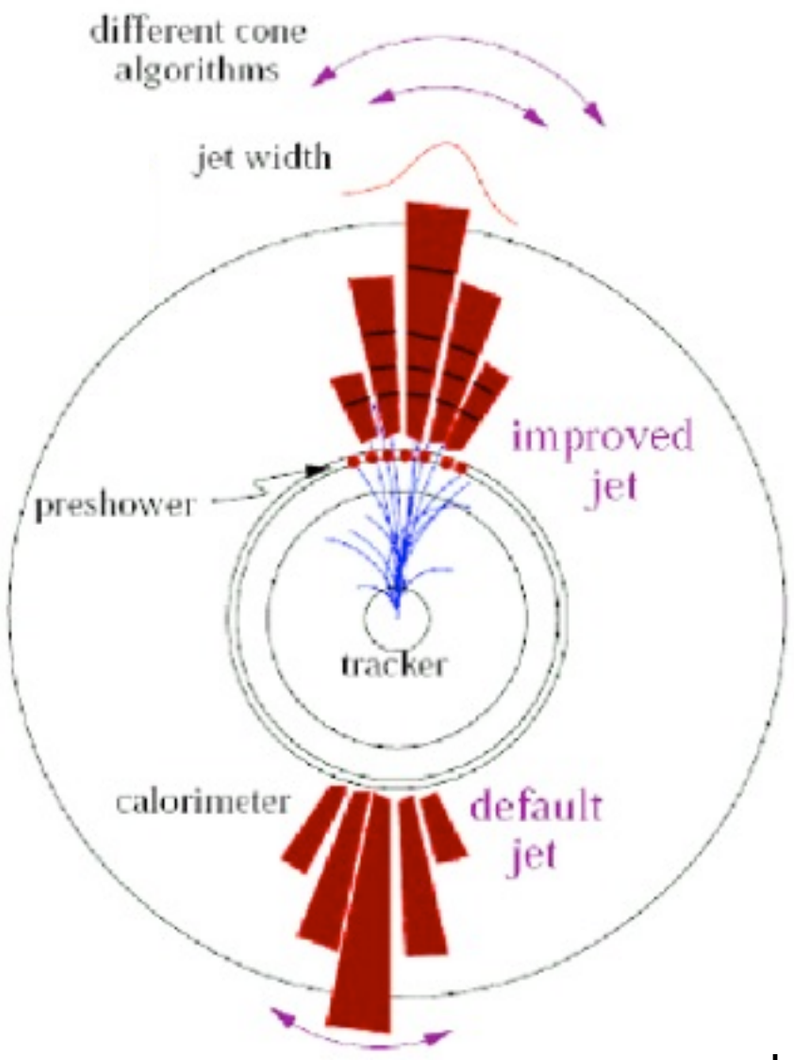
Jet Energy Resolution studies

Jonathan Brown (LPNHE)
D0 France meeting
05/04/2010

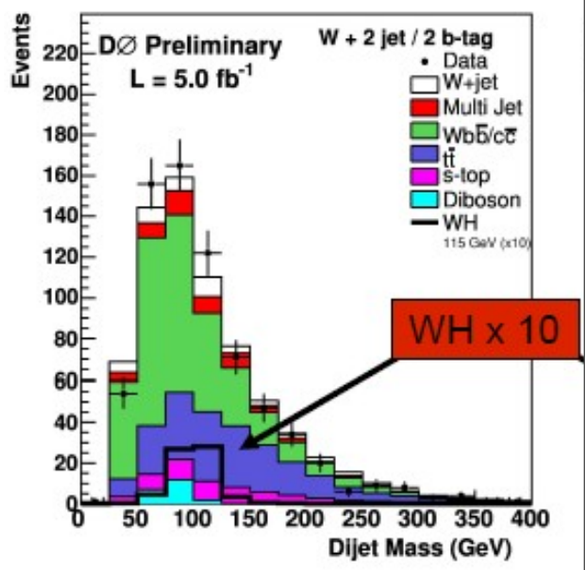


OUTLINE

- JER corrections
- Effect on Mjj
- Jet width correction

Overview

- Jet energy resolution is a crucial aspect for low mass Higgs search (H→bb),
- In theory we expect a narrow dijet resonance from the Higgs decay, but not observed at the analysis level
- We could use more information from the detector to have a more accurate jet energy measurement
- Existing set of JER corrections are now implemented in the WH framework,
- Existing corrections:
  - CPS
  - Trkcal
  - HMatrixlight
  - HMatrix
  - Semileptonic
  - Light combination
  - Heavy combination
- Investigating also an additional correction based on the jet width



Using:

- p21.13
- vjets 3.5.0
- jet_resolution_util package

❖ p20 Electron channel :
- p_T > 15 GeV
- $|\eta_{det}| < 2.5$
- Missing E_T > 20 GeV

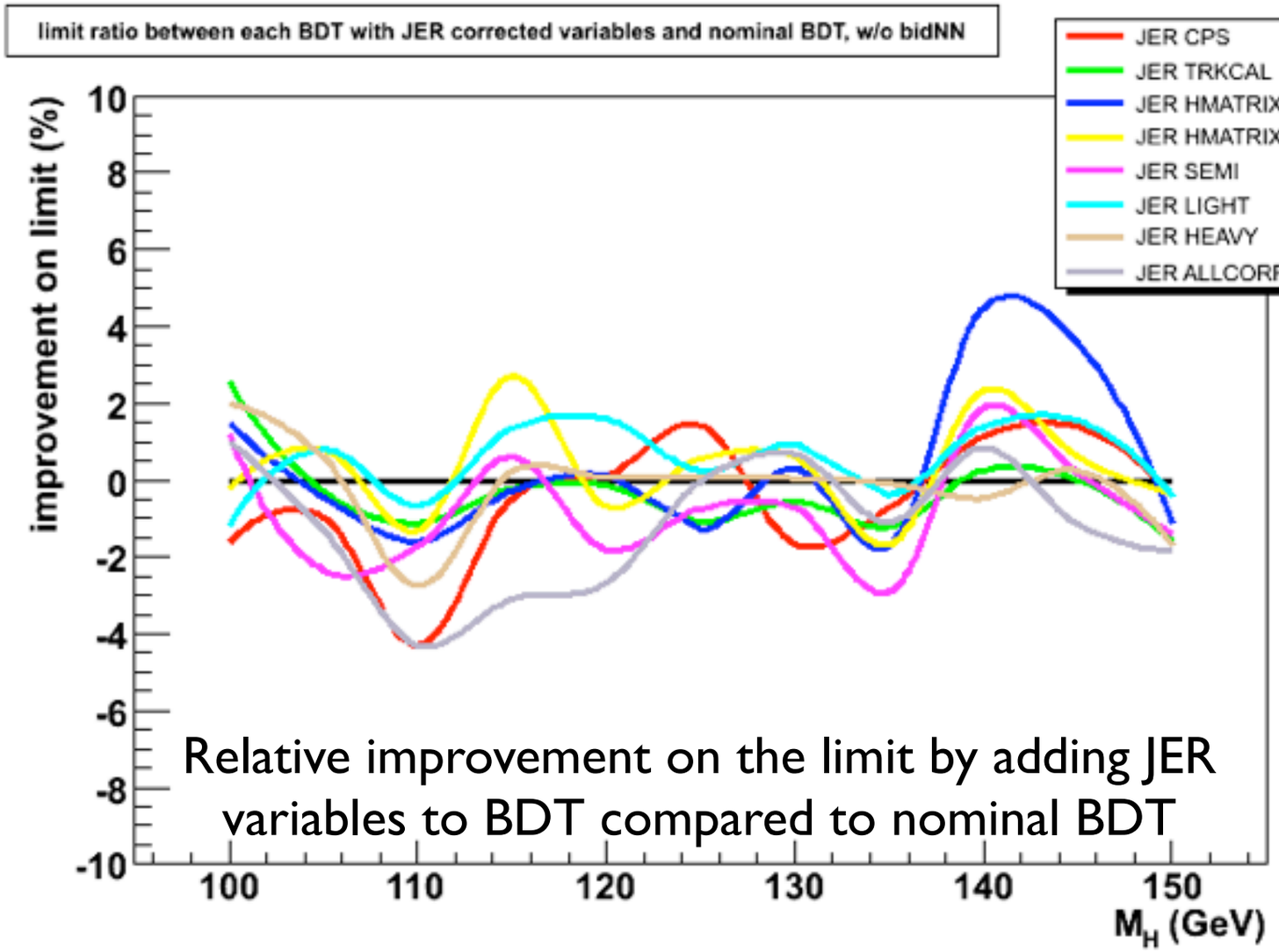
❖ Jets:
- Njets = 2
- p_T > 20 GeV
- $|\eta_{det}| < 2.5$
- H_T > 60 GeV

❖ Triangular cut

❖ $0 \leq |PVz| \leq 40$

Impact of JER variables : Limits

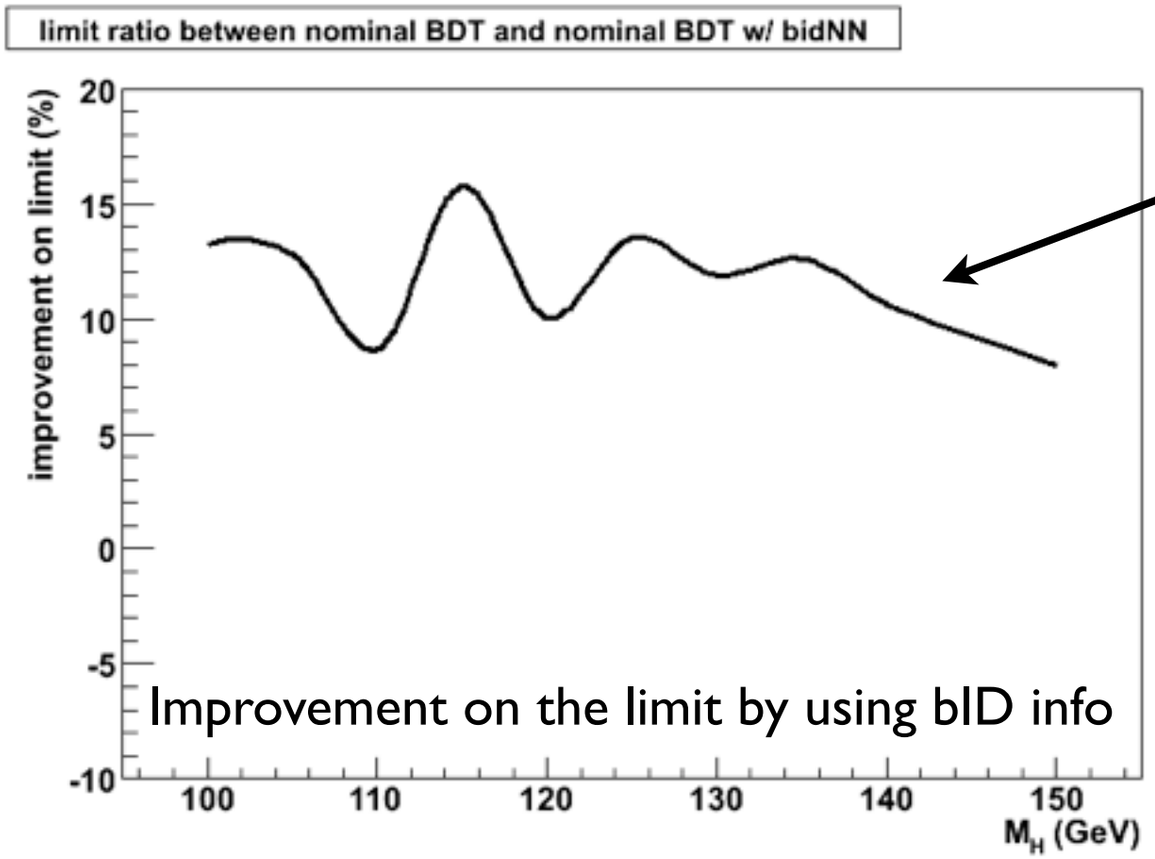
The plan is to train our final discriminant with these JER variables, hopefully improving the final limit



Relative improvement on the limit by adding JER variables to BDT compared to nominal BDT

Nominal BDT (19 variables) + set of 5 JER corrected variables (p_T^{j1}, p_T^{j2}, p_T^{jj}, H_T, Mjj) added separately for each correction (19+5 variables) or full combination (all corrections, 19+35)

Limits with bID info

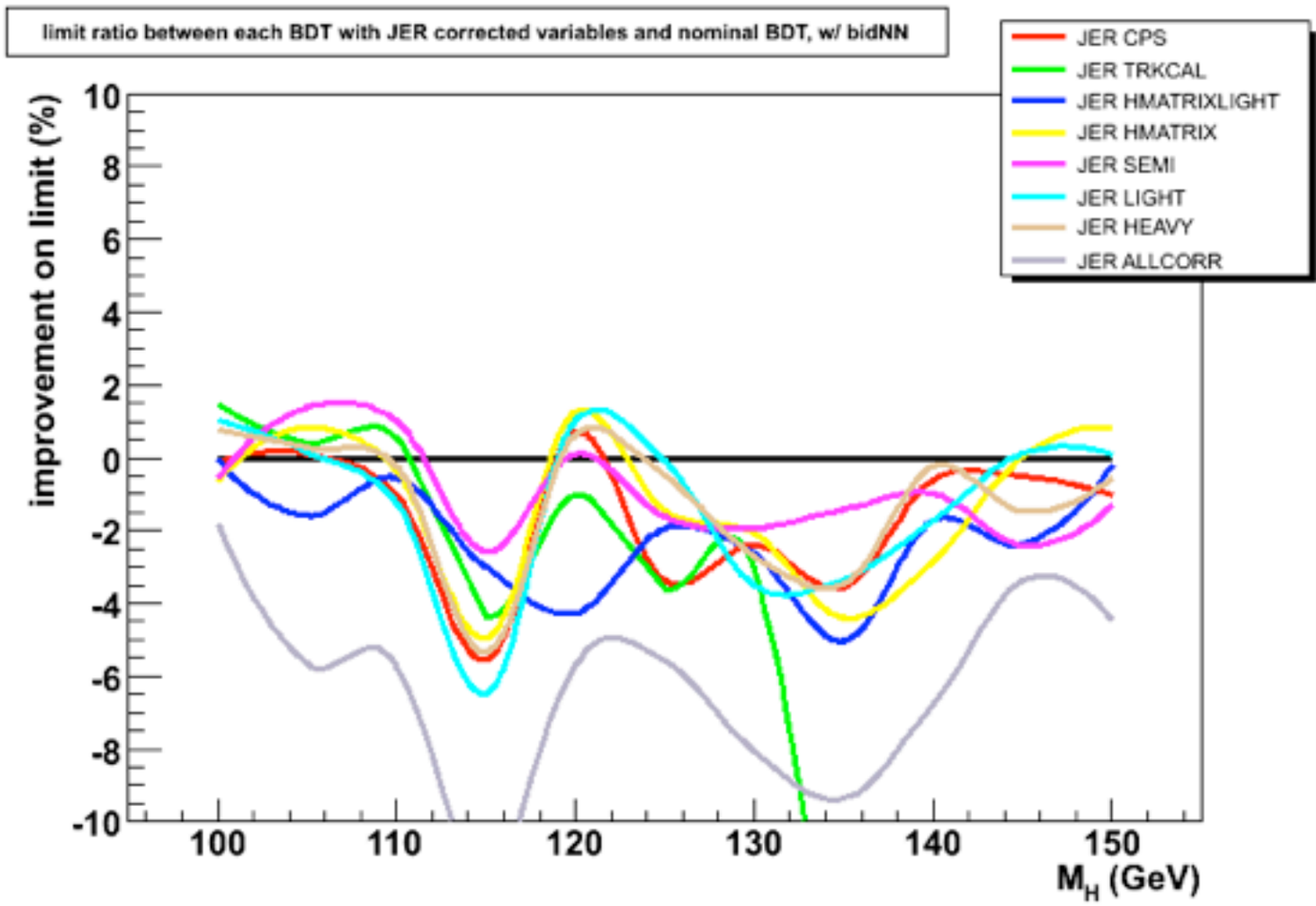


Improvement on the limit by using bID info

looks promising for continuous bID output

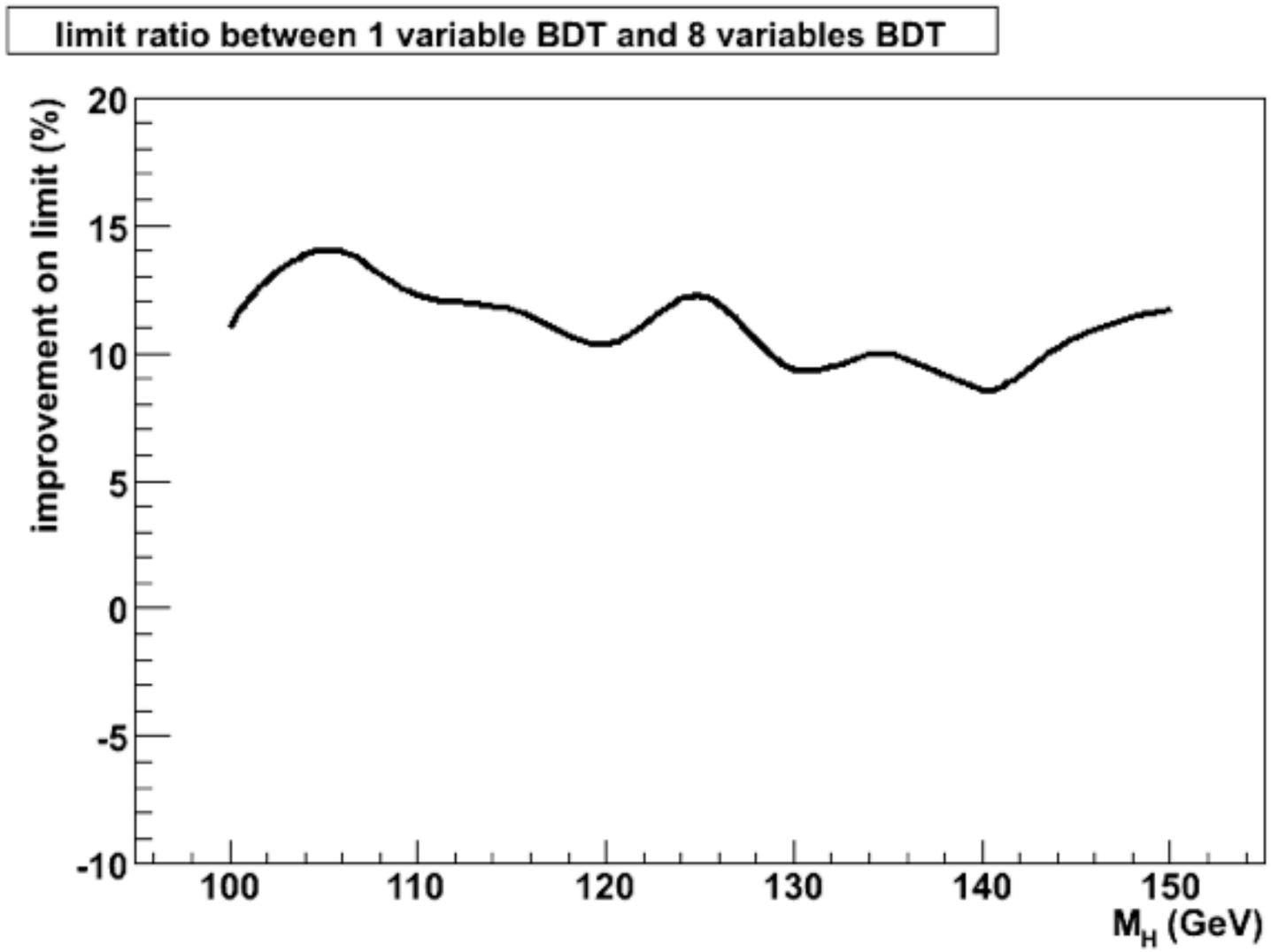
Note: the intent here is not quantify improvement from bID NN info but just to see if correlations between bID NN AND JER variables brings improvement

but doesn't help here....



Improvement on the limit by using bID info + JER variables in BDT compared to nominal + bID info BDT

Simplest BDT - ratio of limits



- BDT with only 1 variable (Mjj)
- Compared to BDT 1+7 JER Mjjs
- Is this the improvement we're looking for....???
- Why not visible in BDT 19 variables + JER variables?

Potential of JER variables

- By how much can the Mjj resolution be improved?
- Next slides show comparison between corrected Mjj and nominal for each correction for WH signal and MC backgrounds (no QCD)
- only for Trkcal, Semileptonic and Heavy correction, others are in backup slides

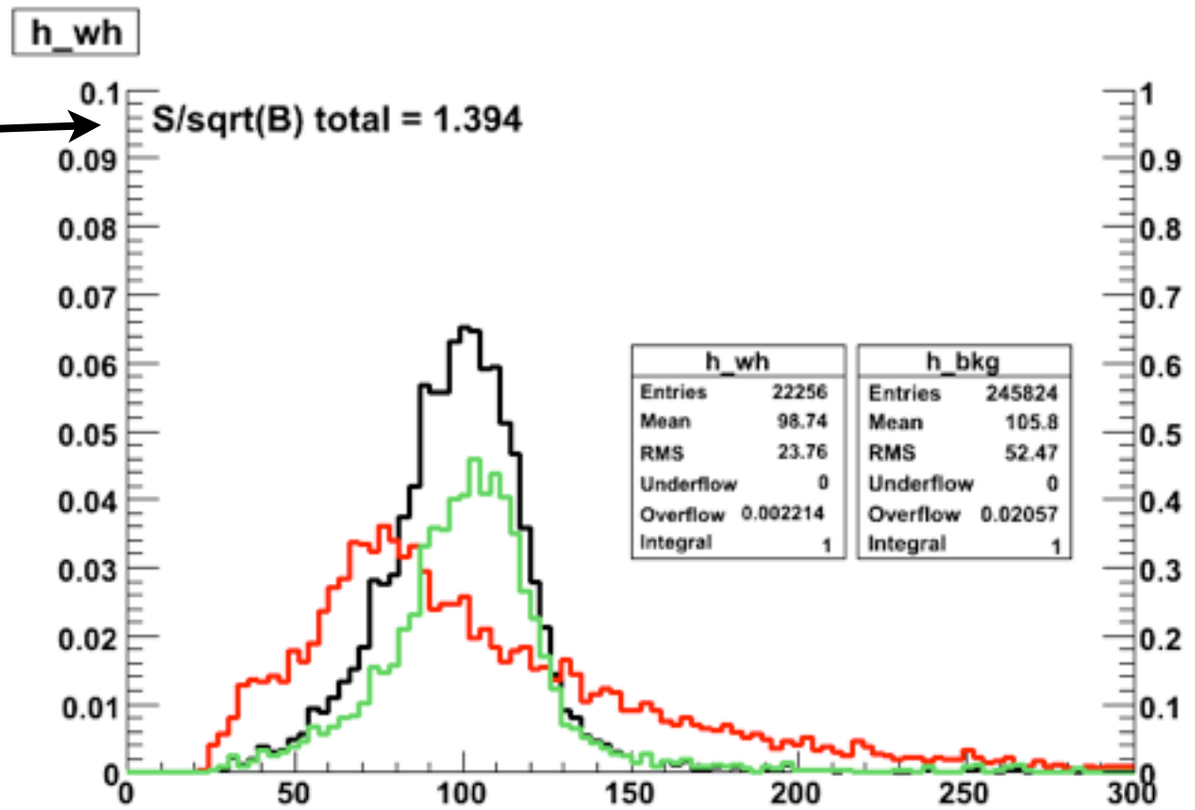
Plots are normalized to unit area

Signal WH 115
MC Backgrounds
S/sqrt(B)

Default Mjj

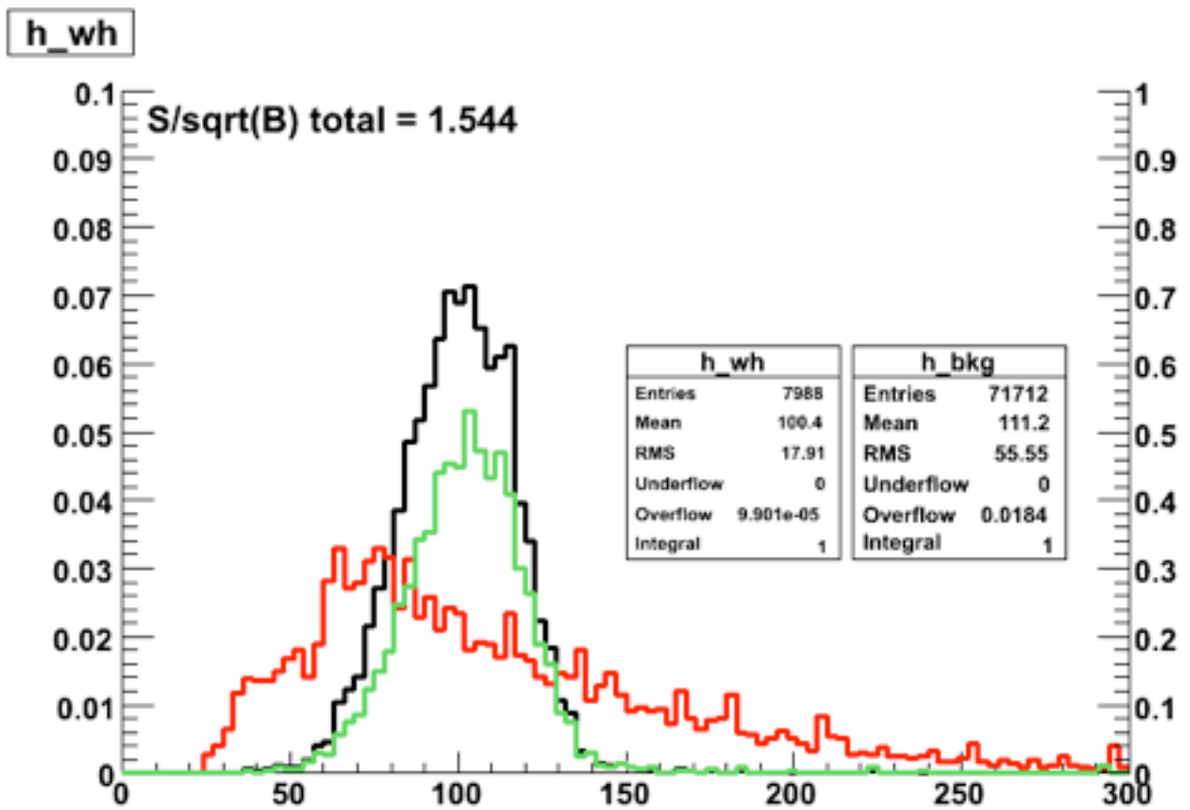
square root of sum of $(S/\sqrt{B})^2$ over all bins

Default ST



NO QCD here!
(and next slides...)

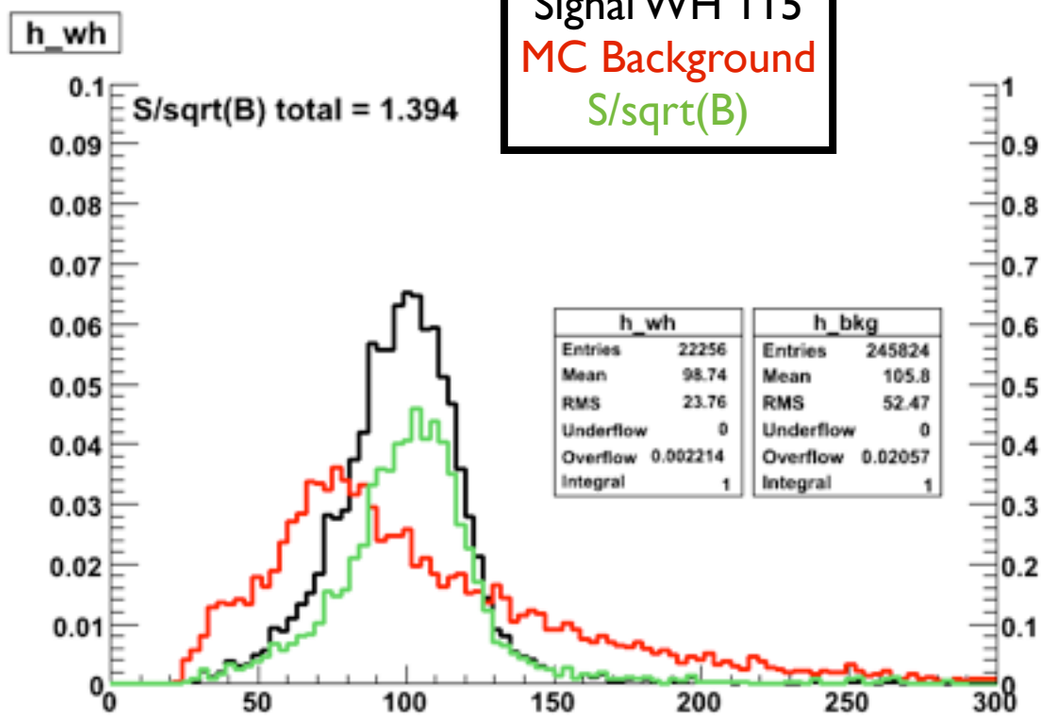
Default DT

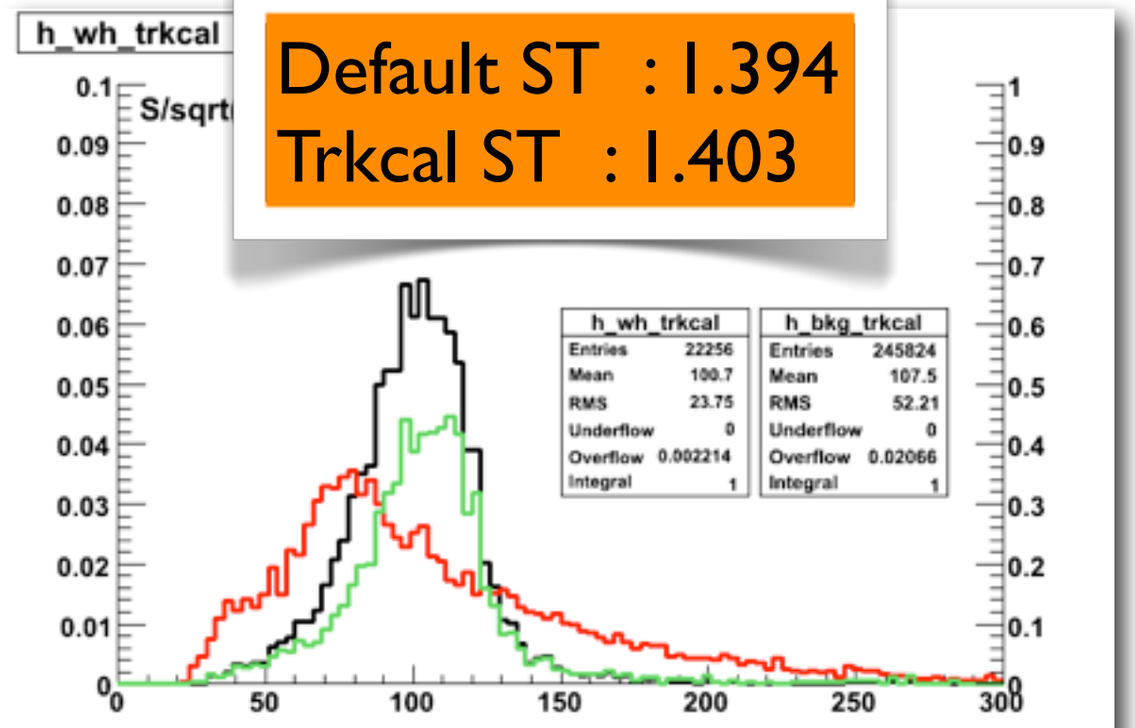


Signal WH 115
MC Backgrounds
S/sqrt(B)

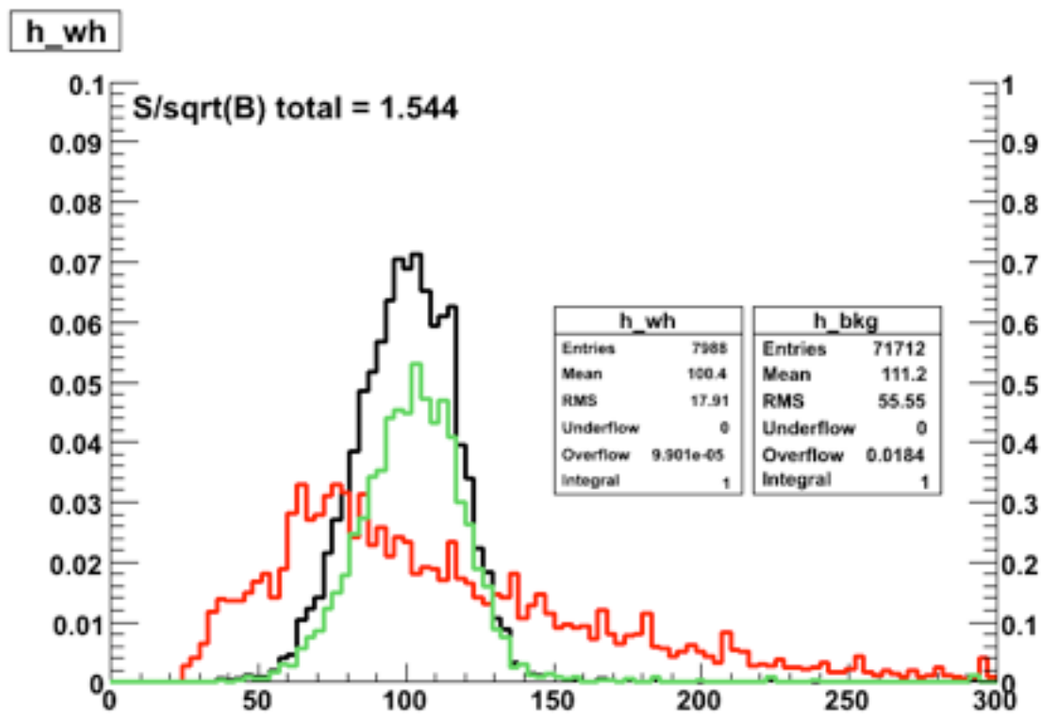
Trkcal Mjj

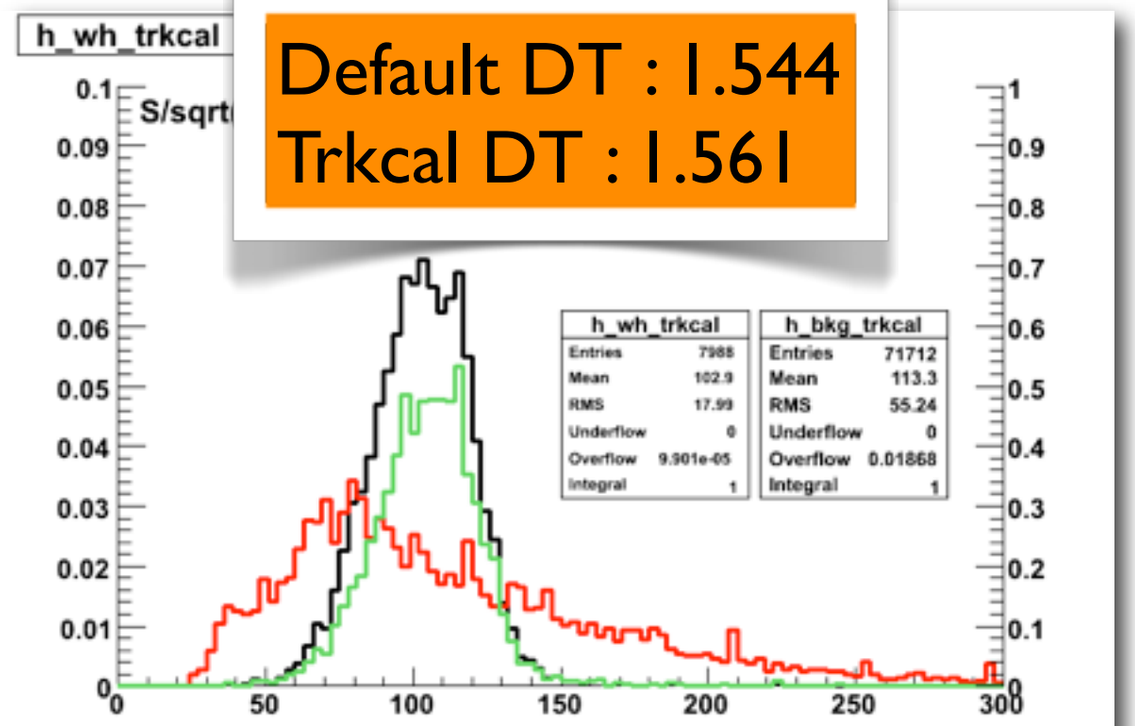
ST





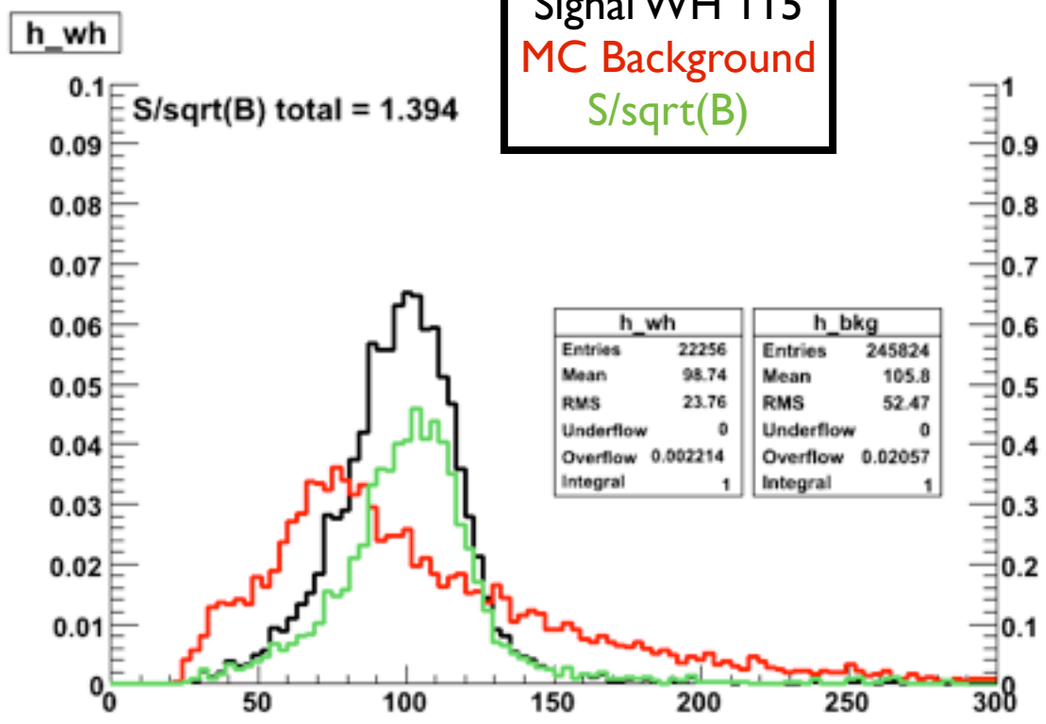
DT

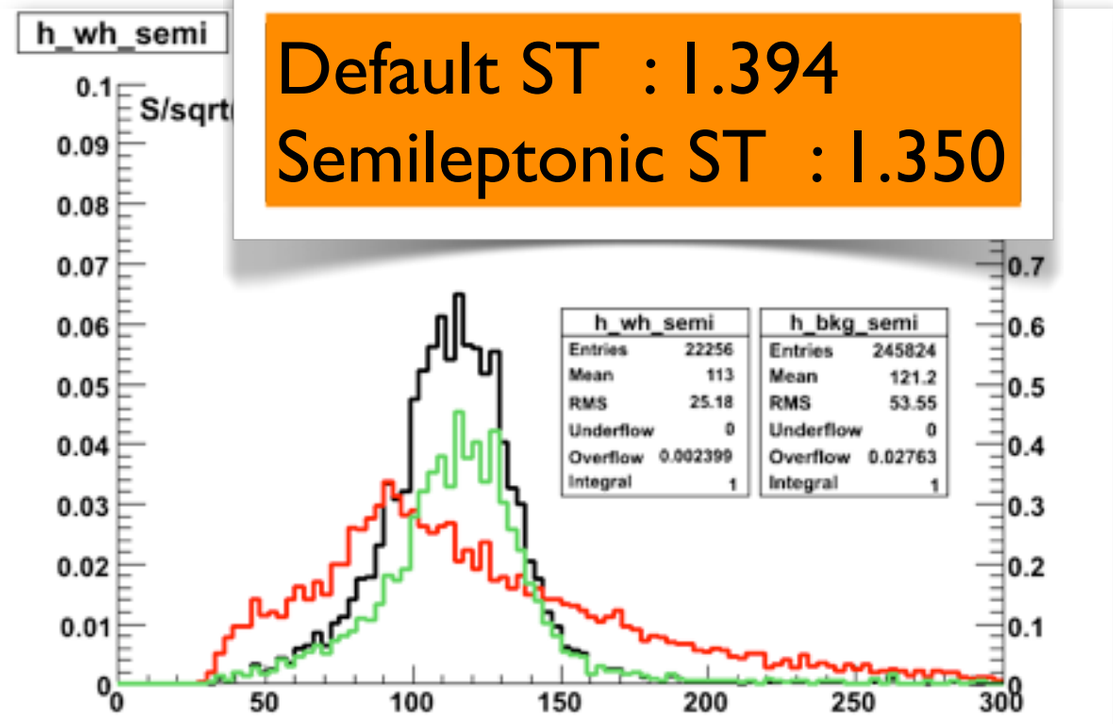




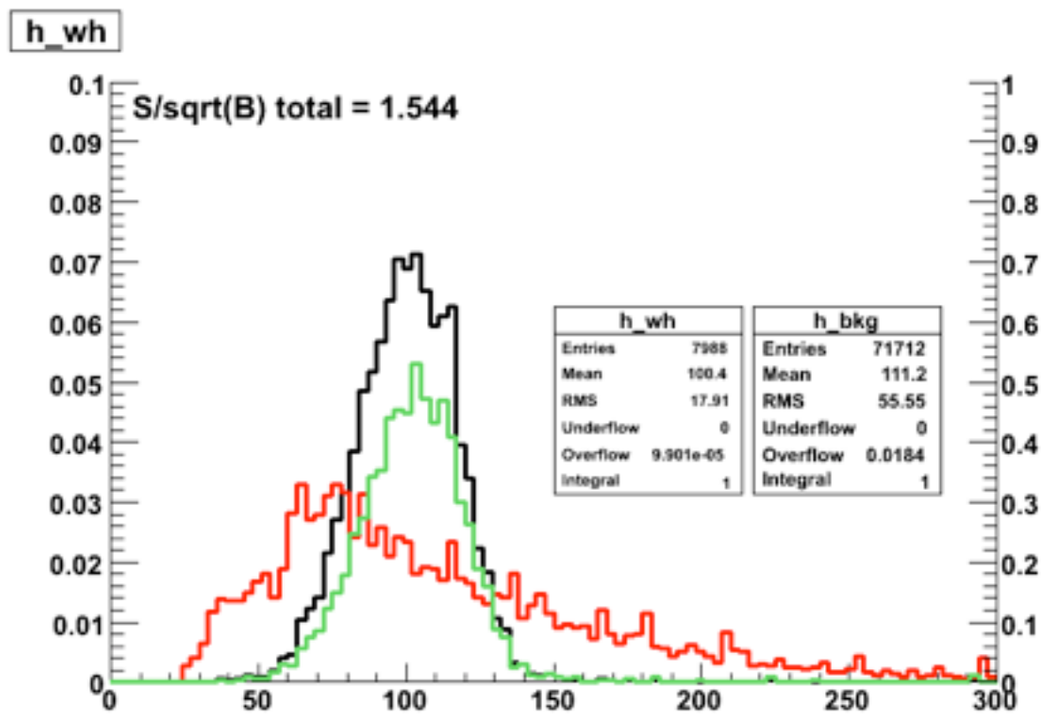
Semileptonic Mjj

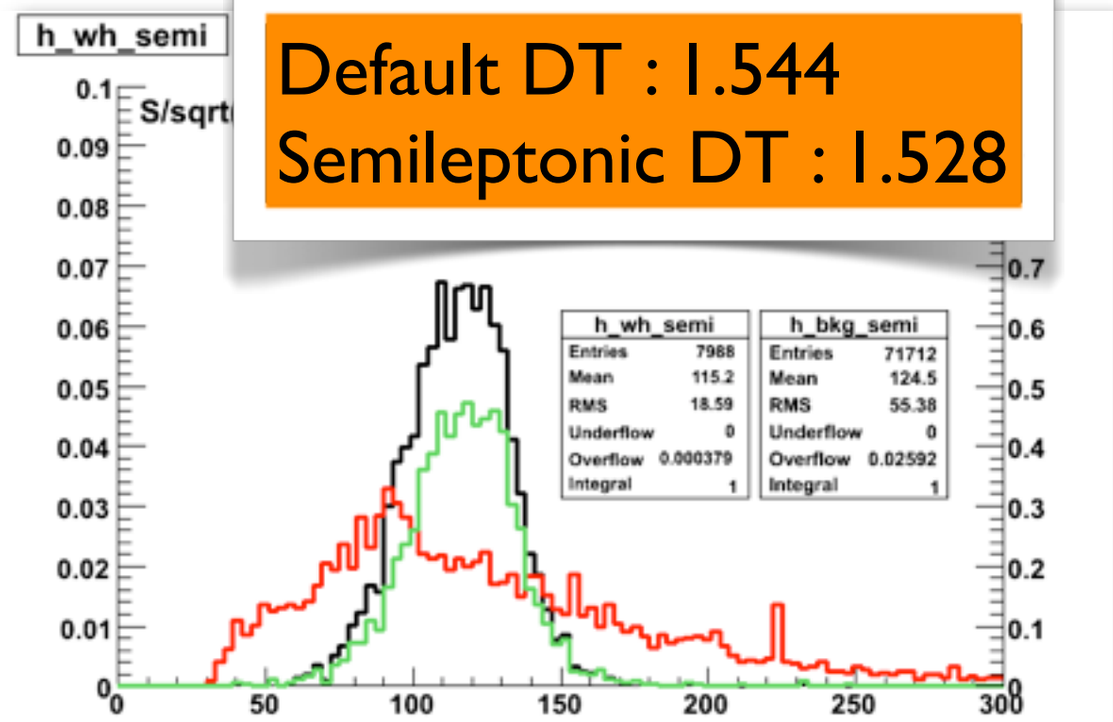
ST





DT

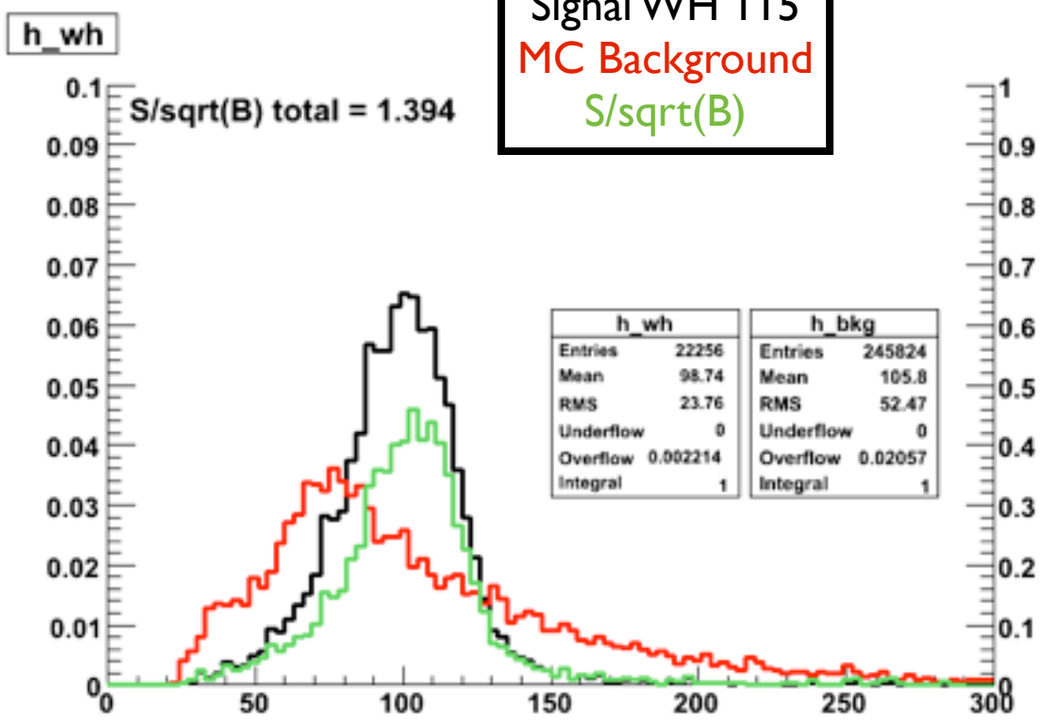


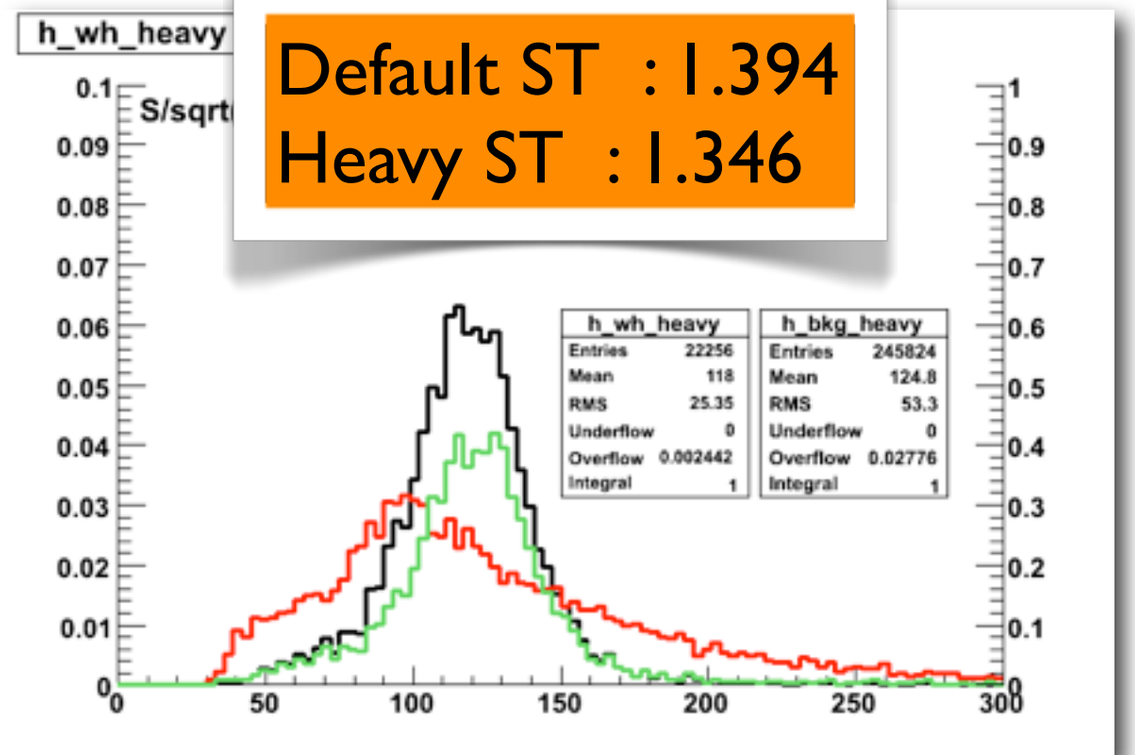


Heavy Mjj

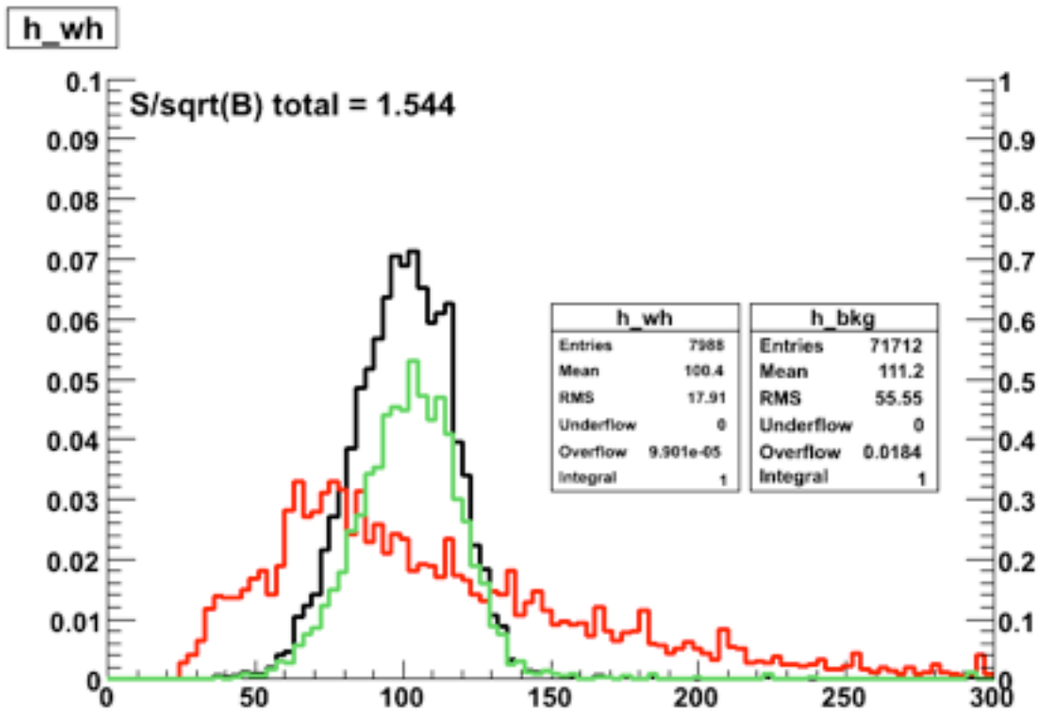
Signal WH 115
MC Background
S/sqrt(B)

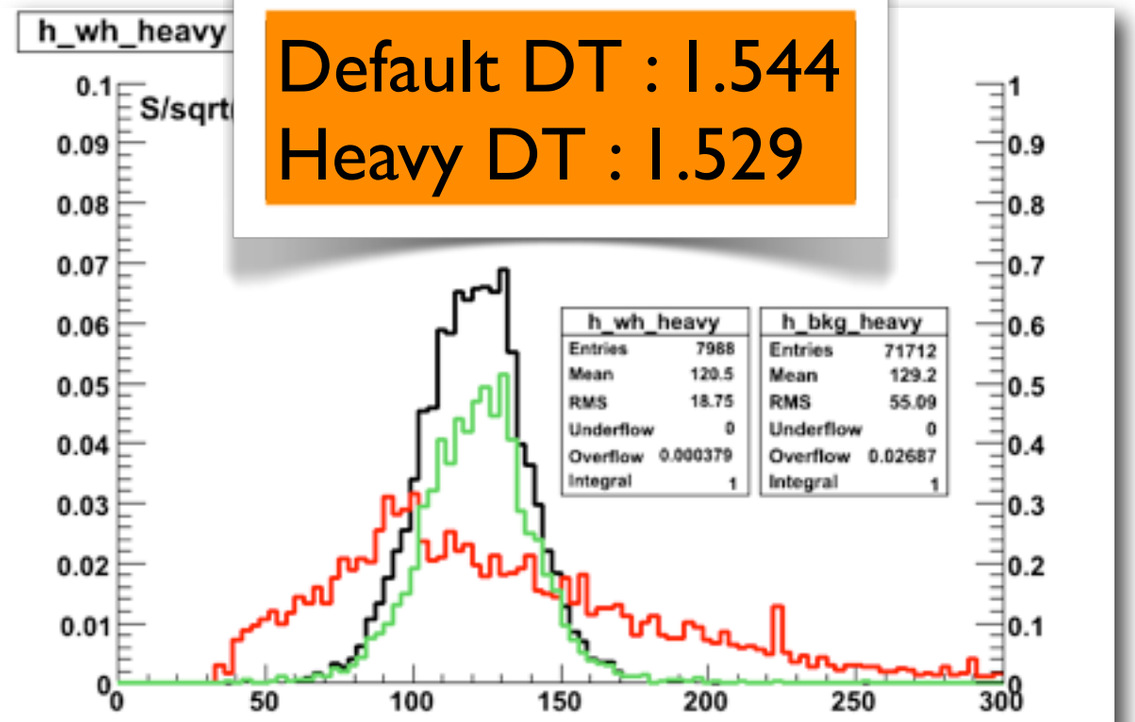
ST





DT





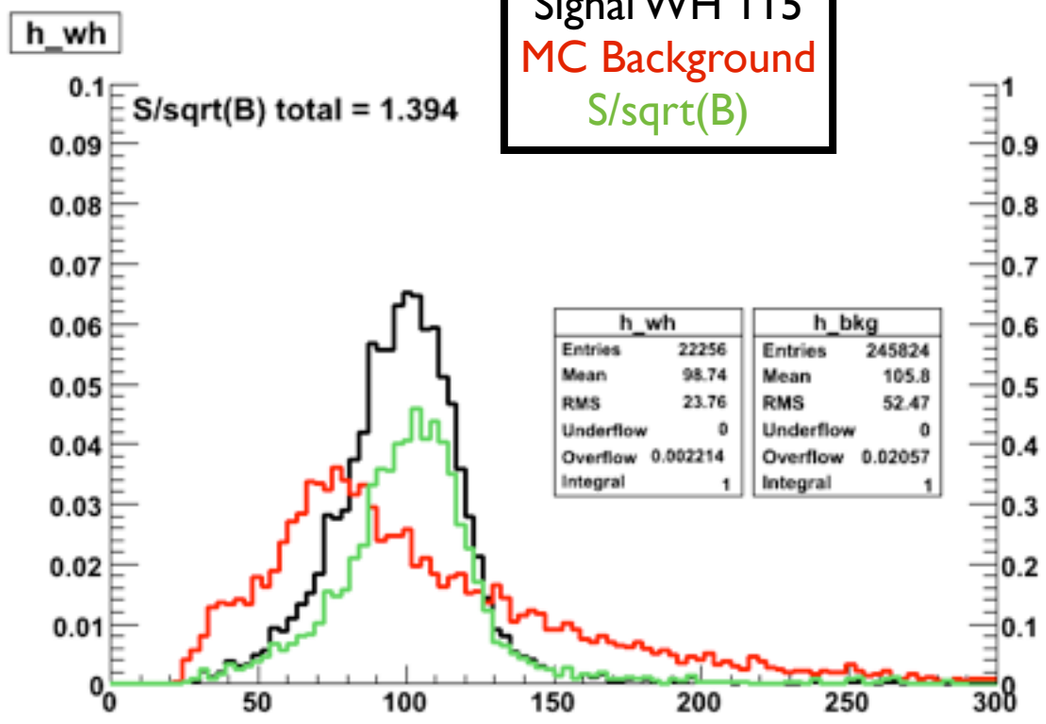
Potential of JER variables

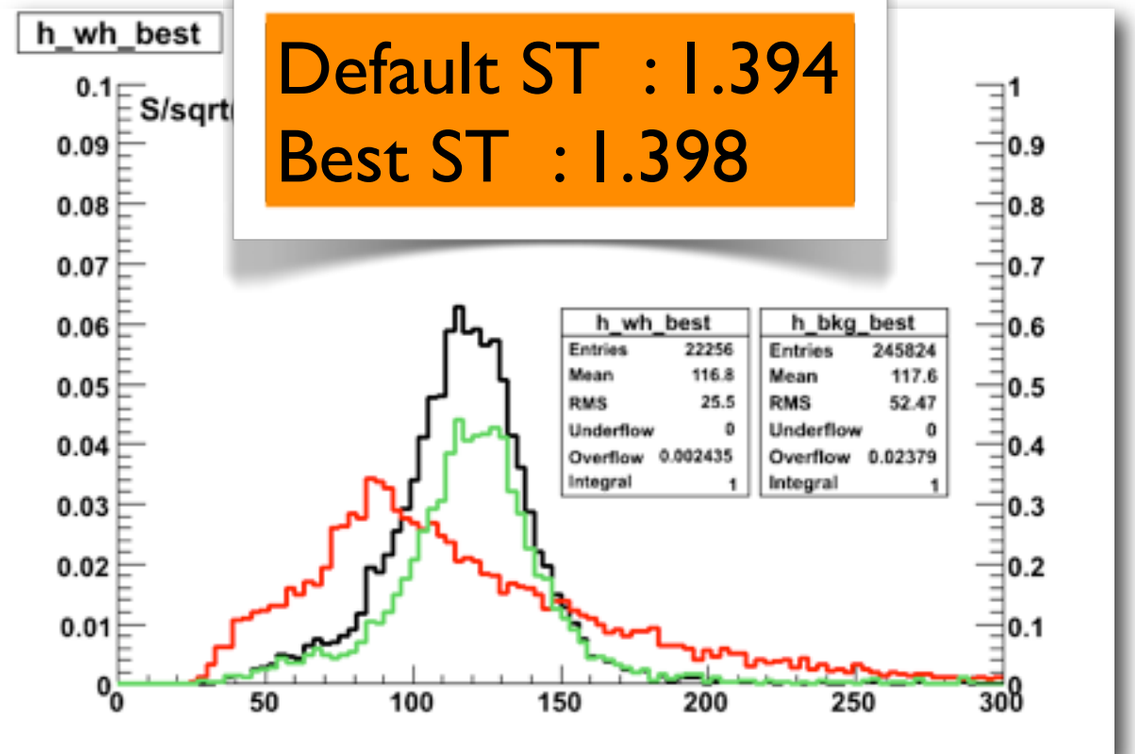
S/sqrt(B)	Single Tag	Double Tag
Default	1.394	1.544
CPS	1.390	1.544
Trkcal	1.403	1.561
HMatrixlight	1.391	1.561
HMatrix	1.391	1.561
Semileptonic	1.354	1.521
Light	1.394	1.549
Heavy	1.346	1.529

- If improved, only at the percent level
- "Let's cheat" to find out if these corrections can provide better S/sqrt(B) using MC truth information
- Redo the exercise, but for each jet
  - If MC flavor ID is "b/c-quark" → apply HEAVY correction
  - If MC flavor ID is NOT "b/c-quark" → apply LIGHT correction
- Virtually, this should be the optimal way how these corrections should be applied

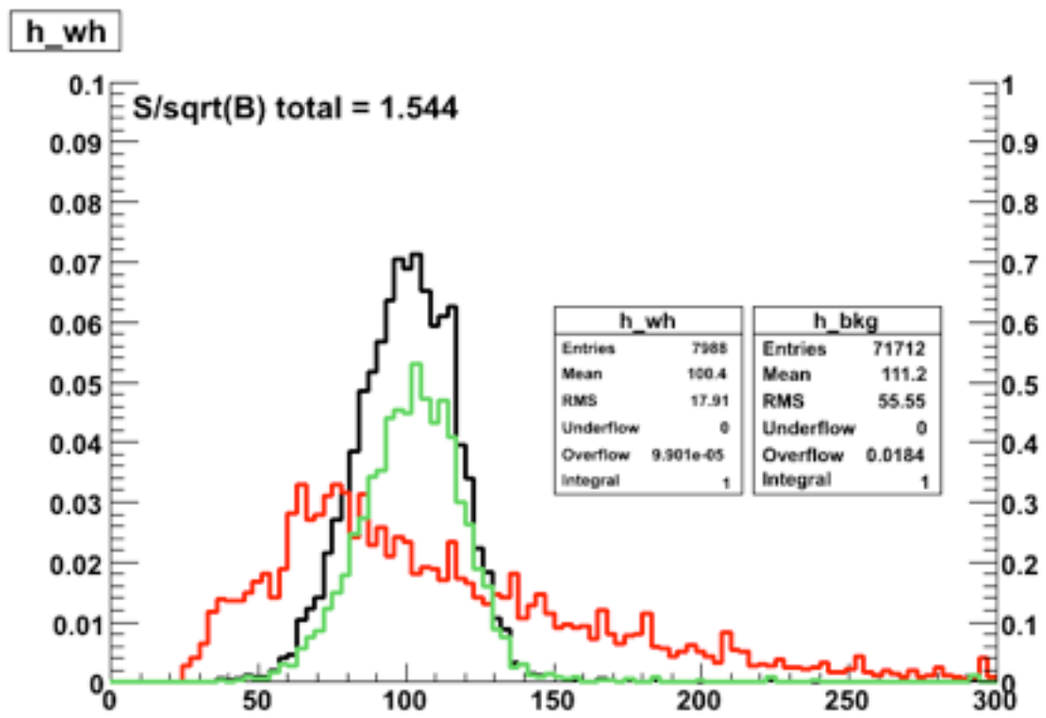
Best Mjj

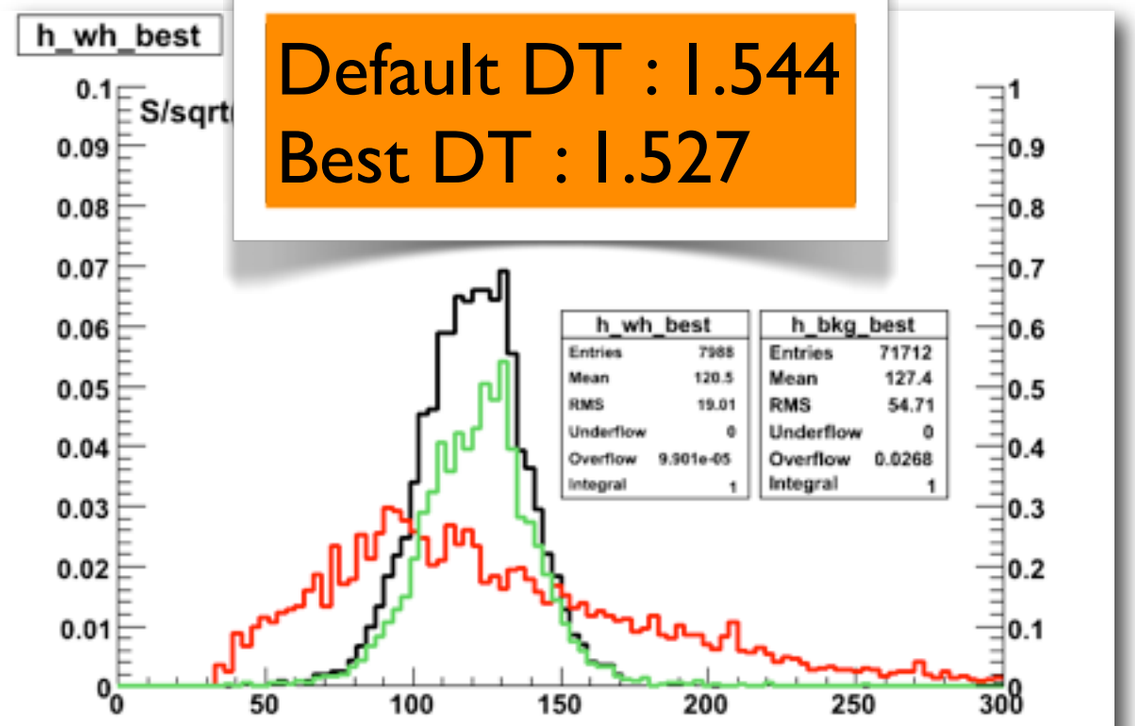
ST





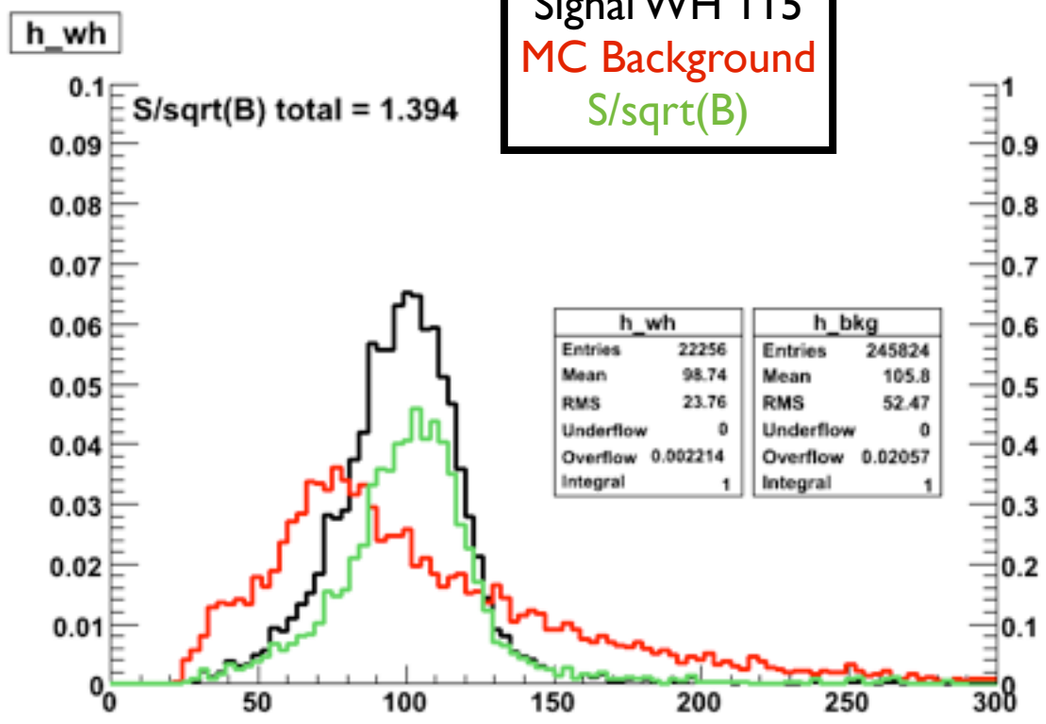
DT

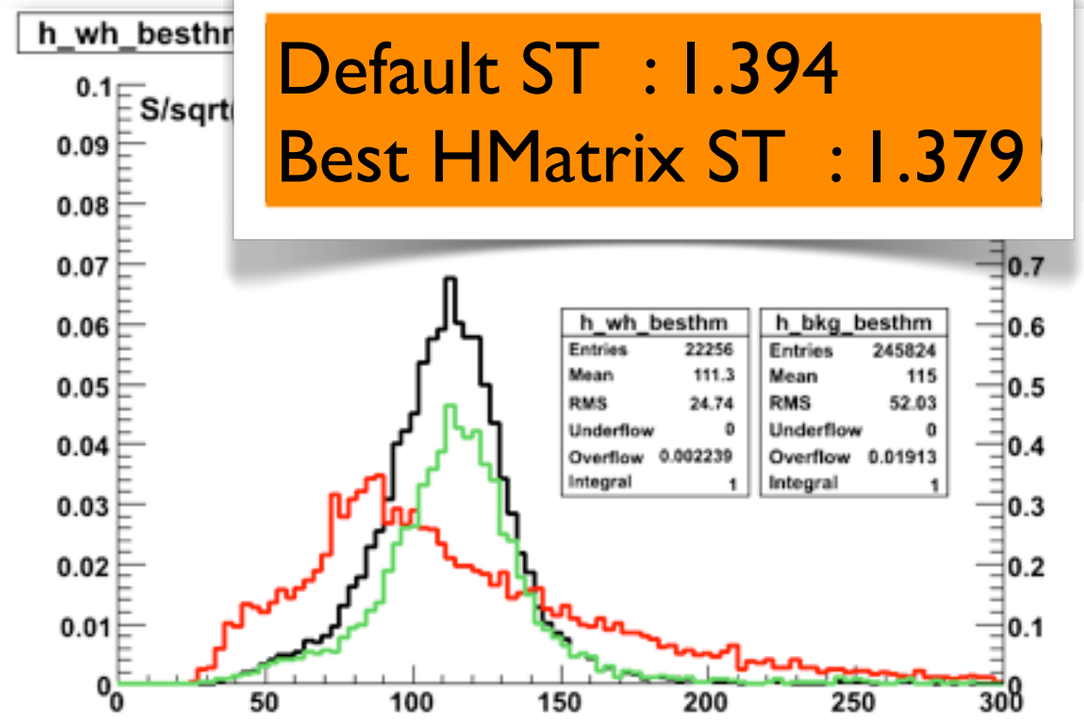




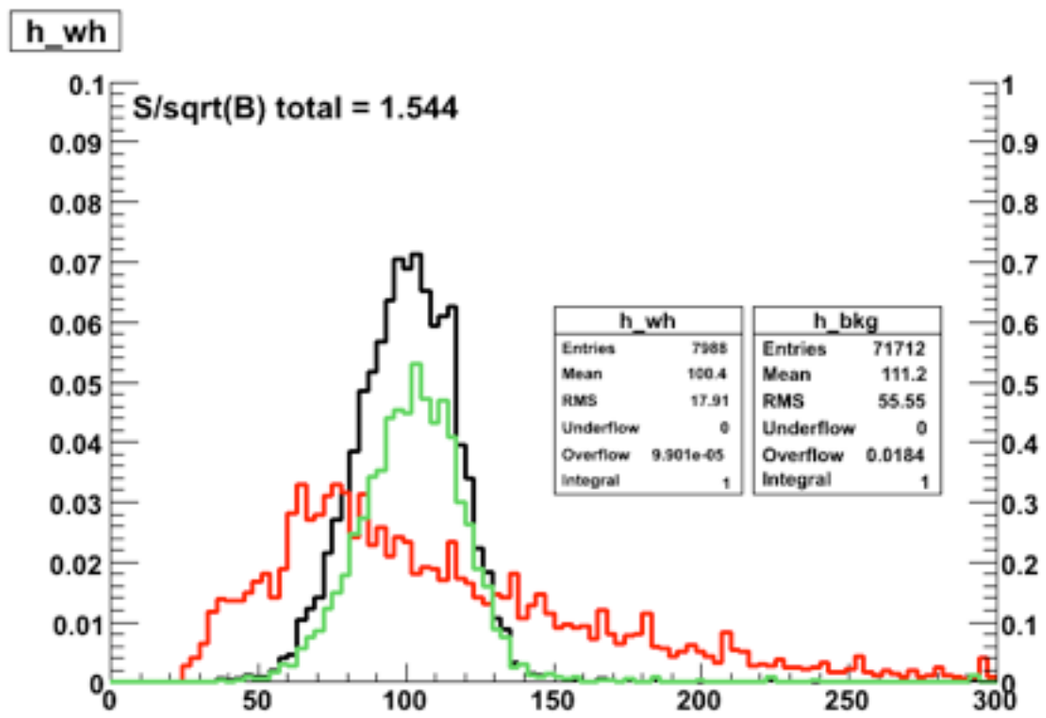
Best HMatrix Mjj

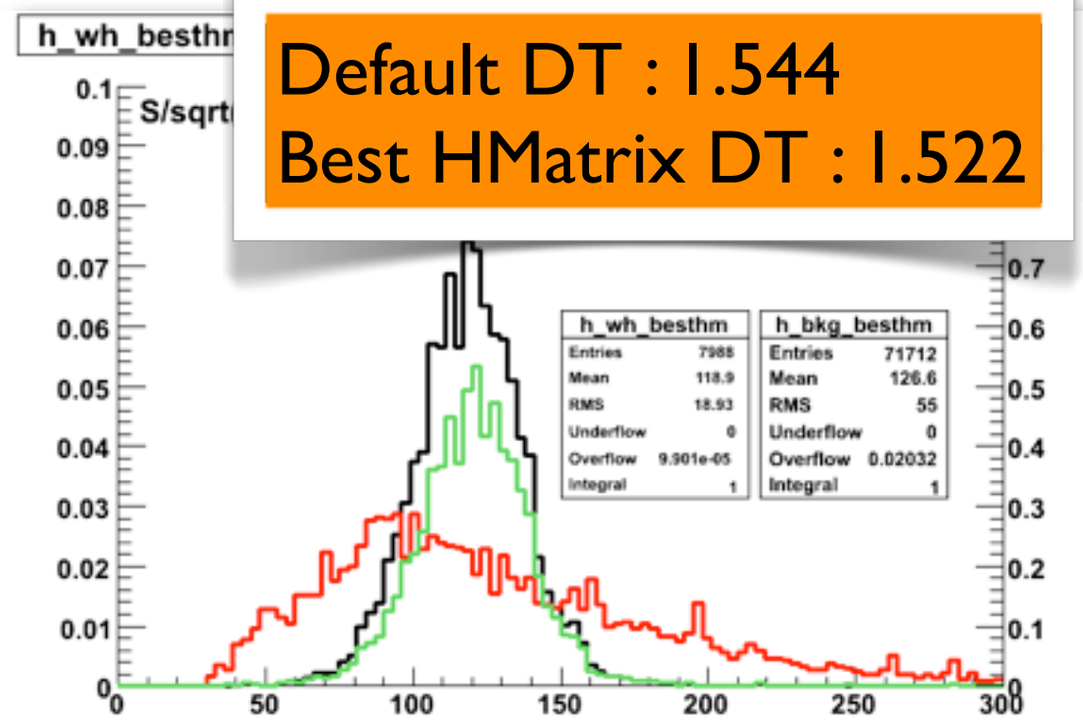
ST





DT





Jet width correction

- Ongoing work on an additionnal correction:
- Jet Energy Resolution is measured by the RMS of the following expression :
  $\mathbf{A = \sqrt{2} \times sign(\eta^{probe} - \eta^{tag}) \times (p_T^{tag} - p_T^{probe})/(p_T^{tag} + p_T^{probe})}$

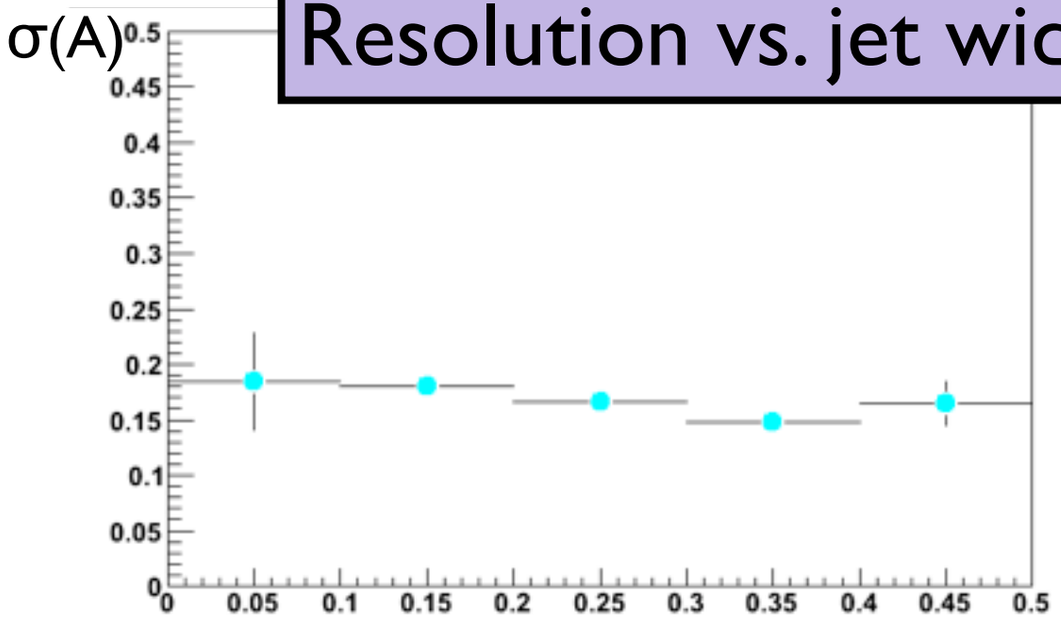
  where $\mathbf{\sigma(A) = \sigma(p_T)/p_T}$

tag & probe jets picked randomly

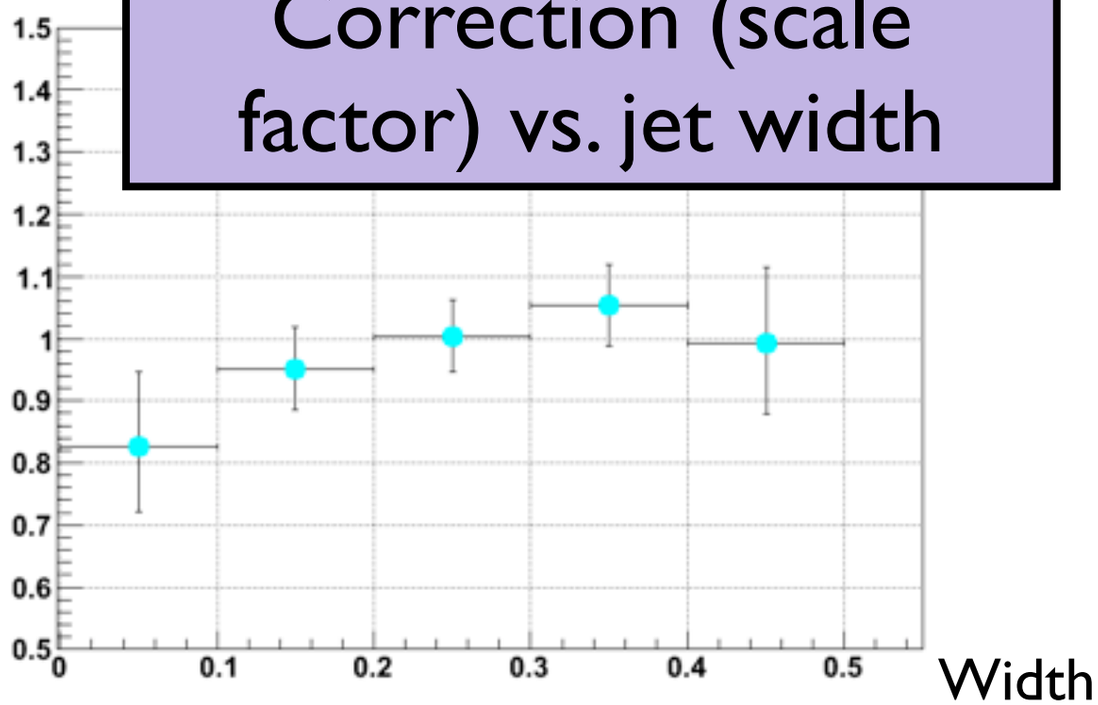
- By how much p_T^{probe} has to be scaled to minimize σ(A)?
  - → determine a scale factor k_w to correct any difference between 2 jets back-to-back ($\mathbf{p_T^{probe} \rightarrow k_w \times p_T^{probe}}$)
- (How) does it depend on the probe jet width (in different p_T bins, maybe btagging...)?
- Selection :
  - Exactly 2 CC jets, back to back
  - In different ranges of $(p_T^{tag} + p_T^{probe})/2$
- In order to look at the asymmetry in the most unbiased way, we're looking now at a data sample, events firing **ZBMB trigger** → low p_T ($15 < (p_T^{tag} + p_T^{probe})/2 < 25$ GeV) & low statistics (highly prescaled trigger)

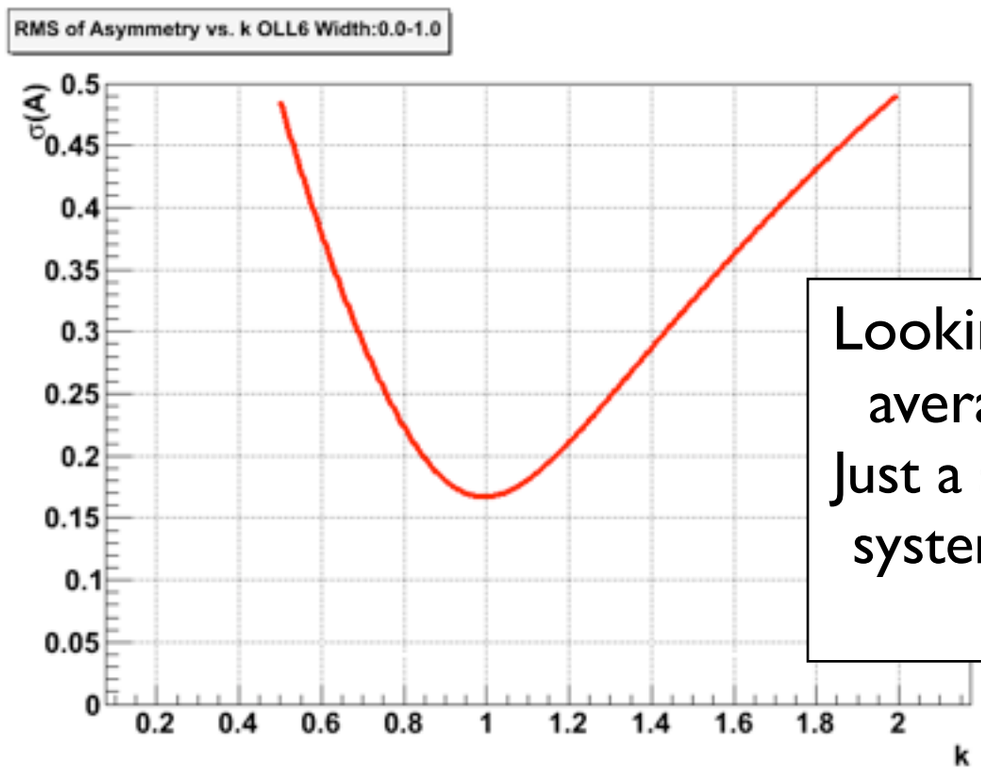
Jet width correction - **15 < (p_T^{tag} + p_T^{probe})/2 < 25 GeV**

Resolution vs. jet width



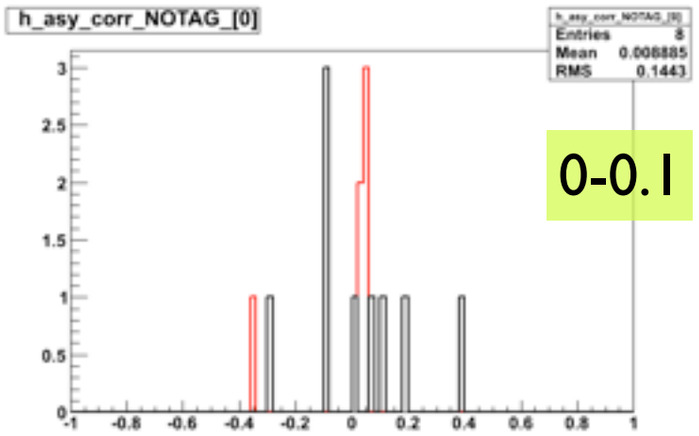
Correction (scale factor) vs. jet width



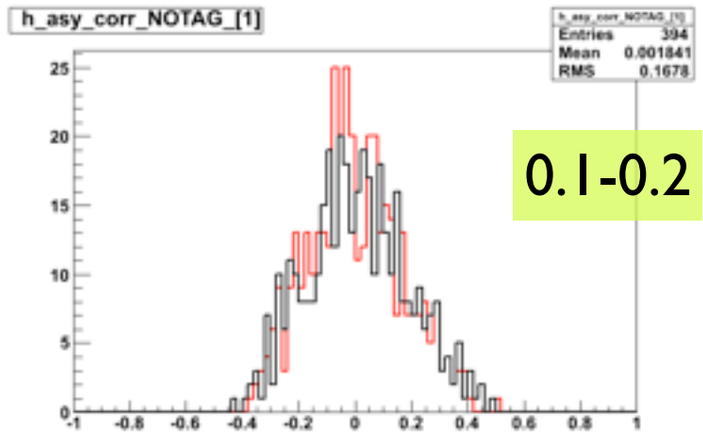


Looking at all jet widths inclusively, the average correction is 1, as expected
Just a sanity check, it show that the dijet system is balanced (good news ;)) and no trigger bias

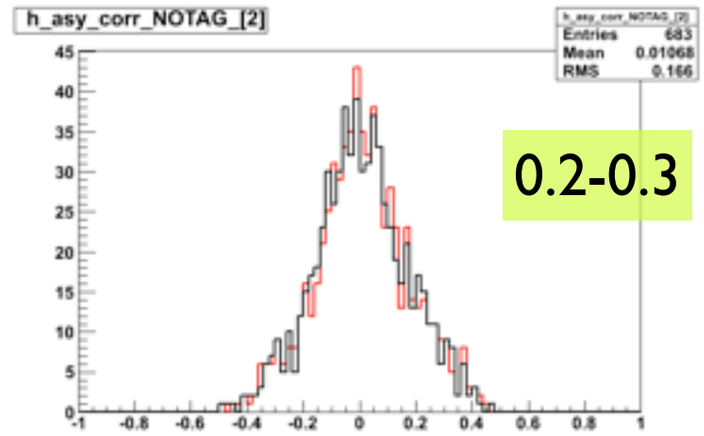
Jet width correction - $15 < (p_T^{tag} + p_T^{probe})/2 < 25$ GeV



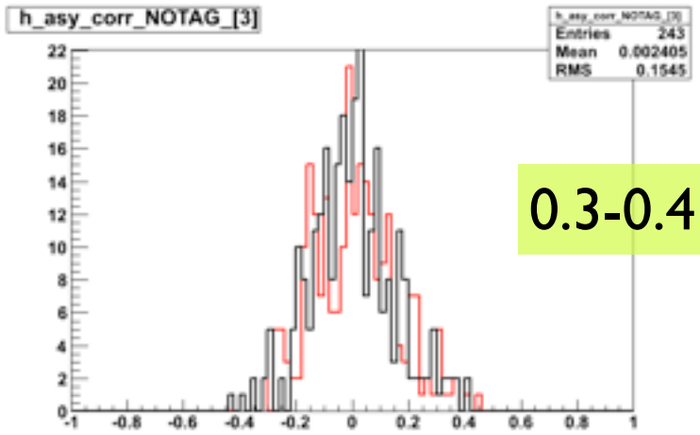
0-0.1



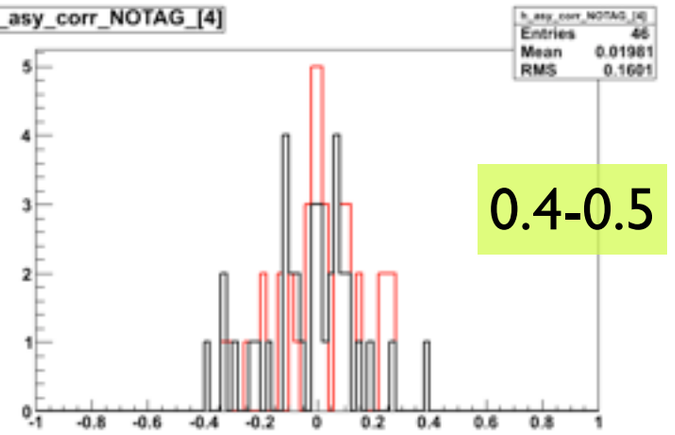
0.1-0.2



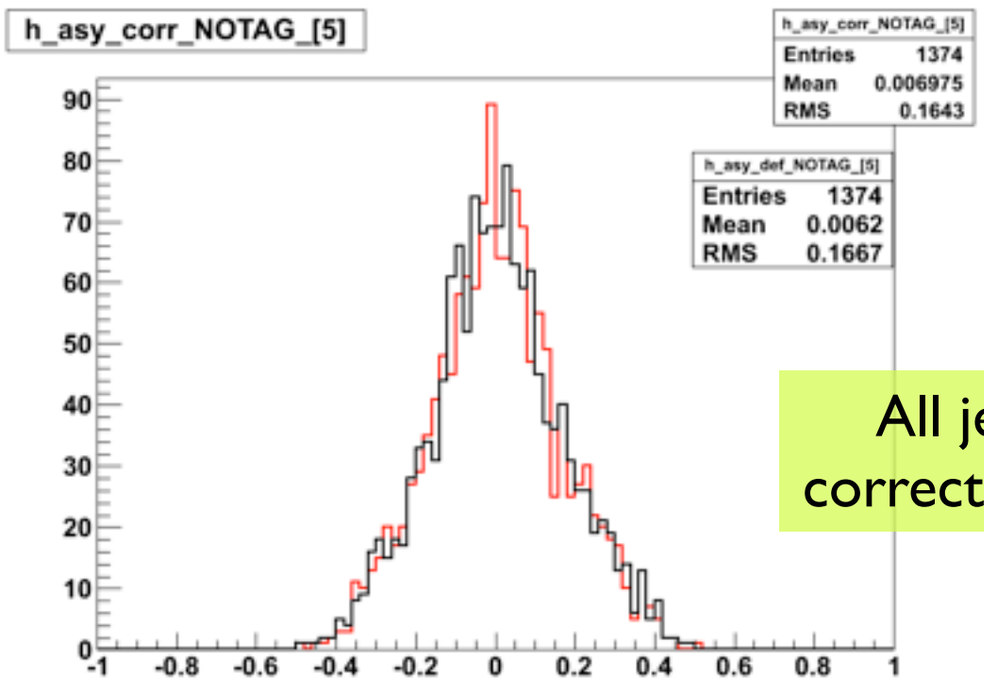
0.2-0.3



0.3-0.4



0.4-0.5

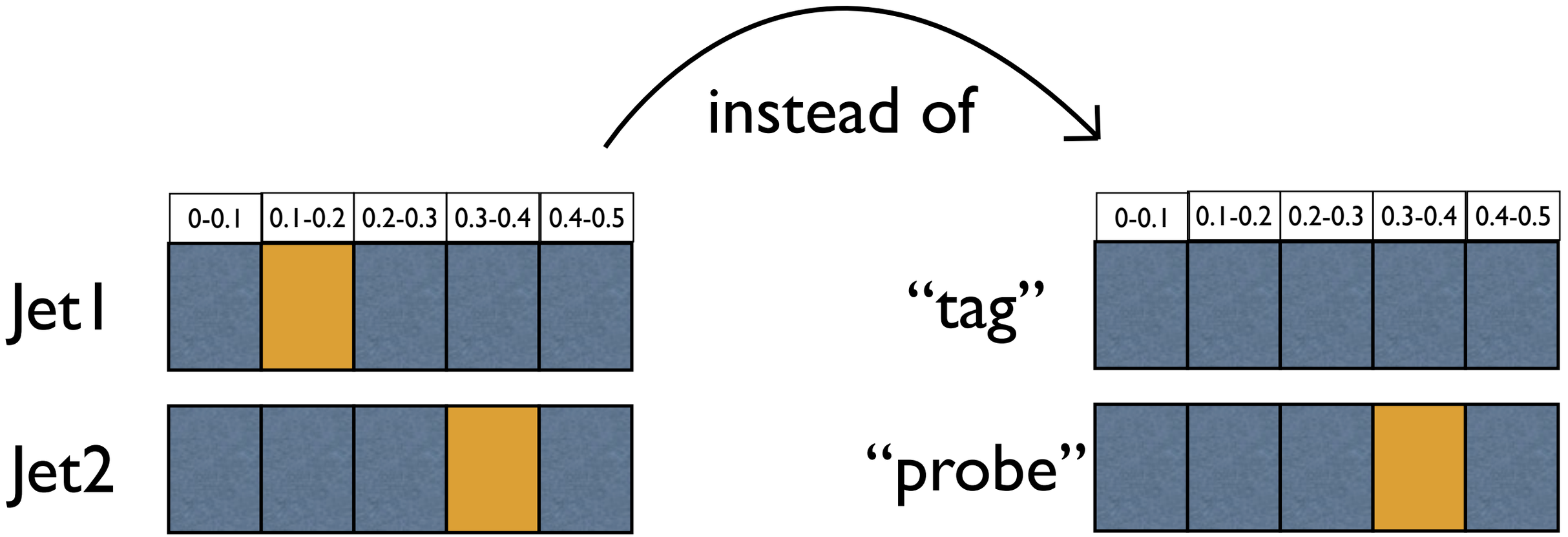


All jet widths corrections applied

Asymmetry before/ after correction, applied only on "probe" jet

Jet width correction

- Now, picking randomly a jet in the event as "tag" or "probe"
- Another solution instead of correcting one jet picked randomly is to correct both jets
- **$A = \sqrt{2} \times \mathrm{sign}(\eta^{j1} - \eta^{j2}) \times (k_{j1} \times p_T^{j1} - k_{j2} \times p_T^{j2})/(k_{j1} \times p_T^{j1} + k_{j2} \times p_T^{j2})$**, if jet1 and jet2 are in different width bins ($k_{j1} \neq k_{j2}$) → on average doubles the statistics in every jet width bin to derive the correction
- For higher p_T bins, use inclusive jet trigger sample where both jets fire the same trigger (JT15?), and check if the corrections is the same when extrapolated to lower p_T as the one derived using ZBMB trigger



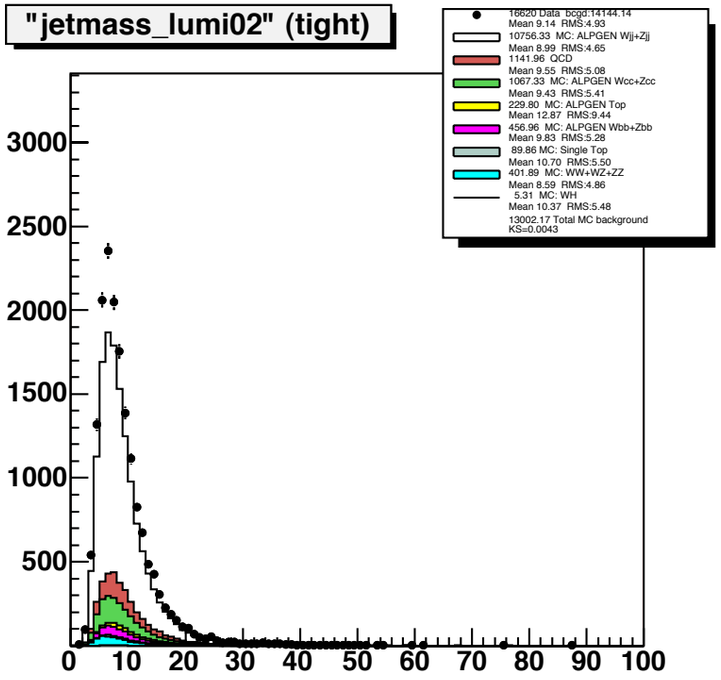
Conclusion

- ZHvvbb and ZHllbb groups see about 5% improvement seen from JER variables in MVAs
- Study on JER variables potential indicates that the discriminating power is not so big, and they are highly correlated
- So far, no significant improvement seen in WH analysis, ongoing studies to confirm this
- Investigate the BDT parameters to make sure we are not just missing the point
- A decision whether to use these corrections for the summer should be made at the Higgs Workshop, estimate systematics (the jet mass which is an input for semileptonic & heavy correction is not properly modeled, may need to correct this)

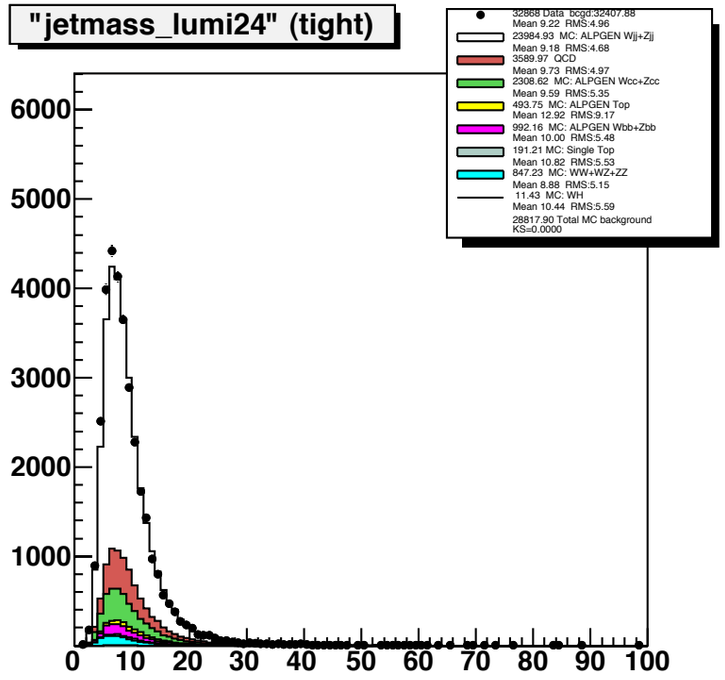
- Jet width correction under investigation, many things to look at still :
  - real dependence of the jet width vs. p_T, η, quark/gluon composition, lumi
  - for ZBMB, not so much of a problem but we will need to understand possible trigger biases when switching to jet inclusive triggers
- about ~5% improvement expected from preliminary studies

Backup Slides

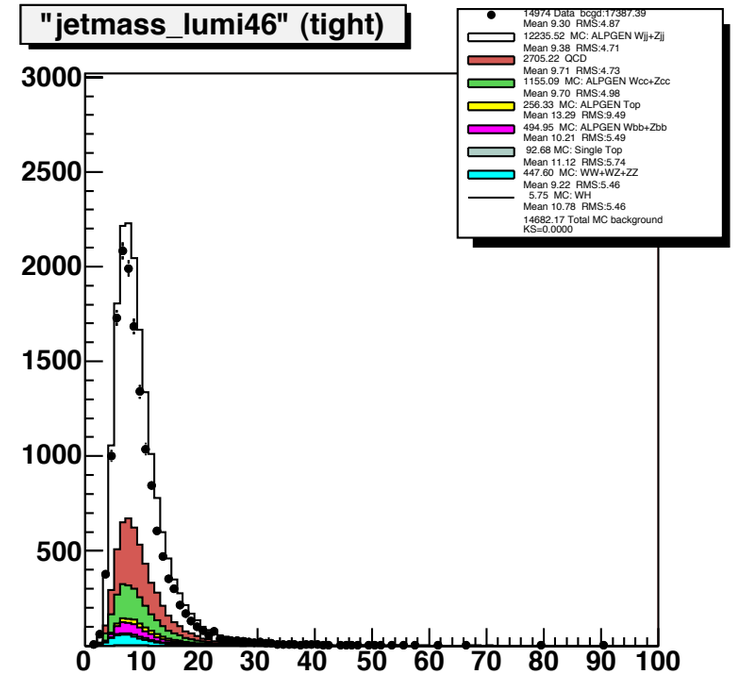
Jet Mass - lumi bins



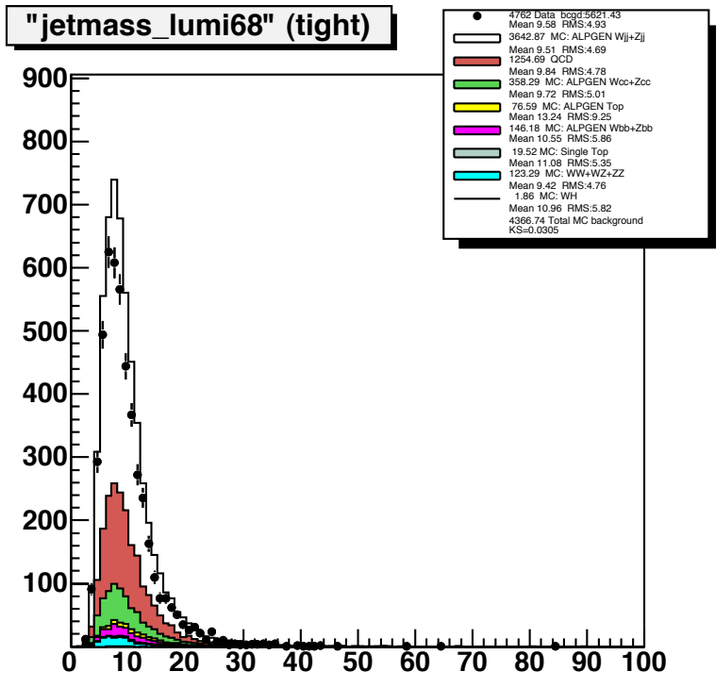
lumi per tick 0-2



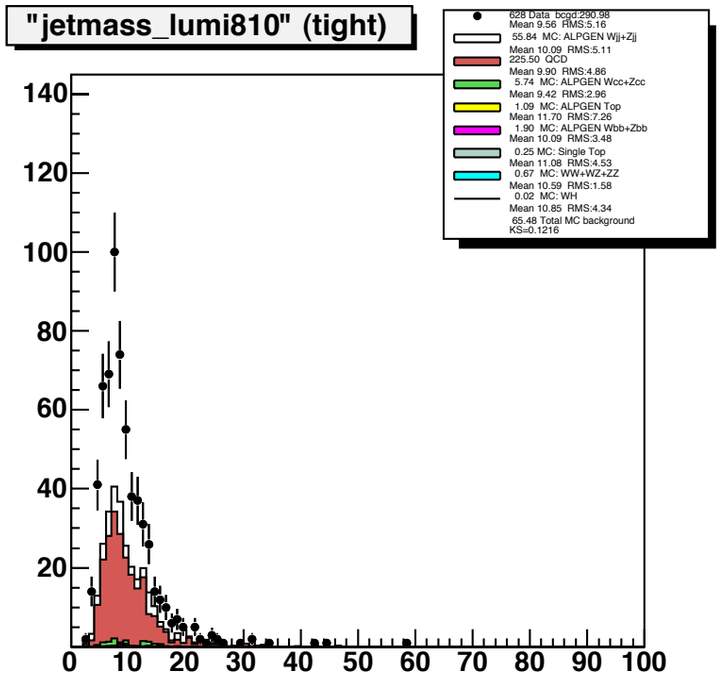
lumi per tick 2-4



lumi per tick 4-6

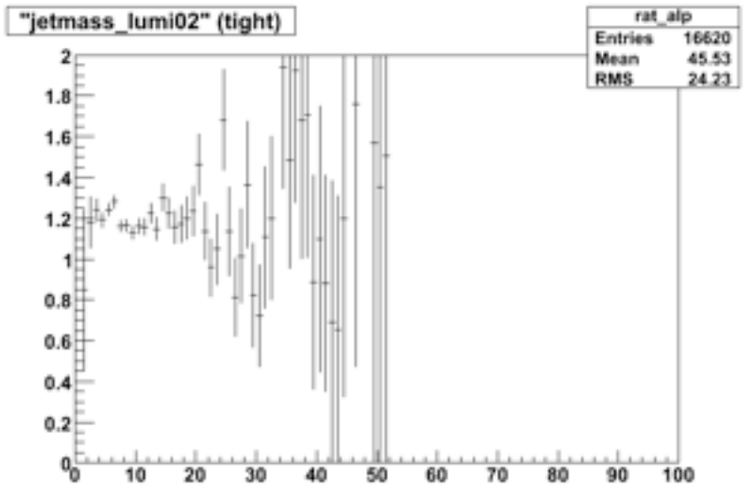


lumi per tick 6-8

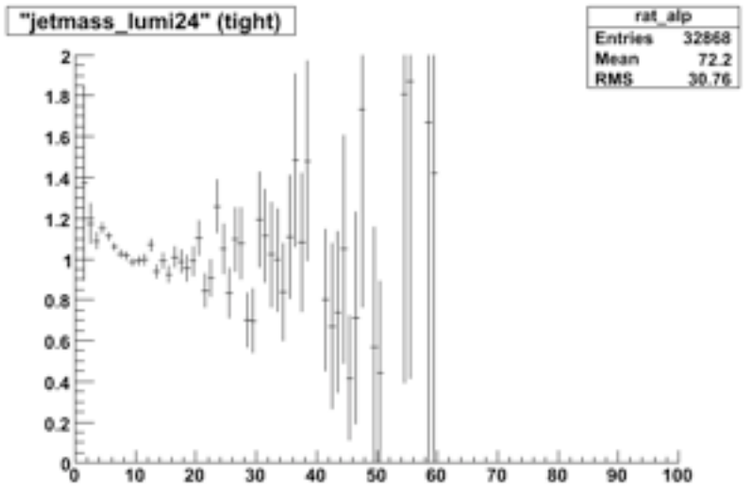


lumi per tick 8-10

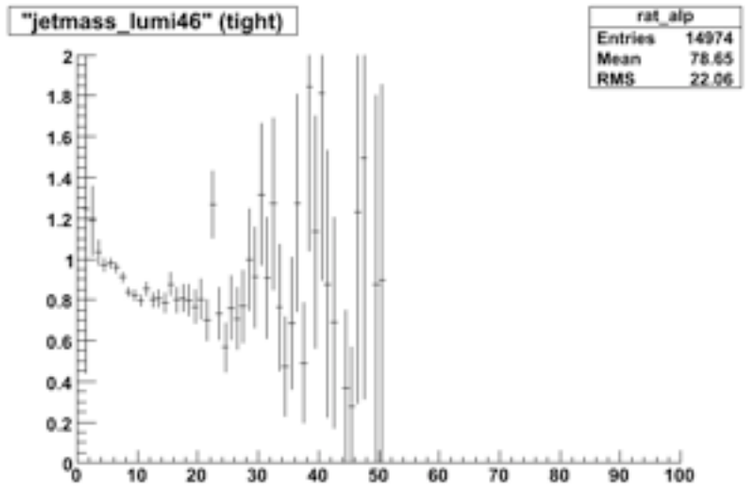
Jet Mass - lumi bins



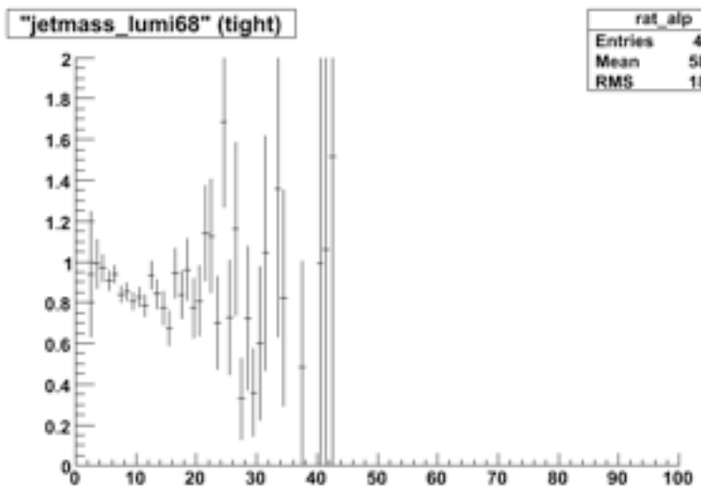
lumi per tick 0-2



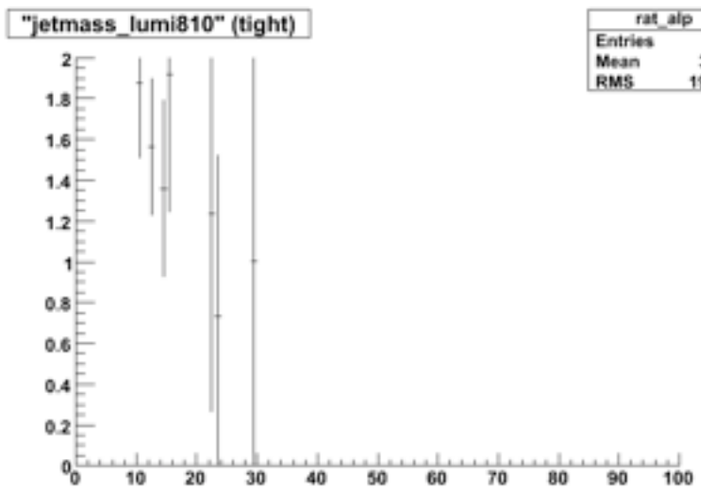
lumi per tick 2-4



lumi per tick 4-6

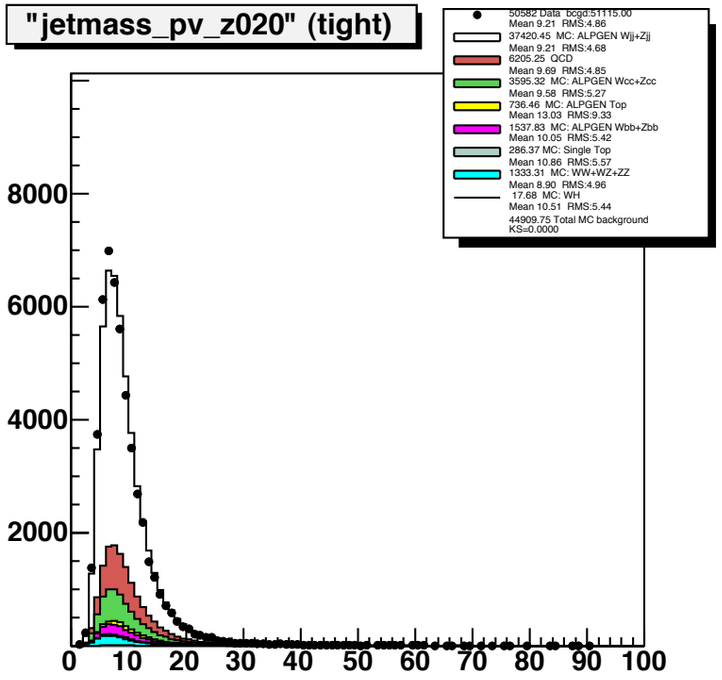


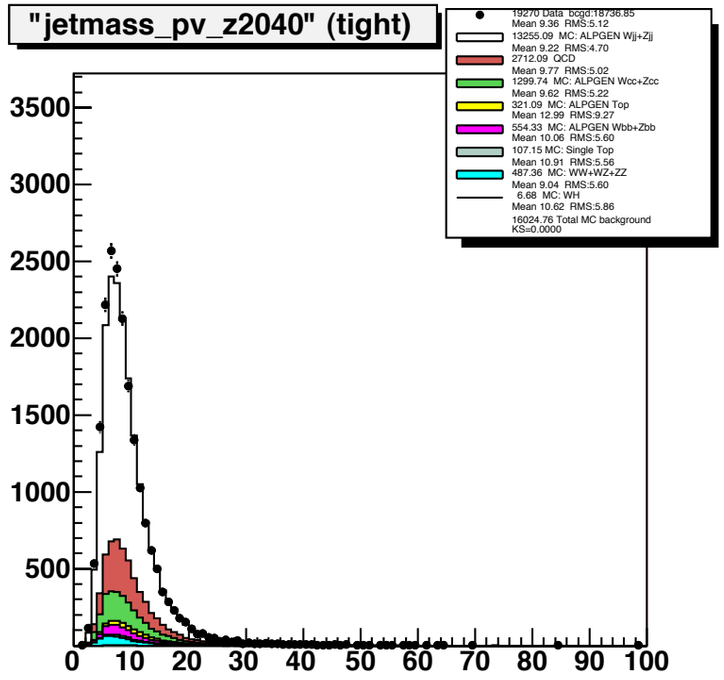
lumi per tick 6-8

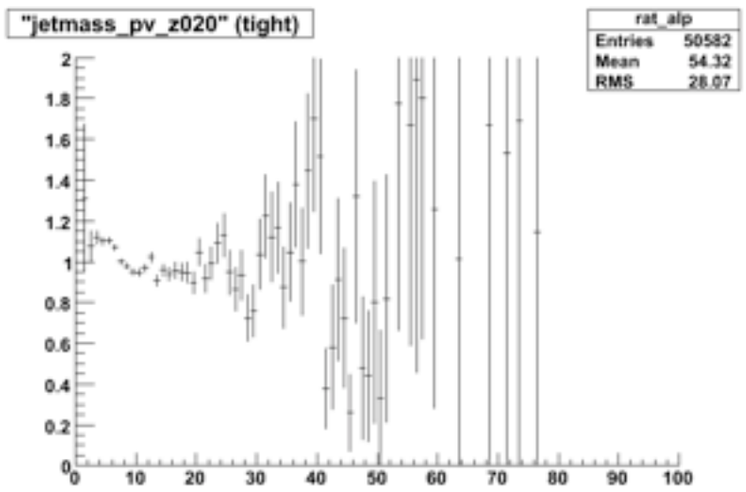


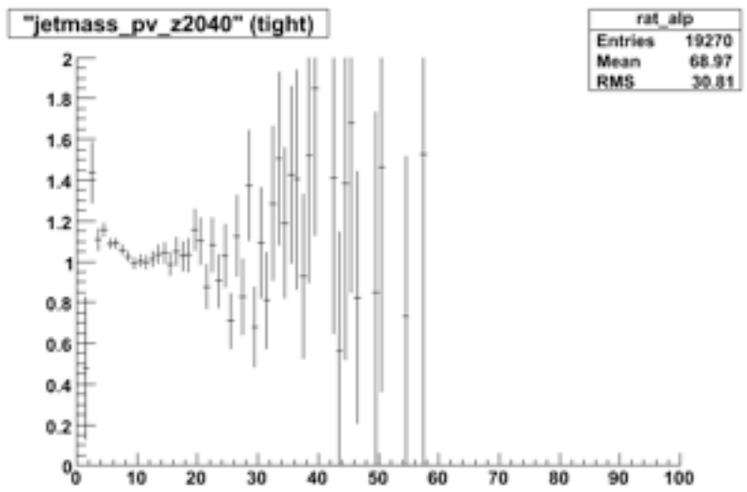
lumi per tick 8-10

Jet Mass - PVz bins





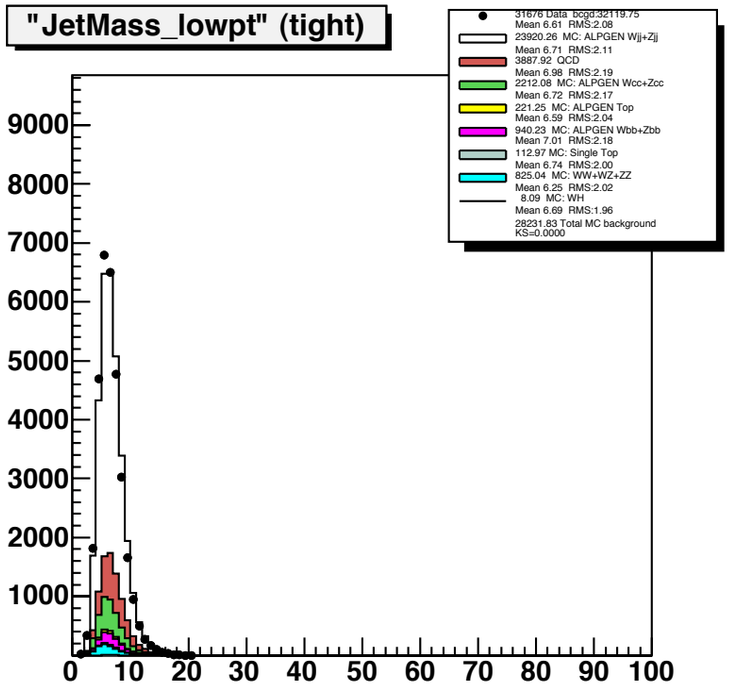


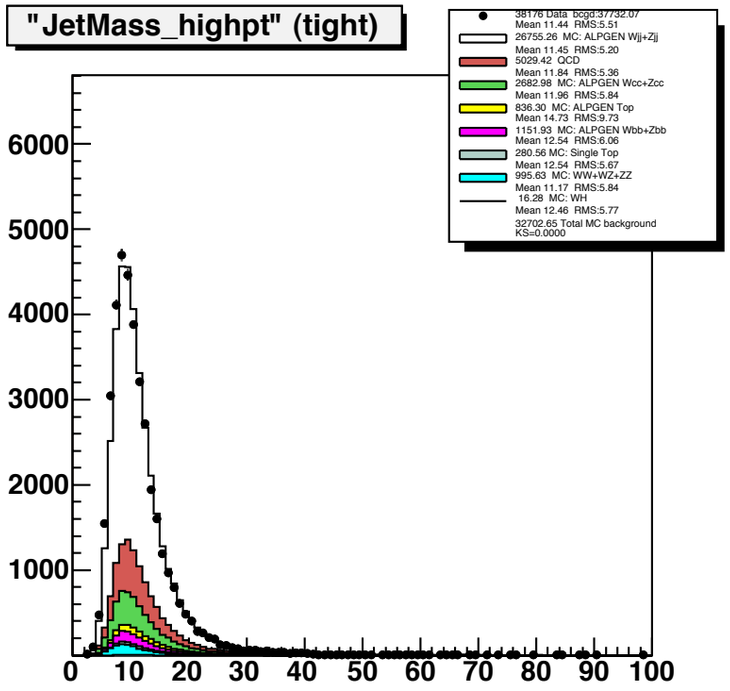


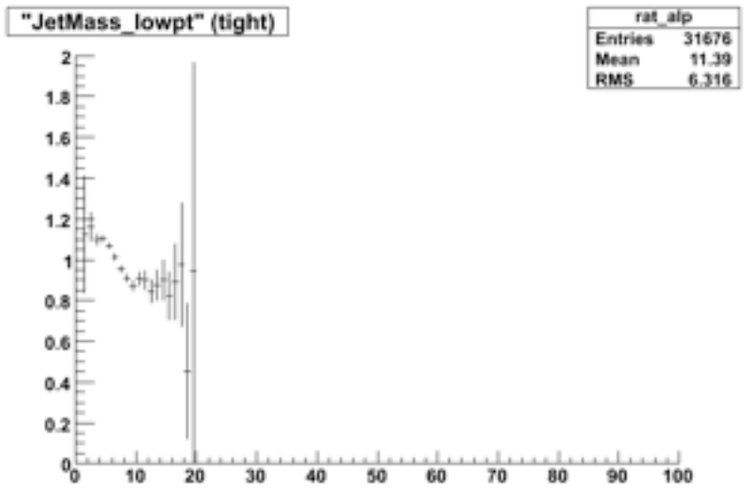
|PVz| 0-20

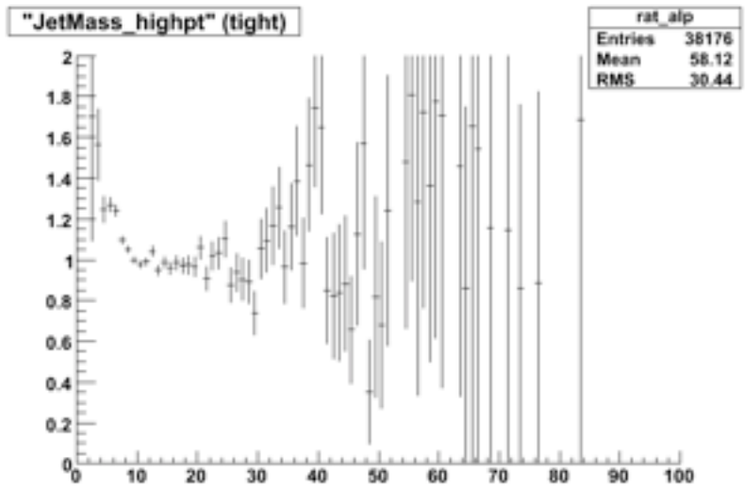
|PVz| 20-40

Jet Mass - pT bins





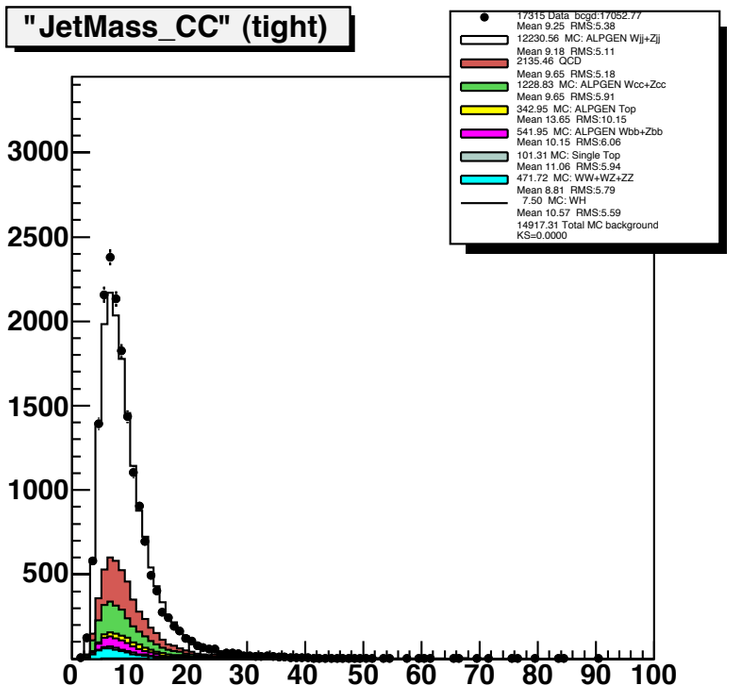


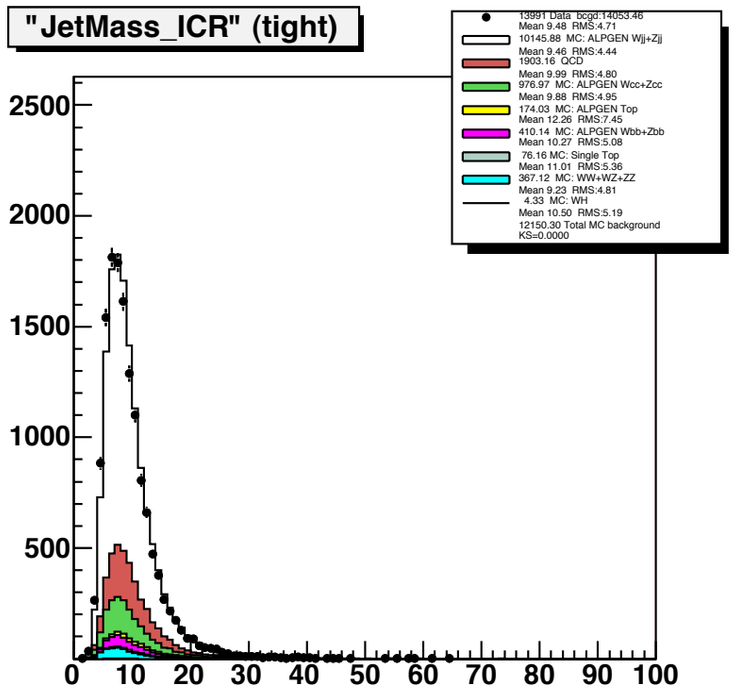


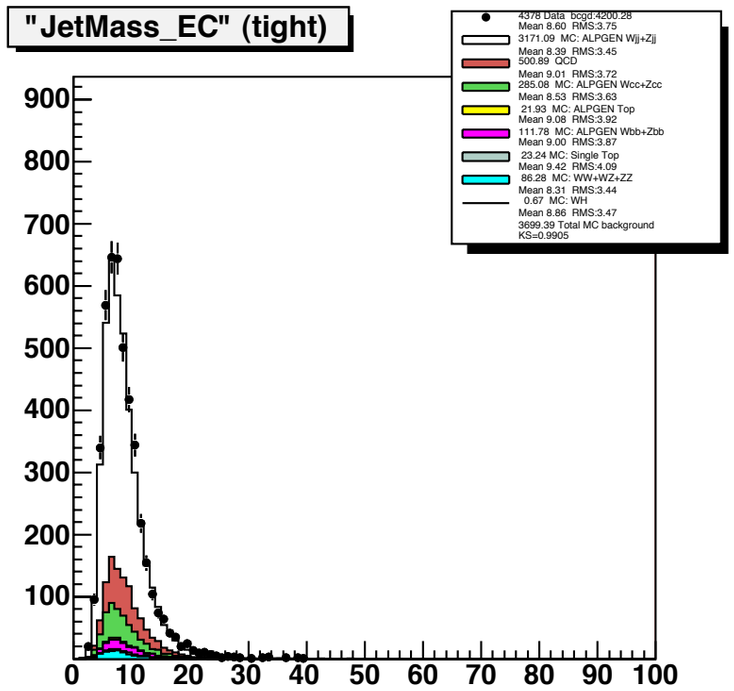
pT<40

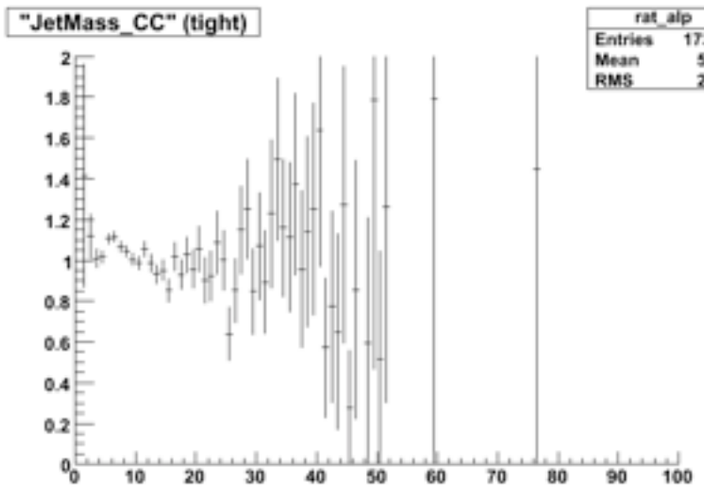
pT>40

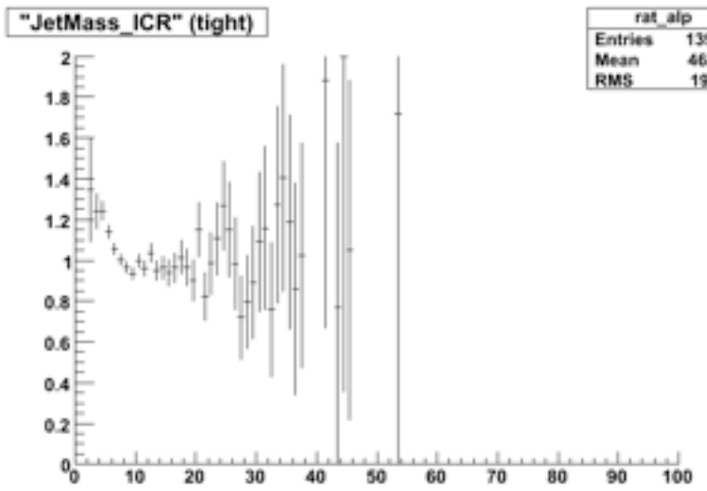
Jet Mass - eta bins

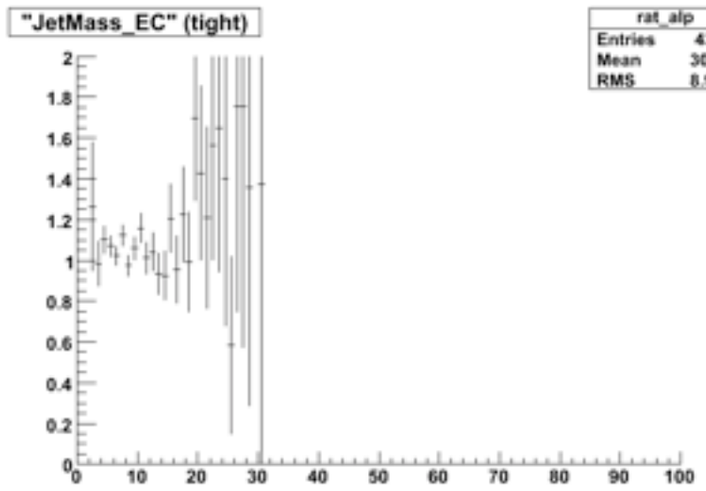






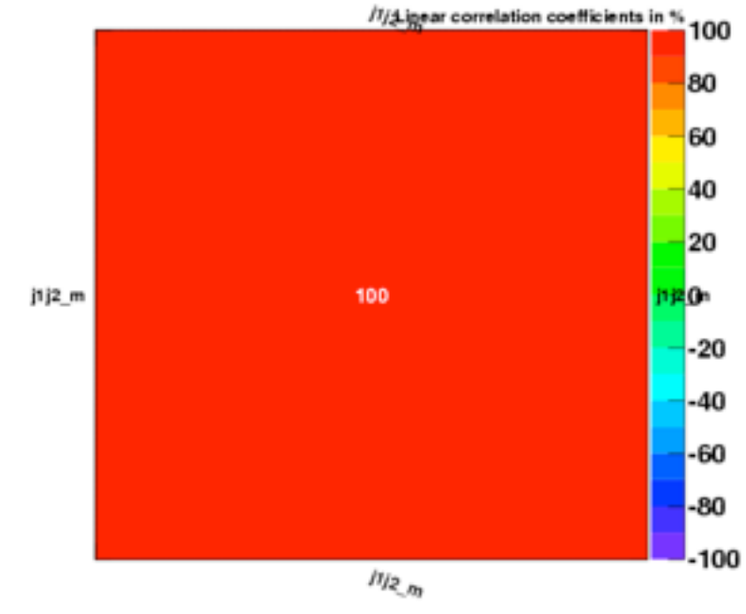


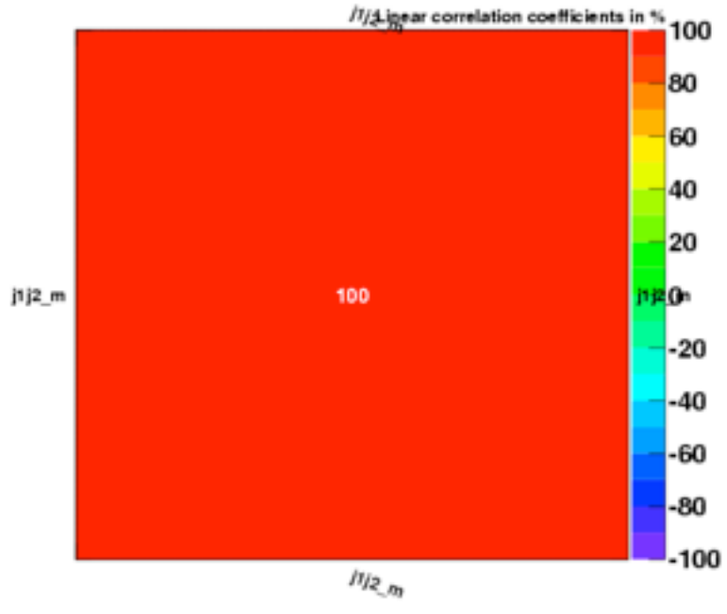


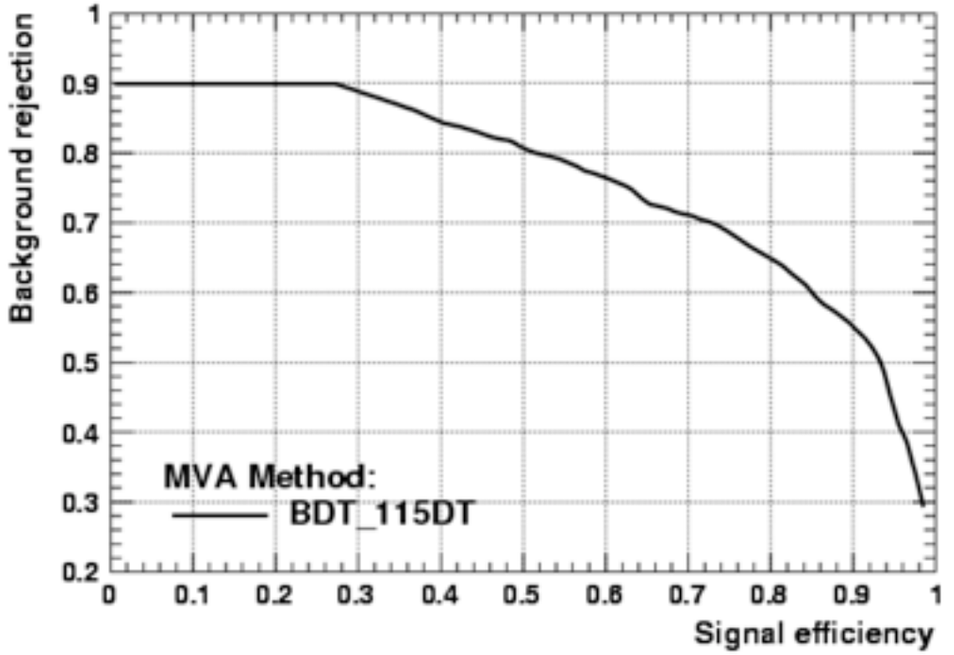


CC ICR EC

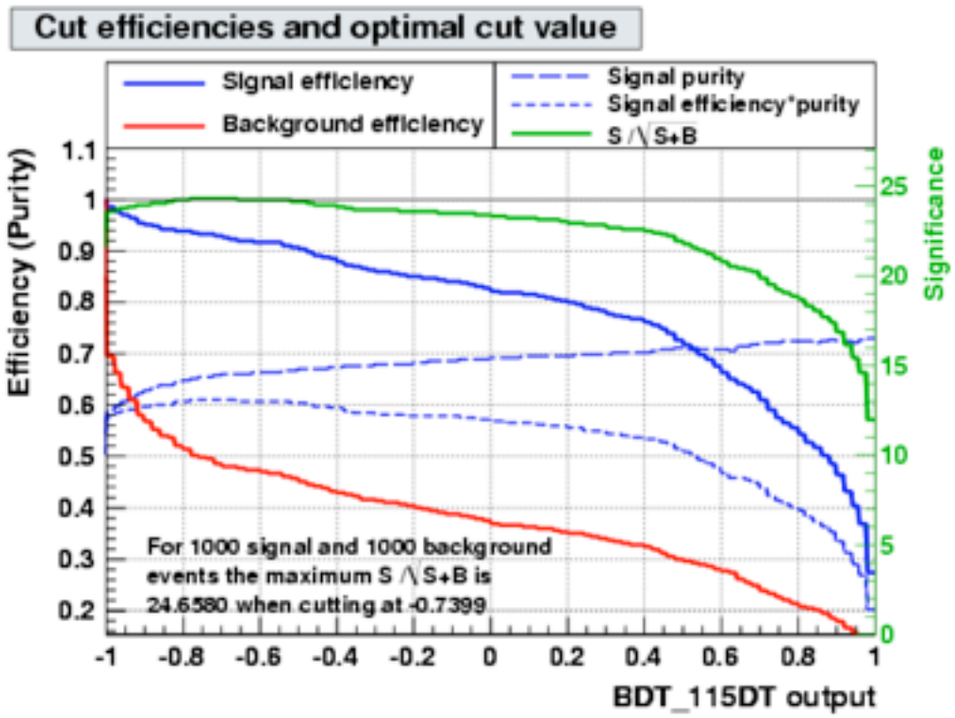
Control plots BDT 1 variable

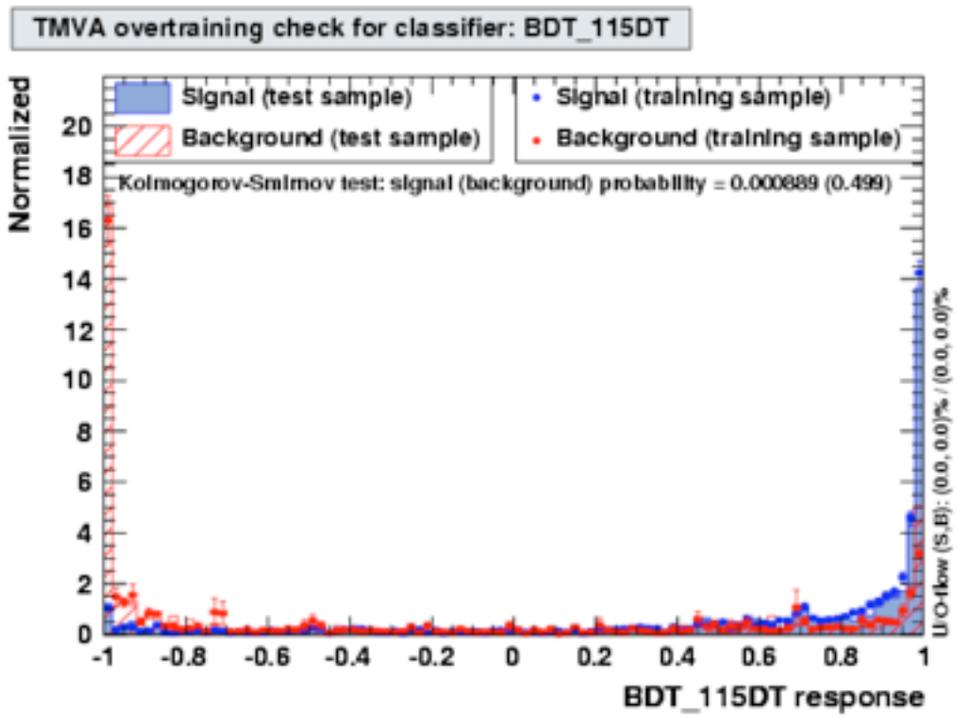




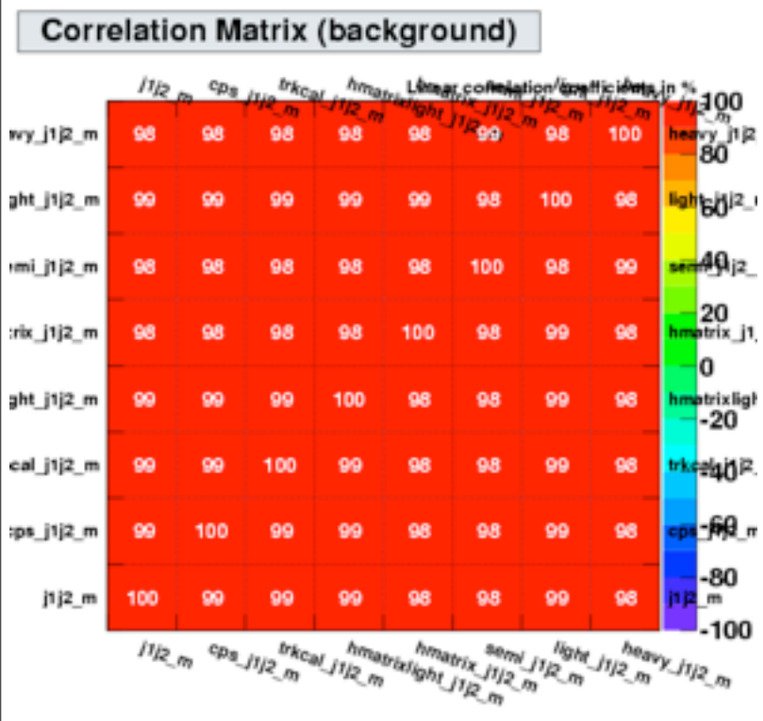


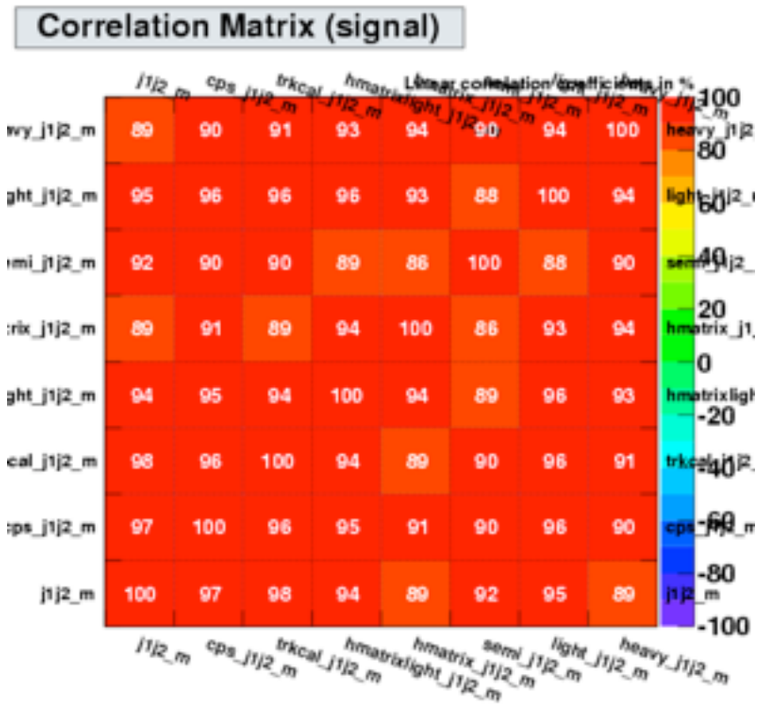
Area : 0.760 - Separation : 0.281 - Significance : 0.875

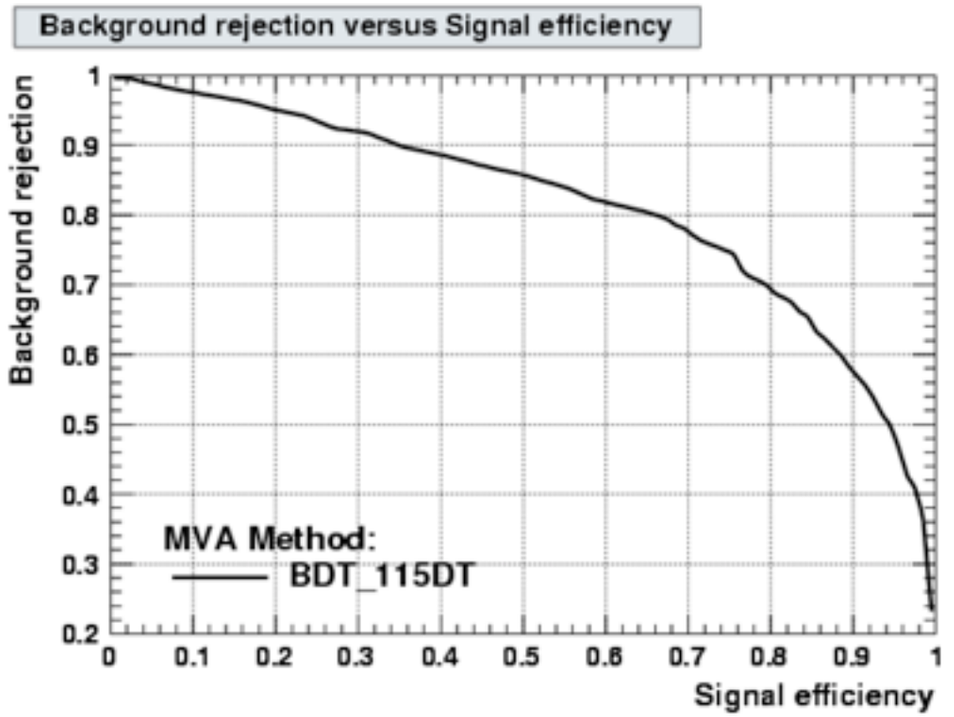




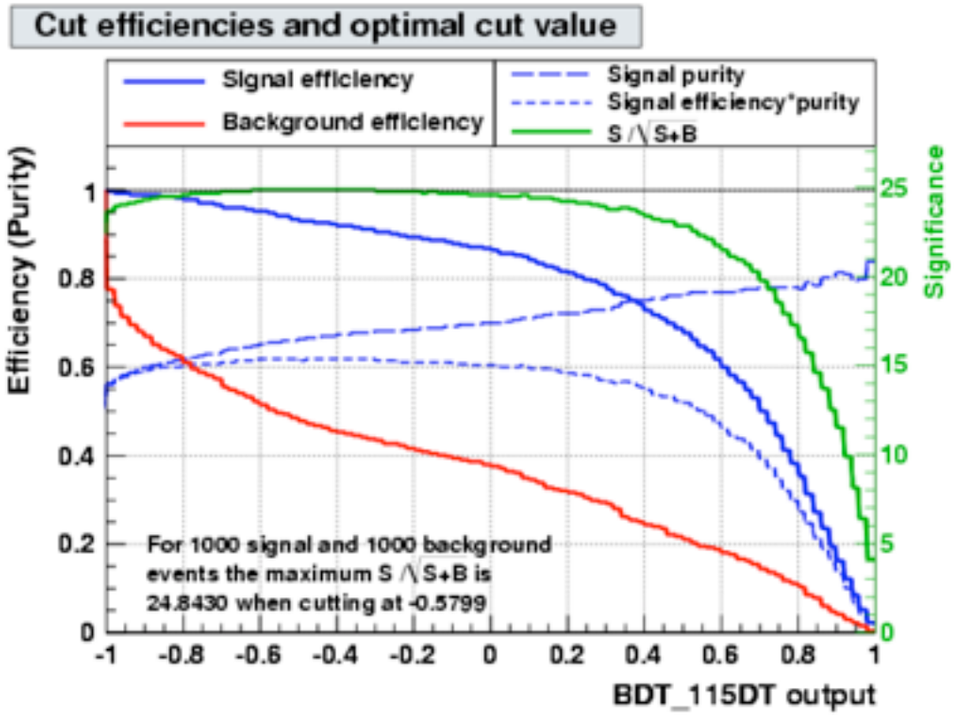
Control plots BDT 1+7 JER variables

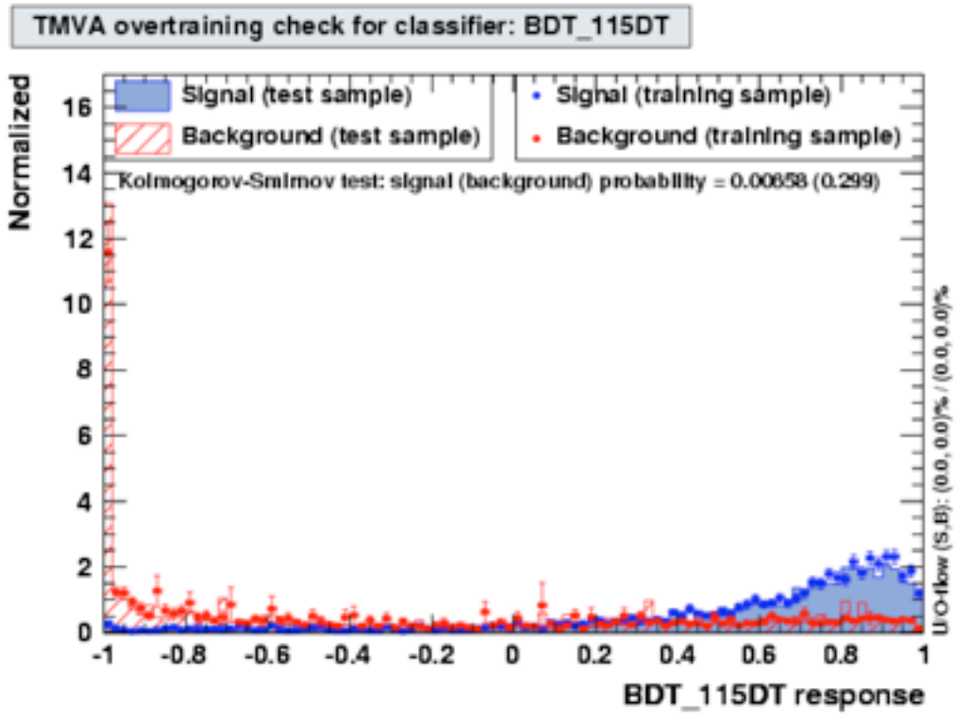




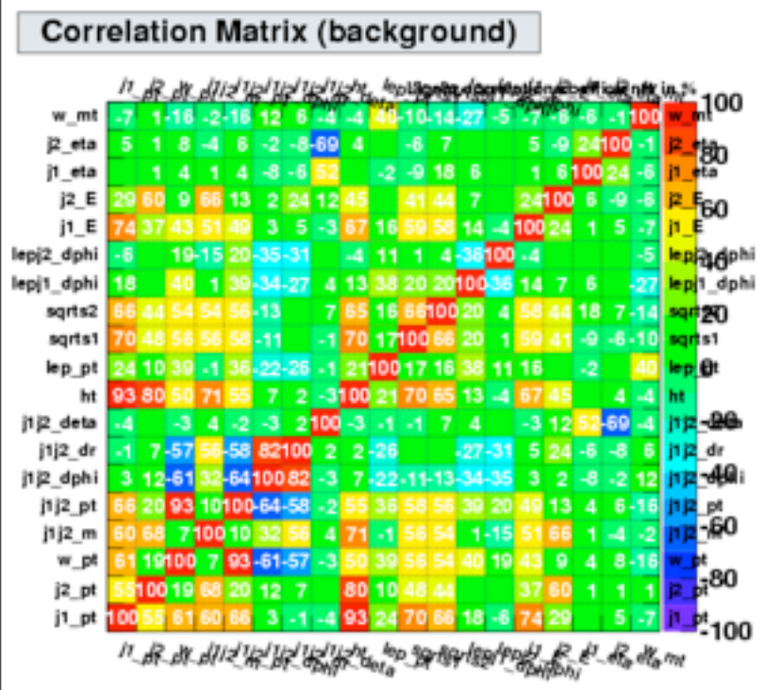


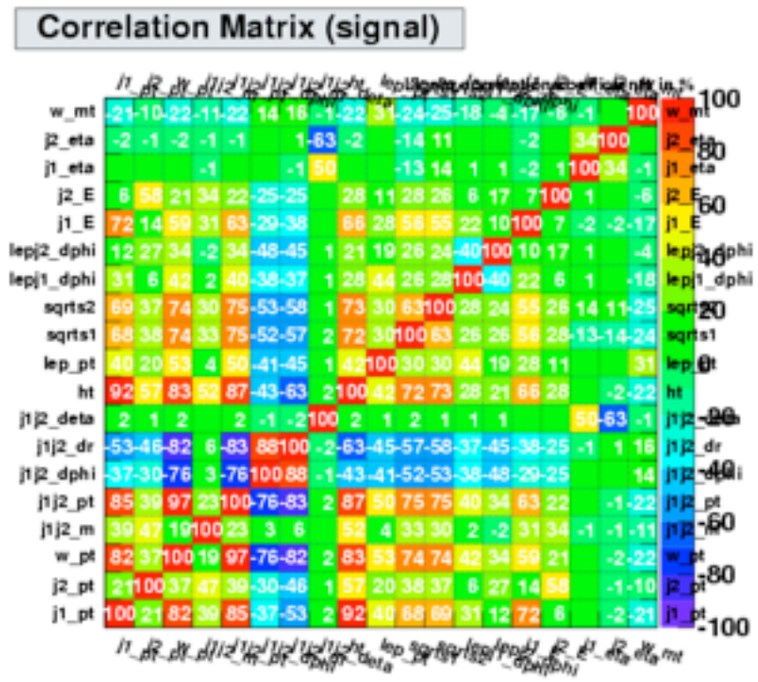
Area : 0.813 - Separation : 0.342 - Significance : 0.981

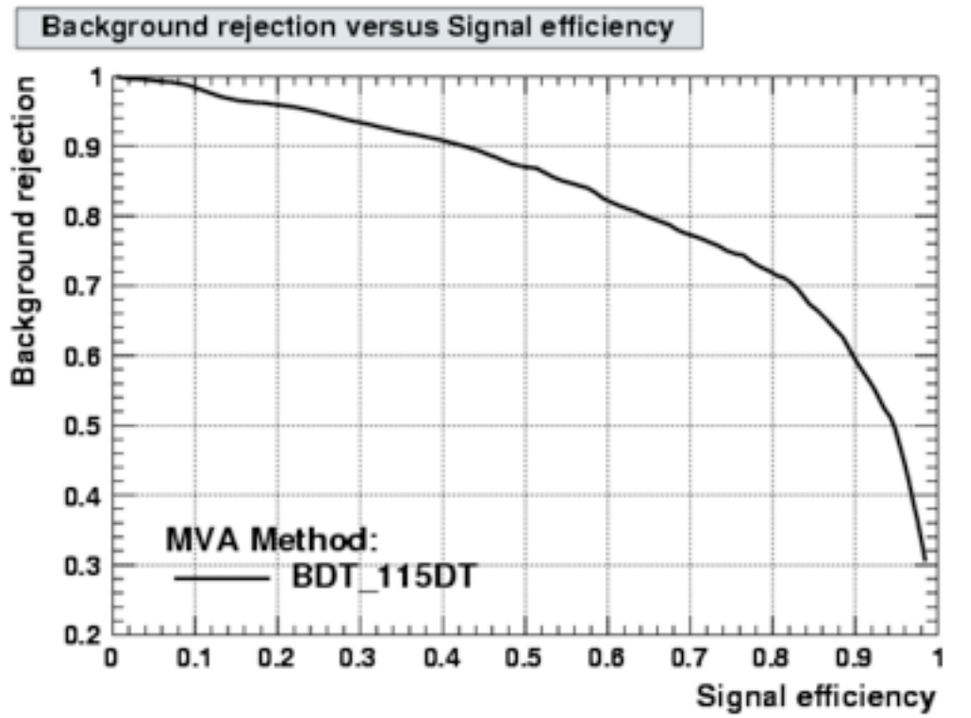




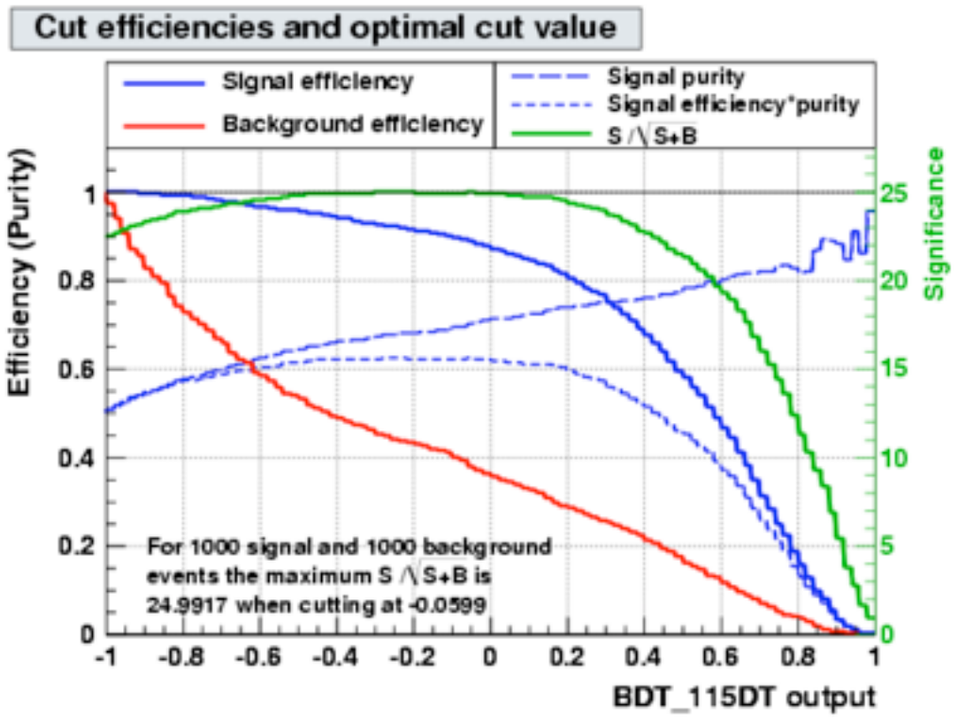
Control plots nominal BDT 19 variables

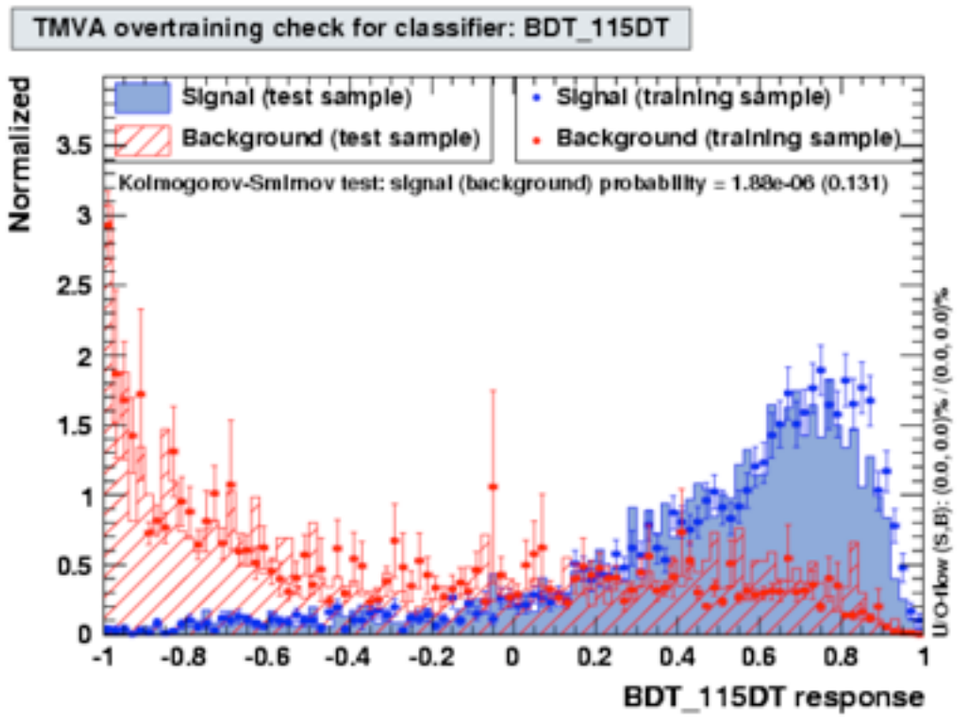






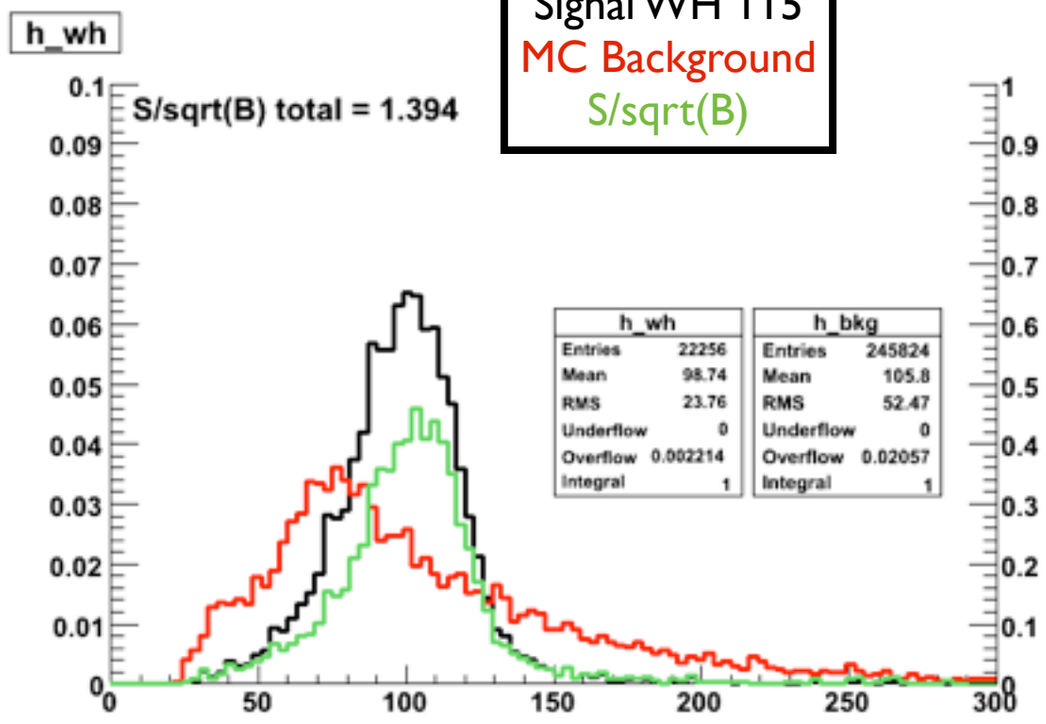
Area : 0.823 - Separation : 0.352 - Significance : 1.078

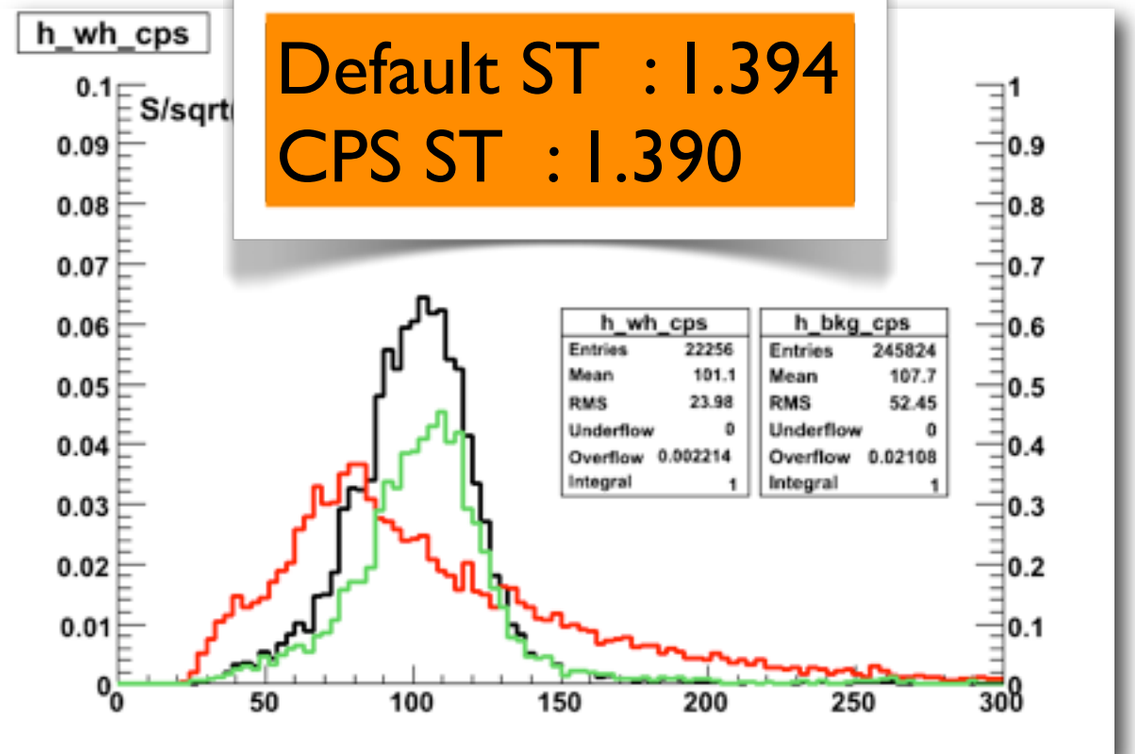




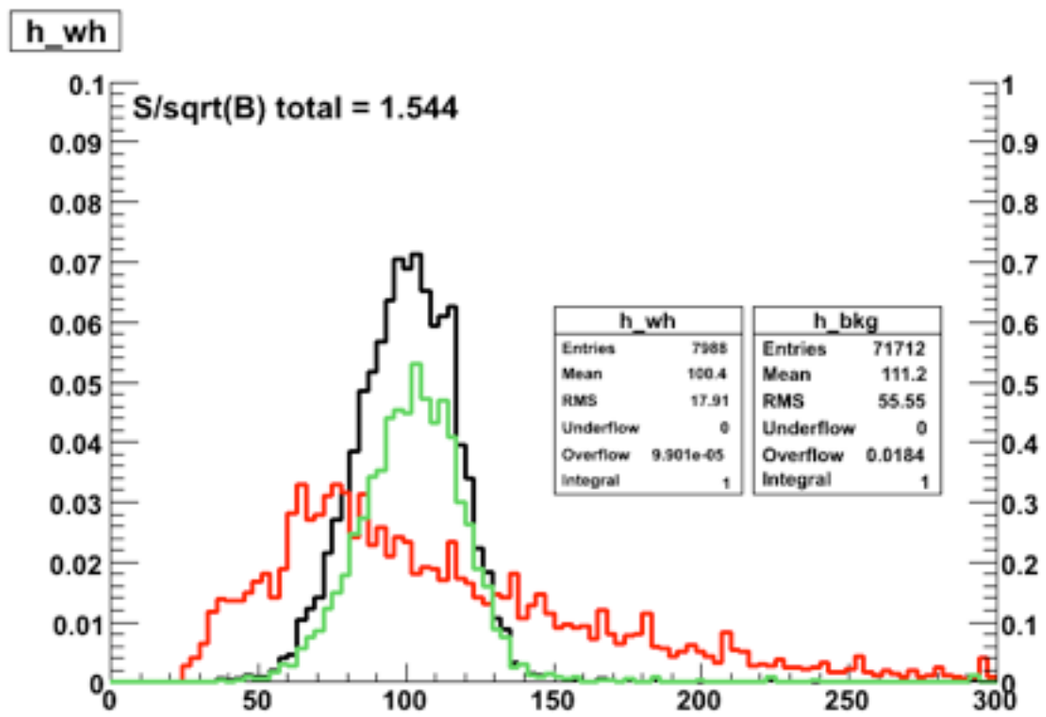
CPS Mjj

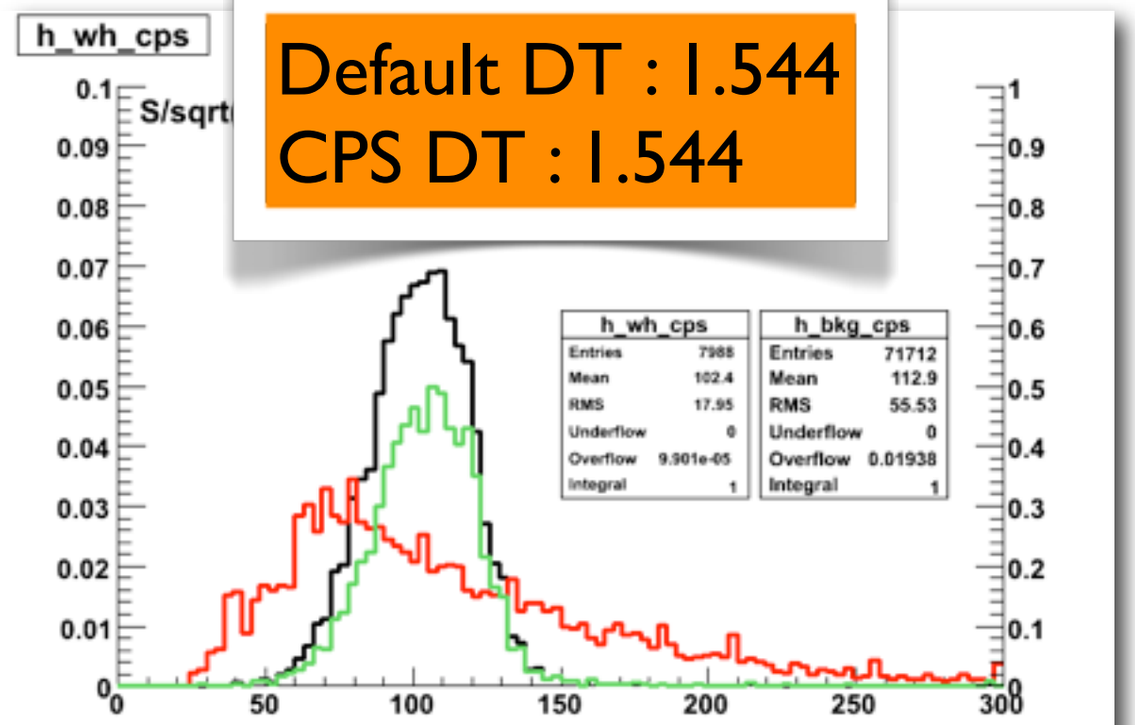
ST


h_wh
S/sqrt(B) total = 1.394
Signal WH 115
MC Background
S/sqrt(B)
h_wh
Entries 22256
Mean 98.74
RMS 23.76
Underflow 0
Overflow 0.002214
Integral 1
h_bkg
Entries 245824
Mean 105.8
RMS 52.47
Underflow 0
Overflow 0.02057
Integral 1


h_wh_cps
Default ST : 1.394
CPS ST : 1.390
h_wh_cps
Entries 22256
Mean 101.1
RMS 23.98
Underflow 0
Overflow 0.002214
Integral 1
h_bkg_cps
Entries 245824
Mean 107.7
RMS 52.45
Underflow 0
Overflow 0.02108
Integral 1

DT

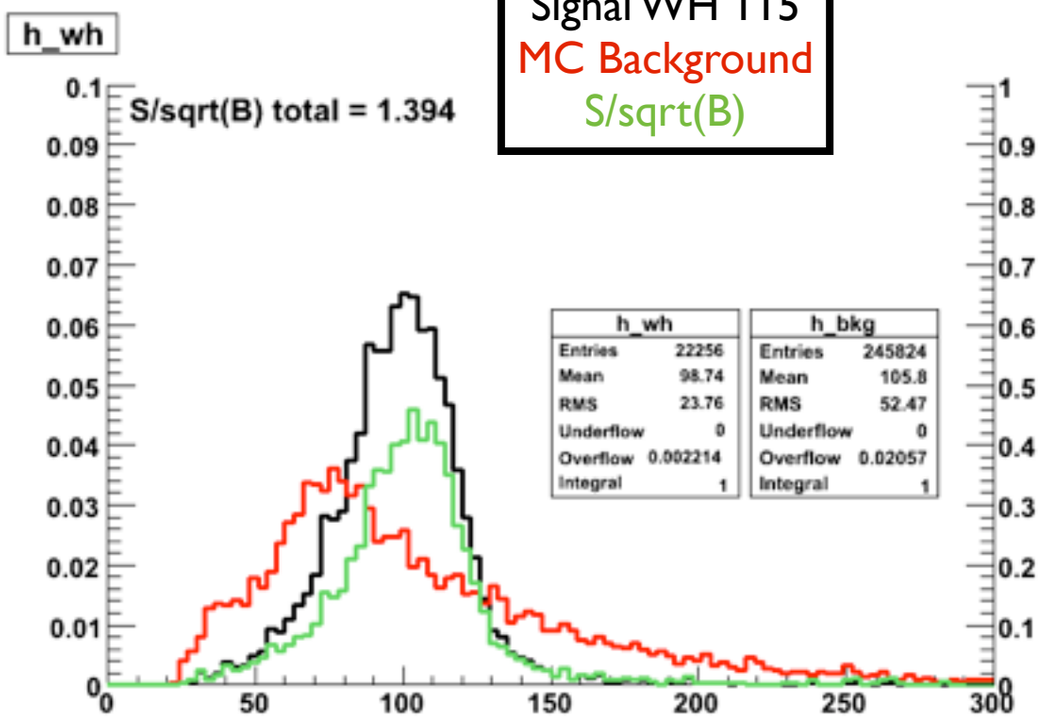

h_wh
S/sqrt(B) total = 1.544
h_wh
Entries 7988
Mean 100.4
RMS 17.91
Underflow 0
Overflow 9.901e-05
Integral 1
h_bkg
Entries 71712
Mean 111.2
RMS 55.55
Underflow 0
Overflow 0.0184
Integral 1

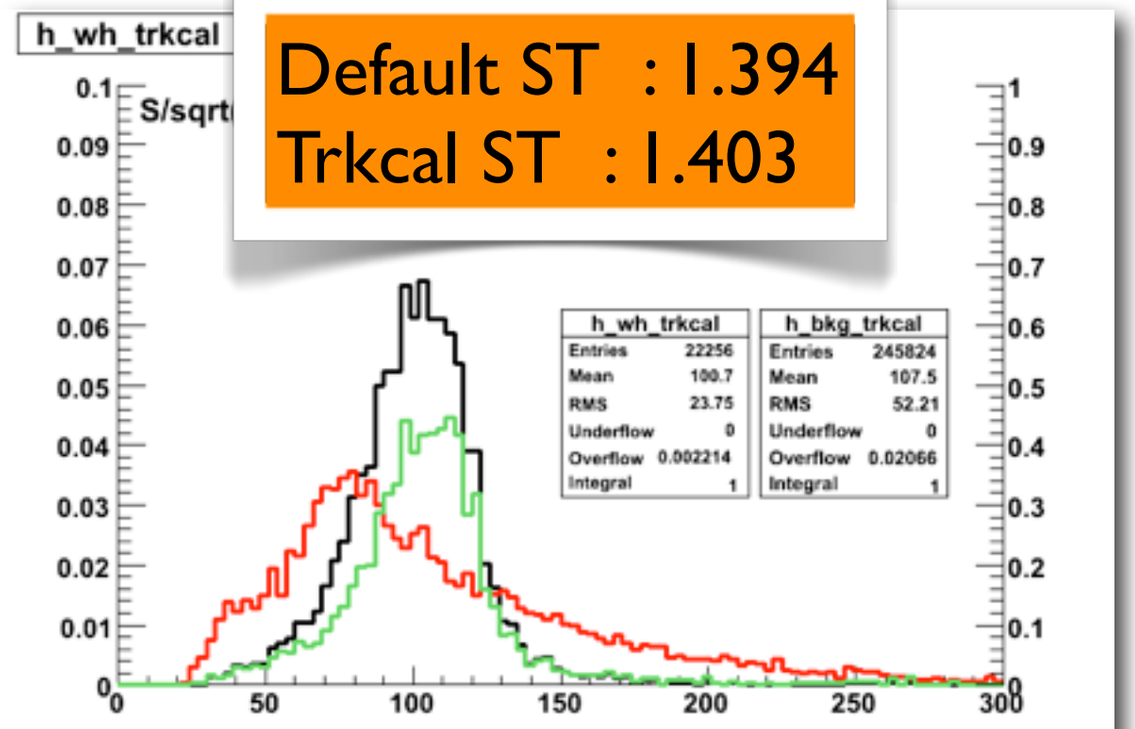

h_wh_cps
Default DT : 1.544
CPS DT : 1.544
h_wh_cps
Entries 7988
Mean 102.4
RMS 17.95
Underflow 0
Overflow 9.901e-05
Integral 1
h_bkg_cps
Entries 71712
Mean 112.9
RMS 55.53
Underflow 0
Overflow 0.01938
Integral 1

Trkcal Mjj

Signal WH 115
MC Background
S/sqrt(B)

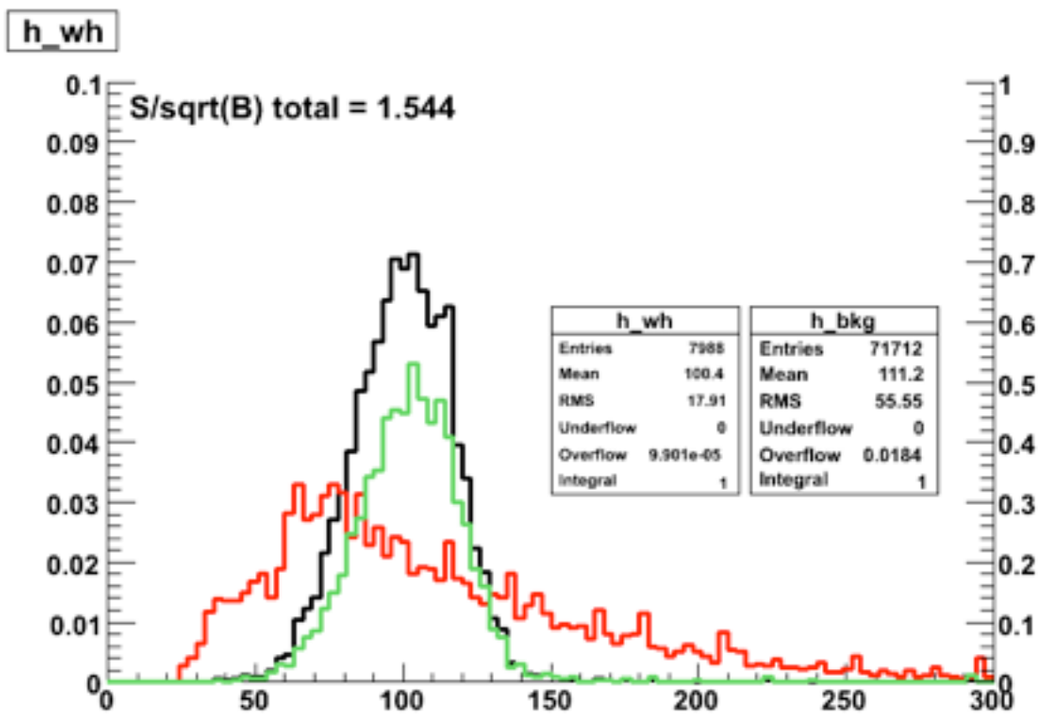
ST

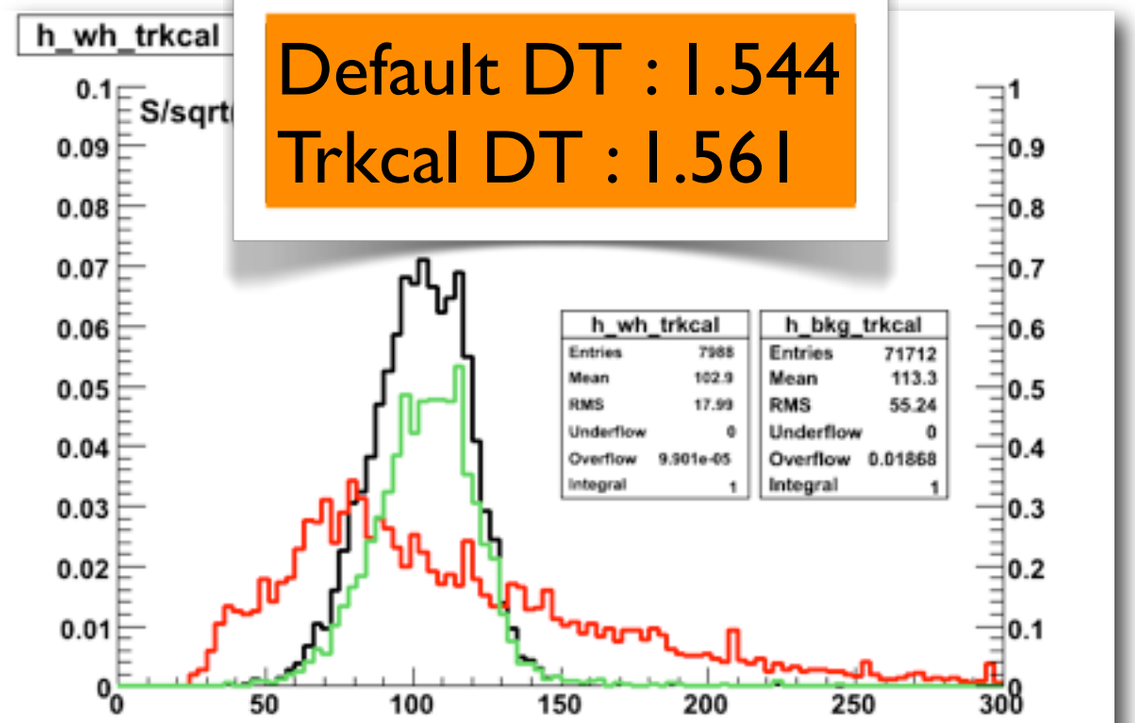




Default ST : 1.394
Trkcal ST : 1.403

DT

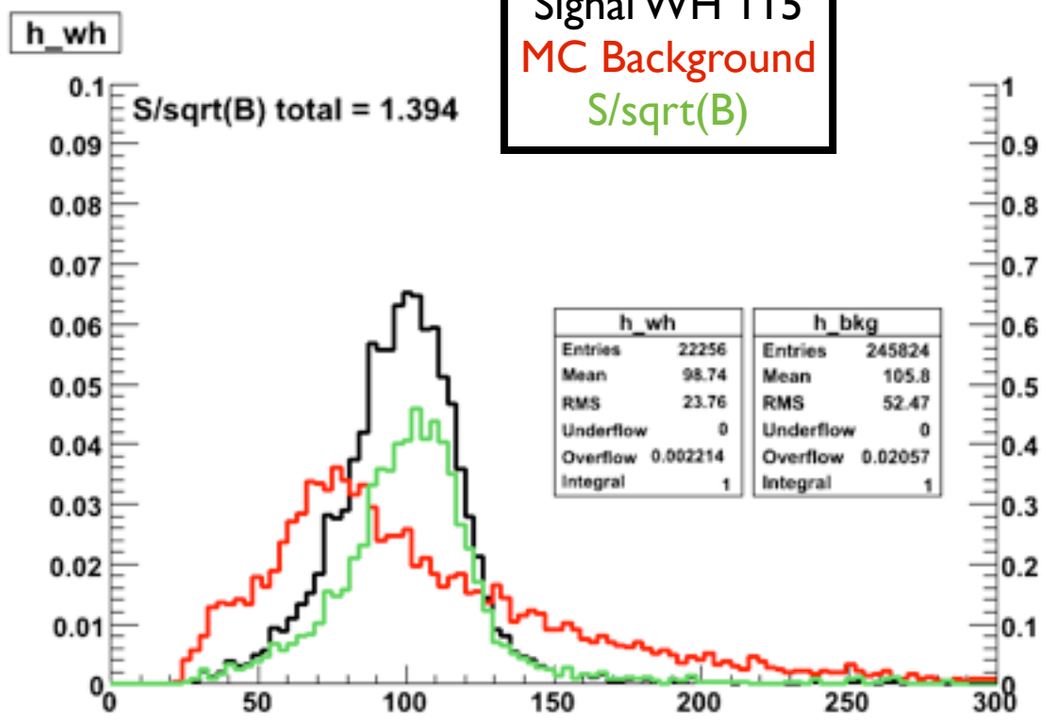


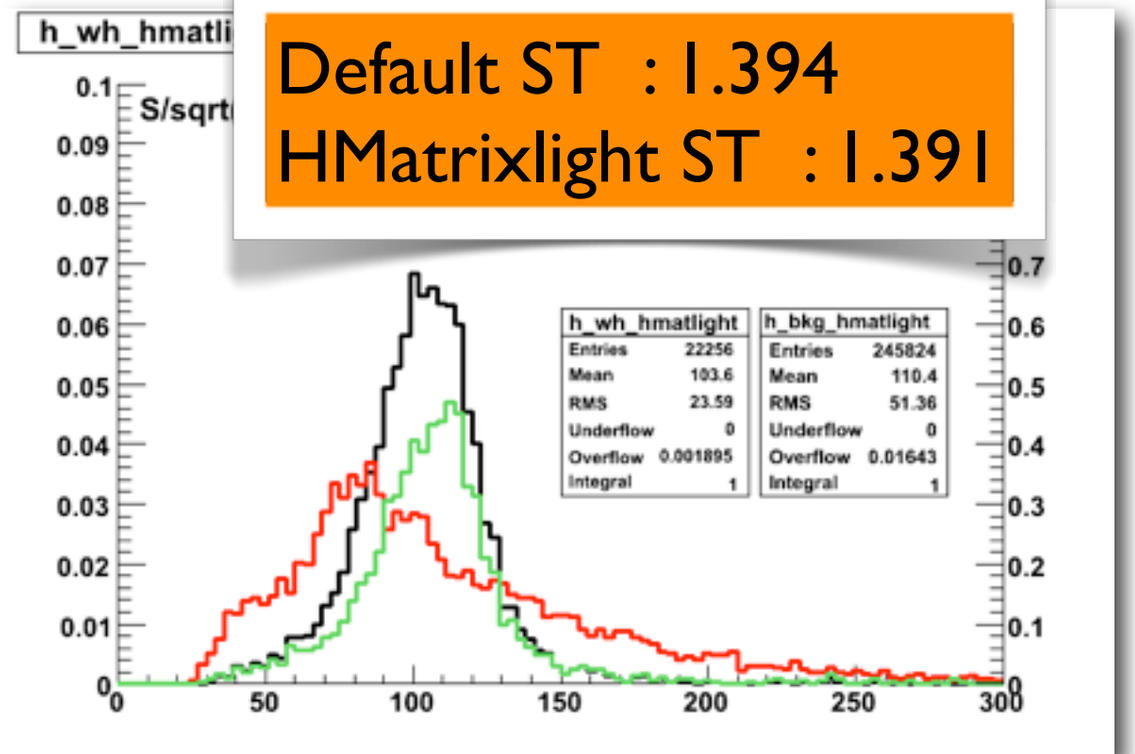


Default DT : 1.544
Trkcal DT : 1.561

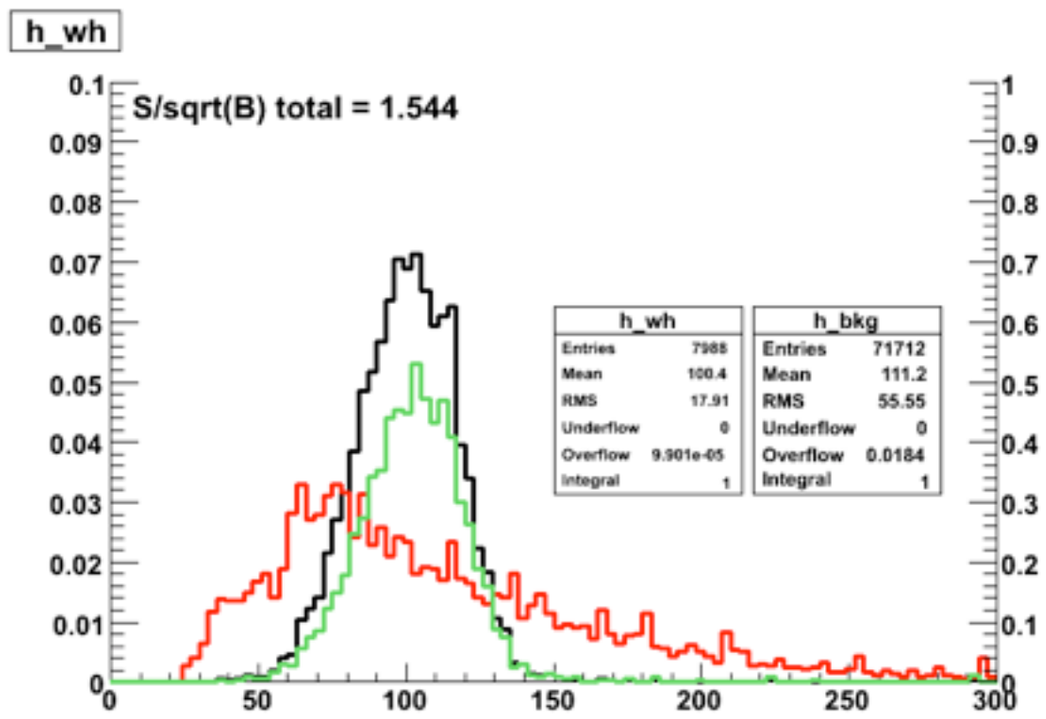
HMatrixlight Mjj

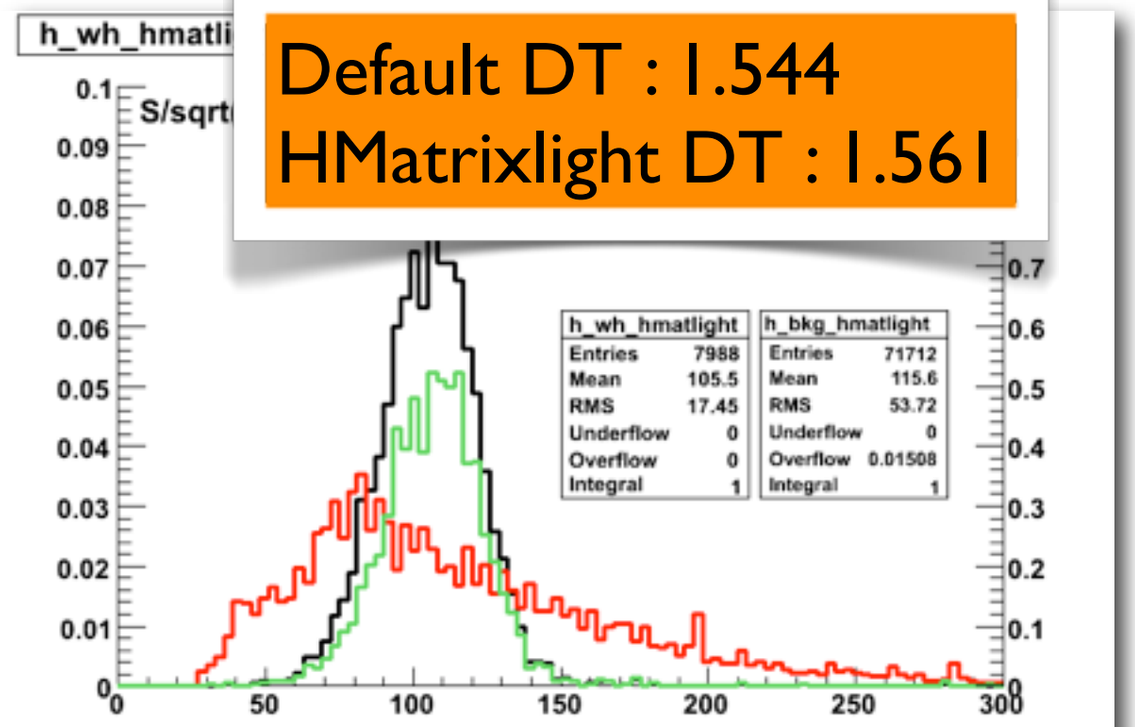
ST





DT

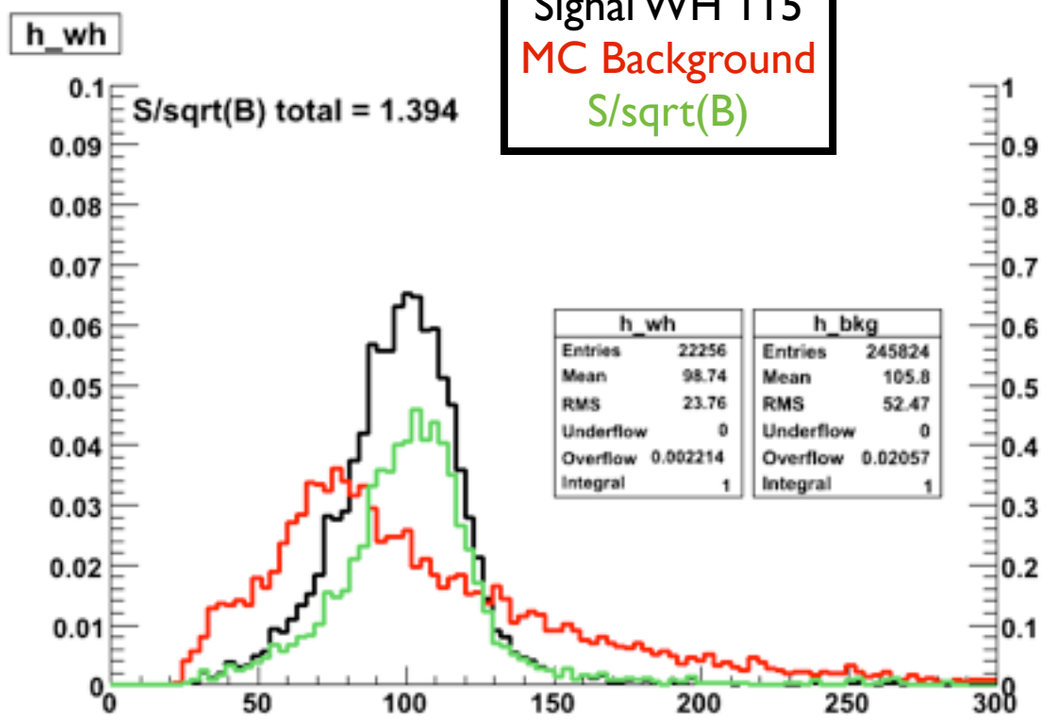


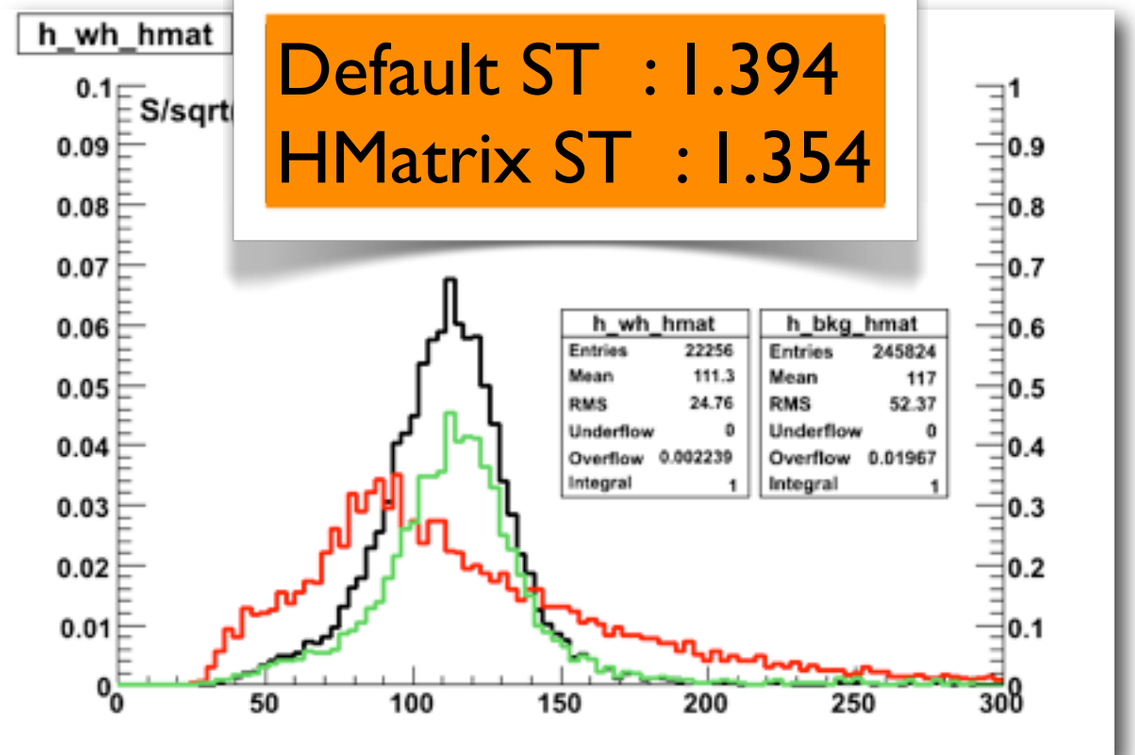


HMatrix Mjj

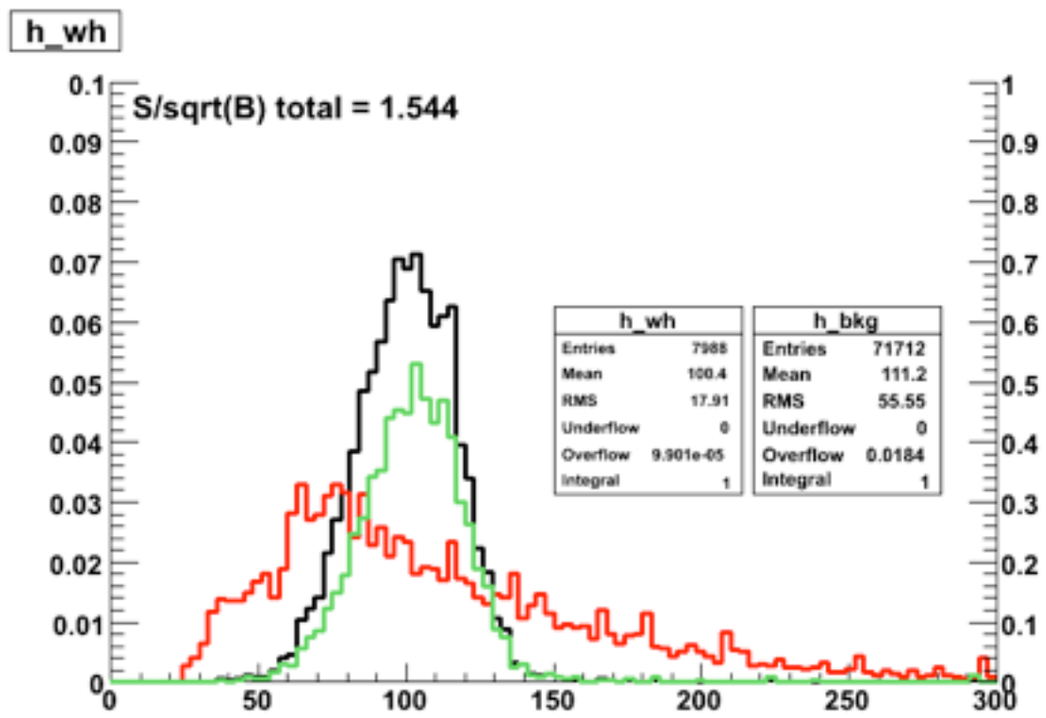
Signal WH 115
MC Background
S/sqrt(B)

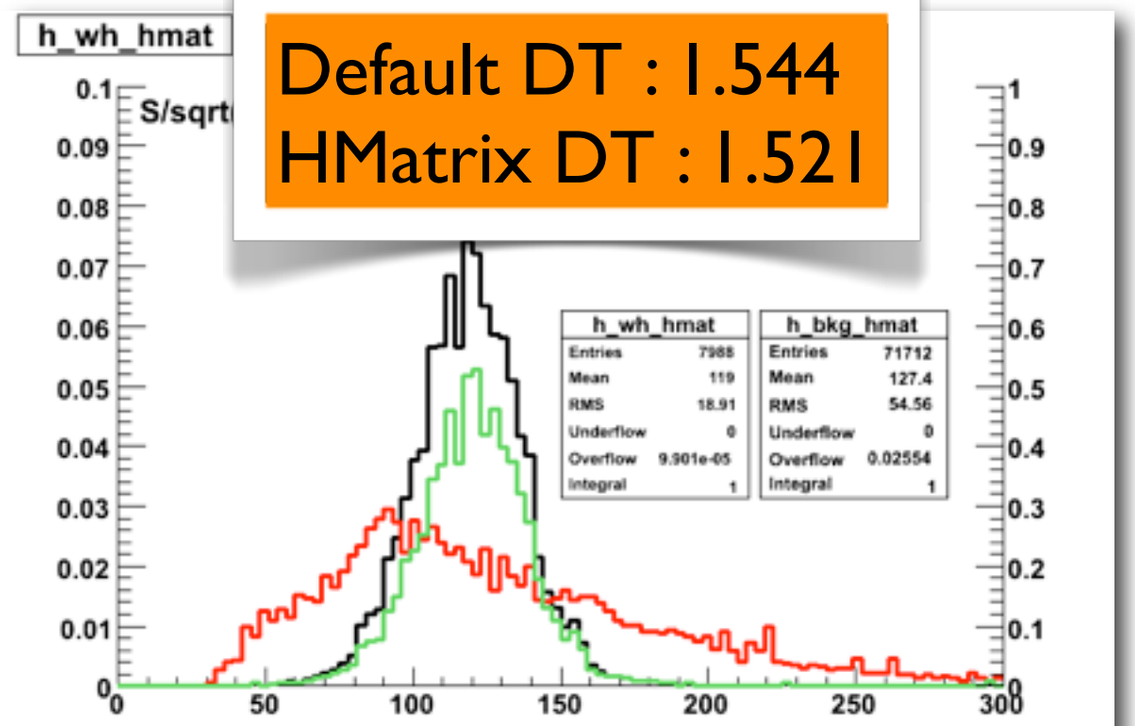
ST





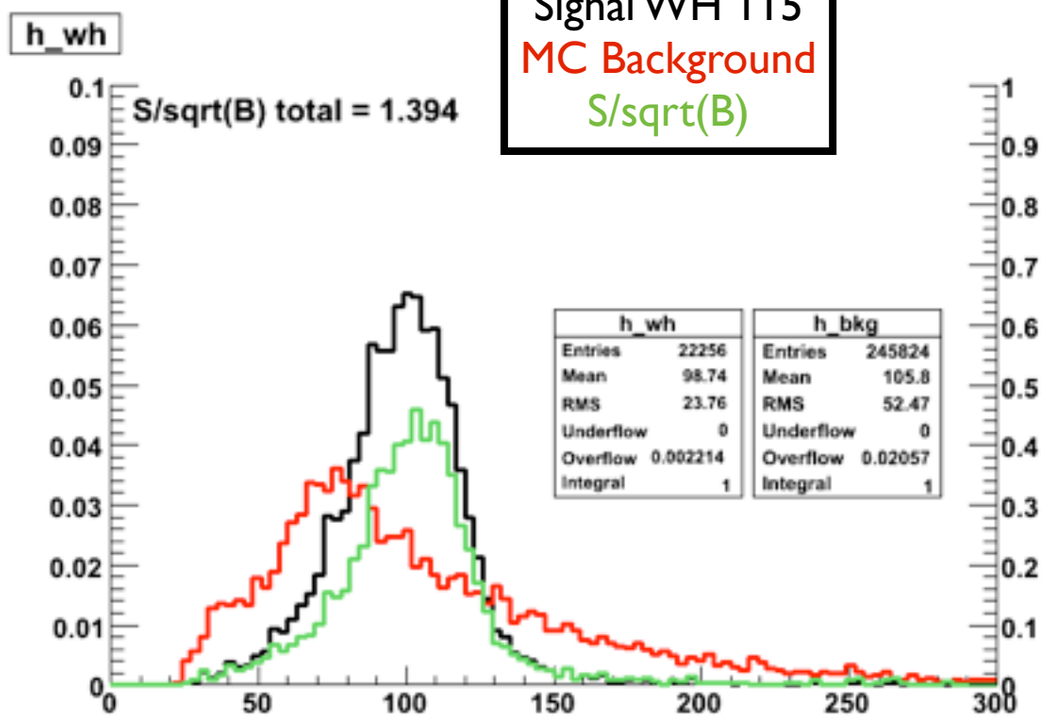
DT

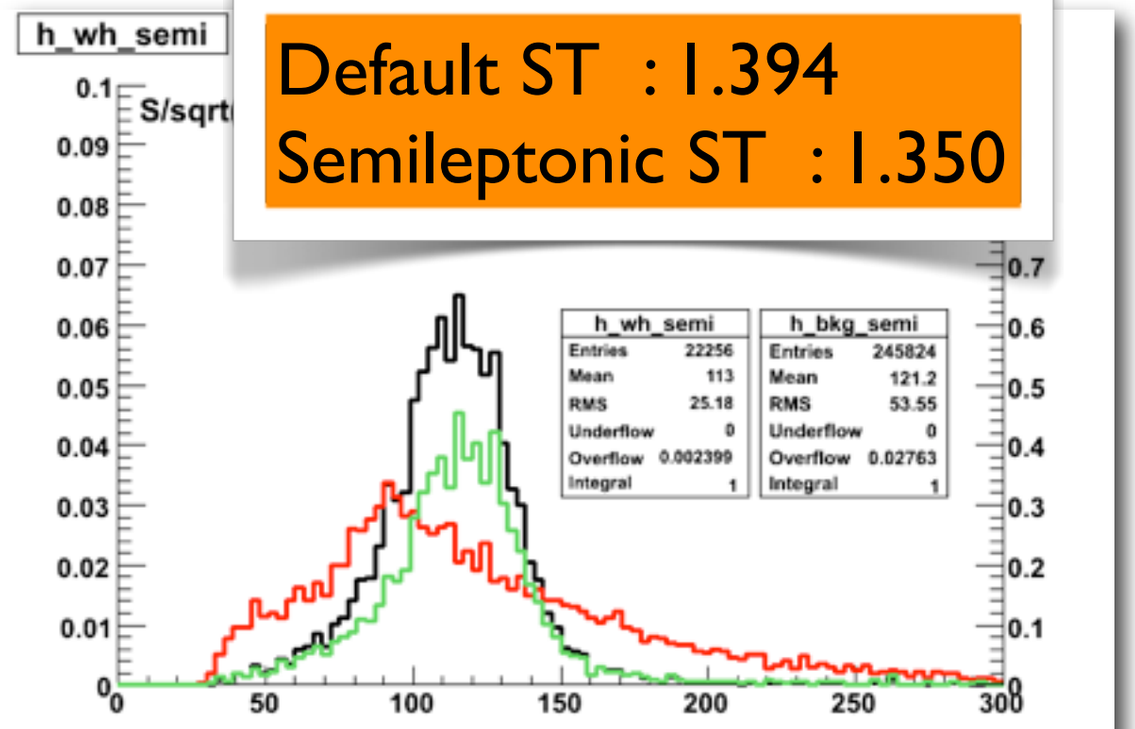




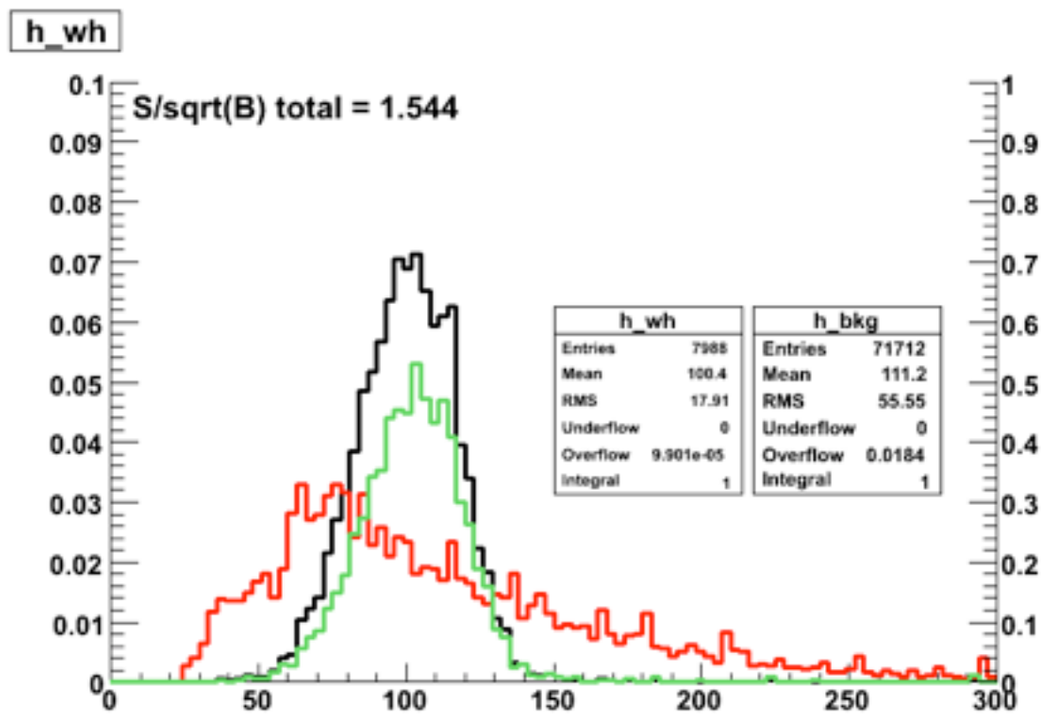
Semileptonic Mjj

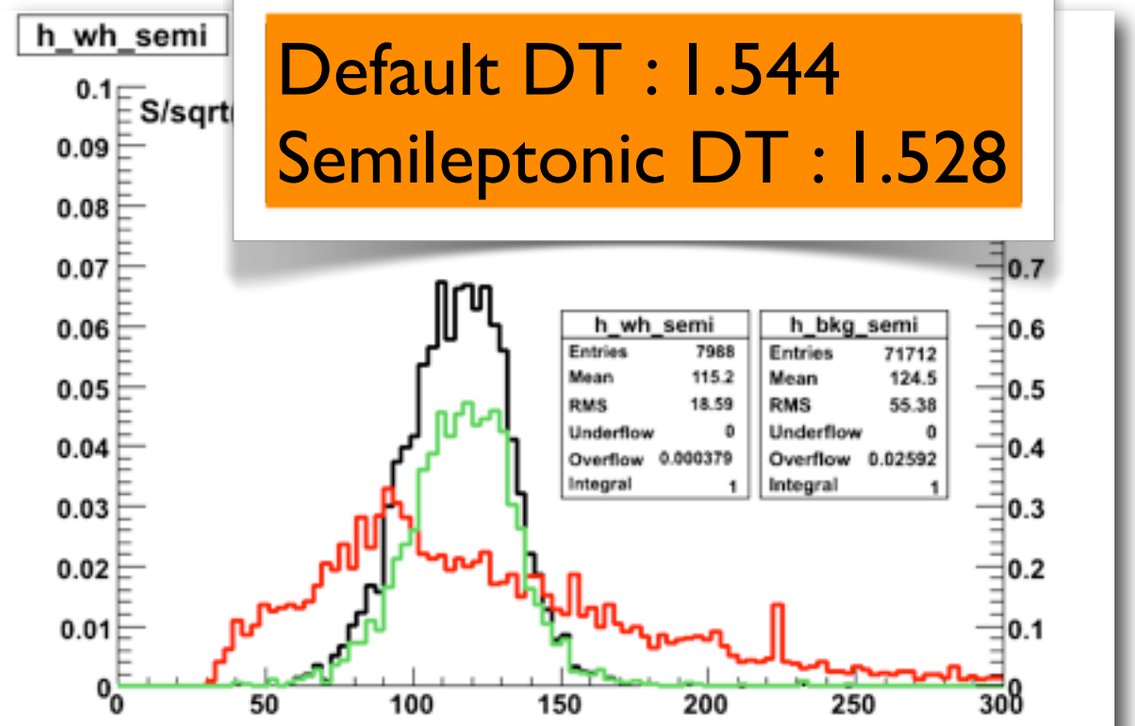
ST





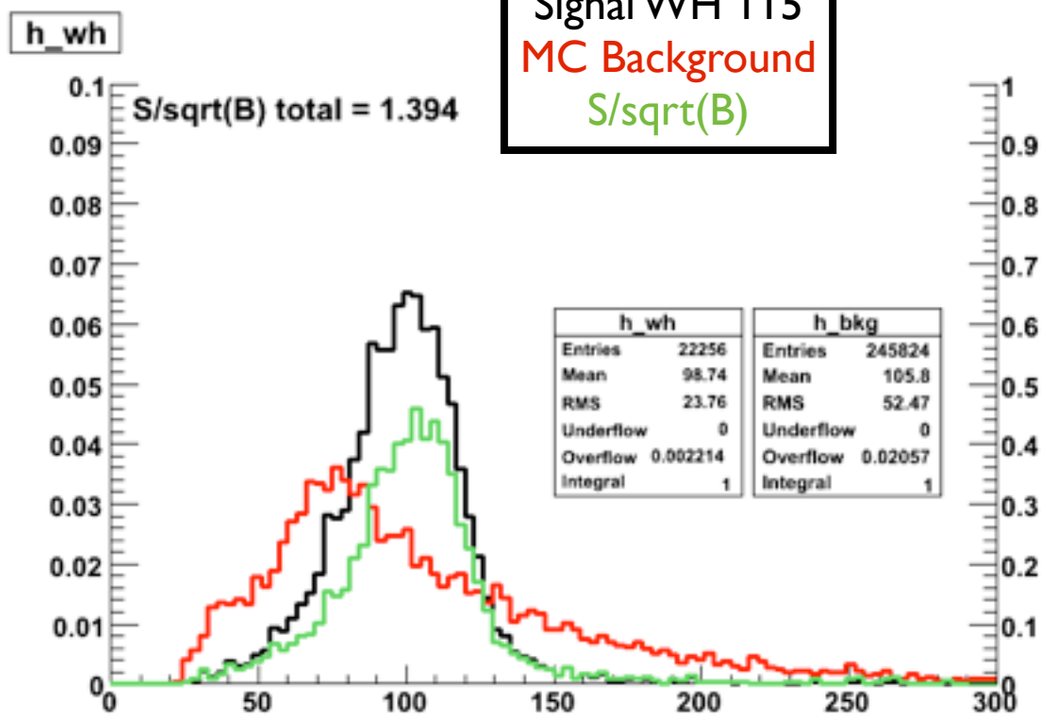
DT

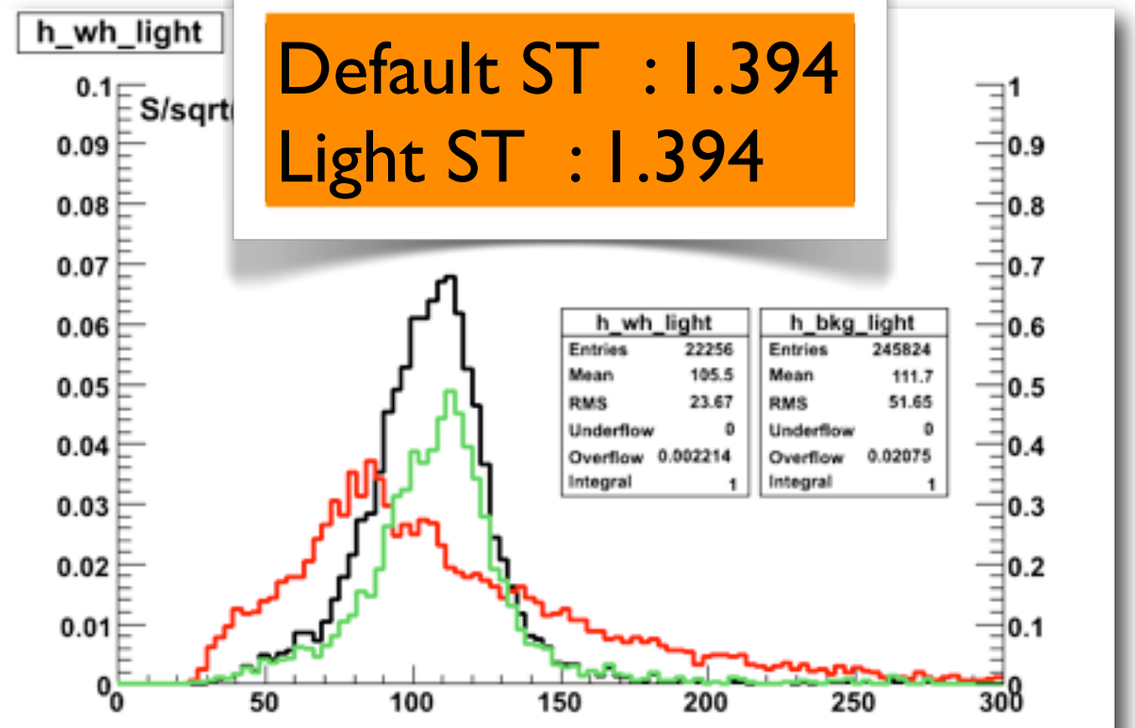




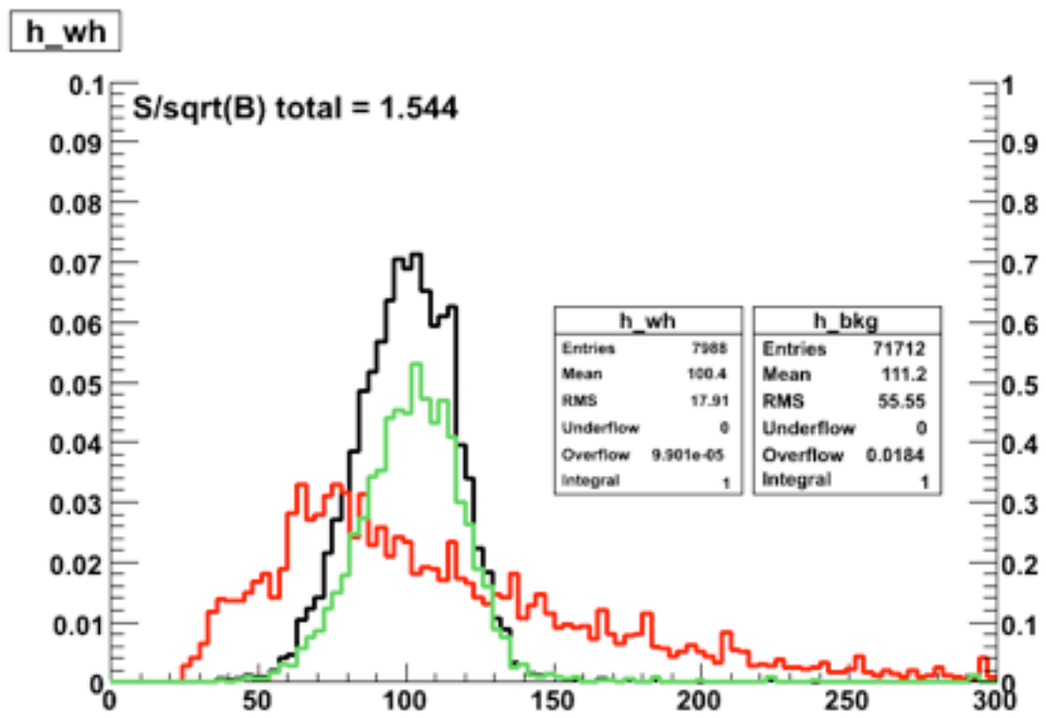
Light Mjj

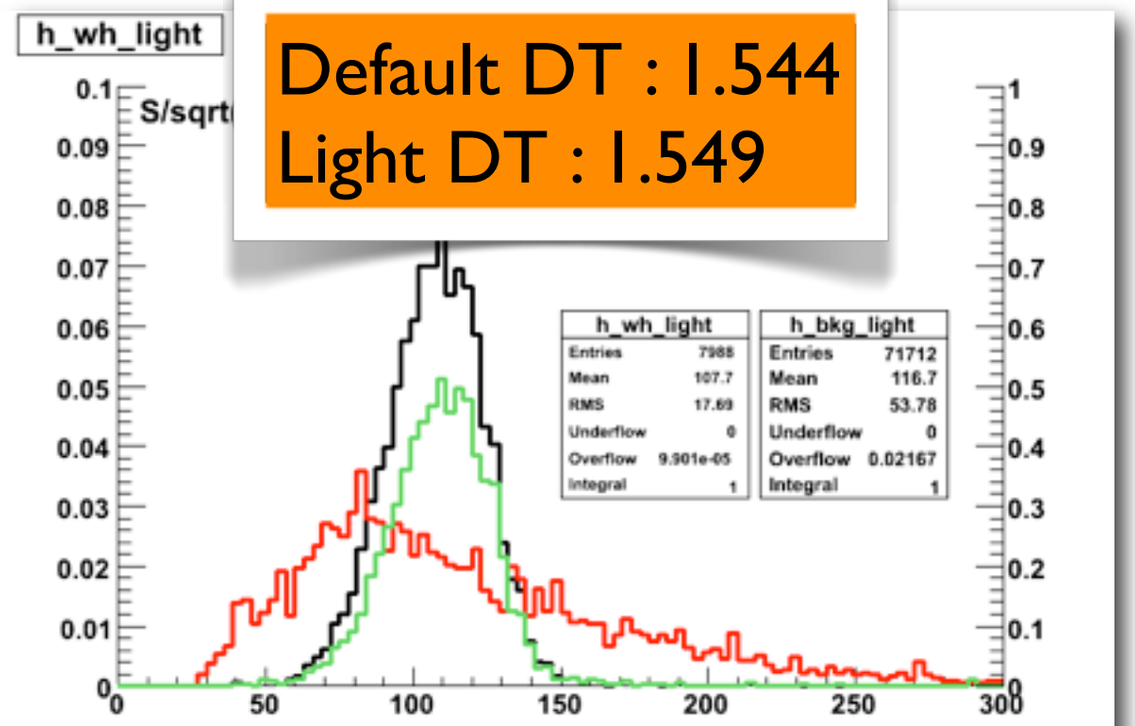
ST





DT

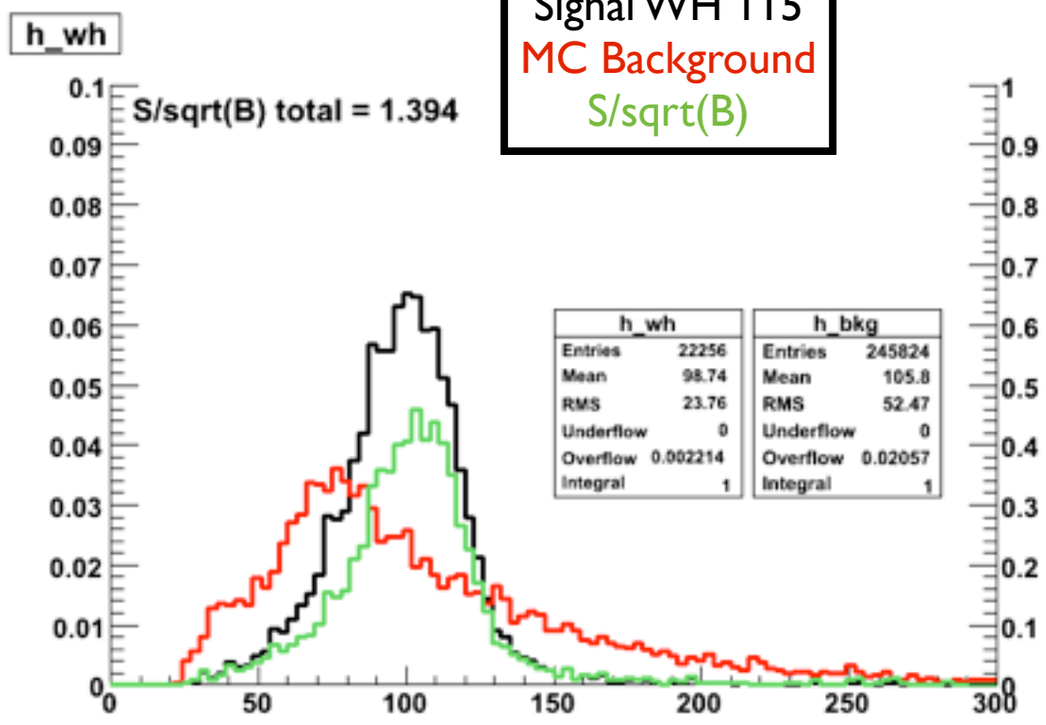


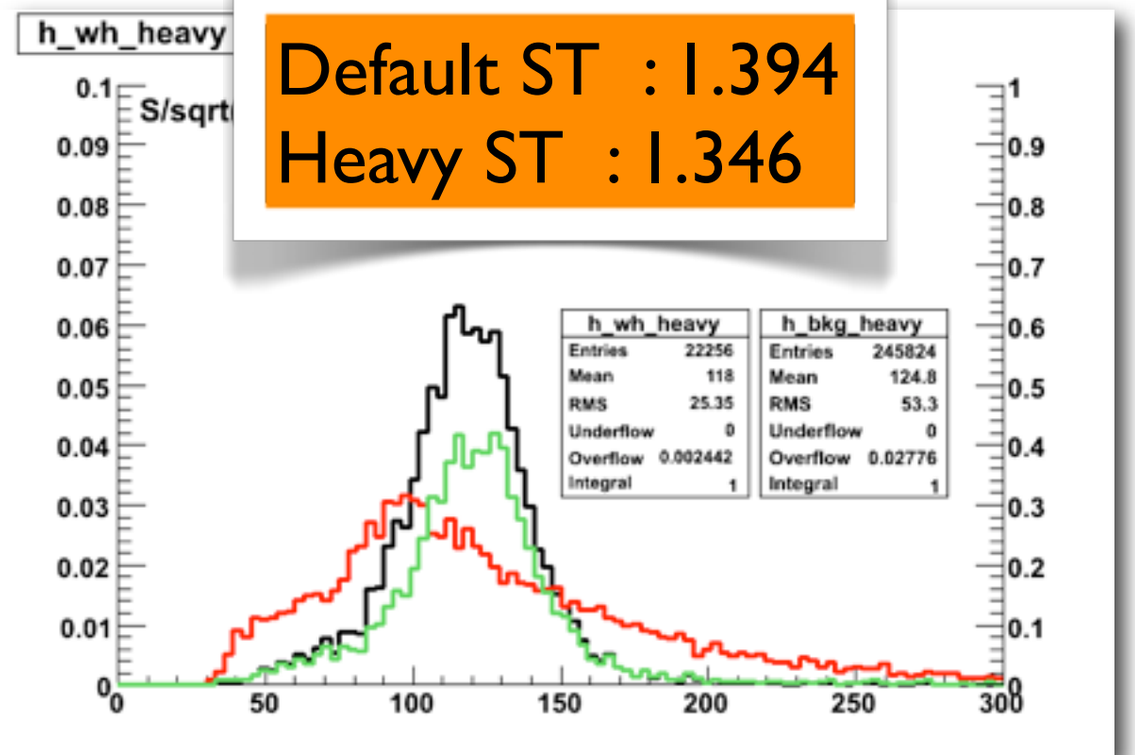


Heavy Mjj

Signal WH 115
MC Background
S/sqrt(B)

ST





DT