Tau Reconstruction and Identification Performance at ATLAS

Björn Gosdzik
on behalf of the ATLAS Collaboration
Deutsches Elektronen-Synchrotron, Notkestrasse 85, D-22607 Hamburg, GERMANY

For many signals in the Standard Model including the Higgs boson, and for new physics like Supersymmetry, τ leptons represent an important signature. This work shows the performance of the ATLAS τ reconstruction and identification algorithms. It will present a set of studies based on data taken in 2010 at a center-of-mass energy of $\sqrt{s} = 7$ TeV. We measured some of the basic input quantities used for these identification methods from selected reconstructed τ candidates and compared the results to the prediction of different Monte Carlo simulation models. For early data taking a cut-based identification method will be used. We also measured the background efficiency for the cut-based τ identification.

1. Introduction

In the Standard Model, a large number of τ leptons is expected from the decay of Z and W bosons with 100 pb$^{-1}$ of data. τ leptons also play an important role in searches for new phenoma like the Standard Model Higgs boson, MSSM Higgs bosons and SUSY with large tan β since τ leptons can differentiate between SUSY models based on polarization information. They have a mass of $m_\tau = 1.78$ GeV and decay $\approx 65\%$ of the time hadronically.

τ leptons in ATLAS typically have a collimated calorimetric cluster, 1 or 3 charged decay products and a displaced secondary vertex in the case of 3-prong decays.

2. Reconstruction and Identification

The studies are based on data collected with the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 7$ TeV and correspond to an integrated luminosity of approximately $\mathcal{L} = 244$ nb$^{-1}$. Dedicated cuts on the data, haven been applied to select events with back-to-back jets and to enrich the sample with fake τ candidates from QCD processes that form the primary background in searches with τ lepton final states.

To compare the distribution of the variables used for the τ reconstruction and identification we used predictions from QCD jets Monte Carlo (MC) samples, generated with the Pythia DW tune. The reconstruction of hadronically decaying τ leptons starts from either calorimeter or track seeds:

- Track-seeded candidates start with a seeding track of $p_T > 6$ GeV, $|\eta| < 2.5$, and satisfy quality criteria on the impact parameter with respect to the interaction vertex ($|d_0| < 2$ mm and $|z_0| \times \sin \theta < 10$ mm).
Calorimeter-seeded candidates consist of calorimeter jets reconstructed with the anti-Kt algorithm (using a distance parameter \( D = 0.4 \)) starting from topological clusters with a calibrated \( E_T > 10 \text{ GeV} \) and \( |\eta| < 2.5 \).

Candidates are labeled double-seeded when a seed track and a seed jet are within a distance of \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.2 \).

Seven variables are currently used as inputs for the identification algorithms to distinguish \( \tau \) leptons from QCD jets. The variables electromagnetic (EM) radius and track radius are shown in Fig. 1.

### 3. Background Rejection in QCD events

![Figure 2](image-url)

To identify \( \tau \) leptons after the reconstruction, three independent identification (ID) algorithms were studied: simple cuts, boosted decision trees (BDT) and a projective likelihood (LL).

The cut-based ID uses three variables:

\[
\begin{align*}
R_{\text{EM}} &= \sum_{i} \frac{E_{\text{EM},i} \Delta R_{i}}{\sum_{i} E_{\text{EM},i} \Delta R_{i}}, \\
R_{\text{track}} &= \sum_{i} \frac{p_{\text{T},i} \Delta R_{i}}{\sum_{i} p_{\text{T},i} \Delta R_{i}} \\
f_{\text{trk},1} &= \frac{p_{\text{T},1}}{p_{\tau}}
\end{align*}
\]

Three selections corresponding to signal efficiencies of 30\% (tight), 50\% (medium), and 60\% (loose) are optimized to maximize the rejection of QCD jets. The background efficiency for all three selections is shown in Table 1. The background efficiency \( \varepsilon'_{\text{bkgd}} \) requires \( n_{\text{track}} = 1 \) or \( n_{\text{track}} = 3 \). Figure 2 shows the background efficiency from Data/MC and signal efficiency from \( Z \rightarrow \tau \tau \) MC for the cut-based ID. Figure 3 (left) shows the BDT jet score and Fig. 3 (right) the likelihood score. The number of \( \tau \) candidates in the MC samples is normalized to the number of \( \tau \) candidates in the data. Very good agreement between the data and the prediction of QCD MC is observed.

| Selection  | \( \varepsilon_{\text{bkgd}} \) (data) | \( \varepsilon_{\text{bkgd}} \) (MC) | \( \varepsilon'_{\text{bkgd}} \) (data) | \( \varepsilon'_{\text{bkgd}} \) (MC) |
|------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| loose      | \((3.2 \pm 0.2) \times 10^{-1}\)  | \((3.4 \times 10^{-1}\)   | \((9.4 \pm 0.6) \times 10^{-2}\) | \((10 \times 10^{-2}\)   |
| medium     | \((9.5 \pm 1.0) \times 10^{-2}\) | \((9.9 \times 10^{-2}\)   | \((3.1 \pm 0.4) \times 10^{-2}\) | \((3.3 \times 10^{-2}\)   |
| tight      | \((1.6 \pm 0.3) \times 10^{-2}\) | \((1.9 \times 10^{-2}\)   | \((5.6 \pm 0.9) \times 10^{-3}\) | \((6.8 \times 10^{-3}\)   |

Table 1: Background efficiencies for loose, medium, and tight selection cuts. The measured background efficiencies in data are compared to the MC DW tune prediction.

Two effects contribute to systematic uncertainties:

- **The transverse momentum calibration**: Two calibration schemes have been compared, a global cell energy-density weighting (GCW) and a simple \( p_{\text{T}} \) and \( \eta \) dependent calibration (EM+JES). The ratio of the background efficiency for both calibration schemes as function of \( p_{\text{T}} \) is shown in Fig. 4 (left).
Figure 3: BDT jet score (left) and LL score (right) for \( \tau \) candidates in data and MC samples [2].

- **The pile-up effect:** During the data taking period, the beam intensity has increased significantly. The number of vertices \( n_{\text{vtx}} \) is highly correlated with pile-up activity. The background efficiency as function of \( n_{\text{vtx}} \) is shown in Fig. 4 (right).

Figure 4: Ratio of background efficiencies using EM+JES and GCW calibration as a function of \( p_T \) (left); background efficiencies as a function of \( n_{\text{vtx}} \) (right) [2].

4. Conclusion

All variables used in the \( \tau \) ID algorithm are well described by MC predictions and show good separation power between \( \tau \) leptons and fake \( \tau \) candidates from QCD jets. Altogether the commissioning of the tau ID was successful.

Acknowledgments

Work supported by the Helmholtz Young Investigators Group VH-NG-303.

References

1. The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.
2. The ATLAS Collaboration, *Tau Reconstruction and Identification Performance at ATLAS*, ATLAS-CONF-2010-086.
3. TeV4LHC QCD Working Group, *Tevatron-for-LHC Report of the QCD Working Group*, [arXiv:hep-ph/0610012v1](https://arxiv.org/abs/hep-ph/0610012v1).

Proceedings of HCP2010 – Toronto