Identification of Hadronic Decays of τ Leptons

2012 Data with the ATLAS Detector

Wouter Van De Pontseele & Renato Pani
Tutor: Elzbieta Richter-Was

HASCO 2014

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

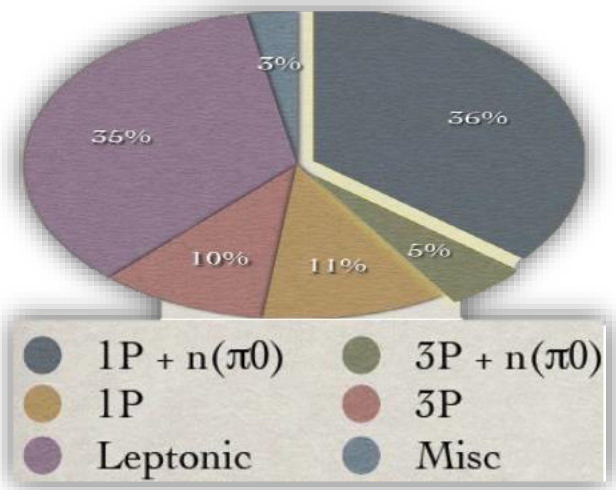
Contents

LEPTONS
ν_e
ELECTRON-NEUTRINO
ν_μ
MUON-NEUTRINO
ν_τ
TAU-NEUTRINO
e^-
ELECTRON
μ
MUON
τ
TAU

Introduction

- τ leptons are the heaviest known leptons, $m_\tau = 1.777\ GeV$. Their decay length is 87μm ⟶ identification trough reconstruction

- Only lepton with hadronic decay channels
  - Leptonic ≈ 35%
  - Hadronic ≈ 65%

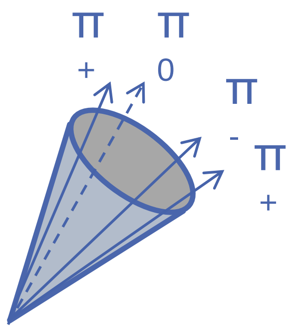


- Leptonic mode is not considered for τ identification cause they cannot be distinguished from prompt electrons or muons ⟶ they do not contain any information specific to the τ!

- Hadronic mode consists predominantly of pions, rarely of kaons
  - One track: 1-prong decays
  - Three tracks: 3-prong decays

Reconstruction of Hadronic τ Decays

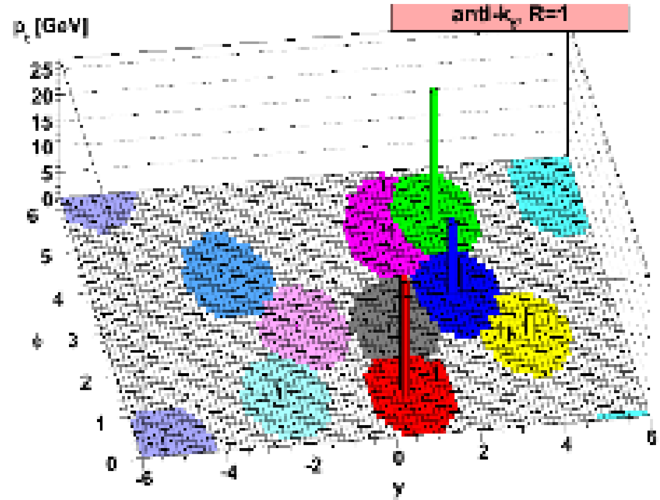
- Distinguish between hadronic τ decays and gluon- or quark-initiated jets
- 8 TeV pp collisions, ATLAS detector
- Use combined data from many sub-detectors
  - p_T, η depends on detector
  - EM and hadronic calorimeter to reconstruct tracks
- Each jet is a τ candidate at the start
  - Select on kinematic quantities
  - Group tracks and clusters with anti-k_t
  - Select collimated jets (R is small)



Example 3-prong jet

Reconstruction Methods: Anti-k_t algorithm

- The Anti k_t algorithm is the most common and used algorithm for jets reconstruction.



- It is based upon the definition of a distance between particles and a distance between jets. Through this the (inclusive) clustering proceeds by identifying the smallest of the distances and if it is a d_{ij} recombining entities i and j, while if it is d_{iB} calling i a jet and removing it from the list of entities.

$$d_{i,j} = min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,i}^2}\right)\frac{\Delta R^2}{D^2}$$

$$d_{i,Beam} = \frac{1}{p_{T,i}^2}$$

Identification of Hadronic τ Decays

The main problem in the identification of these hadronic a decays is separate the tracks and signals we'd like to study from the background which is composed by:

- gluon- or quark- initiated jets
- electrons that can mimick the characteristic 1-prong signature

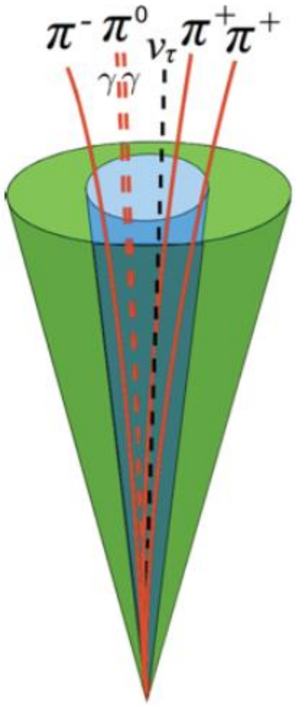
In ATLAS τ jets are reconstructed considering each jet-object as a τ candidate, and then they're discriminated using variables given by the calorimeter and tracking informations.

Then these variables are combined into multi-variate discriminants to reject fake signals

Statistical Methods: Multi-variate determinants

For the identification of hadronically decaying τ leptons, cut-based and multi-variate techniques are used. Specifically are used:

- Projective likelihood method (LLH)
- Boosted decision trees. (BDT)



Three categories of candidates are defined, requiring one reconstructed track (1-prong), three reconstructed tracks (3-prong) and two or three reconstructed tracks (multi-prong).

Statistical Methods: Multi-variate determinants

Different variables are used to applied different statistical methods.

Not always variables used for the 1-prong analysis are good for the 3-prong analysis

Variable	LLH tau ID		BDT tau ID		e-veto	muon veto
	1-prong	3-prong	1-prong	3-prong	1-prong	1-prong
$f_{\mathrm{core}}^{\mathrm{corr}}$	•	•	•	•	•	
$f_{\mathrm{track}}^{\mathrm{corr}}$	•	•	•	•	•	
f_{track}					•	•
R_{track}	•	•	•	•	•	
$S_{\mathrm{lead\ track}}$	•		•			
$N_{\mathrm{track}}^{\mathrm{iso}}$	•		•			
ΔR_{max}		•		•		
$S_{\mathrm{T}}^{\mathrm{flight}}$		•		•		
m_{tracks}		•		•		
f_{EM}					•	•
f_{HT}					•	
$E_{\mathrm{T,max}}^{\mathrm{strip}}$					•	
$f_{\mathrm{HCAL}}^{\mathrm{leadtrk}}$					•	
$f_{\mathrm{ECAL}}^{\mathrm{leadtrk}}$					•	
f_{PS}					•	
$f_{\mathrm{EM}}^{\pi^{\pm}}$					•	
f_{iso}					•	
R_{Had}					•	

Statistical Methods: Likelihood

The projective likelihood method generates a function based on the products between the probability density functions of each variable.

The same set of variables as for the BDT ID is used to create one-dimensional probability density functions (PDFs)

Three working points: ***loose, medium, tight*** , are defined, corresponding to target efficiencies of 70%, 60% and 40% for 1-prong and 65%, 55% and 35% for multi-prong $\tau_{\text{had-vis}}$ candidates, respectively.

$$
\begin{aligned}
L(x_1, x_2, x_3, \cdots, x_N; a) &= P(x_1; a)P(x_2; a)\cdots P(x_N; a) \\
&= \prod P(x_i; a)
\end{aligned}
$$

Statistical Methods: Boosted Decision Trees

- Decision trees:
  - Easy to understand & Implement
  - Binary
  - Computationally simple

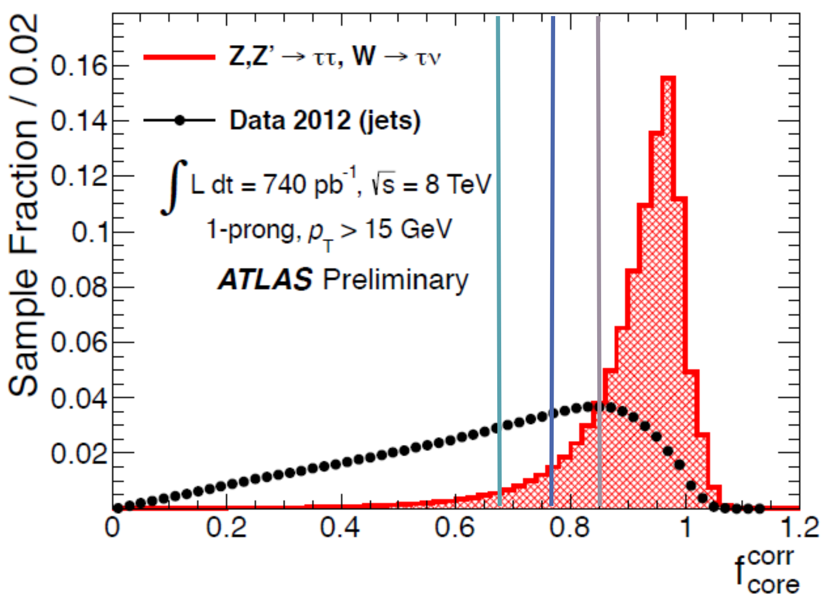
  …But, not al decision trees are equally large,
  choosing the most efficient one is hard, proven to be NP complete!

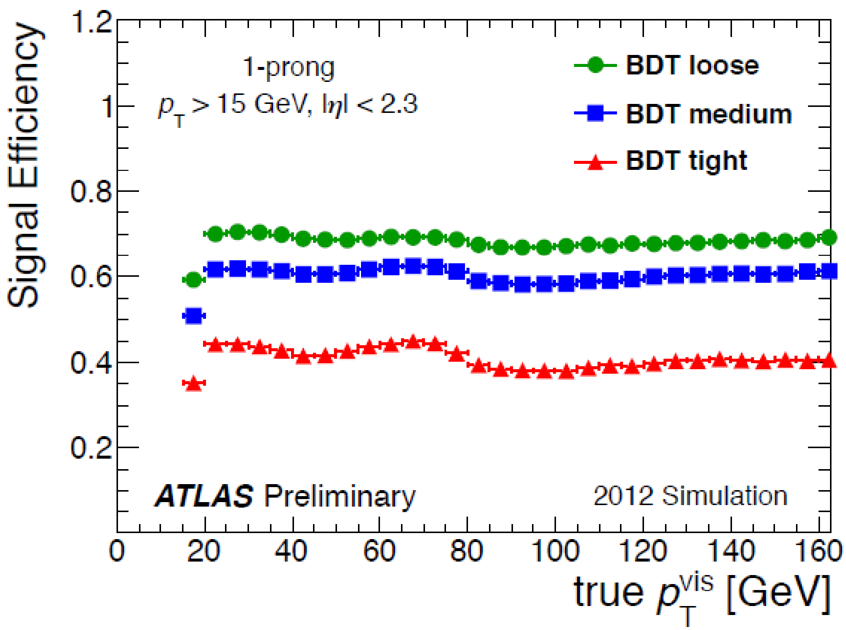
- Boosted:

  Method in which you do not remove measurements but give them a certain weight at each split. This also involves an iteration method to make the tree more efficient.

Statistical Methods: Boosted Decision Trees

Pile-up-corrected core energy fraction: Pile-up-corrected fraction of transverse energy in the central region (#R < 0.1) of the τ candidate:

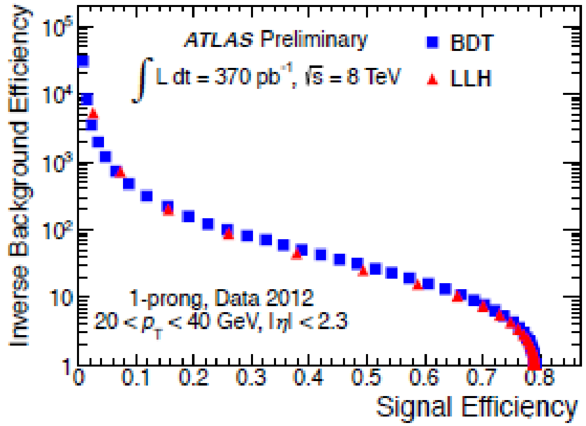


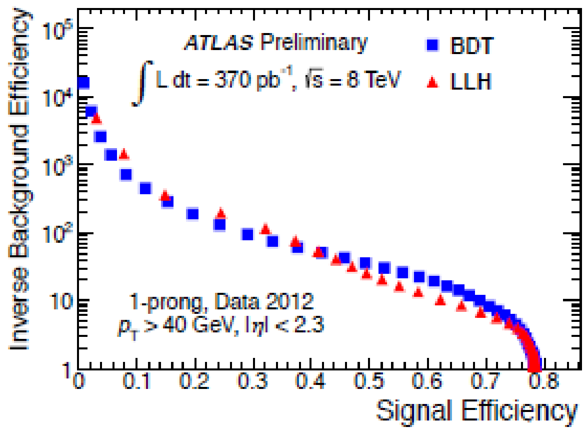


Statistical Methods: LLH & BDT Comparison

Both are stable methods and delivered good results, but at high signal efficiency, the boosted decision trees had less background.

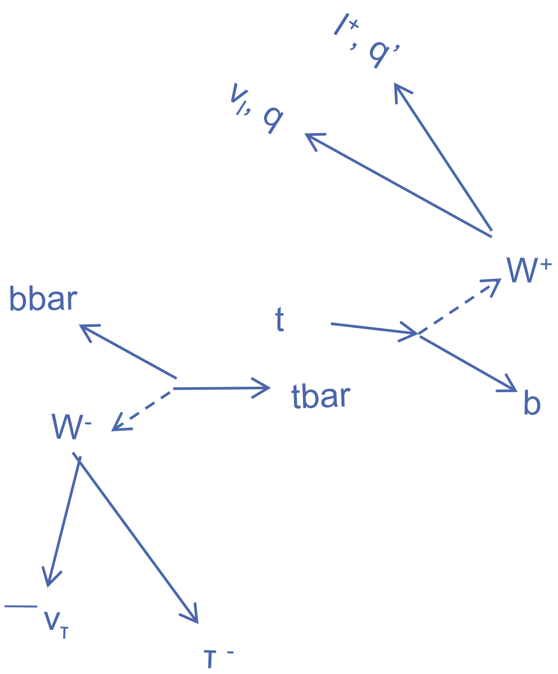
$\longrightarrow$ BDT used for further analysis



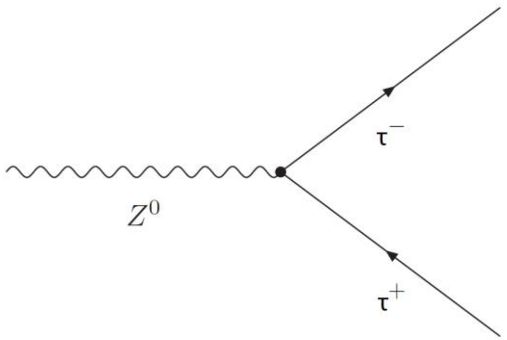


Sources leading to τ leptons

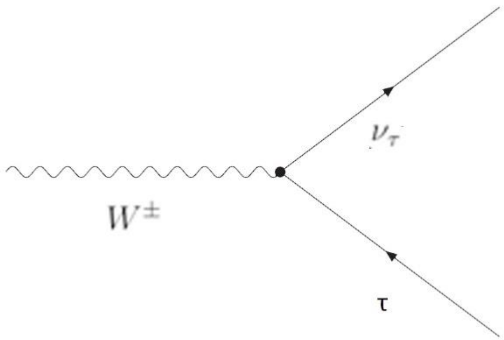
$\bar{t}t \longrightarrow \tau$ + jets
$40 < p_T < 100$ GeV



$Z \longrightarrow \tau\,\tau$
$20 < p_T < 50$ GeV



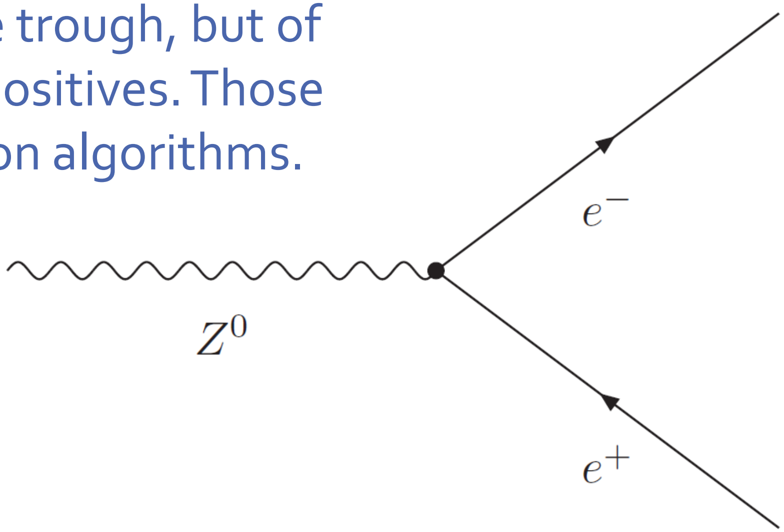
$W \longrightarrow \tau\nu$
$20 < p_T < 60$ GeV



Use of Z $\longrightarrow$ ee

- Background with respect to Z $\longrightarrow$ ττ, especially for 1-prong decays

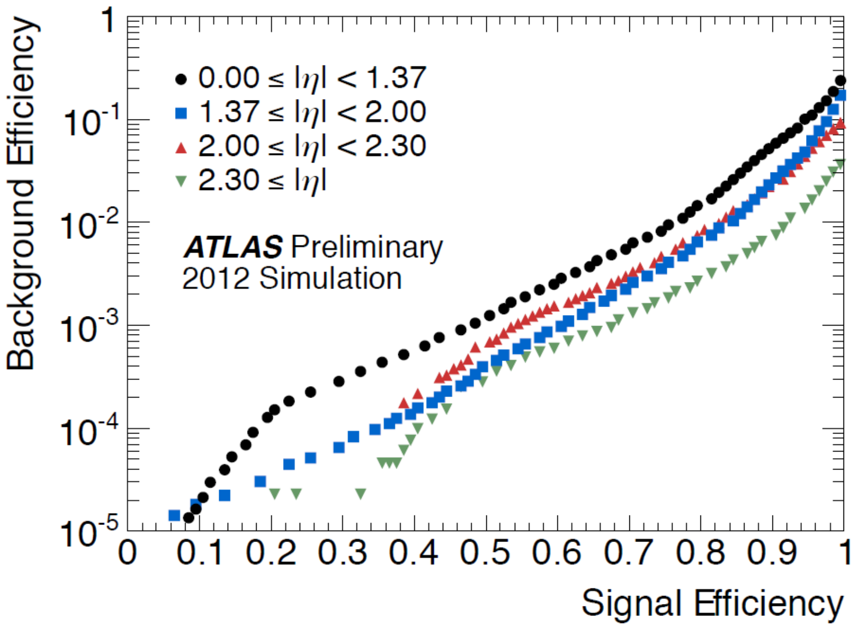
- We take some Z $\longrightarrow$ ee events and use the τ identification algorithms on them, in a perfect world none of them would come trough, but of course we will have some false positives. Those help to optimize the identification algorithms.



Discrimination against Electrons & Muons

The characteristic signature of 1-prong τ can be mimicked by electrons. This creates a significant background contribution after all the jet related backgrounds are suppressed.

The background discrimination from the real signal is based different variables which are related to electron and hadronic properties such as the shower shape or the emission of transition radiation.

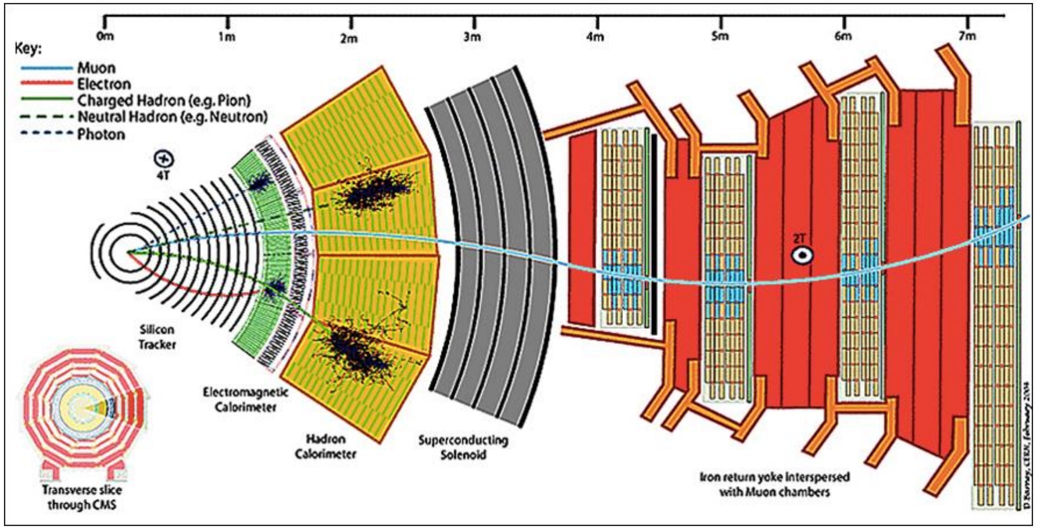


The electron veto BDT (e-veto BDT) is optimized using simulated $Z \rightarrow \tau\tau$ events for the signal and $Z \rightarrow ee$ events for the background.

Discrimination against Electrons & Muons

The signal candidates are required to be matched to true 1-prong hadronic τ decays, while background candidates are matched to true electrons.

As minimum ionizing particles, muons are unlikely to deposit enough energy in the calorimeters to be reconstructed as a τ candidate. However, when a sufficiently energetic cluster in the calorimeter is associated with a muon, the muon track and the calorimeter cluster together may be mis-identified as a τ.



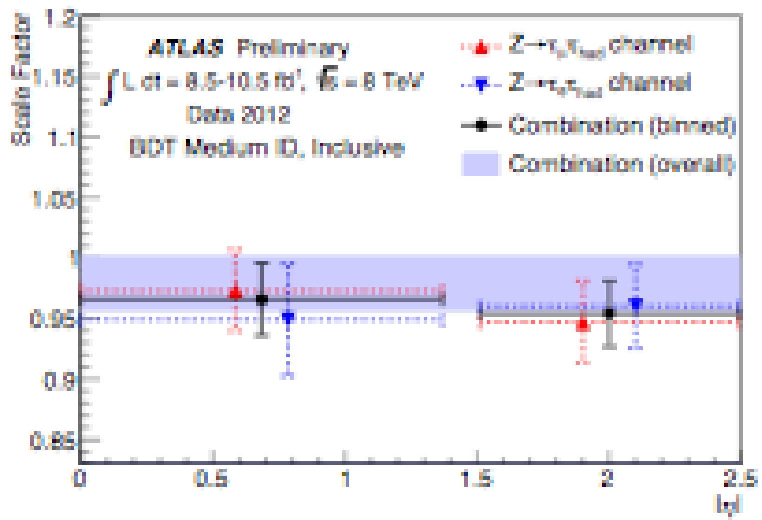
Efficiency Scale Factors: τ Hadronic Identification

It is important to verify that these identification algorithms perform comparably in both the predictions from simulated samples and in data.

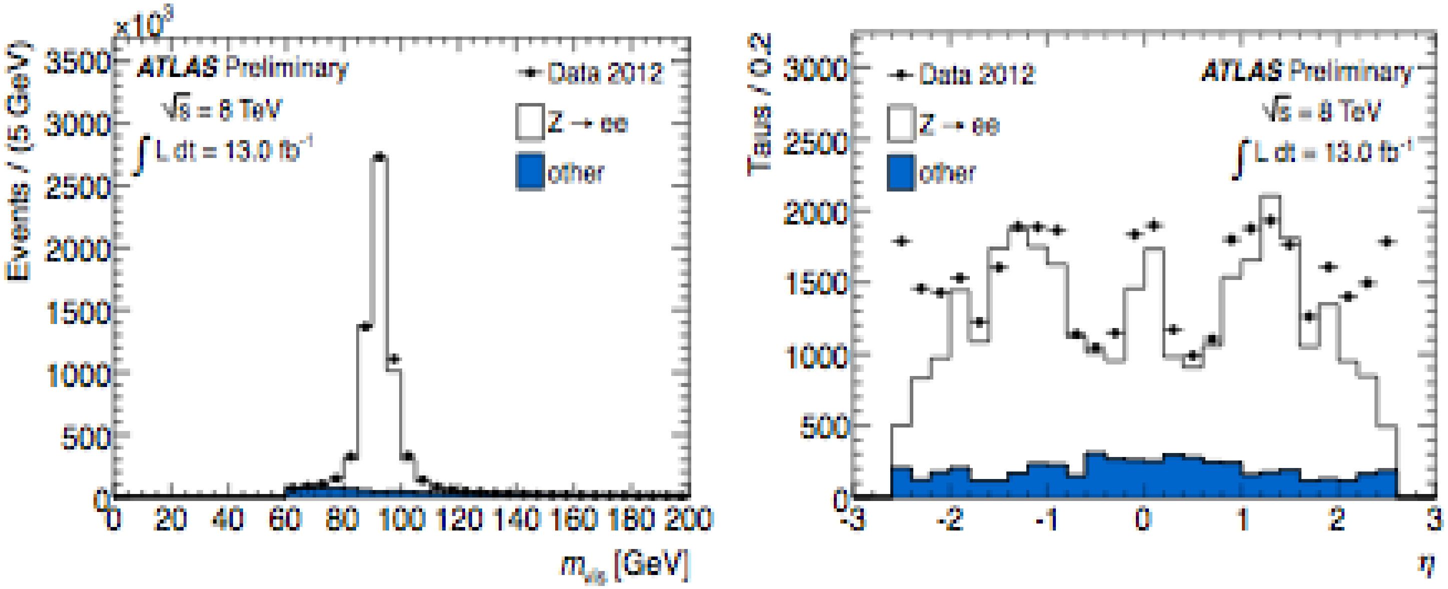
$$\mathrm{SF} = \frac{\varepsilon_{\mathrm{data}}}{\varepsilon_{\mathrm{simulation}}}$$

By comparing these results to the same performance figures in simulated event samples, scale factors are derived. These scale factors are then used in analyses to account for the differences between data and simulation due to modeling of the input variables.

These scale factors therefore account for the differences between data and simulation due to the modeling of the input variables for the identification algorithms.



Efficiency Scale Factors: Electron Veto



	Electron veto scale factors and uncertainties					
	medium BDT tau ID, medium e-veto, loose electron overlap removal					
$\|\eta(\tau_{\text{had-vis}})\|$	0.00 – 0.05	0.05 – 1.37	1.37 – 2.00	2.00 – 2.30	2.30 – 2.47	2.47+
SF	**0.86**	**1.12**	**1.40**	**1.58**	**2.70**	**21.69**

Conclusion

- The τ identification has been re-optimized to be robust with the increased amount of pile-up present in the 2012 data.

- Measurements of the $\boldsymbol{\tau}$ identification algorithm efficiencies were performed using the $Z \longrightarrow \tau\,\tau$, $W \longrightarrow \tau\nu$ and $\bar{t}t \longrightarrow \tau$ + jets channels with the goal of providing data to simulation scale factors.

- Good agreement was found between the performance of the identification algorithms in both simulated events and in data.

Outlook & Questions

Recent studies have shown that 3-prong τ can also be mimicked by electrons. This has not been considered in this note, but will be investigated in future electron veto studies.

Thank you for listening!