

Reconstruction and Identification Techniques for Tau Leptons at the ATLAS Detector at LHC

Felix Friedrich
on behalf of the ATLAS Collaboration

Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany

E-mail: felix.friedrich@tu-dresden.de

Abstract. Tau leptons play an essential role in the physics program at the LHC. They are used in the studies for the recently-observed Higgs boson as well as in electroweak measurements and are one of the key ingredients to search for processes predicted by theories beyond the Standard Model. Identifying hadronically decaying tau leptons with a good performance is an essential part of these analyses. The techniques of tau lepton reconstruction and identification used with the ATLAS detector are presented. The misidentification probabilities of QCD jets and electrons are determined from ATLAS data collected in 2012 and from Monte Carlo simulation.

1. Introduction

Tau leptons play an important role in the physics program of the ATLAS detector at the Large Hadron Collider (LHC) [1] at CERN. They are important signatures for Standard Model processes as well as for new physics. With a mass of $1.777 \text{ GeV}/c^2$ the tau lepton is the heaviest lepton and due to its short lifetime of $2.9 \times 10^{-13} \text{ s}$ ($c\tau = 87 \mu\text{m}$) it decays inside the beam pipe. The tau is the only lepton which decays into both leptons and hadrons, decaying leptonically in 35% of the cases and hadronically in the remaining 65%. The majority of hadronic tau decays are characterized by one or three charged pions which may be accompanied by neutral pions. The kinematics of QCD jets are similar to that of hadronically decaying tau leptons, leading to a potentially high probability for misidentifying them as taus. In addition the cross-sections of most of the Standard Model and new physics processes with tau leptons in the final state are small compared to the overwhelming background from QCD processes at LHC. Therefore a well performing tau lepton identification is crucial for the physics program. In ATLAS [2], tau reconstruction and identification [3] concentrates on the hadronic decay modes of a tau lepton. They are classified according to the number of charged decay particles (also called ‘prongs’ in this text). These decays can be distinguished from QCD jets via their characteristics, such as low track multiplicity, collimated energy deposits, and, in case of 3-prong tau decays, the displacement of the secondary vertex.

2. Tau Lepton Reconstruction

Calorimeter jets with a transverse energy larger than 10 GeV and within the detector acceptance are used as a seed for the reconstruction of tau lepton candidates. The tau energy is calculated using refined calorimeter clusters within a core of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.2$ around the 4-vector sum of calorimeter clusters associated with the jet seed. The final tau energy is then obtained from the reconstructed energy by a simulation-based calibration method [4]. The systematic uncertainties on the tau energy scale are



estimated by measuring single particle response function and their propagation to the tau decay products. They were found to be at the 3%–4% level. Tracks passing certain quality criteria and lying within a cone of $\Delta R = 0.2$ around the tau candidate axis are used to define the core region and to classify the tau candidate into single- or multi-prong (number of charged tracks) categories. Variables, later needed for the tau identification methods, are calculated based on both calorimeter and tracking information.

3. Tau Lepton Identification

During the tau reconstruction process no attempt is made to separate tau leptons from QCD jets. Therefore a dedicated identification procedure is applied, using one of the three following methods: a cut-based approach, placing rectangular cuts on variables, a projective likelihood (LLH) method, using the log-likelihood-ratio of signal and background, and boosted decision trees (BDT), to find the optimal separation in a multi-dimensional phase space. While dedicated working points for predefined signal efficiencies are provided in the cut-based approach, the likelihood (Figure 1) and BDT return a continuous response. All algorithms are based on variables known to have discrimination power between QCD jets and tau leptons. While the charged tracks from the tau lepton decay are collimated in a narrow cone, tracks from QCD jets tend to spread out (Figure 2). In addition, the energy deposit in the calorimeter is collimated in a small area around the tau axis, while for QCD jets a larger area is involved (Figure 3). For the training of the selection algorithms the QCD background was obtained from data, while the tau decay signal is simulated in $W \rightarrow \tau\nu$, $Z \rightarrow \tau\tau$ and $Z' \rightarrow \tau\tau$ Monte Carlo samples. The latter is a hypothetical heavy Z boson and ensures the algorithm is also trained for high tau p_T ranges.

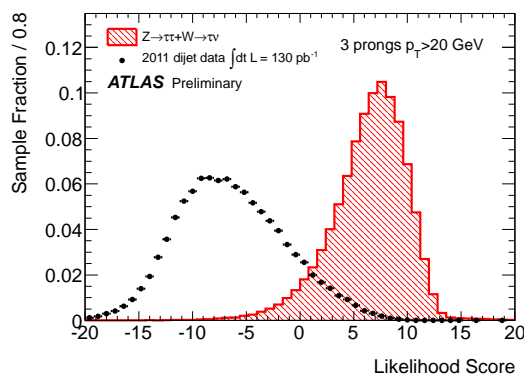


Figure 1: Output score of the projective likelihood tau identification method. The score is calculated with the 7 TeV data set. [5]

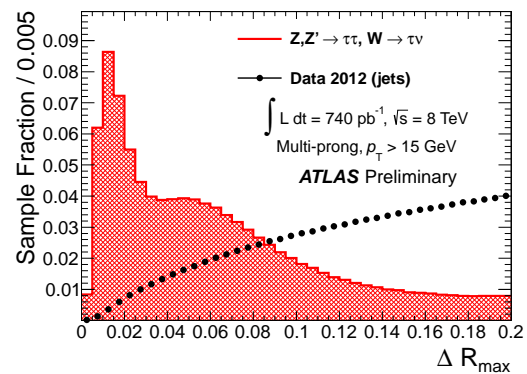


Figure 2: Maximal distance between a track and the tau candidate axis. Only tracks inside a cone of $\Delta R = 0.2$ considered. [3]

3.1. Performance for $\sqrt{s} = 7$ TeV data taking period

In the 2011 data taking period with a center-of-mass-energy of 7 TeV, beside the LLH- and BDT-based algorithm a simple cut-based method is also used to identify hadronically decaying tau leptons. One of the reason for following this approach is to compare the performance of the LLH- and BDT-based tau identification with a method which is well understood. For this data taking period all three methods are trained using different sets of input variables, separately for 1-prong and multi-prong tau candidates as well as parametrized in ranges of p_T of the tau candidate. The LLH and BDT are trained for various number of vertices to account for event pile-up. The performance for 3-prong tau candidates is shown in Figure 4. An inverse background efficiency of about 100 means that only 1 jet out of 100 is identified as a tau lepton.

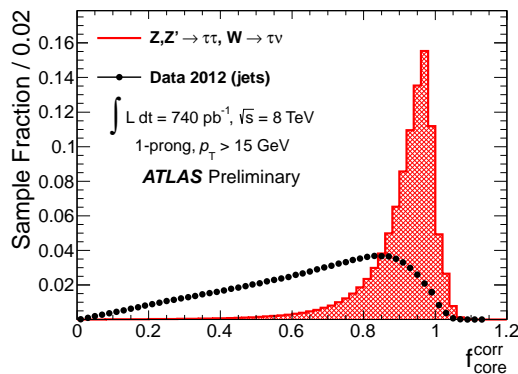


Figure 3: Fraction of the transverse energy of calorimeter cells deposited in a cone of $\Delta R = 0.1$ around the tau candidate axis to those deposited in the region of $\Delta R = 0.2$. [3]

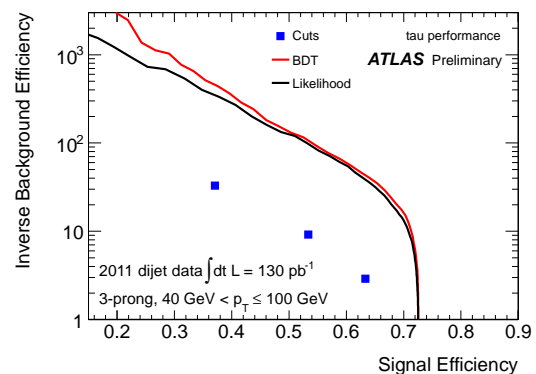


Figure 4: Signal efficiency versus inverse background efficiency for the different tau identification methods shown for 3-prong tau candidates with $40 \text{ GeV} < p_T < 100 \text{ GeV}$. The plot is obtained using the 7 TeV data set. [5]

3.2. Performance for $\sqrt{s} = 8 \text{ TeV}$ data taking period

For the 2012 data taking period with a center-of-mass energy of 8 TeV only the LLH- and BDT-based tau identification are optimized. The cut-based approach is dropped for performance reasons. Both multi-variate methods use the same set of identification variables and are separately trained for single- and multi-prong tau candidates using 5 and 6 variables, respectively. To account for the p_T dependence of the input variables the methods are parametrized in ranges of p_T of the tau candidate. Given the large difference in the composition of background and signal, either flat signal efficiency or flat background rejection can be obtained, but not both. In ATLAS the tau identification is parametrized to be flat in tau p_T for the signal efficiency (Figure 5). In addition some input variables are corrected for the number of reconstructed vertices in the event in order to take event pile-up into account. It is found that correcting the input variables directly results in a better performance than introducing another parametrization in the training of the multi-variate methods. This leads to a flat signal efficiency as well as a flat background rejection (Figure 6). The inverse background efficiency versus signal efficiency for both methods, LLH and BDT, is shown for multi-prong high- p_T tau candidates in Figure 7.

3.3. Rejection against leptons

Electrons can also be misidentified as tau leptons. Due to the signature of the electron in the detector, they will be reconstructed mostly as 1-prong tau candidates. To distinguish between electrons and tau leptons a multi-variate veto based on boosted decision tree (BDT) is used. For the training, $Z \rightarrow ee$ Monte Carlo samples are used as background. The signal samples are the same as for the training against QCD jets. The performance for the $\sqrt{s} = 8 \text{ TeV}$ data taking period of the electron veto is shown in Figure 8. Muons are unlikely to be reconstructed as tau leptons because of their low energy deposition in the calorimeter. However in the presence of high-energy clusters, muons are able to fake tau leptons. In this case most of the energy is deposited in the hadronic calorimeter. To reject those candidates, a muon veto based on simple cuts on a few discriminating variables, like the fraction of energy in the electromagnetic and hadronic calorimeter is used.

4. Summary and Conclusion

ATLAS has a rich physics program involving tau lepton final states, and a well performing tau identification is an essential part of these analyses. Various techniques are used to separate tau leptons from quark- and gluon-initiated jet background as well as from electrons and muons. The multivariate

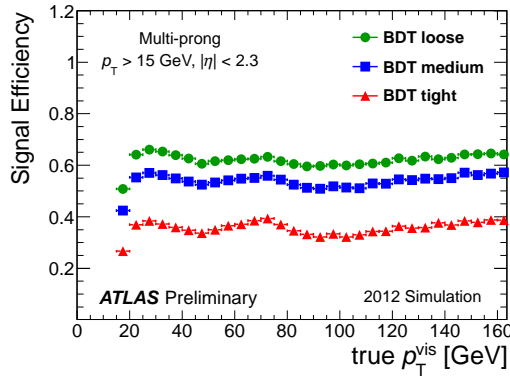


Figure 5: Signal efficiency for multi-prong tau candidates for different BDT working points as a function of the true visible tau p_T for signal candidates. The efficiencies were obtained using the 8 TeV data set. [3]

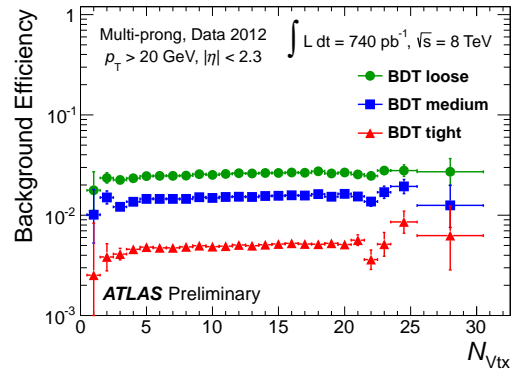


Figure 6: Background efficiency for multi-prong tau candidates for different BDT working points as a function of the number of vertices in the event. The efficiencies were obtained using the 8 TeV data set. [3]

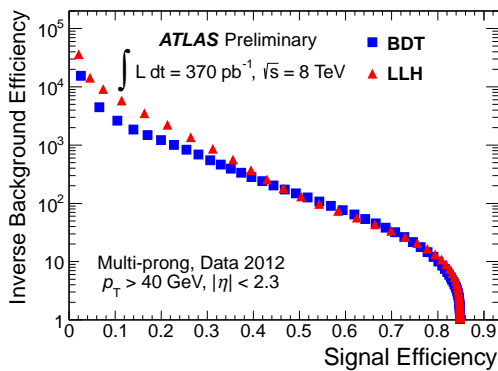


Figure 7: Signal efficiency versus inverse background efficiency for the different tau identification methods shown for multi-prong tau candidates. The plot was obtained using the 8 TeV data set. [3]

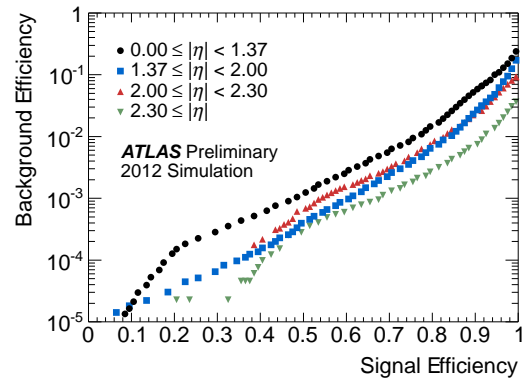


Figure 8: Signal efficiency versus background efficiency for the BDT-based tau electron veto method shown for 1-prong tau candidates with $p_T > 20$ GeV and for different parts of the detector. [3]

methods achieve more performance than a simple cut-based approach. A continuous effort is necessary to keep the performance of the methods at a high level over a wide tau p_T range and for different data taking conditions.

References

- [1] Evans L and Bryant P (ed) 2008 LHC Machine, *JINST* **3** S08001
- [2] The ATLAS Collaboration 2008 The ATLAS Experiment at the CERN Large Hadron Collider, *JINST* **3**, (2008) S08003
- [3] The ATLAS Collaboration 2013 Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector, Conference Note, ATLAS-CONF-2013-064
- [4] The ATLAS Collaboration 2013 Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012, ATLAS-CONF-2013-044
- [5] The ATLAS Collaboration 2011 Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS, Conference Note, ATLAS-CONF-2011-152