Electron and Photon Reconstruction and Identification with the ATLAS Detector

Ashfaq Ahmad
On behalf of the ATLAS collaboration

Stony Brook University

11th ICATPP, Villa Olmo 5-9 October 2009, Como Italy
Electrons and photons @ LHC

- At the LHC electrons and photons are expected to be produced in many physics channels of interest
  - And within a large energy range, typically from few GeV to 5 TeV

- Some important sources of Electrons/photons:
  - Electrons from J/Ψ, Y, W and Z bosons decays
  - Non-isolated electrons from heavy flavor decays
    - Used for performance studies and detector calibration, also above is background to new physics
  - Electrons from higgs decay, e.g. H→ZZ*→4e
  - BSM: electrons form Z' decay, SUSY and extra dimensions
  - Isolated photons from H→γγ and G→γγ

- But a lot of background; e.g. Electron to QCD jet ratio~10^{-5} with p_{T} from 20-50 GeV

- Requires excellent electron/photons reconstruction and identification capability
Object Reconstruction

Will discuss this part
Inner Detector (Tracker)

- **Pixel**:  
  - barrel - 3 layers, 67M pixels  
  - end-cap – 3 layers 6.6M pixels

- **SCT (Semi Conductor Tracker)**:  
  - barrel - 8 layers, ~2M channels  
  - end-cap – 9 layers, ~2M channels

- **TRT (Transition Radiation Tracker)**:  
  - barrel - 73 layers, ~53k channels  
  - end-cap - 160 layers, ~123k channels

Inside 2 Tesla Solenoid  
- Precision tracking for \( p_T > 0.5 \text{ GeV} \) inside \(|\eta| < 2.5\)
- Tracker radiation length \(\sim 0.5-1.5\)
- Relative transverse momentum resolution \(\sim 1.5-4\%\) for 20-100GeV tracks in the barrel
Liquid Argon (LAr) Calorimeter

Good energy resolution:

$$\frac{\sigma_E}{E} = \frac{a}{E} \oplus \frac{b}{\sqrt{E}} \oplus c_{\text{tot}}$$

With $b =$ sampling term $\sim 10\% / \sqrt{E}$

$a =$ noise term $\sim 0.3$GeV and $c_{\text{tot}} \sim 0.7\%$

<table>
<thead>
<tr>
<th>Layer</th>
<th>$\Delta \eta \times \Delta \phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>preSampler</td>
<td>0.025 x 0.1</td>
</tr>
<tr>
<td>Strip</td>
<td>0.003 x 0.1</td>
</tr>
<tr>
<td>middle</td>
<td>0.025 x 0.025</td>
</tr>
<tr>
<td>back</td>
<td>0.05 x 0.025</td>
</tr>
</tbody>
</table>

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

Coverage, $|\eta| < 3.2$
Electron/Photons reconstruction

- **Algorithms:**
  - **Calorimeter seeded (Standard egamma):**
    - Starts from reconstructed **clusters** in the EM calorimeter, match cluster with tracks in the Inner detector (tracker) within a window $\Delta \eta \times \Delta \phi = 0.05 \times 0.1$ and $E(\text{cluster})/P < 10$
    - To determine if particle is electron (track match), photon (no track) or converted photon (with associated conversion)
    - Early classification allows to apply different corrections to electrons and photons
    - Build identification variables based on EM calorimeters and inner detector
    - Used for high $p_T$ isolated electrons/photons
  
  - **Track seeded algorithm (Soft-electron reconstruction):**
    - Starts from good-quality tracks in the tracker, matching with energy deposition in calorimeter
    - Build identification variables as above
    - Used mainly for low $p_T$ electrons up to few GeV e.g. electrons from $J/\Psi$ and b and c quarks
Calibration/correction steps

Cell-Level Calibration

Cluster level Calibration

In situ Calibration

Electron reconstruction

MC based calibration

Intercalibration

Calibration loop

Calculate initial cluster position and energy

Correct position measurements

Correct energy for lateral and longitudinal shower shapes

Correct for \( \eta, \phi \) energy modulations

Cluster corrections steps

How to combine energy deposited in each layer
Clustering algorithm

- Electrons/photons deposit energy in many calorimeter cells
- Clustering algorithms group cells together and sum the total deposited energy within each cluster
- Energies are calibrated to account for energy deposited outside the cluster and dead material

- Sliding-Window algorithm:
  - Build a pre-cluster using a fixed size 5x5 window (seed) of cells
  - The 5x5 window is moved over the tower grid
  - Position of the window is adjusted so to contain a local maximum in energy
  - A threshold of 3 GeV on transverse energy is applied
  - Build clusters of different sizes based on particle type and calorimeter region
  - Apply cluster calibrations
Cluster calibration: energy reconstruction

- Combine energy deposits in each layer and the presampler
- Compute corrections by using special simulations (Calibration Hits) where energy deposited in all material (active + inactive) is recorded

- Energy depositions in the inactive material can be correlated with the measurable quantities
- Corrections are derived for electrons and photons separately

\[ X = \sum_{i=1}^{n} E_i^{\text{long}} \times X_i \]

- \( E^{\text{rec}} = E^{\text{acc}} \cdot E^{\text{ps}} \cdot S^{\text{acc}}(X, \eta) \cdot \left( \sum_{i=13}^{i} E_i^{\text{long}} \right) \cdot C_{\text{out}}(X, \eta) \cdot f_{\text{leak}}(X, \eta) \)

- \( X = \) Shower depth
- \( X_i = \) long. depth of layer “i”
- \( E_i = \) energy deposit in layer “i”
- \( S_{\text{acc}}(X, \eta) = \) calib. factor

ICATPP, Como Oct. 5-9, 2009
Performance (linearity and resolution)

- At low $|\eta|$ resolution similar for electron and photon
- At large $|\eta|$ resolution worse for electron due to more material
- Sampling term for electron goes from $\sim 9\%$ at low $|\eta|$ to $\sim 21\%$ at high $|\eta$.
- For photons maximum sampling term $\sim 12\%$
- Constant term $< 0.6\%$
- Linearity within $0.5\%$ for both electron and photons for energies from 25-500 GeV

Energy resolution

\[ \frac{\sigma_E}{E} = \frac{a}{E} + \frac{b}{\sqrt{E}} + c_{\text{tot}} \]

- $|\eta| = 0.3$
- Electron: $b = 9.3\% \sqrt{E}$, $c = 0.5610\%$
- Photon: $b = 8.6\% \sqrt{E}$, $c = 0.6097\%$
- $|\eta| = 1.65$
- Electron: $b = 18.4\% \sqrt{E}$, $c = 0.43\%$
- Photon: $b = 12.2\% \sqrt{E}$, $c = 0.59\%$
Efficient electron/photon identification methods are needed to reject huge QCD background.

Several methods developed in ATLAS namely;

- Cut-based identification (Standard)
- Multivariate techniques:
  - Log-likelihood ratio based identification
  - Covariance-matrix-based identification (H-matrix)
  - Boosted decision tree and neural network techniques

Here I’ll discuss cut-based identification

Discriminating variables used for electron/photon Id. and jet rejection

- Hadronic leakage ($E_{had1}/E_{clus}$)
- Shower shape in the middle layer of EM calorimeter
- Shower shape in the strip layer (search for 2nd maximum for $\pi^0$ rejection)
- Isolation in calorimeter
- Track isolation to reject low track multiplicity jets containing $\pi^0$

Variables used for Electron Only:

- Track quality (number of hits, impact parameter)
- Track match
- Fraction of high threshold TRT hits
Identification of Electron and Photon (discriminating variables)

Signal: \( H \rightarrow \gamma\gamma \)
Background: QCD jets

In first layer (strips)
\[ \Delta E_s = E_{\text{max}} - E_{\text{min}} \]

Middle layer
\[ \frac{E_2(3\times7)}{E_2(7\times7)} \]

Signal: \( H \rightarrow \gamma\gamma \)
Background: QCD jets

Signal (blue): \( Z \rightarrow \text{ee} \)
Background: QCD di-jets

Middle layer
\[ \frac{E_2(3\times3)}{E_2(3\times7)} \]

Ashfaq Ahmad

ICATPP, Como Oct. 5-9, 2009
Performance of the cut-based Identification (electron)

- Three sets of identification cuts are defined by combining discriminating variables discussed before
  - Namely; Loose, medium and tight cuts

- Electron identification efficiency in H→eeee decays

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency (%)</th>
<th>Jet Rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loose (Had + middle)</td>
<td>88</td>
<td>560</td>
</tr>
<tr>
<td>Medium (Loose + strips + # Si. Hits)</td>
<td>77</td>
<td>2200</td>
</tr>
<tr>
<td>Tight (Medium + TRT + #vert. hits)</td>
<td>64</td>
<td>(10^5)</td>
</tr>
</tbody>
</table>

Recent work has resulted significant improvement in electron efficiency
Performance of the cut-based Identification (jet rejection)

- Expected differential cross section after tight cuts from W/Z, QCD di-jets and minimum bias simulated samples @ 100pb⁻¹
- QCD hard jets with $E_T > 17$ GeV
- Minimum bias $E_T > 8$ GeV
- Shapes of electrons from non-isolated and residual jet background are similar
- ~100k electrons from b and c decays with $E_T > 10$ GeV per pb⁻¹
- ~5k electrons with $E_T > 20$ GeV per pb⁻¹
Performance of the cut-based Identification (photon)

- Photons from $H \rightarrow \gamma \gamma$
  - $E_T > 25$ GeV
- QCD jet background
- Overall efficiency ~84% with jet rejection ~8000

Fake rate = 1/jet rejection

Efficiency vs. $E_T$

Efficiency vs. $\eta$
Comparison of later shower shapes between cosmic data and MC

Selection:
- Require EM cluster of $E_T > 5\text{GeV}$
- Additional cuts to match the difference in acceptance between data and MC
  - At least one Si track $|d_0| < 220\text{mm}$ with $p_T > 5\text{GeV}$
  - $E_{\text{strips}}/E_{\text{cluster}} > 0.1$

After selection sample has 1200 photon candidates in data (out of 3.5 million) and 2161 from MC (out of ~11.7 million)

Good agreement between data and MC
Lateral shower shapes

\[ F_{\text{side}} = \frac{(E_{\pm 3} - E_{\pm 1})}{E_{\pm 1}} \]

Different shower development between top (\( \phi > 0 \)) and bottom (\( \phi < 0 \))

Good agreement between data and MC

For detail cosmic results please see Christian Schmitt talk on Wednesday
Energy resolution is parameterized as

$$\sigma_E / E = a / E \oplus b / \sqrt{E} \oplus c_{tot} \quad \text{with} \quad c_{tot} = c_L \oplus c_{LR} \leq 0.7$$

From the test beam the local constant ($c_L$) term ∼0.5% ⇒ Hence the “long range” zone to zone non-uniformity ($c_{LR}$) must be ≤ 0.5% ⇒ zone is $\Delta \eta \times \Delta \phi = 0.2 \times 0.4$

In-situ calibration also has to establish absolute EM scale to an accuracy ∼0.1%

Long range non-uniformities can be corrected using electrons from Z boson decays

Parameterize electron energy in zone “i” as $E_{i \text{reco}} = E_{i \text{true}} (1 + \alpha_i)$

$\alpha$’s are obtained from likelihood fit by constraining the measured di-electron invariant mass to the Z boson line shape

![Graph showing before and after correction of events and invariant mass distribution](image)
Need to measure efficiency from data to scale MC prediction.

Use "Tag and Probe" method on $Z \to \text{ee}$ and $J/\Psi$ events:

- Select one good quality electron (tag, passing set of cuts)
- Constrain with Z mass
- Measure the efficiency of passing cuts by second electron (probe)
Conclusion

- Understanding of the electron and photon reconstruction and identification are essential for many SM physics measurement and for new physics searches at the LHC

- Different algorithms have been developed in ATLAS for this purpose and thoroughly tested with beam tests, with simulations and cosmic data

- The performance of ATLAS Electromagnetic (EM) calorimeter and tracker has been extensively studied
  
  - The absolute energy scale of EM calorimeter is known to be~1% and linearity better than 0.5% for wide range of energies

- The MC studies show that the reconstruction and identification efficiencies and background rejection are adequate for physics measurements
Backup Slides
Cluster corrections

- Position corrections ($\eta/\phi$ corrections):
  - Correct for bias in $\eta$ position due to the finite granularity of the readout cells
    - depends on particle impact point within cells
    - give rise to S-shape in $\eta$ position
    - a few percent difference between electrons/photons corrections
    - corrections are $\eta$ and energy dependent
  - $\eta$ position resolution for photon is $\sim 3 \times 10^{-4}$ in strips (layer 1) and $\sim 6 \times 10^{-4}$ in middle (layer 2)
  - $\phi$ corrections:
    - correct for small bias in $\phi$
    - $\phi$ resolution $\sim 0.5-1.5 \times 10^{-3}$
\( \Delta \theta \) is defined as the angle between the direction of the shower and the direction defined from the centre of the detector to the centre of the cluster.
<table>
<thead>
<tr>
<th></th>
<th>Barrel</th>
<th>End-cap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EM calorimeter</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of layers and $</td>
<td>\eta</td>
<td>$ coverage</td>
</tr>
<tr>
<td>Presampler</td>
<td>1 (1.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>Calorimeter</td>
<td>3 (1.375 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>2 (1.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>2 (2.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>**Granularity $\Delta \eta \times \Delta \phi$ versus $</td>
<td>\eta</td>
<td>$**</td>
</tr>
<tr>
<td>Presampler</td>
<td>0.025 x 0.1 (1.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>Calorimeter 1st layer</td>
<td>0.025/8 x 0.1 (1.425 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025 x 0.1 (1.425 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025/8 x 0.1 (1.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025/6 x 0.1 (1.8 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025/4 x 0.1 (2.0 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025 x 0.1 (2.4 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.1 x 0.1 (2.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>Calorimeter 2nd layer</td>
<td>0.025 x 0.025 (1.40 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.075 x 0.025 (1.40 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.025 x 0.025 (1.40 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>0.1 x 0.1 (2.5 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>Calorimeter 3rd layer</td>
<td>0.050 x 0.025 (1.35 &lt; $</td>
<td>\eta</td>
</tr>
<tr>
<td>Number of readout channels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presampler</td>
<td>7808</td>
<td>1536 (both sides)</td>
</tr>
<tr>
<td>Calorimeter</td>
<td>101760</td>
<td>62208 (both sides)</td>
</tr>
</tbody>
</table>