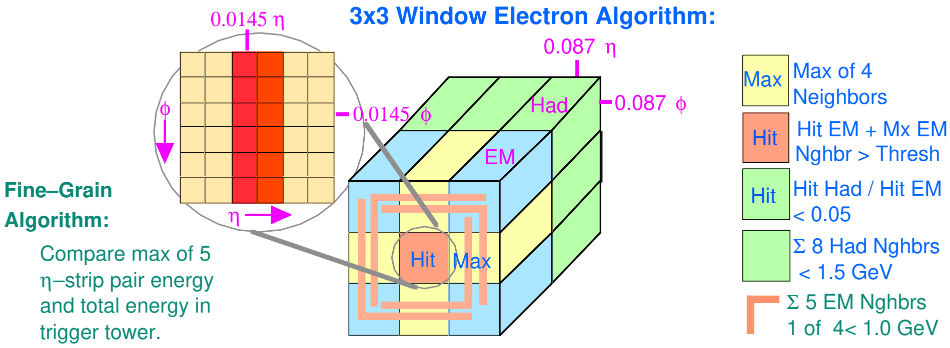

Fig. 1. Schematic view of the 1st level trigger algorithm

This analysis, performed at 40 MHz rate by L-Neuro2.3 chips[3], consists formally of a single neuron ω, dynamically chosen according to the six-dimensional pattern p on which it operates. One L-Neuro (LNSUM) computes the sum $\Sigma 6$ of all 6 strips; a bit $\mathrm{bit}_{\mathrm{sum}}$ will be set if $\Sigma 6$ exceeds a given threshold T. A second L-Neuro (LNMAX) determines which sum of strips m and $m+1$ gives the highest value, $\Sigma 2$, among all 5 possible pairs. This index m is used to select, within a third L-Neuro (LNELECT), the appropriate weights ω to process this pattern. Simulation results have shown that arbitrary ω provide no significant improvement compared to $\omega_i = \delta_{im} + \delta_{im+1} - t$. This is equivalent to comparing $R = \Sigma 2/\Sigma 6$ with a threshold t. Electrons and photons, selected by $R > t$, ($t \sim 0.9$), will have $\mathrm{bit}_{\mathrm{electron}} = 1$, whereas jets give $\mathrm{bit}_{\mathrm{electron}} = 0$.

1.1 Implementation; Beam tests

Prototype boards of the CMS 1st Level Trigger Primitive Generation were tested during the summer of '96. The boards, collectively designated by TPB (Trigger Primitives Boards)[4], consist of the System Board (developed by LIP-Lisbon) and of the Processor, Optical-Link and Transition Boards, developed at LPNHE-Palaiseau.[1] The input data to the Boards were generated by a prototype of the *FERMI microsystem*[5] (which included a 40 MHz sampling ADC and a *Filter Module*), based on the energy deposited by electrons with 15 GeV–150 GeV momenta in a 7×7 $PbWO_4$ crystal matrix. Independent measurements of the energy were performed by charge integrating ADCs, allowing a thorough study of the FERMI's analog compressor and digital filters.

The calorimeter digitized data arrives at the SB at 40 MHz frequency; the SB sends these data to, and accepts the outputs from, the PB card, where the L-Neuro chips are housed. PB includes modules for internal synchronization and computing a trigger bit, depending on the values of m, $\mathrm{bit_{sum}}$ and $\mathrm{bit_{electron}}$. These data flow through FIFOs of programmable depth, located in the SB; upon reception of the external trigger signal, data input is stopped after a given delay, so that at stop time, the time profile of the shower, along with the L-Neuro outputs, are recorded. When ready to output, SB allows the VME readout of all FIFOs, see figure 2.

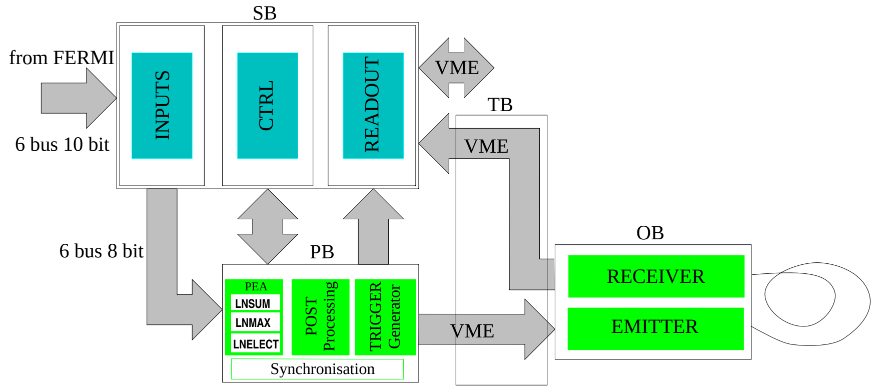

Fig. 2. Synoptic of TPB

The main aspect of data analysis, from the trigger point of view, is the validation of the fine grain trigger concept. Figure 3a shows the reasonable agreement between a reduced R variable computed from beam and simulated data; and in figure 3b, we see the R variable computed from on-line data: a crude decompression LUT was charged in FERMI, and no off-line corrections were applied to the filtered values. The distribution stands well above the 90% level, allowing a safe operation of the trigger.

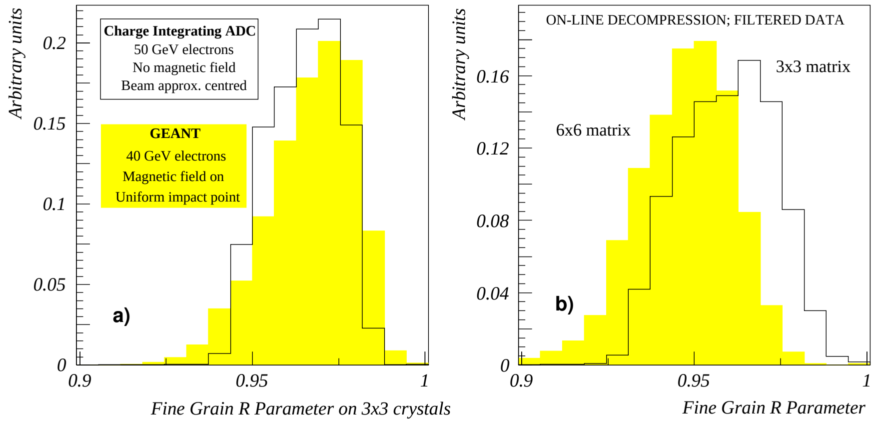

Fig. 3. Some beam test results. a) comparison of R obtained with `GEANT` and beam data. b) R for on-line filtered data, in 3×3 and 6×6 matrices

[1] A similar set of SB+PB boards was already considered for triggering on non-asymptotic Cherenkov rings [6], generated in a gaseous RICH.

1.2 2^{nd} level e/γ trigger

Coping with the huge background rejection ratios required at LHC asks for an exhaustion, in as early times as possible, of the information provided by the detectors. We studied the possibility of implementing a NN-based 2^{nd} level calorimeter trigger on commercial CPUs, using the full calorimeter granularity. It is expected that this level of trigger will reduce by one order of magnitude the trigger rates after the 1^{st} level trigger.

The background on e/γ after the 1^{st} level trigger is predominantly due to $\pi^0 \to \gamma\gamma$: we base our trigger strategy on the assumption that it is possible to enhance the cluster's tails, in order to resolve the nearby photon. Starting from the region indicated by the L1 trigger, we focus our attention on the highest crystal, attaching axis (η, ϕ) to its centre, and operate a transformation, inspired by the symmetries present in the calorimeter, that reduces four quadrants into one. We take $n = 6$ and index a matrix of $2n+1 \times 2n+1$ crystals by $(\eta, \phi) = (-n \ldots n, -n \ldots n)$. Each crystal in $\eta \geq 0$ is added to its $-\eta$ partner; each crystal in $(\eta \geq 0, \phi \geq 0)$ is added to $(\eta, -\phi)$; only $(\eta \geq 0, \phi \geq 0)$ crystals are kept. The 49 *square roots* of the crystals' energies provide a pattern that is processed by a modular NN[8], whose four specialized neurons deal with the four kinds of showers shown in figure 4.

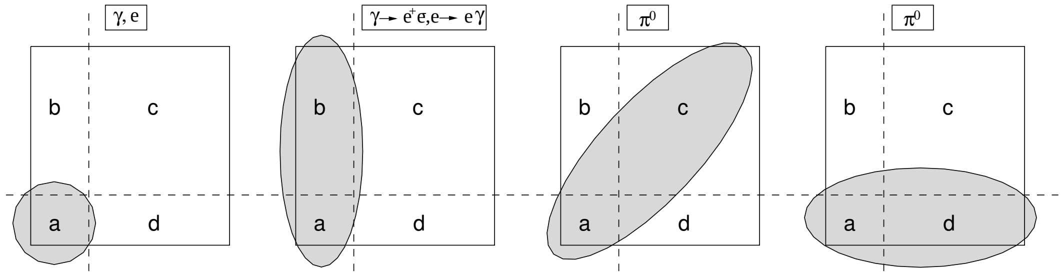

Fig. 4. Separation of tasks in the modular neural network. According to the relative weights of the regions a, b, c, and d, a different neuron will classify the pattern

The results of this study are summarized in table 1. The L2 trigger excludes close to 80% of jets passing the L1 trigger; further, it enhances the trigger energy resolution, allowing a sharper E_T cut. These two effects combined allow for a factor 7 reduction in trigger rates at high $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$ luminosity. In terms of speed of execution, we observed that an implementation of this algorithm took about 15 μs/event on an RS/6000 machine.

Table 1
Background removed by the L2 trigger at low luminosity ($\mathcal{L} = 10^{33}\text{cm}^{-2}\text{s}^{-1}$). The label "Linear" means that SQRT was not applied to the patterns. The case with one neuron corresponds to a single classifier for the whole calorimeter; when two neurons were considered, the endcap and the barrel were processed independently. Finally, with eight neurons, the full modular network was used

Energy scale	Number of neurons	QCD removed (%)
Linear	1	58
Linear	2	65
Linear	8	70
SQRT	8	78 (71 at high $\mathcal{L}$)

References

[1] M. Bercher et al., CMS TN/96-10

[2] CMS collaboration, CERN/LHCC 94-38

[3] M. Bercher et al., CERN/LHCC 95-56

[4] Ph. Busson et al.,CMS IN 1996/8

[5] H. Alexanian et al., Nucl. Instr. and Meth. A 357 (1995) 306

[6] R. Nóbrega and J. Varela, in: Proc. AIHENP, April 95, Pisa

[7] R. Nóbrega and J. Varela, CMS NOTE/96-10

[8] S. Haykin, Neural Networks (Macmillan College Pub. Company, Inc., 1994)