New Cut-Based Electron Id

Matteo Sani (UCSD)
with collaborators from UCSD, UCSB, FNAL

Introduction

- Electrons can be categories according to ECAL and tracker properties, two classification methods proposed so far:
  - tracker based, p_{in} vs p_{out} comparison (fBrem):
    http://indico.cern.ch/conferenceDisplay.py?confId=18239
  - ECAL cluster shape (singal/multi cluster) and tracker based (fBrem):
    CMS Note 2006-040
- At the startup period we want a robust and simple electron selection until we have data to verify and tune.
- We have so designed a simple electron selection:
  - relying on most predictable and stable electron variable
  - aiming for high selection efficiency, 97%
  - improving electron categories idea already implemented

Cut-based Electron Id

- Three level of e-ID has been defined:
  - "robust", based on four simple cut
  - "loose", based on five category dependent cuts
  - "tight", like "loose" but with tighter thresholds (currently under study)
- The levels are defined such that events passing "higher" levels are a subset of those passing the previous.

PTDR Electron Id

- We compare the results of the new selection criteria to the old cut-based e-ID, used in PTDR studies.
- The PTDR e-ID relies on:
  - Match in ϕ and η between supercluster and track
  - Loose E_{SC}/p_{in} and E_{seed}/p_{out} cuts
  - Shower shape requirements
  - H/E cut
- Electrons are divided into four different categories.

Remove some producer cuts

➔Some producer cuts are too tight:

```
double maxEOverPBarrel = 3.
double maxEOverPEndcaps = 5.
double minEOverPBarrel = 0.35
double minEOverPEndcaps = 0.35
double maxHOverE = 0.2
double maxDeltaEta = 0.02
double maxDeltaPhi = 0.1
double ptCut = 5. ▶ 1.5
```

keep these

In future the last cut will be replaced by E_T(SC) > 5 GeV (proposed by C. Charlot):

- track p_T has higher chance of being underestimated
- fakes are more likely to have low E_T then low p_T (due to significant population with E/p<1)

This increases the denominator for selection efficiency calculations and therefore decreases the efficiency calculated.

"Robust" Selection

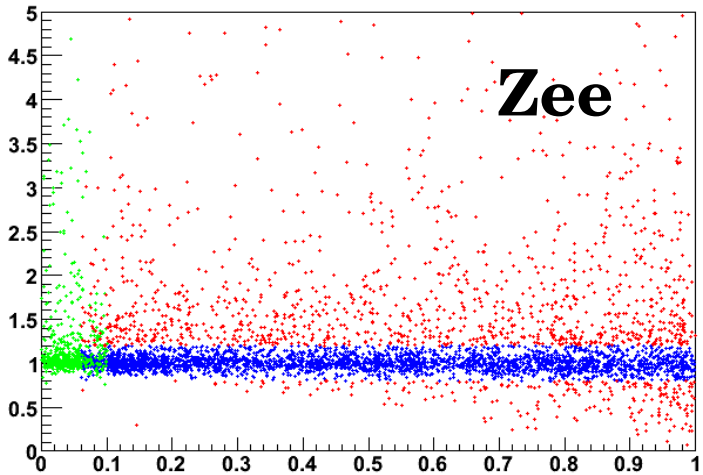
- Simple and robust e-ID for startup.
- No classification, use variables which are independent of measured brem fraction and insensitive to tracker misallignement.
  - $\sigma_{\eta\eta}$: shower shape σ along η
  - $\Delta\eta_{in}$: delta eta between track ans supercluster
  - $\Delta\phi_{in}$: delta phi between track and supercluster
  - H/E: hadronic over electromagnetic energy

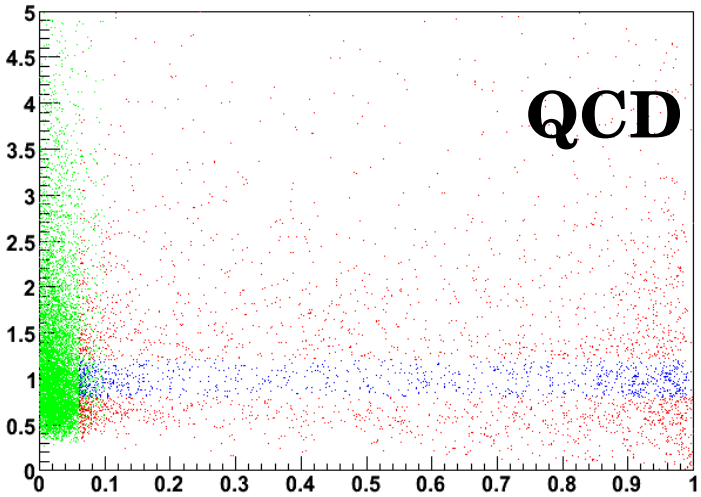
With straight cuts on these quantities, we get ~97% selection efficiency and ~4% of fakes (dijet 50-80 events).

"Loose" Selection

➔ Introduces classification with three classes based on the E/p and fBrem:

- bremming electrons with E/p ~ 1 (little contamination from fakes)
- low brem electrons (high population from both real and fake)
- bad track, E/p ≠ 1





Physics motivations

1. E/p is often well measured for electrons
2. Electrons usually radiate a good deal of energy in the tracker
3. E/p is not often measured to be less than 1 for electrons

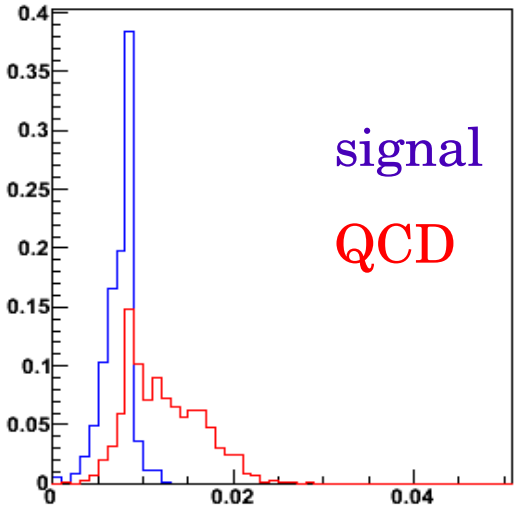
1. Fakes from jets usually have fBrem around 0 (just charged pion tracks…)
2. Many fakes from jets have E/p<1 partly because of the low response of ECAL to charged pions...

The categories

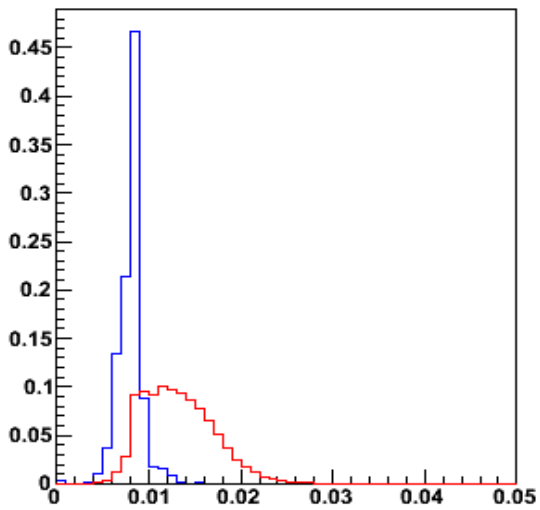
- The new categorization allows:
  - robust separation of regions since p_{in} tracks should have fBrem ≈0.
  - no dependence on clustering algorithms
- The three categories are just used for the selection.
- There is no reason we need to use them for later analysis.
- Old categorization is still accessible through electron **`Data Format`** and used for other purposes (for example energy corrections).

$\sigma_{\eta\eta}$, H/E in categories

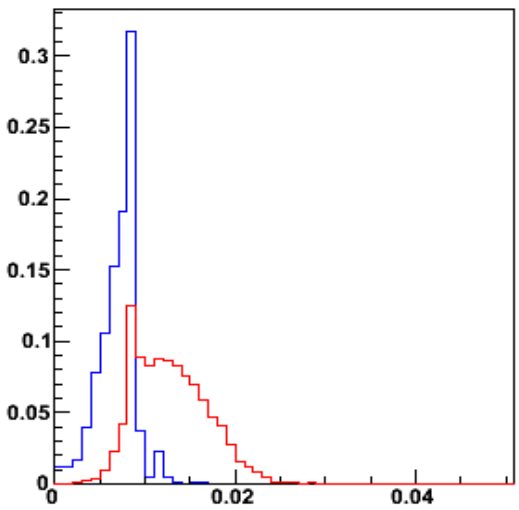
$\sigma_{\eta\eta}$



0

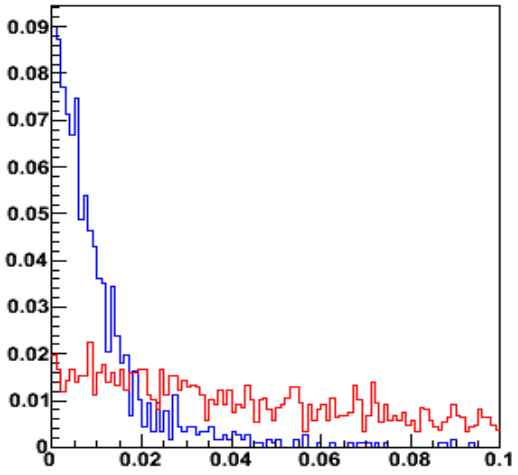


1

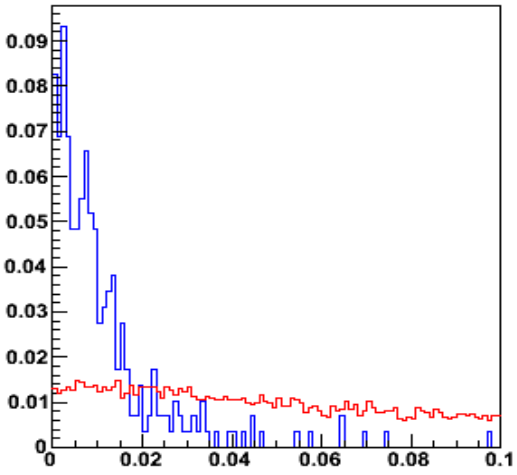


2

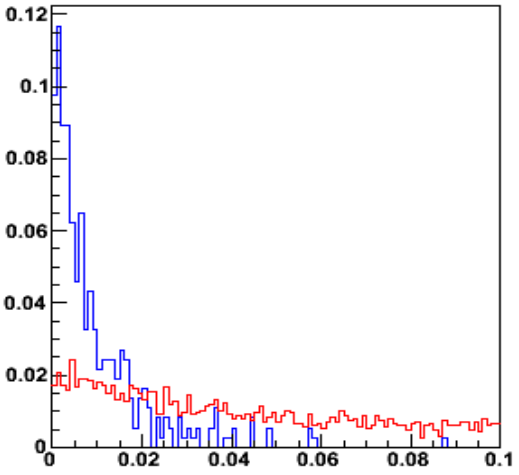
H/E



0



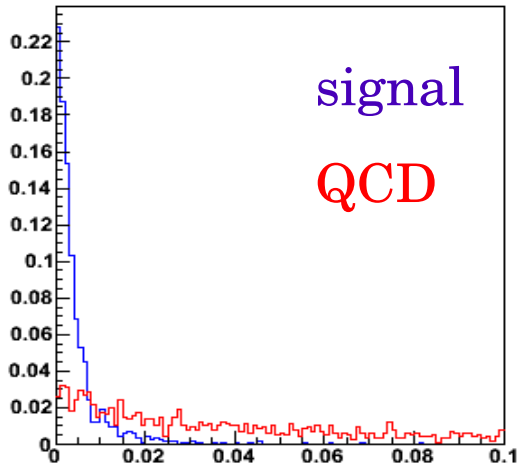
1



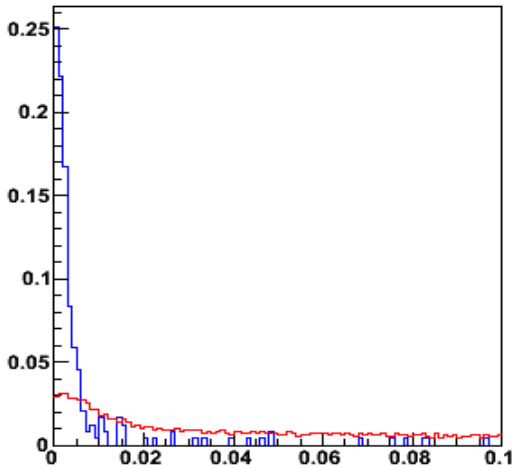
2

$\Delta\phi_{in}$, $\Delta\eta_{in}$ in categories

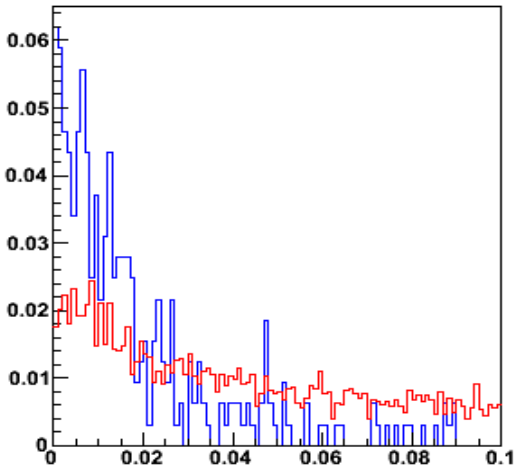
$\Delta\phi_{in}$



0

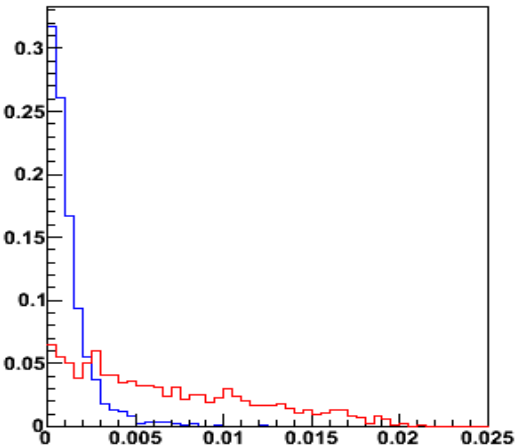


1

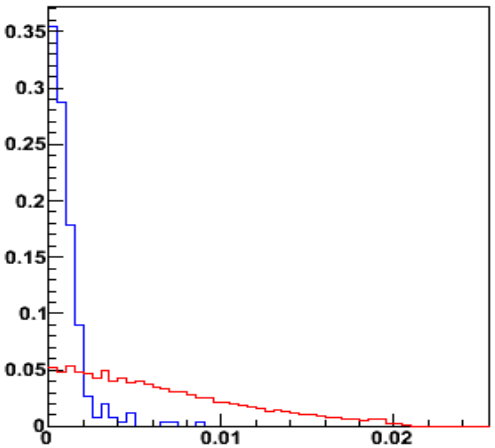


2

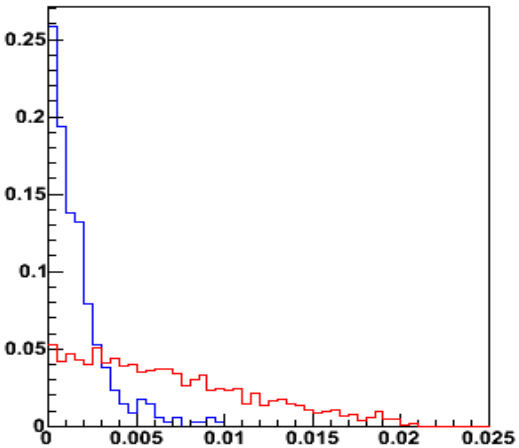
$\Delta\eta_{in}$



0

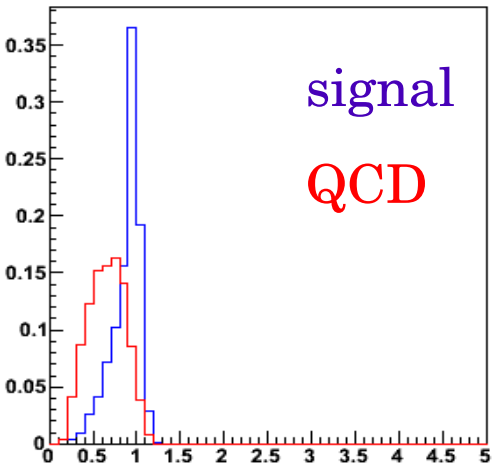


1

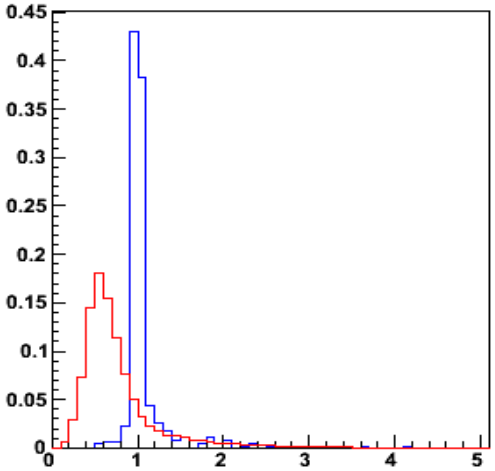


2

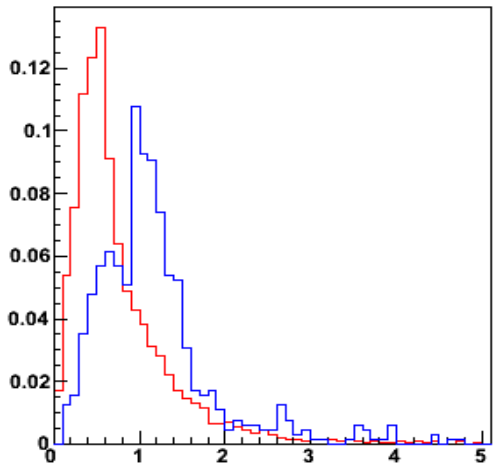
E_{seed}/p_{in} in categories



0

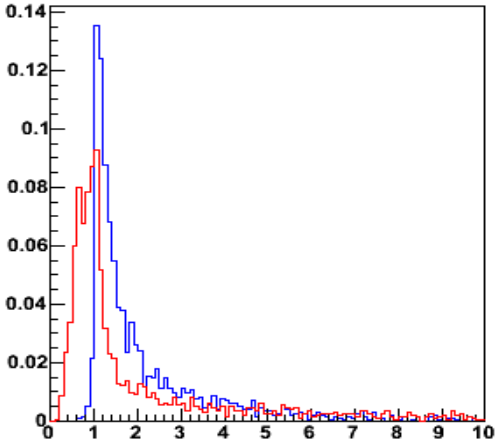


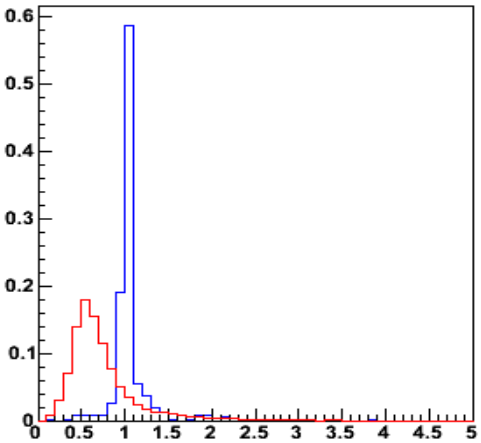
1

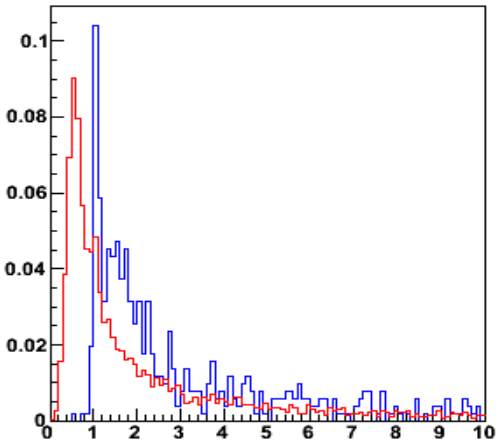


2

E_{seed}/p_{in}







E_{seed}/p_{out}

E_{seed}/p_{in} has a lower discriminant power but the cuts are set such that this one does much to the overlap category where $p_{out} \sim p_{in}$.

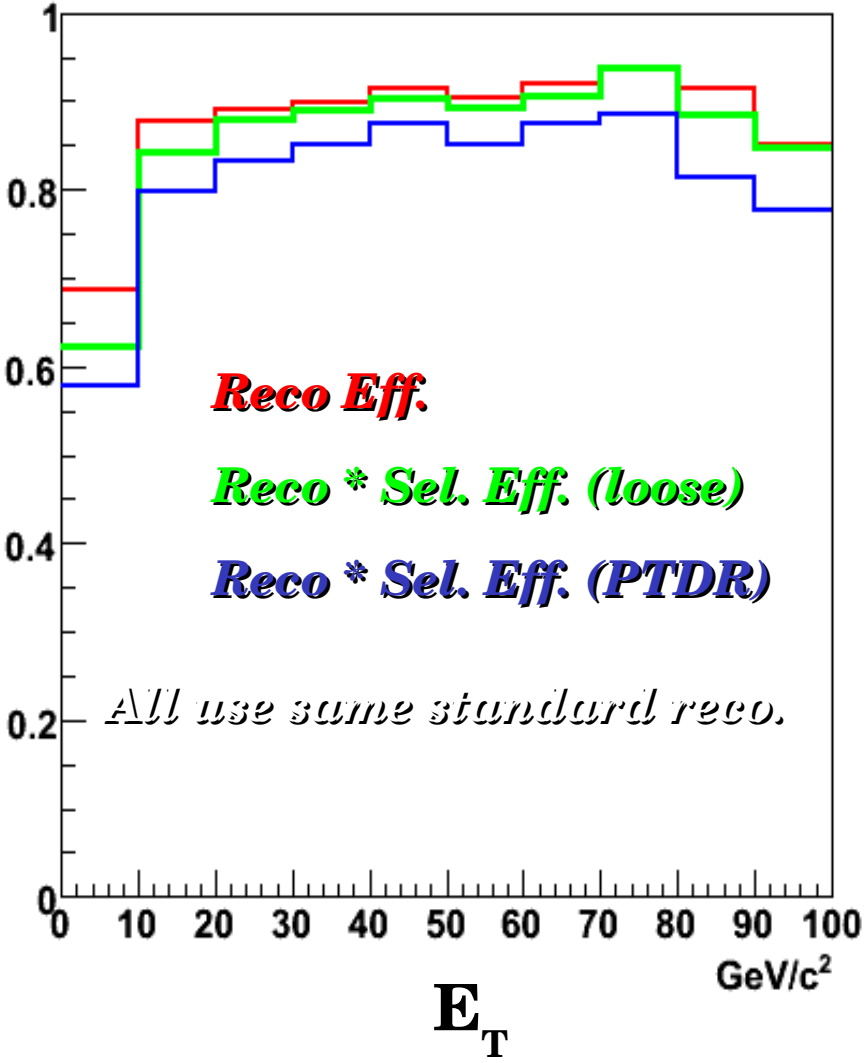
Results (e-ID eff.)

	"robust"	"loose"	PTDR loose
Efficiency from Z→ee events	97.7%	97.9%	93.6%
Fake rate in QCD_50_80	4.1%	1.5%	6.5%

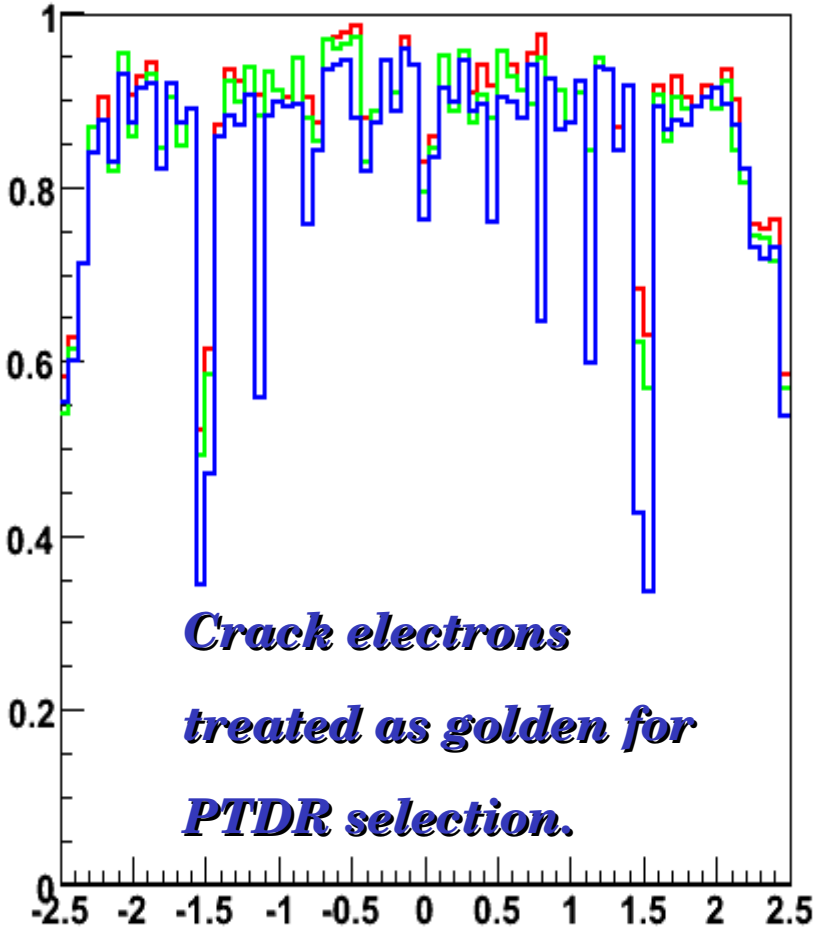
Not completely fair: all crack electrons are accepted in the standard selection here are treated as golden.

NO ISOLATION CUT APPLIED !!!

Efficiency



E_T



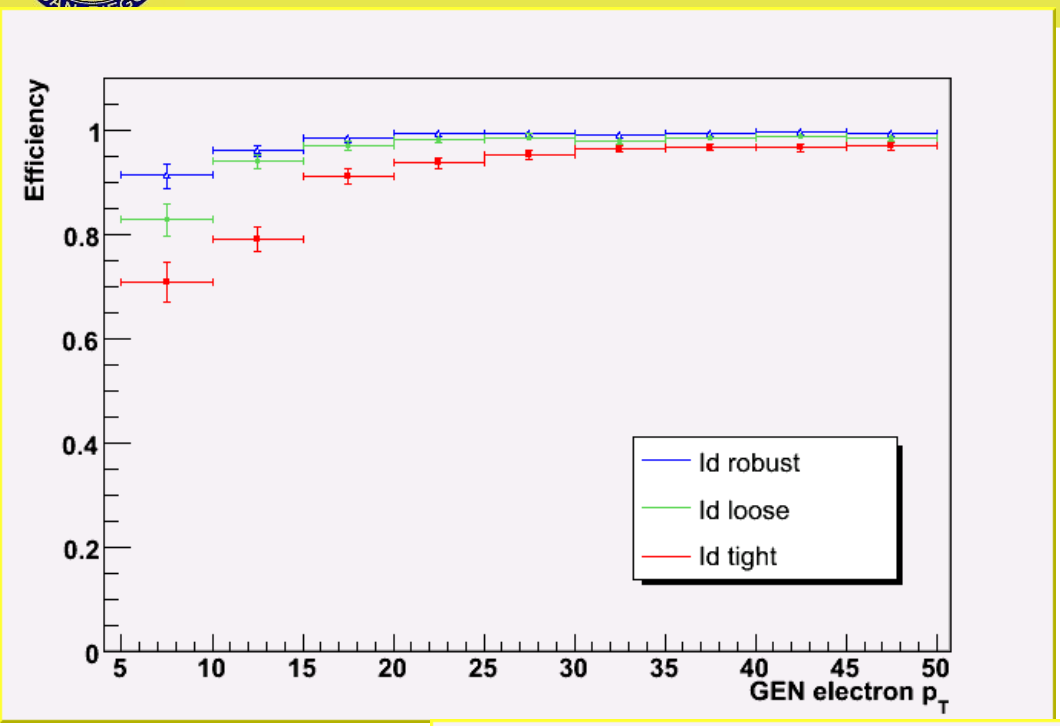
η

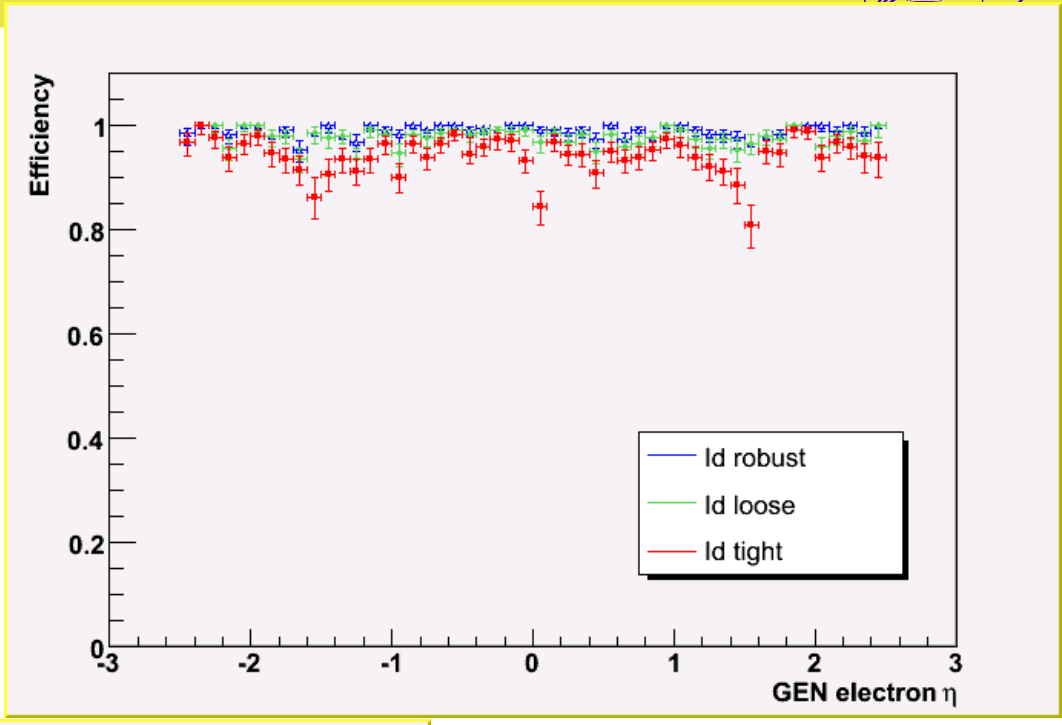
Tight selection

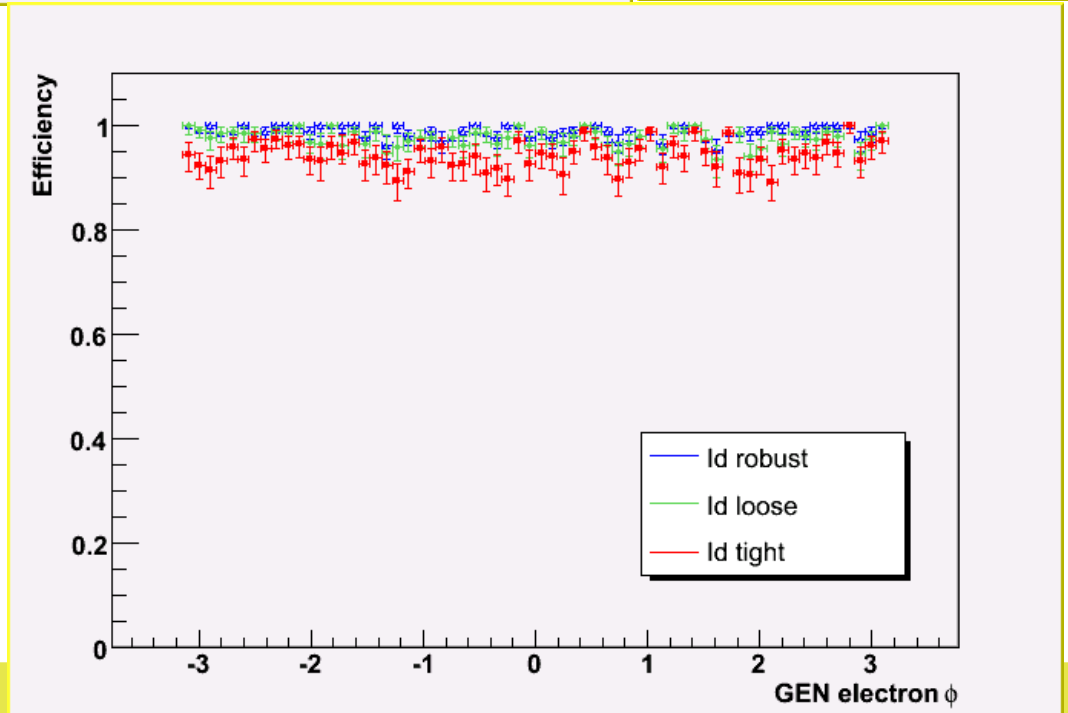
➔Since this presentation I've defined the thresholds for a tight selection which achieve:

- factor 2 rejection
- ~3% signal reduction

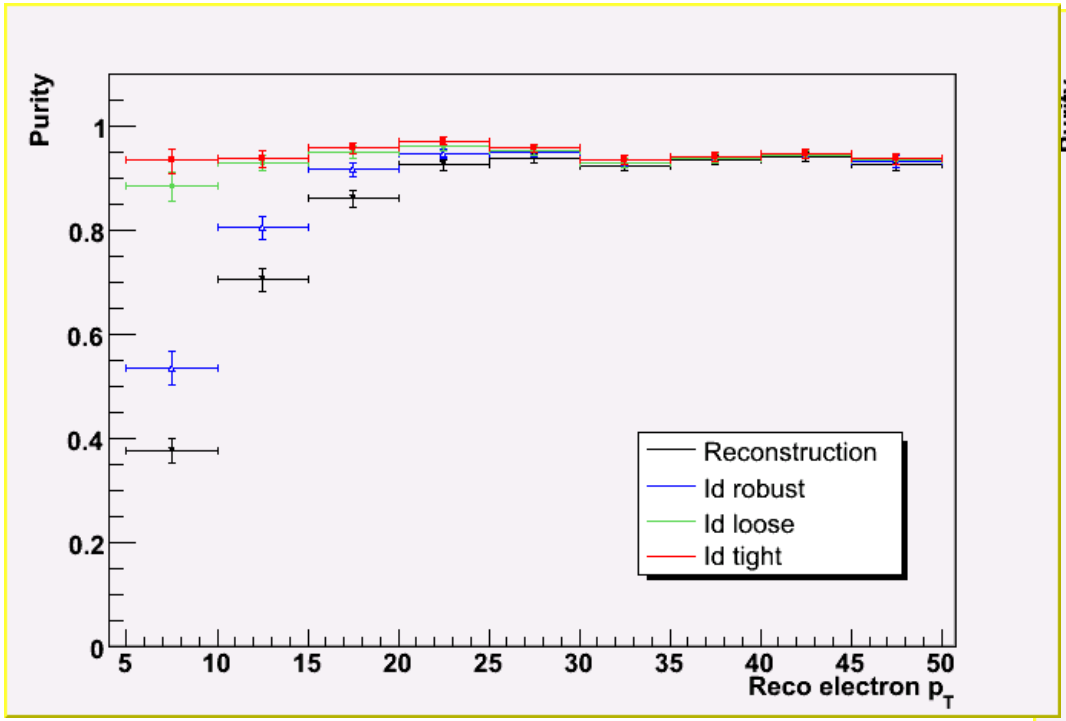
Pedro's tools

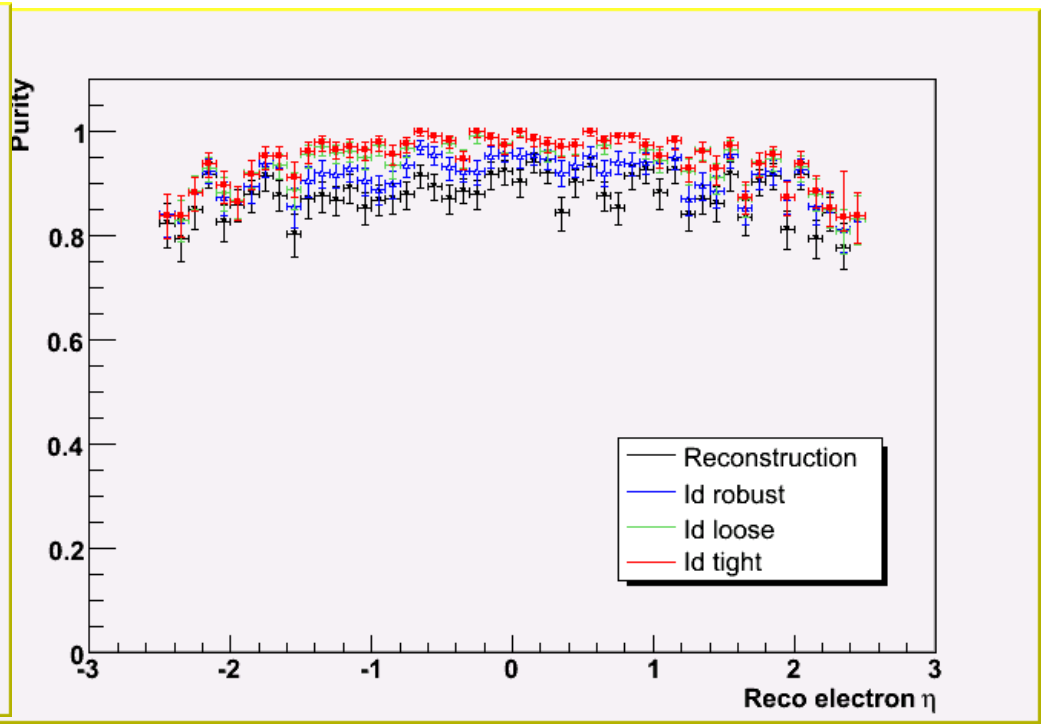






Pedro's tools



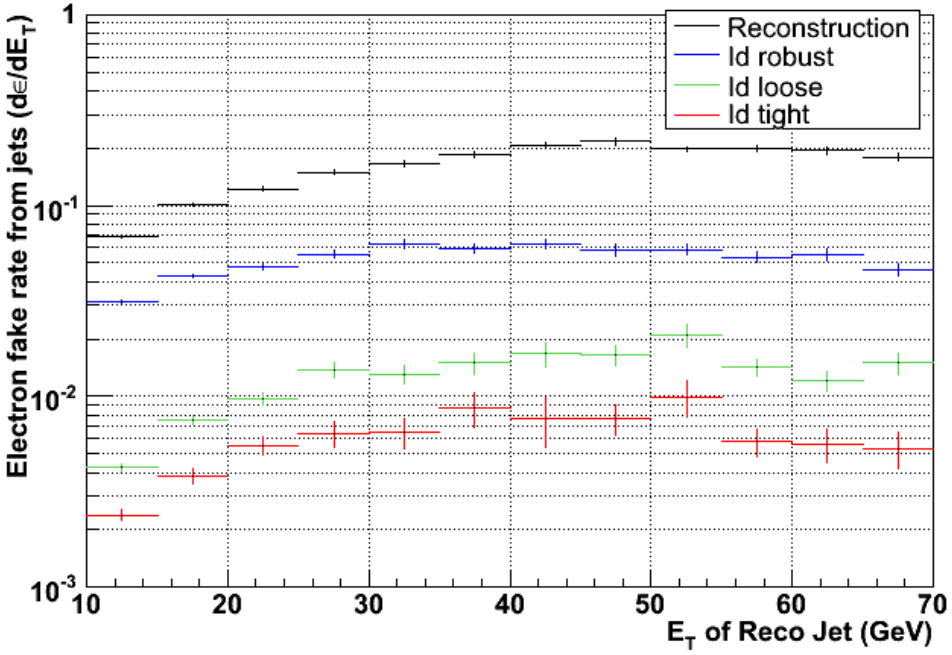


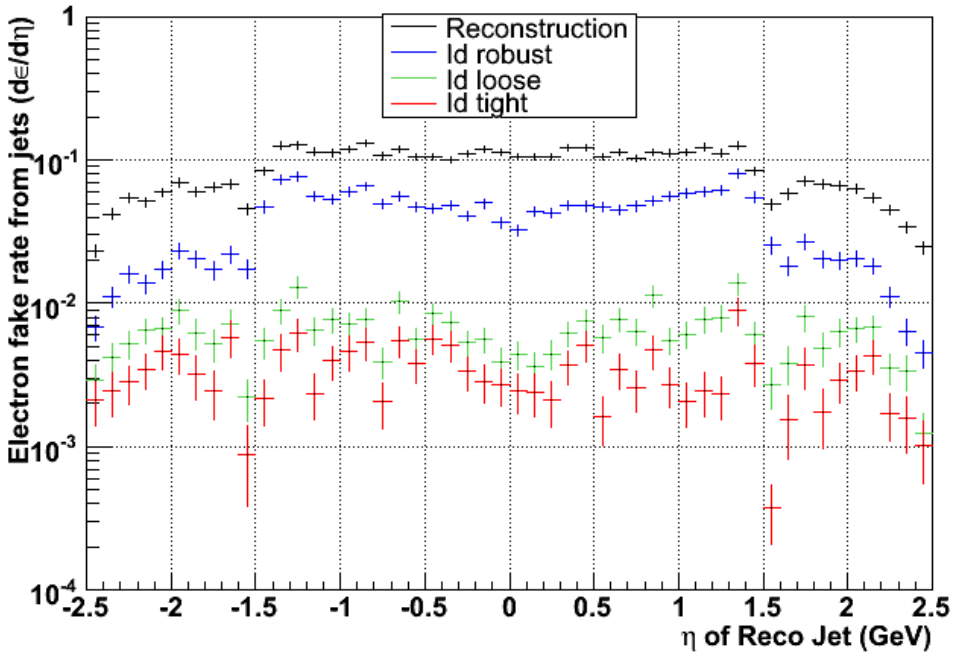
Pedro's tools

Sample : Pythia 2->2 QCD , 20 Gev/c <pt_hat < 120 GeV/c

(CMSSW_1_5_2)

Fake rate per reco jet (iterative cone = 0.5, pT > 10 GeV)





Conclusions

- Simple cuts over variables independent of measured brem fraction and insensitive to tracker misallignement are use with the robust e-ID selection (already good results).
- The loose e-ID based on a categorization uses E/p and fBrem:
  - Electrons and fakes separate to a great extent in this categorization
  - 98% efficiency for electrons and low fake rate

eID News

- Thursday 14 there shoukd have been a discussion on eID issues in view of the CMSSW2 release.
- Skipping technicalities the discussed items were:
  - definition of a startup “tight” electron.
  - implement cut on IP (proposed 0.2 mm from Puneeth study)
  - strategy to reject converted photons
  - study sources of fakes and their relative importance: heavy flavours, photon conversion, pi0/pi+...

From Pascal mail.

eID News

- Fake rate definition (fake as a function of SC ET looks better).
- Study Ecal based variables alternative to f_brem to measure brem fraction (sigma_phi/sigma_eta).
- Crack electrons: define sensible eID strategy.
- Document performance of new Cut-Based ID into a note:
  - I've been contacted to write the algorithm description (in the next days I'll send you the first draft)

From Pascal mail.