

Online Particle Detection by Neural Networks Based on Topologic Calorimetry Information

T. Ciodaro¹, D. Deva¹, J.M. de Seixas¹, D. Damazio²

¹Federal University of Rio de Janeiro

²Brookhaven National Laboratory

E-mail: ciodaro@cern.ch, dhiana.deva@cern.ch, seixas@lps.ufrj.br, damazio.cern.ch

Abstract. This paper presents the last results from the Ringer algorithm, which is based on artificial neural networks for the electron identification at the online filtering system of the ATLAS particle detector, in the context of the LHC experiment at CERN. The algorithm performs topological feature extraction over the ATLAS calorimetry information (energy measurements). Later, the extracted information is presented to a neural network classifier. Studies showed that the Ringer algorithm achieves high detection efficiency, while keeping the false alarm rate low. Optimizations, guided by detailed analysis, reduced the algorithm execution time in 59%. Also, the payload necessary to store the Ringer algorithm information represents less than 6.2 percent of the total filtering system amount.

1. Introduction

Pattern recognition based on artificial neural networks is widely used in many scientific areas achieving good performances [1]. For example, neural networks can also be found in several high-energy physics experiments, for particle identification [2], particle tracking [3], and other related issues [4]. The Large Hadron Collider experiment (LHC), in CERN [5], is a suitable environment for employing artificial neural networks [6, 7].

The LHC accelerates and collides proton beams at an event rate of 40 MHz, with up to 14 TeV of energy at the center of mass (currently, it operates at 7 TeV). At this unprecedented energy level, interesting physics process never seen before, such as the Higgs boson, can be studied.

ATLAS is the largest LHC particle detector [8]. In order to provide detailed information about the collision subproducts, ATLAS is split into three main subdetectors: the inner detector, which is responsible for tracking particles; the calorimeters, responsible for measuring their energy; and the muon chambers, which aim at tracking and measuring the momentum of muons. Figure 1 shows the ATLAS detector.

Altogether, ATLAS furnishes around 1.5 MB of information per collision event [9], which must be stored and analyzed. Considering the LHC collision rate, the total ATLAS data flow is around 60 TB/s. As much of the events don't contribute to the new physics which is the goal of the experiment, a three level online filtering system (trigger) was designed to select interesting events [9].

This article presents the latest results of the proposed Ringer algorithm [10], which combines both intelligent feature extraction, from the ATLAS calorimetry system, and neural networks to select electrons from the background at the ATLAS second level trigger system. Optimizations are applied for improving the processing speed and the whole detection performance is evaluated through algorithm implementation at the trigger framework using simulated data, which include pileup effects.



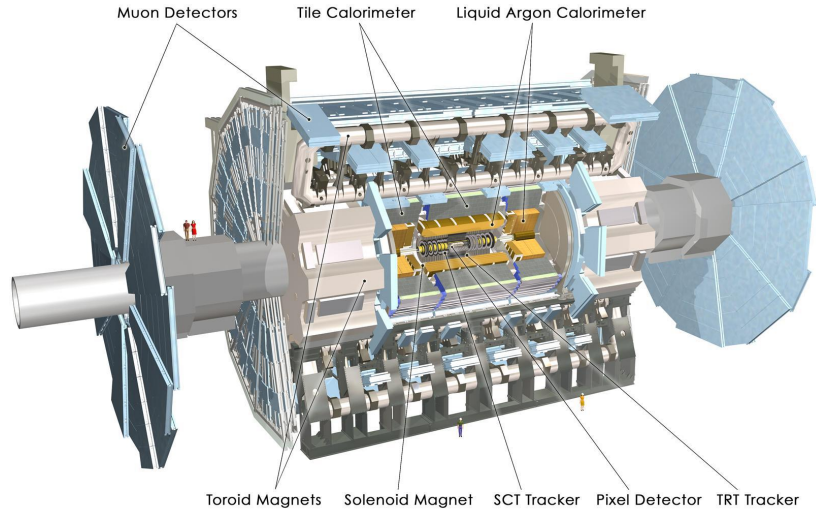


Figure 1. The ATLAS detector.

The article is divided as the following. Section 2 describes the ATLAS calorimeters and Section 3 details the ATLAS trigger system. The Ringer algorithm implementation is presented in Section 4. Section 5 discusses the methodology and the results achieved, while conclusions are derived in Section 6.

2. Calorimetry

ATLAS calorimetry system [11] is divided into two sections: electromagnetic (e.m.) and hadronic (had.), as can be seen on Figure 2(a). The calorimeters are segmented in-depth into layers. The resultant granularity (number of cells per layer) is layer dependent and it was designed to efficiently characterize the desired particles.

The e.m. calorimeter uses liquid Argon technology [12] and is divided into two parts: the barrel and the end-caps. Besides three layers, the e.m. calorimeter also possess a pre-sampler, which is used to correct for energies lost in dead materials in front of the calorimeter (solenoid, cryostat, ...), totalizing four layers. Because of its granularity, the e.m. section comprises hundreds of thousands of readout cells.

The hadronic calorimeter, however, uses two technologies: scintillating tiles (TileCal) [13]

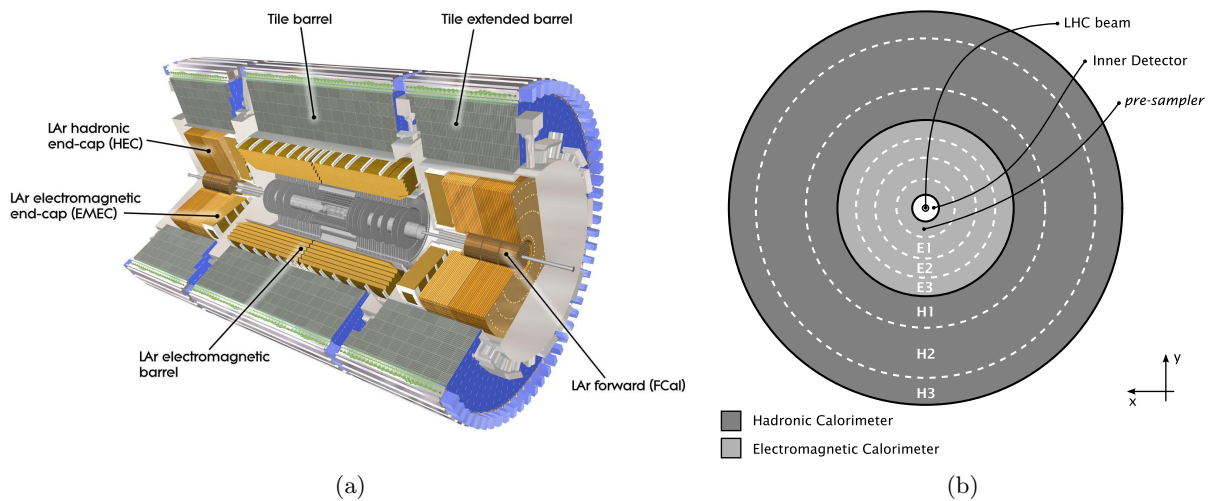


Figure 2. The ATLAS calorimeters (a) and the particle interactions throughout the detector (b)

at the barrel region, and also liquid Argon at the hadronic end-cap regions. The TileCal is segmented into three layers. The e.m. and hadronic calorimeters cover a region of $-\pi < \varphi < \pi$ and $-2.5 < \eta < 2.5$. The region $|\eta| > 2.5$ is covered by the forward calorimeter, which, however, is not used in this work.

3. Trigger System

The trigger system is split into three cascaded levels, each possessing its own maximum event rate and latency requirements [9]. The decision of each level is confirmed, or rejected, by the subsequent level. Eventually, if the event is accepted by all three levels, data from the entire ATLAS detector are stored for offline analysis [9]. Figure 3 shows the ATLAS trigger system diagram.

3.1. The L1-Trigger

The first-level of the ATLAS trigger system (L1) is based on highly programmable devices and its main objective is to reduce the 40 MHz event rate to nearly 100 kHz. The L1 marks the regions at the detector with possible interesting physics (RoI's, Regions of Interest) [14]. The latency in this level is $2.5\mu\text{s}$.

In order to achieve the highest event rate reduction, in the shortest time possible, L1 has access to compacted information from the detector. Only calorimetry and muon information is used. In the calorimeters, trigger tower regions are defined on the $\eta \times \varphi$ plane (0.1×0.1). The energy of the cells inside a trigger tower is summed up, for both e.m. and had. sections. For the muon chambers, low precision information from the muon transverse momentum is used. Electrons and other e.m. objects are the most frequent signature accepted by L1 [14], and this is the reason for their study in this work.

3.2. The High-Level Trigger

The ATLAS High Level Trigger (HLT) is implemented in software [15]. It is responsible for accepting or rejecting the decision of the L1 trigger [9] and it is formed by the second level trigger (L2) and the event filter (EF). HLT operates over the RoI selected by L1, accessing the detector information in its full resolution.

Large computer farms are used to support the HLT framework. The event processing is split into several processing units according to their usage. For the L2, the mean processing time available is 40 ms, and for the EF, it is 4 s.

Differently from L1, where only a rough classification is possible, the L2 runs specialized algorithms, which use information from the different subdetectors (calorimetry, tracking and muon chambers). The algorithm's complexity is further increased at the EF, where adapted versions of the offline reconstruction algorithms are executed. Accessing the full event, the EF better exploits the physics properties of each event.

3.2.1. Calorimeter HLT: The calorimeter High-Level Trigger (HLTCalo) is a set of algorithms that selects particles by performing different clustering methods at the cell level. An optimized infrastructure, common for L2 and EF, for data retrieving is available for all trigger algorithms.

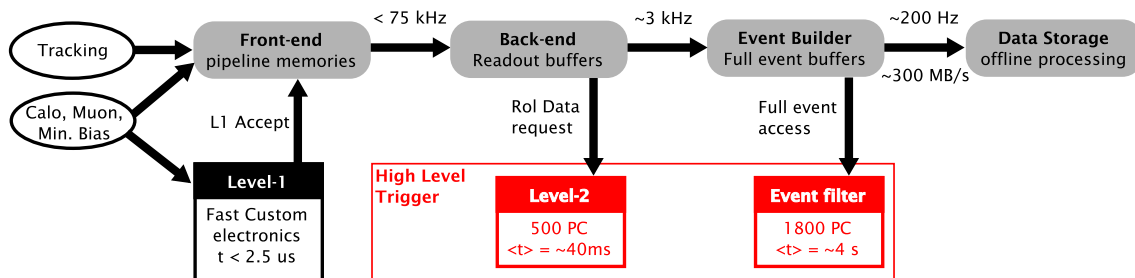


Figure 3. The ATLAS trigger system.

The HLTCalo algorithms are split into two parts: feature extraction (Fex) and hypothesis testing (Hypo). The FeX part access the detector data, through the common HLTCalo infrastructure, and calculates relevant variables needed for particle identification. These variables are used by the Hypo part to classify the RoI and to confirm or reject the previous level decision.

3.3. Athena

All levels of processing of ATLAS data, from high-level trigger to event simulation, reconstruction and analysis, take place within the Athena framework [16]. Athena is fully implemented in oriented object language (C++ and python) and it is structured into projects and packages.

The HLT community maintains a set of algorithms to emulate the trigger system, which runs after the L1 event acceptance. Also, common tools are provided for analyzing the trigger performance for the associated physics. New algorithms can be developed for running at the HLT, as long as they are implemented within the Athena framework, following some standard procedures and common methods.

3.4. Electron Identification

According to the expected event rates in L1, the detection of e.m. particles (electrons and photons) is extremely important for the identification of interesting physics channels. A slice of the trigger system is fully dedicated to the detection of such particles (egamma slice) [17].

In order to control the egamma trigger flow, linear thresholds are applied to the quantities extracted from egamma showers in HLT. These thresholds can also change during the experiment, in order to allow both detection efficiency and false alarm to be controlled. If less background is desired, tight criteria for electron identification are applied, but also more electrons are discarded. Otherwise, loose criteria are selected and more electrons (and also fake electrons from the background), are triggered. Through such selection, the experiment can tune the trigger rate and other physics demands optimizing the trade-off between particle identification efficiency and speed [15].

Offline algorithms, seeded by fine tracking information, also extract standard quantities for characterization of egamma showers, which can be used as guidelines for trigger performance evaluation. In particular, the R_{core} quantity extracted offline is used in this study, which is defined as follows:

- $R_{core} \rightarrow$ characterizes the shower spread at the second e.m. layer: $R_{core} = E_{3 \times 7} / E_{7 \times 7}$, where $E_{m \times n}$ is the energy deposited in a region of $m \times n$ cells in the $\eta \times \varphi$ plane around the seeded egamma tracking extrapolation.

4. The Ringer Algorithm

The Ringer is an alternative L2 algorithm for electron identification [10]. Feature extraction is performed through data description by concentric rings extracted from seven calorimeter layers of a selected L1 RoI. Hypothesis testing is applied through an artificial neural network [18] operating over the extracted calorimetry features (ring sums).

4.1. Feature Extraction

Concentric energy rings from a given RoI are extracted. At each calorimeter layer, the algorithm searches for the most energetic cell, whose energy is considered the first ring. The energies of the outer cells with respect to the previous ring are summed up to form the next ring. This procedure is implemented until the RoI is covered, or a maximum number of rings is reached.

Figure 4(a) shows some ring examples. It can be seen that, in some cases, the ring is completely conceptual, as some are not closed rings. For RoI description, a total of 100 rings are used and each layer contributes with a fixed number of rings (see Table 1). When a given RoI layer description does not provide enough ring sums, the ring vector is filled up with zeros.

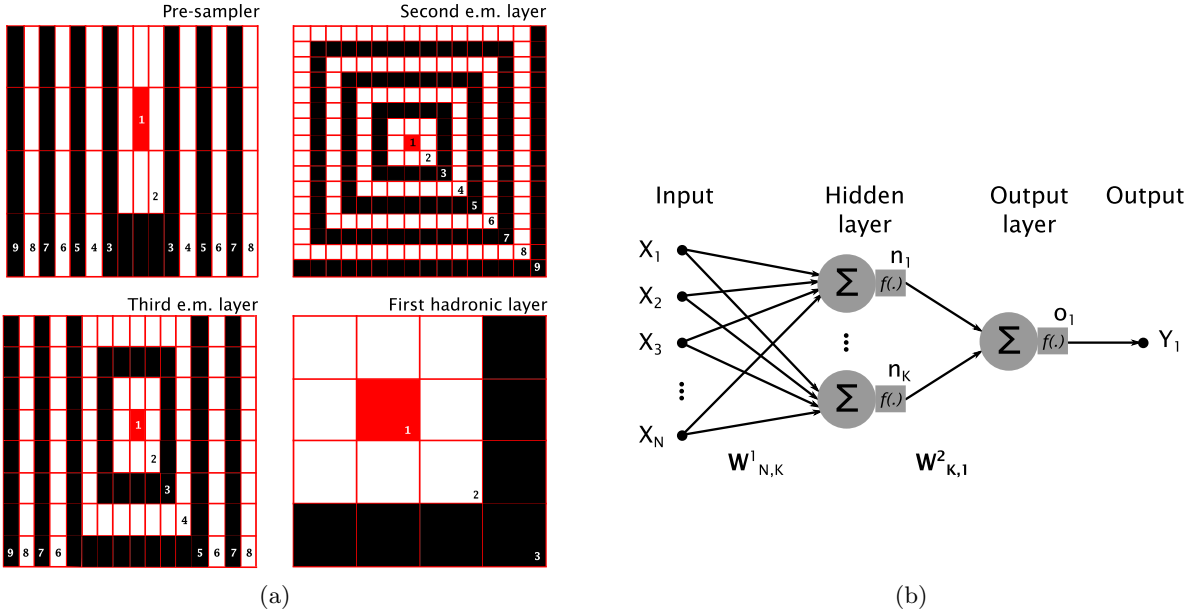


Figure 4. Illustration of possible rings from the Ringer feature extraction (a) and an example of a feedforward neural network for the Hypothesis testing (b).

4.2. Hypothesis testing

The extracted rings are presented to a three-layer feedforward neural network (Multi-Layer Perceptron, MLP), whose architecture is illustrated on Figure 4(b). The network is trained offline and its parameters are set before the start of the experiment. Several normalizations have been tested to improve the discrimination performance [10]. The present work implemented an event normalization, which consists of dividing each ring sum component by the total event energy.

4.3. Implementation

The Ringer algorithm is implemented on the Athena framework, which allows the algorithm to run on both real and simulated data. Several optimizations have been performed to assure that the Ringer implementation respects the time constraints of the L2 trigger. Also, the total amount of data required to store the Ringer information must be reasonable. Results from tests on such requirements are discussed in the next sessions.

5. Performance

The Ringer algorithm performance is first evaluated using 2010 14 TeV Monte Carlo simulations of electrons and particle jets (background). Electrons are from filtered simulated $Z \rightarrow ee$ events. In both simulations, pileup is considered (the current event might have information from previous events), which tends to deteriorate the trigger performance.

Time performance is evaluated through detailed analysis using the Valgrind tool [19], which is a framework specially designed to analyze code implementation. Through this framework, it is possible to reduce code bottlenecks, optimize function calls and spot unnecessary memory consumption and memory leaks. Athena possess a patch for Valgrind analysis, such that the

Table 1. Number of extracted rings per calorimeter layer

Layer	E.M. Calorimeter				Hadronic Calorimeter		
	pre-sampler	EM 1	EM 2	EM 3	HAD 1	HAD 2	HAD 3
# Rings	8	64	8	8	4	4	4

Ringer implementation can be analyzed within the ATLAS software framework.

For the hypothesis testing, (trained) neural network parameters (weights and biases) are exported and uploaded to the Athena framework.

5.1. Methodology

As in Z boson simulations other physics activities are also considered, the RoI must be matched to the electrons from the $Z \rightarrow ee$ decay, by terms of a $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.05$ between the RoI and the electron positions on the $\eta \times \varphi$ plane. For the jet simulation, all fake electron RoI's (accepted by simulated L1 trigger) are considered background. In total, nearly 7k events are selected for electrons and 5k for jets, which are split into half, one forming the training set, and the other forming the test set. While the train set is used to adapt the MLP parameters, the test set is used to evaluate the MLP generalization. Also, the test set is used as validation set to avoid overtraining [18].

The resilient back-propagation algorithm [20] is used for training the MLP in batch mode, because of its fast convergence. Also, the MLP is built with one hidden layer only, in order to keep its architecture simple (interesting for online operation, because of its speed). The hyperbolic tangent is used as activation function in both hidden and output layers, where electrons are targeted to the output value '1' and jet '-1'. The training cost function is the mean squared error between the network output and the desired target. Previous studies have shown that 10 neurons at the hidden layer seems to be the best trade-off between performance and architecture complexity. In order to avoid local minima, the training algorithm is initialized randomly 100 times, where only the network with the best performance is considered.

The SP index [10] is used to evaluate the network performance:

$$SP = \sqrt{\sqrt{P_e \times P_j} \times \left(\frac{P_e + P_j}{2}\right)}$$

where P_e and P_j are the electron and jet detection efficiency, respectively. The MLP output is compared to a threshold in order to finally classify the input as electron or jet. The discrimination threshold is selected in order to maximize the SP index.

5.2. Relevance Mapping

A relevance mapping [21] can reduce the MLP complexity by selecting only the input variables (rings) which are most relevant for the discrimination process. Here, the relevance of a given input variable is measured from the relative difference between the MLP SP index and the SP index achieved when the input variable is fixed to its mean value:

$$R_i = \frac{SP_{MLP} - SP_i}{SP_{MLP}}$$

Rings having relevance values below a given threshold are removed from the training and test sets and a new MLP is designed. Notice that, if negative relevant values are found, it means that the given input variable confuses the MLP, and its removal may increase performance.

5.3. Tag-and-Probe

During normal operation, with proton collisions, it is not known if the given RoI is truly from an electron or a jet. Thus, a data-driven method for efficiency estimation must be employed. Tag-and-probe is often used, as it is extremely suitable for physics processes characterized by a double-object final state signature (like the $Z \rightarrow ee$ process) [22].

The method consists in evaluating the efficiency of detecting 'Probe-like' objects within a properly 'tagged' sample of events. For the L2 online trigger, it consists of selecting possible electron pair candidates which reconstruct the Z boson mass and which have been accepted by the egamma offline algorithm. The event is tagged if one of the electrons (at least) is accepted by a L2 tight electron criteria. The probe object is selected according to the Ringer algorithm

hypothesis testing, considering the discriminator threshold that maximizes the SP index. Only the test set is used to evaluate the Ringer performance with the Tag-and-Probe method.

5.4. Results

The Ringer implementation was evaluated with respect to the total execution time and amount of data needed to save its information. The Valgrind tool did not report any memory leak after several events. The total time needed to run the entire L2 egamma algorithm was measured. After an accumulation of ~ 17 k RoI's, the new Ringer implementation (after the Valgrind analysis) is estimated to take, on average, 1.075 ms to fully execute (feature extraction and hypothesis testing), while before the Valgrind analysis it took, on average, 1.445 ms. The detailed analysis led to a 59% faster execution time. Also, the new Ringer implementation adds an overhead of ~ 0.3 ms per RoI when compared to a standard L2 algorithm. With respect to the amount of data needed, the Ringer information represents 6.2% of the total payload for electron/photon algorithms.

The neural network performance for the training set indicates a 5.18% false alarm probability (jets misclassified as electrons), for a fixed electron detection probability of 98.00%. Figure 5 shows the neural network output for the training set. Notice that most of the electrons were correctly mapped to the output value '1', while jets were mapped to '-1'. The threshold needed to maximize the SP index is also shown, together with the threshold needed to achieve 98.00% electron detection. The maximum SP index translates into a 97.28% electron detection probability, against a 4.00% false alarm probability.

In order to reduce the network complexity, the input variables (rings) were ranked according to their relevance. Figure 6(a) shows the relevance for each ring. The full vertical lines indicate the end of a calorimeter layer and the beginning of the next one. By changing the relevance threshold, the number of rings considered relevant for the discrimination can be adjusted. In order to achieve a reduction of $\sim 70\%$ of the total amount of rings, for instance, a relevance threshold of 0.05% is required (dashed horizontal line). In terms of performance, Figure 6(b) shows the ROC (Receiver Operating Characteristics) [23] curve achieved after training a MLP using only the relevant rings as inputs (also initializing 100 times to avoid local minima). The ROC curve for the MLP using the entire ring set is also shown for sake of comparison. It can be seen that the $\sim 70\%$ information reduction increased the false alarm in less than 3 percent points (fixing the electron detection efficiency to 98%), while the SP index is slightly

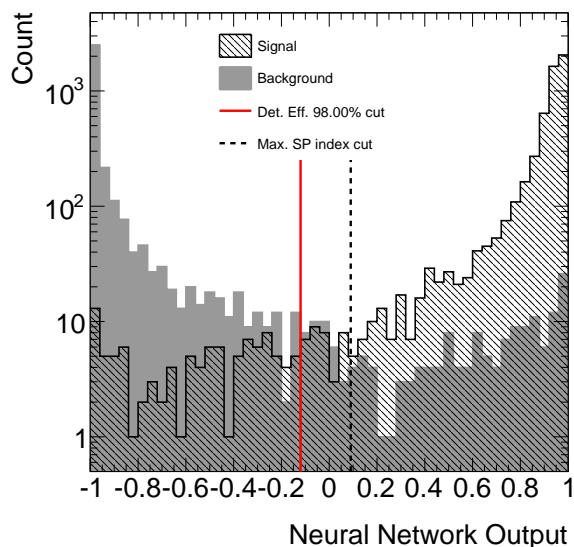


Figure 5. Neural network output for the training set over the MLP trained with full ring description.

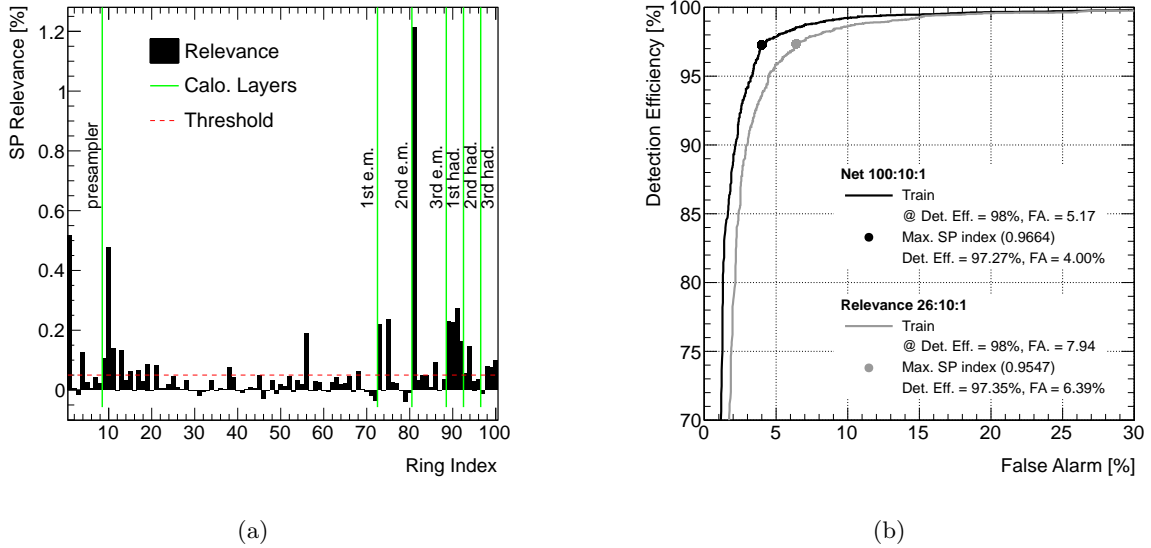


Figure 6. Relevance mapping considering the training set (a) and the resulting ROC curve for a MLP fed from relevant rings and full ring information (b).

reduced (~ 0.1). The relevance threshold value can be used as a knob to adjust the MLP performance in a compromise with the Ringer hypothesis test speed and the total false alarm rate. Also, Figure 6(a) shows some rings with negative relevance, meaning that these rings confuse the particle classification, instead of assisting it.

In order to extract physics conclusions, the neural network output can be compared to some key offline egamma variables for a given RoI. Figure 7 shows the MLP trained without the relevance cut (full ring information) with respect to the R_{core} offline variable (see Section 3.4), for both electrons and jets. The results are considering the test set, in order to evaluate the MLP generalization. The R_{core} measures the e.m. shower spread. Electrons produce narrower showers, leading to a R_{core} value close to unity, while jets tend to form wider showers, thus with a R_{core} value far from unity. It can be seen from Figure 7(a) that the R_{core} value is close to unity for most of the electrons that pass MLP cut. This is particular interesting and useful when data from proton collisions is considered, as the actual RoI classification is unknown. The same result can be seen for the rejection of jets, on Figure 7(b), where most of the jet RoI's are correctly classified by the Ringer hypothesis test, as their R_{core} value is mostly around 0.8.

Finally, the performance with respect to the Tag-and-Probe method was evaluated considering the test set. As described before (see Section 5.3), the Tag-and-Probe is a data-driven method for efficiency estimation, and it is an important performance estimator used in proton collisions. Figure 8(a) shows the Ringer performance and false alarm as a function of the RoI transverse energy. RoI's near the crack region ($|\eta| \approx 1.5$) and the end-cap ($|\eta| \approx 2.5$) are excluded. It can be seen that the Ringer performance over the $Z \rightarrow ee$ channel is very good for most of the energy range. For the false alarm rejection estimation, the probability that a jet RoI, with the respective transverse energy, is accepted by the Ringer algorithm is also shown on Figure 8(a). On average, for the transverse energy range $20 \text{ GeV}/c < E_t < 50 \text{ GeV}/c$, the Ringer algorithm achieves an electron detection efficiency of $98.24 \pm 2.54\%$, for a $15.78 \pm 10.16\%$ false alarm probability. More statistics is needed to increase the measurement accuracy.

Similarly, the performance for the electron detection based on the Tag-and-Probe method and the false alarm rejection as a function of η is shown on Figure 8(b). Only RoI's whose respective transverse energy is greater than $20 \text{ GeV}/c$ are considered. It can be seen that the performance over the $Z \rightarrow ee$ channel is almost uniform in η ($97.34 \pm 2.82\%$ electron detection efficiency against $8.08 \pm 2.46\%$ false alarm probability, considering $|\eta| < 1.5$). For the false alarm rejection, the regions around the calorimeter crack ($|\eta| \approx 1.5$) and at the far end of the

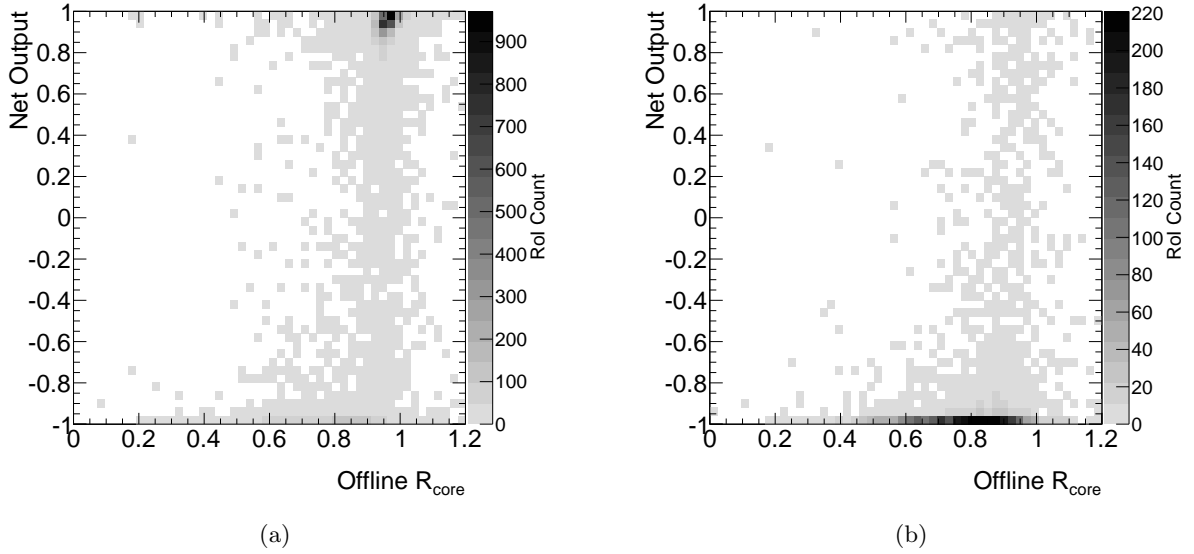


Figure 7. MLP output as a function of the egamma offline R_{core} discriminant variable, considering the test set, for electrons (a) and jets (b).

calorimeter suffer from a higher false alarm probability. This is expected due to the missing information and poor energy resolution at these regions.

6. Conclusions

Results from the ATLAS HLT Ringer algorithm performance were shown, with respect to the implementation in the ATLAS Athena framework, and considering the detection performance over the $Z \rightarrow ee$ channel, including pileup. In addition, a methodology for reducing the neural network complexity, by ranking the input variables according to their relevance, was studied. Finally, the relevance study reduced the MLP input dimension in $\sim 70\%$, with a small impact on detection performance.

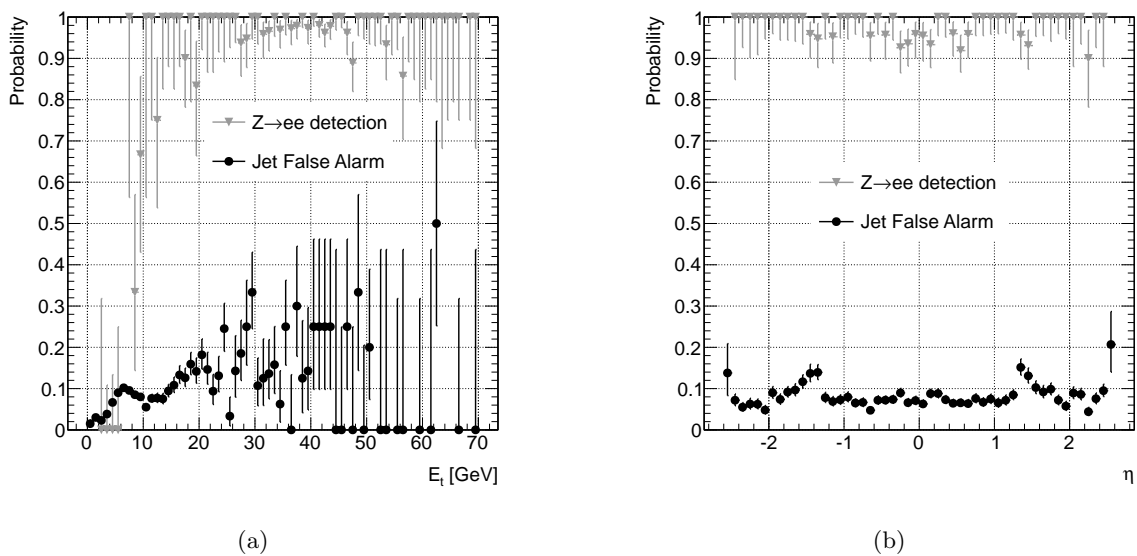


Figure 8. MLP performance considering the test set and the Tag-and-Probe method as a function of the Rol offline transverse energy (a) and η (b).

The detailed Valgrind analysis led to an optimized implementation, which became 59% faster, while the total amount of data needed to store the Ringer information was kept low (6.2% of the total egamma payload).

The detection performance was evaluated from 2010 14 TeV Monte Carlo simulations. The Ringer algorithm showed high performance also when a data-driven method for performance estimation is considered (Tag-and-Probe). Crosscheck with the offline R_{core} discriminant variable shows good agreement with the MLP response for electrons and jets.

The results presented here shows that it is promising to use the Ringer algorithm for electron identification in the trigger system for collider experiments.

Acknowledge

The authors would like to thank the Brazilian funding agencies CNPq, CAPES, FINEP, FAPERJ and RENAFEA, as well as CERN and the European Union, for the support to this work. We also thank our colleagues from the ATLAS/TDAQ collaboration for the interesting discussions concerning this work.

References

- [1] Bishop C M 1994 *Review of Scientific Instruments* **65** 1803
- [2] Abramowicz H *et al.* 1995 *Neural network based electron identification in the ZEUS calorimeter* vol 1 p 763
- [3] Badalà A, Barbera R *et al.* 2003 *Nucl. Instrum. Methods Phys. Res. A* **502** 503
- [4] Bortolotto C, Angelis A D, Groot N D and Seixas J 1992 *Int. J. Mod. Phys. C* **3** 733
- [5] Evans L *et al.* 2008 *Journal of Instrumentation* **1**
- [6] Prevotet J C, Denby B *et al.* 2003 *Nucl. Instrum. Methods Phys. Res. A* **502** 511
- [7] McGlone H M 2009 *Neural Network Analysis in Higgs Search using $t\bar{t}H, H \rightarrow b\bar{b}$ and TAG Database Development for ATLAS* Ph.D. thesis University of Glasgow Glasgow
- [8] The ATLAS Collaboration 2008 *Journal of Instrumentation* **3** 1
- [9] Kordas K *et al.* 2007 *Nuclear Physics B - Proceedings Supplements* **172** 178
- [10] dos Anjos A R, Torres R C, J M de Seixas B C F and Xavier T C 2006 *Nucl. Instrum. Methods Phys. Res. A* **559** 134
- [11] Wigmans R 2000 *Calorimetry: energy measurement in particle physics* (Oxford: Clarendon Press)
- [12] Zhang H *et al.* 2010 *The ATLAS Liquid Argon Calorimeter: Overview and Performance* vol 1 pp 1–6
- [13] Ariztizabal F *et al.* 1994 *Nucl. Instrum. Methods Phys. Res. A* **349** 384
- [14] Berge D, Aielli G, Andrei V *et al.* 2007 *Nucl. Instrum. Methods Phys. Res. A* **581** 476
- [15] Stelzer J *et al.* 2010 *The ATLAS High Level Trigger Configuration and Steering* vol 1
- [16] Cranmer K 2008 *Nuclear Physics B, Proceedings Supplements* **177** 126
- [17] The ATLAS Collaboration 2010 Performance of the electron and photon trigger in p-p collisions at a centre of mass energy of 900 GeV Tech. Rep. ATLAS-CONF-2010-022 CERN Geneva
- [18] Haykin S 2008 *Neural Networks and Learning Machines* (Prentice Hall)
- [19] Nethercote N and Seward J 2007 *Valgrind: A Framework for Heavyweight Dynamic Binary Instrumentation* vol 1 pp 1–12
- [20] Riedmiller M and Braun H 1993 *A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm* pp 586–591
- [21] Gruber A *et al.* 1994 *New Computing Techniques in Physics Research* **1** 429
- [22] Straessner A and Schott M 2010 *J. Phys. Conf. Ser.* **219** 032023. 9 p
- [23] van Trees H 2001 *Detection, estimation, and modulation theory part I* (John Wiley & Sons)