Performance of Tau Trigger and Tau Reconstruction in ATLAS in pp Collisions at $\sqrt{s} = 7$ TeV *

Marcin Wolter †
On behalf of the ATLAS Collaboration

Henryk Niewodniczanski Inst. of Nucl. Physics, PAN, Cracow, Poland

Tau leptons provide a useful signature in searches for new physics phenomena in the ATLAS experiment, like Higgs bosons or supersymmetry. The Standard Model processes with tau leptons are important backgrounds in such searches and also can be used to calibrate the detector and demonstrate the performance of tau identification.

The data collected at centre-of-mass energy of $\sqrt{s} = 7$ TeV with the ATLAS detector are used to study the reconstruction and identification algorithms for hadronic tau decays. Their performance in data and Monte Carlo simulations is compared in dijet sample and good agreement is observed.

The first observation of $W \to \tau\nu$ decays in ATLAS is also presented. The observed yield over the total background is compatible with Standard Model signal expectation.

PACS numbers: 14.60.Fg, 29.90.+r, 07.05.Kf, 07.05.Tp

1. Introduction

Tau leptons play an important role in the LHC physics programme, for example in searches for Higgs bosons or supersymmetry [1]. Decays of Standard Model gauge bosons to tau leptons, $W \to \tau\nu$ and $Z \to \tau\tau$, are important background processes in such searches. They give also a unique opportunity to demonstrate the performance of tau identification and to calibrate the reconstruction algorithm. The cross sections for these processes were never measured at such high energies, so their measurement is an interesting task by itself.

* Presented at the Cracow Epiphany Conference on the first year of the LHC, Cracow, Poland, January 10-12, 2011

† Work supported in part by Polish-DAAD collaboration 764/N-DAAD/2010/0 and the Polish Government grant NN202127937 (years 2009-2011)

Tau leptons decay leptonically to an electron or muon (and associated neutrinos), but such decays are very difficult to distinguish from prompt leptons. Therefore in the following, we concentrate on hadronic tau decays, which represent about 65% of the tau lepton branching ratio. Such a decay is characterized by a small number of collimated tracks (typically one or three, coming from charged pions) in the tracking detectors with no track activity in an isolation region around them. The sizable lifetime $c\tau = 87\mu m$ [2] generates a noticeable transverse flight path. Decaying tau leptons leave also well collimated energy deposits in the calorimeter, often associated with strong electromagnetic (EM) component from π^0 produced in tau decays. Typically the energy deposit in the isolation region around them is small.

Since March 2010 the ATLAS [3] experiment at the LHC has been collecting proton-proton collisions events at a centre-of-mass energy of $\sqrt{s} =$ 7 TeV. The collected data are used to study the performance of the reconstruction and identification of hadronic tau decays, as well as the trigger selection for hadronically decaying tau leptons.

The tau trigger is described in Section 2, while the offline tau reconstruction and identification are presented in Sections 3 and 4. The first observed processes with hadronically decaying tau leptons with the ATLAS detector are reported in Section 5.

2. Tau trigger

In order to ensure the efficient selection of interesting events at data taking, the trigger system [4] of the ATLAS experiment consists of three steps: a fast hardware-based Level 1 trigger (L1), and the software High Level Trigger (HLT), composed of the Level 2 trigger (L2) and the Event Filter (EF).

The L1 tau trigger finds regions of interest (RoI) in the detector. It uses 0.1×0.1 ($\Delta\eta \times \Delta\phi$) calorimeter towers (sums of several cells) to determine the local maximum above E_{T} threshold in a 0.2×0.2 region. The outer cells from the broader 0.4×0.4 region are optionally used to define an isolation region. The HLT uses RoIs defined by L1 trigger for partial detector readout. At L2 tracking information is combined with jets made out of calorimeter cells and the tau identification variables are built. The algorithm run at EF level is similar to the offline reconstruction procedure (described in sections 3 and 4), using calorimeter energy clusters with proper calibration and noise suppression applied. At HLT the selection is based on rectangular cuts on track and calorimeter cluster variables.

The trigger menu is a complete set of triggers covering the full spectrum of tau physics. It contains:

- single tau triggers with increasing energy thresholds and identification tightness, which are used for heavy $H \to \tau\tau$, $Z' \to \tau\tau$ and $H^{\pm} \to \tau\nu$ identification;

- di-tau triggers designed for heavy resonances;
- triggers combining taus with another object (to reduce rates): tau+e/μ for $Z \to \tau\tau$, $t\bar{t}$, $H \to \tau\tau$, SUSY, tau+$E_{\mathrm{T}}^{\mathrm{miss}}$ for $W \to \tau\nu$, $H^{\pm} \to \tau\nu$, SUSY and tau+(b)jets for $t\bar{t}$ and SUSY.

The single tau trigger efficiency is presented in Fig. 1 showing a good agreement between data and MC for L1, L2 and EF for a minimum bias background sample.

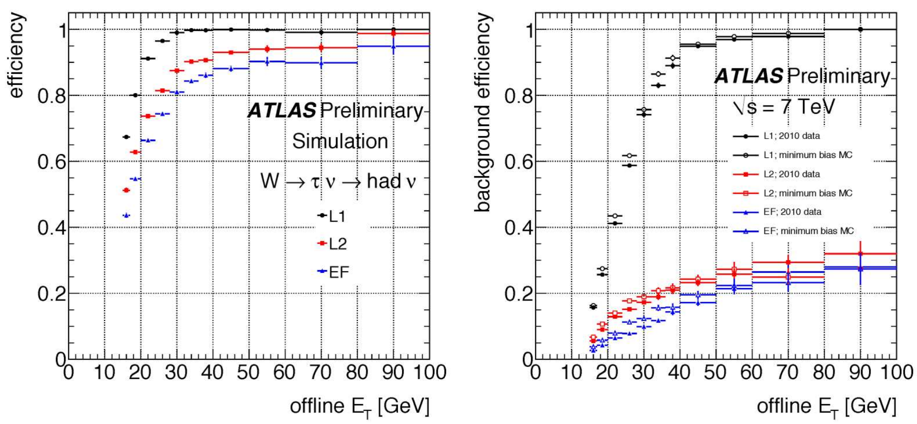

Fig. 1. Fraction of reconstructed tau candidates (no identification applied) passing L1 (5 GeV), L2 (7 GeV) and EF (12 GeV) loose trigger conditions as a function of E_{T} of the offline candidate. Left: signal efficiency on $W \to \tau\nu$; right: data-MC comparison on minimum bias background.

3. Tau reconstruction

The data collected at centre-of-mass energy of $\sqrt{s} = 7$ TeV recorded with the ATLAS detector with integrated luminosity of 244 nb^{-1} [5] are used to study the reconstruction and identification algorithms for hadronic tau decays.

All events must satisfy the Level 1 trigger condition requiring a tau-trigger object passing a $p_{\mathrm{T}} = 5$ GeV threshold. In order to select events with back-to-back jets and therefore enrich the sample with fake tau jets originating from QCD processes additional selection criteria are applied. At least one tau candidate with $p_{\mathrm{T}} > 30$ GeV and another one with $p_{\mathrm{T}} >$

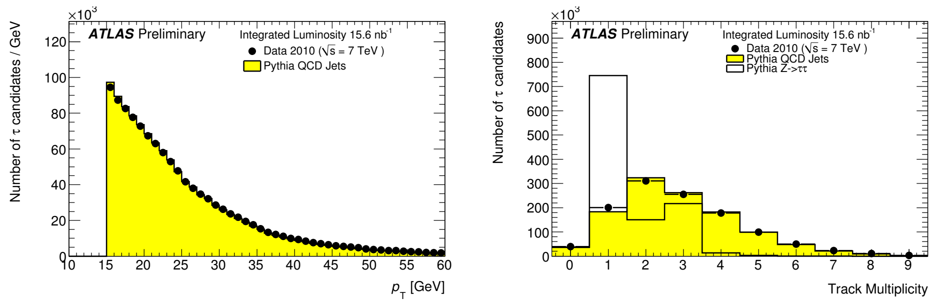

Fig. 2. Transverse momentum distribution (left) and number of associated tracks of τ candidates (right). The number of τ candidates in MC samples are normalised to the number of τ candidates selected in data. The data correspond to an integrated luminosity of 15.6 nb^{-1}.

15 GeV are required. They should be separated by at least 2.7 radians in azimuthal plane. Also the leading tau candidate is excluded to remove any trigger bias. Data sample selected contains about 2.9 million events with 3.9 million tau candidates.

The obtained distributions are compared to the simulated QCD samples. The transverse momenta of the outgoing partons are restricted to be between 8 and 280 GeV. These samples are generated with Pythia [6] using the DW tune [7] and passed through a Geant4 simulation of the ATLAS detector [8]. When showing distributions for true tau candidates, a $Z \to \tau\tau$ MC sample with the MC09 tune [9] is used.

The reconstruction of hadronically decaying tau leptons starts either from calorimeter or track seeds [1]. Reconstruction of calorimeter-seeded tau candidate begins with calorimeter jets reconstructed with the anti-k_t algorithm [10] (using a distance parameter $R = 0.4$) starting from topological clusters [11]. The candidate is required to have $p_T > 10$ GeV. Track-seeded candidates are required to have seeding track with $p_T > 6$ GeV and the tracks with $p_{\mathrm{T}} > 1$ GeV are collected around it in a cone $\Delta R < 0.2$. If jet seeds are found within $\Delta R < 0.2$, such a candidate is labeled as double-seeded.

Only a small percentage of tau candidates are track-seeded only. In the studies presented here, candidates with both seeds and candidates with only a calorimeter-seed with at least one associated track are considered.

Figure 2 shows transverse momentum distribution and the number of associated tracks of the tau candidate (a real tau lepton is expected to have mostly one or three such tracks). Monte Carlo simulation and data agree very well.

4. Tau identification

The tau reconstruction algorithm does not provide large rejection against QCD jets. Therefore an additional identification (ID) step is necessary. Tau leptons are difficult to identify and therefore require the full power of ID variables. A simple cut-based ID as well as more advanced likelihood and boosted decision tree (BDT) multivariate techniques are used [12]. While discriminating variables, multivariate techniques and detailed systematic studies are described in detail in [13], here only the basic identification based on rectangular cuts is presented. This robust identification method is used in the analysis of $W \to \tau\nu$ decays (Sec. 5).

Discriminating variables used by the cut-based ID include the EM radius (E_{T}-weighted shower width in EM calorimeter), the track radius (p_{T}-weighted track width) and the leading track momentum fraction (ratio of the p_{T} of the leading track and the total transverse momentum of the tau candidate). Different cuts are applied for tau candidates with one or with more tracks. The optimization is done for 30% (tight), 50% (medium) and 60% (loose) signal efficiency. The performance of the tau identification is evaluated in terms of signal and background efficiencies. Signal efficiency is defined as $\varepsilon_{\mathrm{s}} = N^{\tau}_{\mathrm{pass,match}}/N^{\tau}_{\mathrm{match}}$, where $N^{\tau}_{\mathrm{match}}$ is the number of reconstructed tau candidates that are matched within a cone of $\Delta R < 0.2$ with a true, hadronically decaying tau lepton with visible transverse momentum $p_{\mathrm{T}}^{\mathrm{vis}} > 15$ GeV and visible pseudorapidity $|\eta^{\mathrm{vis}}| < 2.5$, reconstructed with the correct number of associated tracks; while $N^{\tau}_{\mathrm{pass,match}}$ is the number of these reconstructed candidates that pass the identification criteria. A simulated sample of $Z \to \tau\tau$ decays is used to evaluate the signal efficiency. The background efficiency is defined as $\varepsilon_{\mathrm{b}} = N^{\mathrm{b}}_{\mathrm{pass}}/N^{\mathrm{b}}_{\mathrm{total}}$, where $N^{\mathrm{b}}_{\mathrm{pass}}$ is the number of the τ candidates that pass the identification criteria, and $N^{\mathrm{b}}_{\mathrm{total}}$ is the number of tau candidates in the dijet selection described earlier.

The signal and background efficiencies for the loose, medium and tight settings of the cut-based ID are shown in Fig. 3 as a function of p_{T}. The agreement between data and MC is reasonable. Figure 4 shows background efficiencies as a function of the number of vertices, which is correlated with the beam intensity. Increased beam intensities lead to different pile-up conditions. The stability of the simple cut ID against the presence of pile-up is satisfactory.

5. Observation of real taus in $W \to \tau\nu$ decays

At next-to-next-to-leading order (NNLO), the $W \to \tau\nu$ signal is predicted to be produced with a cross section times branching ratio of $\sigma \times BR =$ 10.46 nb [14, 15], which is about ten times higher than for $Z \to \tau\tau$ events. Events from $W \to \tau\nu$ production produce predominantly low p_{T} tau lep-

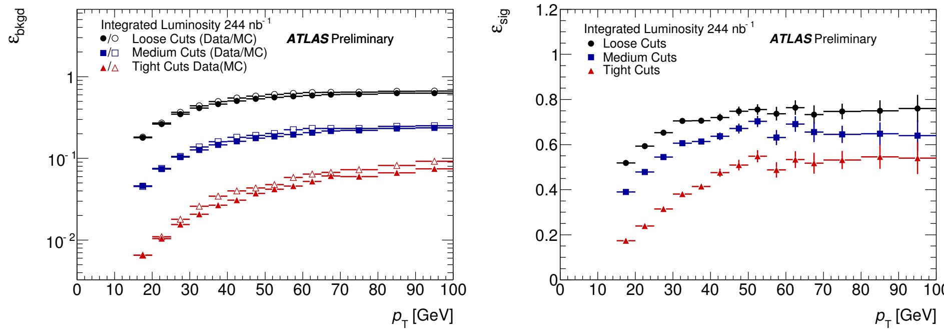

Fig. 3. Background efficiencies obtained for dijet data and MC samples as a function of the reconstructed p_{T}^{τ} (left). Signal efficiencies predicted by a $Z \to \tau\tau$ MC sample as a function of the reconstructed visible p_{T}^{τ} (right).

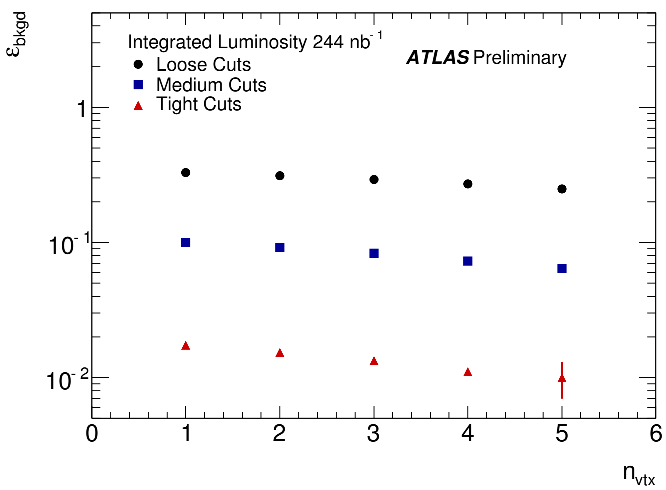

Fig. 4. Background efficiencies as a function of number of vertices n_{vtx}.

tons with typical visible transverse momenta between 10 and 40 GeV. In addition, the distribution of the missing transverse energy, associated with the neutrinos from the W and tau decays, has a maximum around 20 GeV and a significant tail up to about 80 GeV.

The analysis, described in detail in [16], has been performed on data collected between March and mid-August 2010. Only data taken during periods with stable beams and with a good data quality for all the tracking and calorimeter sub-detectors are used. With these basic data quality criteria, the total integrated luminosity available for the analysis amounts to 546 nb^{-1}.

Beside additional quality criteria the events are further required to have the typical $W \to \tau\nu$ signature, i.e., a tau jet accompanied by missing energy due to the undetected neutrinos. A missing transverse energy of $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV is required. Tau candidates must be both-seeded (track and calo-seeded) and identified as tight tau candidates by cut ID. The highest-p_{T} candidate of these is selected and required to have a visible transverse momentum between 20 and 60 GeV. The event is rejected if the selected tau candidate is reconstructed in the pseudorapidity range $1.3 < |\eta| < 1.7$. Electron and muon vetoes are applied to suppress the electroweak backgrounds ($W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu$, $Z \to ee$, $Z \to \mu\mu$ and $Z \to \tau\tau$). Events with identified loose electrons [17] or combined muons [1] with $p_{\mathrm{T}} > 5$ GeV are rejected. The cut-based tau identification provides additional suppression of electrons and muons.

Finally, the event selection includes a requirement on the significance of the missing transverse energy, defined as $S(E_{\mathrm{T}}^{\mathrm{miss}}) = E_{\mathrm{T}}^{\mathrm{miss}}/(0.5 \cdot \sqrt{\sum E_{\mathrm{T}}})$, on the basis of the expected $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of $\sum E_{\mathrm{T}}$ reported in [18]. Events are rejected if $S(E_{\mathrm{T}}^{\mathrm{miss}}) < 6\ \mathrm{GeV}^{1/2}$. This requirement is essential for the rejection of QCD background, for which lower $S(E_{\mathrm{T}}^{\mathrm{miss}})$ values are expected than for $W \to \tau\nu$ events. Figure 5 shows the two-dimensional distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\sqrt{\sum E_{\mathrm{T}}}$ for simulated signal, QCD background and data, together with the $S(E_{\mathrm{T}}^{\mathrm{miss}})$ requirement. The discriminating power of this requirement is clearly visible.

The selection described above results in 78 events. From Monte Carlo simulation, the expected number of signal events that pass the selection is 55.3 ± 1.4 events. The electroweak background from other W and Z decays is 11.8 ± 0.4 events, where the error is the Monte Carlo statistical uncertainty.

A data-driven method is used to estimate the QCD background. It is based on the selection of four independent data samples, three in QCD background-dominated regions (control regions) and one in a signal-dominated region (signal region). The samples are selected with criteria on $S(E_{\mathrm{T}}^{\mathrm{miss}})$ and on the tau identification, which are assumed to be uncorrelated. The following four regions are used in this analysis:

- Region A: events with $S(E_{\mathrm{T}}^{\mathrm{miss}}) > 6$ and tau candidates satisfying the tight tau ID using cut-based method
- Region B: events with $S(E_{\mathrm{T}}^{\mathrm{miss}}) < 6$ and tau candidates satisfying the tight tau ID
- Region C: events with $S(E_{\mathrm{T}}^{\mathrm{miss}}) > 6$ and tau candidates satisfying the loose tau ID but failing the tight ID

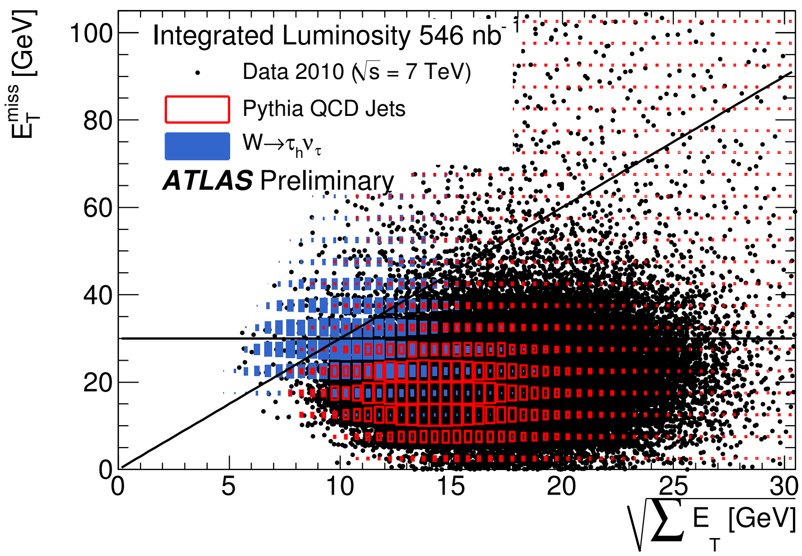

Fig. 5. Distribution of events in the $E_{\mathrm{T}}^{\mathrm{miss}}$ vs. $\sqrt{\sum E_{\mathrm{T}}}$ plane after the trigger requirement for data, simulated signal events and QCD background. The applied $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance cuts are indicated as solid lines.

- Region D: events with $S(E_{\mathrm{T}}^{\mathrm{miss}}) < 6$ and tau candidates satisfying the loose tau ID but failing the tight ID.

This background prediction is based on two assumptions, namely that the shape of the $S(E_{\mathrm{T}}^{\mathrm{miss}})$ distribution for QCD background is the same in the combined regions AB and CD and that the signal and electroweak background contribution in the three control regions is negligible. The estimate for QCD background in the signal region A is then obtained by: $N_{QCD}^{A} = N^{B} N^{C} / N^{D}$ where N^{i} represents the number of observed events in region i.

The estimated QCD background is corrected for electroweak backgrounds in the signal and control regions as well as for the non-negligible signal contribution in the control regions. To confirm the signal observation, the distributions of the tau track multiplicity, $\Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$, the electric charge of the tau candidates, $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T} are compared (see Fig. 6). Here, the data distribution corresponds to the signal region A and the QCD background to the control region C after subtraction of the EW and signal contributions based on Monte Carlo simulation. The distributions are consistent with data.

Of the selected 78 events, $11.1 \pm 2.3_{(stat.)} \pm 3.2_{(syst.)}$ events are estimated

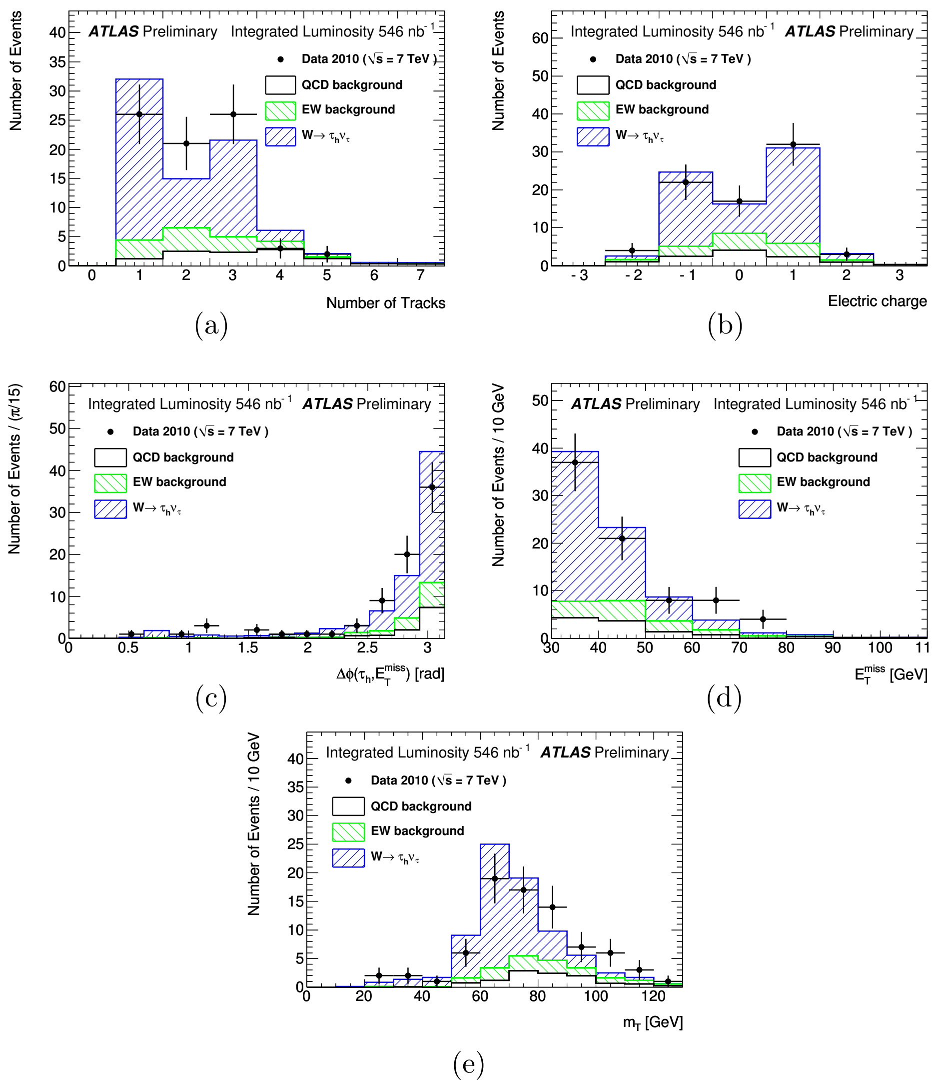

Fig. 6. Distributions of the tau track multiplicity (a), electric charge (b), $\Delta\phi(\tau_{\rm h}, E_{\rm T}^{\rm miss})$ (c), $E_{\rm T}^{\rm miss}$ (d) and transverse mass $m_{\rm T}$ (e) for the data in signal region A, the scaled QCD background from control region C, and the contributions from signal and EW background in region A. The QCD background distribution is normalized to the estimated number of QCD background events in region A ($\rm N^{A}_{QCD}$).

from data to be due to QCD processes. With a remaining background from W and Z decays of $11.8 \pm 0.4_{(stat.)} \pm 3.7_{(syst.)}$ events, estimated from Monte Carlo simulation, this leaves an observed signal of $55.1 \pm 10.5_{(stat.)} \pm 5.2_{(syst.)}$ events. It is compatible with a Standard Model expectation of $55.3 \pm 1.4_{(stat.)} \pm 16.1_{(syst.)}$ events from $W \rightarrow \tau\nu$ decays. This is the first observation of $W \rightarrow \tau\nu$ decays and of hadronically decaying tau leptons in ATLAS.

6. Conclusions

Different tau reconstruction and identification algorithms have been developed by the ATLAS collaboration. During the first period of data taking the focus was on robust performance and understanding the discriminating variables rather than optimal performance. The good agreement between data and Monte Carlo in all identification variables and in background rejection rates motivates the use of more sophisticated multivariate techniques (projected likelihood, boosted decision trees) and more identification variables to improve the tau selection performance.

The first observation in ATLAS of $W \rightarrow \tau\nu$ decays confirms the detector capability to observe hadronic tau decays. An observation of the $Z \rightarrow \tau\tau$ process will be a further confirmation of the ATLAS ability to detect hadronically decaying tau leptons and will be used to further study tau lepton identification at ATLAS.

REFERENCES

[1] ATLAS Collaboration, Expected performance of the ATLAS experiment: detector, trigger and physics, CERN-OPEN-2008-020, 2008.

[2] K. Nakamura et al. (Particle Data Group), JP G 37 075021, 2010.

[3] The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 (2008) S08003.

[4] ATLAS Collaboration, Performance of the ATLAS tau trigger in p-p collisions at $\sqrt{s} = 7TeV$, ATLAS-CONF-2010-090, 2010.

[5] The ATLAS Collaboration, Luminosity Determination Using the ATLAS Detector, ATLAS-CONF-2010-060 (2010).

[6] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP 05 (2006) 026.

[7] R. Field, Min-Bias and Underlying Event at the Tevatron and the LHC, talk at the Fermilab MC Tuning Workshop (Oct 2002).

http://www-cdf.fnal.gov/physics/conferences/cdf8547_RDF_TeV4LHC.pdf

[8] The ATLAS Collaboration, The ATLAS Simulation Infrastructure, Eur.Phys.J.C70:823-874,2010.

[9] The ATLAS Collaboration, ATLAS Monte Carlo tunes for MC09, ATL-PHYS-PUB-2010-002 (2010).

[10] M. Cacciari and G. P. Salam. Phys. Lett. B 641 (2006) 57; M. Cacciari, G. P. Salam and G. Soyez, http://fastjet.fr/.

[11] The ATLAS Collaboration, Calorimeter clustering algorithms: Description and performance, ATL-LARG-PUB-2008-002 (2008).

[12] ATLAS Collaboration, Tau reconstruction and identification performance in ATLAS, ATLAS-CONF-2010-086, 2010.

[13] See talk of A. Zemla, Performance of tau lepton identification in ATLAS 7 TeV data, in these proceedings.

[14] W. Lampl et al., Calorimeter clustering algorithms: description and performance, ATL-LARG-PUB-2008-002, 2008.

[15] M. Cacciari, G. P. Salam and G. Soyez, The anti-kt jet clustering algorithm, JHEP 04 063, 2008.

[16] ATLAS Collaboration, Observation $W \to \tau\nu$ Decays with the ATLAS Experiment, ATLAS-CONF-2010-097, 2010.

[17] ATLAS Collaboration, Electron and photon reconstruction and identification in ATLAS: expected performance at high energy and results at $\sqrt{s} = 900$ GeV, ATLAS-CONF-2010-005.

[18] ATLAS Collaboration, Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Colisions at a Center-of-Mass Energy of 7 TeV with The ATLAS Detector, ATL-CONF-2010-057.