5 different MLP networks
In 5 different regions

Eta Region	Training region	Variable set
$\|\eta\|<0.8$	$\|\eta\|<0.8$	28671
$0.8<\|\eta\|<1.4$	$0.8<\|\eta\|<1.4$	28671
$1.4<\|\eta\|<1.5$	$1.2<\|\eta\|<1.7$ (B) $1.3<\|\eta\|<1.6$ (S)	25143
$1.5<\|\eta\|<2.0$	$1.3<\|\eta\|<2.2$ (B) $1.5<\|\eta\|<2.0$ (S)	28671
$2.0<\|\eta\|<2.4$	$2.0<\|\eta\|<2.4$	28669

Signal training samples:
z->ee,W->ev/+jets, ttbar

Background samples:
Jn.root n=1,..,8

Loosen precuts to have enough statistics to train

variable set 28671

AbsTrack_El_d0
EM37_El_DeltaEta
El_Ehad1OverEt
El_EoverP
El_calRatio
El_e2tsts1
El_emins1
El_etcone
El_f1
El_fracs1
El_weta1
El_weta2
El_wtots1
n_El_TRTratio

variable set 28669

AbsTrack_El_d0
EM37_El_DeltaEta
El_Ehad1OverEt
El_EoverP
El_calRatio
El_e2tsts1
El_emins1
El_etcone
El_f1
El_fracs1
El_weta1
El_weta2
El_wtots1

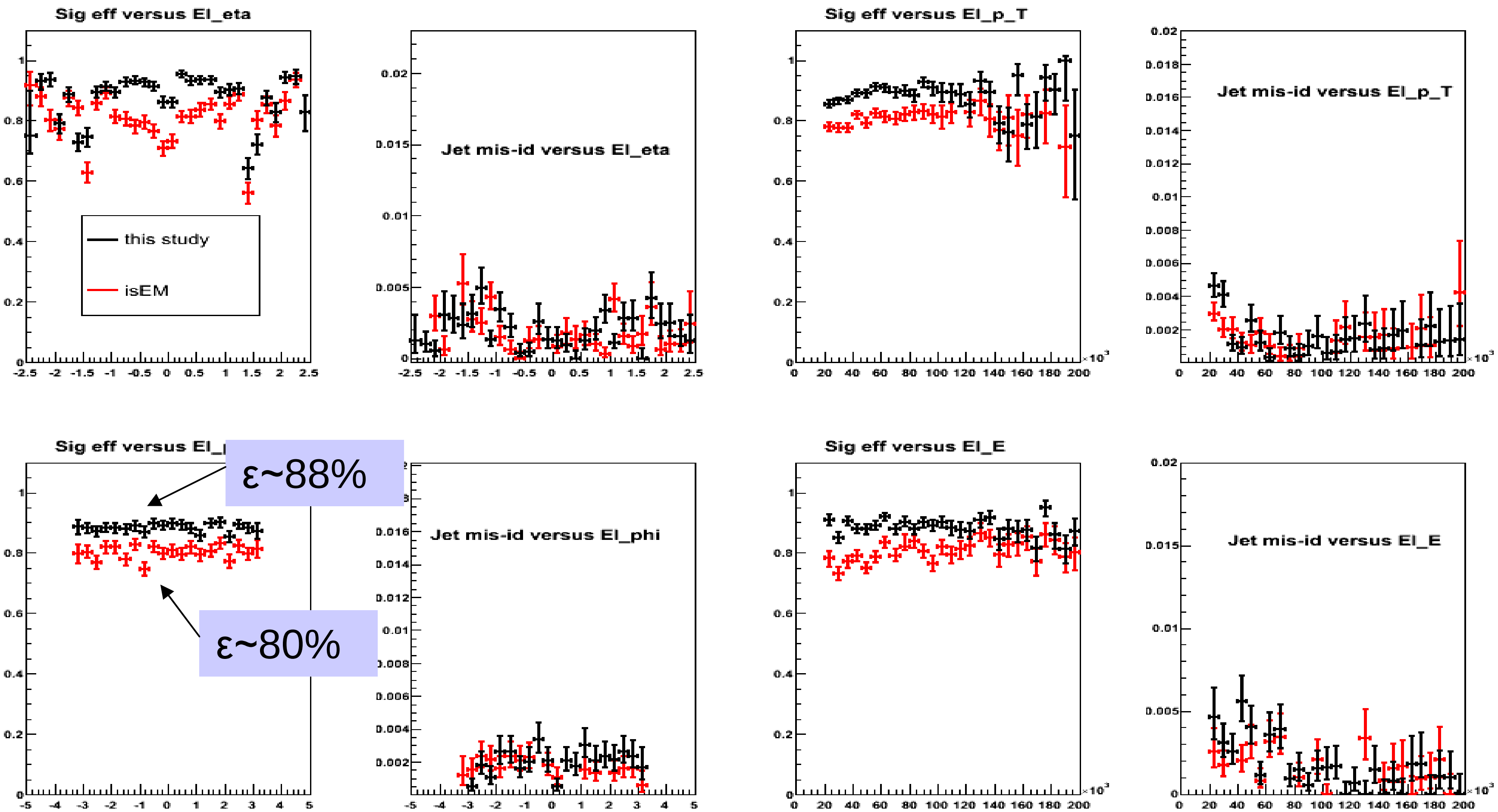
variable set 25143

AbsTrack_El_d0
EM37_El_DeltaEta
El_Ehad1OverEt
El_EoverP
El_calRatio
El_etcone
El_weta1
El_weta2
n_El_TRTratio

Precuts:

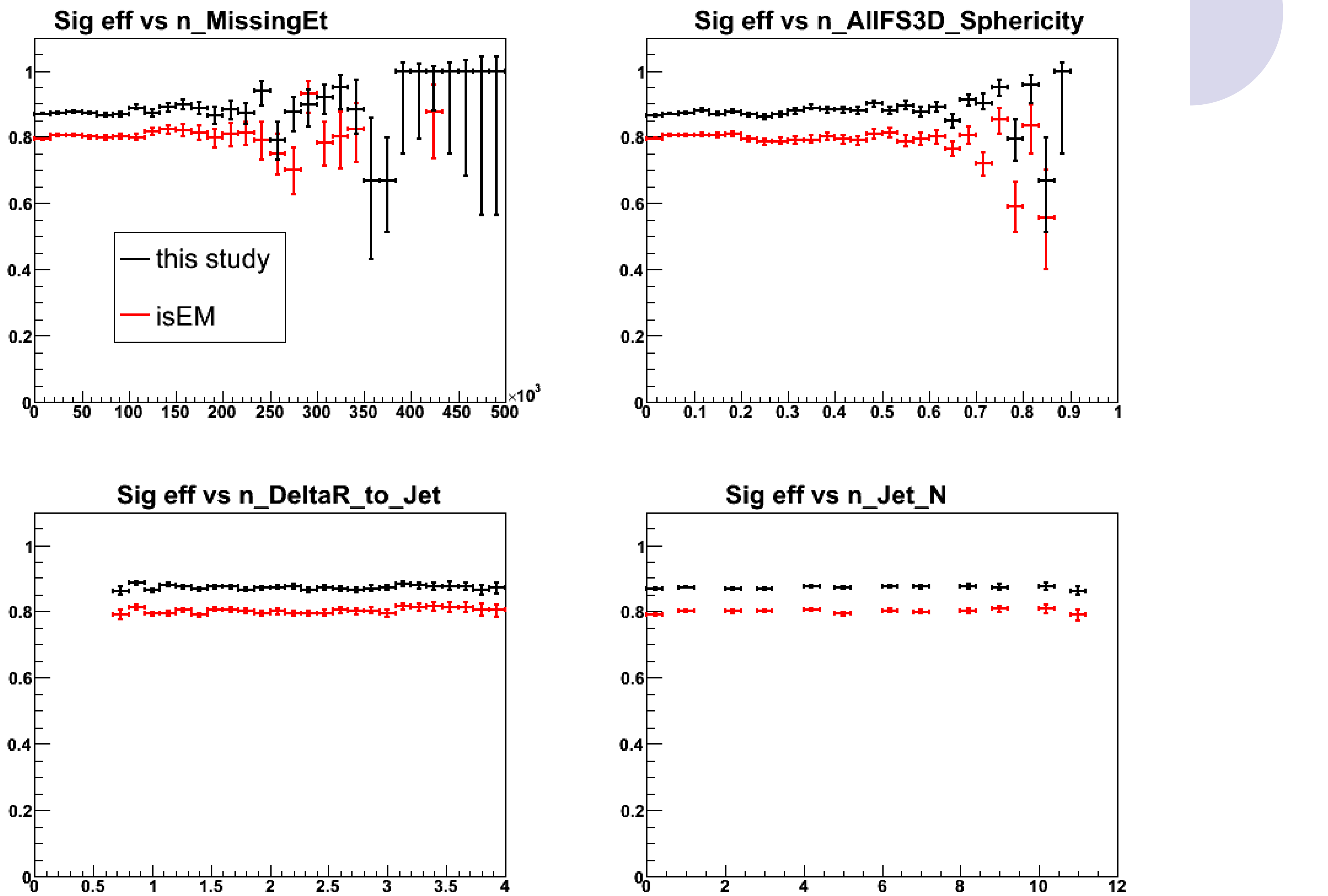
- Eta region dependent
- Require variables input to TMVA have not default values
- Additional precuts = obvious straight cuts

No overlap b/w training and estimate sample

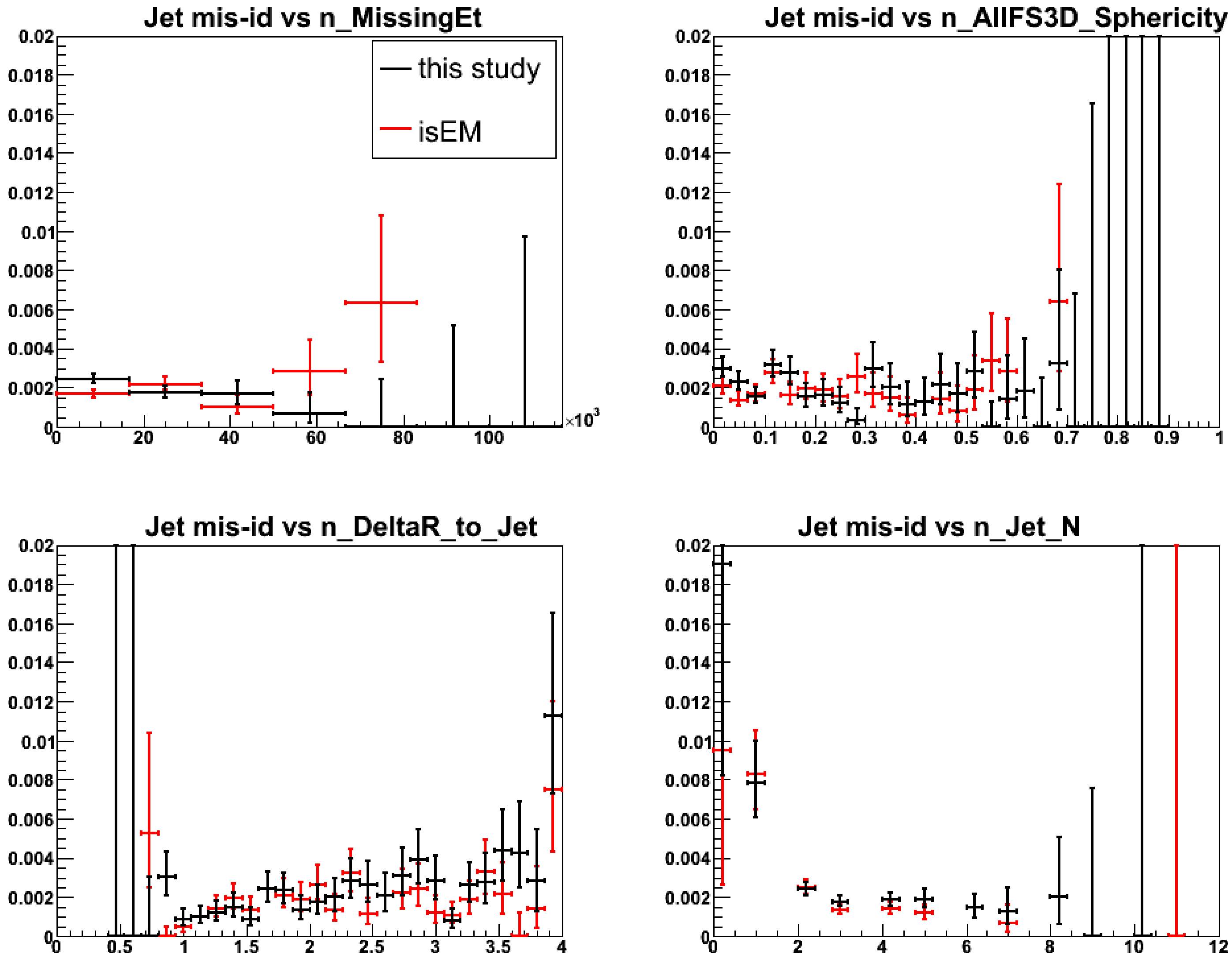



```
---> Region : BIN1   TMVA cut      = 0.583333   TMVA fake rate = 0.000993049   isEM fake rate = 0.00102408
---> Region : BIN2   TMVA cut      = 0.363333   TMVA fake rate = 0.00295287    isEM fake rate = 0.00306644
---> Region : BIN3   TMVA cut      = 0.81       TMVA fake rate = 0.00227273  ▯ isEM fake rate = 0.00227273
---> Region : BIN4   TMVA cut      = 0.69       TMVA fake rate = 0.00244714    isEM fake rate = 0.00244714
---> Region : BIN5   TMVA cut      = 0.623333   TMVA fake rate = 0.00133649    isEM fake rate = 0.0015421
```

Dependence of efficiency versus Njet, MET, S, ΔR(e,jet)



Dependence of fake rate versus Njet, MET, S, ΔR(e,jet)



Sample dependence (would be nice to have for next week!)

• Produce the INTEGRAL efficiency on ttbar, Z->ee, W->ev, W->ev + jets, SU2, SU3, SU8

• These INTEGRAL efficiencies are the result of the convolution of ε(pT,η,other parameter? DR(e,jet)) with the distributions of pT, eta, other parameter.

•Show that we can go from Z->ee INTEGRAL efficiency to SU2 INTEGRAL efficiency by reweighting the Z->ee events. The weights are derived from the ratios of pT, eta, (other parameter) distributions in SU2 divided by the distributions in Z->ee

•The remaining difference in INTEGRAL efficiencies gives us the systematic error on the electron efficiency.

•Needs to be done for isEm and our TMVA-based discriminant.

•I cannot work on this until saturday, because I am travelling...

-- finalize 5 eta region work. CC / **Done**

-- add track information into the ntuple (so we can use track isolation and ntracks in a narrow cone into the discriminant) JB

-- code for efficiency evaluation: automatically determine in each eta region what the TMVA cut should be determination to operate at same rejection as isEM. CC / **Done**

-- train the discriminant on SUSY samples, how does the performance depend on the sample used for training. JB/CC

-- make sure that there is no overlap b/w the sample used to train discriminant and the sample used to determine the performance of the discriminant. CC / **Done**

-- understand the strong sample dependence of isEM and our eID (high priority) JB

-- 5802 filtered jet sample: understand if this can be used to increase our statistics for training discriminant. Produce ntuples out of the AOD. JB

-- need to think seriously about how to determine fake lepton rate from data. so we need to seriously think about the distinction between the electron and isolation electron, here we are concerned really about finding electrons from W/Z/sparticle decays. The isolated electrons should be a subset of our general electron definition...!

-- how hard do we need to cut on TMVA for l+jet (JB) / ll (CC) analysis - when all this is settled put into AOD?

Definition of efficiency and fake rate

- Compare performance with official isEM
- Find TMVA cut that gives in each eta region the same rejection as isEM and compare resulting efficiency.

$$\varepsilon(\text{isEM}) = \frac{N(\text{signal after precutA \& isEM})}{N(\text{signal precutA})}$$

$$f(\text{isEM}) = \frac{N(\text{bkg after precutA \& isEM})}{N(\text{bkg after precutA})}$$

$$\varepsilon(\text{TMVA}) = \frac{N(\text{signal after precutA \& precut \& TMVA>X})}{N(\text{signal precutA})}$$

$$f(\text{TMVA}) = \frac{N(\text{bkg after precutA \& precut \& TMVA>X})}{N(\text{bkg after precutA})}$$

precutA =
1. El_author==1
2. isEM_ClusterEtaRange>-98
3. El_p_T>25GeV

Adjust X cut so that f(TMVA) = f(isEM)

Definition of variables

1. El_author Algorithm used to generate electron candidate (=1 for E/gamma group)
2. El_ethad Et in the HCAL behind EM cluster
3. El_ethad1 Et in the 1st sampling of HCAL behind EM cluster
4. El_etcone Et in a DR=0.45 cone around shower (shower energy not included)
5. El_etcone20 Et in a DR=0.20 cone " "
6. El_etcone30 Et in a DR=0.30 cone " "
7. El_etcone40 Et in a DR=0.40 cone " "
8. El_emins1 E of strip with min E
9. El_emaxs1 E of strip with max E
10. El_wtots1 Total width in 20 strips
11. El_f1 fraction of energy in the 1st sampling
12. El_f1core e131/(e033+e1153+e335) so in 1st sampling
13. El_f3core e333/(e033+e1153+e335) so in 3rd sampling
14. El_pos7diff. b/w shower cell and predicted track in +/- 7 cells
15. El_iso ratio of energy in 3x3/3x7
16. El_weta1 corrected lateral width with 3 strips
17. El_weta2 corrected lateral width in sample 2
18. El_widths2 uncorrected width in sample 2
19. El_e2ts1 energy in group of 3 adjacent strips, this 3 strip cluster must be the 2nd most energetic one
20. El_e2tsts1 energy in 2nd most energetic strip
21. El_fracs1 fraction of energy outside core in S1
22. El_widths1 width with 5 strips
23. El_NTRTHits number of TRT hits
24. El_NHighThresTRTHits number of TRT hits above high threshold
25. Track_El_eta, momentaX/Y/Z, p_T, phi, qOverP : fitted track parameters
26. Track_El_d0 distance of closest approach (xy), wrt to PV or (0,0,0)? **CHECK**
27. Track_El_z0 distance to the PV (z), wrt to PV or (0,0,0) ? **CHECK**
28. Track_El_ij Track error matrix