Universidad Técnica Federico Santa María
Departamento de Electrónica
Valparaíso - Chile

“Performance of a Multivariable Calibration for Electrons and Photons with the ATLAS High-Level Trigger”

Javier Salazar Loyola

Profesor Guía William Brooks
Profesor Correferente Ryan White

Acknowledgements

This program and research have been supported by CONICYT-PCHA/Magíster Nacional/2013-22130526.

Abstract

A multivariate calibration technique is implemented for online electron and photon candidates to account for energy loss upstream of the electromagnetic calorimeter and for lateral and longitudinal shower leakage in the ATLAS experiment. The online calibration is necessary to improve the energy resolution of the electrons and photons with respect to the offline objects to ensure a sharp turn-on response of the trigger system. In this case, boosted decision trees have been applied to calibrate the response of the calorimeters in the detector for electrons and photons, and the performance of the calibration has been measured through determination of the resolution of the electron and photon energy for both online and offline.

Contents

List of Figures

Chapter 1

Introduction

The ATLAS experiment[1] is a multipurpose particle physics experiment which is one of the two high luminosity experiments of the LHC [2] —CMS[3] being the other one— and one of its four major experiments. It is used for a great variety of measurements and physics analyses where its subsystems measure a several variables, as will be seen. Due to the low cross section of interesting physics events compared to the overall inelastic cross section, it is necessary to filter those events before recording them using a trigger system.

The precise determination of the energy of final state particles measured in the calorimeter is critical to many physics analyses. In particular, the Higgs boson discovery reported by ATLAS and CMS experiments on 2012[4], [5] was a major milestone on Physics research and the LHC performance, so it is of primary importance to further study its properties in order to finally define if it corresponds to the Standard Model Higgs boson or to New physics. The particle itself was expected to exist thanks to the gauge symmetry breaking mechanism that gives mass to $W^{\pm}$ and Z bosons and, through Yukawa interactions, to the fermions[6], [7]. The reported Higgs boson discovery was accomplished mainly via two decay channels, with final states including two photons ($H \to \gamma\gamma$) and four leptons ($H \to ZZ^{*} \to 4l$), where the lepton (l) can be an electron (e) or muon (μ). In this discovery, the measured mass was reported to be 125 GeV, a measurement of utmost importance, since the Higgs mass was not predicted by the theory, but rather a free parameter, so it must be determined experimentally.

The study of the Higgs boson decaying to electrons or photons relies on the ATLAS Calorimeter system and its performance, since while muons are detected in the Muon Spectrometer, electron and photon reconstruction, as will be seen, depends on the calorimeter, especially for energy determination, and the calorimeter has an excellent resolution at high transverse energy for both particle types. This energy, in turn, is what allows for the determination of the invariant mass of the particle which originated the measured final states, so it is a priority for providing a reliable measurement.

While the Higgs boson measurement is one of the most important in ATLAS and one of the purposes of the LHC, it is far from being the only notable analysis being carried with ATLAS. A long term goal of LHC is the precise measurement of the W boson mass[8], which requires an excellent accuracy of the electron energy calibration. Moreover, a recent report stated an excess of events around 750 GeV in the $\gamma\gamma$ channel on both ATLAS [9] and CMS [10], with lots of papers trying to describe that new phenomenon (see, for instance, [11]–[14]), and thanks to the new measurements, it was discarded as statistical fluctuations [15].

The measurements of the $Z \rightarrow ee$ decay channel are important for a number of reasons. It is an important probe for QCD predictions[16] and the mentioned measurement W boson mass, but also the detector relies on the well understood shape of the Z boson to understand the energy response of the calorimeter. If the detector can accurately and precisely measure the Z line shape, in turn the mass of the Higgs boson —which as has been stated above, depends on this decay on one of its channels— can be precisely measured.

The Z shape is a standard candle for the determination of the quality of the calorimeter calibration. Since it is desirable that the energy response is the same between the data and the Monte Carlo simulation, *in situ* corrections are used to fix any remaining differences at the reconstruction stage.

Since the event selection is made in the trigger system, as mentioned above, it is important that a correct measurement is already available at that stage, so the calorimeter clusters, which work at the electromagnetic (EM) scale, must already be calibrated at the trigger level in order to take into account the inefficiencies, as the energy loss before the Calorimeter or the lateral and longitudinal leakage. The EM scale accounts correctly for the energy deposited by the electromagnetic showers in the calorimeter, since it is designed in such a way that the entire shower is deposited in it. As a result of the fact that the electromagnetic interaction is a well understood process and the simulations describe very well the ATLAS calorimeter, cluster energies are calibrated on this scale at the reconstruction stage. This reconstruction is done offline and include several calibration stages[1], but only the simulation based calibration is applied to the trigger, so it must be simple enough to be executed online and precise enough to neither under-calibrate —thus not having too much background noise— nor over-calibrate —thus not losing events to unaccounted inefficiencies.

The mentioned simulation-based calibration of the measured energy in the Electromagnetic Calorimeter is the purpose of this work and has been implemented at the trigger level for the Run 2[2] of the LHC. It is based on Monte Carlo simulations and uses multivariate analysis (MVA) techniques. This is one of the most important changes for Run 2 in the e/gamma trigger. This calibration has already been implemented and approved for the offline analysis[8] (at the reconstruction level) and an update of the trigger level calibration was required in order to be in accordance with the offline level, where the physics analysis is made. The method uses multivariate techniques to derive calibration factors based on simulations of the detector to electromagnetic showers.

One of the advantages of the MVA calibration is that is easy and relatively quick to get a new set of corrections (from a few hours to a few days, depending on how much data are being used for the training and testing of the new set) in the case of changes in the detector or geometry description, which is the case for the Run 2, since after the end of the Run 1, the LHC and particularly the ATLAS detector went through a major upgrade. Another advantage is that the technique allows an arbitrary number of input variables for the calibration which can be changed (added, removed) at any time and moreover, it takes into account the correlation between the variables.

This document is organized as follows. The ATLAS detector is described in Chapter 2, as well as its most important subsystems for this work, namely, the trigger and calorimeter systems. In Chapter 3, the implemented calibration is described. The testing results and the resolution studies are presented in Chapter 4. Finally, Chapter 5 presents conclusions for this development.

---

[1] See, for instance, Chapter 3 of this work, or reference [8].

[2] The second operational run is called the Run 2, while the first one is called the Run 1

Chapter 2

The ATLAS Detector

The Large Hadron Collider (LHC) at CERN [2] is the world's most powerful collider so far, and has been designed to collide proton beams with center of mass energy of up to 14 TeV at a design luminosity of 10^{34} cm^{-2}s^{-1}, and heavy ions, in particular lead nuclei, at 5.5 TeV per nucleon pair, at a design luminosity of 10^{27} cm^{-2}s^{-1}. It is mounted in the nearly 27 km circular LEP tunnel and it consists of eight arcs and eight straight sections, between 45 and 170 m below the surface. The accelerator has 1232 superconducting dipole magnets, each of which produce a magnetic field of approximately 8.4 T. The bunches of protons, separated by 25 ns intervals (40 MHz), intersect at the four interaction points where the detectors are placed. These interaction points correspond to ATLAS and CMS, which are general purpose experiments designed for precision measurements and search for new physics; LHCb, a B physics oriented detector; and ALICE, a heavy ion experiment.

The ATLAS detector[1] (A Toroidal LHC ApparatuS) at the LHC is a multipurpose particle physics experiment, forward-backward symmetric with respect to the interaction point, with cylindrical geometry and a coverage in solid angle of nearly 4π. The detector is 25 m in height and 44 m in length with an overall weight of approximately 7000 tonnes. A schematic drawing of the detector is shown in Figure 2.1.

The coordinate system used in ATLAS is a spherical one defined in terms of the beam. The origin is set at the nominal interaction point, while the beam direction defines the z-axis and the $x-y$ plane is transverse to it. In this plane, the transverse momentum p_T and energy E_T are defined. The positive x-axis is defined pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The azimuthal angle ϕ

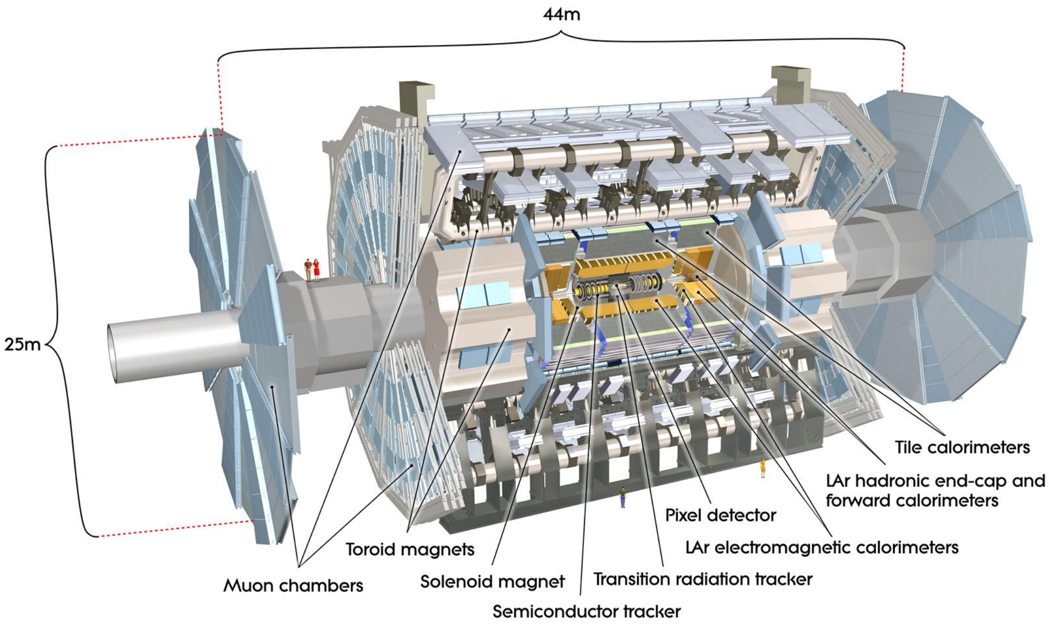


Figure 2.1: The ATLAS Detector

is measured around the beam axis, the polar angle θ is measured from the beam axis, and the pseudorapidity, which is the usual way to express the polar coordinate in ATLAS, is defined as

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right)$$

Its main subdetectors are the Inner Detector, the Calorimeters and the Muon Spectrometer.

The Inner Detector (ID) is the tracking device located in the inner-most part of the detector, which can be seen in Figure 2.2. It has been designed to measure the trajectory of charged particles (track). It is surrounded by a thin superconducting solenoid which provides a 2 T magnetic field and allows electron identification, pattern recognition, momentum and vertex measurement (particularly the primary vertex or interaction point) and reconstruction of unstable particle decay length in the range of $|\eta| < 2.5$ via three subsystems, which are the Pixel detector, the silicon microstrip tracker (SCT) —both of which are types of silicon semiconductor-based detectors— and the transition radiation tracker (TRT). This configuration allows high precision measurements.

The calorimeter (Figure 2.3) is divided into two systems, which are the liquid-argon (LAr) electromagnetic sampling calorimeter and the hadronic calorimeter, which together

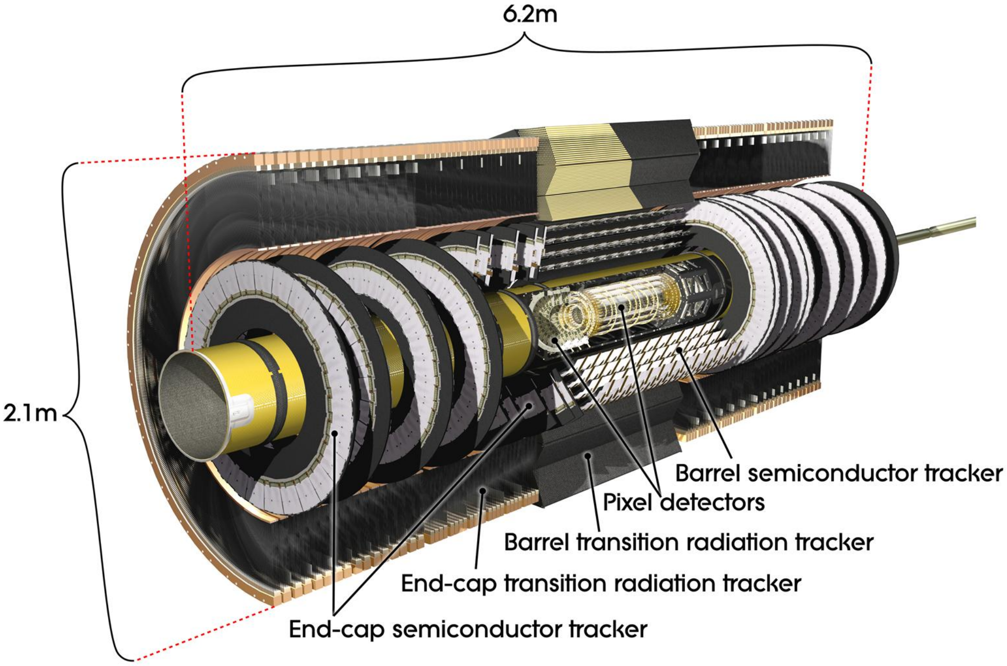


Figure 2.2: The Inner Detector

cover the range $|\eta| < 4.9$. Over the η region of the inner detector, the fine granularity of the electromagnetic calorimeter (EM) is suited for precision measurements of electrons and photons, while the coarser granularity of the rest of the calorimeters is enough for the jet reconstruction and E_T^{miss} (missing transverse energy) measurements. The electromagnetic calorimeter will be reviewed in section 2.1.

The hadronic calorimeter is compresied of two technologies, scintillator-tile calorimeter for $|\eta| < 1.7$ (divided in a barrel for the region with $|\eta| < 1$ and two extended barrels for the range $0.8 < |\eta| < 1.7$) and LAr for the end-caps in the range $1.5 < |\eta| < 3.2$, and for the forward calorimeters (FCal), covering the region in $3.1 < |\eta| < 4.9$, divided into three compartments.

The Muon Spectrometer (MS, Figure 2.4) is a tracking detector for precise measurements of muon trajectory and p_T which surrounds the calorimeter and defines the overall dimensions of ATLAS. It consists of three air-core superconducting toroid systems with three layers of high precision tracking chambers for muons for $|\eta| < 2.7$ and fast detectors for triggering for $|\eta| < 2.4$ with a timing resolution of the order of 1.5-4 ns.

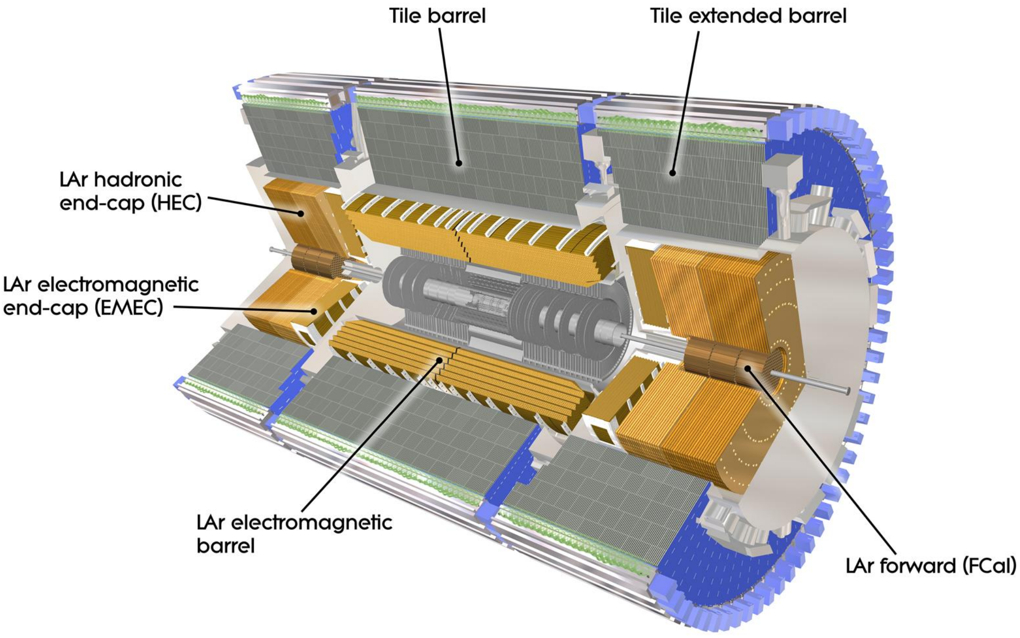


Figure 2.3: The ATLAS Calorimeter system

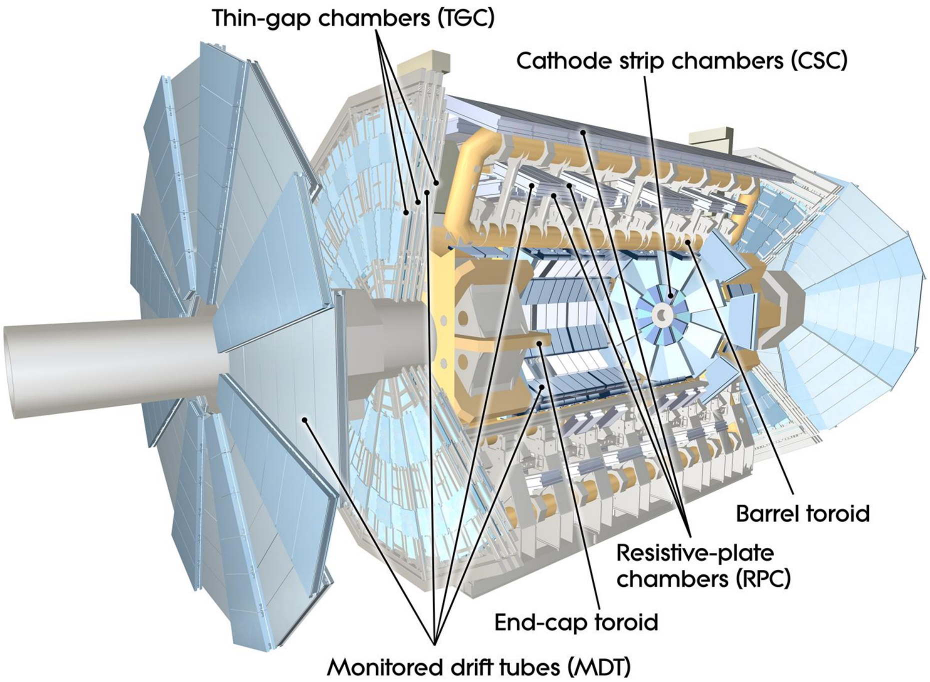


Figure 2.4: The ATLAS Muon spectrometer

2.1 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (EM) is a liquid argon sampling calorimeter with an accordion geometry, divided into a barrel part for $|\eta| < 1.475$ and two end-cap components on each side of the barrel ($1.375 < |\eta| < 3.2$). The barrel calorimeter consists of two identical half-barrels, separated by a 4 mm gap at $z = 0$ and the end-caps are divided into two coaxial wheels, the outer one covering the region $1.375 < |\eta| < 2.5$ and the inner, the region $2.5 < |\eta| < 3.2$. The accordion geometry provides complete ϕ symmetry without azimuthal cracks. For the end-cap inner wheel, the calorimeter is segmented into two sections and has a coarser granularity than the precision physics region ($|\eta| < 2.5$), where the calorimeter is divided into three layers and is preceded by a presampler detector, in $|\eta| < 1.8$.

Each half of the barrel (EMB) is divided into 16 modules in ϕ, each covering a $\Delta\phi = 22.5°$, while each of the end-cap wheels (EMEC) is divided into 8 modules in ϕ, for ease of construction, but once assembled, neither presents discontinuity along the azimuthal angle. The whole barrel is 6.8 m long and its radii are 1.15 (inner) and 2.25 (outer) meters; and each half-barrel weighs 57 tons. The EMEC has the same outer radius as the barrel and the length of one wheel is 3.17 m.

For the precision physics region, the first calorimeter sampling layer has a fine granularity in order to reject π^0 up to $E_T = 50$ GeV or more[17], while the main sampling for measuring energy is the second layer, configured as “square towers.” In this region, the first layer has a thickness of about 4.4 radiation lengths (X_0) and granularity (cell size) of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ for the most central region ($|\eta| < 1.8$), 0.004×0.1 for $1.8 < |\eta| < 2.0$ and 0.006×0.1 for $2.0 < |\eta| < 2.5$; the second layer has a thickness of about $17X_0$ and granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$; and the third layer, $\Delta\eta \times \Delta\phi = 0.050 \times 0.025$ with a thickness of about $2X_0$. The third layer is used to correct leakage beyond the EM for high-energy showers. The presampler is an active LAr layer with a thickness of 1.1 cm in the barrel (0.5 cm in the endcap) and a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ and is used to correct for the energy lost in front of the calorimeter. The first, second and third calorimeter layers are named strip, middle and back layer, respectively.

2.1.1 Energy Reconstruction in the EM Calorimeter

The incoming particles deposit their energy in many calorimeter cells and the clustering algorithms are designed to group the cells and sum the deposited energy within each cluster, so the measured energy can be calibrated later — depending on the particle type —, to account for energy outside the cluster or in dead material[18]. There are two clustering algorithms used in ATLAS, the "sliding-window" and the topological algorithm. The second one consists in grouping neighboring cells with high energies into clusters with variable sizes and it is not used online.

Sliding-Window Clustering

This algorithm has three steps: tower building, seed cluster finding (precluster), and cluster filling.

The first step, the tower building, consist of dividing the calorimeter $\eta - \phi$ space into a grid of $N_\eta \times N_\phi$ elements of size $\Delta\eta \times \Delta\phi$ and the energies of all the cells in all layers are summed into the "tower" energy.

The second step creates a window of fixed size (optimized to obtain the best efficiency for finding seed clusters and limit the noise-induced seeds) in units of the tower size which is moved across the elements of the previously defined grid, in steps of $\Delta\eta$ and $\Delta\phi$. If the window transverse energy (sum of the transverse energy of the towers in the window) is a local maximum and above a threshold (optimized as the window size), a seed is formed. Then, the position of the seed cluster is computed and the duplicates removed.

Finally, EM clusters are formed by taking all cells within a rectangle of size $N_\eta^{\text{cluster}} \times N_\phi^{\text{cluster}}$ centered on a layer-dependent position, where the several layers are processed with a specific order and slightly different cluster sizes for every layer: first the middle layer, then the strips, presampler and finally, the back layer. Different cluster sizes are built for different particle types and the location of the cluster in the calorimeter, where the size is optimized with a compromise between two effects: the cluster should be large enough to contain most of the energy deposited in the calorimeter by the particle, and small enough to remove the noise which arises from too many cells. Table 2.1 lists the cluster size for different particle types.

Particyle Type	Barrel	Endcap
Electron	3×7	5×5
Converted photon	3×7	5×5
Unconverted photon	3×5	5×5
Trigger	3×7	3×7

Table 2.1: Cluster size for different particles in the barrel and endcap of the calorimeter

2.2 Trigger System

During the Run 1 of the LHC, between 2009 and 2013, the trigger system[19] was based on three levels of event selection for deciding the physics of interest. The first level, named simply Level 1 (L1) was hardware based, while the level 2 (L2) and event filter (EF) were software based and run in farms of commercial computers dedicated to their opeartion. During that run, the trigger selected events with high efficiency at centre of mass energies up to 8 TeV for a wide range of physics processes, reducing the input from an initial crossing rate of 40 MHz to around 200 Hz of events which were kept for further analysis. But for the Run 2, the conditions are more challenging for the trigger system, since interaction rates are expected to increase by a factor of around 5 or 6; both the peak instantaneous luminosity and the interactions per bunch crossing will be increased [20]. For this reason, the trigger system has been updated, so it can improve or, at least, preserve, the previous efficiencies.

As noted earlier, the trigger system included three levels, which in Run 2 have been merged into two, described as follows.

- The Level 1 (L1) is a hardware-based trigger with custom electronics having a latency of 2.5 μs and which uses reduced granularity from the calorimeter and dedicated muon chamber information to perform a preliminary selection.
- The High Level Trigger (HLT) consists of a single processing farm with an average processing time of 0.2 s per event over regions of interest (RoI, about 1-4% of the data of each detector) defined by the L1.

In the Run 2, the L1 trigger consists of several sub-systems, namely the L1 calorimeter trigger system, (L1Calo), the L1 muon trigger system (L1Muon), the Central Trigger Processors (CTP) and a new L1 topological trigger module (L1Topo). At this level, the event rate is reduced from 40 MHz to 100 kHz.

With the merging of the two Run 1 software based triggers into the HLT came a reduction in complexity and at the same time, an improvement in flexibility. Because of this, the new software of the trigger uses offline reconstruction and identification algorithms as much as possible, in order to provide high efficiency and low background rates. The reduced trigger rates permit the use of more precise algorithms, but with higher latency. This added precision improves the data analysis, since the acceptance of reconstructed physics objects is improved when the online objects are similar to the offline. By the end of 2016, a hardware based fast track finder was expected to be fully integrated and it now provides tracks for the algorithms of the HLT at the same rate as the L1[21].

The process of configuring the trigger makes use of trigger menus, which are tables of signatures specified by thresholds and selection criteria for different physics objects at each level of the trigger.

2.3 Electron and Photon Reconstruction

2.3.1 Electron reconstruction and identification

There are mainly two methods for the reconstruction of electrons in ATLAS [22], namely the standard one, seeded from the electromagnetic calorimeters and a second one, track based, seeded from the inner detector tracks.

The standard reconstruction[23] is based on reconstructed clusters in the EM calorimeter, which are then associated to charged particles tracks reconstructed in the ID. The reconstruction algorithm has been optimized to provide the smallest background levels for proton-proton scattering. First of all, a set of seed clusters is created with energies above 2.5 GeV (formed by the sliding window algorithm), where duplicate clusters are removed. In the ID region ($|\eta| < 2.5$), an electron is defined by one or more tracks matched to a seed cluster. The matching is done by extrapolating the track from the last measurement point to the second layer of the calorimeter (where, as already mentioned, the main measuring of energy is done). The impact point in η and ϕ are then compared to the cluster's ϕ and η and a match is declared if the difference is below a given threshold.

In the case of several track matches, the tracks are ordered according to the quality of the match via the difference $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The smallest ΔR value is considered to be the best match.

The track based reconstruction algorithm is used for low p_T electrons (of a few GeV) and uses the reconstructed track, extrapolating it to the EM calorimeter and building a cluster using the track impact point as seed. For electrons outside the ID region (forward electrons), the candidates only contain the EM calorimeter cluster information.

For the identification of electrons ATLAS relies on information coming from shower shape variables, track-to-calorimeter matching information and tracking information, in order to separate between isolated electrons and jets.

2.3.2 Photon reconstruction and identification

The main aspects of the reconstruction of photons are the same as the ones for the electrons[24], since both objects are treated in a similar way in the reconstruction algorithm. However, an electron is reconstructed mainly from the presence of a track matching an electromagnetic cluster, while the photons are classified into two main categories, namely converted and unconverted. In order to discriminate between electrons and photons, if the reconstruction finds a track matched to the electromagnetic cluster, it checks for an associated conversion[1]. If no conversion is found, an electron candidate will be created. Otherwise, a photon candidate is created, which can be later checked in order to obtain its conversion variables.

Because of the above strategy, almost all converted photon candidates will be electron candidates. This is remedied through a verification of whether a reconstructed conversion vertex can be matched to a cluster and after the reconstruction of electrons, a recovery of converted photons (and misidentified unconverted photons) is performed. This recovery is done verifying different conditions over the electron candidate, meant to discriminate whether the conversion vertex candidate match is valid or not.

---

[1] An e^+e^- pair production in the tracker material is known as photon conversion[24]. In this process, the momentum of the photon is not equally shared between the electron and positron, so it is possible to obtain highly asymmetric photon conversions which might lead to one of the particles not having enough energy to produce a track in the ID. In that case, only one track will emerge from the reconstructed conversion vertex.

Any cluster without any track matched to it (either primary or from conversion candidate vertices) is taken as unconverted photon candidate.

The photon identification algorithms in ATLAS rely on rectangular cuts using variables from the calorimeter, optimized to give a good separation between isolated photons and fake signatures from QCD jets. This identification uses the same information as that used for electrons when dealing with converted photons, but at the online (trigger) level, only the shower shape information is used.

2.3.3 Electron and Photon Trigger Selection

At the L1[25], photons and electrons are selected using calorimeter information with reduced granularity, given by the Trigger Towers with size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and with the full depth of either the electromagnetic or hadronic calorimeter, creating RoIs for each identified electromagnetic object, with associated position and energy. The trigger is considered to contain an e/γ candidate if a local maximum of energy is found which passes some electromagnetic cluster (which has a size of 4×4 at the L1).

In order to reduce the rates early, in the HLT a fast reconstruction and efficient preselection is performed, which uses a subset of the detector in the neighbourhood of the η and ϕ position given by the L1 and applies trigger specific reconstruction algorithms (fast and robust) with the full granularity of the detector (which is available at this point). The calorimeter reconstruction is similar to the offline version[23], with the difference that the offline version uses the sliding window algorithm, while the online version gets the cluster seed from the most energetic cell in the second EM layer. If the electron criteria for selection are fulfilled, a search for tracks is performed in front of the cluster and the candidates are created by the presencce of a matching reconstructed track. For that purpose, fast tracking algorithms are used. If no track is found, the candidate is marked as unconverted photon, since conversion algorithms have not been run at this point[24].

The HLT uses offline algorithms as much as possible, with the main difference that while offline runs reconstruction once per event, accessing the whole detector, the HLT runs per RoI, only in the subsample of the detector (thus, it can run multiple times per event, once for each defined RoI). The photon clusters are searched and reconstructed in RoIs of size $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$. Due to timing constraints, non-essential parts for the identification or reconstruction of e/γ candidates are omitted, as the Bremsstrahlung recovery for electrons, but the identification performed is very similar in the trigger to the offline and the cuts are selected in accordance with the offline ones.

Chapter 3

Calibration

The reconstructed clusters are calibrated in order to correct for energy lost from several sources inside the detector, such as in the material upstream of the calorimeter and for lateral and longitudinal shower leakage[8], [26].

3.1 Overview of the Calibration Procedure

The full calibration of calorimeter clusters for electrons and photons in ATLAS [8] can be seen in Figure 3.1. A brief description of every step follows:

1) MC-based training Through multivariate techniques, the EM cluster measurements are calibrated using simulated Monte Carlo samples. The constants are determined from a multivariate algorithm (MVA), separately for electrons, converted and unconverted photons (though at the trigger level, there is no distinction between converted and unconverted photons, so the calibration is done as if there were only unconverted photons). In order to achieve this, a precise description of the detector geometry and interactions of particles with matter is required in the simulation level.

2) Longitudinal layer calibration The different longitudinal layers of the EM calorimeter need to be equalized in the data with respect to the energy, prior to the determination of the energy in order to ensure a correct extrapolation of the response in the full p_T range.

3) MC-based calibration The trained MC calibration is now applied to the reconstructed energies in the clusters in both data and simulation samples.

4) Uniformity corrections These are correction that are not included in the simulation which take into account response variation in specific detector regions, such as non-optimal HV regions, geometric effects or biases from the LAr calorimeter electronic calibration.

5) Electron response smearing and calibration $Z \to ee$ samples are used to determine the scale factor, which is applied to data to agree with the simulation. The simulation is then smeared to have the same resolution as in the data.

6) Data-driven validation The calibrated energy scale is validated with candidates from $J/\psi \to ee$ (for electrons) and $Z \to ee\gamma$ (for photons) events in data.

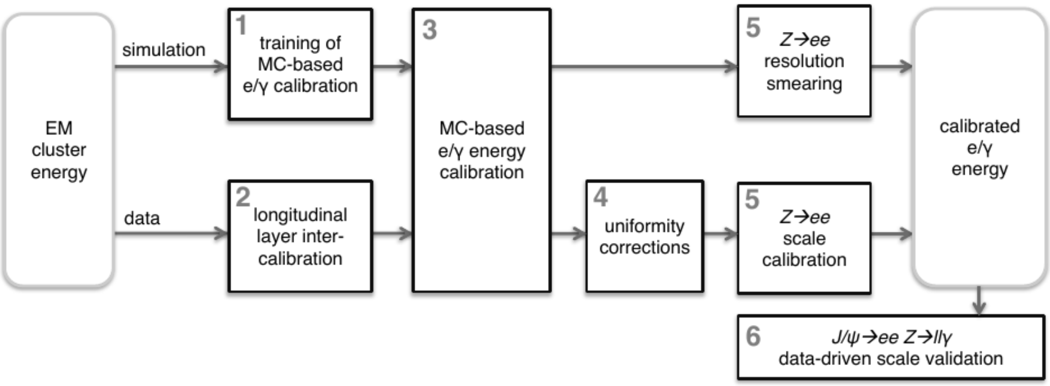


Figure 3.1: Schematic overview of the calibration procedure for electrons and photons.

Online calibration uses a simplified version of the full procedure[27]. There are no data-driven corrections online, and the energy calibration is entirely based on Monte Carlo, with a calibration which relies on boosted decision trees (the selected MVA algorithm) which determines the corrections for both electrons and photons independently but, as stated above, photons are not separated to converted and unconverted, since conversion information is not available at the HLT.

3.2 Implementation of the MVA Algorithm

The calibration is implemented with the Toolkit for Multivariate Analysis framework (TMVA) [28], which is integrated into the analysis framework ROOT[29]. The toolkit includes a large variety of multivariate classification and regression algorithms, alongside with training, testing, performance and evaluation tools for all of them via interfaces, so the user does not need to take care of the underlying details.

A multivariate algorithm is a statistical technique used to determinate the contribution of a set of variables to a single outcome of interest. A typical algorithm takes several input variables and outputs a single value (for regression) or label (for classification) and consists of two distinct stages, the training and application, where the training includes the learning stage and the testing.

In the training, the simulated data are split into two subsets, one for the training itself and another for testing the performance of the recently trained algorithm. One could in principle train and test with the same sample, but that would estimate an optimistic error (often called “resubstitution error”)[30]. The user of the algorithm defines a set of input variables and the MVA algorithm to be used. Among the many options available in the framework, the Boosted Decision Tree (BDT, described in section 3.2.1) was the selected algorithm because of its simple nature, which requires no special configuration of any kind. An independent training is performed for each of the several defined bins (see section 3.2.4).

The calibration has already been implemented in the offline analysis via a dedicated software package, and some additions have been done to it in order to account for online calibration.

3.2.1 Decision Tree

The Decision Tree is a well-known classifier or regressor structured as a binary tree, where the data are split into several non-overlapping regions. Basically several yes/no decision are taken on a single variable at a time until a stop criterion is met in a way that each one of the regions either is classified (for the classification; which usually goes as signal or background) or represents a value of the target variable (for the regression; this variable is the function the regression is trying to estimate). A simple example of this tree can be seen in Figure 3.2.

They are similar to a rectangular cut based selection, but where the cuts allow the selection of one hypercube in the phase space, the tree splits the phase space into a large number of hypercubes, where each one of them can be attributed to a constant value for the regression tree or can be identified as “signal” or “background” for the classification. Since for this work the important use of the decision tree is as a regressor, in the following *decision tree* will refer to the regression tree.

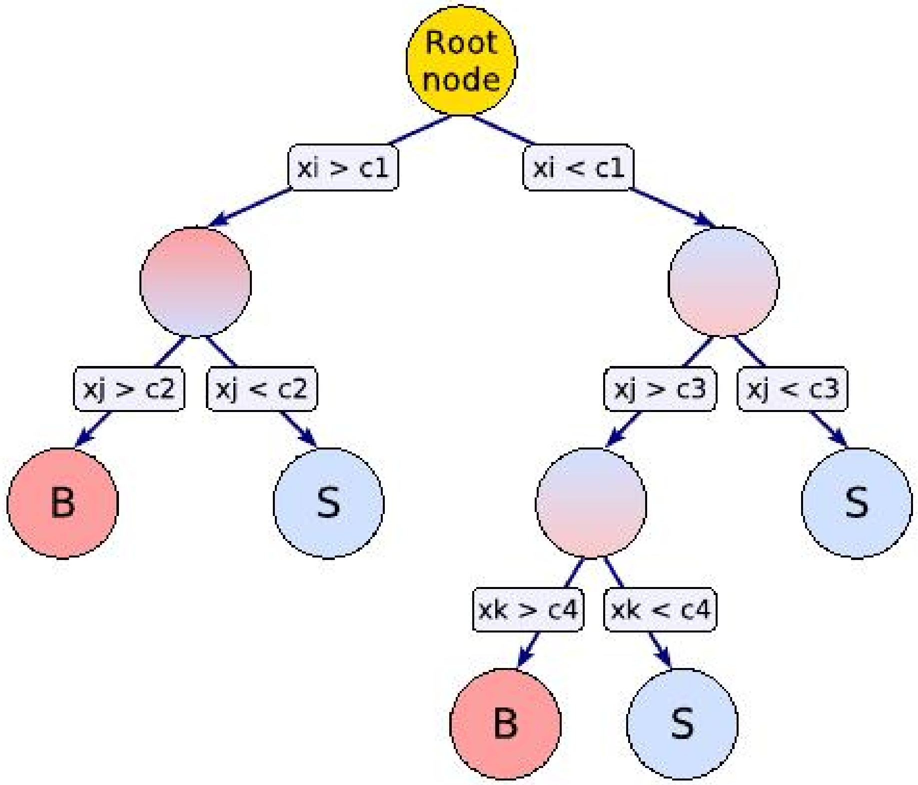


Figure 3.2: Simple schematic view of a decision tree for classification. The simplicity of this scheme makes its interpretation rather straightforward. From the root node, the data is split using x_i as discriminant variable. Each split uses the variable which gives the best separation between signal (S) and background (B), thus the same variable can be used in several nodes, while others may not be used at all.

Decision trees have the big flaw that they can be unstable with respect to statistical fluctuations in the training sample so the usual method for working with them is through boosting. Boosting is a way of enhancing the performance and increasing the stability of otherwise weak MVA methods. It consists in sequentially applying an MVA algorithm to reweighted ("boosted") versions of the training sample. This way, the boosting of a decision tree extends the single tree to a set of many trees, simply called a *forest*. The forest is combined into a single regressor which is given by a weighted average of the individual decision trees. A disadvantage of this method of improving the training is that the straightforward interpretation of the decision tree is lost. For a few trees, one can interpret the result, but it is difficult to be able to do it for hundred of trees in a forest.

The response of a boosted decision tree is given by a weighted sum of the responses of individual trees (or more generally speaking, "weak learners") [28] and the boosting is performed in a way which minimizes the deviation (measured through a *loss function* $L(F, y)$) between the estimation F and the true value y obtained from the training sample. The several boosting methods differ in the loss function used to calculate the deviation.

3.2.2 Target variable

The purpose of the calibration is to estimate the true energy of the particle from the measurements done by the detector, so the MVA target should be either the true energy itself or some quantity which allows to calculate the true energy (e.g. as a constant to apply to the measured energy), so the targets considered have been [31] the true energy (E_{true}, directly estimated by the MVA output), the ratio between the true energy and the measured energy ($E_{\text{true}}/E_{\text{acc}}$, "acc" stands for accordion) and the ratio between the true energy and the previous calibrated energy.

One important feature of the MVA is that they work better if the target range is small with respect to the input range. As explained in section 3.2.1, the response of the decision tree is given by a sum of the individual trees and this sum is finite, despite the large number of trees used (and it should not be too large, considering that a compromise between performance and used resources such as memory, disk usage, processing power, etc., must be done[1]), so it cannot cover with enough precision the full range of expected energy for photons and electrons, which goes from a few GeV to a few TeV, therefore the estimate of the true energy is not a good option.

---

[1] Moreover, since this calibration should be *online*, it is required that it does not take too long to be processed, since there is a limited latency time between processed events

On the other hand, the other two options, which estimate a rate, are close to unity, but correlations of target with inputs are very complex with the standard calibration, and it *depends* on the standard calibration for obvious reasons. Moreover, the intention is to avoid the use of that calibration. This means that the better option is the ratio $E_{\text{true}}/E_{\text{acc}}$ as a correction factor to E_{acc} which, moreover, assures independence from the original calibration.

3.2.3 Input variables

The MVA algorithms can use an arbitrary number of input variables and the process of training a new version of the algorithm is not too expensive, so for different processes different sets of input variables can be used. The offline calibration takes advantage of this feature using a different set of variables for electrons and photons in order to take into account conversion vertex reconstruction. In the online case, since there is no such reconstruction at that point, the set of variables is the same for both kinds of particles:

Energy in the accordion E_{acc}, the sum of the energies of the accordion layers, strips (1), middle (2) and back (3).

Ratio of the energy in the presampler to energy in the accordion E_0/E_{acc}, used only for cluster where the presampler is present ($|\eta| < 1.8$, see section 2.1).

Ratio of the energy in the first two layers E_1/E_2, which replaces the previous choice in [31], where the shower depth was used, defined as $X = \sum X_i E_i / \sum E_i$, with X_i the amount of material in the layer i. The performance is similar between both choices, as can be seen in sections 4.3 and 4.3.1, but E_1/E_2 is cheaper in terms of resources.

Pseudorapidity in the ATLAS frame η_{cluster}, which takes into account the misalignment of the detector and other effects.

Cell index defined as the integer part of $\eta_{\text{calo}}/\Delta\eta$, where η_{calo} is the pseudorapidity of the cluster in the calorimeter frame and $\Delta\eta$ is the size of the cell in the middle layer (see section 2.1). This results in a number between 0 and 99.

Pseudorapidity with respect to cell edge defined as η_{calo} modulo $\Delta\eta$.

ϕ **with respect to the lead absorbers** defined as ϕ_{calo} modulo the periodicity of the lead sheets in each region ($2\pi/1024$ in the barrel, $2\pi/768$ in the endcap)

It is worth noting that for the offline calibration, some additional variables are needed for photons, since a different calibration is required for converted photons. Such variables include the radius of the conversion, the ratio of the conversion p_T to the E_T in the calorimeter and the fraction of the p_T carried by the more energetic conversion track. But since the online version does not make the distinction, those variables are left out for the final calibration.

3.2.4 Binning of the sample

In order to help the MVA optimise the response in the different regions of the phase space, the sample is divided into several bins in $|\eta_{\text{cluster}}|$ and transverse energy in the accordion E_T^{acc}, chosen to match approximately the non-uniformities of the detector and significant changes in the response, and according to particle type (electron or photon; offline calibration includes the division of whether the photons come from a conversion vertex or not). For each region an independent training is performed, so the application of the calibration starts with the selection of the corresponding bin.

A rectangular grid of 10×9 bins in $|\eta_{\text{cluster}}| \times E_T^{\text{acc}}$ has been defined, with another 4×6 "special bins" for regions close to the edges of the half-barrel modules and for the "crack region", where the barrel and endcap modules meet, between $1.37 < |\eta_{\text{cluster}}| < 1.52$. The bins are defined as follows:

- Bins in $|\eta_{\text{cluster}}|$ (normal): 0.05, 0.65, 0.8, 1.0, 1.2, 1.37, 1.55, 1.74, 1.82, 2.0, 2.2 and 2.47, $1.37 - 1.52$ is excluded.
- Bins in $|\eta_{\text{cluster}}|$ (special): $0 - 0.05$, $1.37 - 1.52$, $1.52 - 1.55$, and $2.47 - 2.5$
- Bins in E_T^{acc} (normal): 0, 10, 20, 40, 60, 80, 120, 500, 1000 and $50\,000$ GeV
- Bins in E_T^{acc} (special): 0, 25, 50, 100, 500, 1000 and $50\,000$ GeV

Chapter 4

Results and Statistical Analysis

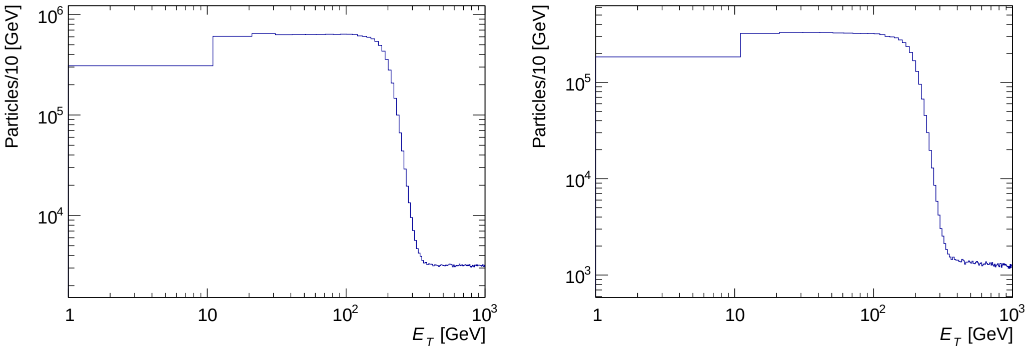
4.1 Sample selection

4.1.1 Samples

Several sets of simulated samples have been prepared in order to perform this work. The main sets are two, one for resolution studies prior to the calibration training and testing, and another set for the training and testing itself.

For the resolution studies, large validation samples of $Z \rightarrow ee$ and $H \rightarrow \gamma\gamma$ events were used. These samples have been generated with PYTHIA 8[32] and interfaced with POWHEG[33]. After generation, the samples have undergone the full simulation of the ATLAS detector [34], which is based on GEANT4[35]. This simulated sample is then reconstructed the same way the data samples are. The validation samples use a center of mass energy of 14 TeV.

For the production of the e/γ calibration itself, single particles have been generated with no pile-up and an E_T distribution ranging from 1 GeV to 3 TeV. The electron sample has 20 million single electrons and the photon sample, 40 million single photons. Their energy distributions can be seen in Figure 4.1



(a) Photons

(b) Electrons

Figure 4.1: Transverse energy distribution for the single particle samples

4.1.2 Selection

In order to select the objects for the training, some minimum cuts have been applied, namely that a match has been found between the online object and the simulation (truth particle), that the particle is in the high precision physics region ($|\eta| < 2.5$), a loose identification, and that a difference between the reconstructed energy and the true energy of the particle not greater than 50%, in order to remove outliers that could reduce the performance of the MVA.

The matching between the online objects and the truth particle is carried out as follows. The online object is matched to an offline object, which is done requiring $\Delta R < 0.1$, with ΔR calculated as

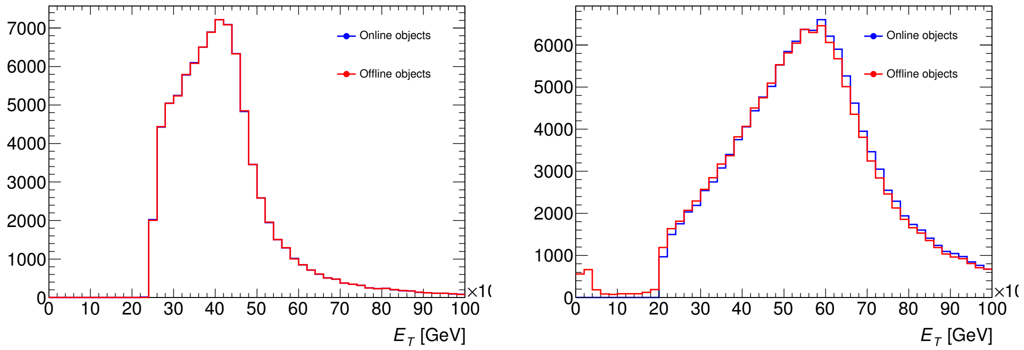
$$\Delta R^2 = \Delta\eta^2 + \Delta\phi^2 \tag{4.1}$$

In the data model of ATLAS, offline e/γ objects have a link to a matched truth particle, if any, so if an offline match is found, the online object is matched to its truth particle match.

4.2 Resolution Studies before MVA Calibration

Preliminary studies have been done over the performance of several important variables used in the calibration process, in order to compare the behaviour in an online environment with respect to the referential offline behavior. The resolution studies have been done with the $Z \to ee$ and $H \to \gamma\gamma$ samples described in section 4.1.1.

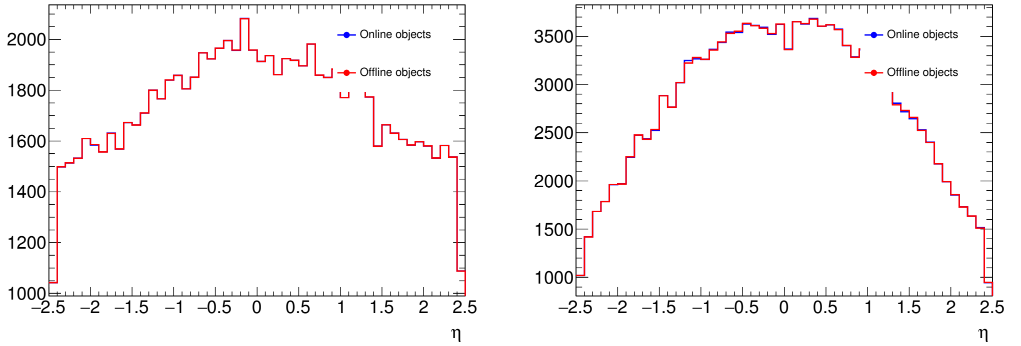
Figure 4.2 shows the transverse energy distribution for the used electron samples. These distributions have been created after finding the match between online and offline object. As well, Figures 4.3 and 4.4 show distributions for η and ϕ angle. No scaling has taken place, since because of the matching, the number of online and offline entries is the same. There is a small bump at low energy for the offline photons, which could arise from a mismatch with low energy particles. The threshold for reconstructed offline electron is 25 GeV, and 20 GeV for photons.



(a) Electrons E_T distribution

(b) Photons E_T distribution

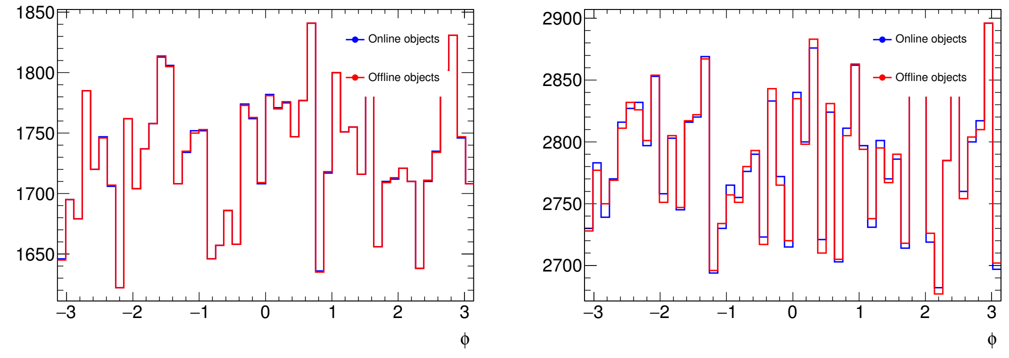
Figure 4.2: Electrons (left) and photons (right) transverse energy distribution for validation samples $Z \rightarrow ee$ and $H \rightarrow \gamma\gamma$.



(a) Electrons η distribution

(b) Photons η distribution

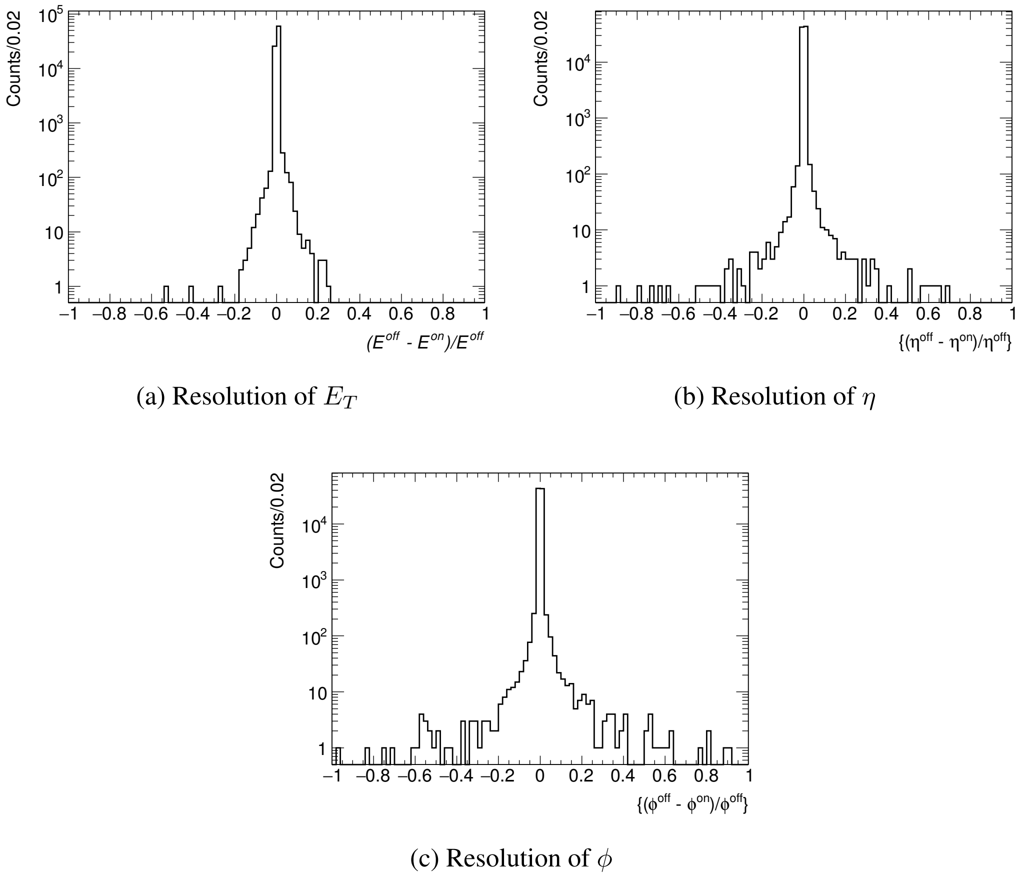
Figure 4.3: Electrons (left) and photons (right) pseudorapidity distribution for validation samples $Z \rightarrow ee$ and $H \rightarrow \gamma\gamma$.



(a) Electrons ϕ distribution

(b) Photons ϕ distribution

Figure 4.4: Electrons (left) and photons (right) azimuthal angle distribution for validation samples $Z \to ee$ and $H \to \gamma\gamma$.



(a) Resolution of E_T

(b) Resolution of η

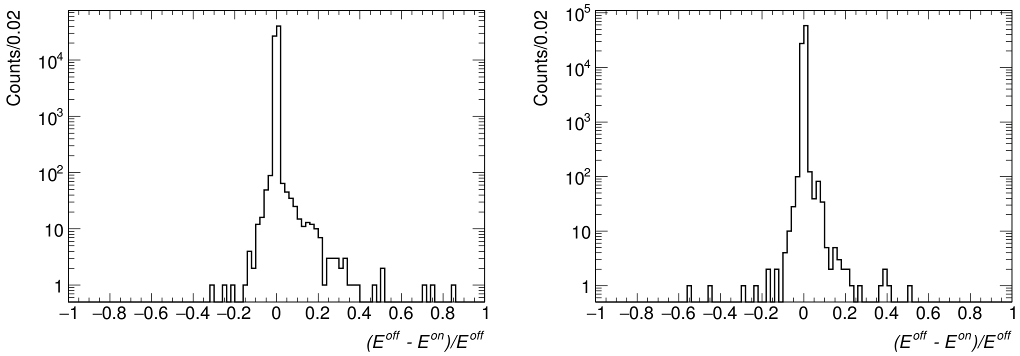
(c) Resolution of ϕ

Figure 4.5: Resolution for electrons

A good agreement can be found in the shapes of the offline and online objects, and Figure 4.5 shows the resolution of those quantities for the electrons. It can be seen that the shape agreement corresponds to a good resolution for this. The resolution for quantity X has been measured as

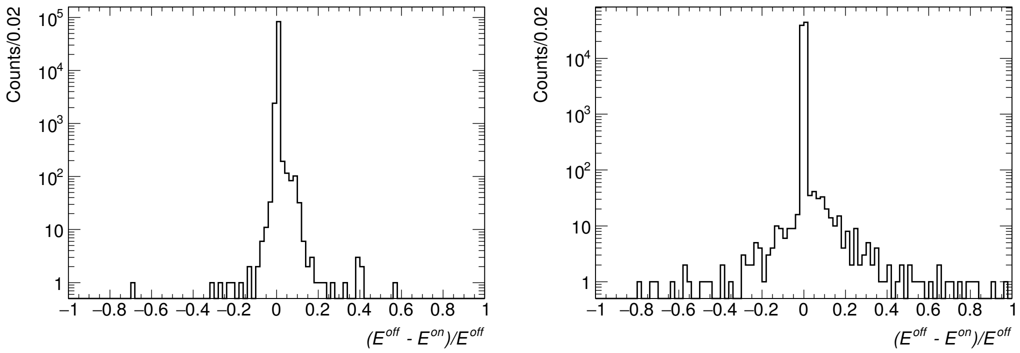
$$r = \frac{X_{\text{off}} - X_{\text{on}}}{X_{\text{off}}} \tag{4.2}$$

Figures 4.6 and 4.7 show the resolution in the calorimeter layers for electron and photon samples, respectively.



(a) Presampler

(b) Strips layer

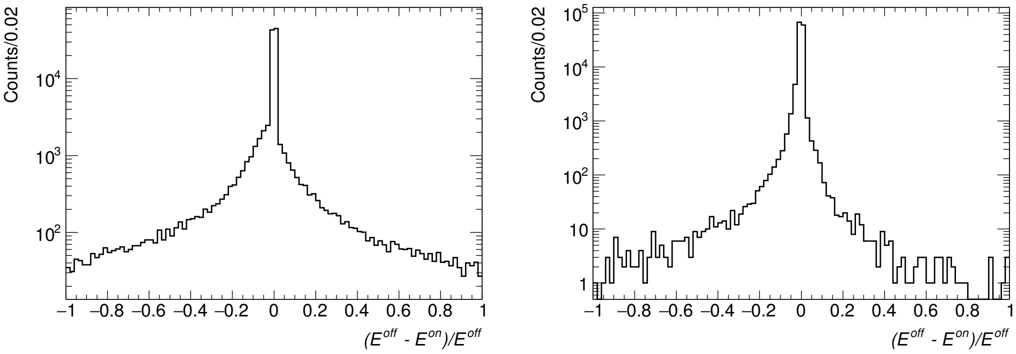


(c) Middle layer

(d) Back layer

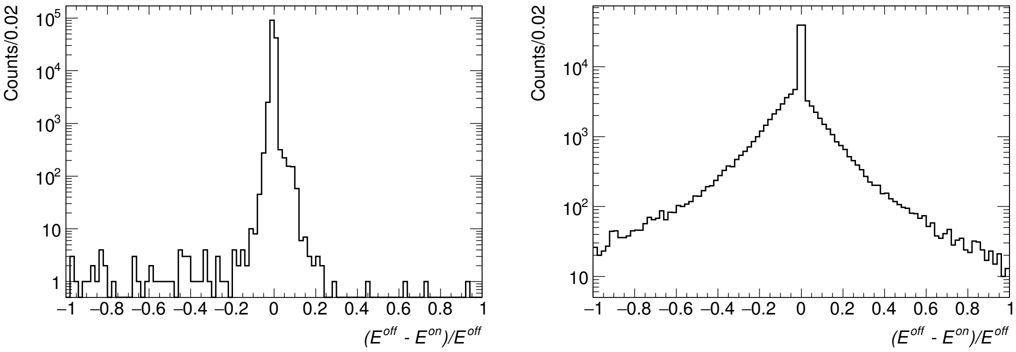
Figure 4.6: Resolution in calorimeter layers for electron sample

In the Figures, an enormous peak around 0 can be seen, which is a good signal that the online performance compares well to the offline performance. Both electrons and photons



(a) Presampler

(b) Strips layer



(c) Middle layer

(d) Back layer

Figure 4.7: Resolution in calorimeter layers for photon sample

also show a degradation in the resolution in the back layer, which is understandable since there is less energy measured in that layer and the statistical accuracy is worse in those regions of the calorimeter.

4.3 Performance of the MVA Calibration

The main idea of the calibration is to have E/E_{true} as close to unity as possible, and its performance is measured in terms of its linearity. The linearity is quantified as the peak position of the quantity E/E_{true} as a function of E_T^{true}, estimated by the mean value of a Gaussian function fitted to the core of the distribution in each $(E_T^{\text{true}}, |\eta_{\text{true}}|)$ bin, restricted to [-1,+2] standard deviations to avoid biases from the tails. In order to do this evaluation, the set of events marked as the test sample were used, which, as stated in section 3.2, were not included in the training sample, in order to avoid biases.

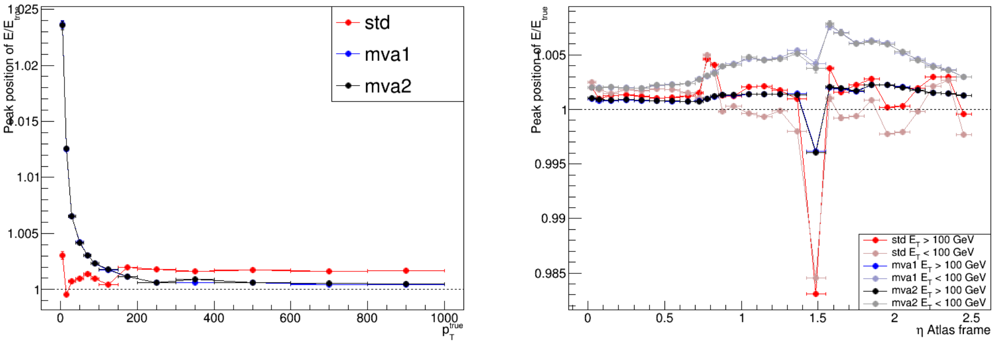
Figure 4.8, where the y axis corresponds to the peak position of E/E_{true}, presents the linearity studies for electrons and photons with the test sample of the MVA algorithm. On these plots, "std" is referred to the Run 1 original cluster calibration (sometimes referred to as the "standard" calibration) —that is, the previously used calibration. The label "mva1" stands for the calibration using the shower depth as input to the calibration and "mva2" stands for the calibration using the E_1/E_2 ratio as input.

The first observation is that there is no important difference between using shower depth or the E_1/E_2 ratio, which is an advantage since the second is cheaper in terms of calculational resources. As can be seen, in η, the distribution is rather flat, except for the crack region, where a large amount of material is present upstream of the calorimeter, however, the low energy distributions are farther from unity than the high energy ones, an effect which is more vissible in the distribution over p_T on the left side. This effect can be seen with the example distributions for electrons in Figure 4.9, where the shower depth was used as input variable. The previous (the standard) calibration tried to adjust the peak of the E/E_{true} to be near one, and in order to mimic that behaviour, a set of corrections (shifts) must be computed. Figure 4.10 shows the performance with inclusive η and p_T after the mentioned corrections have been applied. An immediate improvement in the low energy regions can be seen for every case. Finally, it can be said that the non-linearity is in the order of the per-mil percent, being worst for the crack region, where the overall resolution of the calorimeter deteriorates.

4.3.1 Stability

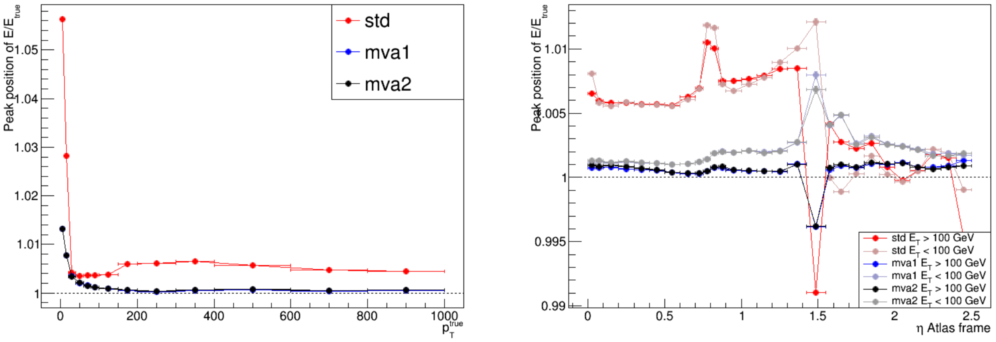
The stability, measured as the behaviour of the response with respect to the input variables, has been studied in order to check for the proper use of variables and potential improvements where they might appear.

Figures 4.11 and 4.12 show the fraction of the total energy for the measured energy in every layer of the calorimeter for electrons and photons, respectively. Since most of the energy is deposited in the second layer, there are fewer counts in the lower region of the spectrum, hence the inefficiencies in that region, and the same effect, but the other way around, i.e. for higher regions of the spectrum, happens for the deposited energy in the first



(a) Online electrons, inclusive η

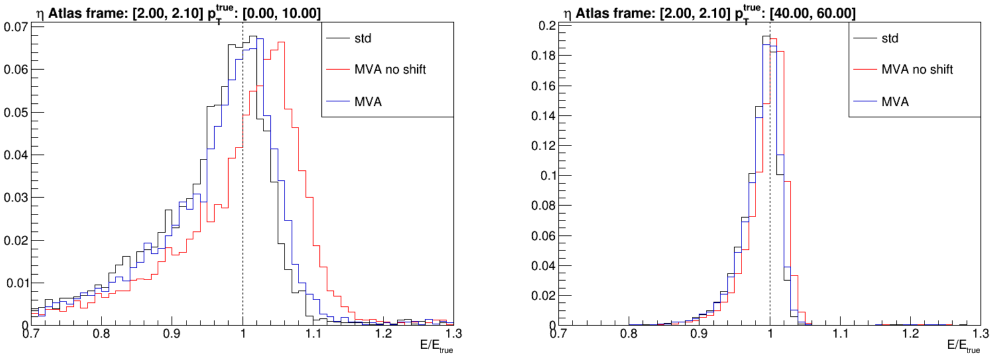
(b) Online electrons, inclusive p_T



(c) Online photons, inclusive η

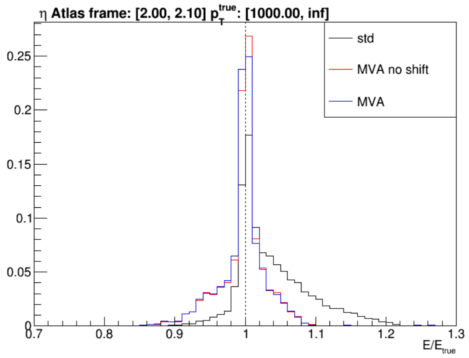
(d) Online photons, inclusive p_T

Figure 4.8: Linearity plots for inclusive η as a function of p_T (left) and inclusive p_T as a function of η (right). The three curves represent the standard measurement (in red for $E_T > 100$ GeV and faded red for $E_T \leq 100$ GeV), the calibration using shower depth as input variable (in blue for $E_T > 100$ GeV and faded red for $E_T \leq 100$ GeV) and the calibration using the ratio between the measurements in the first and second layer of the calorimeter(in black for $E_T > 100$ GeV and faded black for $E_T \leq 100$ GeV). For the full p_T range, the curves are, respectively, red, blue and black.



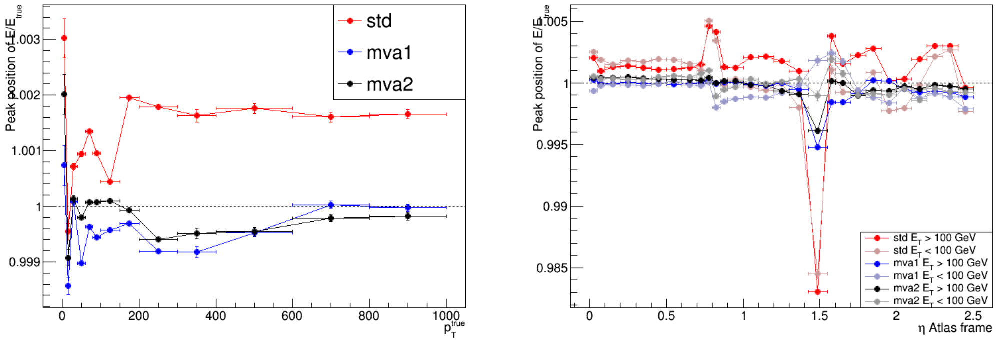
(a) Low energy

(b) Medium energy



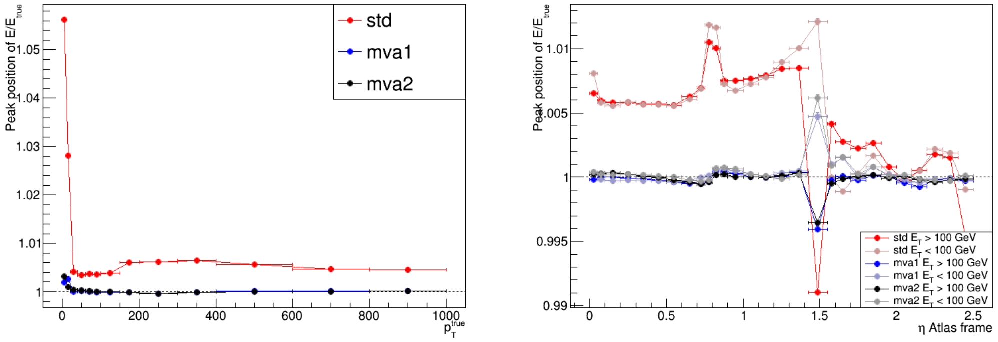
(c) High energy

Figure 4.9: Sample distributions for low, medium and high energies for the electrons for $\eta = [2.00, 2.10]$. As seen, the need for a shift decreases as the energy grows.



(a) Online electrons, inclusive η

(b) Online electrons, inclusive p_T

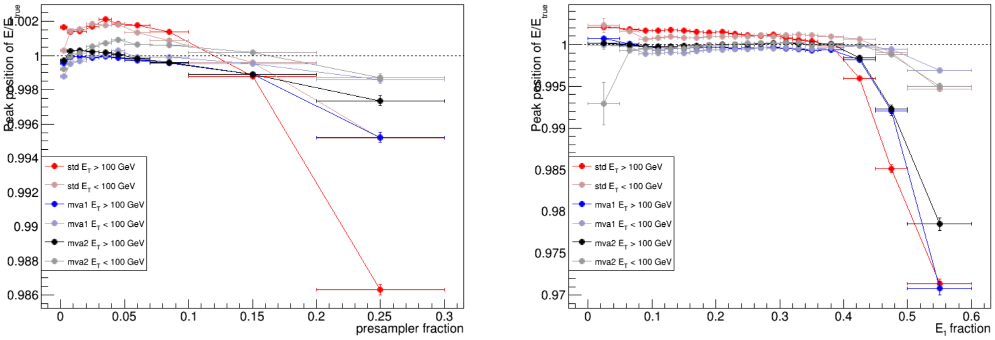


(c) Online photons, inclusive η

(d) Online photons, inclusive p_T

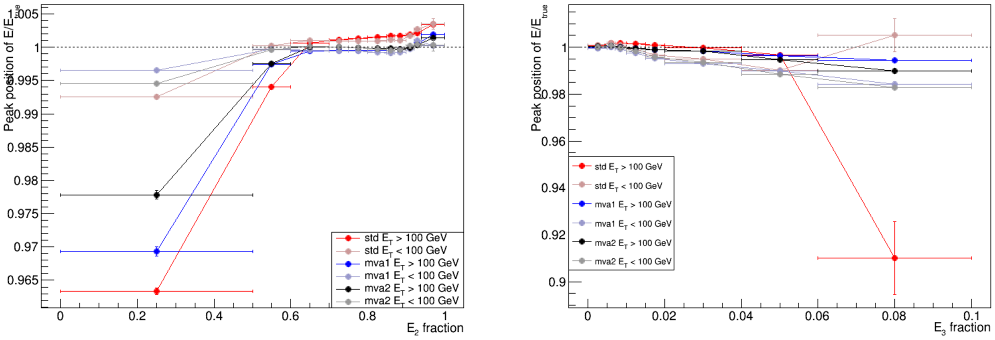
Figure 4.10: Linearity plots for inclusive η as a function of p_T (left) and inclusive p_T as a function of η (right), after the application of corrections. The three curves represent the standard measurement as in Figure 4.8

layer. Figure 4.13 shows the stability for the E_1/E_2 ratio and for the shower depth. In both cases there is an improvement of the stability with respect to the previous calibration of the energy and since there is no significant difference between both options, the ratio E_1/E_2 can be selected as the input variable instead of the shower depth.



(a) Presampler fraction

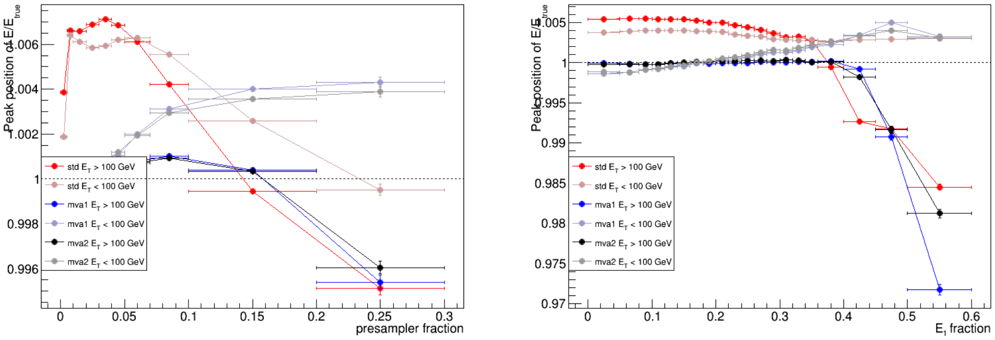
(b) E_1 fraction



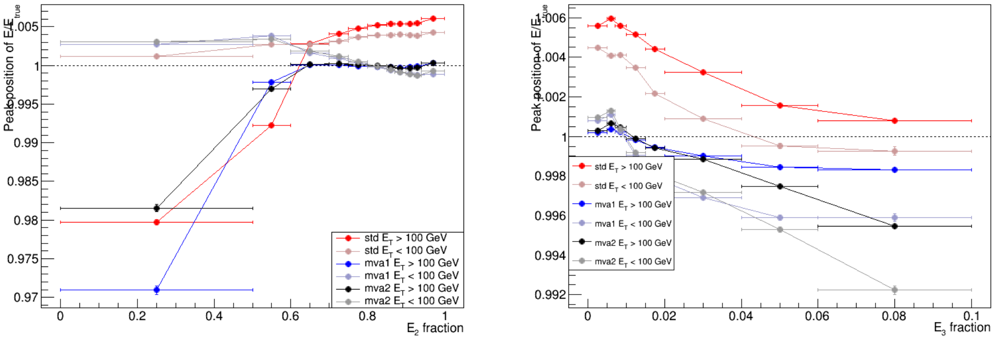
(c) E_2 fraction

(d) E_3 fraction

Figure 4.11: Peak position of E/E_{true} of online electrons as a function of the fraction of the total energy present in the presampler, E_1, E_2 and E_3 layers, compared between the previous calibration, MVA calibration with shower depth as input and MVA calibration with E_1/E_2 ratio in red, blue and black, respectively, for $E_T > 100$ GeV. The faded colors show lower energy, $E_T < 100$ GeV.

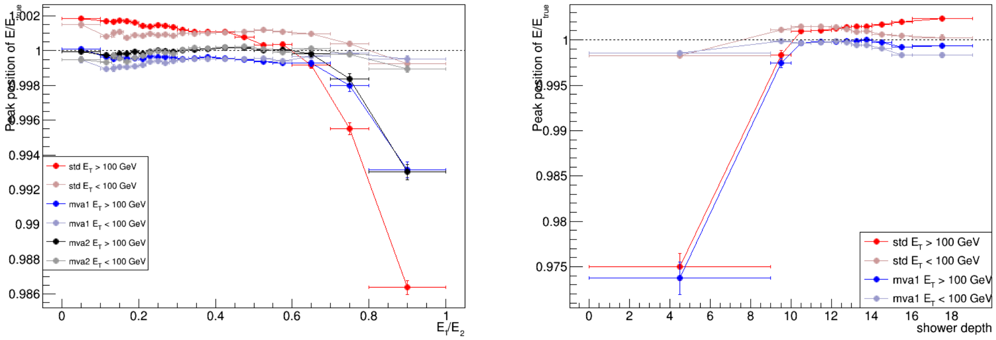


(a) Presampler fraction (b) E_1 fraction



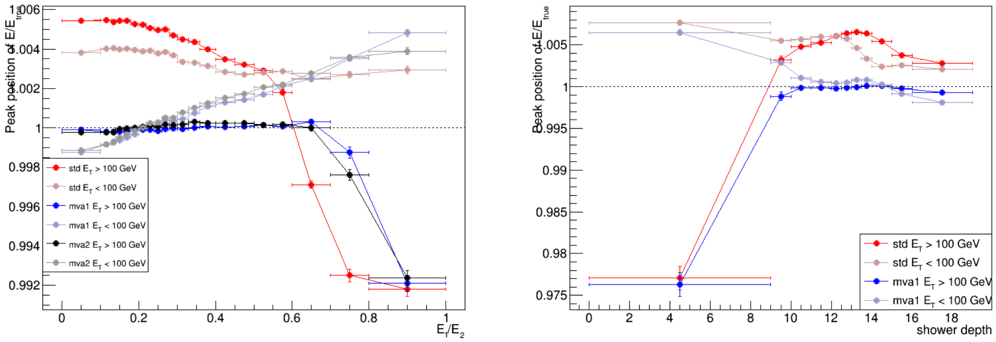
(c) E_2 fraction (d) E_3 fraction

Figure 4.12: Peak position of E/E_{true} of online photons as a function of the fraction of the total energy present in the presampler, E_1, E_2 and E_3 layers, compared between the previous calibration, MVA calibration with shower depth as input, and MVA calibration with E_1/E_2 ratio in red, blue and black, respectively, for $E_T > 100$ GeV. The faded colors show lower energy, $E_T < 100$ GeV.



(a) E_1/E_2 fraction for electrons

(b) Shower depth for electrons



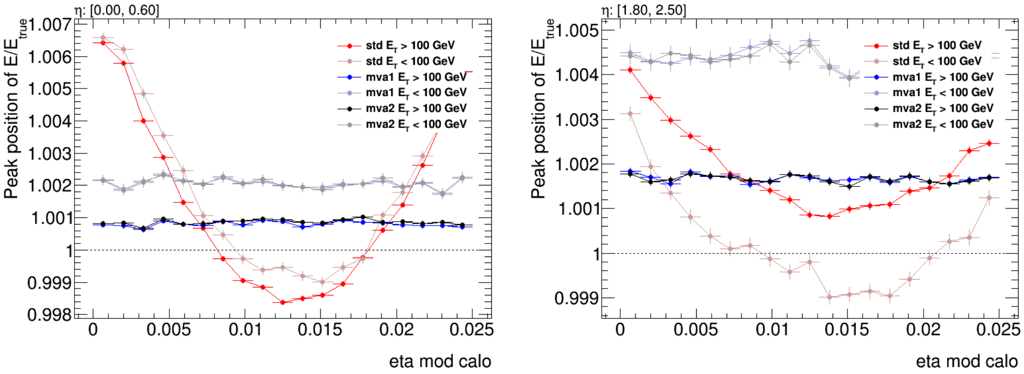
(c) E_1/E_2 fraction for photons

(d) Shower depth for photons

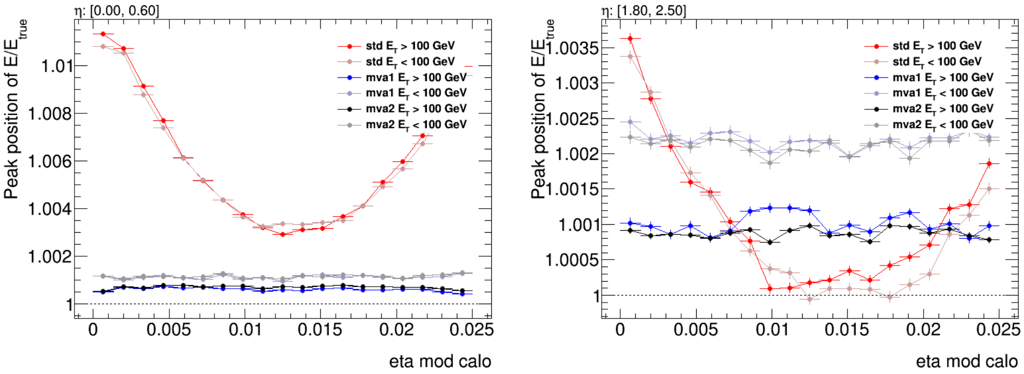
Figure 4.13: Peak position of E/E_{true} of online electrons and photons as a function of the fraction of the energy in the first two layers (left) and of the shower depth (right), compared between the previous calibration, MVA calibration with shower depth as input, and MVA calibration with E_1/E_2 ratio in red, blue and black, respectively, for $E_T > 100$ GeV. The faded colors, show lower energy, $E_T < 100$ GeV.

4.4 Modulation

The peak position of E/E_{true} has been studied as a function of the η position with respect to the cell edge and the ϕ coordinate with respect to the lead absorbers, both variables used as input to the MVA calibration, as can be seen in section 3.2.3. A “modulation” appears in the peak position with respect to these two variables. In previous calibrations for offline analysis, this modulation has been corrected with a parameterisation [22], but for the MVA calibration, they have been included as input variables. Figures 4.14 and 4.15 show this effect. As can be seen, the modulation effect is eliminated in the MVA calibration.

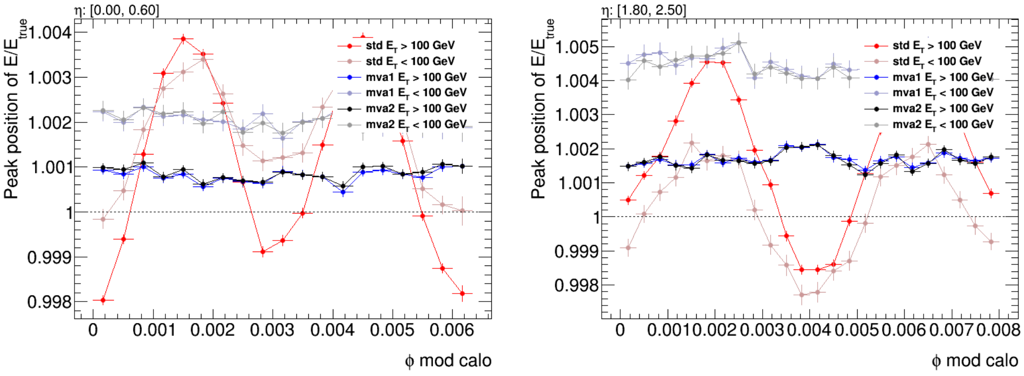


(a) Electrons, $0 < |\eta| < 0.6$ (b) Electrons, $1.8 < |\eta| < 2.5$



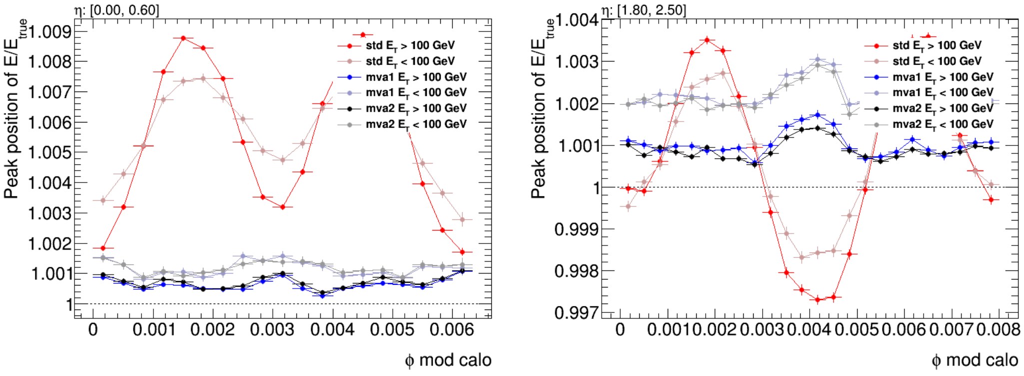
(c) Photons, $0 < |\eta| < 0.6$ (d) Photons, $1.8 < |\eta| < 2.5$

Figure 4.14: Peak position of E/E_{true} of online electrons and photons as a function of η with respect to the cell edge. The comparison is between the original measurement, calibration using shower depth as input, and calibration using E_1/E_2, in red, blue, and black, respectively. Faded colors are for low energy ($E_T < 100$ GeV)



(a) Electrons, $0 < |\eta| < 0.6$

(b) Electrons, $1.8 < |\eta| < 2.5$



(c) Photons, $0 < |\eta| < 0.6$

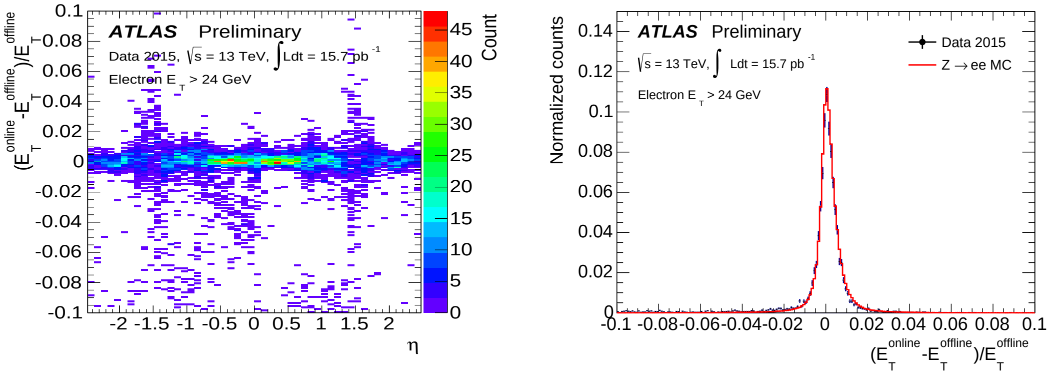
(d) Photons, $1.8 < |\eta| < 2.5$

Figure 4.15: Peak position of E/E_{true} of online electrons and photons as a function of ϕ with respect to the lead absorbers. The comparison is between the original measurement, calibration using shower depth as input, and calibration using E_1/E_2, in red, blue, and black, respectively. Faded colors are for low energy ($E_T < 100$ GeV)

4.5 Resolution Studies on 2015 Data and Monte Carlo

The Figure 4.16 displays the energy resolution in the high level trigger for electrons with respect to the offline calibration (left) and a comparison between the expected resolution from Monte Carlo simulations and the measured resolution (right), integrated over the whole pseudorapidity range[27]. The data have been taken in the Run 2 with center of mass energy of $\sqrt{s} = 13$ GeV and integrated luminosity of 15.7 $1/\text{pb}$.

For these plots, the trigger electron is required to have $E_T > 24$ GeV and pass medium identification requirements. The resolution has been measured with a tag-and-probe method which uses the $Z \to ee$ decay where no background subtraction has been applied. It can be seen that the resolution is excellent, in general, and in most of the pseudorapidity range, with the exception of the crack region, which as been said before, presents a large amount of material and the overall resolution of the calorimeter worsens in that region.



(a) As a function of pseudorapidity

(b) Compared to Monte Carlo simulation

Figure 4.16: Electron energy resolution online with respect to offline reconstruction

Chapter 5

Conclusions

A new calibration has been implemented in the calorimeter to be run at the trigger level in the ATLAS detector, via the application of the new MVA offline calibration in the online analysis, in order to replace the previous ones.

The feature of the ease of getting a new set of corrections has been used in order to try two sets of input variables, namely including the shower depth or the E_1/E_2 ratio as replacement for the shower depth, and no significant difference has been found between the two approaches. Also, the implementation has already been used to get the first official performance measurements for the trigger system with the Run 2 of the LHC.

A comparison of the performance of the new calibration has been made and compared to the already approved offline calibration and the previous trigger calibration, and improvements have been found in linearity and stability with respect to several input variables.

An important aspect of this calibration is that does not require data driven corrections, and it relies entirely on a good simulation of the geometry of the detector and as stated earlier, in the event that an update, improvement or change in the geometry of the detector occurs, the new information can be used to provide a new sets of corrections with similar performance in an easy and relatively fast way.

The main source of improvement for future versions of this calibration would come from two sources: a better understanding of the detector geometry, thus improving the simulation, and a new set of variables that contribute to the output. In particular, if the conversion information for photons is implemented at the trigger stage, it can be included and this way improve the performance for photons.

Finally, there are some questions left open, which have to do with the pile-up dependence of the calibration: how its performance is affected by higher pile-up, a question of particular importance when it comes to study heavy ion collisions, where there are tens or hundreds of particles per event going through the detectors.

Bibliography

[1] The ATLAS Collaboration et al, "The ATLAS Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation*, vol. 3, no. 08, S08003, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[2] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation*, vol. 3, no. 08, S08001, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[3] The CMS Collaboration et al., "The CMS experiment at the CERN LHC", *Journal of Instrumentation*, vol. 3, no. 08, S08004, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[4] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.*, vol. B716, pp. 1–29, 2012. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[5] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.*, vol. B716, pp. 30–61, 2012. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[6] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961, ISSN: 0029-5582. DOI: http://dx.doi.org/10.1016/0029-5582(61)90469-2. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[7] S. Weinberg, "A model of leptons", *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 21 Nov. 1967. DOI: 10.1103/PhysRevLett.19.1264. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[8] The ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data", *Eur. Phys. J. C*, no. arXiv:1407.5063. CERN-PH-EP-2014-153, 74. 51 p, Jul. 2014. [Online]. Available: `https://cds.cern.ch/record/1744017`.

[9] ——, "Search for resonances decaying to photon pairs in 3.2 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector", CERN, Geneva, Tech. Rep. ATLAS-COM-CONF-2015-096, Dec. 2015. [Online]. Available: `https://cds.cern.ch/record/2111338`.

[10] The CMS Collaboration, "Search for new physics in high mass diphoton events in proton-proton collisions at $\sqrt{s} = 13$ TeV", CERN, Geneva, Tech. Rep. CMS-PAS-EXO-15-004, 2015. [Online]. Available: `https://cds.cern.ch/record/2114808`.

[11] Y. Nakai, R. Sato, and K. Tobioka, "Footprints of new strong dynamics via anomaly and the 750 gev diphoton", *Phys. Rev. Lett.*, vol. 116, p. 151 802, 15 Apr. 2016. DOI: `10.1103/PhysRevLett.116.151802`. [Online]. Available: `http://link.aps.org/doi/10.1103/PhysRevLett.116.151802`.

[12] G. Li, Y.-n. Mao, Y.-L. Tang, *et al.*, "Pseudoscalar decaying only via loops as an explanation for the 750 gev diphoton excess", *Phys. Rev. Lett.*, vol. 116, p. 151 803, 15 Apr. 2016. DOI: `10.1103/PhysRevLett.116.151803`. [Online]. Available: `http://link.aps.org/doi/10.1103/PhysRevLett.116.151803`.

[13] C. Petersson and R. Torre, "750 gev diphoton excess from the goldstino superpartner", *Phys. Rev. Lett.*, vol. 116, p. 151 804, 15 Apr. 2016. DOI: `10.1103/PhysRevLett.116.151804`. [Online]. Available: `http://link.aps.org/doi/10.1103/PhysRevLett.116.151804`.

[14] W. S. Cho, D. Kim, K. Kong, *et al.*, "750 gev diphoton excess may not imply a 750 gev resonance", *Phys. Rev. Lett.*, vol. 116, p. 151 805, 15 Apr. 2016. DOI: `10.1103/PhysRevLett.116.151805`. [Online]. Available: `http://link.aps.org/doi/10.1103/PhysRevLett.116.151805`.

[15] "Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s}$=13 TeV in 2015 and 2016 with the ATLAS detector", CERN, Geneva, Tech. Rep. ATLAS-CONF-2016-059, Aug. 2016. [Online]. Available: http://cds.cern.ch/record/2206154.

[16] The ATLAS Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector", *JHEP*, vol. 09, p. 145, 2014. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[17] ——, *ATLAS liquid-argon calorimeter: Technical Design Report*, ser. Technical Design Report ATLAS. Geneva: CERN, 1996. [Online]. Available: https://cds.cern.ch/record/331061.

[18] W. Lampl, S. Laplace, D. Lelas, *et al.*, "Calorimeter Clustering Algorithms: Description and Performance", CERN, Geneva, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, Apr. 2008. [Online]. Available: https://cds.cern.ch/record/1099735.

[19] D. Scannicchio, "Atlas trigger and data acquisition: Capabilities and commissioning", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 617, no. 1–3, pp. 306–309, 2010, 11th Pisa Meeting on Advanced Detectors Proceedings of the 11th Pisa Meeting on Advanced Detectors, ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/j.nima.2009.06.114. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900209013382.

[20] P. Czodrowski, "The ATLAS Trigger System: Ready for Run 2", CERN, Geneva, Tech. Rep. ATL-DAQ-PROC-2015-038, Oct. 2015. [Online]. Available: https://cds.cern.ch/record/2058218.

[21] The ATLAS Collaboration, "Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System", CERN, Geneva, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, Sep. 2013, Final version presented to December 2013 LHCC. [Online]. Available: http://cds.cern.ch/record/1602235.

[22] ——, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, 2009. arXiv: 0901.0512 [hep-ex].

[23] ——, “Expected electron performance in the ATLAS experiment”, CERN, Geneva, Tech. Rep. ATL-PHYS-PUB-2011-006, Apr. 2011. [Online]. Available: https://cds.cern.ch/record/1345327.

[24] ——, “Expected photon performance in the ATLAS experiment”, CERN, Geneva, Tech. Rep. ATL-PHYS-PUB-2011-007, Apr. 2011. [Online]. Available: https://cds.cern.ch/record/1345329.

[25] ——, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, in. 2009, ch. Physics Performance Studies and Strategy of the Electron and Photon Trigger Selection, pp. 619–646. arXiv: 0901.0512 [hep-ex].

[26] ——, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, in. 2009, ch. Calibration and Performance of the Electromagnetic Calorimeter, pp. 44–71. arXiv: 0901.0512 [hep-ex].

[27] G. Pasztor, “The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and their Performance”, CERN, Geneva, Tech. Rep. ATL-DAQ-PROC-2015-053. arXiv:1511.00334, Oct. 2015. [Online]. Available: https://cds.cern.ch/record/2063746.

[28] A. Hoecker, P. Speckmayer, J. Stelzer, *et al.*, “TMVA - Toolkit for Multivariate Data Analysis”, *ArXiv Physics e-prints*, Mar. 2007. eprint: physics/0703039.

[29] R. Brun and F. Rademakers, “{root} — an object oriented data analysis framework”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 389, no. 1–2, pp. 81–86, 1997, New Computing Techniques in Physics Research V, ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/S0168-9002(97)00048-X. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016890029700048X.

[30] I. Narsky and F. C. Porter, “Introduction to classification”, in *Statistical Analysis Techniques in Particle Physics*. Wiley-VCH Verlag GmbH & Co. KGaA, 2013, pp. 165–193, ISBN: 9783527677320. DOI: 10.1002/9783527677320.ch9. [Online]. Available: http://dx.doi.org/10.1002/9783527677320.ch9.

[31] B. Lenzi and R. Turra, “Monte Carlo calibration update for electrons and photons using multivariate techniques”, CERN, Geneva, Tech. Rep. ATL-COM-PHYS-2013-1426, Oct. 2013.

[32] T. Sjöstrand, S. Ask, J. R. Christiansen, *et al.*, “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[33] C. Oleari, “The POWHEG-BOX”, *Nucl. Phys. Proc. Suppl.*, vol. 205-206, pp. 36–41, 2010. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph].

[34] The ATLAS Collaboration, “The ATLAS Simulation Infrastructure”, *Eur. Phys. J.*, vol. C70, pp. 823–874, 2010. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[35] S. Agostinelli, J. Allison, K. Amako, *et al.*, “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003, ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/S0168-9002(03)01368-8. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900203013688.