Tau Performance

ATLAS UK Physics Meeting
9th January 2008

Alan Phillips
University of Cambridge

(on behalf of the Tau Working Group)

Overview

I plan to cover the following topics as a sample of the work being carried out by the Tau Performance WG:

- Tau Reconstruction - Introduction
- Algorithms
  - Calorimeter-Seeded Algorithm (`tauRec`)
  - Track-Seeded Algorithm (`tau1p3p`)
- Plans for early data
  - Fake Rates
  - Tau Fakes from Electrons
  - $W \rightarrow \tau\nu$
  - $Z \rightarrow \tau\tau$
- Other news from the TauWG
- Summary

Tau Properties

The reconstruction of hadronic tau decays is a difficult task in a hadron collider environment – v.large background from QCD multijets

Decay modes	TAUOLA-CLEO
$\tau \rightarrow e\nu_e\,\nu_\tau$,	17.8 %
$\tau \rightarrow \mu\nu_\mu\,\nu_\tau$	17.4 %
$\tau \rightarrow h^{\pm}neutr.\nu_\tau$ (single-prong)	49.5 %
$\tau \rightarrow \pi^{\pm}\nu_\tau$	11.1 %
$\tau \rightarrow \pi^0\pi^{\pm}\nu_\tau$	25.4 %
$\tau \rightarrow \pi^0\pi^0\pi^{\pm}\nu_\tau$	9.2 %
$\tau \rightarrow \pi^0\pi^0\pi^0\pi^{\pm}\nu_\tau$	1.1 %
$\tau \rightarrow K^{\pm}neutr.\nu_\tau$	1.6 %
$\tau \rightarrow h^{\pm}h^{\pm}h^{\pm}neutr.\nu_\tau$ (three-prong)	14.6 %
$\tau \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$	9.0 %
$\tau \rightarrow \pi^0\pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$	4.3 %
$\tau \rightarrow \pi^0\pi^0\pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$	0.5 %
$\tau \rightarrow \pi^0\pi^0\pi^0\pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$	0.1 %
$\tau \rightarrow K^0_S X^{\pm}\nu_\tau$	0.9 %
$\tau \rightarrow (\pi^0)\pi^{\pm}\pi^{\pm}\pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$ (five-prong)	0.1 %
other modes with K	1.3 %
others	0.03 %

1-prong

3-prong

- Hadronic tau decays are typically well collimated, comprising of $\pi^{\pm}$ and π^0 components.
- Typically one or three charged decay products – characteristic “ntrack” spectrum
- The tau direction is well reproduced by charged component
- Taus are reconstructed by matching calorimetric clusters with inner detector tracks
- One can exploit the characteristic shape of hadronic tau decays to provide rejection power against QCD

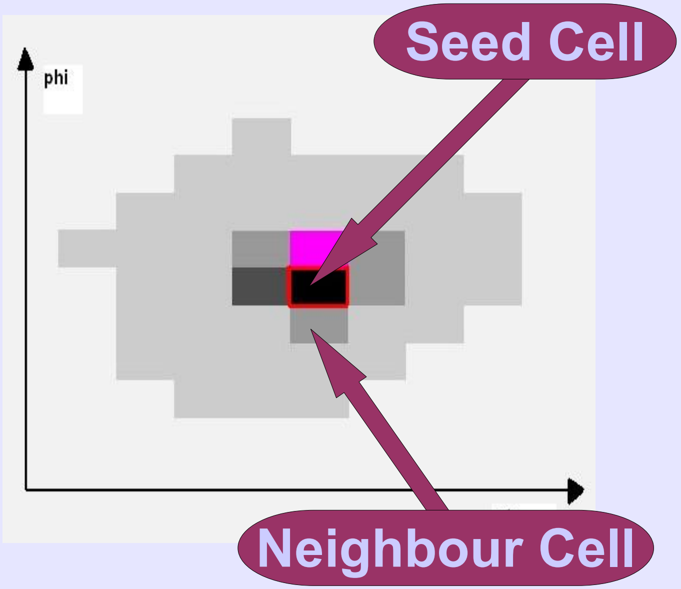
Along with measuring the properties of electroweak bosons and top quarks, one can expect tau leptons to be an important probe in searches for new physics phenomena (Low mass SM Higgs, MSSM Higgs, SUSY...)

Calo-Based Taus

➲The `tauRec` algorithm seeds tau reconstruction from calorimeter clusters. Recently `tauRec` has moved to using Cone 0.4 TopoJets formed from TopoClusters (previously seeded from CaloClusters from a sliding-window alg.).

➲These TopoJets are used as reconstructed tau candidates before identification.

➲TopoClusters have intrinsic noise suppression unlike the “sliding-window” approach – improves the tau reconstruction efficiency at lower p_T.

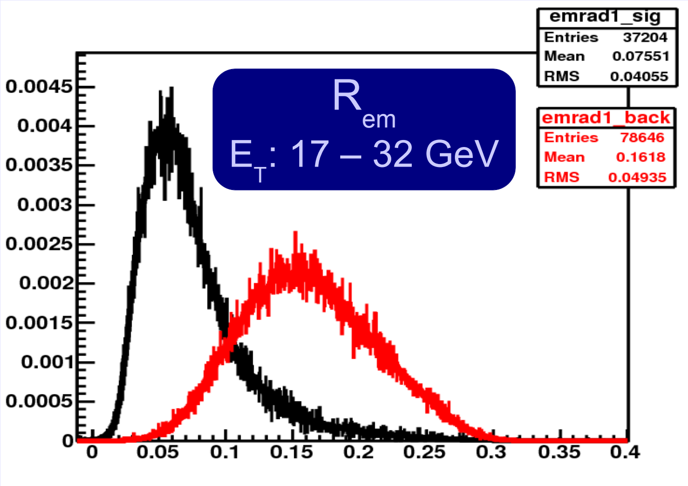


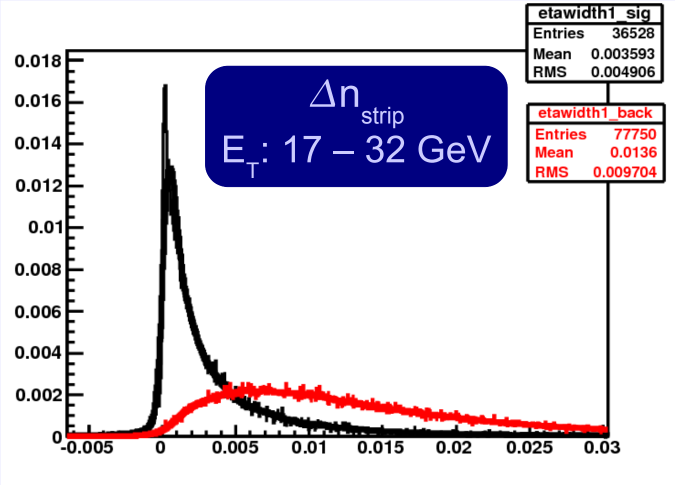
- TopoClusters are seeded from calo cells with energy above 4σ.
- Neighbouring cells with energy above 2σ are added iteratively. Finally all adjacent cells are added.
- A hadronic tau decay can form multiple clusters. A cone algorithm is used to group TopoClusters into TopoJets.

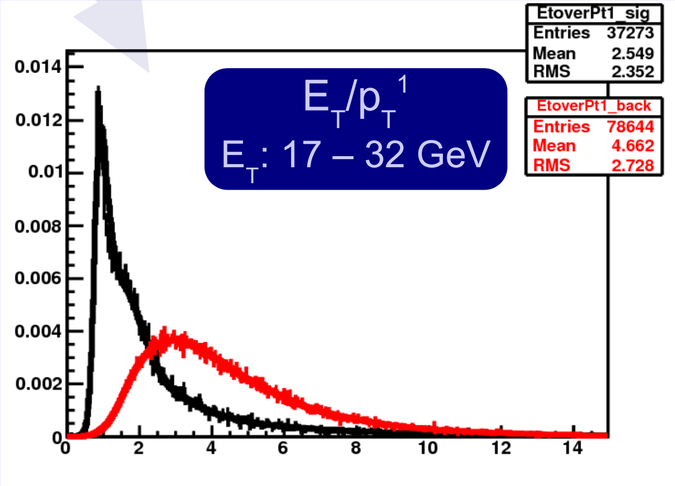
[*Adapted from talks by Stan Lai & Nico Meyer*]

Calo-Based Taus

- A likelihood function is used to identify taus from QCD fakes
- PDFs of discriminating variables are formed in 9 separate E_T bins
- Likelihood Variables:
  - R_{em} : Radius of cluster in EM Cal.
  - ΔE_T^{12} : Fraction of E_T in $0.1<\Delta R<0.2$ around cluster center
  - N_{strip} : Number of hits in η-strip with E_T>200 MeV
  - Δn_{strip} : Width in the η-strip
  - E_T/p_T^1 : Ratio of E_T over p_T of leading track
  - Q_τ :Tau Charge
  - Ntrks: Number of Tracks (within 1 to 3)
  - σ_{IP} : Lifetime signed impact parameter



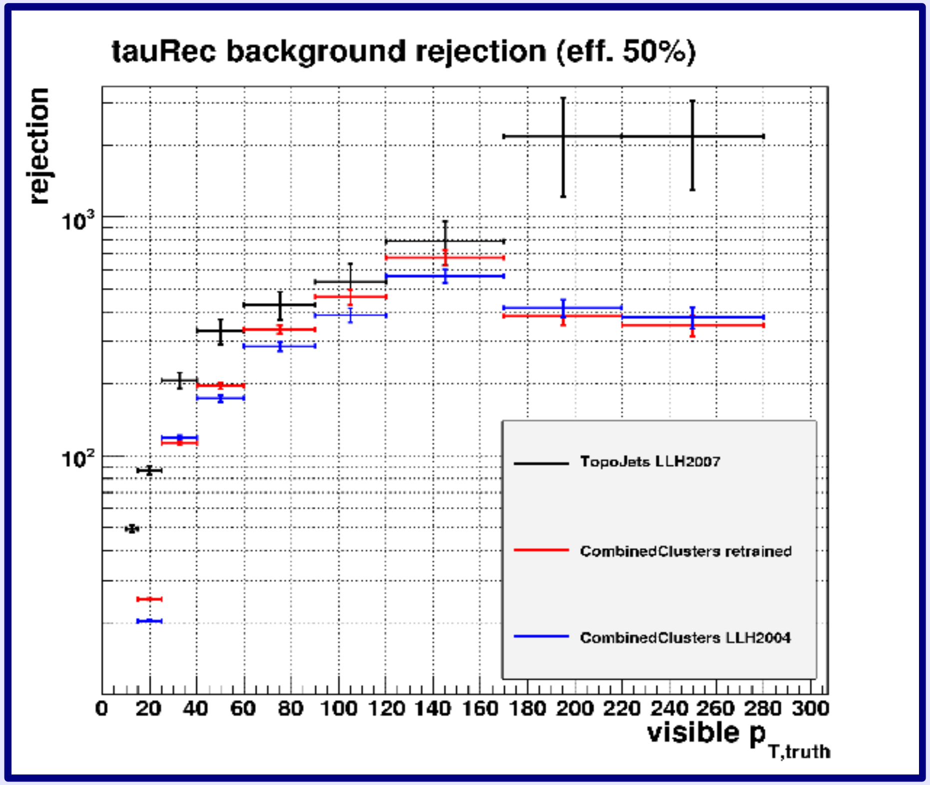




Examples of some of the variables used to construct the likelihood function. Signal taus are shown in black, with background in red

Calo-Based Taus

- A cut on this likelihood is used to make an identification decision.
- For a 50% signal efficiency, one can obtain rejections against QCD fakes of:
  - ~ 300–1000 for p_T>40 GeV
  - ~ 50–200 for p_T<40 GeV
- The move to TopoClusters and new retrained PDFs shows an improved performance in the background rejection (see figure)



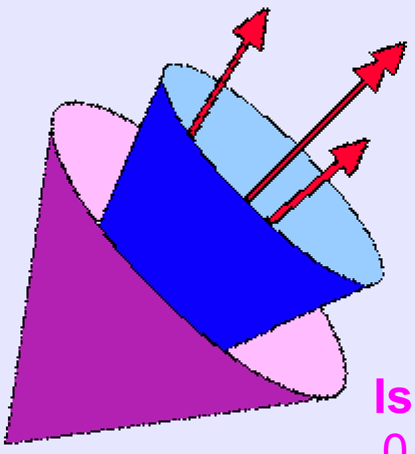
(*for more details see S.Lai's talk during the November 2007 T&P week*)

Track-Based Taus

The complementary `tau1p3p` algorithm seeds tau reconstruction from tracks, optimized for low p_T taus

Reconstruction

- ➲ Starts from tracks...
  - Identify a leading track: only “good quality” tracks with pT > 9 GeV
  - Define track spectrum in “core cone”
    - Ntrk=1 no nearby tracks with pT>1 GeV
    - Ntrk>1 (up to 6) more “good quality” tracks with pT > 1 GeV
    - For Ntrk=3, sum(charge) = +-1 checked
- ➲ (η, φ) of candidate defined by the track at vertex (1P candidates) or weighted barycentre (multi-track candidates)
- ➲ associate calorimeter cells to candidates
- ➲ Energy scale defined using an energy-flow approach (using “core cone”)

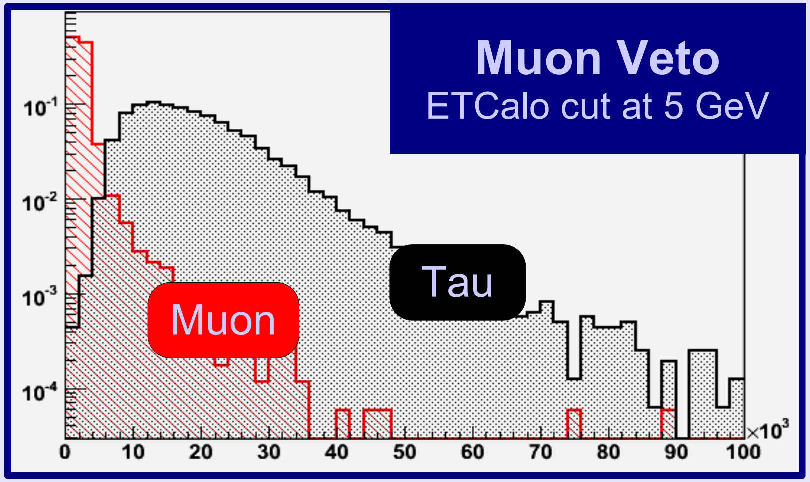


Identification

- ➲ Using tracking and calorimetric discriminanting variables
  - Cut-based discriminant
  - Multivariate Methods:
    - Neural Network (NN)
    - PDRS
- ➲ Also apply lepton vetoes

[*Adapted from talk by Anna Kaczmarska*]

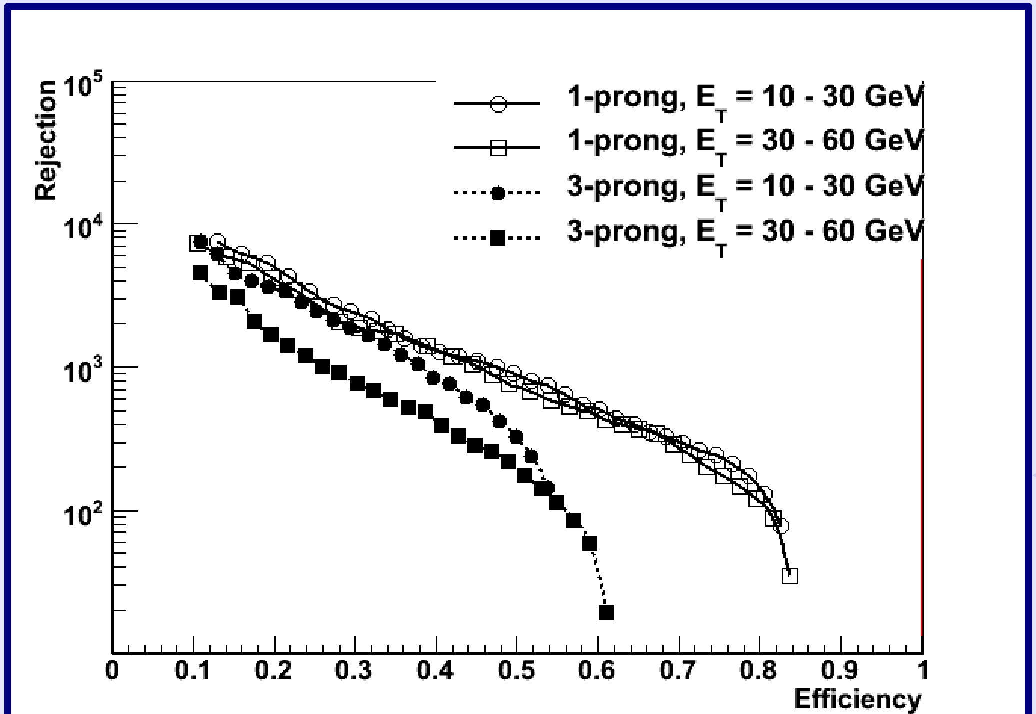
Track-Based Taus



Electron Veto Variables:

- E^{HCAL} – Energy in 1st layer of HCAL around an extrapolated track
- $E^{strip\text{-}max}$ - 1st local maximum non-associated with track
- E_T/P_T – ratio between ECAL energy and track momentum
- HT/LT – ratio High to Low Threshold hits

Electron Veto
using 4 calo / inner detector variables

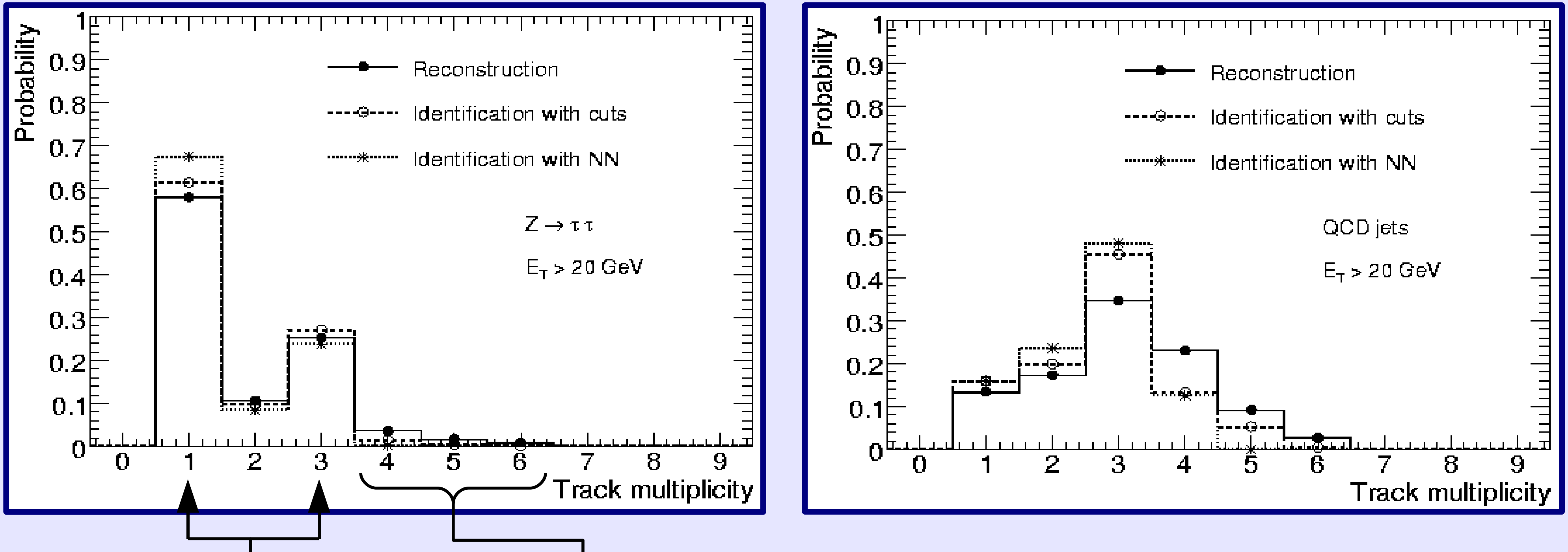


- Discriminants using simple cut-based and multivariate methods (NN, PDRS) are used to reject QCD background.
- Figure left demonstrates the efficiency v rejection curves for a NN discriminant for 13.0.30 (*1-prong and 3-prong taus are shown separately*)

Normalization used: **signal:** true taus with ETvis > 10 GeV, |eta|<2.5, at least one pion with pT>9 GeV **dijets:** MC truth jets

Track-Based Taus

- Normalized track multiplicity spectra for hadronic tau candidates (E_T > 20 GeV). $Z \rightarrow \tau\tau$ (*left*) and QCD (*right*)
- (i) Reco. (ii) Reco+id [cuts] and (iii) Reco+id [NN] are shown separately



Only use 1+3 track candidates as signal

Higher track multiplicities important for bgnd normalization

Aim to not bias the ntrack spectrum through the reconstruction and identification process

Physics With First Data

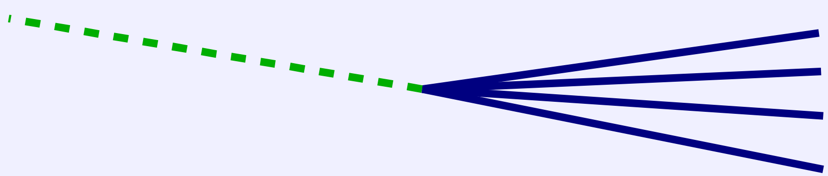
- One of the first tau physics goals in ATLAS will be to obtain a high purity sample of tau-leptons from the data and to measure the tau identification efficiency.
- $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$ and ttbar events will provide sources of taus for study in early data (first ~100pb^{-1}).
- Some recent studies in preparation for early physics include:
  - Calculating Jet Fake rates from data
  - Use of data in electron tau separation studies
  - $W \rightarrow \tau\nu$ analysis
  - $Z \rightarrow \tau\tau$ analysis
- For more information on this work, and for further studies, I encourage people to consult the Tau CSC Note (*ATL-COM-PHYS-2007-066*) and recent TauWG agendas

Tau Fake Rates From Data

- Early data will be used to understand tau fake rates from QCD jets.
- Rather than relying on MC, data driven methods could be used to correct/tune tau reconstruction in MC.
- The abundance of jets in LHC data mean that this can be studied with very early data.

Back-to-back: $|\Delta\phi| = \pi \pm 0.3$

p_T balance: $\Delta p_T < p_T^{max}/2$



Probe Jet

No further selection (to keep whole spectrum)

Check if identified as a tau

Tag Jet

ntrack ≥ 4 (with p_T > 1 GeV)

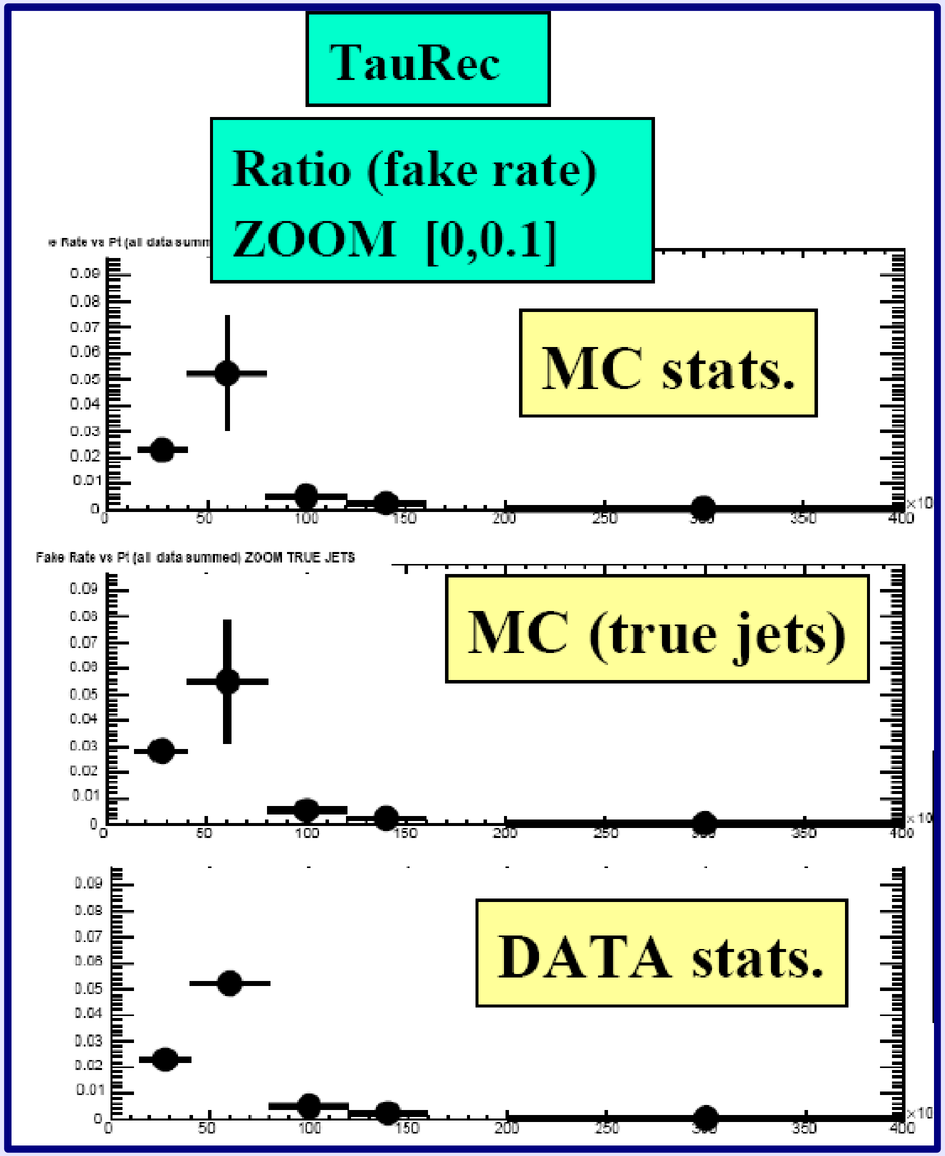
Removes most real taus

- The idea is to select a sample of "very likely" QCD jets and see how many are wrongly identified as taus
- Look for 2 back-to-back objects
- Requiring one "nice" tag jet to remove most real taus
- Then we are confident that we also have a jet on the other side
- Use the other side as a probe to calculate the fake rate.

$$\text{Fake Rate} = \frac{\#\ \text{Probe Jets identified as taus}}{\#\ \text{Probe Jets}}$$

[*S.Brunet, P.Bechtle, D.Cote, S.Johnert*]

Tau Fake Rates From Data



Studies with MC samples have been performed as a proof of principle

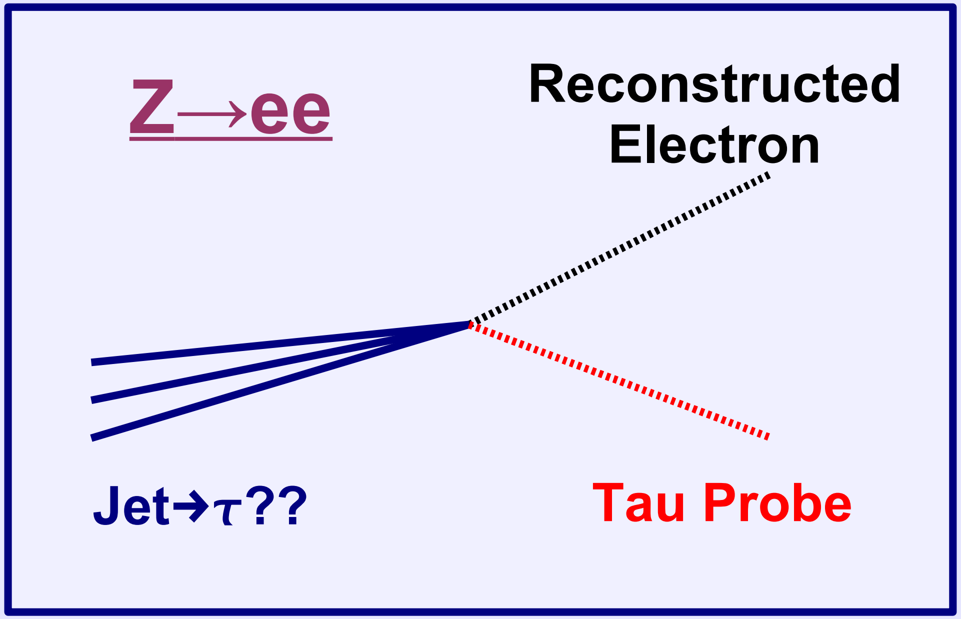
The available MC statistics are too small so the numbers and error bars aren't too meaningful at the moment

However, statistics shouldn't be a problem in data. More important at the moment is the development of the method

[*S.Brunet, P.Bechtle, D.Cote, S.Johnert*]

e/τ Separation From Data

- Early data will also provide the opportunity to study electron tau separation. Z→ee events where one electron fakes a tau are studied.



- One electron is well defined and is used as a reference tag
- The other leg is used to probe electrons being mis-tagged as taus
- The only contamination is from Jet→τ fakes
- Reconstructing the invariant mass of the e-"tau" system allows one to extract the fraction of misidentified taus from data

Tau1P3P	Identified Taus	Electron Veto	Number of Zs	Number of Zs with electron veto
Tight Electron	169053	4583 (2.71%)	133606	3881 (2.90%)

Misidentified electrons in Z→ee

MC studies using a sample of Z→ee.

Numbers here are for tau1p3p (before and after applying the electron veto).

Of electrons faking taus and passing the tau-id, < 3% pass the electron veto

[*S.Dhaliwal, R.Mazini, R.S.Orr*]

W→$\tau\nu$ Analysis

- W→$\tau\nu$ events will be produced abundantly σ×BR = 1.7×10^4 pb
- Very large QCD background with σ ~10^9 pb
- Also expect a significant background contribution from W→eν
  - electron passes selection of the hadronic tau trigger.
- Events will be triggered with a combined *tau+MET* trigger
- It is important to have a carefully tuned procedure to correct for inevitable trigger bias
  - need to understand tau trigger efficiencies
- Expect a S/B ~ 1. Observability will be checked by looking at the track spectrum
  - (Characteristic 3:1 ratio of 1-prong to 3-prong taus, with a suppressed contribution from 2-prong candidates)
- The acceptance of the tau trigger may strongly bias the track multiplicity e.g. Could be more efficient for 1-prong than 3-prong. Need to correct for this.
  - Z→$\tau\tau$ as a control sample... unfortunately for the first 100pb^{-1}, this has a lower rate

Track multiplicity for W→$\tau\nu$ signal and QCD background. (A Preliminary Fast Simulation study. See Tau CSC note for full details of analysis)

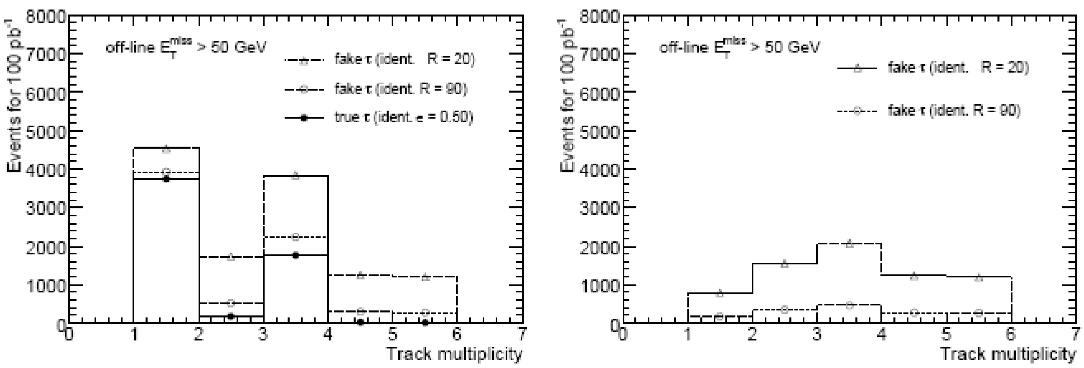


Figure 76: Track multiplicity distribution for signal candidates from $W \to \tau\nu$ and background events from the QCD processes. Residual background from $W \to e\nu$ is not included. Results from early version of the track-based fast-simulation.

Z→ττ Analysis

- The Z→ττ Signature will occur at a rate 10 times lower than that of W→τν.
- Trigger on Z→$\tau_{lep}\tau_{had}$ with a lepton trigger (unbiased sample of hadronic tau decays – Useful to help understand tau-trigger efficiencies).
- The hadronic-tau energy scale can be determined by using the visible invariant mass of tau pairs
- The following results are for 100pb^{-1}, using fully simulated samples from release 12 CSC production. Signal dataset used DS 5189. [*D. Cavalli, C. Pizio, CSC Note*]
- The analysis procedure starts with event selection:
  - Basic set of cuts on lepton and global event quantities (E_TMiss, m_T (lep-E_TMiss), SumE_T, b-tagging)
  - Leptons are required to have p_T>15 GeV, |η|<2.5, and pass isolation and i.d.
  - At this stage, all the cuts are *independent of the tau algorithm used*

Cuts	$Z\to\tau\tau$	$W\to e\nu$	$W\to\mu\nu$	Jall	tt	Zee	$Z\mu\mu$
Isol lepton (e,μ)	$1.5\cdot10^4$	$7.4\cdot10^5$	$9.3\cdot10^5$	$8.1\cdot10^7$	$2.6\cdot10^4$	$1.1\cdot10^5$	$1.3\cdot10^5$
$E_T^{miss} > 20$	4744	$6.3\cdot10^5$	$8.0\cdot10^5$	$4.5\cdot10^6$	$2.4\cdot10^4$	5137	$4.7\cdot10^4$
$m_T^{lep,E_T^{miss}} < 30$	3215	$1.2\cdot10^4$	$1.4\cdot10^4$	$3.5\cdot10^6$	3646	1525	1646
$\Sigma E_T < 400$	2997	$1.1\cdot10^4$	$1.3\cdot10^4$	$3.4\cdot10^6$	1281	1251	1540
b-jet veto	2765	$1.0\cdot10^4$	$1.3\cdot10^4$	$3.1\cdot10^6$	133	1161	1427

After the event selection, most significant bg comes from QCD, also W→eν and W→μν

Z→ττ Analysis

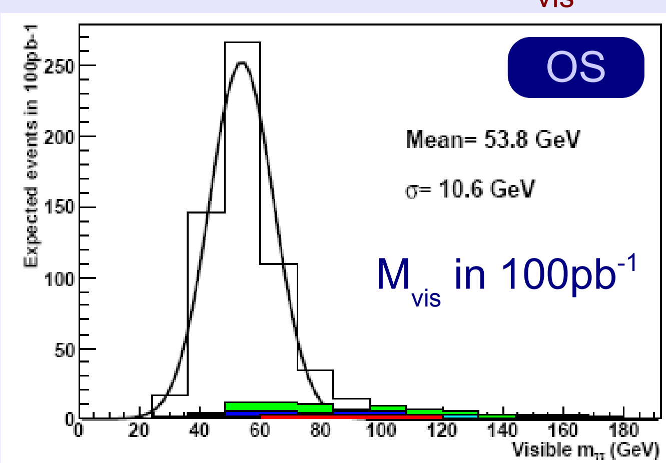
- The next step is to form mass combinations:
  - Each e/μ candidate is combined with a had. τ candidate (either `tauRec` / `tau1p3p`)
  - Apply a 2nd set of cuts to the mass combinations ($\Delta\phi$, τ-Id (factorized for backgds))
  - Reconstruct the visible $\tau\tau$ mass
  - Select Opposite Sign events, subtracting backgrounds using Same Sign events

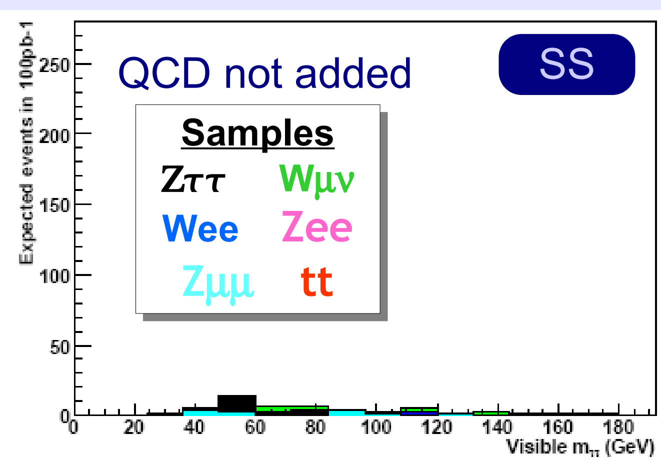
- Numbers here are for `taurec`, (similar analysis for `tau1p3p`)

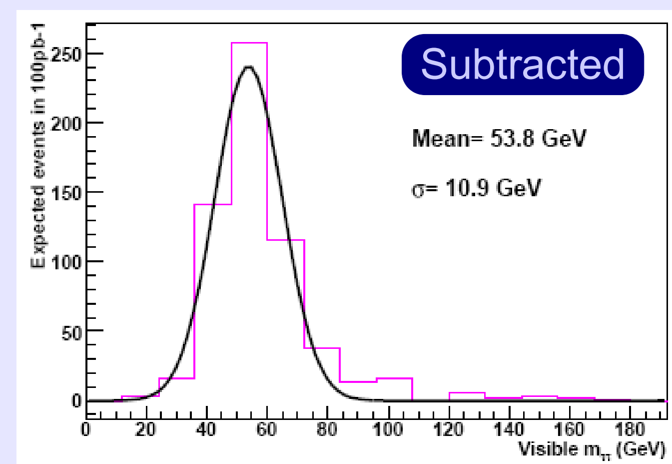
Very high QCD dijet bg with large errors. Low statistics – need to use Fast Simulation

Cuts	$Z \to \tau\tau$	$W \to e\nu$	$W \to \mu\nu$	Jall	tt	Zee	$Z\mu\mu$
Mass comb. with τ-cand	3754±74	10710±315	12818±413	$3.3\cdot10^6$	304±18	1141±19	1157±40
$\Delta\phi < 3.1 \& > 3.2$	3449±71	9970±304	11993±400	$3.2\cdot10^6$	299±18	1050±18	1062±39
$1 < \Delta\phi < 5.3$	3248±69	9636±299	11553±392	$2.0\cdot10^6$	265±17	929±17	998±37
Identified $\tau - jet$	628±30	75±4	139±8	3050 ±870	4±2	6±1	27±6
OS events	596±30	47±3	88±6	1060±502	4±2	4±1	13±4
$m^{lep,\tau-jet}$ 37-75	519±28	15±2	29±3	N.E.	0±0	2±1	6±3
SS events	32±7	28±2	51±5	1990±710	0±0	2±1	14±4
$m^{lep,\tau-jet}$ 37-75	27±6	10±1	14±2	N.E.	0±0	1±1	6±3

- The resulting M_{vis} distribution can be used in defining the tau energy scale.





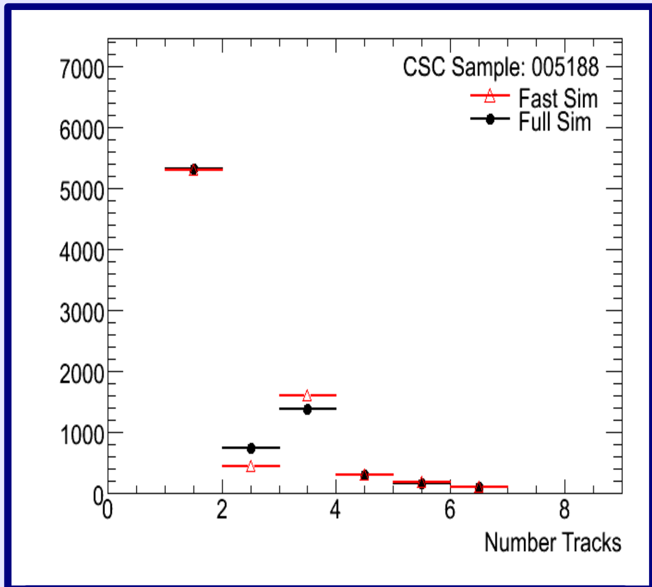


Other Developments In The TauWG...

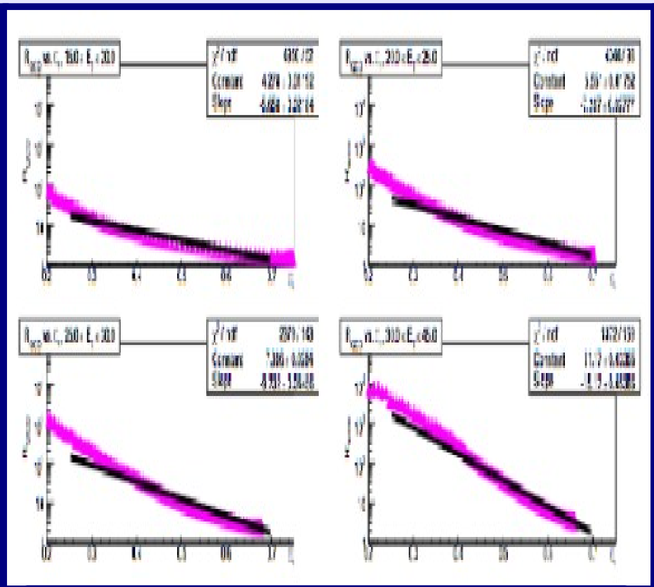
- A Merged Algorithm
  - Combine the advantages of two complementary algorithms into a single merged algorithm,
  - Aims to combine the low-p_T performance of `tau1p3p` with the high-p_T performance of `tauRec`.
  - Benefits to the user (*only one baseline tau object*) and from the trigger point of view (*only need to optimize one alg.*)
- TauDPDMaker [*David Cote*]
  - A prototype for making the primary DPDs of the Tau WG
  - DPD = subset of ESD/AOD in pool.root format
  - Skimming, slimming and thinning technicalities now work
  - There is ongoing effort in exploring the physics content
- TauTools library [*Jyothsna Rani Komaragiri*]
  - Provides a library of useful code, avoiding duplication of effort – simple tools that can be accessed in any analysis environment
  - In CVS... `PhysicsAnalysis/TauID/TauTools`
- Fast Simulation *(See next slide)*

Fast Simulation

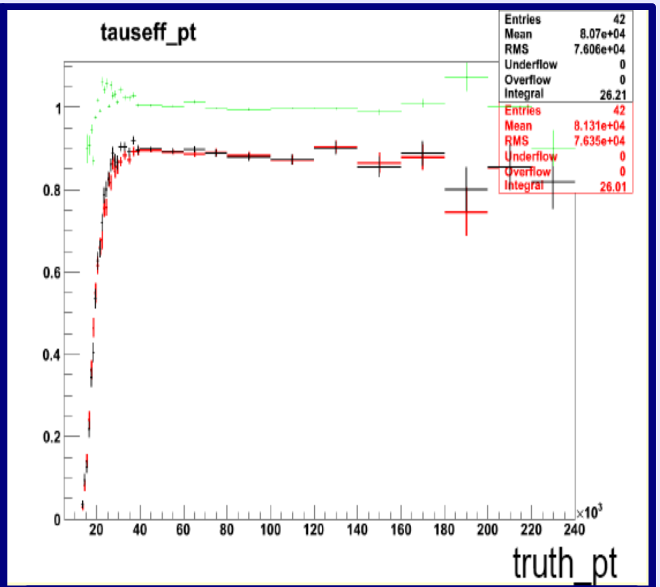
- Many studies require very large statistics due to the huge QCD cross-section. (Full sim samples available for CSC analyses were ~100-1000 times to small)
- These statistics aren't possible with full simulation alone (O(10mins) / event). Many studies need to also make use of fast simulation tools e.g. Atlfast-I and Atlfast-II
- There is ongoing work on updating parametrizations for Atlfast-I and developments with Atlfast-II



Track-based taus in ***Atlfast-I*** – A track seeded approach to fast sim. Planned updates for r14 [*A.Phillips*]



Calo-based taus in ***Atlfast-I*** – Planned updates to tau-tagging parametrization for r14 [*W.Davey*]



Atlfast-II Development. Ongoing validation and prep. of tau efficiency corrections [*E.Schmidt, W.Davey*]

Summary

- Hadronic tau-decays will play an important role in physics at the LHC
- Identifying taus is a difficult task
  - Requires rejecting large background from QCD jets
- Tau reconstruction in ATLAS is served by two complementary tau reconstruction algorithms, one calorimeter seeded, and one track-seeded.
- One of the first tau physics goals in ATLAS will be to obtain a high purity sample of tau-leptons from the data. $W \rightarrow \tau\nu$, $Z \rightarrow \tau\tau$, ttbar will be early sources of taus in data
- Need to measure the efficiency and rejection performance of the tau identification using real data

Extra Slides

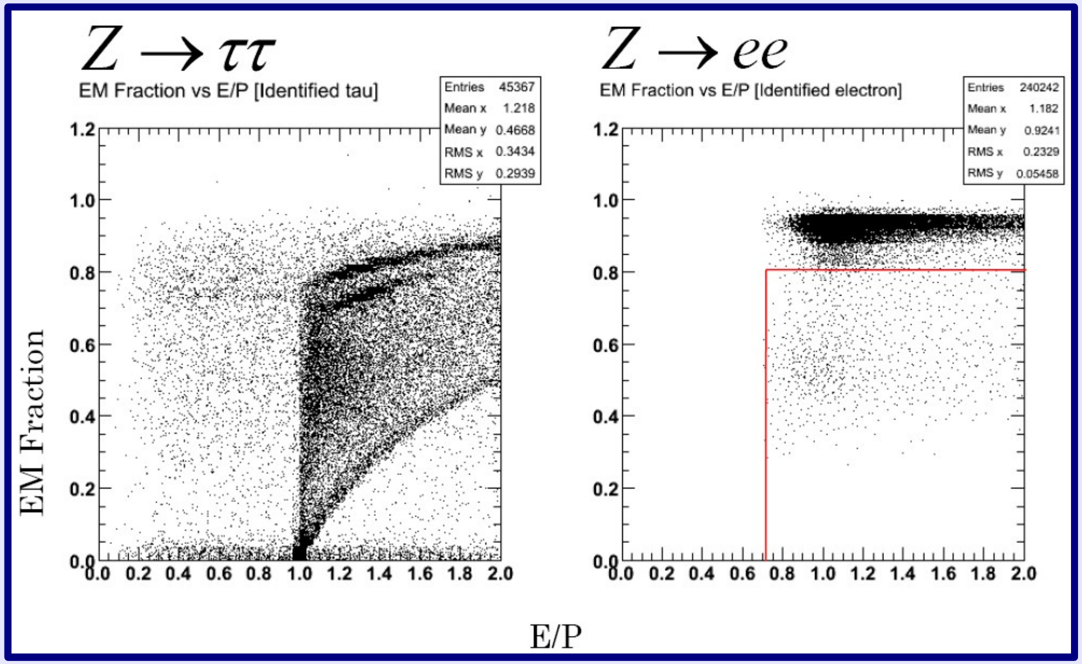
e/τ Separation From Data

TauRec	Identified Taus	Electron Veto	Number of Zs	Number of Zs with electron veto
Tight Electron	189148	32794 (17.3%)	149439	22513 (15.1%)
Loose Electron	172424	27332 (15.9%)	204028	30807 (15.1%)

Tau1P3P	Identified Taus	Electron Veto	Number of Zs	Number of Zs with electron veto
Tight Electron	169053	4583 (2.71%)	133606	3881 (2.90%)
Loose Electron	158204	3888 (2.46%)	182536	5327 (2.92%)

MC studies using a sample of Z→ee events demonstrate the e/τ misidentification rate with existing algorithms.
Results are shown for both tau algorithms, before and after applying their electron veto.

Studies include investigations into variables to discriminate between electrons and taus.
e.g EM Fraction vs E/P shown right



[*S.Dhaliwal, R.Mazini, R.S.Orr*]