



# Some preliminary studies on improving electron selection for inclusive W/Z analysis

**Kalanand Mishra**

*on behalf of Z Signal Extraction team*

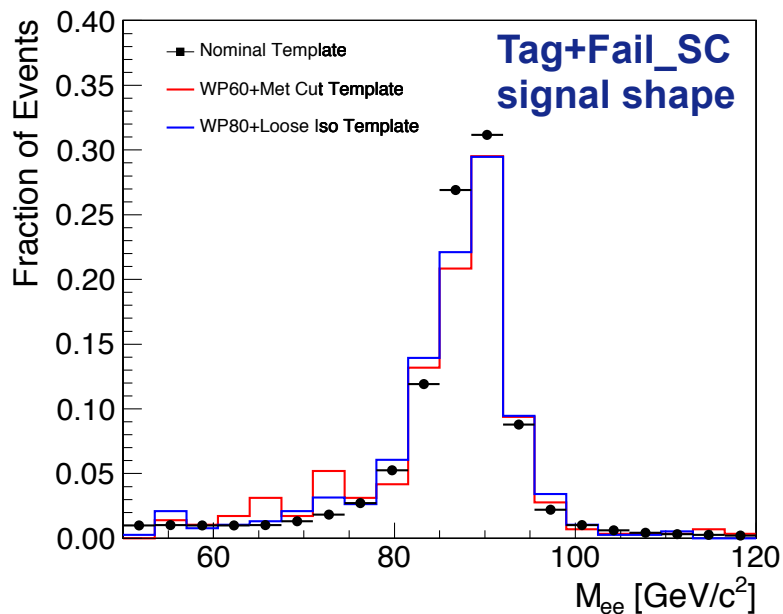
*VBTF meeting  
(December 3, 2010)*

# Motivation



◆ Uncertainty in EWK electron channel measurements is dominated by systematic uncertainty in electron reco/id: [all numbers are in percent]

Source	$W \rightarrow e\nu$	$Z \rightarrow e^+e^-$	$W^+ (e)$	$W^- (e)$	$W^+/W^- (e)$	$W/Z (e)$
Lepton reconstruction & identification	3.9	5.9	5.1	5.1	5.2	3.0
Momentum scale & resolution	2.0	0.6	2.2	1.8	0.4	2.0
$\cancel{E}_T$ scale & resolution	1.8	n/a	1.6	1.9	0.4	1.8
Background subtraction / modeling	1.3	0.1	1.1	1.5	0.7	1.3
PDF uncertainty for acceptance	0.8	1.1	0.9	1.5	1.7	0.9
Other theoretical uncertainties	1.3	1.3	1.3	0.9	1.3	1.0
<b>Total</b>	<b>5.1</b>	<b>6.2</b>	<b>6.1</b>	<b>6.2</b>	<b>5.7</b>	<b>4.4</b>



- ◆ The main source of systematic uncertainty comes from signal yield in “Tag+Fail” sample
  - Needed for computing reco and id efficiency
- ◆ Higher background rejection through an improvement in electron selection will help

# Methodology



- First, look into the performance of the existing electron selection points
  - VBTF working points 95, 85, 80, 70, 60
  - Egamma Cuts-In-Category (V6) working points: “very loose”, “loose”, “medium”, “tight”, “very tight”, “hyper tight”.
- Try adding more variables, like,
  - $E/p$ ,  $f_{\text{Brem}}$
  - $\Delta\eta$ ,  $\Delta\phi$  at Calo surface
  - additional shower shape quantities: R25 ( $=e_{2 \times 5 \text{Max}}/e_{5 \times 5}$ ), R15
  - although most of these are correlated with the existing variables being used
- Try various statistical signal classifiers
  - likelihood, linear discriminant (Fisher), neural network, k-nearest neighbor, boosted decision tree

It is worth pointing out that:

- lowest  $p_T$  unrescaled single electron trigger is tighter than WP95
- WP60/70 are extremely tight in rejecting jets but at the cost of big drop in efficiency. These points are not being used for anything other than cross checks.

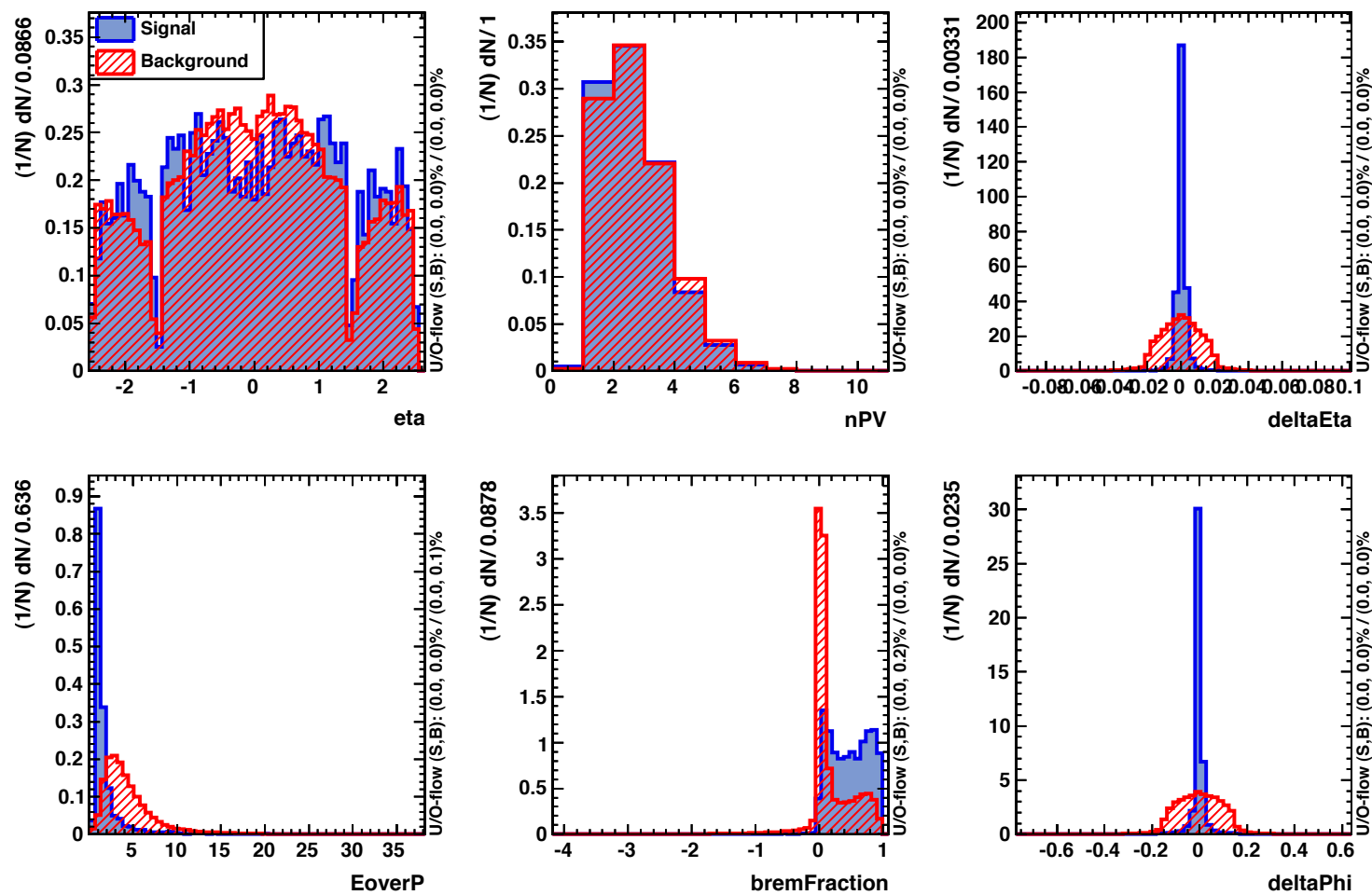


# Some variables of interest - I

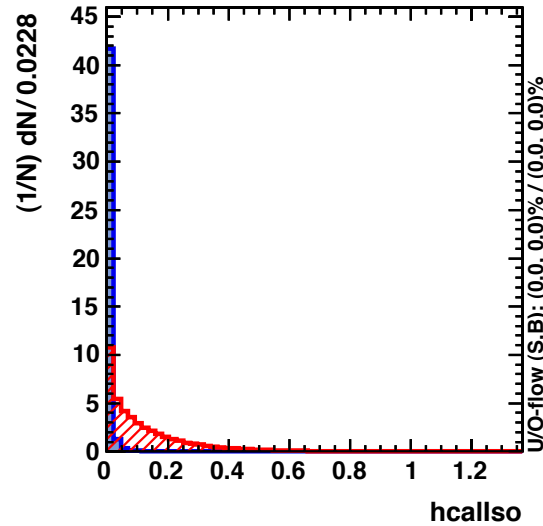
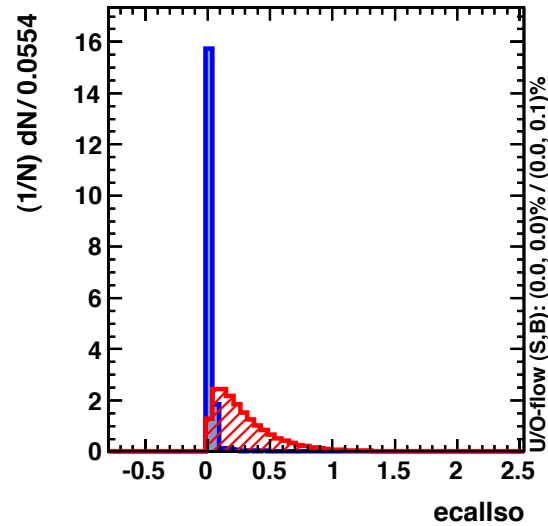
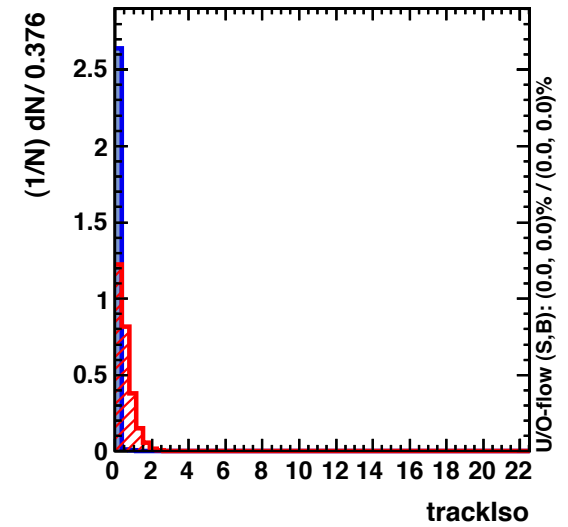
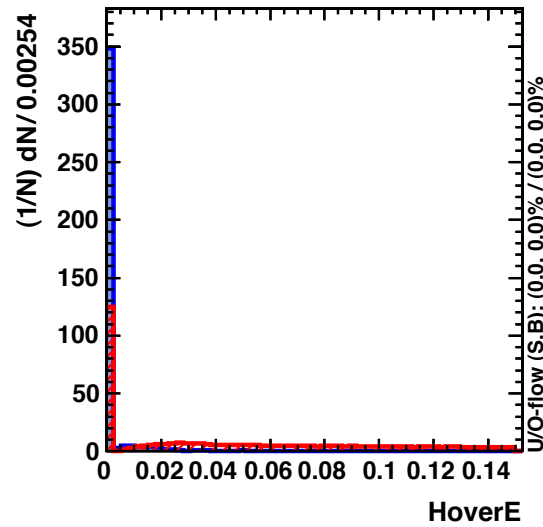
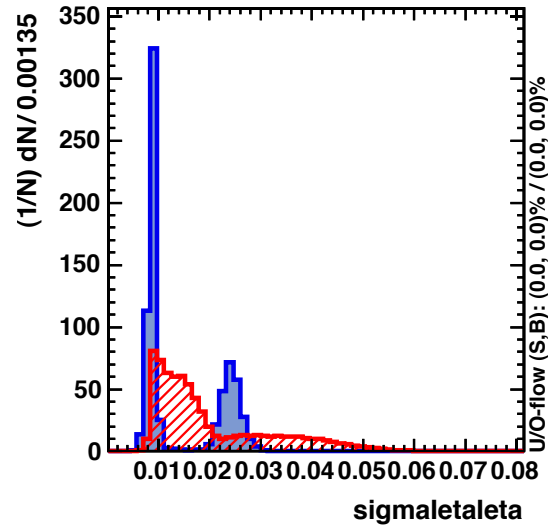
## Comparison of signal and background electrons

Signal  $\equiv$  electrons in Z events (require one leg to pass WP80, choose other leg)

Background  $\equiv$  fake electrons from Jet70 triggered events (remove trigger bias)



# Some variables of interest - II

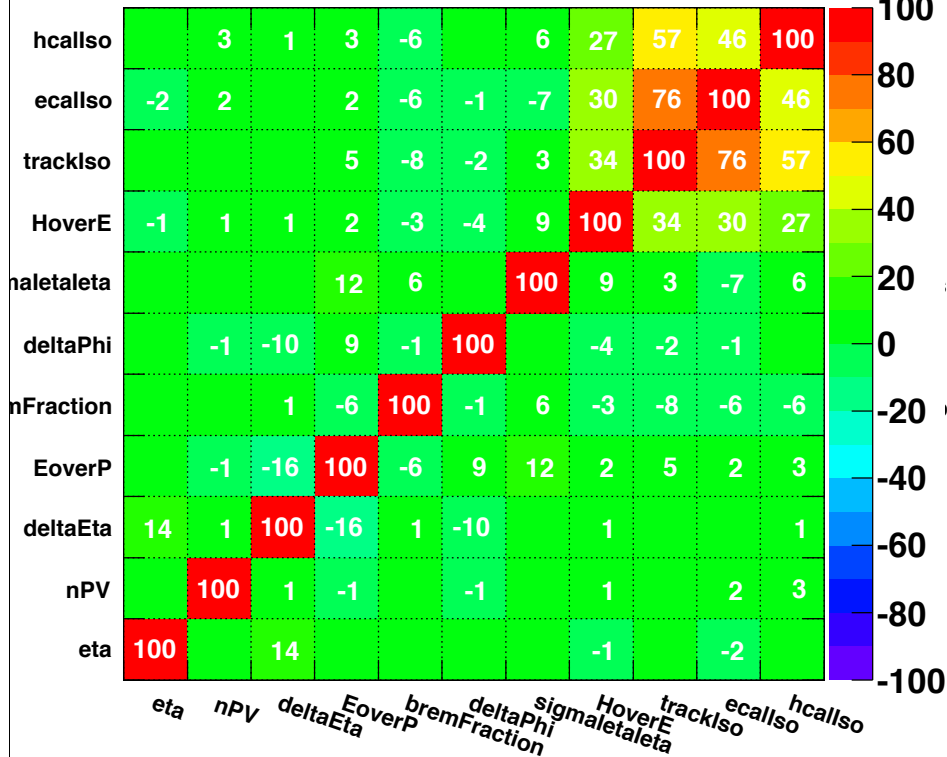


# Correlation among the variables



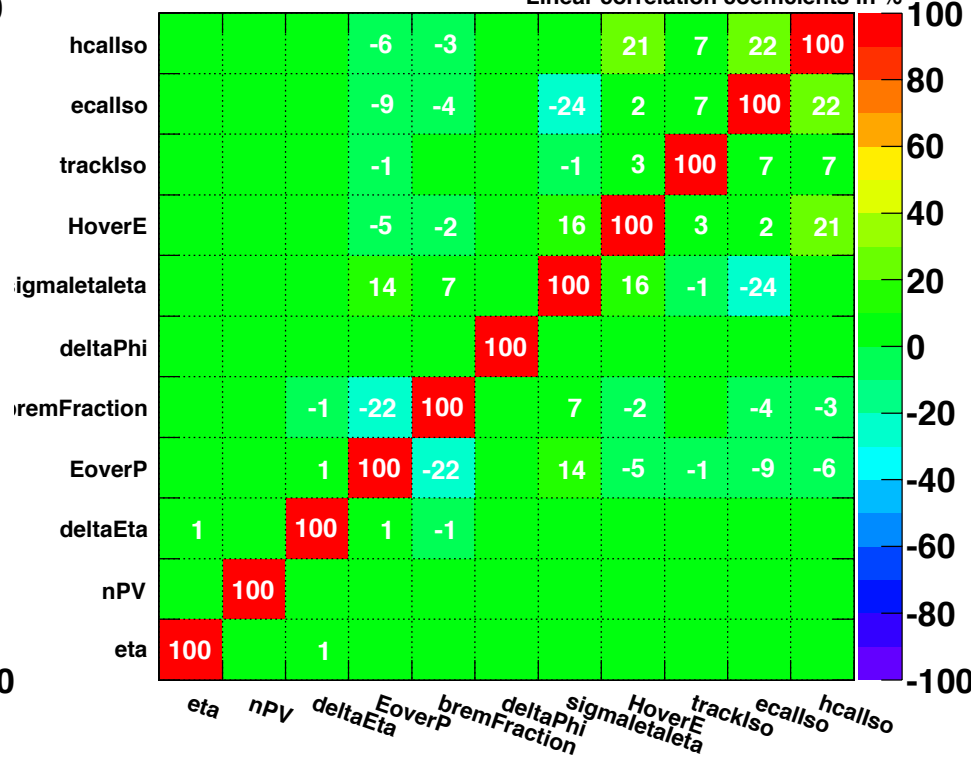
## Signal

Linear correlation coefficients in %



## Background

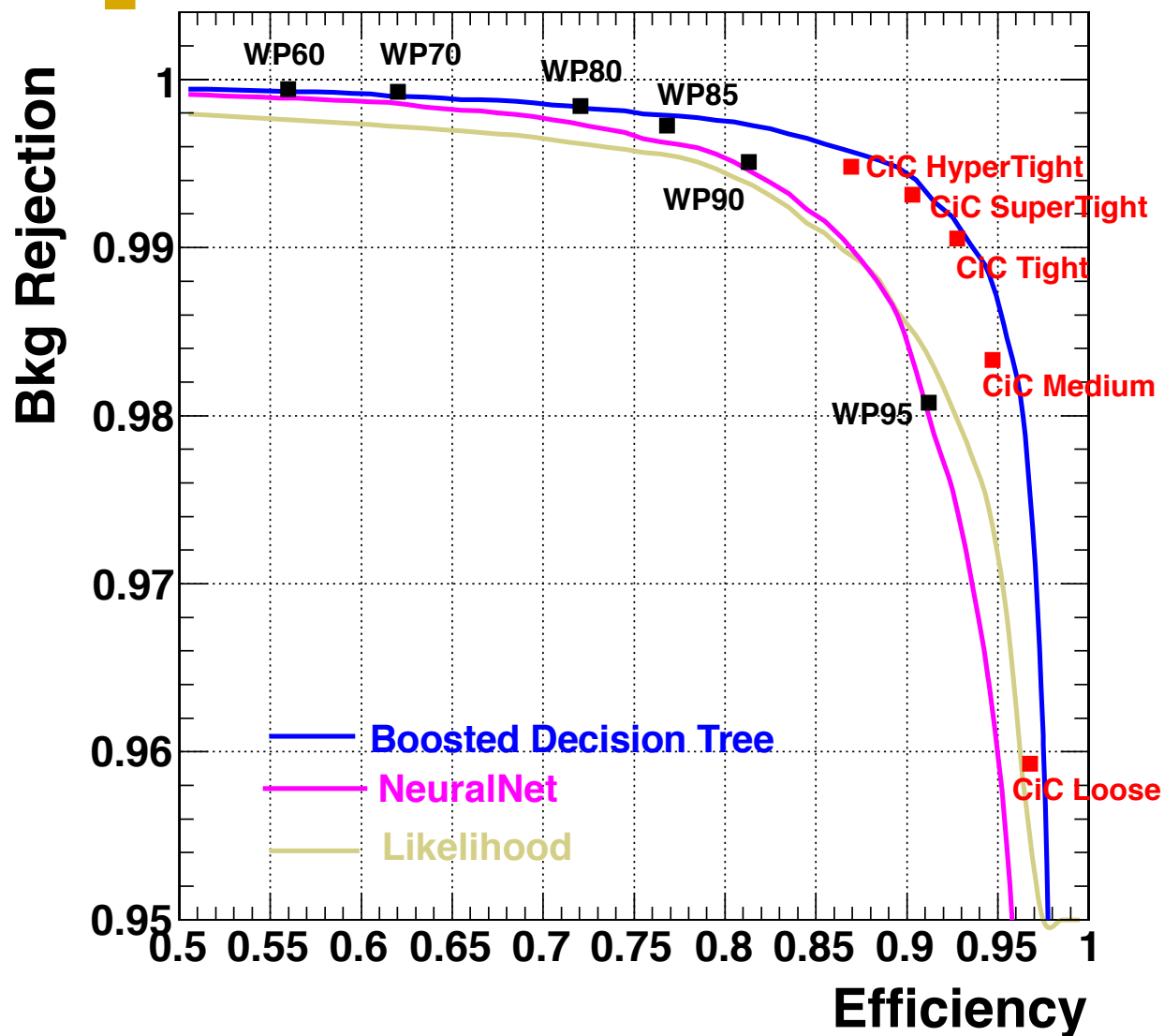
Linear correlation coefficients in %



Most variable pairs are only mildly correlated for background.



# My current findings



- The higher the curve/ point, the better performing it is

- All selections have been treated at equal footing: efficiency is computed in exact same way in all cases

- No conversion rejection cut applied in any selection

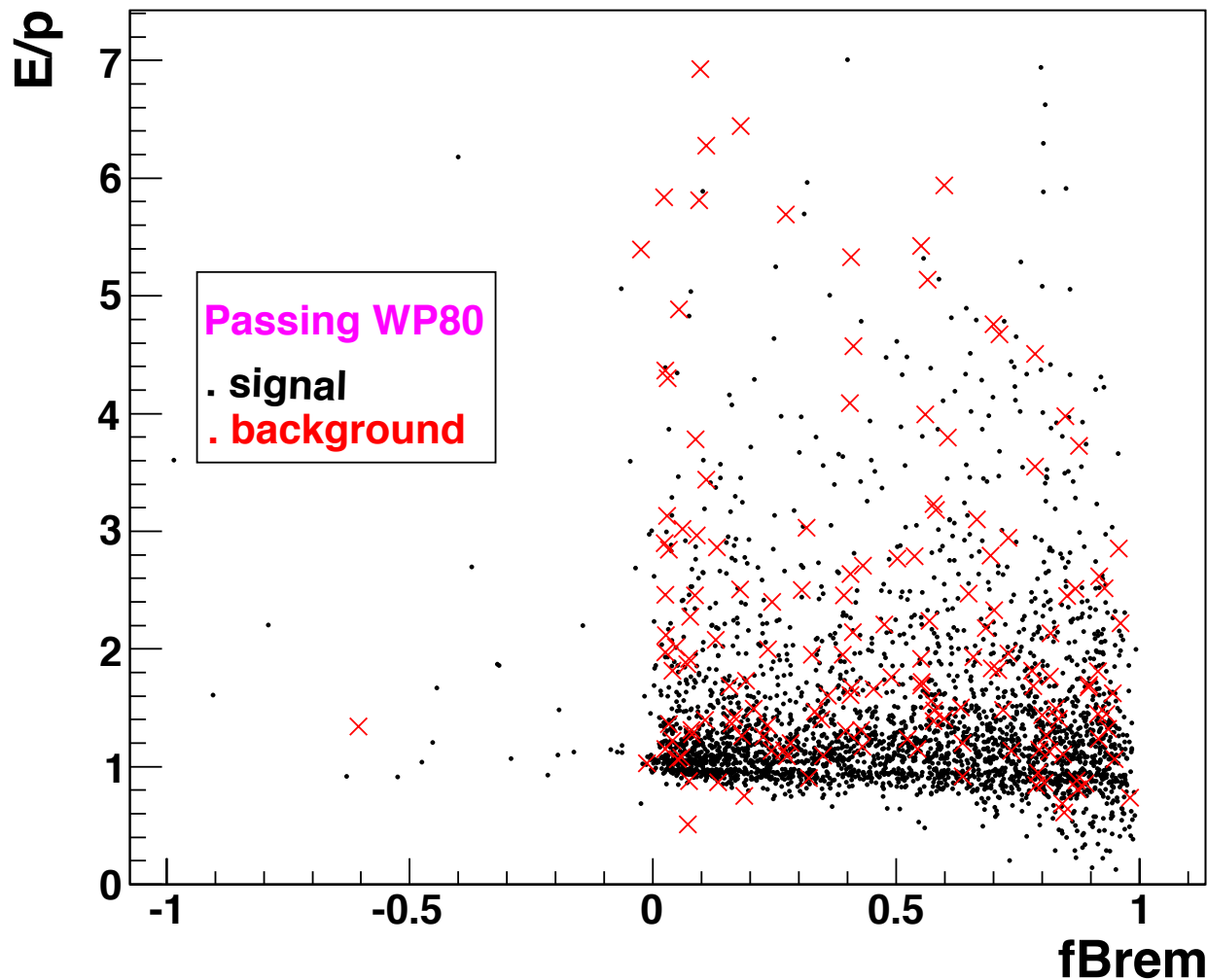
► See some observations on the next slide

## Some Observations



- ◆ One can significantly improve electron selection for electrons which are looser than WP85.
  - This might be of use for electron selection at HLT level, but of marginal interest for physics analysis.
- ◆ I can only marginally improve (by any further tuning of BDT) over WP80.
  - Essentially, at this point all further rejection power comes from H/E and  $\sigma_{\eta\eta}$  which WP80 is already cutting very hard.
  - So, there is very little to gain on reducing background rate.
- ◆ All CiC cuts seem way too loose – they seem to have been tuned for maximum efficiency, not to aggressively reduce fake rate.
- ◆ Among the training methods I tried
  - Likelihood performs well with few uncorrelated input variables: H/E, the three isolations,  $\sigma_{\eta\eta}$ , E/p, fBrem,  $\Delta\eta$ ,  $\Delta\phi$ , nPV,  $\eta$ .
  - NeuralNet, KNN, BDT improve with inclusion of additional kitchen sink variables. I found BDT to be best performing among these.

# Why dividing in categories doesn't help over WP80

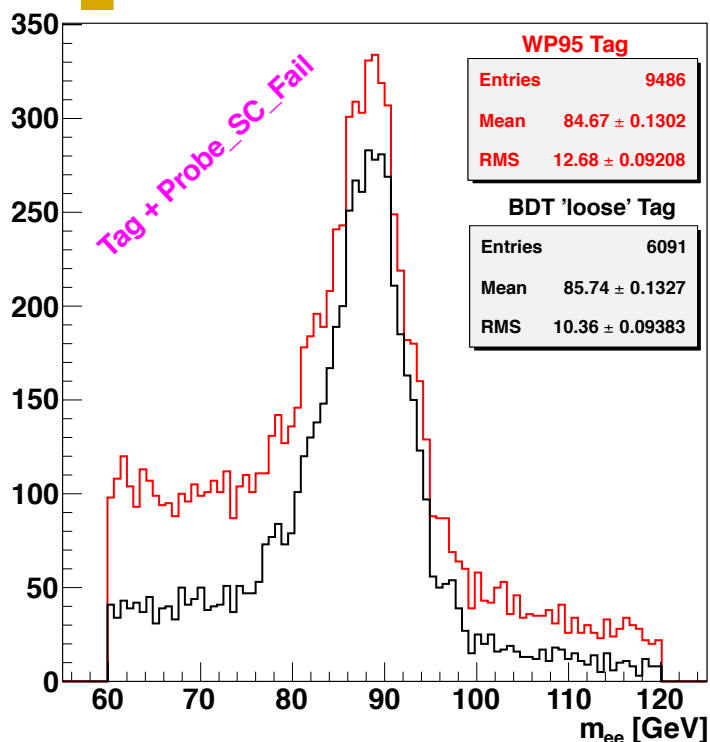


◆ There is some structure to exploit here  
-high  $E/P$  low  $f_{Brem}$  looks background enriched  
-but not nearly enough.

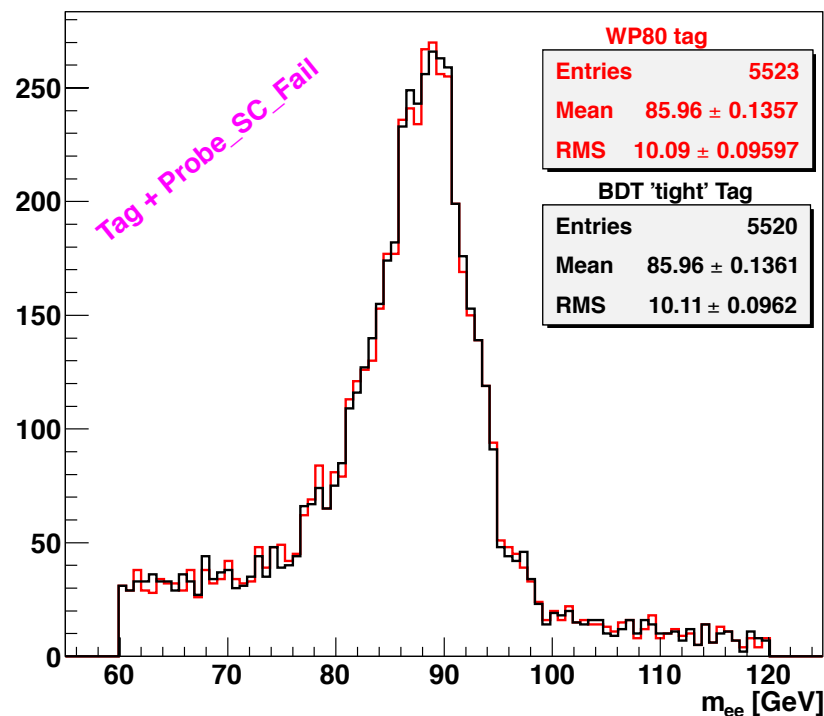
◆ Applying cuts in categories does not help much for already quite pure electrons  
-there is no more low-hanging fruit after WP80 😞



# Bottom line: Look at "Tag+Fail" mass



can improve big over WP95



Not so much over WP80

My recommendation: For now we should stick with WP80. It is NOT worth to change our selection point for few percent improvement. On the other hand, we should aggressively focus on electron selection in HLT: there is room for factors of 2–3 improvement in background rejection there.