Measurement of hadronic tau identification efficiency using $W \rightarrow \tau\nu$ events

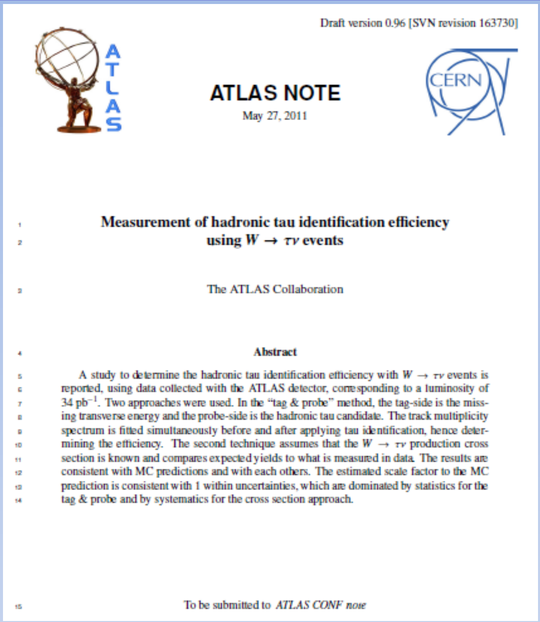
Group approval meeting

1st June,2011

Y. Coadou, K. Hanawa, K. Hara, S. Lai, K. Nakamura,
S. Protopopescu and S. Tsuno
on behalf of the tau performance group

Introduction

- This note describes "Measurement of hadronic tau identification efficiency using $W \to \tau\nu$".
  - **-CONF draft : ATLAS-COM-CONF-2011-085**
    http://cdsweb.cern.ch/record/1349546
  - **-Supporting COM draft : ATL-COM-PHYS-2011-476**
    http://cdsweb.cern.ch/record/1349543



Draft version 0.96 [SVN revision 163730]

ATLAS NOTE

May 27, 2011

Measurement of hadronic tau identification efficiency using $W \to \tau\nu$ events

The ATLAS Collaboration

Abstract

A study to determine the hadronic tau identification efficiency with $W \to \tau\nu$ events is reported, using data collected with the ATLAS detector, corresponding to a luminosity of 34 pb^{-1}. Two approaches were used. In the "tag & probe" method, the tag-side is the missing transverse energy and the probe-side is the hadronic tau candidate. The track multiplicity spectrum is fitted simultaneously before and after applying tau identification, hence determining the efficiency. The second technique assumes that the $W \to \tau\nu$ production cross section is known and compares expected yields to what is measured in data. The results are consistent with MC predictions and with each others. The estimated scale factor to the MC prediction is consistent with 1 within uncertainties, which are dominated by statistics for the tag & probe and by systematics for the cross section approach.

To be submitted to *ATLAS CONF note*

- There are two methods to evaluate tau identification efficiency and scale factor (SF) :
  - ☐ Tag and Probe method.
  - ☐ Cross section method.
- This note shows the result of tau ID SF in six ID working point.

	1-prong	Multi-prong
Looser(CUT/LLH)	Loose	Medium
Tighter(CUT/LLH)	Medium	Tight
Looser(BDT)	Loose	Loose
Tighter(BDT)	Medium	Medium

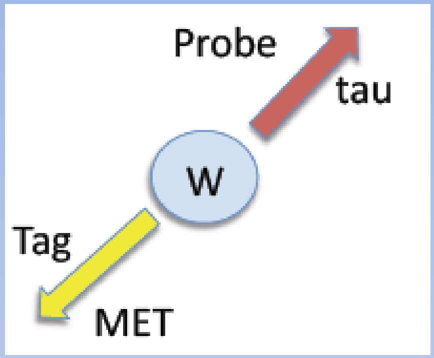
Typical efficiency from MC

Method	*e*-veto	Efficiency
Looser cuts	medium	0.74
Tighter cuts	tight	0.57
Looser LLH	medium	0.80
Tighter LLH	tight	0.63
Looser BDT	medium	0.74
Tighter BDT	tight	0.60

Description of each method

Tag and probe method:

•The efficiency is defined by: $$Eff. = \frac{N^{tau}(AfterID)}{N^{tau}(BeforeID)}$$



•The events "after ID" are always sub-set of those "before ID".

•Key point is how we can collect as many tau candidates as possible "before ID".

Cross section method:

•Assume the W production cross section (from data, lepton universality) to compare the MC acceptance with DATA.

•The deviation is quoted as the "Scale Factor".

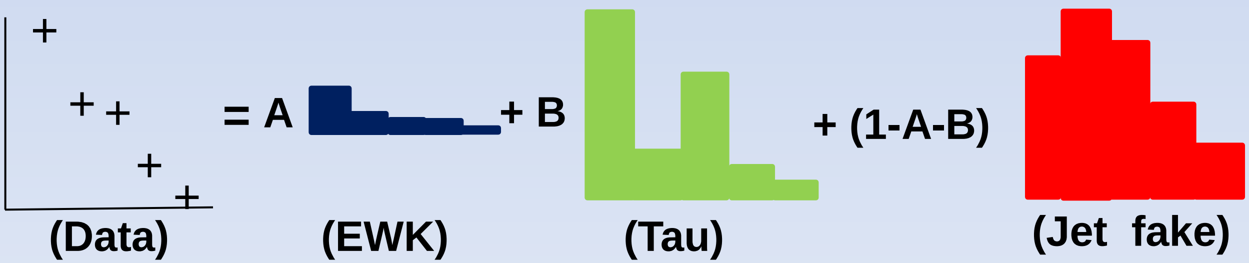
$$SF = \frac{N^{tau}(Data)}{N^{tau}(MC)}$$

•The key point is how we can control the "Acceptance".

Fitting to extract the number of taus from data:

Use three templates for tau ,EWK and jet fake.

Normalization is different in each method , described later.

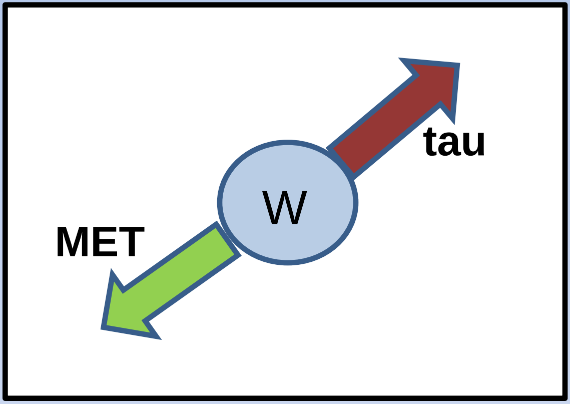


Event selection

Trigger
To avoid trigger bias
・Missing Et Trigger
Tag&Probe: all MET trigger
X-Sec : EF_xe30(40)_noMu(un-prescaled).

Lepton veto
-To reduce W→e/mu v
・veto events if they have at lease one lepton(e with Pt>20GeV/mu with Pt>15GeV)

MET side
To suppress QCD di-jet
・MET>30(40)GeV
・MT(MET,tau)
・MET significance
Tag&Probe:MET/sqrt{0.5(SumET)}
X-Sec :MET/sqrt(SumPt)



Tau side
-Tau candidate Pt>20GeV
-Tau ID is taken at several working points.

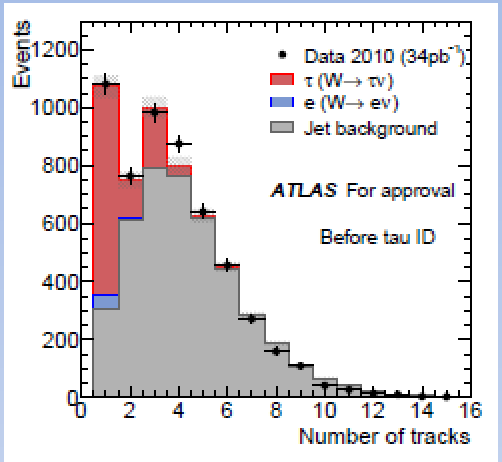
Remaining background after these selections:
•W→enu
•jet (from W+jets /QCD di-jets)

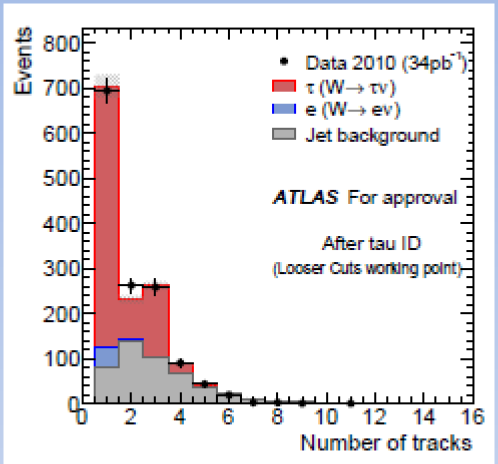
Fitting method to determine efficiency/SF

- Number of tau in both methods are extracted by fitting the track multiplicity .

Tag & probe method :

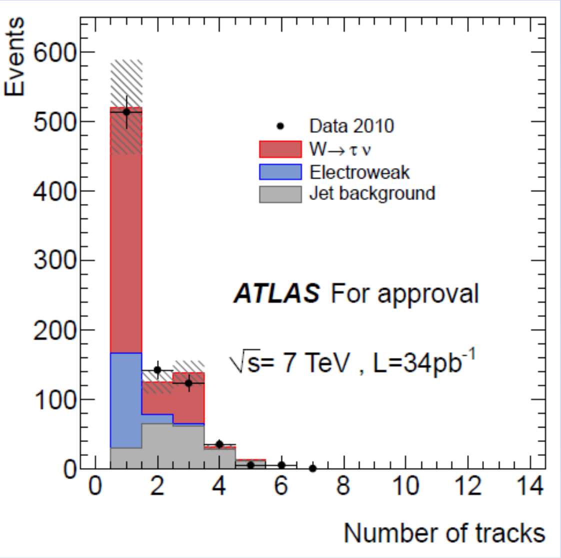
•Fit twice "before" and "after" ID to obtain N_{tau}(after/beforeID).
•Two parameter fitting using f_{signal} and $f_{electron}$
•Statistical error is dominated by the contribution of QCD "before ID".





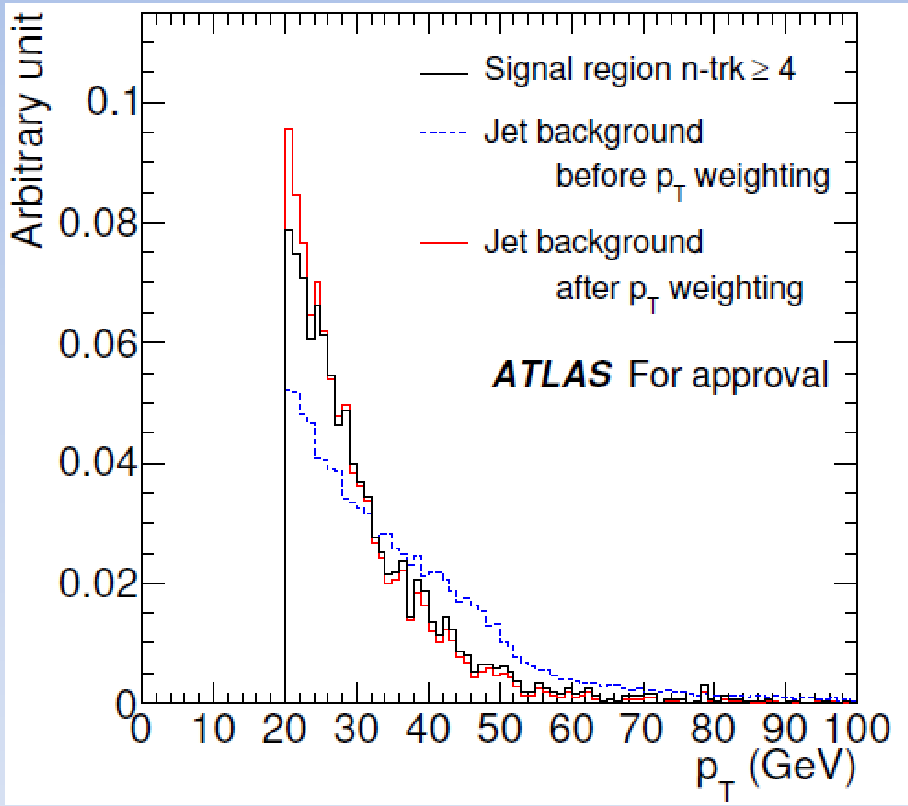
Cross section method :

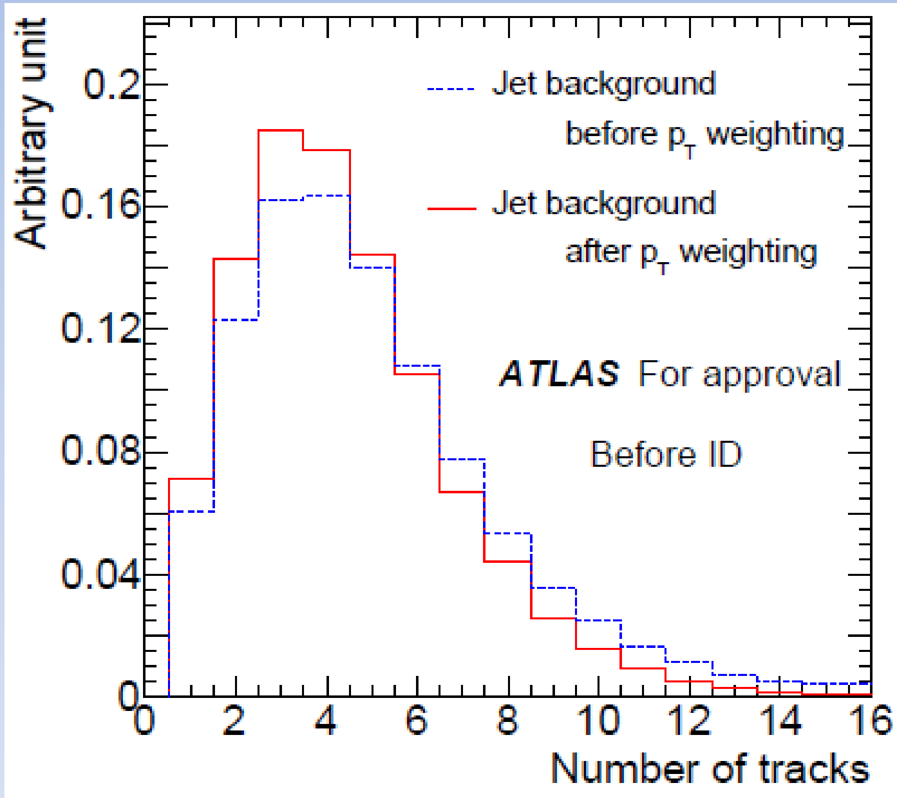
•Fit just one time "after" ID to obtain N_{tau} after ID.
•One parameter fitting using f_{signal} ,where normalization for EWK is MC prediction.
•Statistical error will be smaller than T&P method , while suffer from systematic uncertainty on acceptance determination.



Jet background template for Tag&Probe

- Jet background template for Tag&Probe is created by the shape in low MET significance region. (2< MET significance <4.5)
- The shape is reweighted by pt spectrum to correct the pt difference between CR and SR.

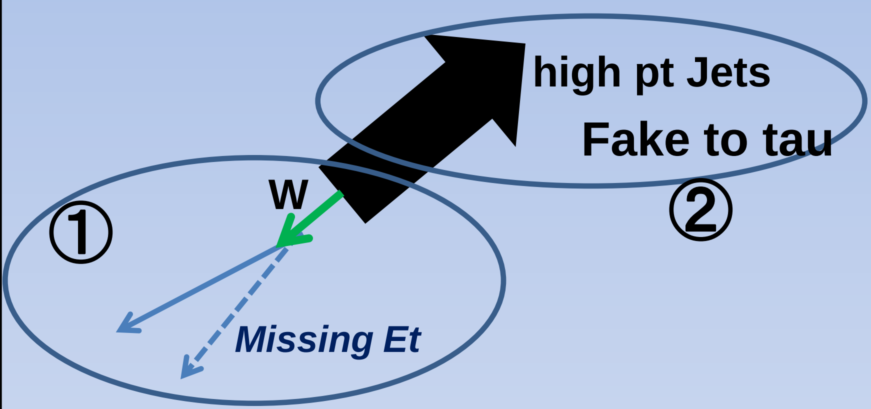




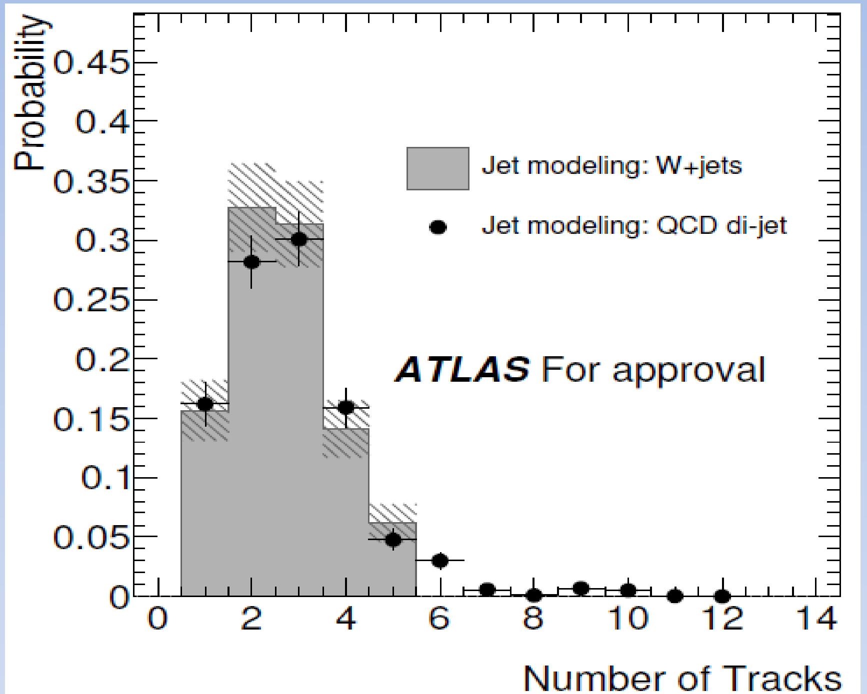
Jet background template for Cross section

•Two models of jet background for cross check.
1.extracted from W→eν +jets events (for centeral value),
selected by single lepton trigger and same selection as SR.

W→eν+jet modeling



①**Require single electron trigger ,electron and MET to obtain W→eν.**
②**use the track multiplicity obtained by fake tau not overlapping electron in this event.**



These model are in good agreement!

2.One is similar to Tag&Probe ,pt reweighting (for systematics).
CR : data in low Mt region and
SR : fake tau of MC in signal region.

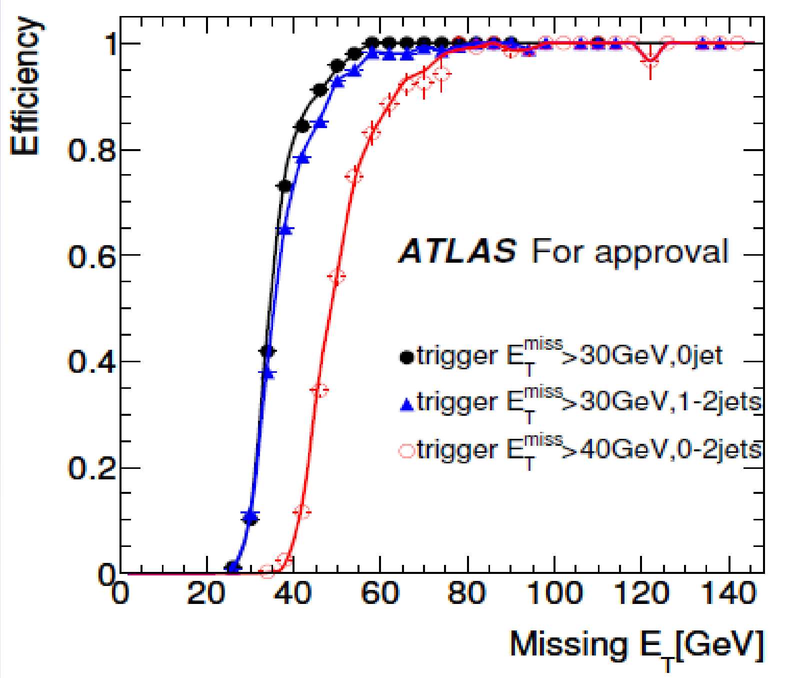
Trigger efficiency for cross section method

Event weight is applied as a weight to the W→τν MC instead of trigger simulation.

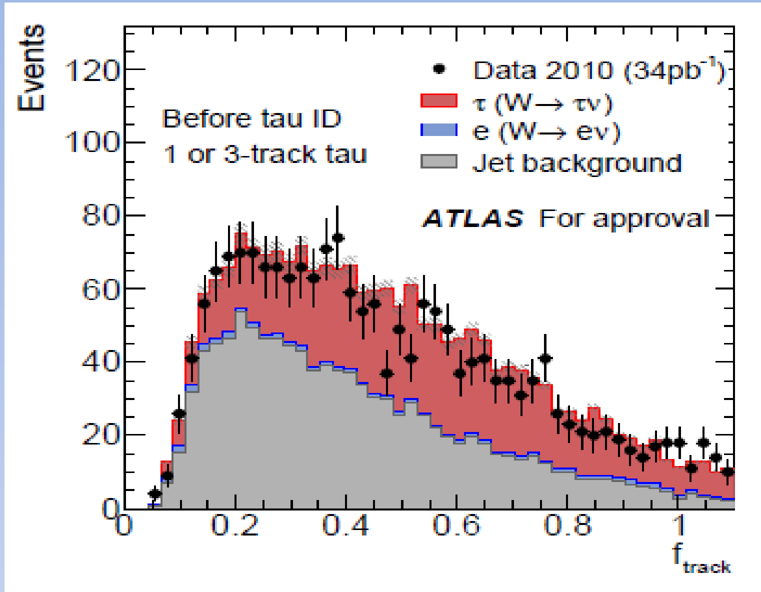
$$\text{EventWeight} = \underbrace{data(W \to e\nu)}_{①} \underbrace{\frac{MC(W \to \tau\nu)}{MC(W \to e\nu)}}_{②}$$

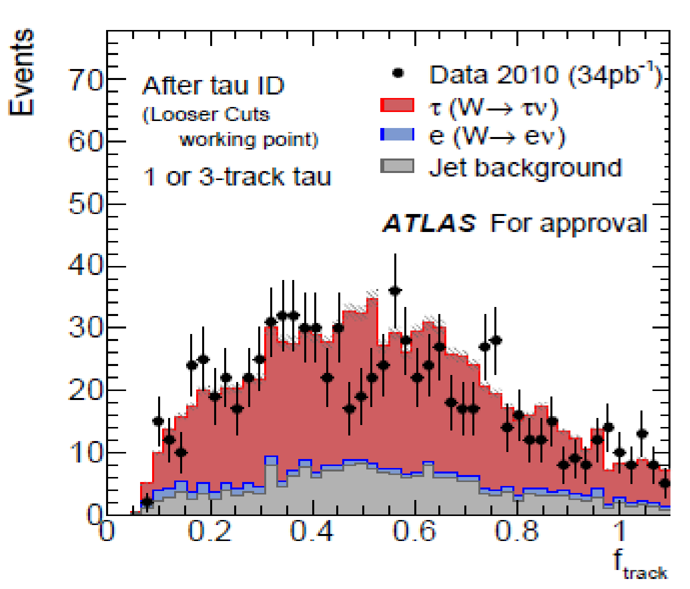
①Trigger efficiency is extracted from W→eν event.

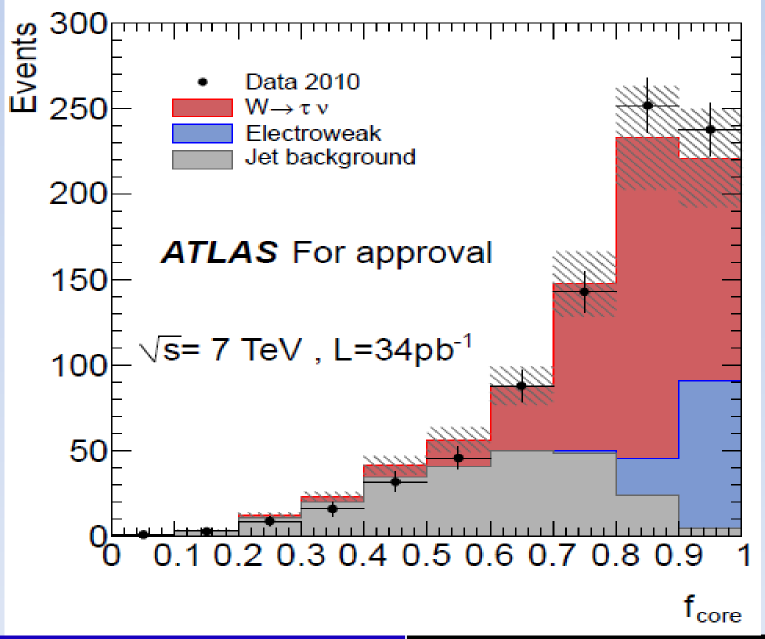
②To consider difference between W→τν and W→eν, we apply correction with the ratio of the efficiency in W→eν and W→τν MC samples.



Verification for full model

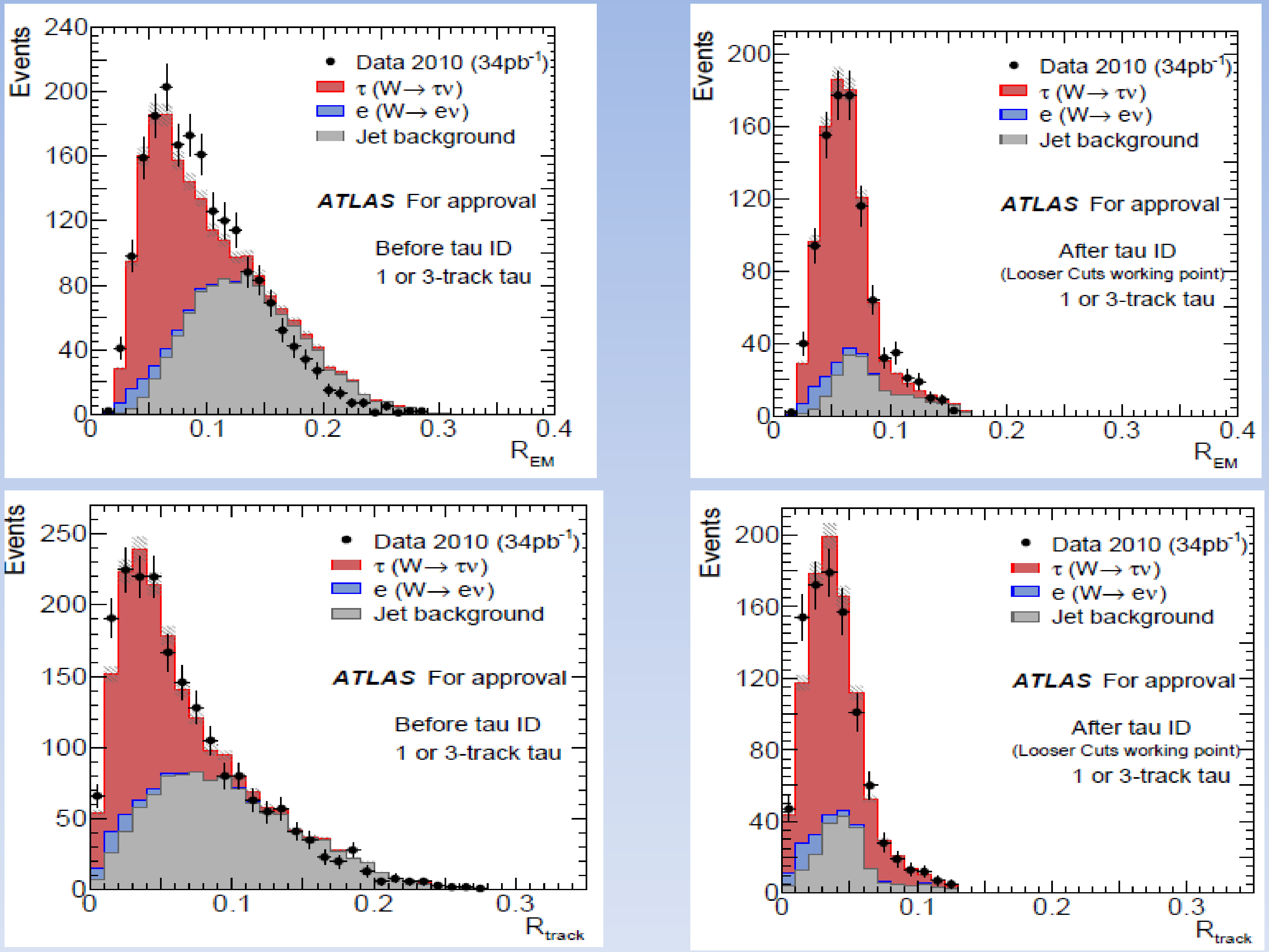






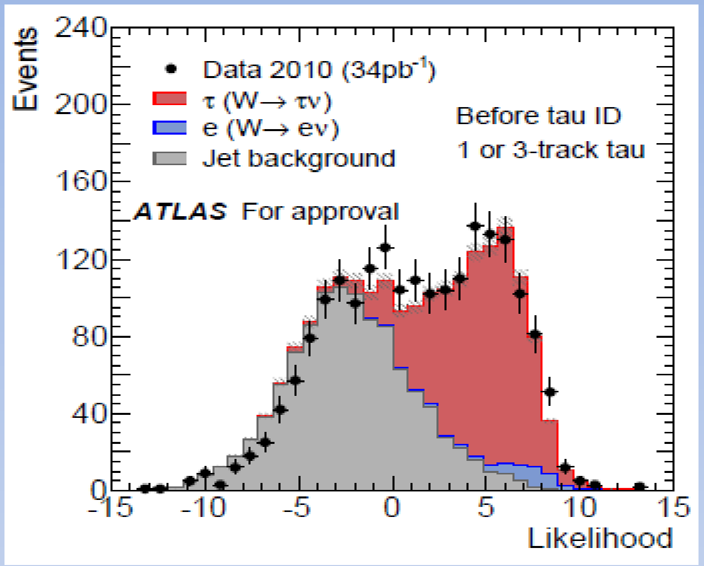
- Identification variables are well modeled
- The different contributions are normalized to their respective number of events as measured by the fit.

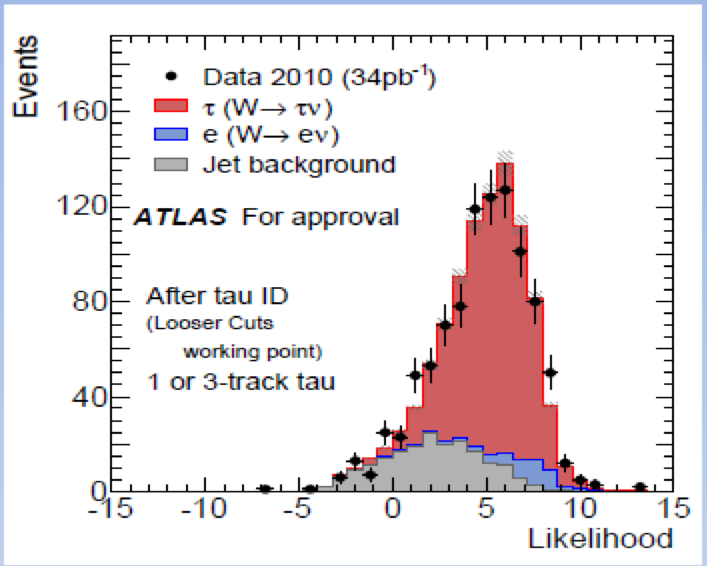
Other variables

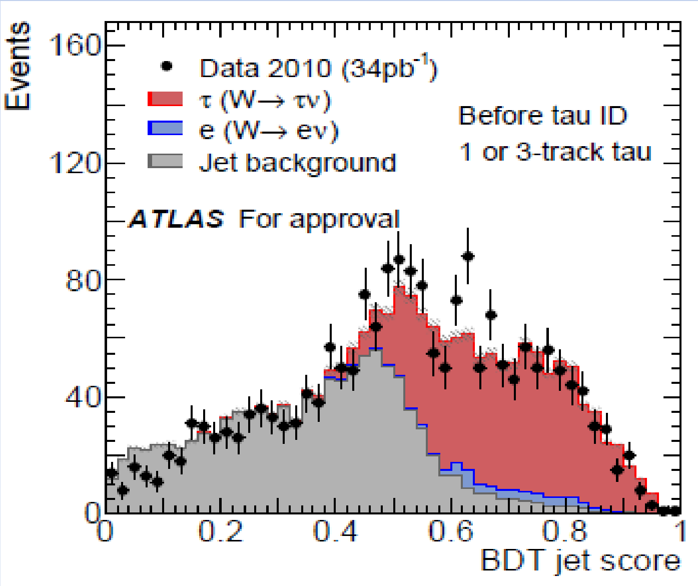


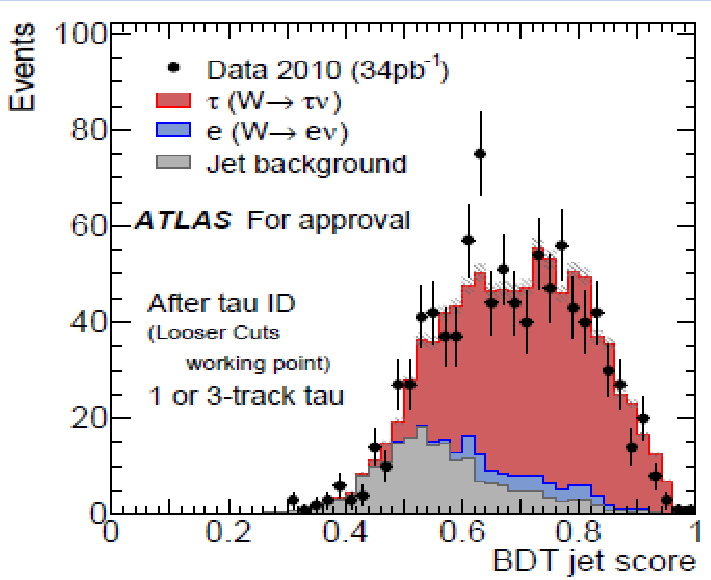
Identification variables are well modeled .

LLH and BDT score









Likelihood and BDT score are also well modeled.

Systematic uncertainties

- The difference in the fit results with nominal fit is taken as a systematic uncertainty.

Tag & probe method :

•The “kinematic (acceptance)” origin systematics are canceled out. That is, no syst. by JES etc.

•QCD modeling and internal shower structure would be the systematics.

E_T^{miss} trigger	0.7%
Jet modelling (p_T-weighting)	0.4%
Jet modelling ($S_{E_T^{miss}}$)	0.6%
Electron fake rate	1.6%
Pileup condition	1.4%
Shower model	2.6%
Detector geometry	0.9%
Underlying event	1.3%
Total systematic uncertainty	3.7%

Cross section method :

•Tau energy scale is dominant source.
(this already inclusively contains the tau internal shower structure.)

•Different event topology (kinematical origin syst.) also comes as “UE uncert”.

Jet modelling	1.1%
W cross section	5.1%
Trigger efficiency	2.7%
Electron fake rate	4.8%
Tau energy scale	9.0%
Jet energy scale	0.1%
Electron energy scale	0.8%
Pileup	0.2%
Underlying event	6.8%
Total systematic uncertainty	13.6%

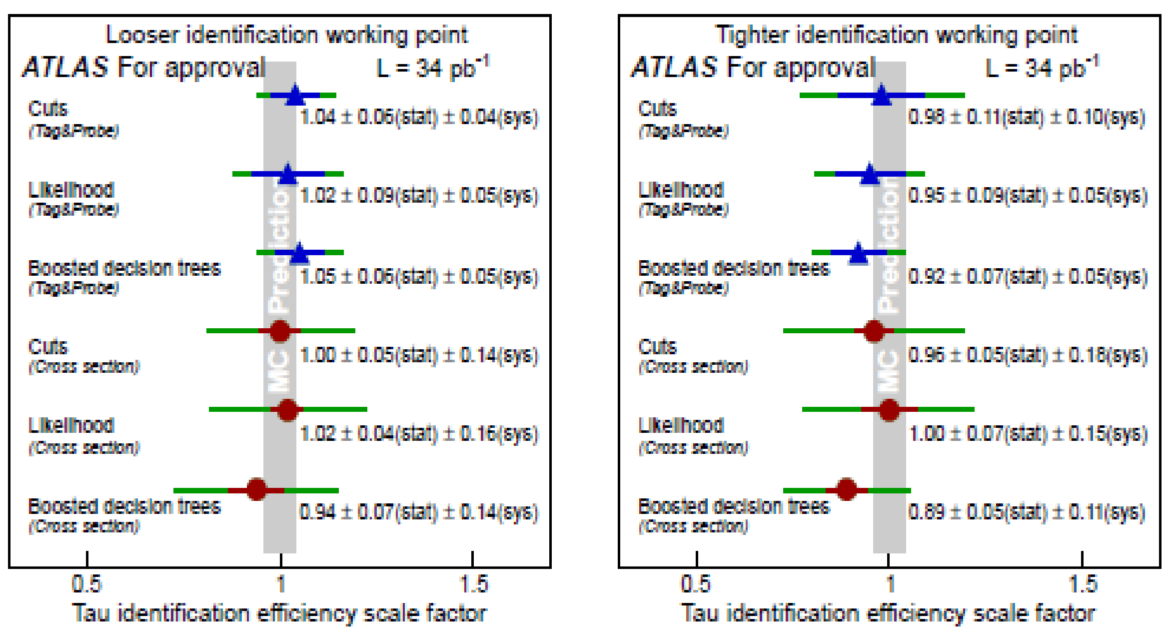
Results

Tag&Probe

Method	Syst.	Efficiency	Scale factor
Looser cuts	3.7%	0.77 ± 0.05 ± 0.03	1.04 ± 0.06 ± 0.04
Tighter cuts	9.9%	0.56 ± 0.06 ± 0.06	0.98 ± 0.11 ± 0.10
Looser LLH	5.0%	0.82 ± 0.07 ± 0.04	1.02 ± 0.09 ± 0.05
Tighter LLH	5.7%	0.60 ± 0.06 ± 0.03	0.95 ± 0.09 ± 0.05
Looser BDT	4.4%	0.78 ± 0.05 ± 0.03	1.05 ± 0.06 ± 0.05
Tighter BDT	5.1%	0.55 ± 0.04 ± 0.03	0.92 ± 0.07 ± 0.05

Cross section

Method	Scale factor
Looser cuts	1.00 ± 0.05 ± 0.14
Tighter cuts	0.96 ± 0.05 ± 0.18
Looser LLH	1.02 ± 0.04 ± 0.16
Tighter LLH	1.00 ± 0.07 ± 0.15
Looser BDT	0.94 ± 0.07 ± 0.14
Tighter BDT	0.89 ± 0.05 ± 0.11



T&P method

- large stat. error + small syst.

Xsec method

- small stat. error + large syst.

Overall, both methods provide “similar size of error (stat+syst)”.

Main questions during reviews

Q:Why is the systematic uncertainty of CUT tighter larger than other working points in Tag&Probe method ?
A:cut-based ID more sensitive to shifts in single variable data-MC comparisons. See R_{EM} on page 10.

Q:Why is the systematics of X-sec method larger than Tag&Probe one ?
A:Because of Tau energy scale uncertainty, this uncertainty is cancelled in Tag&Probe method.

Q:How about JES for Tag&Probe ?
A: It was expected to cancel in the ratio. Verified it: the effect is tiny ,the difference is 0.06 %.

Q:Why is the SF central value scatter so small compared to the uncertainties ?
A:Because all working points and methods are correlated strongly due to using same variables for TauID.

More questions

Q:Why is the ratio 1 track to 3track large in X-sec method ?
A:1tracks increase while 3tracks decrease after track-based met significance because we favor small sumPt.

Q:Why do we use track-based met significance ?
A: To avoid large variation of SumEt.

Q: Isn't the mT selection (tau candidate leading to mT closest to 65 GeV) in T&P biasing your sample?
A: No, because the tau template is using only truth matched candidates.

Q:Size of tau energy scale uncertainty
A:known to be overestimated, but only documented one until last week.
→could we apply the new tau energy scale ?

Conclusion

- We have completed the analysis of hadronic tau identification efficiency on 2010 data.
- We obtained that the TauID SF is almost equal to 1.
- We can believe TauID efficiency and MC prediction.
- This is first measurement of TauID efficiency with ATLAS.
- We have had EdBoard meeting .
- We have addressed all questions from the EdBoard so far.

Back up

Event selection detail

- Both methods are required similar selection .
  - **→MET ,Mt and METsignificance (to suppress QCD di-jet)**
  - **→Lepton veto (to suppress W→lv)**
- But there is some difference due to the difference of each methods.

Tag & probe method :

E_T^{miss} trigger	Use all E_T^{miss} triggers, irrespective of prescales
Event cleaning	Good data quality and at least one primary vertex with ≥ 4 tracks
E_T^{miss}	$E_T^{miss} > 30$ GeV
$\Delta\phi(E_T^{miss}, \mathrm{jet})$	$\Delta\phi(E_T^{miss}, \mathrm{jet}) \geq 0.7$
e/μ veto	Reject electrons with $p_T > 20$ GeV and muons with $p_T > 15$ GeV
E_T^{miss} significance	$S_{E_T^{miss}} \geq 6\ \mathrm{GeV}^{1/2}$
Tau candidate	$p_T > 20$ GeV, leading track $p_T > 2.4$ GeV, m_T closest to 65 GeV
Transverse mass	$m_T \leq 80$ GeV

Different point
→Since the efficiency is defined as the sub-set of events "after ID",there should not cause any bias in the event selection. (relative acceptance)
•Use multiple MET triggers without caring the prescale factors.

Cross section method :

Event cleaning	Good data quality and at least one primary vertex with ≥ 4 tracks		
e/μ veto	Reject electrons with $p_T > 20$ GeV and muons with $p_T > 15$ GeV		
Tau candidate	$p_T > 20$ GeV, $\lvert\eta\rvert < 1.3$ or $1.6 < \lvert\eta\rvert < 2.5$, Exactly one candidate (keep smallest m_T if necessary)		
E_T^{miss} trigger	$E_T^{miss} > 30$ GeV		$E_T^{miss} > 40$ GeV
E_T^{miss}	$E_T^{miss} > 30$ GeV		$E_T^{miss} > 40$ GeV
Jet multiplicity	0	1 or 2	0–2
$\Delta\phi(E_T^{miss}, \mathrm{jet})$	—	$\Delta\phi \geq 0.5$	$\Delta\phi \geq 0.5$
E_T^{miss} significance	$S^{vtx}_{E_T^{miss}} \geq 6\ \mathrm{GeV}^{1/2}$	$S^{vtx}_{E_T^{miss}} \geq 7\ \mathrm{GeV}^{1/2}$	$S^{vtx}_{E_T^{miss}} \geq 8\ \mathrm{GeV}^{1/2}$
Transverse mass	$60 < m_T < 100$ GeV	$30 < m_T < 90$ GeV	$30 < m_T < 80$ GeV

Different point
→How to avoid the possible systematics variation in the selection. (need to estimate the absolute acceptance)
•Use un-prescale MET trigger. The data set is divided by period-by-period.
•The event selection is optimized by trigger-by-trigger.
•Use "track-based MET significance" to avoid large variation of SumEt.