Electron reconstruction

Matteo Sani, Avi Yagil, Boris Mangano
UCSD

Introduction

- To improve physics analysis it would be worth to have the electron reconstruction efficiency as high as possible.
- We tried to develop a new reconstruction "sequence" to achieve this improvement.
- Doing so:
  - we tried not to do too much coding (few new lines of code have been written)
  - re-use existing objects and mainly re-arrangement of already available tools and changes in configurations
  - the resulting sequence is the **`GlobalGsfElectrons(GGE)`**

Goal and Implementation

➔ The main goal of this study is to maximize the efficiency of the producer:

- electrons which are rejected/not reconstructed by the producer cannot be recovered by any subsequent electron ID module or analysis
- the efficiency/fake-rate trade-off can be optimized later according to the specific channel under investigation

➔ Implementation overview:

- use standard CTF mixed seeding inside electron reconstruction
- loosen some of the production cuts
- for more details see talks:
  - Matteo Sani, EGamma POG 08/06/2007
  - Boris Mangano, CMS Higgs WG 10/11/2007

Environment

- The performance of the two sequences have been compared in **`CMSSW_1_8_0_pre6`**:
  - for the signal we have used Z→ee events of the RelVal180pre4
  - as a background QCD_pT_50_80 events from the RelVal180pre4 also.
- Efficiency and fake-rate have been evaluated using Pedro's tool to use “standard” definitions.
- Thanks to Ursula, the Electron producer has been modified so that changing only one cfg parameter we can switch from an algorithm to the other.

Standard CMS Electron Sequence

Super Cluster reconstruction → track-seeds generator (superCluster driven) → CTF trajectory builder ("adapted" to electron) → GSF track fitter → Electron producer

CMS electron sequence

track-seeds generator:

- the track-seeds are produced by pairs of pixel (<u>and only pixel</u>) hits
- a tight matching between seeds and cluster is applied

CTF trajectory builder:

<u>the same as standard CTF</u> tracking, but:

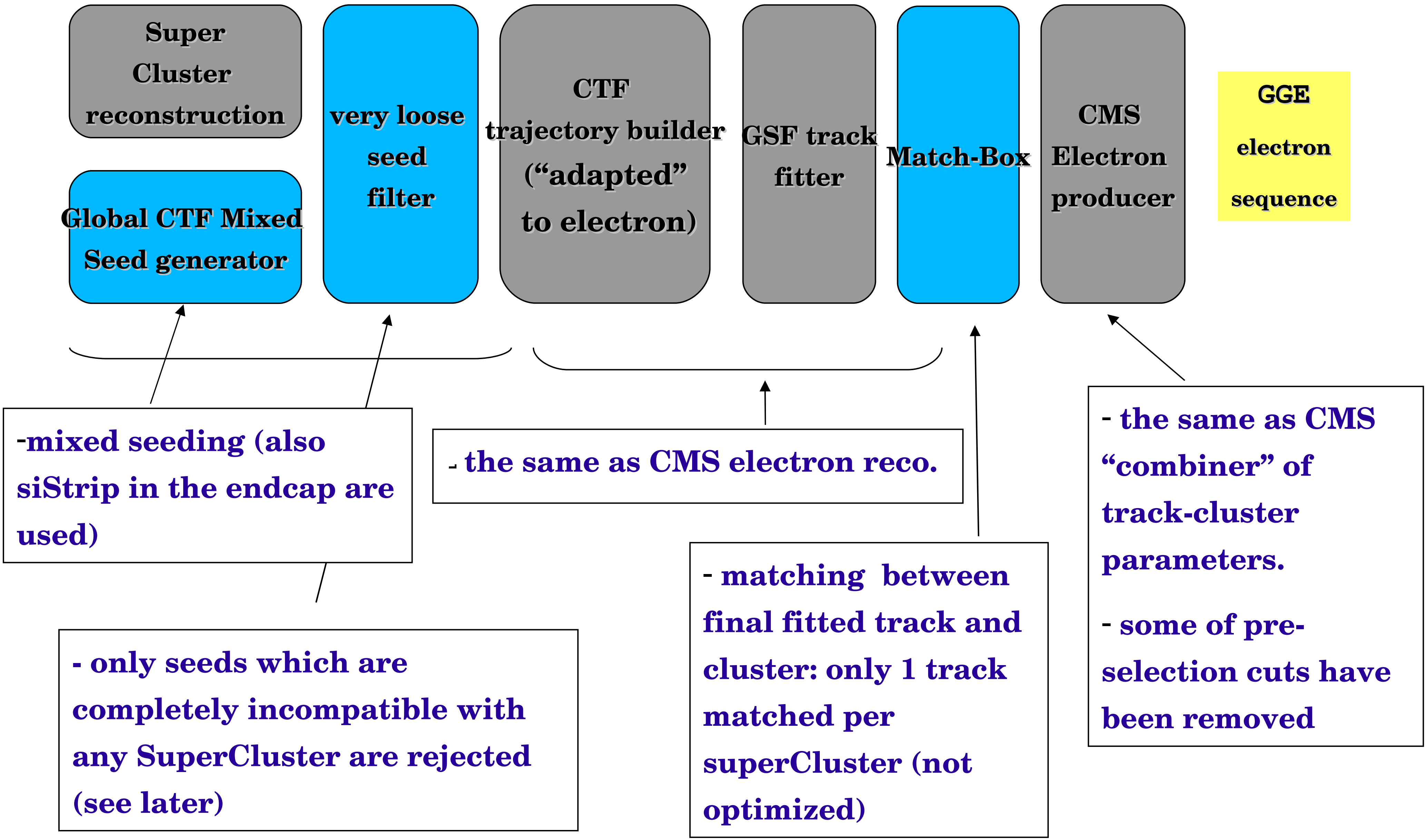
- propagator uses different formula for energy loss
- cut on hit compatibility is loosen; number of branches during building is reduced

<u>No GSF is involved here</u>

Electron producer:

- track's momentum and cluster's energy are combined.
- <u>pre-selection cuts are applied</u>

GlobalGsfElectrons Sequence



Differences between the 2 Sequences

- Mixed seeding: **GGE** reconstruction can recover electrons which cross just 1 or 0 pixel detectors.
- Matching: **GGE** matched the final fitted tracks with the SC. In doing so the more precise final track information was used.
- Removing of selection cuts, some default cuts removed from the producer:

Very loose cut not relevant for signal eff.

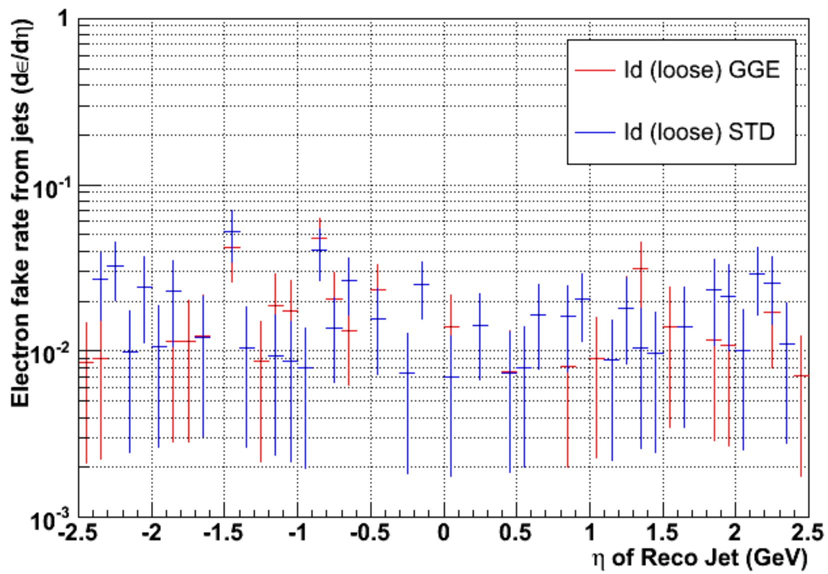
```
double maxEOverPBarrel = 3.
double maxEOverPEndcaps = 5.
double minEOverPBarrel = 0.35
double minEOverPEndcaps = 0.35
double maxHOverE = 0.2
double maxDeltaEta = 0.02
double maxDeltaPhi = 0.1
double EtCut = 5. → 0.
```

We keep these (maxHOverE, maxDeltaEta, maxDeltaPhi)

E_T cut removal

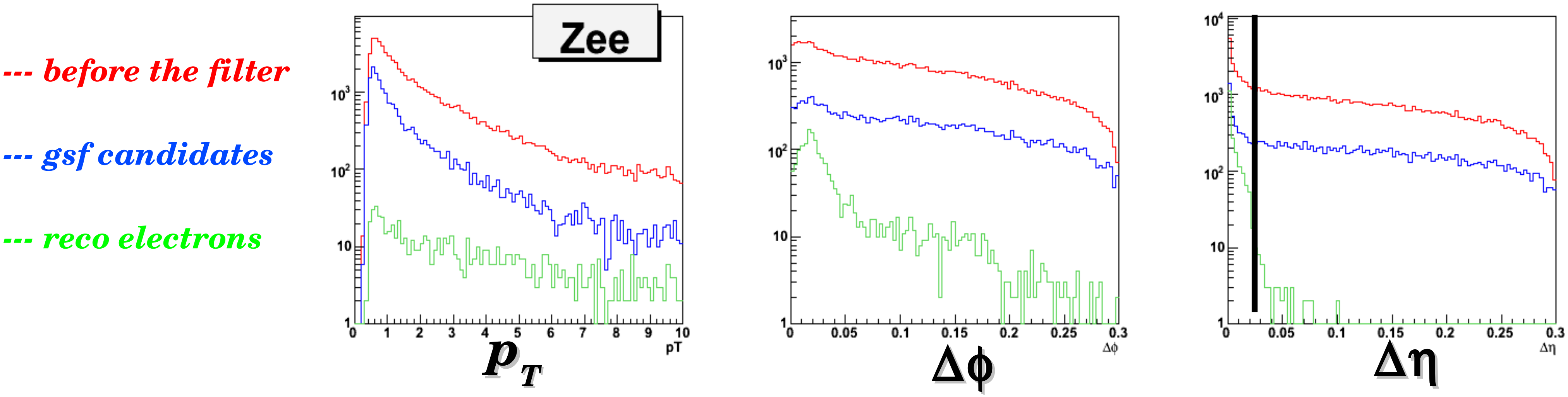
- Optimization of preselection cuts is needed
  - simplest solution seems GlobalGSFElectrons cuts
  - $E_T > 5$ probably needs to be retained to prevent number of candidate (especially for low pt-hat bins)
- For the moment I checked fake rate (without eID) in QCD_30_50 the lowest available and it is under control.

From Seez report.



Timing

- Originally the very loose seed filter (seed to cluster matching) led to a huge number of candidate to be propagate spending too much time in track reconstruction.



- A tighter match in Δη (0.025) greatly reduce the number of candidate so that now the two algorithms have similar timing.

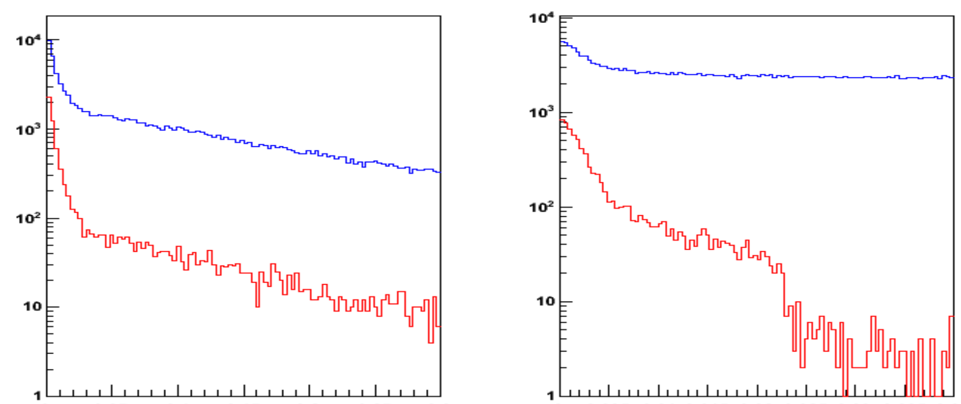
Zee events		
	dEta < 0.3	dEta < 0.025
GGE	0.7 s	0.21 s
STD	0.16 s	0.16 s

QCD events		
	dEta < 0.3	dEta < 0.025
GGE	1.1 s	0.54 s
STD	0.18 s	0.18 s

ϕ, η window

- ➔ We realized that we were not using an extrapolation of the impact point at calorimeter to compute $\Delta\eta$ and $\Delta\phi$ calo/seed.
- ➔ Boris recently implemented this and I'm currently running on Zee:
  - to optimize thresholds again
  - to compute eff, fake and timing (expected same eff with lower timing)

$\Delta\phi$



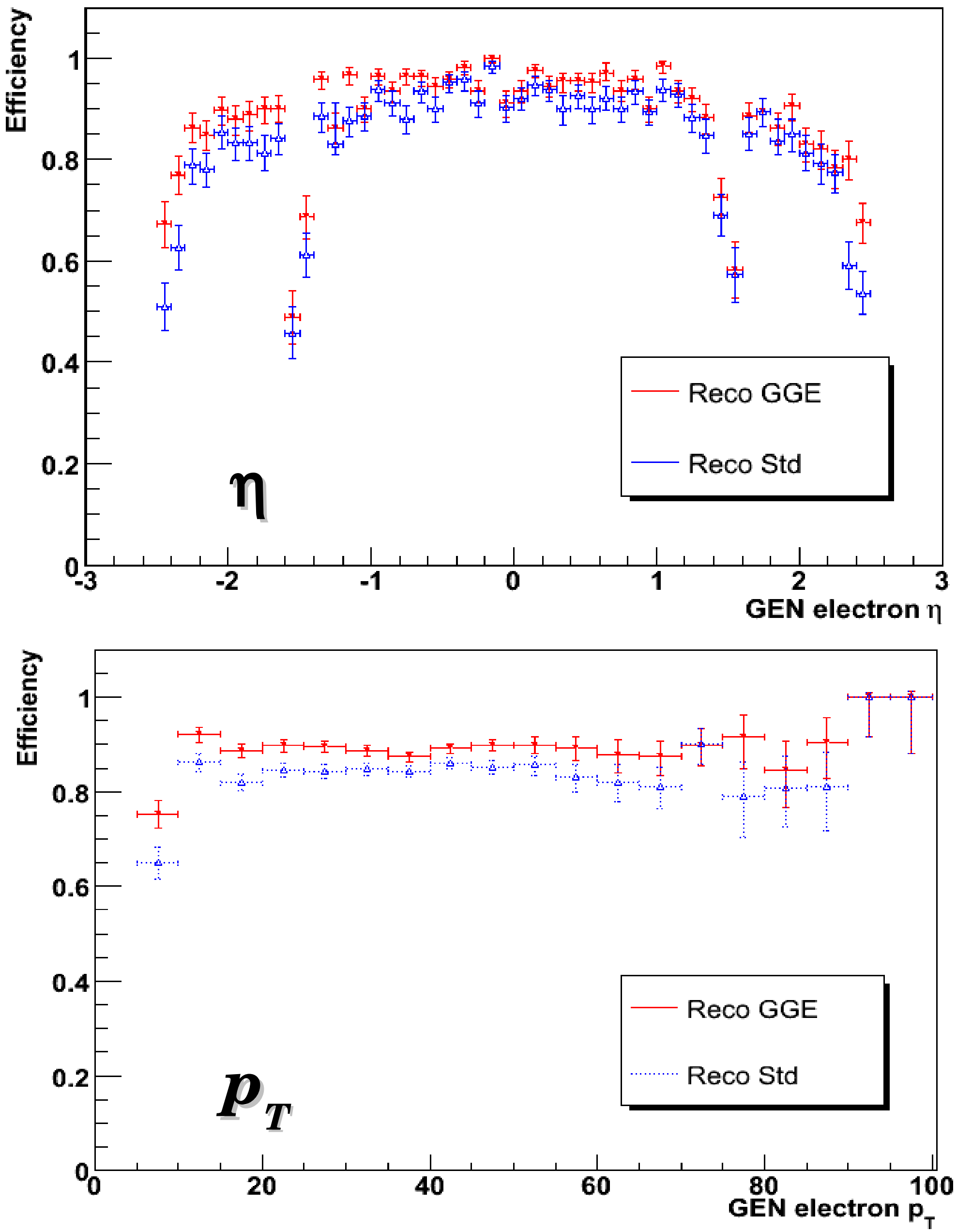
$\Delta\eta$

Efficiency definition

➔ The MC-RECO matching criteria for efficiency calculations:

- a reconstructed electron matches a MC electron if $\Delta R < \Delta R^{max}$ and has the same charge
  - $\Delta R^{max} = 0.05$
- if more than one reconstructed electrons satisfies these conditions, the one with closest p relative to the MC electron is selected.
- only prompt electrons are considered

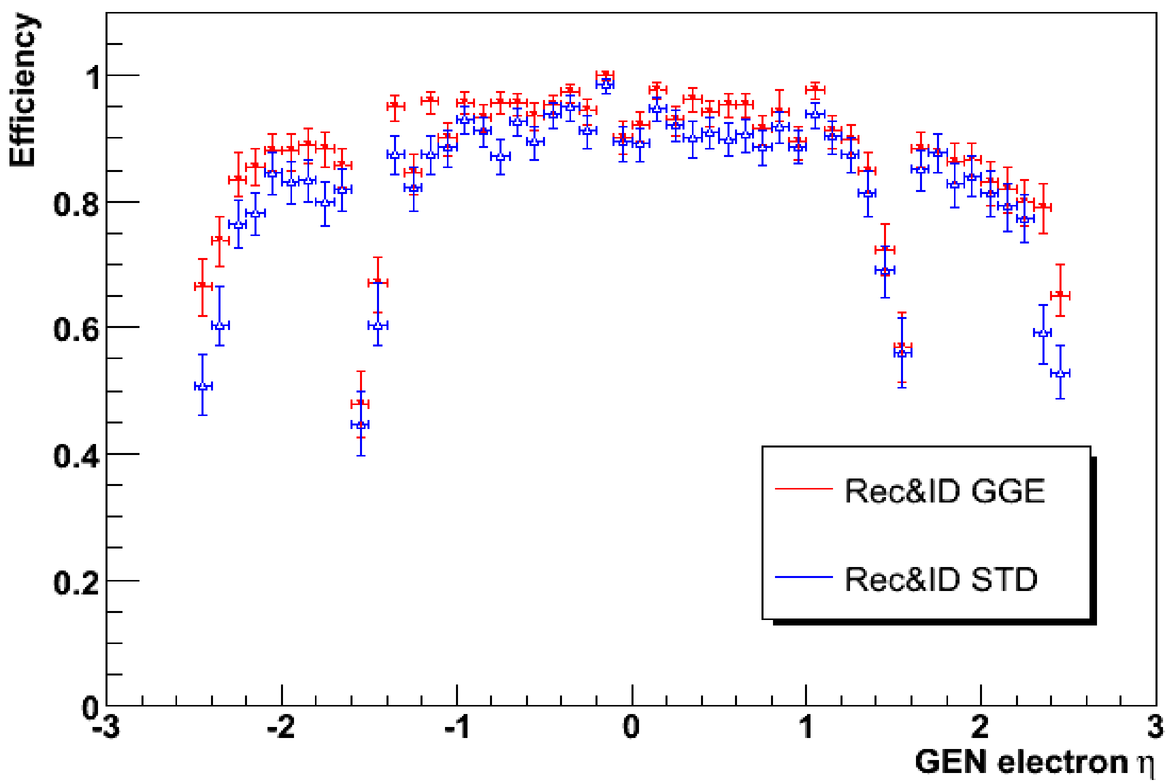
Performance - Efficiency

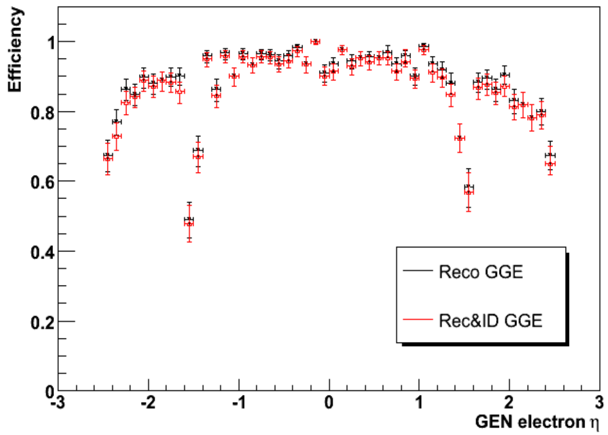


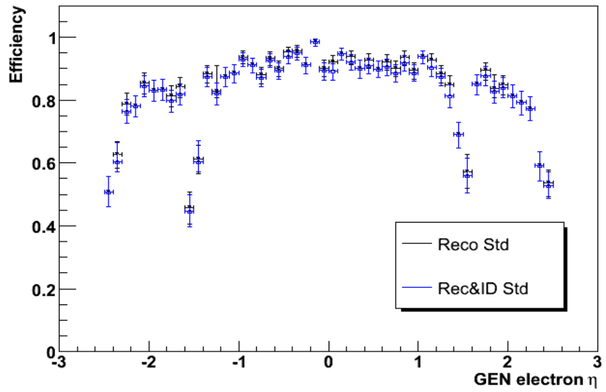
- **GlobalGSFElectrons** is more efficient over full η range
  - GGE = 90%
  - STD = 85%
- Source of improved efficiency:
  - better seeding of electron tracks (use the standard CTF track seeder)
  - removal of some producer level cuts
  - more efficient seed to cluster match

Electron ID (robust)

➔ To be sure that the increase in efficiency is not due to the reconstruction of fakes we compare the performance after the application of robust and loose eID.

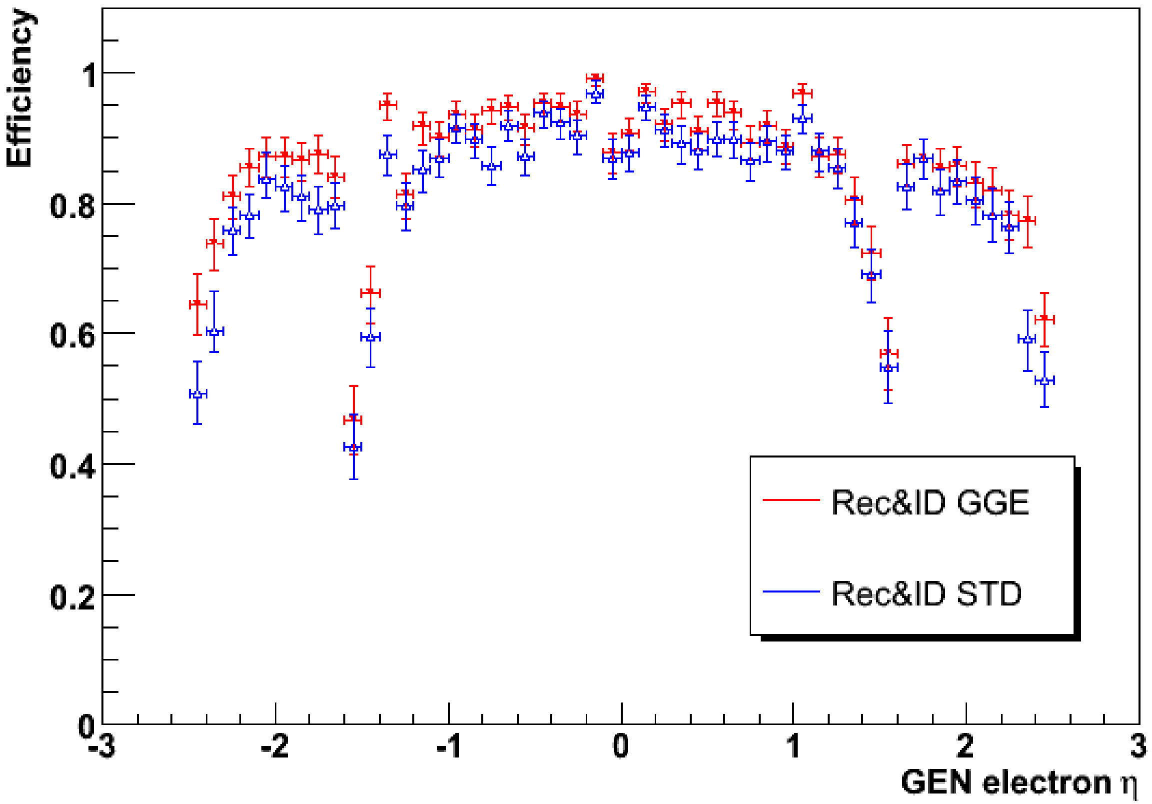


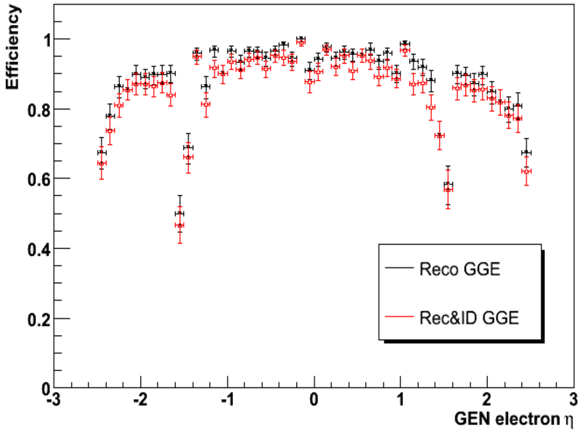


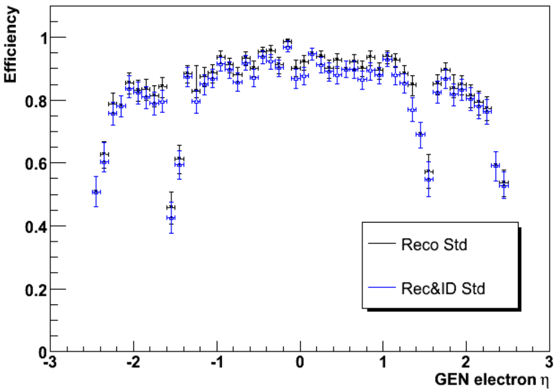


Electron ID (loose)

➔ The efficiency gain is almost the same so we are not reconstructing junk.



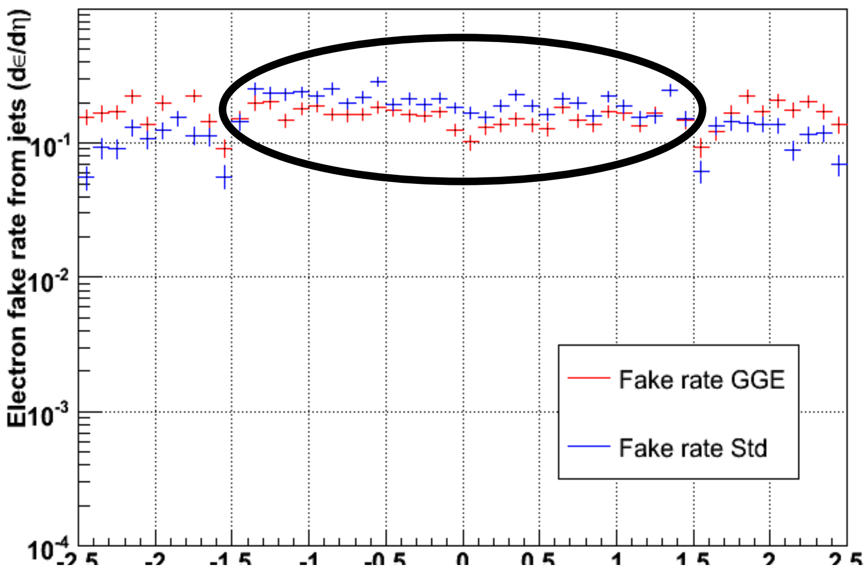


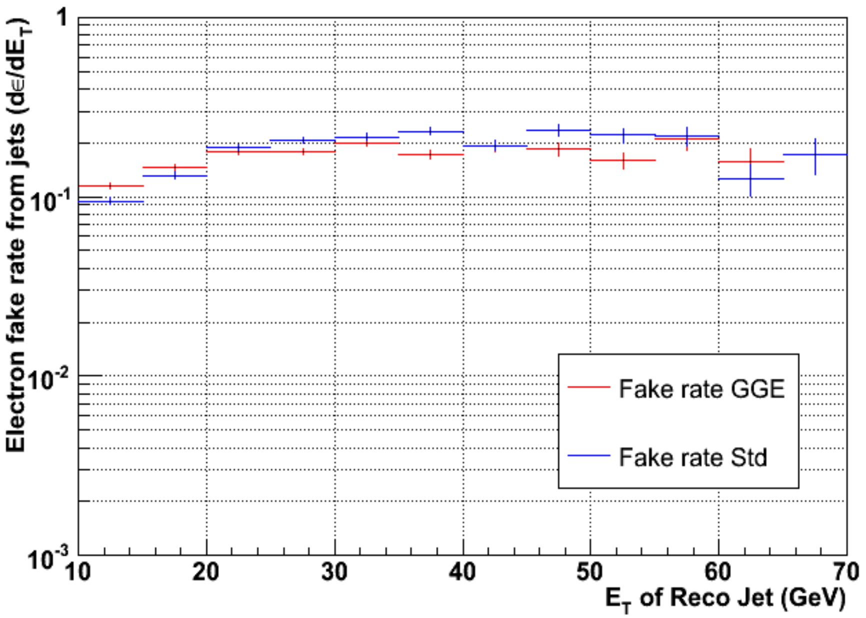


Fake definition

- ➔ A QCD sample (pt_50_80) has been used to evaluate the fake-rate.
- ➔ The fake rate is defined as the number of reconstructed jets that match a reconstructed electron divided by the number of reconstructed jets:
  - jets reconstructed with Iterative Cone (size 0.5) algorithm and $E_T > 10$ GeV
  - jet corrections not considered

Performance - Fake Rate (no eID)





--- STD fake-rate: 15 %

--- GGE fake-rate: 15 %

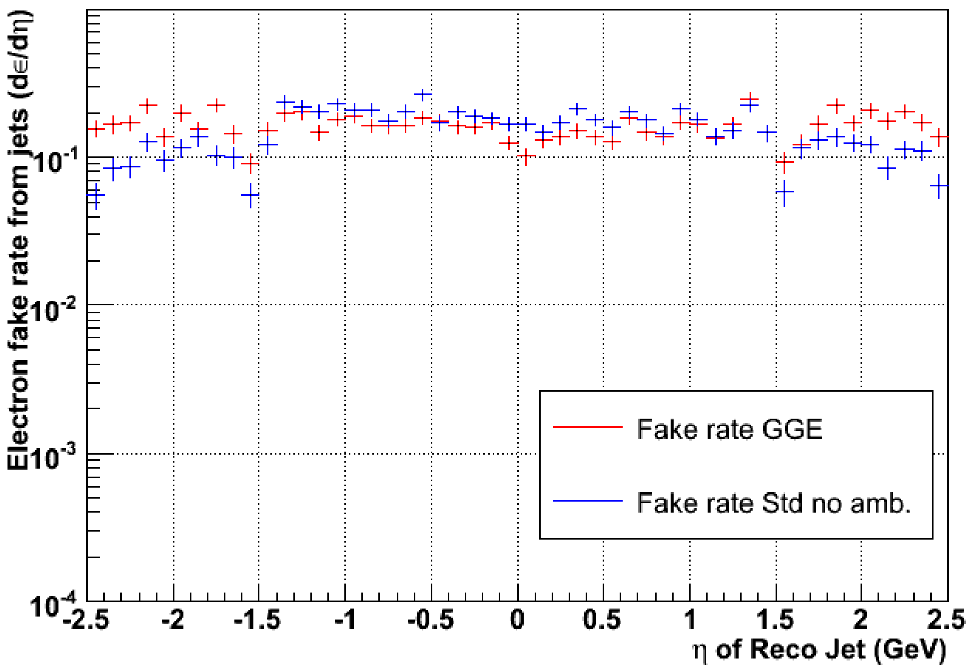
The average fake-rate is the same for both algorithm.

In the endcap GGE fake-rate is higher as the reconstruction efficiency is higher.

That's not true in the barrel.

Fake Rate in Barrel

- ➔ A fraction of the higher fake-rate for STD reconstruction is due to the possibility to have multiple candidate per SC.
- ➔ Resolving the ambiguity by using E/p of the candidates slightly reduces the effect.



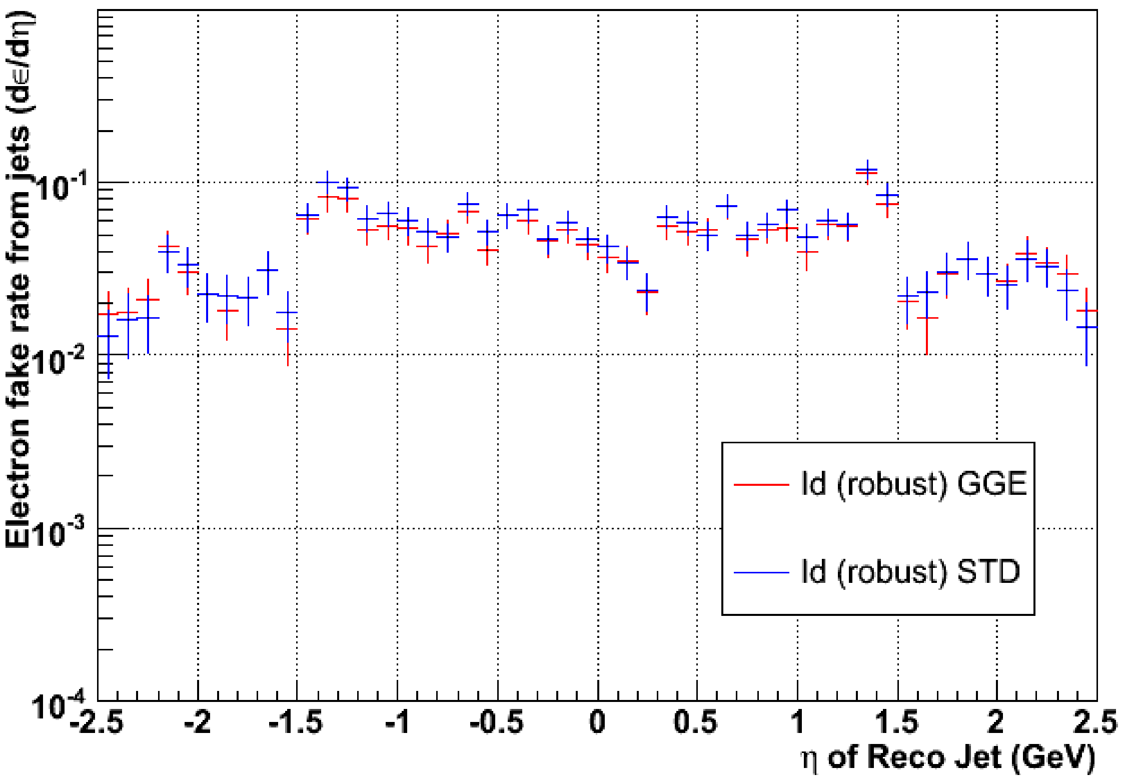
The remaining difference can be explained by the cleaner matching performed in GGE where the full reconstructed track information is used.

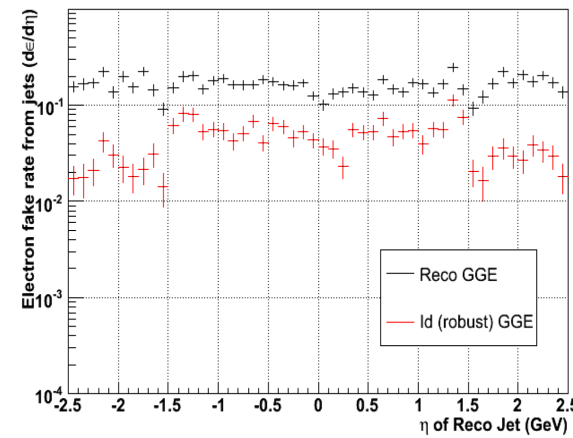
Ambiguity resolution

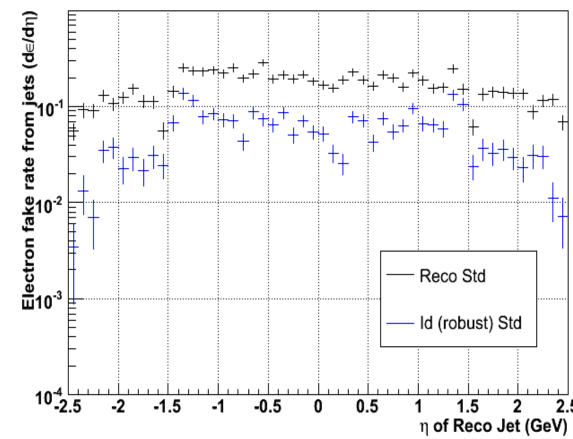
- ➔A small and simple module after electron producer should resolve ambiguity between electrons belonging to the same SC.
- ➔Our choice to select best candidate according to E/p seems reasonable.
- ➔Sent results on old study which showed distributions improvement after E/p selection, more studies needed.

From Seez report.

Performance – Fake Rate (eID)







➔ Fake-rate is still under control after applying electron Identification.

Conclusions

➔ We have set up a new sequence to reconstruct electrons (**`GlobalGsfElectrons`**):

- GlobalMixedSeed
- standard GSF track building and fitting

➔ We have loosened the electron producer cuts and modified the matching between track/seed and SC.

➔ The electron producer has been modified so that it is possible to switch between the two sequence.

➔ We see an increase in efficiency both in the endcap and barrel region.

Algorithm comparison

- General agreement that eff and fake rejection are similar (provided threshold tuning).
- Mixed seeding approach has advantage to use standard tracker code/tools.
- This advantage is incompletely realized if it can't be used in HLT: if special pixel seeding needs to be mantained for HLT then it is arguable that it can/should also used off-line.

From Seez report.