Tau Lepton Reconstruction and Identification with the ATLAS Detector at the LHC

A. Kalinowski
On behalf of the ATLAS Collaboration

a Physics Department, University of Regina, 3737 Wascana Parkway Regina, Saskatchewan, Canada

The Standard Model and many scenarios of physics beyond the Standard Model predict an abundant production of \( \tau \) leptons. In particular \( \tau \) leptons in the final state are important signatures for Higgs boson and SUSY searches. Equipped with excellent tracking and calorimetry, the ATLAS detector allows for the reconstruction and efficient identification of hadronically decaying \( \tau \) leptons. The methods developed by the ATLAS Collaboration for the offline reconstruction and identification of hadronic decays of \( \tau \) leptons are described. The performance of the reconstruction algorithm and the discrimination power of the identification methods against jets and electrons are shown. Ideas for the validation of the \( \tau \) identification methods with early data are also presented.

1. INTRODUCTION

The \( \tau \) lepton branching fraction for decays into one or more hadrons is about 65\%, and about 35\% for decays into muons or electrons [1]. The hadronic \( \tau \) decays are usually divided into two classes: decays with one charged particle (single-prong) — 77\% of all hadronic \( \tau \) decays, and with many charged particles (multi-prong) — 23\% of all hadronic \( \tau \) decays. The \( \tau \) leptons in the final state are important signatures for Higgs boson and SUSY searches [2], but hadronic \( \tau \) decays are difficult to identify in hadron collisions due to the overwhelming QCD background. Efficient and robust reconstruction and identification methods therefore have to be developed by the LHC experiments. This contribution describes the algorithms developed by the ATLAS Collaboration for the reconstruction and identification of hadronically decaying \( \tau \) leptons. The performance of the identification algorithms using a detailed simulation of the ATLAS detector [3] is presented.

2. TAU LEPTON RECONSTRUCTION

The reconstruction of hadronically decaying \( \tau \) leptons is performed by two complementary algorithms, using two types of initial seeds: a) a track with transverse momentum above 6 GeV, passing strict quality criteria, or b) topological clusters [2] in the calorimeter with transverse energies \( E_T > 10 \text{ GeV} \). The \( \tau \) lepton reconstruction flow combines the two algorithms in the following way: 1) reconstruction starts from the track seed, and a cluster seed is searched for in a cone of radius \( \Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.2 \); 2) If a cluster seed is found, the cluster-based reconstruction algorithm is run. In that case the properties of the final \( \tau \) candidate are set using the information from both types of candidates, exploiting the best of the two in the following way: the \((\eta, \phi)\) position is defined by the position of the track-based candidate, the energy is defined by the cluster-seeded candidate, and the associated track multiplicity is defined by the track-seeded candidate; 3) if no calorimeter seed was found, the \( \tau \) candidate is marked as track-seeded only; 4) the remaining clusters are used for seeding the calorimeter based only candidates. Among the reconstructed hadronically decaying \( \tau \) leptons in \( Z \rightarrow \tau \tau \) events, 70\% of all candidates have both types of seeds, 25\% are only calorimeter-seeded, and the remaining 5\% are track-seeded only candidates.

The energy estimate for the calorimeter-seeded candidates uses all calorimeter cells in a cone \( \Delta R = 0.4 \) around its position. The cell ener-
gies are calibrated using the method described in [4]. Further energy calibration is applied to the reconstructed energy, using the calibration scale depending on the \( \tau \) candidate’s \( \eta \) and \( p_T \). The ratio of the reconstructed and the true visible energy of the \( \tau \) lepton for the calorimeter-seeded candidates is close to one for a large range of candidate’s \( E_T \) (Fig. 1). For the track-seeded candidates the energy reconstruction is done using an energy flow approach. The calorimeter deposits from charged particles are substituted by the momentum estimated from the charged particle tracks reconstructed in the inner detector, since the track momentum resolution is much better than the calorimeter energy resolution for the \( p_T \) range of interest. The width of the fractional energy resolution, \( (E_{\text{true}} - E_{\text{rec}})/E_{\text{true}} \) for single-prong events is 8\%, as shown in Fig. 2, and 3\% for three-prong events.

The tracks associated to the \( \tau \) candidate, within a cone of \( \Delta R = 0.2 \), are required to pass looser criteria than required for the leading track. The quality requirements allow for the rejection of a substantial fraction of tracks from photon conversions, where most of the photons come from \( \pi^0 \) decays. Figure 3 shows the reconstructed track multiplicity for \( \tau \) candidates matched to a true hadronically decaying \( \tau \) lepton. The peaks from one and three-prong decays are clearly visible. The contribution to events with two tracks comes mainly from true three-prong events, where one of the tracks was not reconstructed or failed to pass the quality criteria.

The isolated topological clusters found in the electromagnetic calorimeter with \( E_T > 1 \) GeV are interpreted as energy deposits from \( \pi^0 \)'s. Approximately 66\% of the \( \tau \rightarrow \nu \pi \) decays are reconstructed with zero \( \pi^0 \) subclusters, and more than 50\% of the \( \tau \rightarrow \rho \nu (\tau \rightarrow a_1 \nu) \) decays are reconstructed with one (two) \( \pi^0 \) subcluster(s).

3. TAU LEPTON IDENTIFICATION

Efficient and robust identification methods are necessary to be able to discriminate between the overwhelming QCD background and real, hadronically decaying \( \tau \) leptons. It is also necessary to use identification algorithms to reject true electrons and muons reconstructed as \( \tau \) candidates. The ATLAS Collaboration has developed a number of identification methods including a simple cut-based method, and a few multivariate methods: projective likelihood (LL),
neural networks (NN), probability density range searches (PDRS) [5], and boosted decision trees (BDT) [6,7].

In this section only the projective likelihood and the electron veto will be briefly described.

The projective likelihood identification method uses the probability distribution functions (PDFs) to calculate the likelihood discriminant. The individual PDFs for signal and background are created for the $\tau$ candidates depending on the number of tracks associated to the $\tau$ candidate (single- or multi-prong), the number of associated $\pi^0$ clusters (zero or at least one cluster) and the valid seed types. In addition the PDFs are created separately for ten bins in $p_T$ of the reconstructed $\tau$ candidate. The likelihood discriminant is defined as a product of individual likelihood ratios for the signal and background hypotheses in each identification variable. The variables used for QCD jets - $\tau$ leptons discrimination include: radius and profile of EM calorimeter energy deposits, track width distributions, isolation variables calculated from calorimetric energy deposits and tracks, impact parameter significance of the leading track, invariant mass of the associated tracks, ratios of energy deposits to the sum of track transverse momenta, and the transverse flight path significance of the $\tau$ candidate vertex in the case of multi-prong candidates. Figure 4 shows an example distribution of the radius of EM calorimeter energy deposits for hadronically decaying $\tau$ (full dots) leptons and QCD jets (triangles). The neural-network-based identification uses a similar list of input identification variables. The performance of the identification in terms of QCD rejection factors (defined as $1/\epsilon_{QCD} - 1$) versus the efficiency for $\tau$ leptons, including the reconstruction efficiency for both signal and background, are presented in Fig. 5 for the LL, and Fig. 6 for the NN method. The results for various $p_T$ bins, and separately for single- and three-prong events are shown. For a medium ($30 < p_T < 60$ GeV) transverse momentum range the NN method provides a rejection of $1030\pm30$ ($590\pm70$) for single-prong (three-prong) events at 30% signal efficiency, and the corresponding rejection for the LL method is $1130\pm50$ ($187\pm3$) respectively.

The electrons reconstructed as single-prong $\tau$
candidates are vetoed using a cut based selection, based on the following variables: the ratio of the transverse energy deposited in the EM calorimeter to the track transverse momentum ($E_T/p_T$), which tends to be higher for electrons compared to charged hadrons, and the ratio of high threshold hits to low threshold hits in the Transition Radiation Tracker for the track, which also tends to be higher for electrons. With the above selection the efficiency for reconstructing $\tau$ candidates in $W \rightarrow \tau \nu$ events is 1.6%, while retaining an efficiency for real, hadronically decaying $\tau$ leptons from $W \rightarrow \tau \nu$ decays at the level of 94.9%.

4. PLANS FOR FIRST DATA

The $\tau$ reconstruction and identification methods have to be tested and tuned with the first LHC data to provide reliable performance estimates for physics analyses. This section presents plans for the $\tau$ reconstruction and identification activities with the very first LHC data.

Already the first 100 pb$^{-1}$ of data should allow for a “re-discovery” of the $Z \rightarrow \tau \tau \rightarrow$ hadrons + lepton + X and $W \rightarrow \tau \nu$ processes. The analyses [2] using a detailed simulation of the ATLAS detector predict 1550 $W \rightarrow \tau \nu$ events on top of about 700 background events, including $W \rightarrow e\nu$, $W \rightarrow \tau \nu$, QCD di-jets, tt, and $Z \rightarrow e\tau, \tau\tau$ events, after a number of selection steps, including the identification and reconstruction of the hadronically decaying $\tau$ leptons. Figure 7 presents the expected track multiplicity distribution after the $W \rightarrow \mu\nu$ analysis selection. A comparison of this distribution with the results of the selection on real data should allow for an estimation of the identification efficiency for signal events.

In the case of the $Z \rightarrow \tau \tau \rightarrow$ hadrons + lepton + X process the expected number of signal events is 520, whereas the number of background events is about 90. The analysis of the reconstructed visible mass of the (lepton $\rightarrow \tau$ candidate) pair for $Z \rightarrow \tau \tau$ events can be used for the estimation of the $\tau$ candidate energy scale. Figure 8 presents the dependence of the reconstructed visible mass on the $\tau$ candidate’s energy scale.

More details on the $\tau$ lepton physics with the first LHC data, including observation of the $t\bar{t}$,
Figure 7. The track multiplicity spectrum of accepted $\tau$ candidates after the $W \rightarrow \tau \nu$ analysis selection.

and exotic processes, like $Z' \rightarrow \tau\tau$ can be found in [8].

The huge number of QCD di-jets events available already with the first 100 pb$^{-1}$ can be used for the training of the multivariate methods, and for the direct estimation of the background rejection of these methods. The di-jet events will be analysed using a tag-and-probe method to select a clean sample of jets without a significant contamination from true $\tau$ leptons. Events with two back-to-back jets of similar $p_T$ will be selected. One of the jets, called “tag jet” will be selected requiring more than four associated tracks for $p_T^{\text{jet}} < 50$ GeV and one additional track for each additional 50 GeV interval in $p_T$, to ensure that it is not a true hadronic $\tau$ decay. The other jet, called “probe jet” will be used to measure the fake rate. In addition, to avoid a direct dependence on the trigger, the probe jet is required not to be used in the trigger selection for the analysed event. The statistical uncertainty on the estimation of the rate from QCD jets with this method will be of the order of 0.01% for jets with $p_T < 80$ GeV, and 0.001% for jets with $p_T > 80$ GeV. The systematic uncertainty of this method still has to be estimated.

Figure 8. The reconstructed visible mass of the (lepton $\rightarrow \tau$ candidate) pair from $Z \rightarrow \tau\tau$ decays as a function of the $\tau$ candidate’s energy scale. The dashed lines correspond to $\pm 1\sigma$ and $\pm 3\sigma$ error bands with respect to the reconstructed peak position.

5. CONCLUSIONS

To be able to fully exploit the potential of physics processes involving $\tau$ leptons in the final state, the ATLAS Collaboration has developed robust and efficient reconstruction and identification methods allowing for good separation between QCD jets, electrons, and real hadronically decaying $\tau$ leptons. Detailed studies show that it will be possible to observe hadronically decaying $\tau$ leptons with the first 100 pb$^{-1}$ of LHC data, allowing for studies of the reconstruction and identification algorithms. The identification methods will also be studied in detail using QCD jets selected with tag-and-probe methods.

REFERENCES

8. Z. Czyczula, Tau Physics with First Data in ATLAS, poster at this conference