X Congresso Brasileiro de Inteligência Computacional (CBIC'2011), 8 a 11 de Novembro de 2011, Fortaleza, Ceará

Neural *leptonjets*/jet discrimination in highly segmented calorimeters

Danilo Lima de Souza, José Manoel de Seixas
Signal Processing laboratory - LPS/COPPE
Federal University of Rio de Janeiro - UFRJ
{danilo,seixas}@lps.ufrj.br

Abstract – This paper presents a proposal of *leptonjets*/jet discriminator system for operating at the Second Level Trigger of ATLAS. The system processes formated calorimetry data, so that high performance can be achieved. The data processed, based in specialist knowledge, feed linear or non-linear classifiers with somewhat close performance. This implementation resulted on high *leptonjets* detection efficiency for a low false alarm.

Keywords – LHC; ATLAS; Neural Networks; Pattern Recognition; leptonjets

1. Introduction

Data filtering system are used in different fields of research, aiming at isolating signal of interest from patterns related to a given background noise. Actually, in many complex applications, the input data space dimensionality is very high, as well as the incoming data rate. In this case, the difficulty of the input data stream analysis increases significantly. Also, the processing speeding plays a critical role when the filtering system is envisaged for online operation. Finally, the signal of interest may rarely occur, forcing the experiment to keep running for a long period of time in order to acquire a reasonable amount of events for better measure estimation. In general, online filtering system should have the following features:

- High detection efficiency for a low false alarm probability.
- Flexibility in order to accomplish future requirements.
- Execution speed capable of achieving the desire time requirements.
- Robustness, in order to keep its filtering features trough its liftime operation.

To cope with such high-input data dimension, smaller selected area of analysis and feature extraction techniques may be applied in order to isolate the relevant information from the event data description, eventually reducing its dimension. For this, different techniques have been developed using expert information or/and stochastic processing. Inteligent processing schemes may even reduce the complexity of the classifier (signal against background) design.

In many applications, neural networks [1] may play a role in signal classification. On the other hand, reducing the classifier design complexity by means of signal processing, it might be possible to restrict the nonlinear processing implemented by a neural network to perform slight adjustments in the linear signal classification. It may also be the case where signal classification can go linear (through a Fisher Discriminant) [2], as a result of highly-efficient processing scheme.

In experimental high-energy physics, stringent conditions make signal processing a challenge, as there is often a large gap between the experiment requirements and the technology currently avaible, forcing the development of new technologies. This is particulary the case for modern particle collider experiments, in which particles are accelerated at high speed and put in collision route. Analyzing the resulting products of collision, one can probe deeper into the structure of matter [3]. One important aspect in particle collider experiments is that events of interest are typically very rare, since most of the produced events are from background noise. In addition, the fine-grained segmentation of the particle detectors placed around the collision point for the resulting interaction readout produces terabytes per second of information. Therefore, an online filtering system must be applied for selecting only physics channels, while rejecting, as much as possible, the huge amount of background noise.

Presently, the Large Hadron Collider (LHC) at CERN [4] is the largest particle accelerator in the world. LHC has a total length of 27km and will be colliding protons with 14TeV at their center of mass, at a rate of 40MHz and at a luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ [3]. Multiple collisions points occur around the LHC ring. Around each collision point, a detection laboratory is placed to analyze the sub-products of the collisions. Among such detectors, ATLAS [5] is the largest one. It may comprises multiple sub-detectors, such as tracking, calorimeter and muon detection systems. Due to the detector granularity, each collision produces ≈ 1.5MBytes of information, resulting in a total rate of ≈ 60TB/s of information. Therefore, an online filtering system is mandatory for proper ATLAS operation.

This paper is organized as follow: in section 2 the ATLAS filtering system is showed, as well as the proposed procedure. In section 3 the obtained results for such application are discussed. Finally, conclusions are derived in section 4

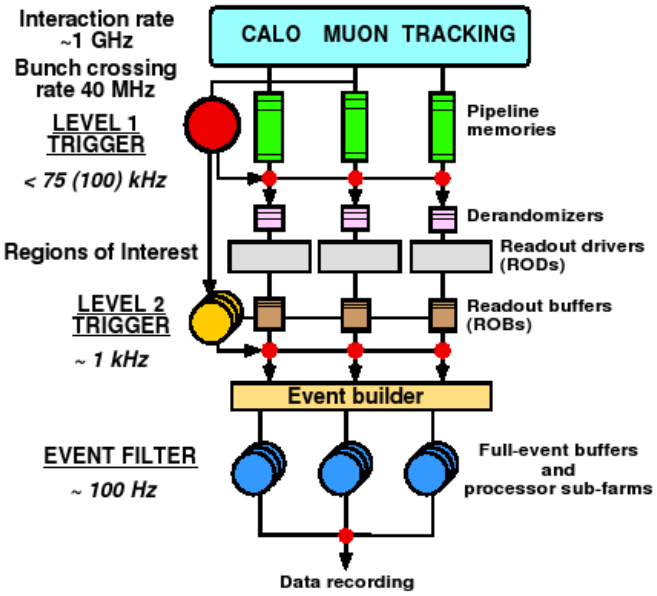


Figura 1: Block diagram of the ATLAS Trigger/DAQ system.

2. The proposed *leptonjets* ATLAS channel

The filtering system comprises three cascaded operation level, applying sucesive cuts to the incoming data [6], as showed in figure 1.

The first level (LVL1) receives full data and reduces the input event rate to $\approx$ 75kHz. It is responsible for marking the regions in the detector that have effectively been excited. These regions are known as *Regions of Interest* - RoI, and are the only information passed to the second level analysis. A RoI selected by the first-level filtering system amounts, in average, to $1,000$ calorimeter cells.

The second level (LVL2) receives the RoIs marked by the first level and it does a more specific analysis on them. For meeting a time latency requirement of 14ms per event, a set of 500 off-the-shelf server processors are employed, providing a multi-processed environment. The third and last filtering level, known as Event Filtering - EF, takes the final decision on events approved by previous levels.

One of the ATLAS main research goals is to experimentally prove the Higgs boson existence [7]. Being the Higgs boson highly unstable, it soon decays into more stable particles. Therefore, the physicists will prove its existence not by detecting the Higgs boson directly, but by analyzing its decaying signatures. The same occurs with other unexperimentally proved theories.

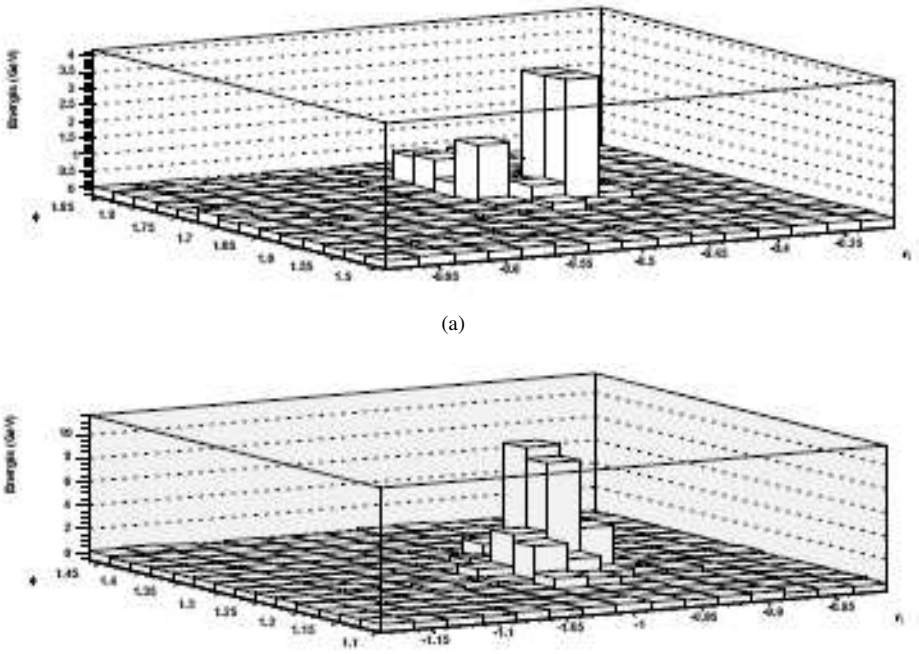
An interesting issue in Experimental Physics is Dark Matter (DM). In light of recent astrophysical observations, the authors of [8] have proposed a broad class of theories in which the annihilation of $\approx$TeV scale dark matter in the galactic halo accounts for the anomalous excess of cosmic ray leptons. While this dark matter is probably inaccessible to colliders, it is accompanied by a $\approx$GeV scale dark sector which couples, albeit very weakly, to the standard model (SM). The dark sector states are relatively light and can be produced at high energy colliders, only to cascade decay through the dark sector and ultimately return to the visible sector as electrons, muons, and possibly pions. The outgoing leptons emerge in the detector as *leptonjets* [9], which are highly collimated multiple muons or electrons that result from the decay of these highly boosted dark sector states. Recently, experimental effort in searching for such objects has been reported in [10].

On the other hand, a cascade of quark and anti-quark pairs can be produced during the vector decay, which quickly merge into more stable particles, producing a pattern known as jet. These jet may deposit their energy in a manner very similar to electrons (leptons), making the correct identification of *leptonjets* a tricky process. Our analysis will focus on the *leptonjets*/jet separation problem at the second level of the ATLAS filtering system.

The system is based in calorimeter information. Calorimeters are total absorption detectors [7]. Typically, they use a (passive) material for absorbing the energy of the incoming particle and sample the energy being deposited in the detector by using a active material. Calorimeter play a major role in collider experiments. They provide fast response, their energy resolution improves with increasing energy, and they interact with charged and non-charged particles.

The ATLAS calorimetry system is composed by two calorimeter sections [7]: the electromagnetic (EM) calorimeter is responsible for detecting electrons, positrons and photons. The hadronic (HAD) calorimeter is responsible for detecting hadrons (kaons, pions etc) and it is placed on top of the electromagnetic calorimeter. Both detectors comprise 3 sequential layers with distinct granularity and depth providing detailed information of incoming paticles. The electromagnetic calorimeter has, in addition, a very thin layer in front of it, which is called Pre Sampler (PS). The figure 2 displays the energy deposition profile for *leptonjets* and jet interaction at the eletromagnetic middle layer.

Isolated electrons have the property of depositing their energy in a punctual way, differently from jets, which, for LVL2 data, tend to slightly spread their energy over multiple cells within a layer. For *leptonjets*, the collimated feature of their leptons would



(a)

(b)

Figura 2: Example of segmnted calorimeter information obtained from the eletromagnetic middle layer: (a) an incoming *lepton-jets*; (a) an incoming jet.

lead to properties similar to isolated eletrons events but the multiplicity of leptons in it results in energy over multiple cells as well as the jets events.

Therefore, the relevant information relies not at the impact point center, but at its surrounding area. Aiming at exploiting this feature, a topological procedure based on ring sums has been tried by some LVL2 algorithms for data formatting [11, 12]. In this approach, the cell that samples the highest energy value (also known as the hottest cell) is considered the center of our RoI in each calorimeter layer (seven in total). Then, a set of concentric rings are built around this hottest cell in a pattern similar to the one presented in figure 3. The ringer procedure is performed on a per layer basis, resulting, at the end, in a total of 100 rings per RoI (table 1).

Tabela 1: Number of rings per layer of calorimeter.

Layers	Rings			
	RoI	$\Delta R < 0,1$	$\Delta R < 0,2$	$\Delta R < 0,3$
Pre Sampler	8	1	2	3
EM *front layer*	64	1	2	3
EM *middle layer*	8	3	6	8
EM *back layer*	8	1	2	3
HAD *layer* 0	4	1	2	3
HAD *layer* 1	4	1	2	3
HAD *layer* 2	4	1	1	2

Even with the ringer procedure, the number of rings represents a high dimension space of analysis. To further reduce this number others studies are being investigated. According to the literature [13] a smaller part of the RoI could be used in the process of detection/classification. This small concentric region, based on the energy profile around the center of the RoI, would be enough to separate the patterns. This region is determined through

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{1}$$

where η and ϕ are the coordinates of the ATLAS detector, and $\Delta\eta$ and $\Delta\phi$ are the respective granularities in each layer. The figure 4 ilustrates these regions for diffents values of ΔR. The table 1 shows the number of rings per layer cosidering three diferents values of ΔR. All of them represents a considering smaller space of analysis.

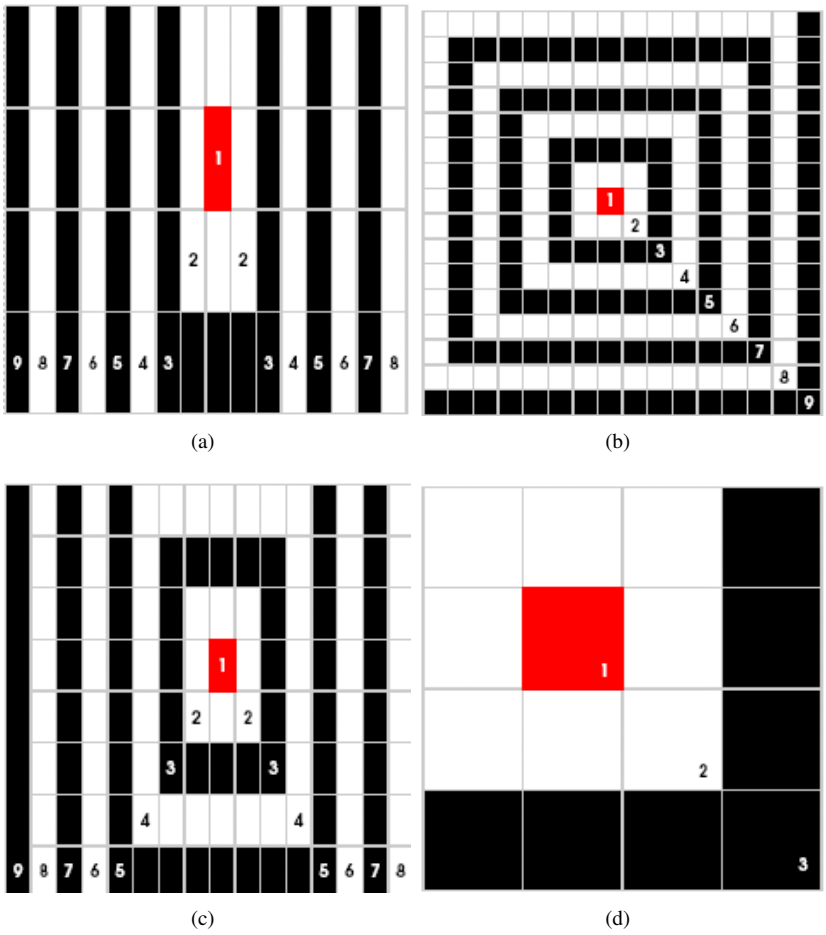


Figura 3: Ring formatting for calorimetry: (a) *Pre Sampler*; (b) EM *middle layer*; (b) EM *back layer*; (d) HAD *layer* 0.

3. Results

The available data set was obtained through Monte Carlo simulation and comprises apprimately $450,000$ jets and $1,000$ *leptonjets* (4 electrons) signatures. The simulation considers the detector characteristics and the first-level filtering operation. The jet data was equally split into trainning, validation and testing sets. The *leptonjets* were split in two sets. The first one, with 75% of the events, was replied until achieve the number of jets used in the trainning stage. The procedure adopted to equalize the *a priori* probabilities of the two events, since the simulation of *leptonjets* is a hard task. The rest of the *leptonjets* were used in the validation and test stages.

It is presented in figure 5 the block diagram of the *leptonjets*/jet discriminator. The raw calorimeter data is received and the topological procedure on ring sums is performed. Next, the rings feed the classifier algorithm for the hypothesis making. In this work the Fisher Linear Discriminant (FLD) and the Supervised Multi-Layer Perceptrons (MLP) neural classifiers were used to perform the particle identification (hypothesis testing) over processed calorimeter information. The neural network were trained using the backpropagation algorithm [1]. In order to compare the discrimination efficiency for the proposed classifiers, the SP index [1, 14] was applied. The SP index is computed through

$$\mathrm{SP} = \sqrt{\frac{\sum_{i=1}^{j} \mathrm{PD}_i}{j} \times \sqrt[j]{\prod_{i=1}^{j} \mathrm{PD}_i}} \tag{2}$$

where PD_i is the efficiency for the ith class of event. The threshold value that maximizes the SP provides both high PD and low PF.

The table 2 shows the final performance, where $\mathrm{PD}_{\mathrm{lj}}$ is the *leptonjets* acceptance rate, PF_{j} is the false alarm rate and, $\mathrm{SP}_{\mathrm{lj/j}}$ is the SP index considering the *leptonjets*/jet separation. Accordind to the numbers, we can say that the Fisher classifier can not correctly separete the patterns by the rings from the RoI. The poor $\mathrm{PD}_{\mathrm{lj}}$ prevents its use. The neural classifier worked much better, mainly because its nonlinear properties can contour the problems where rings presents close cluster to the signal and the background. The neurons in the hidden layer ranged from 2 to 15. The best topology was obtained with 7 neurons.

As mentioned in previous sections, the available calorimeter signatures are topologically formated, generating 100 rings for an incoming event. This number is much smaller than the avareged number of cell in RoI, but it is still considered high. The investigation of the potencial of separation with a smaller number of rings presents interesting results. The Fisher Linear

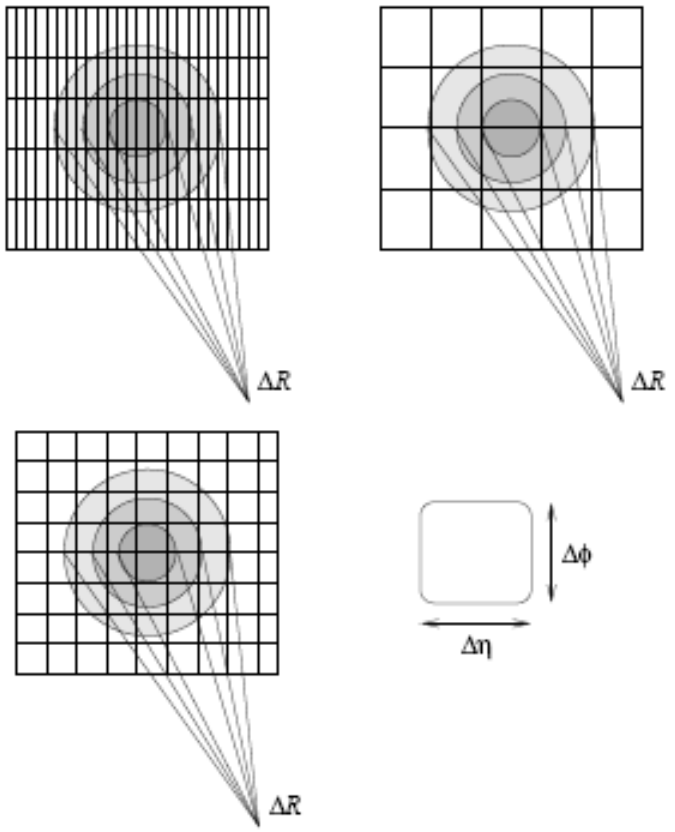


Figura 4: ΔR regions of the RoI per layer.

Figura 5: Processing chain of the *leptonjets*/jet separation system.

Discriminant has a huge improvement with the three values of ΔR studied. It means, the rings in the edge of the RoI (almost no energy) were, as we supposed, the responsible by the confusion between the patterns. We can observe that increasing the ΔR implies a better performance, as shows the $SP_{lj/j}$ index. We consider the maximized functional $SP_{lj/j}$ as the indicative of the best classifier. From this, with a $SP_{lj/j}$ value of 0.7243, the best Fisher classifier can accept 58.56% of the *leptonjets* with a false alarm rate of 12.22%.

In order to verify whether a non-linear classifier suffices, the signals were also used to feed a neural network. We varied the number of hidden neurons from 2 only to 5, because of the smaller number of rings. Analogously, to that observed with the FLD, increasing the ΔR implies a better performance.

The best neural classifier presents the topology $25 - 5 - 1$. The use of only 25 rings (belonging to the region of $\Delta R = 0.3$) means 75% of compactation. The final $SP_{lj/j}$ index, indicates, also, a somewhat close performance to obtained with the processing of the RoI. Another interesting observation is that for all the ΔR scenarios investigated, the final performance of the FLD was close to the presented by the neural classifier. As neural classifiers are much more fast to train and perform, and much more easy to investigate, it is a positive issue.

It can be seen that the ΔR processing was able to improve the discrimination obtained trough both classifiers (linear and non-linear) providing more separated patterns. Through the proposed processing chain, the results of the linear classifier indicate that is not necessary a neural classifier (non-linear) to achieve good detection efficiency. An important issue is the computational cost, which, for a linear classifier is smaller. This might be a striking advantage for a online filtering. However, if the detection of *leptonjets* is mandatory a less complex neural network topology could be employed.

Tabela 2: Performance of the Fisher and Neural classifiers.

Region of Analysis		Classifier	Hidden Neurons	PD_{lj}	PF_j	$SP_{lj/j}$	threshold
RoI		Fisher		46.39	46.42	0.4992	-0.09
		Neural	7	68.44	17.28	0.7541	0.07
ΔR	0.1	Fisher		66.92	27.17	0.6984	-0.19
		Neural	4	72.62	31.95	0.7032	0.11
	0.2	Fisher		63.50	21.42	0.7084	0.01
		Neural	5	62.36	16.20	0.7268	0.25
	0.3	Fisher		58.56	12.22	0.7243	0.29
		Neural	5	68.06	20.46	0.7369	0.13

4. Conclusions

The hard triggering task of high-dimensional data find parallel aplications in many areas. Combining accumulated knowledge by experts about the problem with inteligent signal processing techniques is being shown to be an efficient design approach. Among the benefits, high signal compactation rates, relevant feature extraction and reduced computational load are accomplished.

It was shown that processing the calorimeter data with the knowledge provided by expert play a interesting role for additional applications to be developed at the actual infrastructure of the ATLAS detector. Efficient feature extraction, about 25% of the RoI rings data, reduced significantly the computational load, wich is attractive due to the low latency time required by the application.

5. Acknowledgment

The authors would like to thank the CNPq (Brazil) and CERN (Switzerland) for the support. We also thank the ATLAS Trigger/DAQ collaboration at CERN for providing the simulated calorimeter data as well as Bilge Dermikoz for the interesting discussions about *leptonjets*.

References

[1] S. Haykin. *Neural Networks: A Comprehensive Foundation (2nd Edition)*. Prentice Hall, second edition, July 1998.

[2] R. O. Duda, P. E. Hart and D. G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, second edition, November 2000.

[3] L. Evans and P. Bryant. "LHC Machine". *Journal of Instrumentation*, vol. 3, no. 08, pp. S08001, 2008.

[4] CERN. "LHC - The Large Hadron Collider". `http://public.web.cern.ch/public/en/LHc/LHC-en.html`, 2008.

[5] CERN. "The ATLAS Experiment, Mapping the Secrets of the Universe". `http://atlasexperiment.org/`, 2007.

[6] The ATLAS TDAQ Collaboration. "ATLAS High-Level Trigger Data Acquisition and Controls Technical Design Report". Technical report, CERN/LHCC/2003-022, May 2003.

[7] G. Aad *et al.*. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics". 2009.

[8] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer and N. Weiner. "A Theory of Dark Matter". *Physical Review D*, vol. 79, pp. 015014, 2009.

[9] N. Arkani-Hamed and N. Weiner. "LHC Signals for a SuperUnified Theory of Dark Matter". *JHEP*, vol. 0812, pp. 104, 2008.

[10] D. C. V. Abazov. "Search for NMSSM Higgs bosons in the h->aa->mumu mumu, mumu tautau channels using ppbar collisions at sqrt{s}=1.96 TeV". *Physical Review Letters*, vol. 103, pp. 061801, 2009.

[11] M. R. Vassali and J. M. Seixas. "Principal component analysis for neural electron/jet discrimination in highly segmented calorimeters". volume 583, pp. 89–91. AIP, 2001.

[12] D. L. de Souza and J. M. de Seixas. "Statistical Processing on High Resolution Calorimetry Information". In *WCI 2010 - III Workshop on Computational Intelligence*, pp. 1–6, 2010. ISBN:978-85-7669-250-8.

[13] C. Cheung, J. T. Ruderman, L.-T. Wang and I. Yavin. "Lepton Jets in (Supersymmetric) Electroweak Processes". *JHEP*, vol. 04, pp. 116, 2010.

[14] A. dos Anjos, R. Torres, J. Seixas, B. Ferreira and T. Xavier. "Neural triggering system operating on high resolution calorimetry information". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 559, no. 1, pp. 134 – 138, 2006. Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research - ACAT 05.