Draft version 0.01

ATLAS CSC Note

July 31, 2007

Tau overall performances: comparison of various algorithms for tau identification and reconstruction

edited by E. Richter-Was and J. Tanaka

Abstract

This note discusses ATLAS detector overall performances for identification of true hadronically decaying taus and rejection against candidates build from jets or around electron or muon tracks in the very large dynamic range of transverse momenta, spanning from 10-15 GeV to 500 GeV. Tau candidates will be reconstructed as a narrow calorimetric shower with strong electromagnetic component and low tracks multiplicity.

Reviewed are base-line algorithms for hadronic tau reconstruction, calibration and identification implemented in ATLAS software framework *Athena*: calorimeter seeded and track seeded options of *tauRec* package and also summarised are additional developments. The reference performance figures for efficiencies and rejections against jets and electron and muon tracks are given based on the CSC simulated data from different signal and background samples. Briefly discussed is status of parametrisations for fast simulation and summarised is expected performance of the HLT tau trigger.

In the last part of the Note outlined are prospects for tau identification and calibration with early data of $100 pb^{-1}$ and tau leptons produced in decays of W and Z bosons.

1 Tau's in early data

1.1 Tau's in $t\bar{t}$ events

(contribution from: S. Cabrera-Urbán, S.M. Demers, and M. T. Pérez Garcia-Estañ)

1. Motivation.

   The decay chain $t\bar{t} \to W(e/\mu, \nu_e/\nu_\mu)W(\tau_{had}, \nu_\tau)b\bar{b}$ in the first 100 pb^{-1} of ATLAS data is interesting for both its physics potential and the possibility of using the channel for understanding tau identification. Due to the increased center of mass energy available at the LHC the cross section for $t\bar{t}$ production increases by nearly two orders of magnitude over what is available at the TeVatron. This results in sufficient statistics for observation and a cross-section measurement with a relatively small amount of integrated luminosity. While the low yeild of events will result in a larger cross-section uncertainty than with other $t\bar{t}$ channels, this sample has a different sensitivity to new physics and beyond-the-Standard Model effects in top quark decays. The decay chain is also important to understand because it is a background in many SUSY and Higgs analyses. In the case of an MSSM Charged Higgs, for example, an excees of the final state could be evidence for $t\bar{t} \to W(e/\mu, \nu_e/\nu_\mu)H^{\pm}(\tau_{had}, \nu_\tau)b\bar{b}$.

   The decay chain $t\bar{t} \to W(e/\mu, \nu_e/\nu_\mu)W(\tau_{had}, \nu_\tau)b\bar{b}$ in the first 100 pb^{-1} of ATLAS data is interesting for both its physics potential and the possibility of using the channel for understanding tau identification. Due to the increased center of mass energy available at the LHC the cross section for $t\bar{t}$ production increases by nearly two orders of magnitude over what is available at the TeVatron. This results in sufficient statistics for observation and a cross-section measurement with a relatively small amount of integrated luminosity. While the low yeild of events will result in a larger cross-section uncertainty than with other $t\bar{t}$ channels, this sample has a different sensitivity to new physics and beyond-the-Standard Model effects in top quark decays. The decay chain is also important to understand because it is a background in many SUSY and Higgs analyses. In the case of an MSSM Charged Higgs, for example, an excees of the final state could be evidence for $t\bar{t} \to W(e/\mu, \nu_e/\nu_\mu)H^{\pm}(\tau_{had}, \nu_\tau)b\bar{b}$.

   The goals for early tau physics at ATLAS include aquiring a sample of taus from the data with as high a purity as is possible so the tau identification efficiency can be measured and the simulation can be tuned. The optimal channel in which to perform this analysis is $W \to \tau\nu_\tau$ due to high statistics and the possibility for an excellent signal-to-background ratio. $Z \to \tau\tau$ is also a useful channel due to the high cross-section and potential for an excellent signal-to-background ratio. We study this $t\bar{t}$ channel with an eye toward supplementing the taus expected from Ws and Zs. In spite of a lower cross-section and likely worse signal-to-background ratio, there are some benefits to this channel. The events can be triggered on with high efficiency, even at high luminosity, with a single 20 GeV electron or muon trigger. If necessary, the event jet multiplicity can be taken advantage of in the trigger. Unile the case with Ws and (potentially) Zs, a tau trigger will not be required, leaving the resulting sample of taus less biased. Note that the standard model cross-section for the signal would need to be assumed to use this final state to study tau identification efficiency.

   This analysis closely follows the philosophy of early $t\bar{t}$ analyses at CDF [**?**] [**?**], though many of the selection values used here are higher due to the higher center of mass energy available at the LHC.

2. Event selection strategy.

   The following event selection is performed:

- In the $e-\tau$ case only: An electron with $P_T > 20$ GeV, IsEM=0, tracking isolation: $\Sigma P_T(tracks, \Delta R < 0.2) < 5$ GeV, and calorimeter isolation.
- In the $\mu-\tau$ case only: A muon with positive MuonID, tracking isolation: $\Sigma P_T(track, < 0.2) < 5$ GeV, and $\chi^2 > 0$.
- A reconstructed tau with visible $P_T > 15$ GeV and the tau1p3p discriminant equal to 1.
- E_T of the first jet > 50 GeV. Figure 1 shows the jet E_T spectrum in signal and background events.
- E_T of the second jet > 35 GeV. Figure 1 shows the jet E_T spectrum in signal and background events.
- $\not\!\!E_T > 25$ GeV. Figure 2 shows the $\not\!\!E_T$ spectrum in signal and background events.
- Event $H_T > 250$ GeV, where H_T is defined as the sum of the E_T of the electron(muon), P_T of the tau, E_T of both jets, and $\not\!\!E_T$. Figure 2 shows the H_T spectrum in signal and background events.
- The event is required to not have the electron(muon) and tau reconstructed to the mass of the Z when the event matches the kinematics of $Z \to \tau\tau + jets$ events, with the $\not\!\!E_T$ falling between the electron(muon) and the tau.
- the electron(muon) and the tau must have opposite charge.

	N_{evt} $(100\ pb^{-1})$					
	$t\bar{t}(e\tau_h)$	$W(e\nu_e)+3p$	$Z(\tau_e\tau_h)+2p$	$t\bar{t}(\mu\tau_h)$	$W(\mu\nu_\mu)+3p$	$Z(\tau_\mu\tau_h)+2p$
	1193 ± 20	12390 ± 123	741.9 ± 7.5	1213 ± 11	6468 ± 52	764.2 ± 7.6
1 e/μ $P_T > 20$ *GeV*	602 ± 14	5560 ± 82	250.5 ± 4.4	783 ± 9	3711 ± 39	377.4 ± 5.3
1 τ $P_T > 15$ *GeV*	232 ± 9	1827 ± 47	98.0 ± 2.7	327 ± 6	1298 ± 23	148.6 ± 3.4
Tau ID	83 ± 5	377 ± 21	49.5 ± 1.9	124 ± 4	226 ± 10	69.6 ± 2.3
$E_T^{1st\ jet} > 50$ *GeV*	77 ± 5	297 ± 19	21.7 ± 1.3	119 ± 4	169 ± 8	30.8 ± 1.5
$E_T^{2nd\ jet} > 30$ *GeV*	73 ± 5	269 ± 18	14.6 ± 1.1	112 ± 3	140 ± 8	20.4 ± 1.2
$\not\!\!E_T > 25$	65 ± 5	214 ± 16	6.9 ± 0.7	101 ± 3	113 ± 7	9.1 ± 0.8
$H_T > 250$ *GeV*	65 ± 5	200 ± 16	5.5 ± 0.6	99 ± 3	98 ± 6	5.5 ± 0.6
Z veto	64 ± 5	198 ± 16	3.6 ± 0.5	95 ± 3	95 ± 6	3.6 ± 0.5
Opposite Charge	55 ± 4	86 ± 10	2.2 ± 0.4	84 ± 3	45 ± 4	1.9 ± 0.4
b-tagging	44 ± 4	33 ± 6	2.2 ± 0.4	68 ± 3	19 ± 3	1.9 ± 0.4

Table 1: The number of events remaining after each event selection requirement in the left column is shown for both signal ($t\bar{t}$) and background samples.

The effect of each of these requirements on the signal and major backgrounds are shown in Table 1. The largest decreases in signal efficiency come from the electron(muon) and tau selection. These requirements also significantly reduce the backgrounds.

3. B-tagging.

   We can take advantage of the presence of two b quarks in the decay chain by requiring at least one of the jets in the event to be tagged as a b-jet. While this significantly improves the signal-to-background ratio, as can be seen in Table 1, the efficiency of one of the jets being tagged as a b-jet is on the order of 80%, so we lose a significant amount of signal. There is also the possibility that b-tagging will not be available during the very early days of ATLAS analysis. For these two reasons it would be preferable to not need to rely on b-tags in the analysis with the first 100 pb^{-1}.

	N_{evt} $(100\ pb^{-1})$							
	$t\bar{t}(e\tau_{h:1-\text{prong}})$	$t\bar{t}(e\tau_{h:1-\text{prong}})$	$t\bar{t}(\mu\tau_{h:3-\text{prong}})$	$t\bar{t}(\mu\tau_{h:3-\text{prong}})$	$t\bar{t}(\tau_e\tau_h)$	$t\bar{t}(\tau_\mu\tau_h)$	$Acc_(\tau_e\tau_h)$	$Acc_(\tau_\mu\tau_h)$
	930.0 ± 17.4	923.1 ± 17.3	68.4 ± 4.7	95.1 ± 5.6	194.9 ± 8.0	194.9 ± 8.0	0.16 ± 0.01	0.16 ± 0.01
1 e/μ $P_T > 20$ *GeV*	519.1 ± 13.0	654.3 ± 14.6	35.8 ± 3.4	66.1 ± 4.6	47.6 ± 3.9	62.2 ± 4.5	0.08 ± 0.01	0.08 ± 0.01
1 τ $P_T > 15$ *GeV*	204.0 ± 8.2	277.9 ± 9.5	15.6 ± 2.3	24.1 ± 2.8	12.4 ± 2.0	24.8 ± 2.8	0.05 ± 0.02	0.08 ± 0.02
Tau ID	71.0 ± 4.8	100.0 ± 5.7	7.5 ± 1.6	12.1 ± 2.0	4.9 ± 1.3	12.1 ± 2.0	0.06 ± 0.03	0.10 ± 0.03
$E_T^{1st\ jet} > 50$ *GeV*	65.2 ± 4.6	95.5 ± 5.6	7.2 ± 1.5	11.4 ± 1.9	4.9 ± 1.3	11.7 ± 2.0	0.06 ± 0.03	0.10 ± 0.03
$E_T^{2nd\ jet} > 30$ *GeV*	61.9 ± 4.5	89.9 ± 5.4	6.8 ± 1.5	11.4 ± 1.9	4.6 ± 1.2	11.1 ± 1.9	0.06 ± 0.03	0.10 ± 0.03
$\not{E}_T > 25$	53.8 ± 4.2	80.8 ± 5.1	6.2 ± 1.4	10.1 ± 1.8	4.6 ± 1.2	9.8 ± 1.8	0.07 ± 0.03	0.10 ± 0.03
$H_T > 250$ *GeV*	54.4 ± 4.2	79.5 ± 5.1	6.2 ± 1.4	10.1 ± 1.8	4.6 ± 1.2	9.4 ± 1.8	0.07 ± 0.03	0.10 ± 0.03
Z veto	53.1 ± 4.2	76.6 ± 5.0	6.2 ± 1.4	10.1 ± 1.8	4.6 ± 1.2	8.8 ± 1.7	0.07 ± 0.03	0.09 ± 0.03
Opposite Charge	47.6 ± 3.9	70.7 ± 4.8	4.2 ± 1.2	6.8 ± 1.5	3.3 ± 1.0	6.2 ± 1.4	0.06 ± 0.03	0.07 ± 0.03
b-tagging	38.1 ± 3.5	57.3 ± 4.3	3.9 ± 1.1	5.9 ± 1.4	2.3 ± 0.9	4.9 ± 1.3	0.05 ± 0.03	0.07 ± 0.03

Table 2: The number of events remaining after each event selection requirement in the left column is shown for signal ($t\bar{t}$) samples distinguishing the different contributions: 1,3 prong and ditau leptonic channels and its relative acceptances .

The b-tag is currently implemented using impact paramenter and secondary vertex information combined in weights available in the AOD. For the distribution of jet weights in the signal and background events see Figure 3.

4. CSC samples used.

$t\bar{t}$ signal	trig1_misal1_mc12.005200.T1_McAtNlo_Jimmy.recon.AOD.v12000601
$W \to e\nu_e + \geq 3j$	trig1_misal1_mc12.008241.AlpgenJimmyWenuNp3_pt20_filt3jet.recon.AOD.v12000601
$W \to \mu\nu_\mu + \geq 3j$	trig1_misal1_mc12.008245.AlpgenJimmyWmunuNp3_pt20_filt3jet.recon.AOD.v12000601
$Z \to \tau\tau + \geq 2j$	trig1_misal1_mc12.008156.AlpgenJimmyZtautauNp2LooseCut.recon.AOD.v12000601

Table 3: The physics samples and corresponding Monte Carlo Datasets used in this study are shown.

The Monte Carlo samples used for this analysis are all CSC mc12 samples. Details are given in Table 3.

5. Conclusions and results.

In the first 100 pb^{-1} of ATLAS data we expect to have 55 ± 2 signal events in the $e - \tau$ channel, with a signal-to-background ratio (S:B) of 1:1.6. If we require a b-tag the expected number of signal events decreases to 44 ± 2 but S:B improves to 1:0.8. In the $\mu - \tau$ channel we expect 84 ± 3 signal events with S:B of 1:0.6. When a b-tag is required the number of signal events expected drops to 68 ± 3 and S:B improves to 1:0.3.

6. Prospects this be a place-holder for the time being till end of august The current background estimates for the $W \to e\nu + 3j$ and $W \to \mu\nu + 3j$ rely on the jet to tau fake rate in the monte carlo. For the actual analysis we plan to use a data-driven approach using non-trigger-biased jets from dijet events as is done in the current CDF analysis. [?]. The fake rate can be parametrized with the jet transverse energy, the event jet multiplicity, and the sum of the E_T deposited in the event. This study can currently be performed on dijet monte carlo.

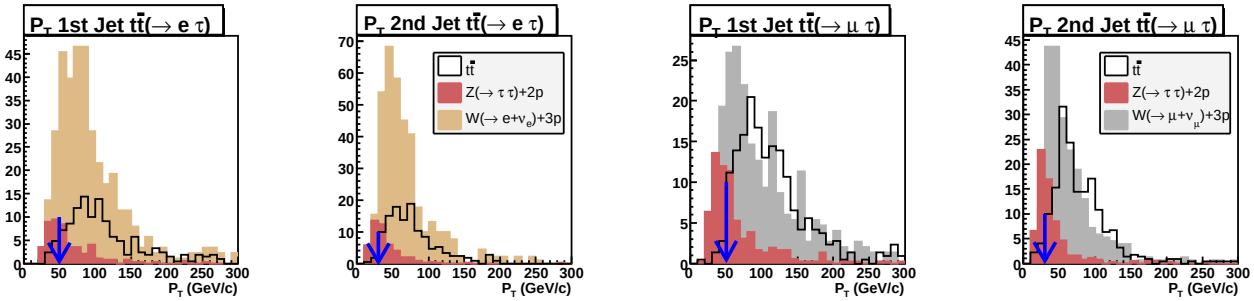

Figure 1: The upper left plot shows the E_T spectrum of the first jet in the signal and background events for the $e-\tau$ channel. The upper right plot shows the E_T spectrum of the second jet in the signal and background events for the $e-\tau$ channel. The lower left plot shows the E_T spectrum of the first jet in the signal and background events for the $\mu-\tau$ channel. The lower right plot shows the E_T spectrum of the second jet in the signal and background events for the $\mu-\tau$ channel.

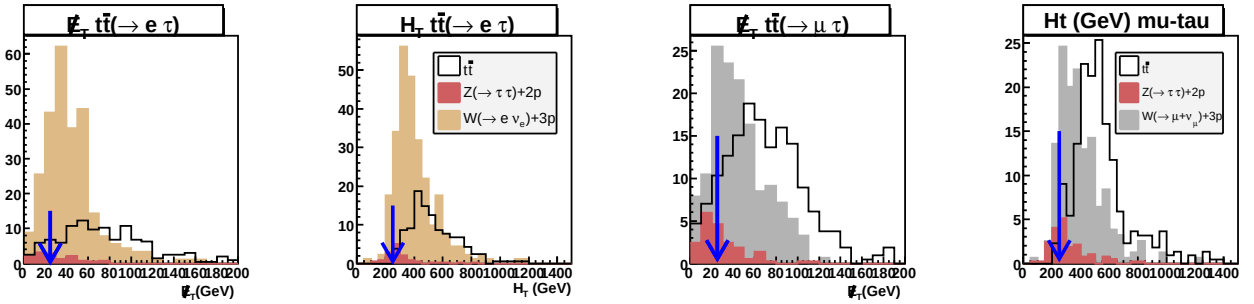

Figure 2: The top two plots show the $\not\!\!E_T$ spectrum in signal and background events for the $e-\tau$ channel (on the left) and $\mu-\tau$ channel (on the right). The bottom two plots show the H_T spectrum in signal and background events for the $e-\tau$ channel (on the left) and $\mu-\tau$ channel (on the right).

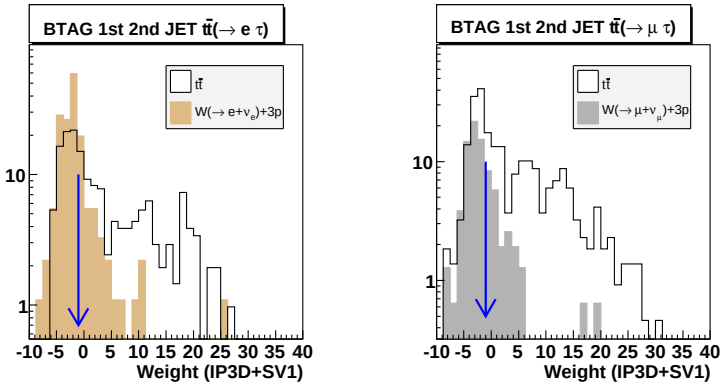

Figure 3: The plots show the b-tagging weights of the first two jets in the signal and W-background events for the $e-\tau$ channel (on the left) and $\mu-\tau$ channel (on the right.) The cut value of -1 is indicated with the arrows.