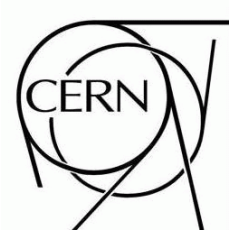
ATLAS NOTE

ATL-PHYS-PUB-2007-xxx

August 7, 2008



Expected Performance for the Identification of Photons

The ATLAS Collaboration

This note presents the descriptions and performance of photon identification methods in ATLAS. The reconstruction of an electromagnetic object begins in the calorimeter, and the inner tracker information determines whether the object is a photon - either converted or unconverted - or an electron. Three photon identification methods are presented: a simple cut based method, a log likelihood ratio method and a covariance matrix based method. The shower shape variables based on calorimeter information and track information used in all three methods are described. The efficiencies of the three methods for the single photon and the benchmark $H \rightarrow \gamma\gamma$ signal events, as well as the rejection of the background from jet samples, are presented. The performance of the cut based method on high p_T photons from a graviton decay process $G \rightarrow \gamma\gamma$ is discussed.

1 Introduction

Isolated photons with large transverse momentum, p_T, in the final state are distinguishing signatures for many physics analyses envisaged at the LHC. The Higgs particle has been sought over several decades in many high energy experiments, including those currently running at the Tevatron. It is understood that if the Standard Model Higgs particle exists, and unitarity is not violated, its mass is within the reach of LHC. As described in detail in other parts of this work [1], while the expected cross section times branching ratio of the Higgs particle decaying into the two photon final state is relatively small, given its distinct signature, isolated high p_T photons may play a significant role in discovering the Higgs particle in the low mass region. In addition, very high p_T photons are also signatures of more exotic particles, such as the graviton predicted in Ref. [2], which is expected to have mass larger than 500 GeV. These photons appear as a single, isolated objects with most of their energy deposit in the electromagnetic compartment of the calorimeter. Thus the primary source for background to these photons, namely fake photons, result from jets that fluctuate highly electromagnetic which contain a high fraction of photons from neutral hadron decays, such as $\pi^0 \to \gamma\gamma$.

Since the ATLAS electromagnetic calorimeter [3] is highly segmented with a three-fold granularity in depth and with an $\eta \times \phi$ granularity in the barrel of 0.003×0.1, 0.025×0.025, and 0.05×0.025, respectively, in the front, middle and rear compartments assisted by a pre-sampler in front of the calorimeter, photon identification methods in ATLAS should be much more powerful that those used in past experiments. The experiment also employs elaborate trigger systems that select electrons and photons efficiently, as described in detail in Ref. [4].

This paper presents three ATLAS photon identification methods and their performance for single, isolated photons as well as for photons from physics processes.

2 Data Samples

The $H \to \gamma\gamma$ ($M_H = 120$ GeV) process is used as the primary signal benchmark sample for medium p_T photons and with the pile-up that corresponds to the instantaneous luminosity 10^{33} cm^{-2}s^{-1}. Rejection studies were conducted using a pre-filtered jet sample (described in details in Ref. [5]), containing all relevant hard-scattering QCD processes with $p_T > 15$ GeV. A filter is applied at the generator level, requiring the summed transverse energy of all stable particles (excluding muons and neutrinos) in a region of $\Delta\phi \times \Delta\eta = 0.12 \times 0.12$ to be above 17 GeV. A total number of 3 million events were used in rejection studies. Two additional samples with 150 GeV $< p_T <$ 280 GeV (Jet5) and 280 $< p_T <$ 400 GeV (Jet6) were also employed for high p_T photon rejection studies. Finally, an additional 300,000 event $\gamma +$ *jet* sample has been used for rejection and fake rate studies.

In addition to these signal and background samples, the three identification methods described in this paper were developed using single photon samples - events with no activity except the photon - with full detector simulation in the energy range 10 – 1000 GeV with flat rapidity distributions in $|\eta| < 2.5$. For high p_T photons, graviton samples with masses of 0.5 and 1.0 TeV were employed.

All the samples used in this report were generated using PYTHIA and its fragmentation scheme and were passed through the full detector simulation. Some of the simulations were done with the nominal geometry and material distribution ("ideal") and others with a realistic alignment and additional material added ("misaligned").

In order to maintain the consistency between different studies, the following requirements and definitions are used for efficiencies and rejections.

- Truth match: The reconstructed photons must lie within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.2$ of the true photons in the simulation.

- The reconstructed photons must be within the fiducial volume, pseudo-rapidity $0 < |\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ to avoid the overlap between the barrel and endcap calorimeters.

Using the base samples that satisfy the above requirements, the efficiency is defined as follows:

$$\varepsilon = \frac{N_\gamma^{reco}}{N_\gamma^{truth}} \tag{1}$$

where N_γ^{truth} is the number of true photons in the simulation that satisfy all the requirements above with the true E_T greater than either 25 GeV or 40 GeV and N_γ^{reco} is the number of reconstructed photons that satisfy all the requirements with the true E_T greater than either 25 GeV or 40 GeV and that pass the threshold for one of the three methods.

Similarly, the rejection from the pre-filtered jet sample is computed as follows:

$$R = \frac{N_{jet}}{N_{fake\gamma}} \frac{N_1}{N_2} \frac{1}{\varepsilon_{\gamma-filter}} \tag{2}$$

where N_{jet} is the total number of jets reconstructed in the normalization sample (same generation as the reconstructed sample but without the filter requirements) using particle four-momenta from the generator hadron level within a cone size ΔR = 0.4, and $N_2 (= 400,000)$ is the number of events used in this normalization sample. The values for N_{jet}/N_2 in the fiducial volume of $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$ are 0.226 for jets with $E_T > 25$ GeV and 0.042 for jets with $E_T > 40$ GeV. $N_{fake\gamma}$ is the number of fake photons in the reconstructed (filtered) sample with the candidates that matched to true photons from the hard scatter or from quark bremsstrahlung removed, and $N_1 (= 3,095,900)$ is the number of events analyzed from this sample. Finally, $\varepsilon_{\gamma-filter}$ $(= 0.082)$ is the efficiency of the generator level filter applied to the jet sample.

3 Photon Identification Methods

As discussed in previous sections, three photon identification methods have been developed and are available at present in ATLAS: simple cut-based identification method, Log-Likelihood Ratio based identification method (LLR) and the covariance matrix based identification method (H-matrix). Partial description of basic electromagnetic object reconstruction and their calibration can be found in Ref. [6]

3.1 Characteristics Variables and the Cut-Based Identification Method

In order to separate real photons from fake photons resulting from jets, several discriminating variables are defined using the information both from the calorimeters and the inner tracking system. Cuts on these variables are developed to maintain high photon efficiency even in the presence of pile-up resulting from the overlapping minimum bias events due to high instantaneous luminosity at the LHC. The discriminating variables used in this study are the same as in previous studies [7–11]. Calorimeter information is used to select events containing a high-E_T electromagnetic shower. The fine-grained first compartment allows to reject showers from photons from π^0 decays. Track isolation is used to improve the rejection. Only electromagnetic clusters with $E_T > 20$ GeV are used in this study.

3.1.1 Variables Using Calorimeter Information

In the electromagnetic calorimeter, photons are narrow objects, well contained in the electromagnetic calorimeter, while fake photons induced from jets tend to have a broader profile and can deposit a substantial fraction of their energy in the hadronic calorimeter. Hence, longitudinal and transverse shower shape variables can be used to reject jets.

- **Hadronic leakage :** The hadronic leakage is defined as the ratio of the transverse energy in the first layer of the hadronic calorimeter in a window $\Delta\eta \times \Delta\phi = 0.24 \times 0.24$ to the transverse energy of the cluster in order to avoid boundary effects that could result from using readout cells. Real photons are purely an electromagnetic object, therefore they deposit their energy primarily in the electromagnetic compartment of the calorimeter. Fake photons induced from jets contain hadrons that would penetrate deeper into the calorimeter depositing sizable energy beyond the electromagnetic calorimeter.

- **Variables using the second compartment of the ECAL :** Electromagnetic showers deposit most of their energy in the second layer of the electromagnetic calorimeter. For this reason several variables that measure the shape of the shower are available as follows:

  - The real photons deposit most of their energy in a $\Delta\eta \times \Delta\phi = 3 \times 7$ window (in units of middle cells). The lateral shower shape variables, R_η and R_ϕ, are given by the ratio of the energy reconstructed in 3×7 middle cells to the energy in 7×7 cells and the ratio of the energy reconstructed in 3×3 cells to the energy in 3×7 cells, respectively. Due to the effect of the magnetic field increasing the width of the converted photon contributions in the ϕ direction, R_ϕ is less discriminating than R_η.
  - The lateral width in η is calculated in a window of 3×5 cells using the energy weighted sum over all cells. $w_2 = \sqrt{\frac{\sum(E_c \times \eta_c^2)}{\sum E_c} - \left[\frac{\sum(E_c \times \eta_c)}{\sum E_c}\right]^2}$, where E_c is the energy deposit in each cell, and η_c is the actual η position of the cell represented by the center of the cell in η direction. Therefore, w_2 is given in units of η. A correction is applied as a function of the impact point within the cell to reduce the bias from the finite cell size.

- **Variables using the first compartment of the ECAL :** Cuts applied on the variables in the hadronic calorimeter and the second layer of the electromagnetic calorimeter reject jets that contain high energy hadrons and resulting broad showers. Jets containing single or multiple neutral hadrons such as η and π^0, are the main contribution which can fake photons. The readout of the first layer of the calorimeter uses strips and provides very fine granularity in pseudo-rapidity. Thus, the information from this layer can be used to identify substructures in the showers and distinguish isolated photons from the hard scatter and photons from π^0 decays efficiently. The lateral shower shape in the strips is exploited for $|\eta| < 2.35$ where the strip granularity is sufficiently fine, as long as 0.5% or larger fraction of the total energy is reconstructed in this layer.

  - Since the energy deposit pattern from π^0's is often found to have two maxima due to $\pi^0 \to \gamma\gamma$ decay, showers are studied in a window $\Delta\eta \times \Delta\phi = 0.125 \times 0.2$ around the cell with the highest E_T to look for a second maximum. If more than two maxima are found the second highest maximum is considered. The following two variables are constructed using the information from the identified second maximum:
    - $\Delta E_s = E_{\text{max2}} - E_{\text{min}}$, the difference between the energy associated with the second maximum E_{max2} and the energy reconstructed in the strip with the minimum energy, found in between the first and second maxima, E_{min}.
    - $R_{\text{max2}} = E_{\text{max2}}/(1 + 9 \times 10^{-3} E_T/\text{GeV})$, where E_T is the transverse energy of the cluster in the electromagnetic calorimeter. The value of the second maximal energy deposit is corrected as a function of the transverse energy of the cluster to minimize its sensitivity to fluctuations [9, 10].
  - $F_{\text{side}} = [E(\pm 3) - E(\pm 1)]/E(\pm 1)$, the fraction of the energy deposited outside the shower core of three central strips. The variable $E(\pm n)$ is the energy deposited in $\pm$ n strips around the strip with the highest energy.

- $w_{s3} = \sqrt{\sum E_i \times (i - i_{max})^2 / \sum E_i}$, the shower width over the three strips around the one with the maximal energy deposit. The index i is the strip identification number, i_{max} the identification number of the most energetic strip, and E_i is the energy deposit in strip i. w_{s3} is expressed in units of strip cells and corrected for impact point dependence [9].
- w_{stot}, the shower width over the strips that cover 2.5 cells of the second layer (20 strips in the barrel for instance). It is expressed in units of strip cells.

Figure 1 shows the average values of the calorimeter-based discriminating variables as a function of the absolute value of pseudo-rapidity. Features in the plots can be explained by: upstream material thickness which increases with pseudo-rapidity in the barrel; physical cell size changes in the end-cap to maintain a constant granularity in η-ϕ; and the change in the granularity of the first layer in the end-cap. The dip in the hadronic leakage variable near $|\eta|$=1.1 corresponds to a smaller coverage by the first hadronic layer in this region.

The cut values are tuned separately in six pseudo-rapidity intervals in $|\eta| < 2.37$ to reflect the pseudo-rapidity dependence of these variables. The subdivision is motivated by the varying granularity and material in front of the electromagnetic calorimeter. The quantities calculated using the first compartment can be used only in the regions $|\eta| < 1.37$ and $1.52 < |\eta| < 2.37$ since there are no strips in the crack region or beyond $|\eta| > 2.40$. In addition, up to eight different bins in transverse energy are also used for the cut value adjustment. Figure 2 shows the distributions of the variables in the first η bin and in one energy bin. The dashed vertical lines represent the cut values in this bin. The variables are shown for all reconstructed electromagnetic objects before cuts.

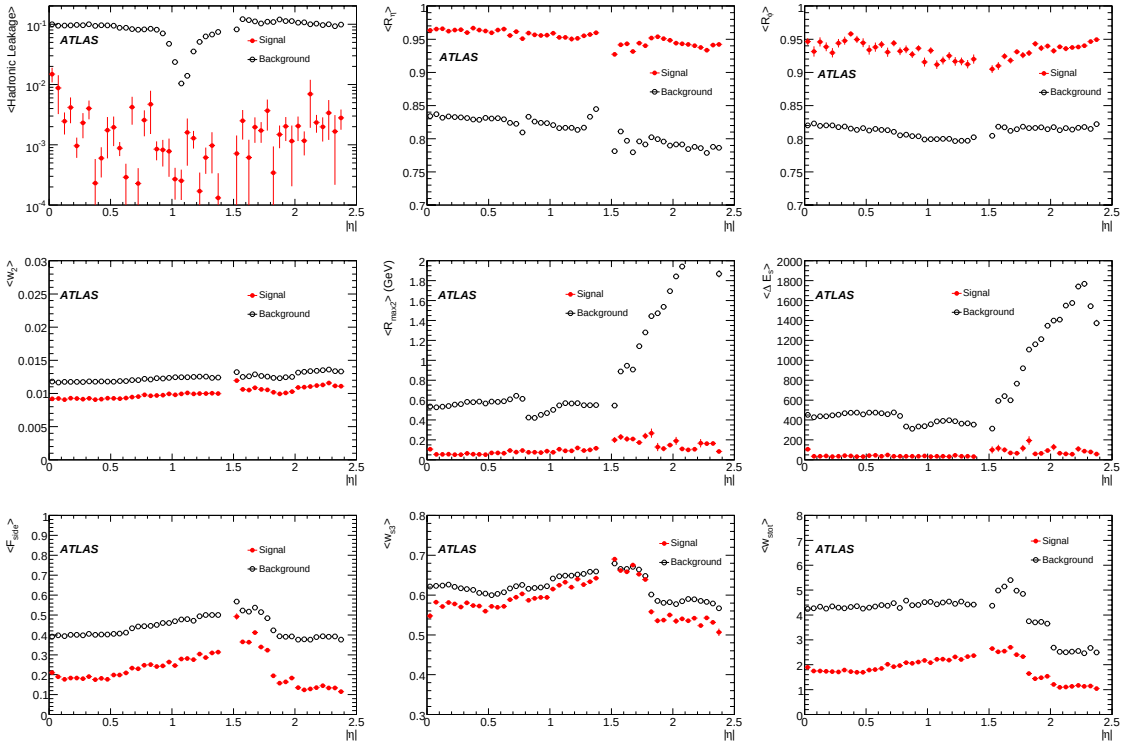


Figure 1: Distributions of the mean of each calorimetric discriminating variable as a function of the pseudo-rapidity $|\eta|$ for true and fake photons (before cuts) with $20 < E_T < 30$ GeV. The samples have been simulated with the geometry under the realistic alignment scenario and additional material.

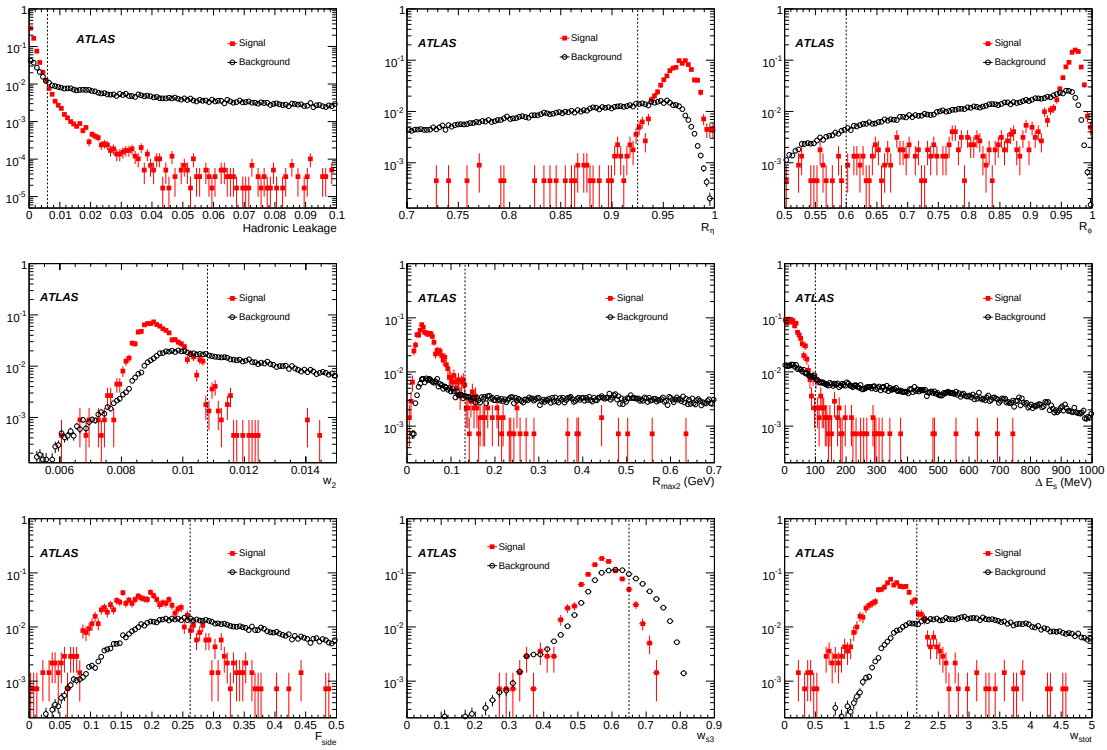


Figure 2: Normalized distributions of the discriminating variable for $|\eta| < 0.7$ for true and fake photons (before cuts) with $20 < E_T < 30$ GeV. The samples have been simulated with the geometry under the realistic alignment scenario.

Figure 3 shows the impact of pile-up and additional material before the calorimeter on the shower shapes for photons from Higgs decays. The impact of the large amount of additional material in the transition region, $1.5 < |\eta| < 1.8$, in the realistic alignment geometry can clearly be seen for two shower shape variables. While pile-up at a luminosity of 10^{33} cm^{-2}s^{-1} does not change the average shower shape significantly as can be seen for the two variables in Figure 3, it is observed that it does increase RMS of the distributions.

At present, the same cuts are applied for converted and unconverted photons. Studies of the γ-π^0 separation, however, have shown that if conversions can be identified efficiently, different cuts can be applied for converted and unconverted photons [12], which could improve rejection by 10-20% while maintaining the same overall photon identification efficiency.

The cuts have been chosen comparing the photons from $H \rightarrow \gamma\gamma$ decays to fake candidates in inclusive jet samples. For this optimization, samples generated with realistic alignment geometry and pile-up have been used. Some improvement in the performances should be possible at higher E_T for further refinement and optimization in some of the variables, such as hadronic leakage. The rejection presented in this paper has been estimated on a sample statistically independent from the one used to tune the cuts.

3.1.2 Track Isolation

After the calorimeter cuts, the contamination of the inclusive signal from charged hadrons is greatly reduced. The remaining background is dominated by low track multiplicity jets containing high-p_T π^0 mesons. In order to further remove fake photons from these jets, the track isolation variable is defined as

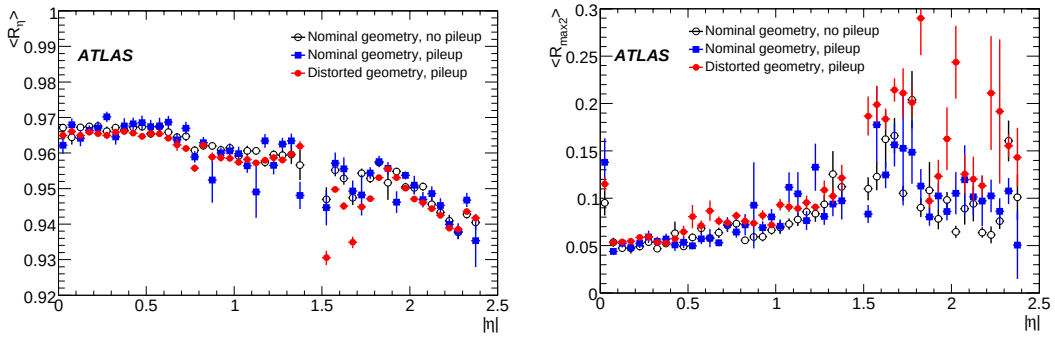


Figure 3: Effect of pile-up and distorted materials on mean values of two shower shape variables for photons from $H \rightarrow \gamma\gamma$ decays. Left: R_η, Right: Energy of the second maximum in the first layer.

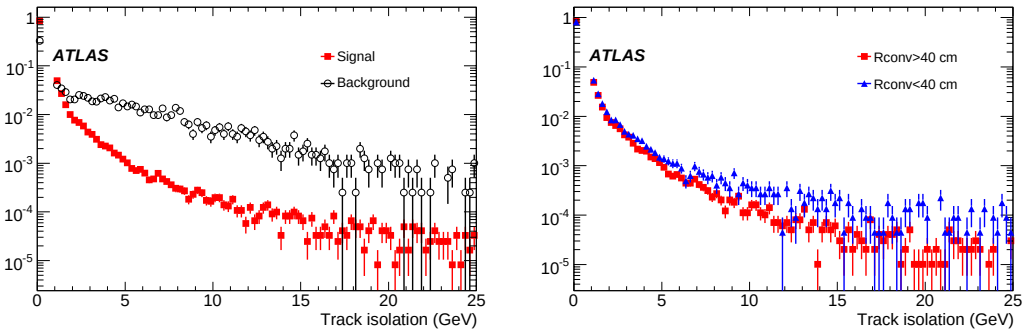


Figure 4: Normalized distribution of the track isolation variable for events passing the calorimeter selection criteria. Left: comparison of true and fake photons. Right: comparison of early conversion (true conversion radius less than 40 cm) and not early conversion for photons from $H \rightarrow \gamma\gamma$ decays.

the sum of the p_T of all tracks with p_T above 1 GeV within $\Delta R < 0.3$, where ΔR is the $\eta - \phi$ distance between the track position at the vertex and the cluster centroid. Track $p_T > 1$ GeV is imposed to minimize the effect of pile-up and underlying events.

Since the tracks from photon conversions should not be included in computing this variable, some additional selections are applied to tracks within $\Delta R < 0.1$ of the cluster centroid. The impact parameter with respect to the beam line must be less than 0.1 mm. The track p_T must not exceed 15 GeV to remove tracks from very asymmetric conversions, must not be part of a reconstructed conversion vertex and must have a hit in the innermost pixel layer.

The plot on the left in Figure 4 shows the distribution of the track isolation variable for true and fake photon candidates, after the calorimeter shower shape cuts. An additional rejection of factor 1.5 to 2 is possible for a relatively small efficiency losses. The plot on the right in this figure shows the track isolation variable for early converted and late converted photons. The difference between the two distributions is rather small, showing that the tracks from conversion have been efficiently removed. At present, a 4 GeV upper cut on the track isolation variable is applied for this method.

3.2 Log Likelihood Ratio Based Photon Identification

In the Log Likelihood Ratio (LLR) based method, the distribution of each of the shower shape variables is normalized to unity to obtain the probability density functions (PDF). The shower shape variables are pseudo-rapidity dependent so they are separated in four regions of η and three bins in p_T for this method. The PDF's are obtained using 1.6 million $\gamma + jet$ events which provided slightly over 100,000 events in each bin. Since the statistics for PDF computation is somewhat low in some kinematic phase space regions, further improvement can be obtained by using the KEYS [13] tool to smooth PDF's to compensate for the low statistics. Once the PDF's are established, the Log-likelihood Ratio parameter is defined as:

$$LLR = \sum_{i=1}^{n} \ln\left(L_{si}/L_{bi}\right), \tag{3}$$

where L_{si} and L_{bi} are PDF's of the i^{th} shower shape variable for the photon and the jet, respectively.

The shower shape variables used for the LLR method were the same as those used for the cut-based method described previously. Track isolation was also included as a discriminating variable in equation 3. Figure 5 shows the LLR parameter distribution for true photons and for jets. The LLR cut can be tuned over η and P_T to obtain an optimal separation between photons and jets.

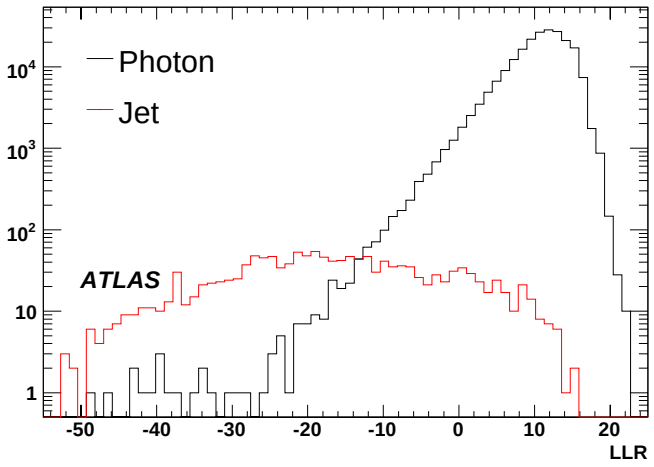


Figure 5: *LLR* cut parameter distributions for true photons (solid histogram) and for jets (dashed histogram)

3.3 Covariance Matrix Based Photon Identification Method

The shower shape variables associated with a photon shower in the calorimeter are correlated. The covariance matrix (H-matrix) technique takes advantage of these correlations. The technique was employed successfully in the DØ experiment at the Tevatron and was used to identify electrons [14].

The ten photon shower shape variables used in the ATLAS H-matrix method are as follows:

- Five longitudinal shower shape variables: fraction of energy deposited in pre-sampler layer; fractions of energy deposited in sampling layers 1, 2 and 3 separately; and the hadronic leakage, the energy leakage into the first layer of the hadronic calorimeter.
- Five transverse shower shape variables: the ratio of the energy in 3×3 cells to the energy in 7×7 in the second sampling layer of the electromagnetic calorimeter; w_{rms3}, the corrected width in 3

strips in sampling layer 1; w_2, the corrected width in a 3×5 window in sampling layer 2; the energy outside of the shower core; R_ϕ, the ratio of energy in a 3×3 to a 3×7 window around the cluster centroid.

Using the above variables, a covariance matrix, M, is constructed as follows:

$$M_{ij} = \frac{1}{N}\Sigma_{n=1}^{N}(y_i^{(n)} - \overline{y}_i)(y_j^{(n)} - \overline{y}_j), \tag{4}$$

where indices i and j run over the ten variables, N is the total number of photons used in the training sample, y_j^n is the j^{th} variable for the n^{th} photon candidate, and $\overline{y}_j$ is the mean value of y_j variable for the control sample electrons/photons. These matrix elements are constructed for each η bin and parameterized for energy dependences. The photon likeness of an object is then measured by the value of the χ^2, defined as follows:

$$\chi^2 = \Sigma_{i,j=1}^{dim}(y_i^{(m)} - \overline{y_i})H_{ij}(y_j^{(m)} - \overline{y_j}) \tag{5}$$

where $H \equiv M^{-1}$, the inverse of the covariance matrix, and the indices i and j run from 1 to the total number of variables (ten) which is the same as the dimension of the matrix, dim.

The mean value of the χ^2 is close to the number of dimensions for a photon shower. The shapes of the distributions of the selected shower shape variables depend on the η and the energy of the incident photon. These effects are taken into account in the construction of the H-matrix using single photon samples of energies 10 – 1000 GeV generated flat in $|\eta|$ and parameterizing each of the covariance terms in the matrix M of Eq. 4 as a function of the photon energy. The parameterization as a function of photon energy is obtained in each of the 12 η bins. The discrimination power of the H-matrix between real photons and jets is well illustrated in Figure 6, where the χ^2 distribution of the H-matrix for the jet sample is contrasted to that obtained from photons from $H \rightarrow \gamma\gamma$ decays.

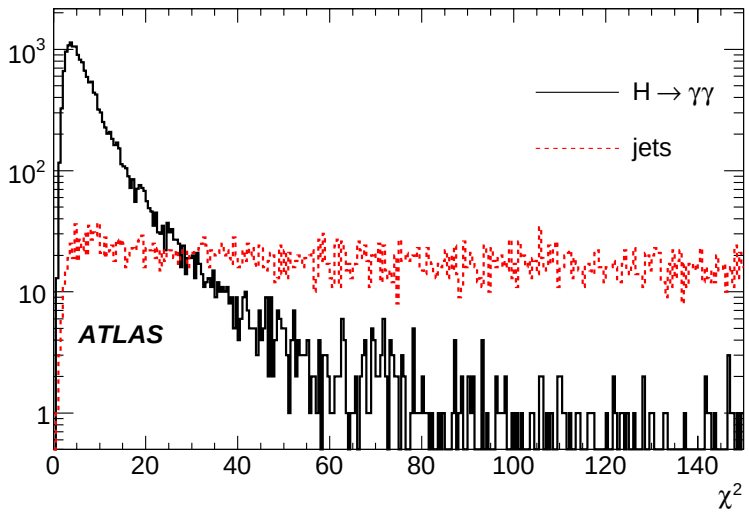


Figure 6: The distributions of H-matrix χ^2 for inclusive jet samples (dashed histogram) and for the individual photon from $H \rightarrow \gamma\gamma$ sample (solid histogram).

Since the H-matrix implementation at this time does not includethe same variables as the other two methods, its performance is currently not directly comparable. Consequently, we do not report the performance here, although the method is decribed for completeness.

4 Photon Identification Performance in Medium p_T Photons

This section describes the performance (efficiencies and rejections) of the cut-based method and the Log Likelihood Ratio method on medium p_T photons, in particular the photons from the $H \rightarrow \gamma\gamma$ decays and the jet background samples described in section 2.

4.1 Performance of the Cut-Based Method

In the performance studies presented in this section all reconstructed electromagnetic objects, including both electron and photon candidates are considered. The efficiency as defined in section 2 includes both reconstruction efficiency and the efficiency of the identification cuts.

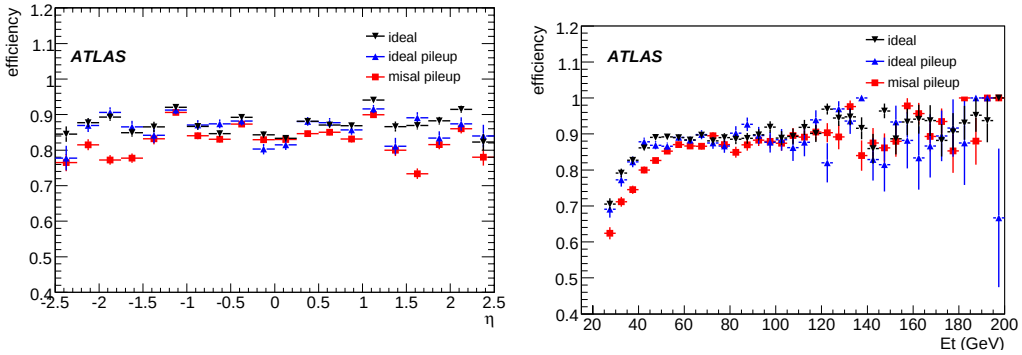


Figure 7: Efficiency of the calorimeter cuts as a function of pseudo-rapidity (left) and the transverse momentum (right) of the true photons from three different samples.

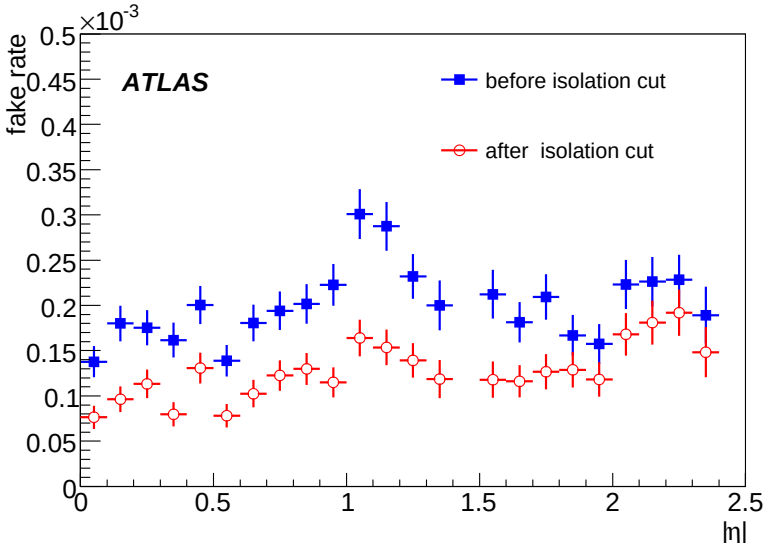


Figure 8: Fake rate as a function of pseudo-rapidity, in the filtered jet sample

Figure 7 shows the efficiency of the calorimeter cuts for photons with E_T >25 GeV as a function of the pseudo-rapidity (left) and the transverse momentum (right) for three different samples of Higgs

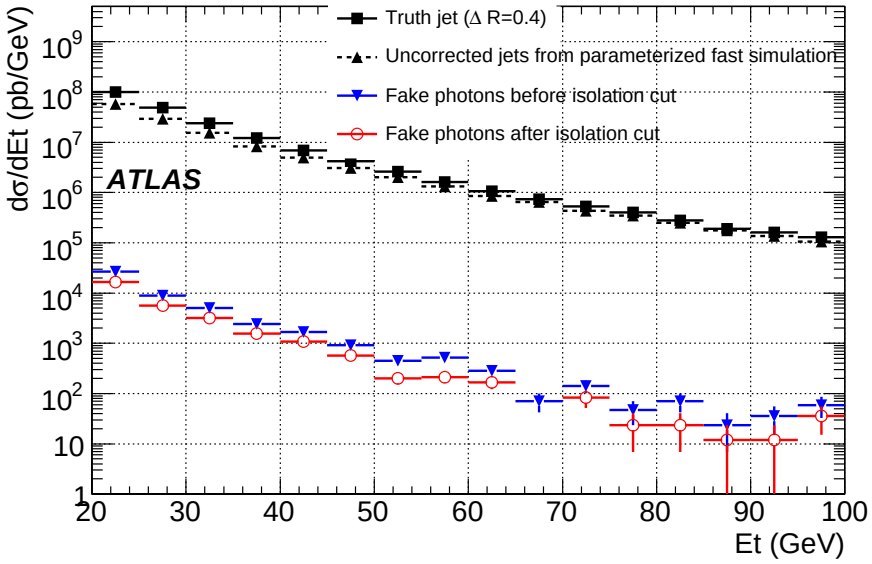


Figure 9: E_T spectra from the inclusive jet sample, for the generated jets (solid square for full simulation and solid triangle for uncorrected jets from parameterized fast simulation) and the fake photon candidates before (upside-down solid triangle) and after (open circle) the track isolation cut. The normalization is that predicted by PYTHIA.

decay: the nominal geometry without pile-up, the nominal geometry with the pile-up at 10^{33} cm^{-2}s^{-1} and the realistic alignment geometry with additional material and with the pile-up at 10^{33} cm^{-2}s^{-1}. The average efficiency of the calorimeter cuts and that of the track isolation cuts are summarized in Table 1.

Sample	ε(calorimeter cuts)	ε(track isolation cut)
Nominal geometry, no pile-up	87.6±0.2%	99.0±0.1%
Nominal geometry, pile-up	86.6±0.5%	98.0±0.2%
Distorted geometry, pile-up	83.6±0.2%	98.1±0.1%

Table 1: Efficiency for photons from $H \rightarrow \gamma\gamma$ decays for three different simulation choices.

The rejection from the pre-filtered jet sample is computed using the equation 2. The rejection is computed separately for all jets, for quark-initiated jets and for gluon-initiated jets. The quark or gluon initiation is defined using the type of the highest E_T parton from the PYTHIA record inside the cone $\Delta R = 0.4$ around the reconstructed jet object. The rejection values are summarized in Table 2 for the three categories of jets. A small fraction ($\approx$ 1-2%) of jet objects are not classified, so the total number of "quarks" and "gluons" is slightly smaller than "All". A cut $E_T > 25$ GeV is applied to both reconstructed photons and jets. Table 3 shows the same computation but for $E_T > 40$ GeV.

Figure 8 shows the fake rate, defined as the inverse of the rejection, as a function of pseudo-rapidity for all jets with E_T greater than 25 GeV. There is a slight increase of fake rate as a function of pseudo-rapidity due to the increase in material in front of the calorimeter, which implies somewhat looser cuts to preserve the consistent efficiency. Some additional increase near $|\eta| = 1.1$ is also visible probably coming from the reduced energy in the first layer of the hadronic calorimeter as pointed out previously. This effect, however, gives less than 10% increase in the overall fake rate.

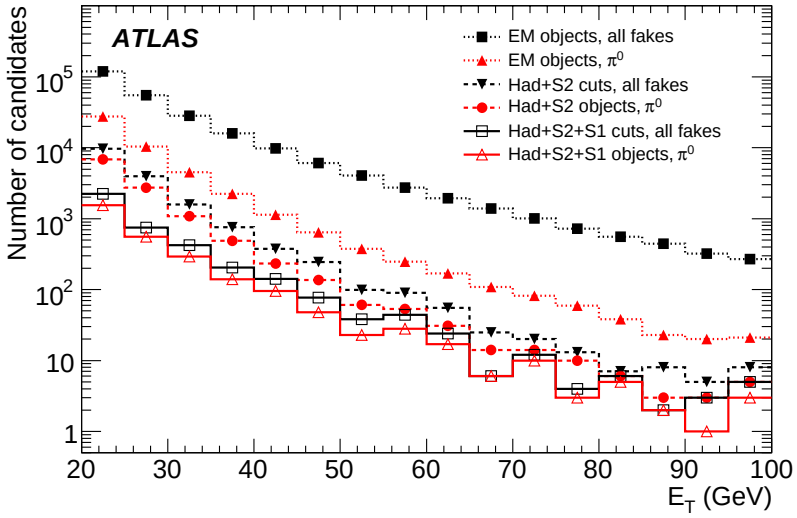


Figure 10: E_T distribution of fake photon candidates in jets after different level of cuts. The contribution from "single" π^0 is also shown

	All	quark-jet	gluon-jet
N(jet)/N(generated events)	0.23	0.056	0.177
	before isolation cut		
N(fake)/N(filtered events)	$(5.43\pm0.13).10^{-4}$	$(3.87\pm0.11).10^{-4}$	$(1.44\pm0.07).10^{-4}$
Rejection	5070±120	1770±50	15000±700
	after isolation cut		
N(fake)/N(filtered events)	$(3.38\pm0.10).10^{-4}$	$(2.47\pm0.08).10^{-4}$	$(0.78\pm0.49).10^{-4}$
Rejection	8160± 250	2760±100	27500±2000

Table 2: Rejection (Equation 2) measured in the inclusive jet sample for $E_T > 25$ GeV

Figure 9 shows the E_T distribution of the jets and of the fake photon candidates before and after the track isolation cut. This figure also shows that the rejection at 25 GeV is ≈ 30% lower if the normalization is based on the uncorrected parameterized jets from the fast simulation, as was done in Ref. [11].

Figure 10 shows the π^0 content of the fake photon candidates at three different cut levels; all reconstructed electromagnetic objects, after the cut on the hadronic leakage and the second layer shower shape variables (Had+S2) and after all the cuts (Had+S1+S2). A fake photon is defined as coming from a π^0 if the energy of the leading π^0 in the cone of 0.2 around the cluster centroid is more than 80% of the reconstructed cluster energy. The figure shows already after the second layer shower shape cuts, the dominant background contribution comes from π^0 as expected. After all cuts, the fraction of π^0 is ≈ 70% of the remaining fake photon candidates.

Figure 11 shows the rejection of the cuts on the first layer variables for candidates from single π^0's passing the cuts on the hadronic leakage and the second layer shower shape variables. As expected, the rejection power against these isolated π^0's decreases with energy, as the opening angle between the two photons from π^0 decays become smaller. The rejection is also better in the central part in the barrel as there is less material than in the higher η part of the barrel, and also opening angle is larger than in the

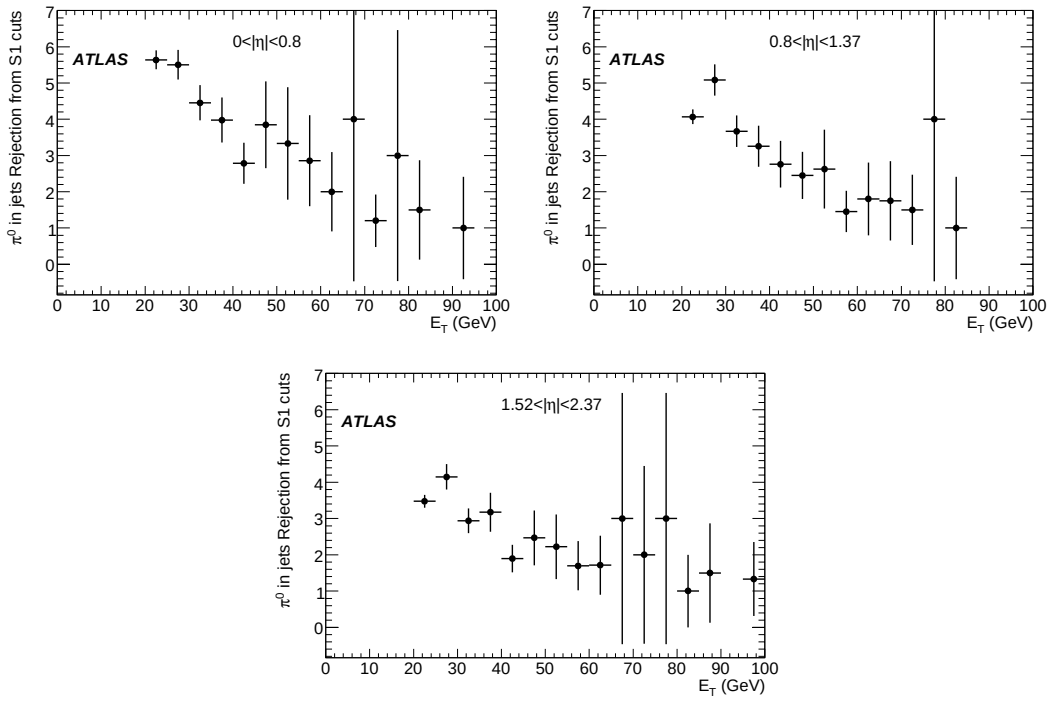


Figure 11: Rejection of the first layer cuts against fakes coming from "single" π^0 in jet sample. as a function of the transverse energy, for three different pseudo-rapidity regions

end-cap for the same p_T . As a cross-check, Figure 12 shows the efficiency of the calorimeter cuts for real single photons and single π^0 of E_T =40 GeV, as a function of pseudo-rapidity. Again, the rejection is slightly higher than 3 in the central part of the barrel calorimeter and is in reasonable agreement with findings from previous studies [15].

The rejections measured in these studies have to be taken with care as they rely strongly on the modelling of the fragmentation tail in PYTHIA and the details of the simulation of the detector response. A discussion of the first effect can be found in Ref. [8] from which one would expect an uncertainty of $50 \sim 100\%$, and where the uncertainty is larger for gluon initiated jets. In addition, a recent investigation on the differences in fragmentation algorithms in PYTHIA and HERWIG shows appreciable differences in π^0 production rates. Some differences in rejection are anticipated if the momentum distributions of

	All	quark-jet	gluon-jet
N(jet)/N(generated events)	0.042	0.011	0.034
	before isolation cut		
N(fake)/N(filtered events)	$(1.16\pm0.06).10^{-4}$	$(8.3\pm0.5).10^{-5}$	$(2.8\pm0.3).10^{-5}$
Rejection	4400±230	1610±100	15000±1600
	after isolation cut		
N(fake)/N(filtered events)	$(6.4\pm0.4).10^{-5}$	$(4.6\pm0.5).10^{-5}$	$(1.5\pm0.2).10^{-5}$
Rejection	7800±540	2900±240	28000±4000

Table 3: Rejection (Equation 2) measured in the inclusive jet sample for $E_T > 40$ GeV

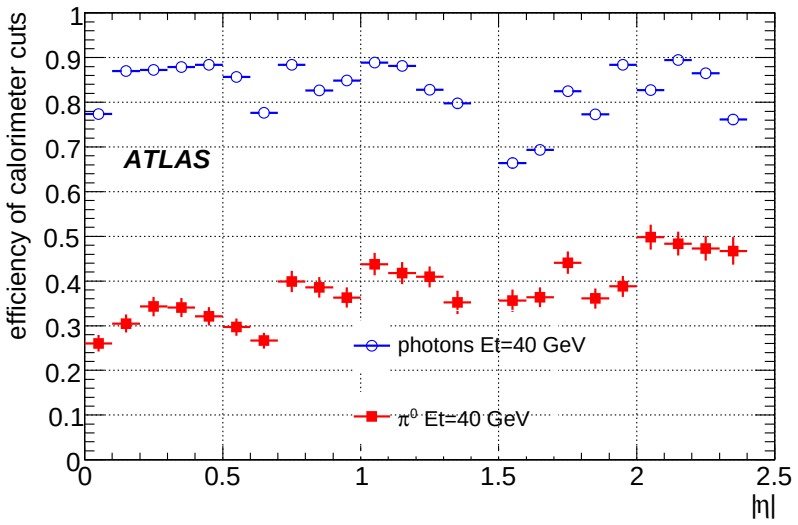


Figure 12: Efficiency of calorimeter cuts vs pseudo-rapidity for 40 GeV E_T single photons and π^0 (distorted geometry without pile-up

the π^0's from the two fragmentation algorithms differ.

4.2 Performance of the Log Likelihood Ratio Method

The efficiency for the Log Likelihood Ratio method is computed for individual photons from the $H \rightarrow \gamma\gamma$ events generated with the nominal geometry. Figure 13 shows the photon efficiency as a function of p_T (left) and η (right) for LLR cut values set at 8, 9 and 10. The overall efficiencies for LLR cuts at 8, 9 and 10 are summarized in Table 4. Jet rejection (left) and photon identification efficiency (right) are shown in Figure 14 as a function of LLR cut parameter values for three different jet p_T regions that corresponds to the three mean jet p_T values indicated.

	$E_T > 25$ GeV			$E_T > 40$ GeV		
LLR cut	$LLR > 8$	$LLR > 9$	$LLR > 10$	$LLR > 8$	$LLR > 9$	$LLR > 10$
Efficiency(%)	87.6 ± 0.3	84.3 ± 0.2	80.0 ± 0.2	86.4 ± 0.3	83.2 ± 0.2	79.0 ± 0.2
Rej.($\gamma +$ *jet*)	1660 ± 170	2190 ± 260	2930 ± 390	1690 ± 140	2170 ± 210	2650 ± 280
Rej. (di-jet)	6820 ± 440	8930 ± 650	12430 ± 1070	6780 ± 1000	7800 ± 1230	11550 ± 2220

Table 4: Overall photon efficiencies and jet rejections with different LLR cut values.

Figure 13 shows the p_T dependence of the photon efficiency. A looser cut on low p_T photon seems to be beneficial in order to retain flat photon efficiency as a function of p_T. Furthermore, it might also be useful to parametrize LLR cut values as a function of photon p_T for further optimization. The jet rejection is also p_T dependent as shown in the plot on the left in Figure 14. A harder cut on LLR for varying jet p_T can help to keep the rejection constant as a function of p_T.

The rejection for jets from $\gamma +$ *jet* and di-jet samples are shown in the fourth and fifth rows in Table 4. The cuts on photon and jet p_T are 25 GeV and 40 GeV, respectively. The rejection for jets from di-jet samples is significantly higher than that on $\gamma +$ *jet* samples. This is largely due to the fact that the jets in

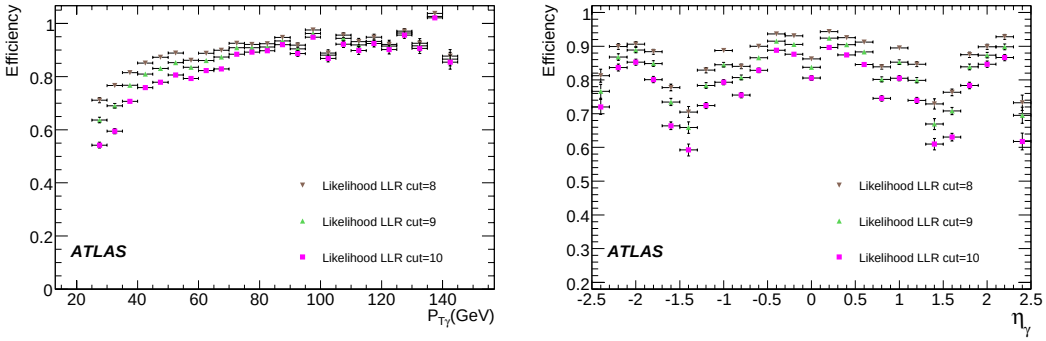


Figure 13: Photon efficiency vs p_T and efficiency vs eta with different LLR cuts. The photon are from $H \to \gamma\gamma$ sample with nominal geometry.

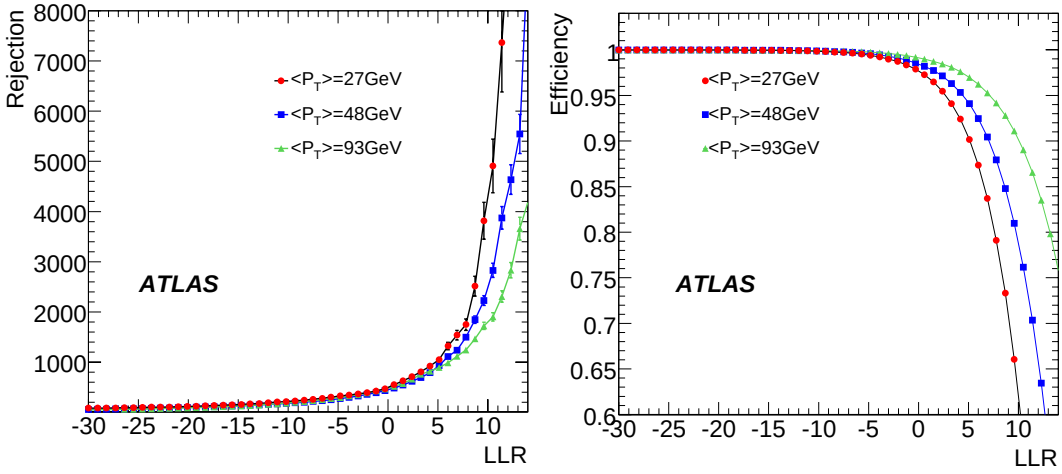


Figure 14: Jet rejection (left) and photon efficiency (right) as a function of LLR cut parameter values.

$\gamma+$ *jet* events are dominated by quark initiated jets while those in di-jet events are enriched with gluon initiated jets.

5 High p_T Photon Performance

Searches for particles of very high mass decaying to photons, such as the Randall-Sundrum graviton, G, decaying via $G \to \gamma\gamma$ [2], require excellent detector and particle identification performance in a kinematic region very different than the benchmark $H \to \gamma\gamma$ process. The p_T-dependent effect caused by differences in kinematics can complicate high mass graviton searches because they modify the shape of background distributions as a function of the two photon invariant mass, $M_{\gamma\gamma}$.

The performance of the cut-based identification method for high p_T photons has been investigated. Studies of the shower characteristics of the photons in the $H \to \gamma\gamma$ and $G \to \gamma\gamma$ ($M_G = 500$ GeV) processes found only minor differences in most of the shower shape variables. In the absence of the track isolation cut, the photon efficiency as a function of p_T in both the barrel and end-cap calorimeters is approximately constant above $p_T = 50$ GeV. The barrel and end-cap photon electromagnetic reconstruc-

tion efficiencies, before applying any identification or isolation cuts, are found to be within 10% of one another for photons from graviton decays. After applying photon identification but no isolation requirements, the efficiencies are 0.829 ± 0.004 in the barrel and 0.639 ± 0.010 in the end-cap calorimeters for $p_T^\gamma > 100$ GeV and $M_G > 500$ GeV.

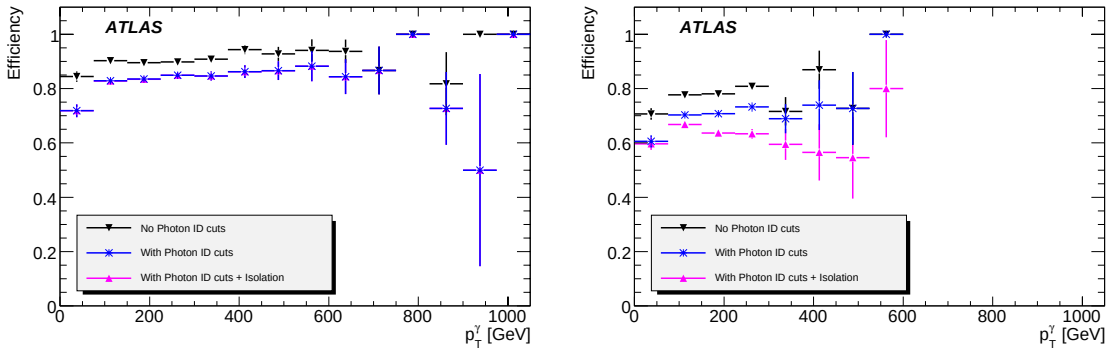


Figure 15: Photon efficiency in the 500 GeV graviton sample as a function of p_T for barrel (left) and end-cap (right) calorimeters.

An isolation variable based on the calorimeter energy in a cone of size $\Delta R = 0.45$ around the cluster centroid was studied. The cut on the calorimeter isolation was observed to produce roughly constant efficiency as a function of p_T, using the Toolkit for Multivariate Analysis (TMVA). A linearly p_T-dependent selection cut was determined for barrel and end-cap photons independently. The efficiencies of these p_T-dependent cuts for barrel and end-cap calorimeters are shown in Figure 15 for photons from 500 GeV graviton decays. As can be seen in the figures, these p_T dependent isolation cuts show about 0.1% reduction in efficiency for photons over the entire p_T range.

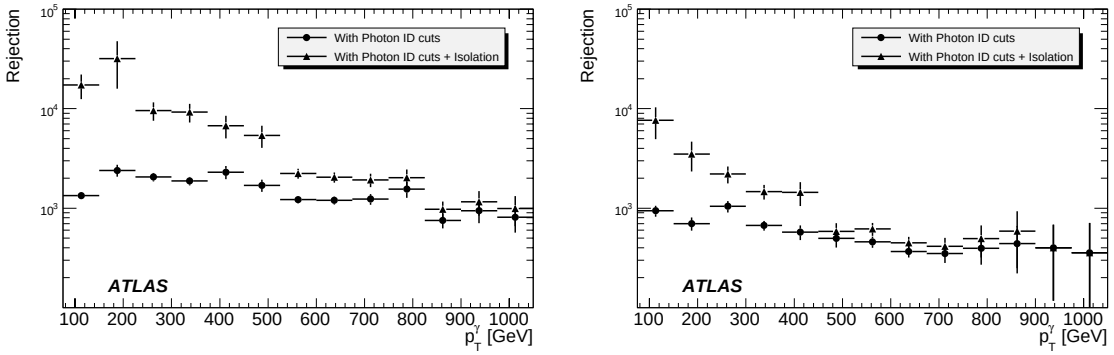


Figure 16: Fake photon rejection as a function of p_T of the reconstructed photon object for high p_T binned di-jet samples in the barrel (left) and end-cap (right) calorimeters.

The Jet5 and Jet6 high p_T jet samples discussed in Section 2 were used for rejection studies. Figure 16 shows the p_T dependence of jet rejection with and without the calorimeter energy isolation cuts. It can be seen that while the efficiency loss is small, employing the isolation cut increases rejection across the full p_T range. In particular, the region below $p_T = 500$ GeV shows a factor 5 - 10 increase in rejection. Table 5 provides the measured rejection in the barrel and end-cap calorimeters using these two di-jet samples.

Table 5: Jet rejections obtained using two binned di-jet samples, with the cut-based photon identification without (left) and with (right) the track isolation cut

Region	Rejection($x10^3$)	Rejection($x10^3$)
barrel	1.54 ± 0.05	3.85 ± 0.03
endcap	0.64 ± 0.03	1.14 ± 0.08
total	1.25 ± 0.03	2.78 ± 0.08

6 Comparison of the Photon Identification Methods

Figure 17 shows the rejection and efficiency curves for two of the three currently available photon identification methods - the cut-based method and the Log Likelihood Ratio method - for $\gamma +$ *jet* generated in specific photon momentum bins and the benchmark $H \rightarrow \gamma\gamma$ samples. Similarly Figure 18 shows the rejection and efficiency curves for these methods for di-jet and $H \rightarrow \gamma\gamma$ samples. Tables 6 and 7 provide numerical comparisons of fake photon rejections for the methods keeping the photon identification efficiencies similar.

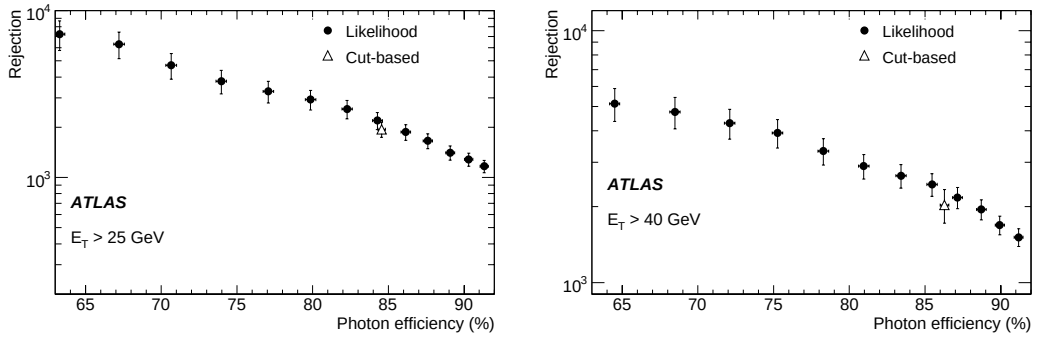


Figure 17: Jet rejection vs photon efficiency for binned $\gamma +$ *jet* and $H \rightarrow \gamma\gamma$ benchmark samples for $p_T^{\gamma}, p_T^{jet} > 25$ GeV(left) and $P_T^{\gamma}, p_T^{jet} > 40$ GeV(right).

	$E_T > 25$ *GeV*		$E_T > 40$ *GeV*	
	LLR	Cut-based	LLR	Cut-based
Efficiency(%)	84.3 ± 0.2	84.5 ± 0.2	87.1 ± 0.2	86.3 ± 0.2
Rejection	2190 ± 250	1940 ± 230	2170 ± 210	2030 ± 190

Table 6: Comparison of jet rejection vs photon efficiency with the cut-based method and the LLR method for $\gamma +$ *jet* sample

The $\gamma +$ *jet* events, whose jets are dominated by quark initiated jets are the largest background to $H \rightarrow \gamma\gamma$ process. It is apparent from Figures 17 and 18 that the methods demonstrate significantly reduced rejections for jets from the $\gamma +$ *jet* samples than for those from di-jet samples whose jets are predominantly from gluons. As discussed in previous sections, this difference in rejection can be attributed to the fragmentation differences between the quark and gluon-initiated jets.

Finally, Figures 17 and 18 also illustrate that, for equal efficiencies, the Log-Likelihood Ratio method and the cut-based method perform comparably in rejecting jets.

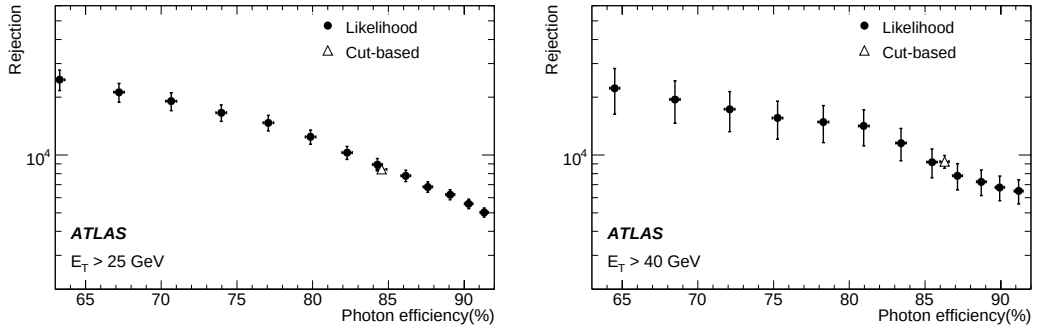


Figure 18: Jet rejection vs photon efficiency of the three methods for filtered di-jet and $H \rightarrow \gamma\gamma$ benchmark samples for $p_T^{\gamma}, p_T^{jet} > 25$ GeV(left) and $P_T^{\gamma}, p_T^{jet} > 40$ GeV(right).

	$E_T > 25\ GeV$		$E_T > 40\ GeV$	
	LLR	cut-based	LLR	cut-based
Efficiency(%)	84.3 ± 0.2	84.6 ± 0.2	85.5 ± 0.2	86.3 ± 0.2
Rejection	8930 ± 650	8240 ± 270	9170 ± 1570	9240 ± 710

Table 7: Comparison of jet rejection vs photon efficiency with the cut-based method and the LLR method for di-jet samples

7 Conclusions

This report presents the three photon identification methods developed in ATLAS - the cut-based method, the Log-Likelihood Ratio (LLR) based method and the covariance matrix based method (H-matrix). The efficiencies and fake photon rejections of the first two methods have been measured using fully simulated $H \rightarrow \gamma\gamma$ ($M_H = 120$ GeV), $\gamma +$ *jet* and filtered electromagnetic di-jet samples. The cut-based and LLR methods show similar rejection factors at equal efficiencies. The strength of the continuous methods such as the LLR and H-matrix is the ability to vary the cuts on LLR or χ^2 values to optimize for specific physics analyses. The performance of the cut-based method for extreme high p_T photons from Randall-Sundrum graviton samples has also been studied and, while the cut selection was optimised at low p_T compared to the signal in the graviton sample the efficiency remains high. While the currently available photon identification methods perform very well in rejecting background, with high efficiency in retaining photons, it is of critical importance to study the performance of the methods with beam collision data.

References

[1] ATLAS Collaboration, "Prospects for the Discovery of the Standard Model Higgs Boson Using the $H \rightarrow \gamma\gamma$ Decay with the ATLAS Detector," this volume

[2] L. Randall and R.Sundrum, Phys. Rev. Lett. 83, 3370 (1999).

[3] ATLAS Collaboration, "The ATLAS Experiment at CERN Large Hadron Collider," accepted for publication by JINST (2008).

[4] ATLAS Collaboration, "Data Preparation for the ATLAS High-Level Trigger Calorimeter Algorithms," this volume.

[5] ATLAS Collaboration, "Expected Performance for the Reconstruction and Identification of Electrons," this volume.

[6] ATLAS Collaboration, "Electromagnetic Calorimeter Calibration and Performance," this volume.

[7] G.Unal and L.Fayard, "Photon identification in $\gamma - jet$ events with Rome layout simulation and background to $H \rightarrow \gamma\gamma$," ATL-PHYS-PUB-2006-025 (2006).

[8] M.Escalier *et al*, "Photon/jet separation with DC1 data," ATL-PHYS-PUB-2005-018 (2005).

[9] M. Wielers, "Photon identification with the Atlas detector," ATLAS-PHYS-99-016 (1999).

[10] M. Wielers, "Isolation of photon," ATLAS-PHYS-2002-004 (2002).

[11] ATLAS collaboration, "TDR ATLAS detector and physics performance, vol. 1," CERN/LHC 99-14, ATLAS-TDR-14 (1999).

[12] ATLAS Collaboration, "Photon Conversions in ATLAS," this volume.

[13] K. Cranmer, "Kernel estimation in high-energy physics," Computer Physics Communications, Volume 136, Number 3, 15 May 2001 , 198-207(10) (2001).

[14] V. Amazov *et al.*, D$\emptyset$ collaboration, "$t\bar{t}$ Production Cross-section in $p\overline{p}$ Collisions at $\sqrt{s} = 1.8$ TeV," Phys. Rev. **D67**, 012004 (2003).

[15] J.Colas *et al*, "Position resolution and particle identification with the ATLAS EM calorimeter," Nucl.Instrum.Meth.A550:96-115 (2005).