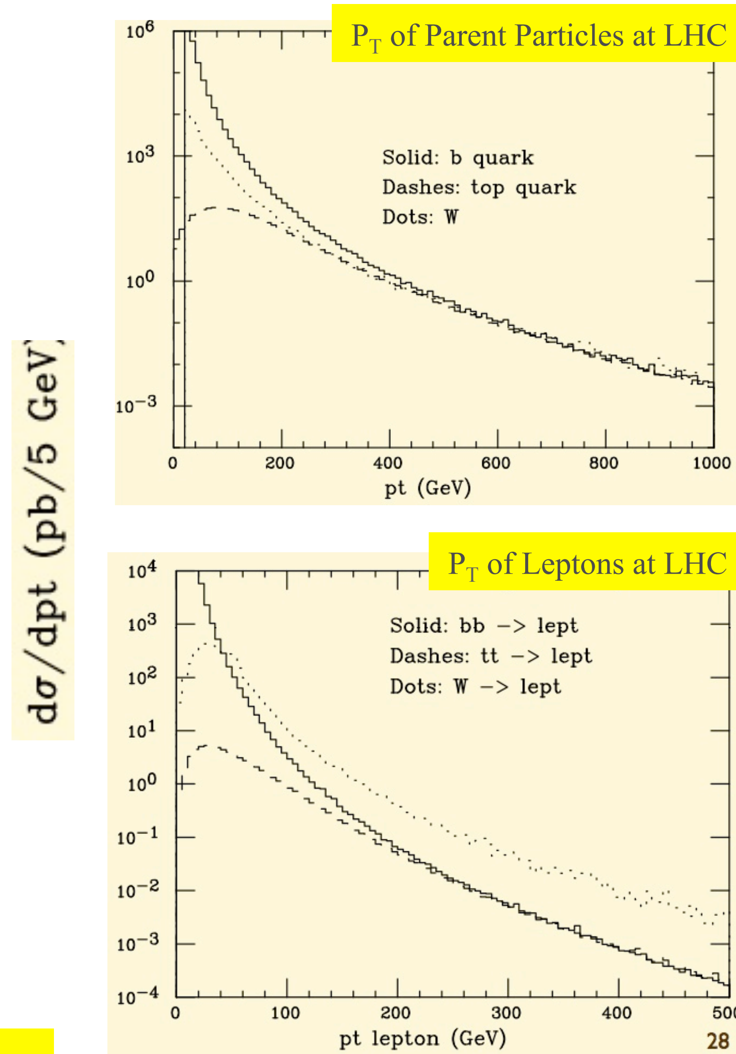
Hard Scattering in Hadron-Hadron Collisions: Physics and Anatomy

Section 5: Identification of Charged Leptons

1. Sources of leptons
2. ID techniques for electrons, muons and taus
3. Identification efficiencies
4. Background considerations
5. Example: Top quark decays to τ leptons

Sources of Leptons

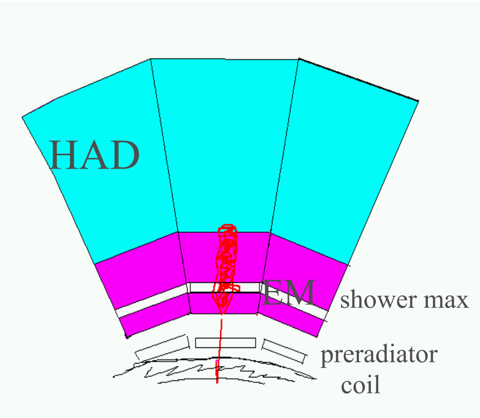
- **In hadron-hadron collisions, leptons arise from**
  - **Semileptonic decay of heavy quarks (t/b/c)**
  - **W and Z boson decay**
  - **Drell-Yan production**
  - **“Onia” production/decay**
- **Various sources of backgrounds**
  - **Electrons**
    - Photon conversions
    - Misidentified jets
  - **Muons**
    - Cosmic rays
    - Decays-in-flight of hadrons
  - **Taus**
    - Misidentified jets

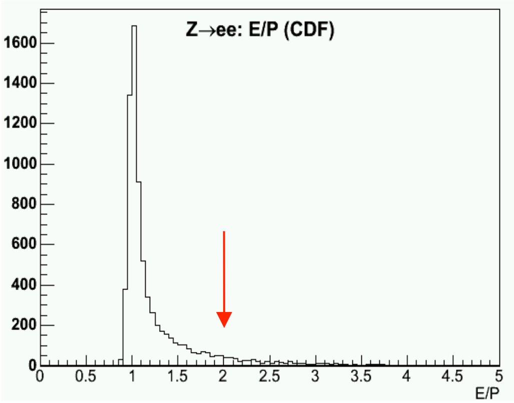


Courtesy of M. Mangano

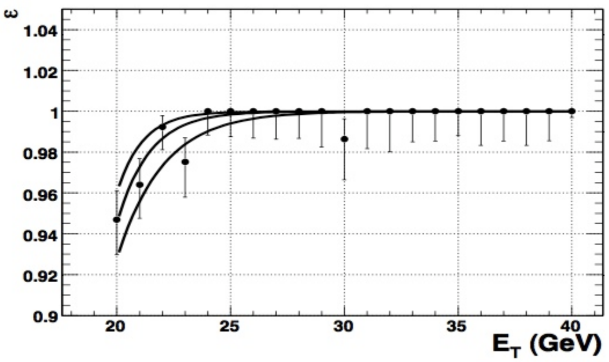
Electron Identification Strategies

- **Identification makes use of**
  - **Calorimeter**
    - Shower shape and location
  - **Charged particle reconstruction**
    - Position matching
    - Energy vs Momentum
  - **TRD and/or dEdX**
    - TRD perhaps has the highest rejection power
- **Strategy works well, but**
  - **Depends on a large number of ID variables**
    - Have to be well-modelled
  - **Requires "isolated" lepton candidates**
    - Electrons from b/c decays difficult to reconstruct
  - **Have correlation between tracks & calorimetry in trigger**

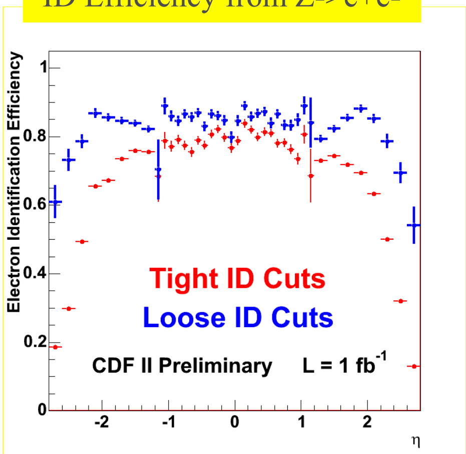




Trigger Efficiency from Z->e+e-

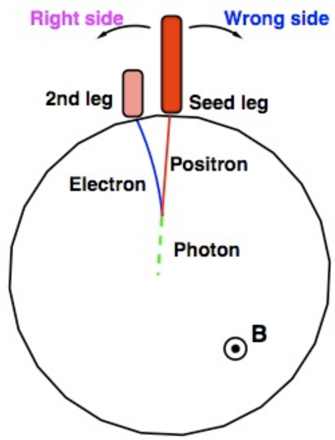


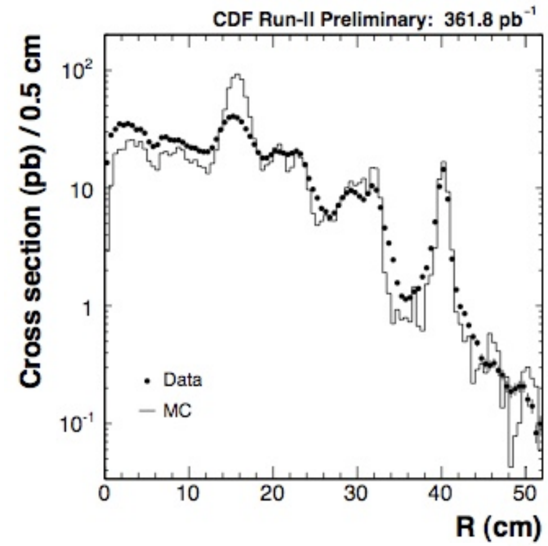
ID Efficiency from Z->e+e-

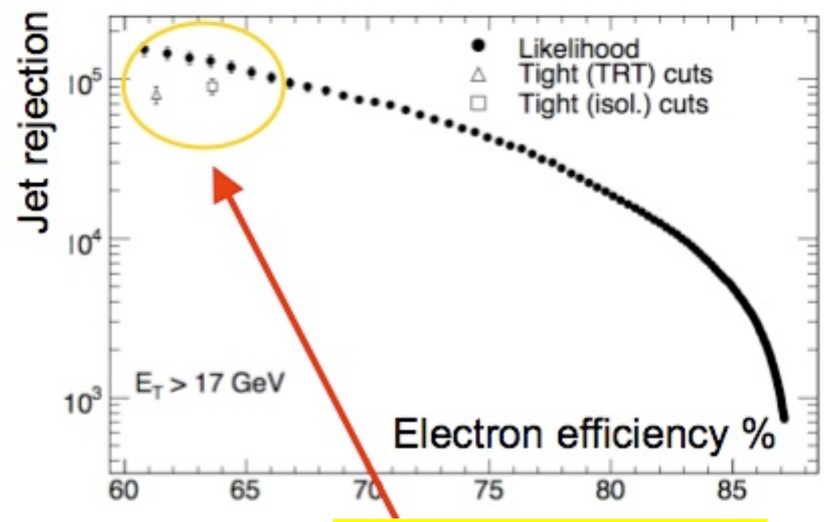


Backgrounds to Electrons

- **Backgrounds have to be measured**
  - **Conversions ($\gamma \to e^+e^-$)**
    - Source of real electrons (about 30-40% of electrons above P_T>12 GeV/c)
      - **Search for partner leg, or**
      - **No charged track**
    - Large background, but also a good control sample
  - **Jets**
    - π^+/π^0 overlap
      - **Two pions overlap & mimic electron signature**
    - Charge-exchange
      - **$\pi^+ p \to \pi^0 n$ early in calorimeter**
- **Can get to relatively pure samples**
  - **S/N > 10-100, depending on process**
  - **ATLAS/CMS expect to be able to do very well**
    - Z->e$^+$e$^-$ provides excellent "standard candle"







Cut-based results vs likelihood

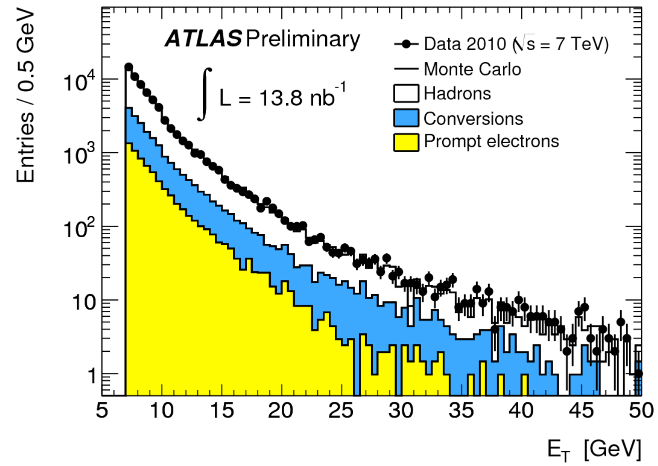
Performance of ATLAS

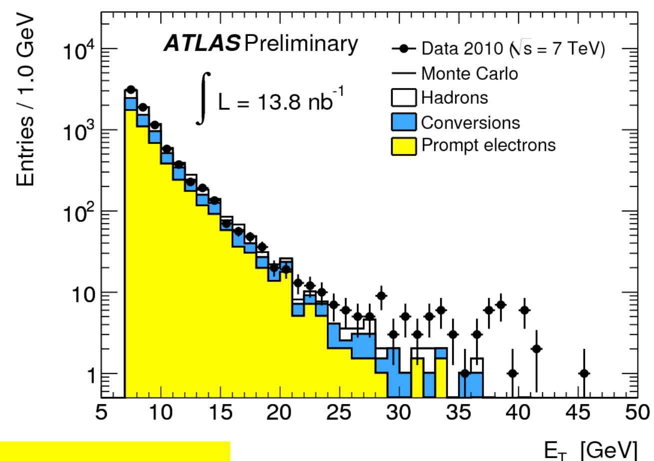
- Developed sets of ID cuts that efficiently select electrons
  - **Key is to separate "prompt" electrons from**
    - Conversions ($\gamma \to e^+e^-$)
    - Candidate from hadrons/jets
  - **Use two variables**
    - Number of hits in pixel detector
      - **Conversions typically have fewer**
    - Fraction of large pulse-height "hits" in TRT
- **Can then solve for the three components**

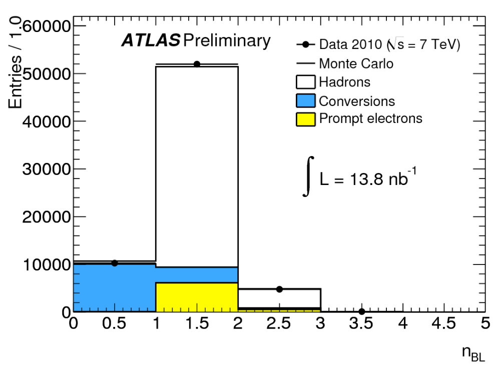
$$N = N^h + N^\gamma + N^Q$$

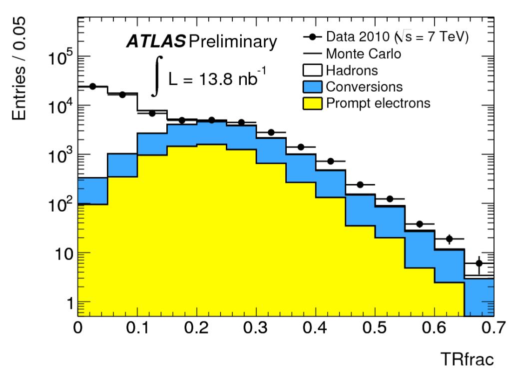
$$N_{TR} = N^h \varepsilon^h_{TR} + N^\gamma \varepsilon^\gamma_{TR} + N^Q \varepsilon^Q_{TR}$$

$$N_{BL,TR} = N^h \varepsilon^h_{BL} \varepsilon^h_{TR} + N^\gamma \varepsilon^\gamma_{BL} \varepsilon^\gamma_{TR} + N^Q \varepsilon^Q_{BL} \varepsilon^Q_{TR},$$







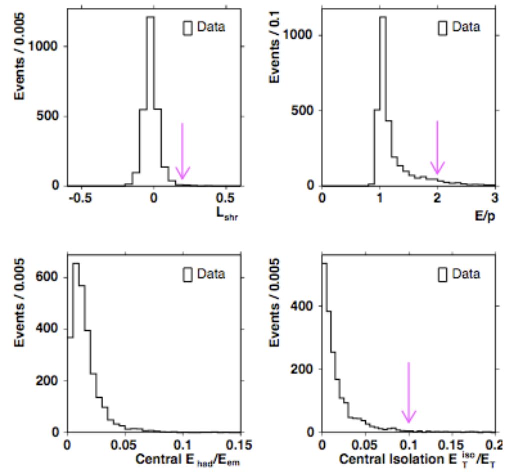


ATLAS-CONF-2010-073

Component	$h \to e$		$\gamma \to e$		$Q \to e$	
Method	Matrix	Likelihood	Matrix	Likelihood	Matrix	Likelihood
Fraction of electron candidates	65.2 ± 0.4	65.4 ± 0.3	19.8 ± 0.2	19.4 ± 0.2	15.0 ± 0.2	15.2 ± 0.2

Cut-Based Lepton ID Selections

- **Electron ID uses a large number of variables**
  - **Traditionally, define selection criteria**
  - **Measure efficiencies of each cut in well-understood control sample**
    - Relax the cut and see how control sample responds
- **Challenge is how to measure efficiencies & backgrounds**
  - **Need to understand correlations between cuts**
  - **Multiple control samples are very helpful**
  - **Remember trigger also performs selection!**



- **Likelihood techniques in principle more efficient**
  - **But more difficult to understand in detail**
  - **Correlations are important to get right**

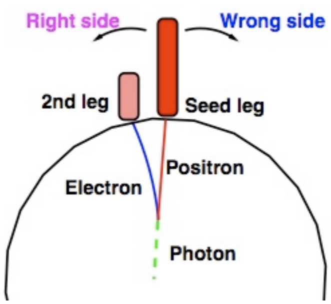
Z Selection and Efficiency

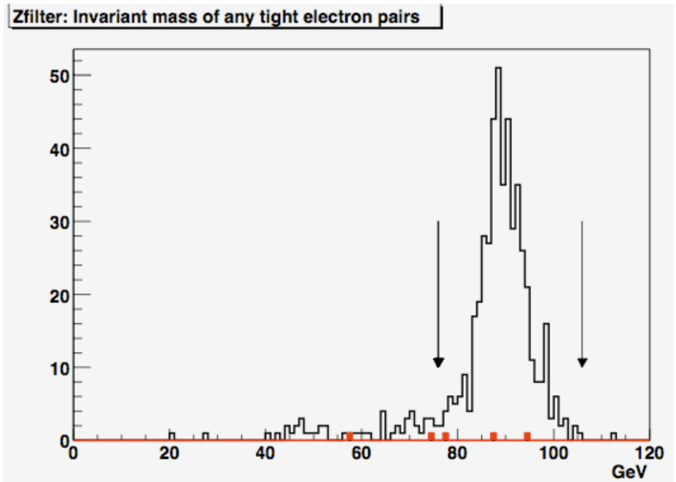
Selection Criteria	Number of Events	Net Acceptance
Total Events	507500	-
$\lvert z_{vtx}\rvert < 60$ cm	490756	-
$66\ \mathrm{GeV}/c^2 < M_{ee}(\mathrm{Gen}) < 116\ \mathrm{GeV}/c^2$	376523	-
Central EM Cluster	363994	0.9667 ± 0.0003
Calorimeter Fiducial Cuts	299530	0.7955 ± 0.0007
Electron Track $p_T > 10$ GeV/c	252881	0.6716 ± 0.0008
EM Cluster $E_T > 25$ GeV	186318	0.4948 ± 0.0008
Second EM cluster (Central or Plug)	176417	0.4685 ± 0.0008
Second Cluster Calorimeter Fiducial Cuts	146150	0.3882 ± 0.0008
Second Electron Track $p_T > 10$ GeV/c (Central)	138830	0.3687 ± 0.0008
Second EM Cluster $E_T > 25$ GeV (Central), 20 GeV (Plug)	125074	0.3322 ± 0.0008
Second EM Cluster $E_{had}/E_{em} < 0.125$ (Plug)	124881	0.3317 ± 0.0008
$66\ \mathrm{GeV}/c^2 < M_{ee}(\mathrm{Rec}) < 116\ \mathrm{GeV}/c^2$	120575	0.3202 ± 0.0008
Opposite Charge (Central-Central)	119925	0.3185 ± 0.0008

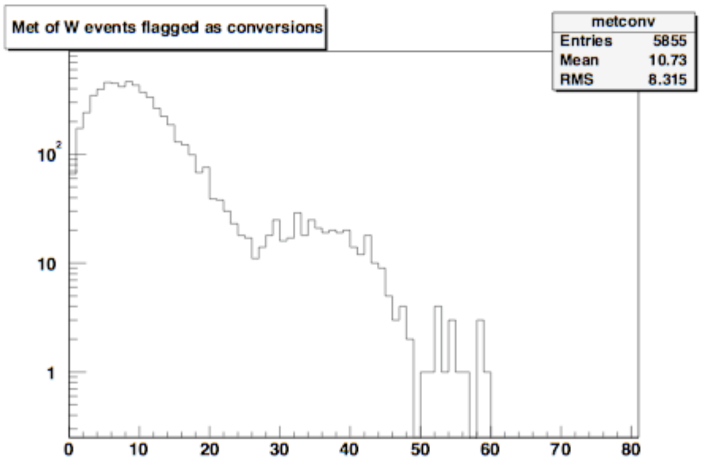
Abulencia et al. (CDF), J. Phys G 34, 2457 (2007)

Example: Conversion Removal

- **To identify conversions, CDF uses presence of 2nd track & SVX hits**
  - **~85% efficiency in identifying conversions in electron sample with P_T>9 GeV/c**
    - Residual background depends on other cuts
  - **Significant "over-efficiency"**
    - Probability of misidentifying a prompt electron as γ
    - Measure this using Z->e^+e^- decays
    - Get between 5-10%, depending on details of algorithm
      - **Measured to be 4.5±0.6%**
  - **Check against W's**
    - "Bump" at high MET are Ws identified as conversions
    - Gives a consistent answer

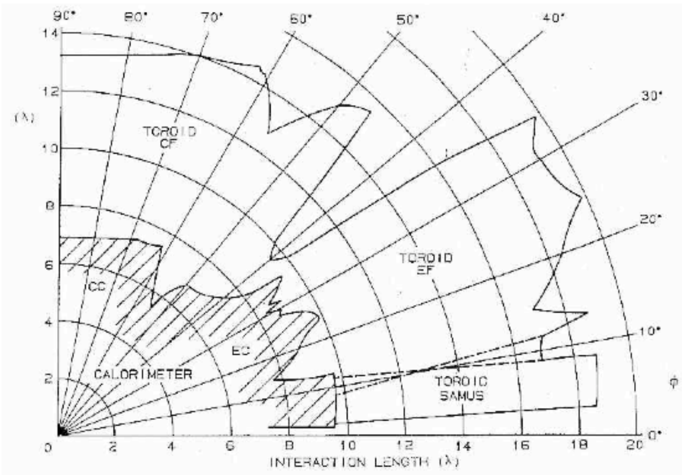




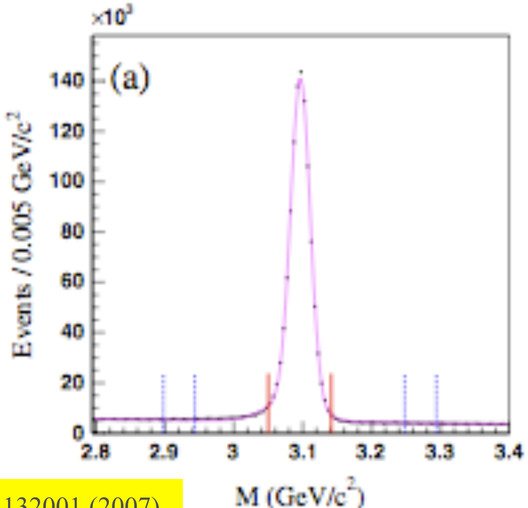


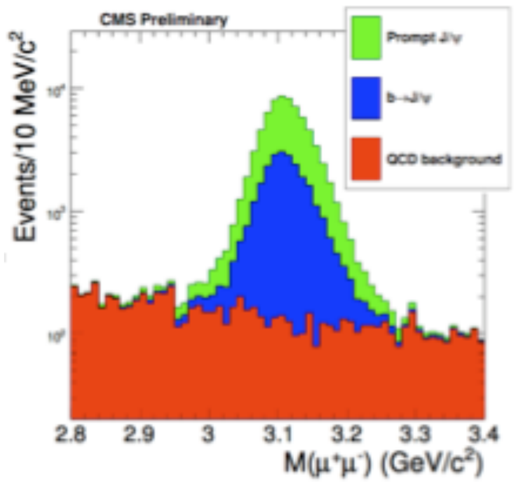
Muon Identification

- All techniques are based on highly-penetrating nature of γ
  - Have large amount of material (>10 λ)
    - Require min-ionizing particle
      - Calorimeter energy deposition
  - Track particles before and after material
    - Momentum analyze
  - Shown to be very effective
    - High rejection factors, especially with isolation
- Backgrounds are primarily
  - Cosmic rays
  - Decay-in-flight for lower momentum candidates
  - Size of background depends critically on other requirements



D0 Material Inventory vs Azimuth (quarter of detector)

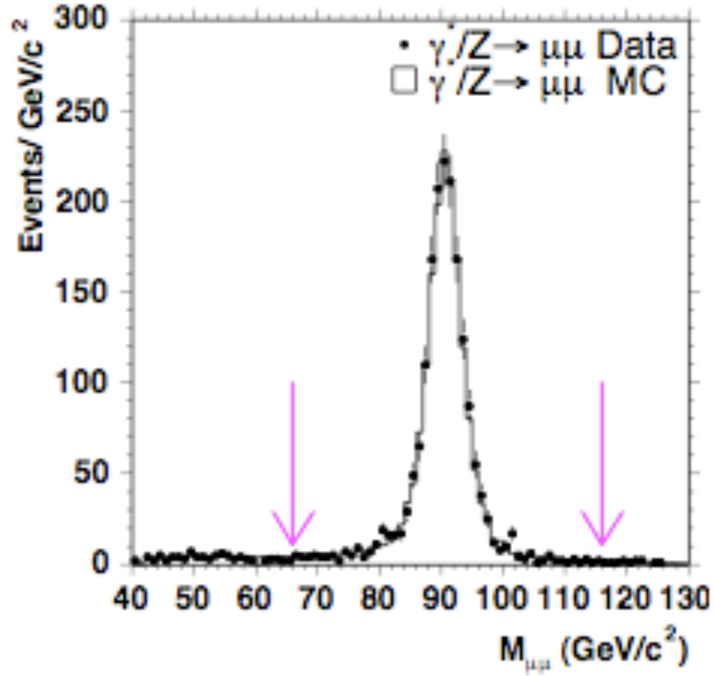




CDF, PRL 99, 132001 (2007)

Muon ID Efficiencies

- **Efficiencies for high-P_T μ determined from $Z \to \mu^+\mu^-$**
  - **Select events requiring one μ candidate $P_T > 20$ GeV/c**
  - **Look at efficiency of reconstructing second leg**
- **Can get MC/simulation to agree approximately**
  - **Predicts 92.1% efficiency, but measure 88.6±0.9%**
  - **Rejection hard to quantify**
    - Key question is "rejection from what?"
      - **In some sense, not relevant if one measures remnant background directly**
    - Limiting backgrounds are
      - **Cosmic rays**
      - **QCD jets "punching through"**

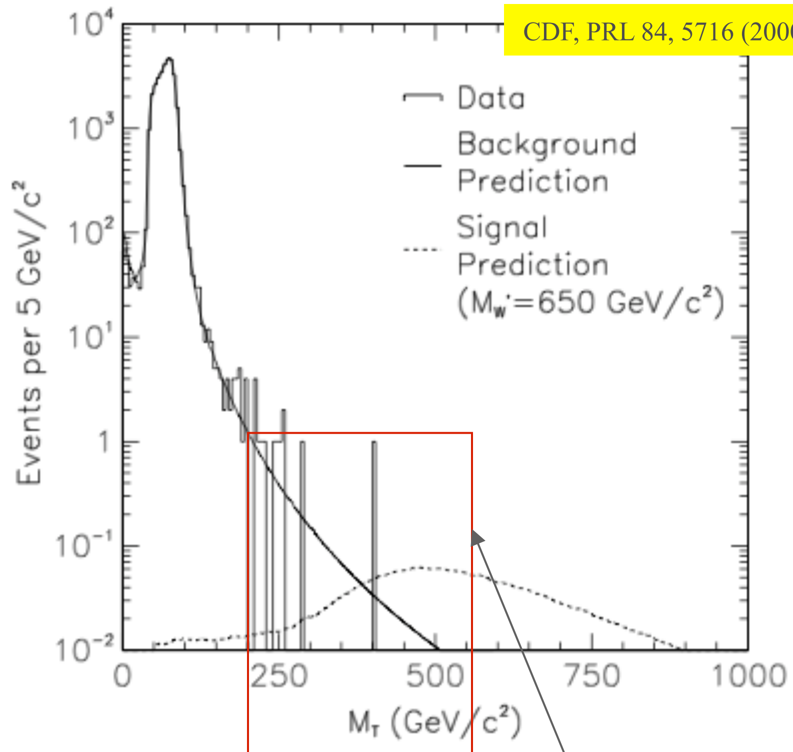


	data		MC	
No cuts applied	1153		28271	
Em Cut	1126	0.9766 ± 0.0044	27203	0.9622 ± 0.0011
Had Cut	1131	0.9809 ± 0.0040	27654	0.9782 ± 0.0009
Cot Cut	1128	0.9783 ± 0.0043	28226	0.9984 ± 0.0002
d0 Cut	1150	0.9974 ± 0.0015	28254	0.9994 ± 0.0001
isol Cut	1126	0.9766 ± 0.0045	27692	0.9795 ± 0.0008
dxemu Cut	1114	0.9662 ± 0.0053	28228	0.9985 ± 0.0002
all Cuts	1022	0.8864 ± 0.0093	26025	0.9206 ± 0.0016
all (w/o isol) Cuts	1044	0.9055 ± 0.0086	26523	0.9382 ± 0.0014

Backgrounds to Muons

- Backgrounds depend on the physics process & other requirements
  - Already see that dilepton signals are very clean
  - Probably most difficult region is high momentum
    - Example is W' -> $\mu\nu_\mu$
      - Only high P_T object in detector is μ candidate
    - Data comes from Run 1 with 100 pb^{-1}
- Difficult to find a signal limited by backgrounds!
  - Most backgrounds at high P_T are "intrinsic" -- ie., have a real μ

CDF, PRL 84, 5716 (2000)

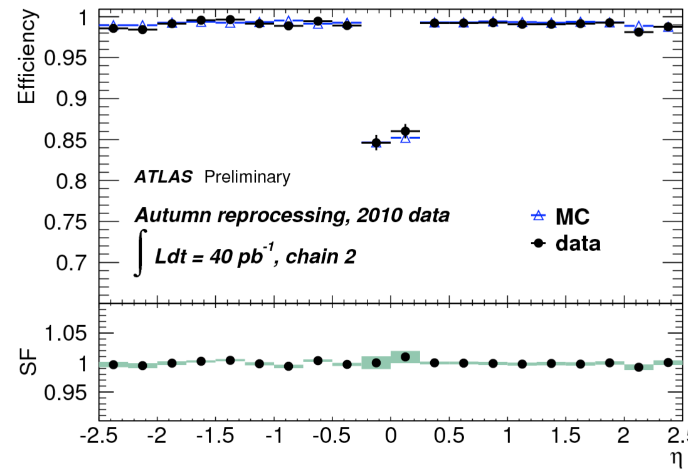


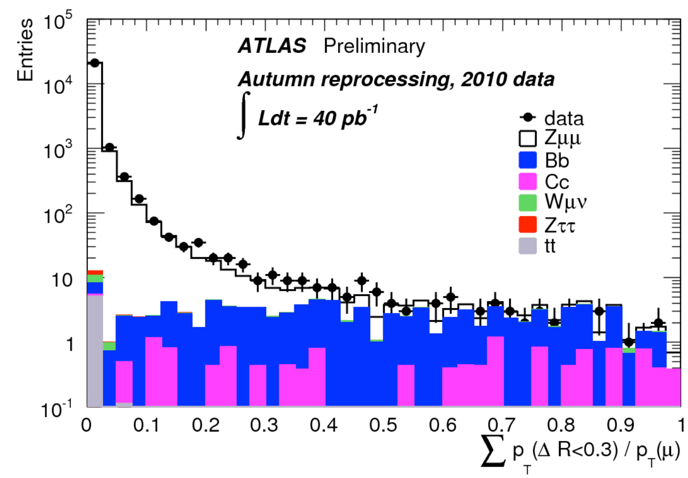
14 events M_T>200 GeV/c^2 and 12 expected (mostly W)

Muon Backgrounds at ATLAS

CDF, PRL 84, 5716 (2000)

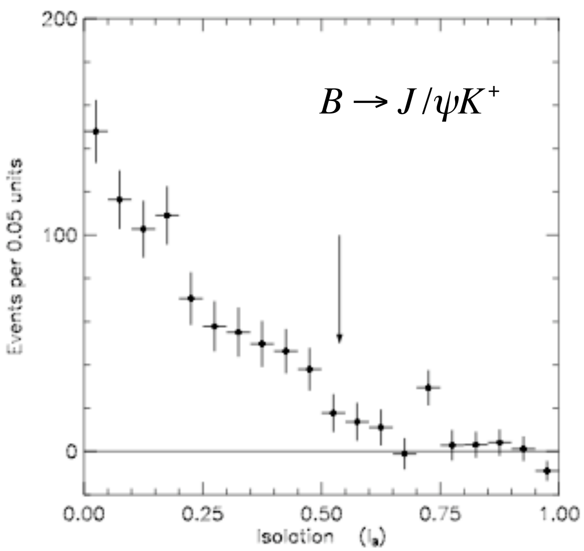
- **ATLAS muon reconstruction is intrinsically cleaner**
  - **The primary background sources are hadrons decaying in flight**
    - Become neglible at large p_T
  - **See this most clearly when looking at Z -> $\mu\mu$ decays**
    - Very clean sample, with less than 1% background from non-m
- **Efficiency overall is ~97%**
  - **Comes from a combined strategy of matching tracks in inner detector with muon spectrometer**
  - **This sample shows how effective "isolation" is in separating out the different sources**
    - Heavy flavour decay
    - Z and ttbar production





Comment on Isolation

- **Isolation requirements on lepton candidates appear to be powerful tool.**
- **Why?**
  - **Think about backgrounds:**
    - Jets faking leptons tend to be associated with additional particles
    - Semileptonic decays of b/c jets also have associated energy
  - **On the other hand, leptons from W/Z decay are generally isolated**
- **But many forms of "isolation cuts"**
  - **Some implicit**
    - Example: lepton ID criteria
  - **Some explicit**
    - Energy (or charged tracks) in a cone $\Delta R=0.2$ or $\Delta R=0.4$
      - **Cut on ratio of E_T in cone to lepton candidate (10% typical)**
- **However, there are many ways to look at isolation**
  - **Example comes from B -> $\psi(2S)\pi\pi$**
  - **Form cone of ΔR=1.0 around B candidate**
    - Sum up tracks not associated with B candidate
    - Reject events with $I_B>7/13$



$$I_B \equiv \frac{\sum_{i \notin B}^{R} \vec{p}_i \cdot \hat{p}_B}{|\vec{p}_B|},$$

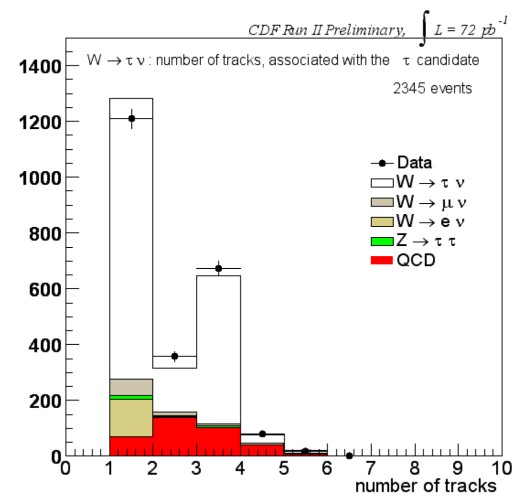
Tau Lepton Identification

Tau Listing, 2008 PDG

- **Tau leptons difficult to identify**
  - **Decay to either**
    - Leptonic final state (μ/e+$\nu\nu$)
      - **37% of time**
    - Hadronic final states
      - **12% with single charged particle**
      - **37% with h^- + neutral hadrons**
  - **Look for low-multiplicity "jets"**
    - Work to reconstruct π^0
    - Shower shape cuts to reduce QCD backgrounds
  - **Use track multiplicity to estimate observed yield**
    - 26 pb of W->$\tau\nu_\tau$
      - **Compare with 500 pb of W->eν_e**
      - **Factor of x20 lower efficiency**
    - Purity also about x10 worse
- **Means that tau physics has been "poor" cousin to electrons & muons**

Modes with one charged particle

particle$^-$ ≥ 0 neutrals $\geq 0 K^0 \nu_\tau$ ("1-prong")		(85.36 ± 0.08) %
particle$^-$ ≥ 0 neutrals $\geq 0 K_L^0 \nu_\tau$		(84.73 ± 0.08) %
$\mu^- \bar{\nu}_\mu \nu_\tau$	[g]	(17.36 ± 0.05) %
$\mu^- \bar{\nu}_\mu \nu_\tau \gamma$	[e]	$(3.6\pm0.4) \times 10^{-3}$
$e^- \bar{\nu}_e \nu_\tau$	[g]	(17.85 ± 0.05) %
$e^- \bar{\nu}_e \nu_\tau \gamma$	[e]	(1.75 ± 0.18) %
$h^- \geq 0 K_L^0 \nu_\tau$		(12.13 ± 0.07) %
$h^- \nu_\tau$		(11.60 ± 0.06) %
$\pi^- \nu_\tau$	[g]	(10.91 ± 0.07) %
$K^- \nu_\tau$	[g]	$(6.95\pm0.23) \times 10^{-3}$
$h^- \geq 1$ neutrals ν_τ		(37.08 ± 0.11) %
$h^- \geq 1\pi^0 \nu_\tau$ (ex. K^0)		(36.54 ± 0.11) %
$h^- \pi^0 \nu_\tau$		(25.95 ± 0.10) %
$\pi^- \pi^0 \nu_\tau$	[g]	(25.52 ± 0.10) %



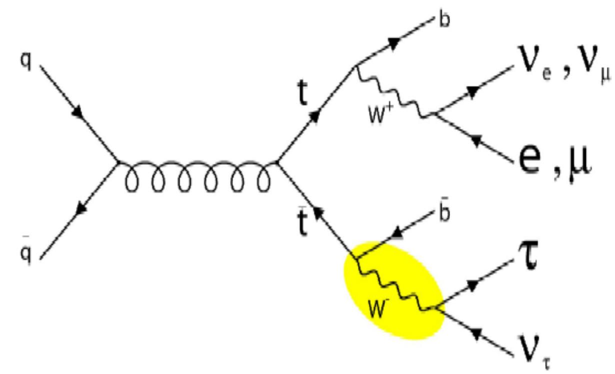
A. Safanov (for CDF Collaboration), Nucl. Phys. B 144, 323 (2005).

Tau Lepton Reconstruction

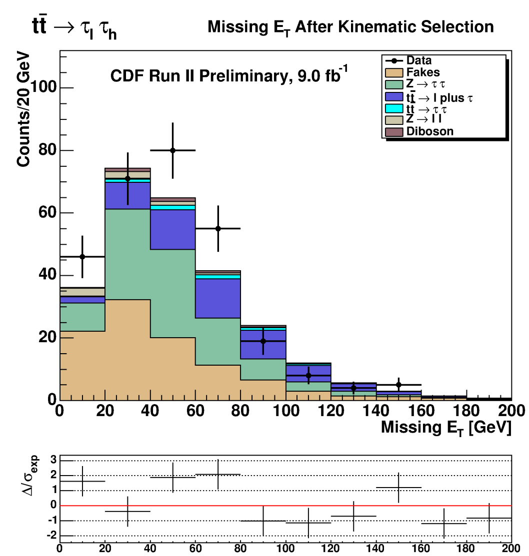
- **Why is τ reconstruction so lousy?**
  - **Have at least one ν, sometimes several**
    - Compromises energy measurement
    - Reduces energy scale (and efficiency)
  - **Reject decays to $l\nu_l\nu_\tau$**
    - Background from leptons too large
  - **Reliance on charged tracking information and π^0 reconstruction**
    - Hit by BR and reconstruction efficiencies
  - **Trigger is less efficient**
    - Presence of ν in effect pushes up the minimum ν P_T
    - Work to add other information
      - **τ + MET trigger**
      - **τ + lepton trigger**
      - **Help but don't solve the fundamental problem**
- **Criteria for τ identification**
  - **1 charged track + evidence of π^0**
    - Requires reconstruction of π^0 in EM calorimeter
    - BR ~ 1/3!
  - **Look for "narrow" jet**
    - Seed tower E_T>6 GeV
    - Seed track P_T>4.5 GeV/c
    - <=6 towers with E_T>1 GeV in cluster
    - Overall efficiency of ~50%
  - **A further "isolation" cut to reduce backgrounds from QCD jets**
    - Typical cut: E_T in cone R=0.4 < 10% of τ candidate E_T
    - About 60-70% efficient
- **Loss of x10 compared with e or μ**
  - **And backgrounds still high**

Example: Top decaying to $\tau\nu_\tau b$

- Top quark decays to τ lepton should be observable
  - Also key signature for H^+
- Analysis strategy
  - Look for isolated e or μ
  - Isolated τ candidate
  - MET > GeV
  - >=2 jets
    - Leading jet $E_T > 25$ GeV and 2nd jet $E_T > 15$ GeV
    - Reduce Z->$\tau\tau$
  - Require significant energy in event
    - $H_T > 205$ GeV



http://www-cdf.fnal.gov/physics/new/top/2012/ttbar_taulep_xsec_9invfb/Publicpage.html

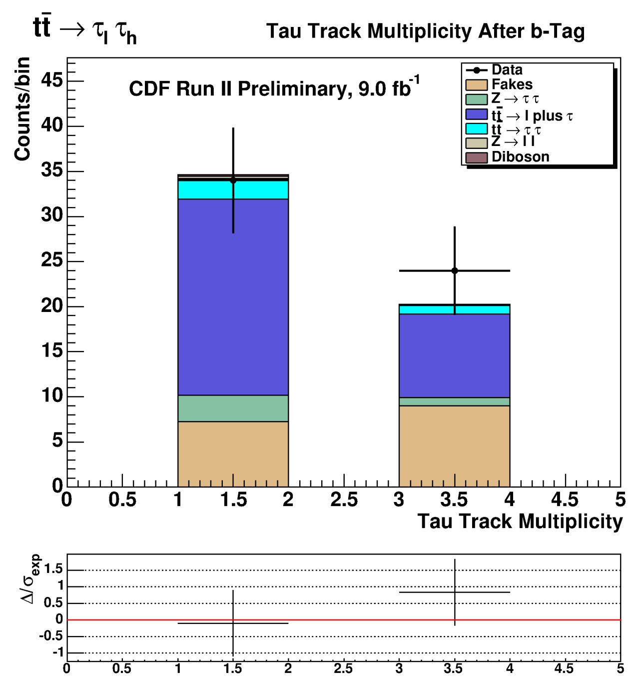


- **MET distribution for electron+τ**
  - Kinematic cuts
  - Require >=2 jet

CDF Results

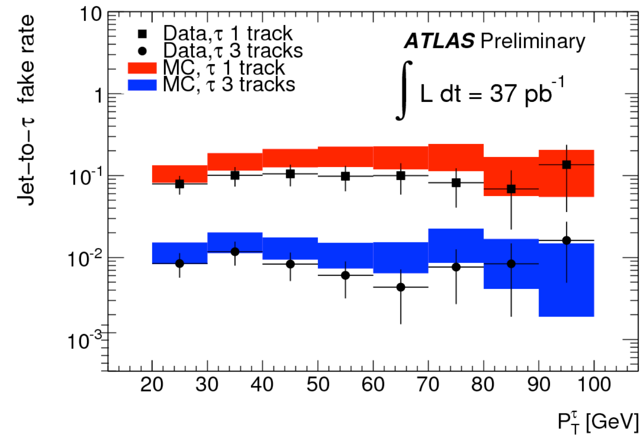
- **Backgrounds dominated by "fake" τ candidates**
  - **To estimate, use dijet data**
    - Create "fake matrix" that gives probability of jet passing τ criteria
    - Have to be careful about "denominator"
      - **Also correlations with rest of event**
      - **Primary background from W+jets**
- **See 290 candidate events in 9 fb^{-1}**
  - **Expect 60 from top quark production**
- **This is hard!**
  - **ee/eμ/$\mu\mu$ + 2 jets (1 b tag) has 80 candidate events with 2.8 fb^{-1}**
    - Estimate 4 background!
  - **Guess that ~8(?) of these are from ttbar -> $\tau\tau$bb**

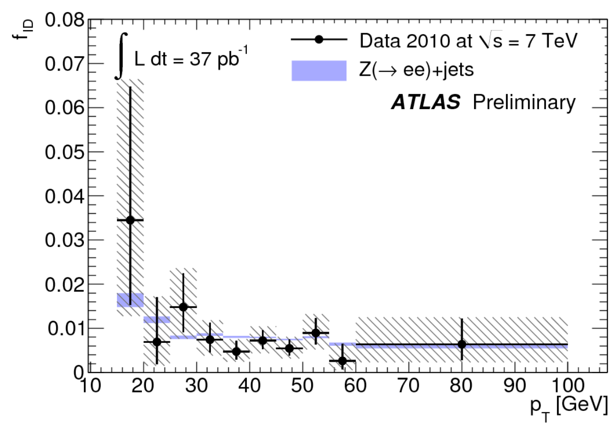
	e+tau (events)	mu + tau (events)
Jet -> tau	65±14	36±8
Drell Yan	54±11	47±10
Top	33±3	26±3
WW	2.4±0.3	1.8±0.3
Total expected	154±20	110±14
Observed	175	115



LHC τ's are not poor cousins

- Use similar techniques to separate τ candidate
  - The width of the EM energy deposition
  - The width of the charged track energy deposition
  - The invariant mass of the charged tracks
- Can measure rejection rates from jets and electrons
  - Get rejection rates of order 20-100
  - With these, can see a clear Z -> ττ -> eμ signal
    - A little bit of a cheat, as it only looks for events with $25 < m_{e\mu} < 80$ GeV/c^2





- A good start allowing one to build toward a full analysis
- Efficiencies only 10-20% that for electrons, muons

ATLAS-CONF-2011-045