Tau trigger and tau reconstruction, efficiency and fake rates in ATLAS

Yann Coadou[*][†]

CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
E-mail: coadou@cppm.in2p3.fr

Tau leptons will play an important role in the physics program at the LHC. In particular, they provide a useful signature in searches for new phenomena like charged Higgs bosons or Supersymmetry. In addition, they are being used for Standard Model electroweak measurements and for detector related studies such as the determination of the missing transverse energy scale.

Due to the huge background from QCD processes, efficient tau identification techniques with large fake rejection are essential. Tau objects appear as collimated jets with low track multiplicity and single variable criteria are not enough to efficiently separate them from jets and electrons.

We report on the commissioning steps and performance of the tau trigger, which is designed to efficiently reject low-energy jets while keeping a high efficiency with respect to hadronic tau leptons identified by the offline algorithms.

We present the current status of tau reconstruction and identification at the LHC with the ATLAS detector. Reconstructed tau candidates in dijet backgrounds are studied in data and compared with predictions from Monte Carlo simulation. The performance of the fake tau rejection is estimated in a dijet data sample. We discuss the plans for measuring tau identification efficiency using $W \to \tau \nu$ and $Z \to \tau \tau$ signal events. Both cut-based and more advanced multivariate techniques which make optimal use of all the information available are presented. These Standard Model measurements are instrumental in validating tau identification for discovery physics.

---

[*]Speaker.

[†]on behalf of the ATLAS collaboration

1. Introduction

The tau lepton plays an important role in the ATLAS [1] physics programme. The heaviest lepton (with a mass of 1.78 GeV [2]) appears in many final states of physics beyond the Standard Model (SM), for instance in Higgs boson decays $H/A/h \rightarrow \tau\tau$, charged Higgs boson decays $H^{\pm} \rightarrow \tau\nu$, supersymmetric multilepton decays and exotic scenarios [3].

In order to use tau reconstruction as a reliable tool for new physics discovery, it is, however, necessary to first understand the detector and observe well-known SM processes involving tau leptons (such as $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$ and $t\bar{t} \rightarrow \tau + X$) to demonstrate tau identification feasibility and to calibrate its performance.

Tau leptons can decay leptonically to an electron or muon (and associated neutrinos) but such decays are very difficult to distinguish from prompt leptons. In the following, only hadronic decays of tau leptons, which represent about 65% of all decays, will be considered. In the detector, such decays are characterised in the inner tracking system by a small number of collimated tracks (typically one or three, coming from charged pions) with no track activity in an isolation region around the core. A small but sizable lifetime (decay length $c\tau = 87 \ \mu$m [2]) also generates a noticeable impact parameter and transverse flight path. Tau leptons also leave collimated calorimetric energy deposits, often associated to a strong electromagnetic (EM) component (from π^0 in tau decays), with little energy deposition in the isolation cone.

The results presented use pp collisions at the LHC at a centre-of-mass energy of $\sqrt{s} = 7$ TeV collected with the ATLAS detector between March and July 2010, with an integrated luminosity up to 244 nb^{-1}. While the number of true tau leptons in this sample is very small, the performance of the tau algorithms is studied on background jets. The tau trigger is described in Section 2. The offline tau reconstruction and identification algorithms and performance in data and Monte Carlo (MC) are presented in Sections 3 and 4, and first signs of tau leptons in data are reported in Section 5.

2. Tau trigger

The ATLAS trigger system consists of three steps: a hardware-based Level 1 trigger (L1) followed by the software High Level Trigger (HLT), composed of the Level 2 trigger (L2) and Event Filter (EF). The L1 tau trigger uses 0.1×0.1 ($\eta \times \phi$) calorimeter towers, finding the local maximum above some E_T threshold in a 0.2×0.2 region. Outer cells from the 0.4×0.4 region are optionally used for isolation. Regions of interest (RoI) are defined by L1 objects for partial detector readout in the HLT. At L2, tracking information is available and combined with jets made from cells (without noise suppression) in RoI's, in order to build tau identification variables. The algorithm run at EF level is similar to the offline reconstruction, using calorimeter energy clusters with proper calibration and noise suppression applied. At HLT level, the object selection is using rectangular cuts on jet shapes and track properties.

The trigger menu includes a complete set of tau-related triggers to cover the full spectrum of tau physics. Single tau triggers with increasing energy thresholds and identification tightness are used to select events from heavy $H \rightarrow \tau\tau$, $Z' \rightarrow \tau\tau$ or $H^{\pm} \rightarrow \tau\nu$. Ditau triggers enable the recording of events from heavy resonances, while triggers combining taus with other objects are often

required to reduce rates and to minimise trigger biases on tau objects selected for tau performance studies. Such triggers include tau+e/μ to select $Z \to \tau\tau$, $t\bar{t}$ or $H \to \tau\tau$, tau+E_T^{miss} for $W \to \tau\nu$ or $H^\pm \to \tau\nu$, as well as tau+(b)jets for instance for $t\bar{t}$ or SUSY processes.

The single tau trigger efficiency is shown in Fig. 1. A good agreement between data and MC is observed [4]. In the near future the trigger efficiency for real tau leptons will be measured with three approaches: from tau-like QCD jets, using a bootstrap method for high p_T items (computing the efficiency for trigger B with respect to trigger A, assuming the trigger A efficiency is known and all events that trigger B also trigger A), and with the tag-and-probe method (from $Z \to \tau\tau$ using e/μ as the tag, or in $t\bar{t}$ triggering on four jets and measuring the tau+E_T^{miss} trigger efficiency).
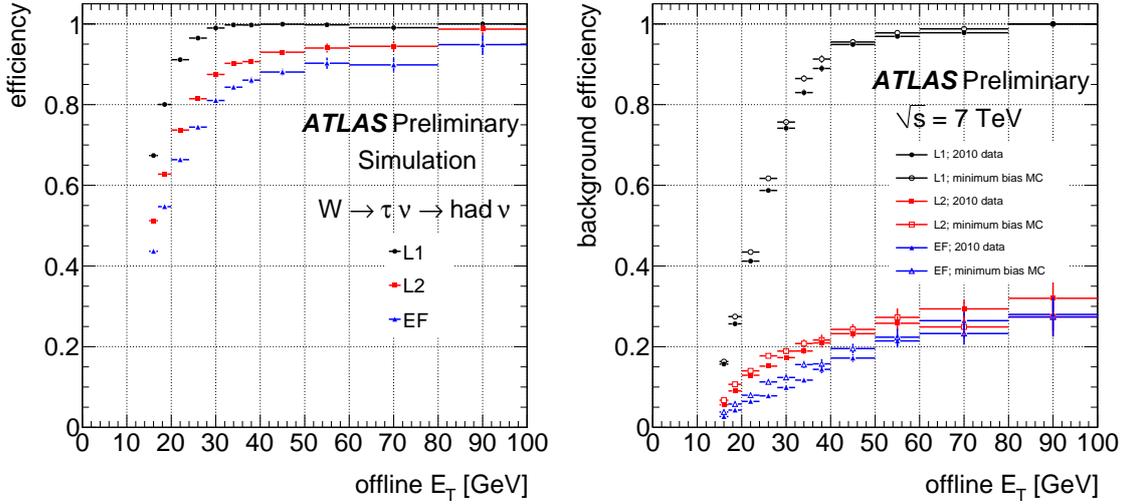

Figure 1: Fraction of reconstructed tau candidates (no ID applied) passing L1 (5 GeV), L2 (7 GeV) and EF (12 GeV) loose trigger conditions as a function of E_T of the offline candidate. Left: signal efficiency on $W \to \tau\nu$ MC; right: data-MC comparison on minimum bias background.

3. Tau reconstruction

Tau lepton decays are reconstructed starting from either a calorimeter seed (calorimeter jet reconstructed from topological clusters [5] with the anti-k_t algorithm [6] using a distance parameter $R = 0.4$, with $p_T > 10$ GeV) or a track seed (good quality track with $p_T > 6$ GeV). Starting from a track seed, tracks with $p_T > 1$ GeV are collected around it in a cone $\Delta R = 0.2$[1]. Jet seeds within $\Delta R = 0.2$ are also associated to this candidate. The transverse energy is computed from calorimetry only (with MC-derived correction factors to calibrate it to the tau visible p_T, defined as the vector sum of all tau decay products except neutrinos) and with an energy flow algorithm that combines tracking and calorimetric information. Such candidates are labelled double-seeded. The remaining jet seeds (with associated looser-quality tracks) define calo-seeded candidates. Only few candidates are track-seeded only. In the following, only double-seeded and calo-seeded candidates are used.

---

[1]$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ where η is the pseudorapidity and ϕ is the azimuthal angle.

Figure 2 shows the good agreement between data and MC (using the PYTHIA DW tune [7]) for the p_T and pseudorapidity spectra of tau candidates [8]. The reconstructed candidates' intrinsic properties like track multiplicity and number of topoclusters are also well reproduced.
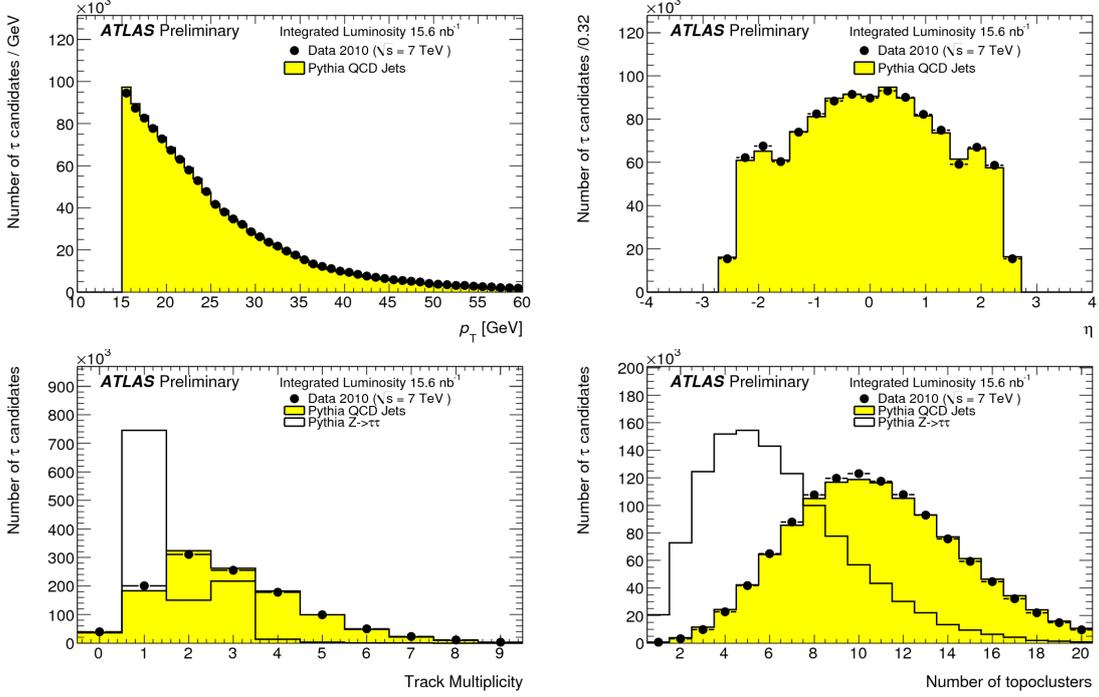

Figure 2: Data-MC comparison of the transverse momentum (top left), pseudorapidity (top right), track multiplicity (bottom left) and number of topoclusters (bottom right) of tau candidates in QCD jets. The open histograms on the bottom plots show the expected shape (arbitrary unit) from real tau decays in $Z \to \tau\tau$ events for comparison.

4. Tau identification

The tau reconstruction algorithm does not provide much rejection against jets. A separate identification (ID) step is necessary, based on discriminating variables. Tau leptons are difficult to identify and 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 presented [9]. The optimisation was done for 30% (tight), 50% (medium) and 60% (loose) signal efficiency, separately for candidates with one track and at least two tracks.

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). The first two are shown in Fig. 3. The multivariate techniques use, in addition, the invariant mass of topoclusters, the invariant mass of the track system, the fraction of E_T in a $\Delta R = 0.1$ cone and the fraction of E_T in the EM calorimeter. All variables show good data-MC agreement and discriminating power.
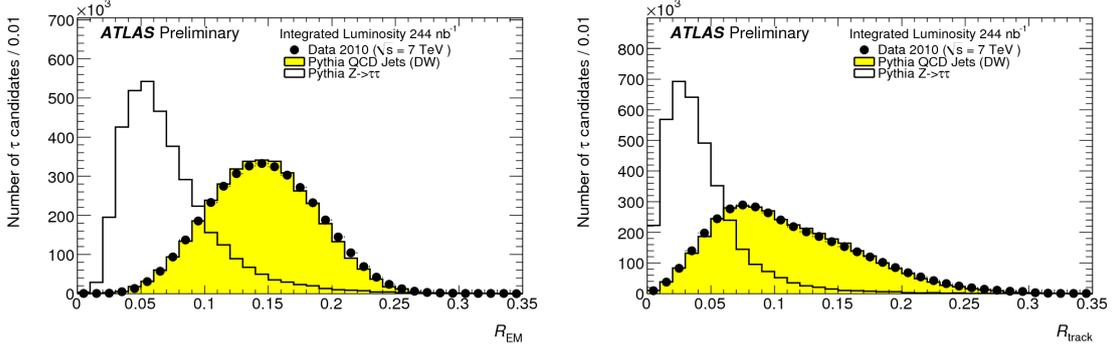
Figure 3: EM radius (left) and track radius (right) for data (black dots), MC QCD jets (yellow histogram) and MC signal from $Z \to \tau\tau$ (open histogram).

The ID performance is evaluated using the signal efficiency $\varepsilon_{\text{sig}} = N^{\tau}_{\text{pass,match}}/N^{\tau}_{\text{match}}$, where $N^{\tau}_{\text{(pass,)match}}$ is the number of tau candidates (that pass the ID criteria) within $\Delta R = 0.2$ of a true hadronically decaying tau with visible $p_{\text{T}} > 15$ GeV and visible $|\eta| < 2.5$, and the background efficiency $\varepsilon_{\text{bkgd}} = N^{\text{bkgd}}_{\text{pass}}/N^{\text{bkgd}}_{\text{total}}$, where $N^{\text{bkgd}}_{\text{total(pass)}}$ is the number of tau candidates (that pass the ID criteria).

The signal and background efficiencies for the loose, medium and tight settings of the cut-based ID are shown in Fig. 4 (top) as a function of p_{T}. The agreement between data and MC is reasonable. Values integrated over the full p_{T} spectrum for data and MC are summarised in Table 1, also reporting $\varepsilon'_{\text{bkgd}}$, which is similar to $\varepsilon_{\text{bkgd}}$ but also requires that candidates have exactly one or three tracks. Systematic effects from transverse momentum calibration and pile-up were estimated to be 2.1–9.6% and 5.7–14.5%, respectively, depending on p_{T}.

Multivariate discriminant performance was also assessed. The performance of cuts, likelihood and BDT is shown in Fig. 4 (bottom) for the medium efficiency setting. There is a clear improvement coming from these techniques, which will be used in future physics measurements.

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

Table 1: Cut-based background efficiency for data and MC for the loose, medium and tight settings, without ($\varepsilon_{\text{bkgd}}$) and with ($\varepsilon'_{\text{bkgd}}$) requiring exactly one or three tracks.

5. Finding real tau leptons in data

At the time of the workshop, there was not yet a claim of observation of $W \to \tau\nu$ or $Z \to \tau\tau$ decays in ATLAS. By now the first observation of $W \to \tau\nu$ events and hadronic tau decays was reported [10]. It corresponds to 55 signal events (55.3 expected) out of 78 candidate events in data, in 546 nb^{-1} of integrated luminosity, with good data-MC agreement as shown in Fig. 5. In the near future real tau leptons from W decays will be used to validate MC efficiency measurements. With
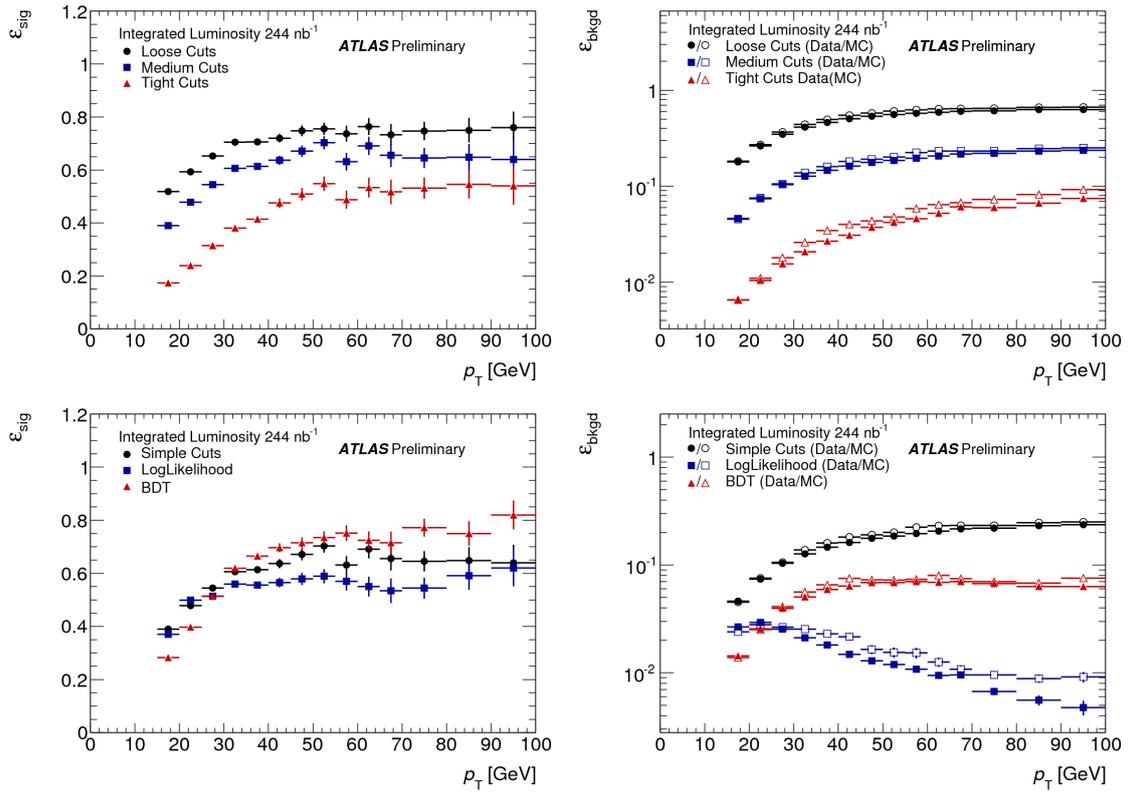
Figure 4: Signal (left) and background (right) efficiencies in data (background only) and MC as a function of p_T. Top: cut-based ID for loose (black circles), medium (blue squares) and tight (red triangles) settings. Bottom: cut-based (black circles), likelihood (blue squares) and BDT (red triangles) comparison for the medium efficiency setting.

more data, the $Z \rightarrow \tau\tau$ channel will soon be observed and studied, allowing to measure trigger and ID efficiencies in an unbiased way, as well as giving a handle on tau energy scale and E_T^{miss} scale [3].
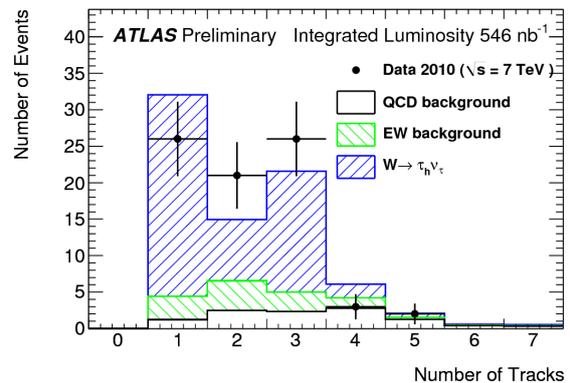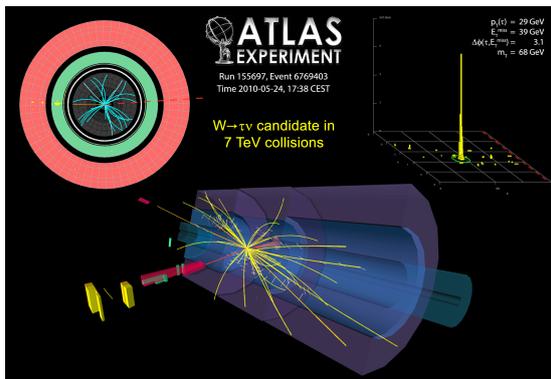

Figure 5: Left: event display of a $W \rightarrow \tau\nu$ candidate. Right: Track multiplicity of tau candidates after selecting $W \rightarrow \tau\nu$ events.

6. Conclusion

The ATLAS collaboration has developed a full suite of tau reconstruction and identification algorithms. With first data the focus was on robust performance rather than optimal rejection, but the good agreement between data and Monte Carlo predictions in all identification variables for background jets, as well as in fake rejection rates, motivates the use of more powerful techniques like likelihood and boosted decision trees. These multivariate techniques show much increased performance and will soon benefit from the introduction of new powerful variables. The first observation of $W \to \tau \nu$ decays confirms the capabilities of the experiment and more data in the $W \to \tau \nu$ and $Z \to \tau \tau$ channels will soon allow to fully validate tau identification in data, bringing it to a stage when it can be used to tackle the full ATLAS physics programme.

References

[1] G. Aad *et al.* (ATLAS Collaboration), *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3**, S08003, 2008.

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

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

[4] ATLAS Collaboration, *Performance of the ATLAS tau trigger in p-p collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-090, 2010; https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ TriggerPublicResults?rev=109#Tau_trigger_performance_in_7_TeV

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

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

[7] M. G. Albrow *et al.* (TeV4LHC QCD Working Group), *Tevatron-for-LHC report of the QCD Working Group*, arXiv:hep-ph/0610012.

[8] ATLAS Collaboration, *Reconstruction of hadronic tau candidates in QCD events at ATLAS with 7 TeV proton-proton collisions*, ATLAS-CONF-2010-059, 2010.

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

[10] ATLAS Collaboration, *Observation of $W \to \tau \nu$ decays with the ATLAS experiment*, ATLAS-CONF-2010-097, 2010.