NNCuts on LRHWW, Mll and MetSig

Richard St. Denis University of Glasgow/CDF

10 December 2007

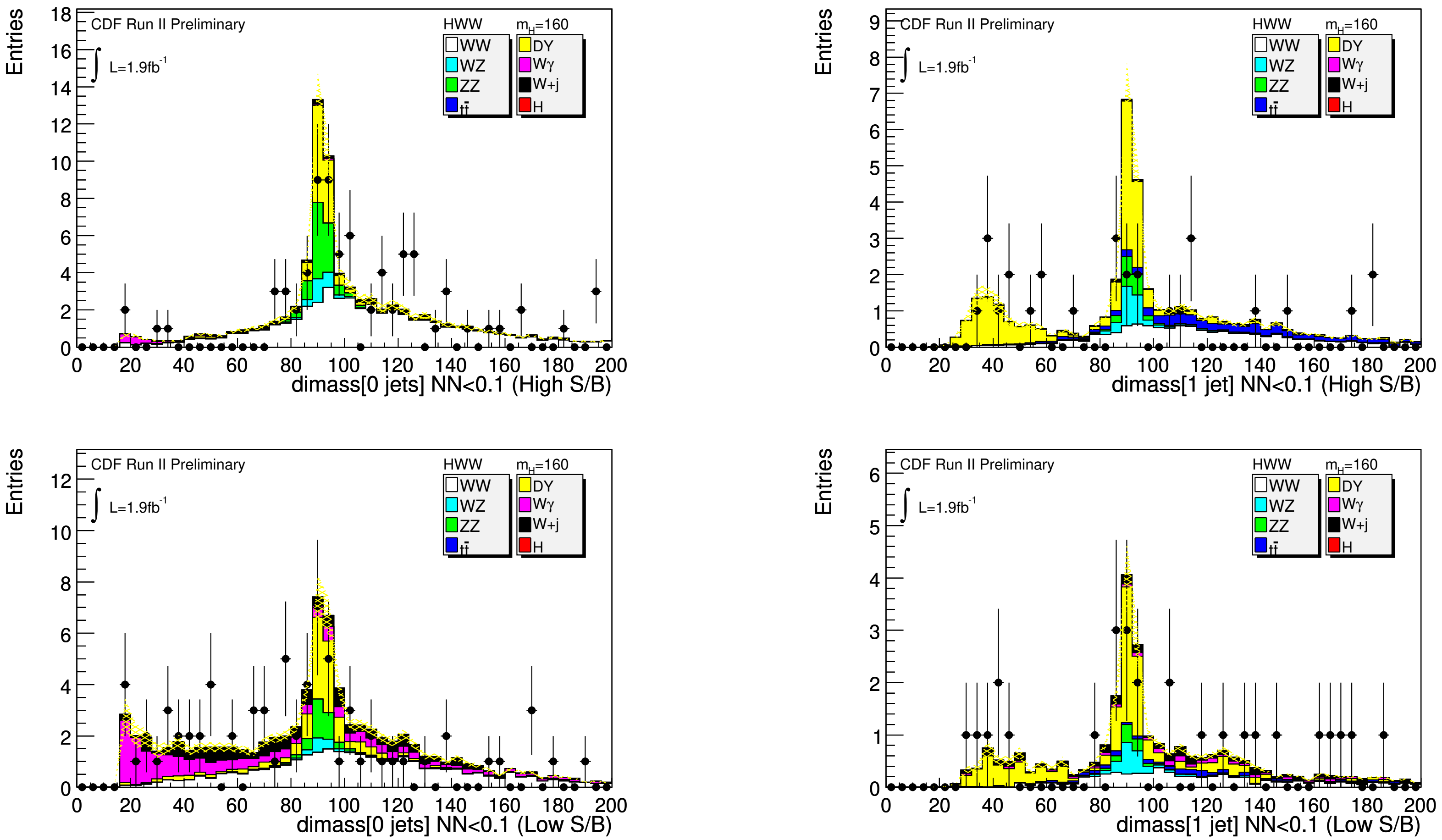
Outline

- The 0.1 cut on NN for Mll, MetSig and LRHWW
- Elliot's test of MetSig, Met and Fakes
- Discussion of control regions

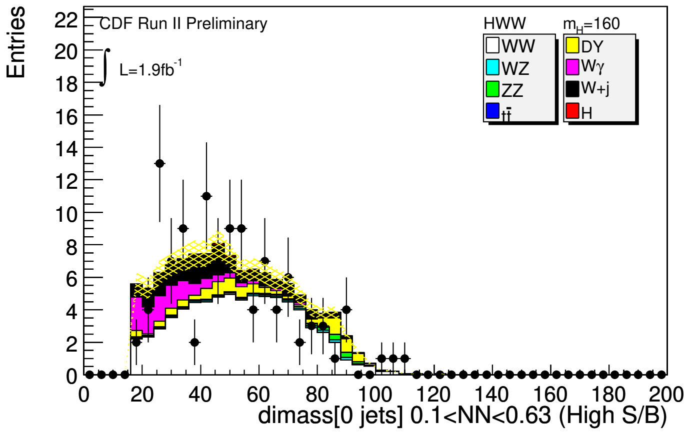
The 0.1 cut on NN

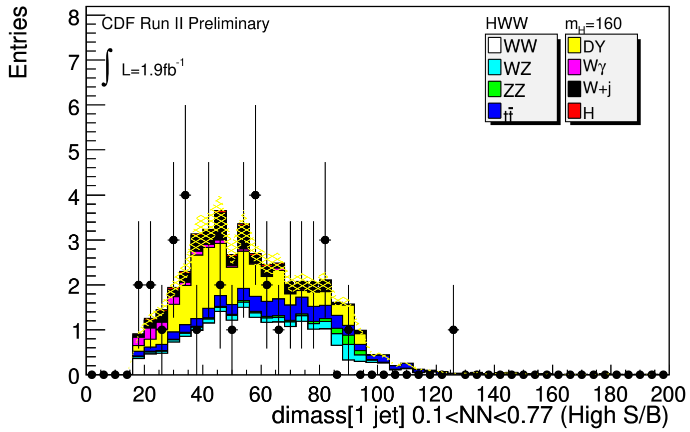
- Have cut up the regions into NN< 0.1, 0.1<NN<80% Cut, and $NN >$ 80% Cut
- Will show that these are the ZZ/WZ, control and signal regions
- Here only for dimass, MetSig and LHRHWW. All variables available as well in a separate posting.

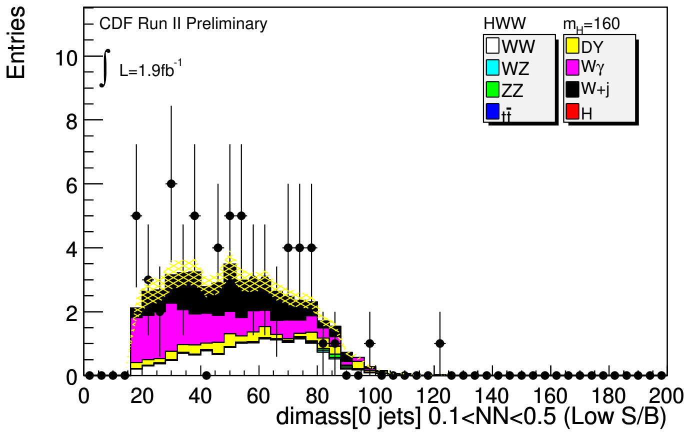
dimass: $NN < 0.1$ cut

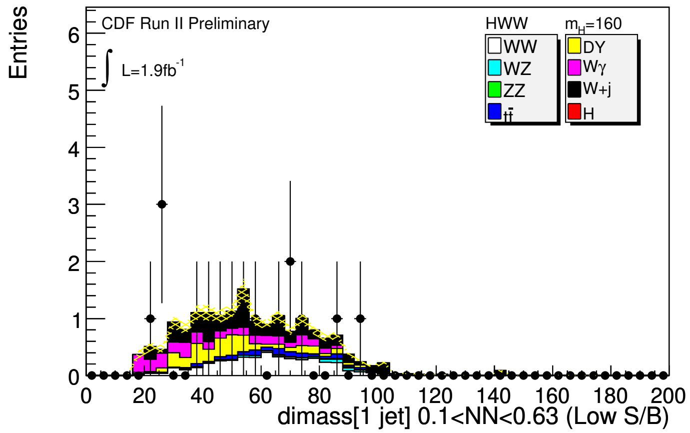


dimass: $0.1 < NN <$ 80% cut

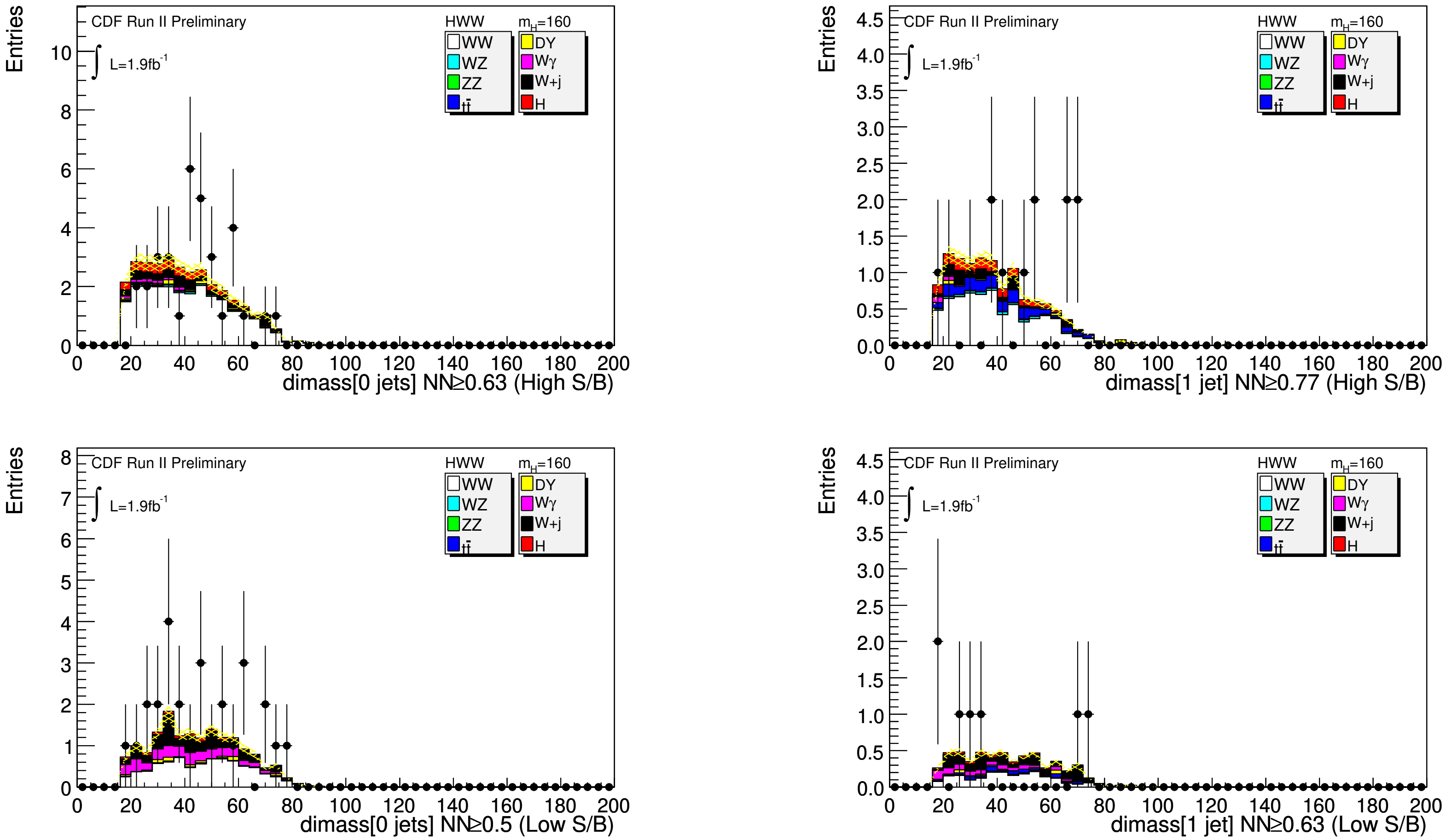




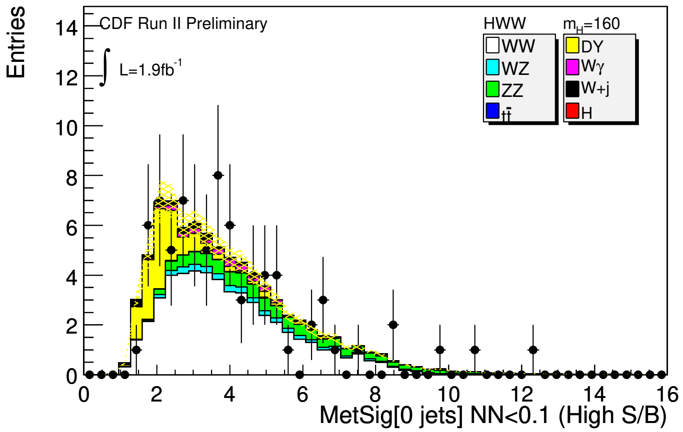


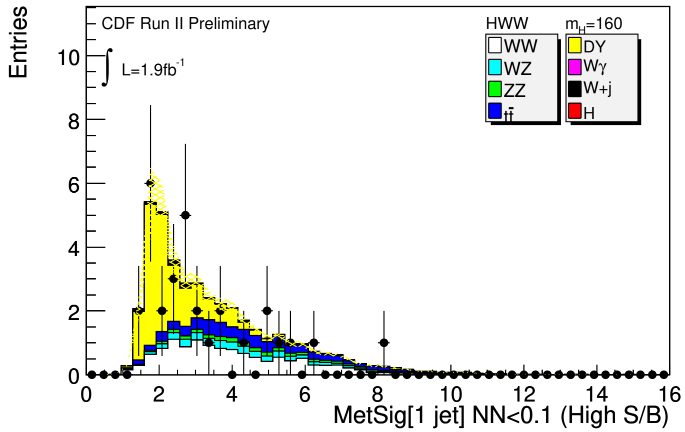


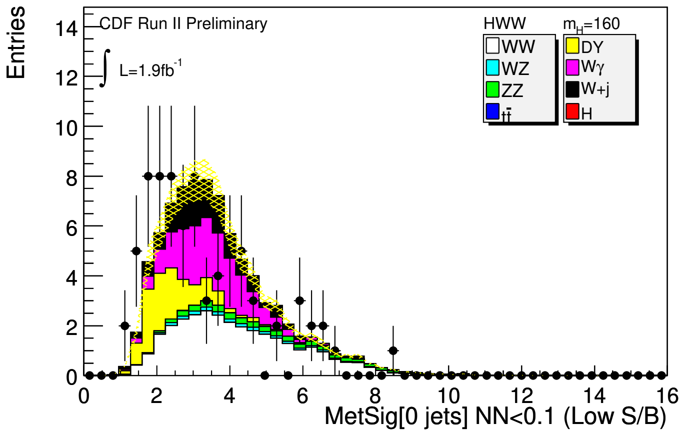
dimass: $NN \geq 80\%$ cut

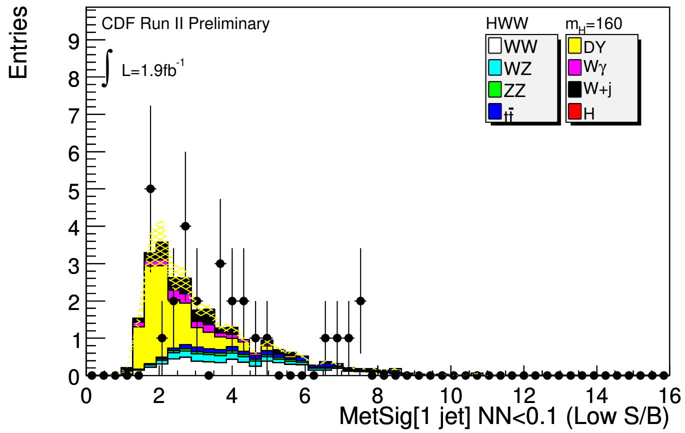


MetSig: $NN < 0.1$ cut

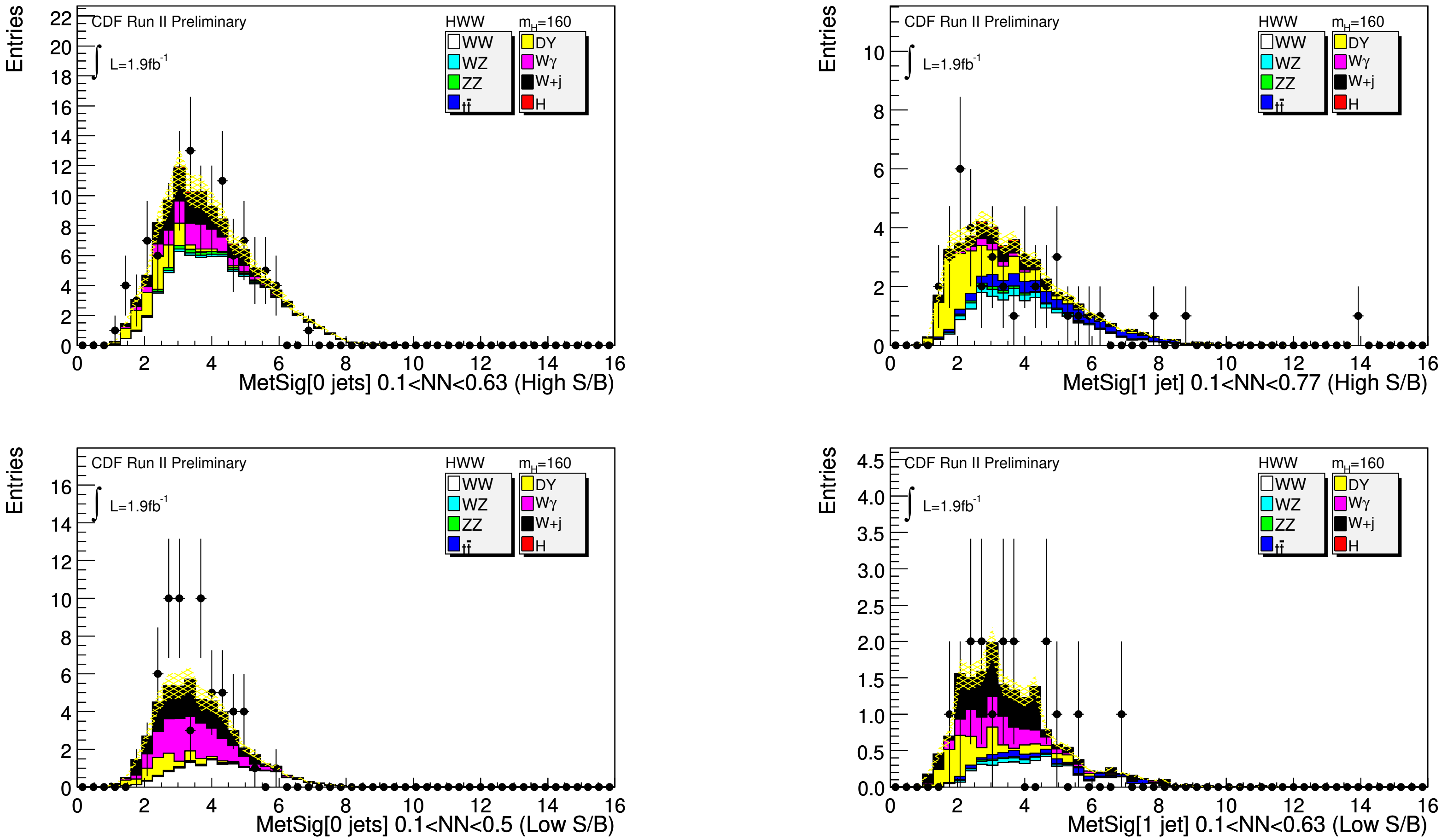




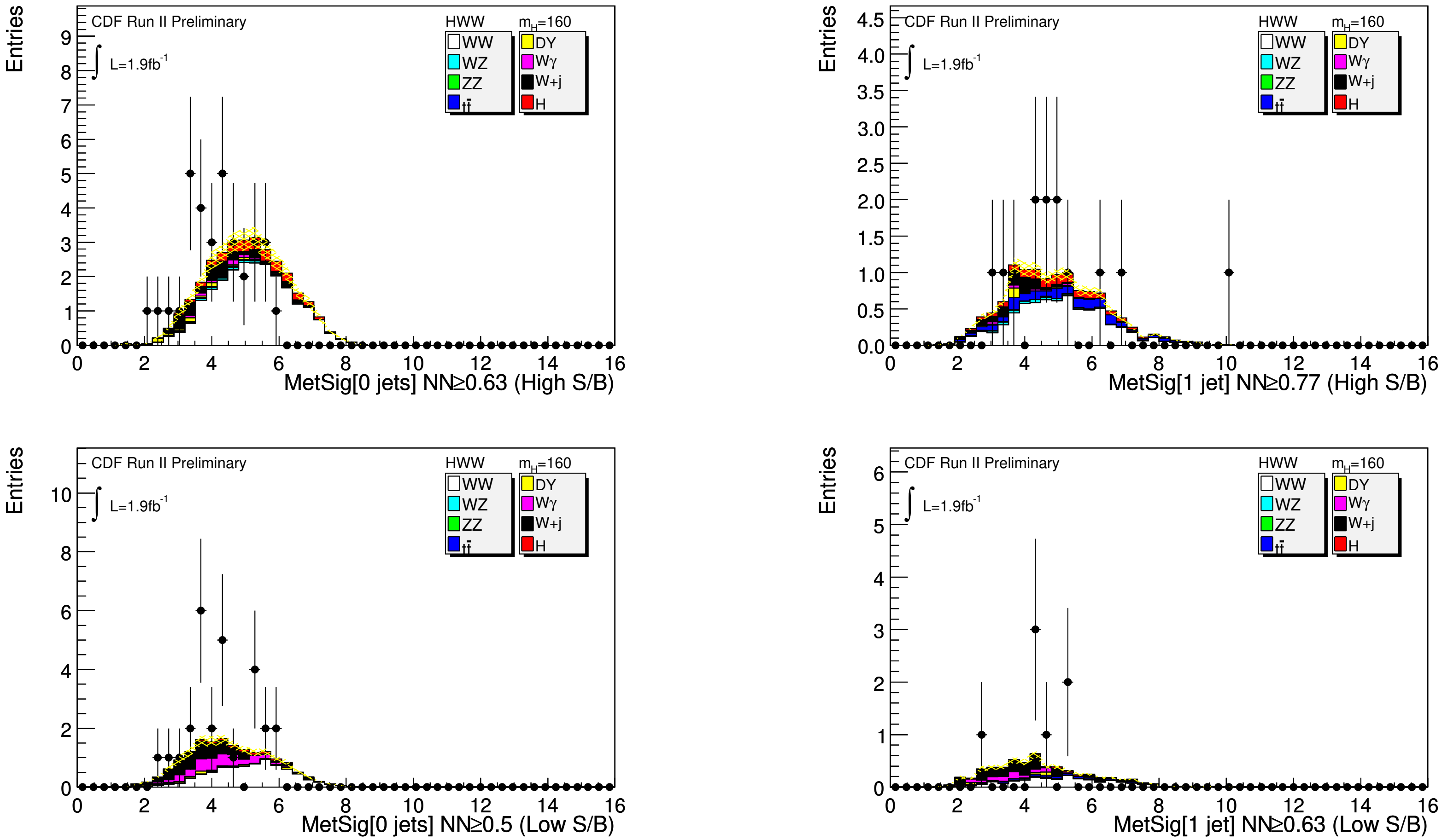




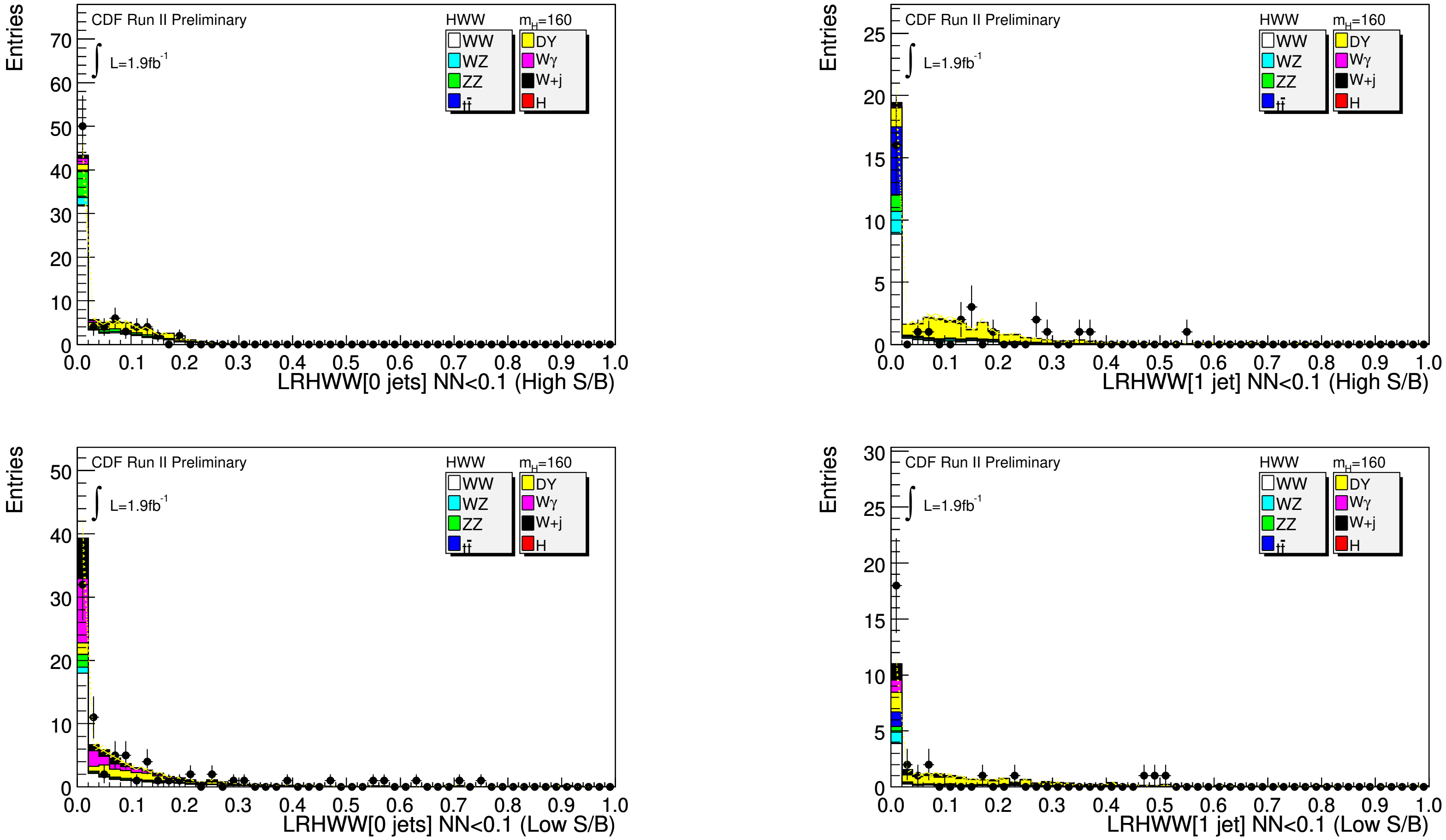
MetSig: $0.1 < NN <$ 80% cut



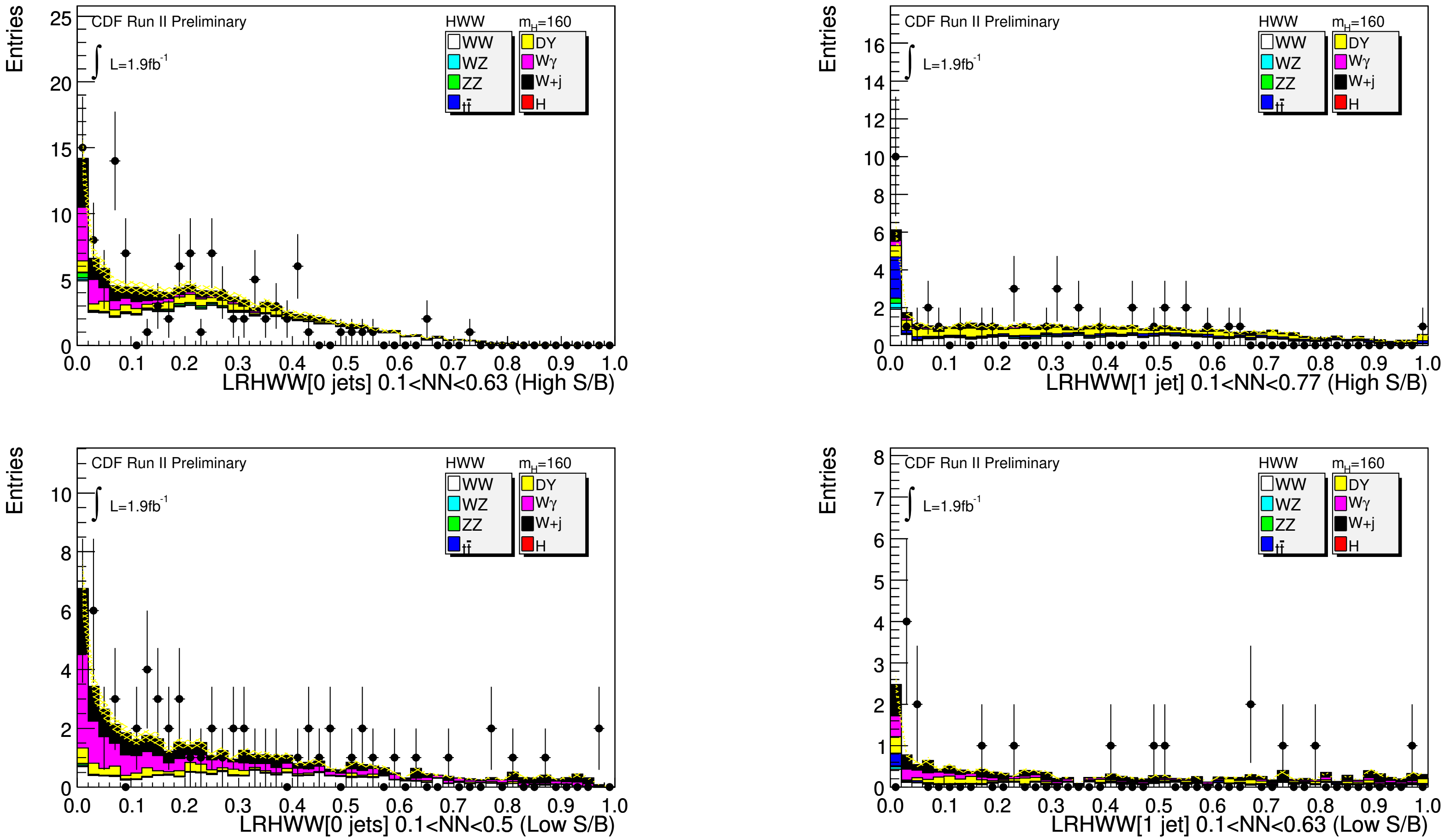
MetSig: $NN \geq$ 80% cut



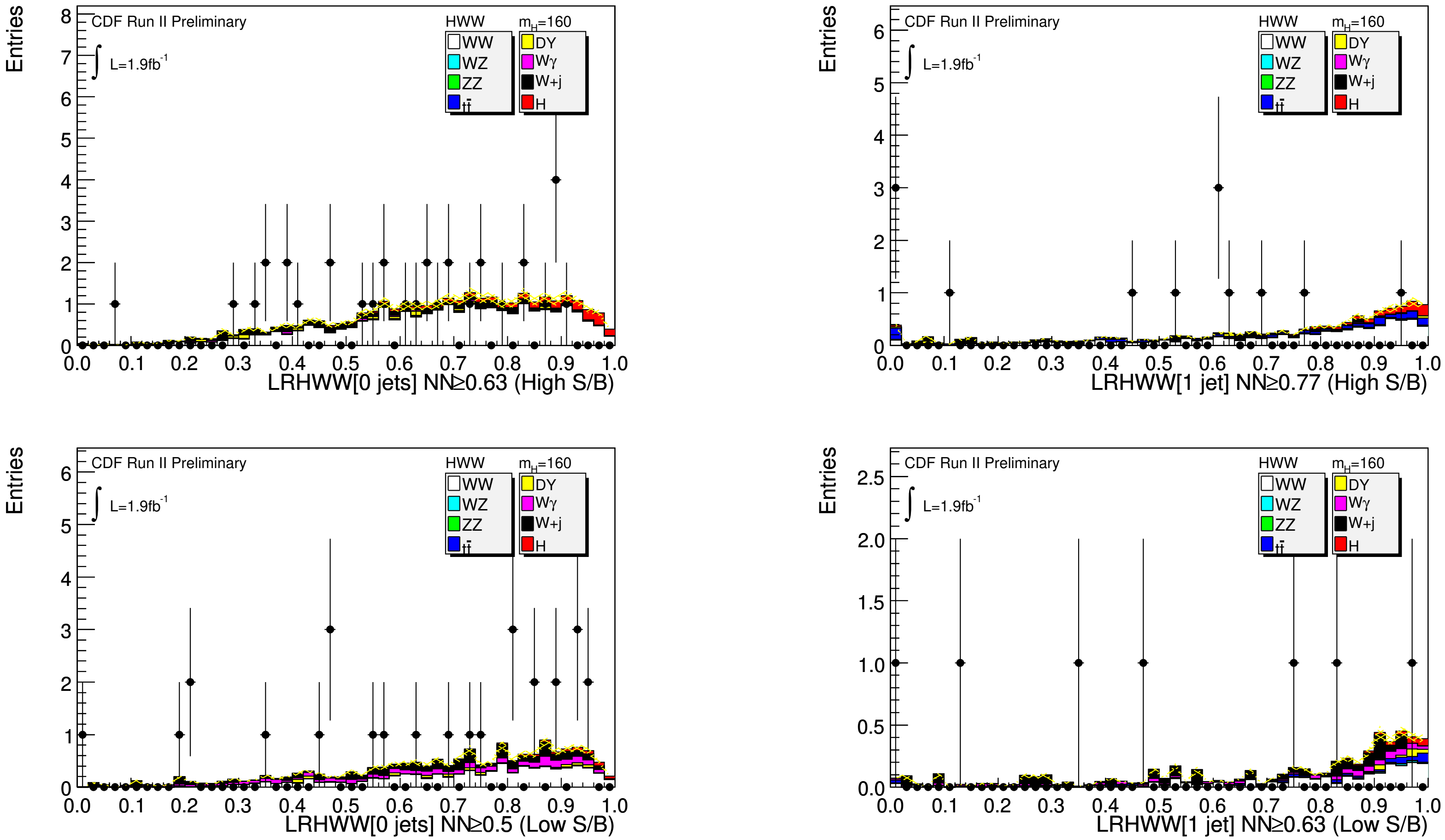
LRHWW: $NN < 0.1$ cut



LRHWW: $0.1 < NN <$ 80% cut



LRHWW: $NN \geq 80\%$ cut



Observations Concerning Dilepton Mass

- The NN<0.1 cut isolates the Z peak.
  - High S/B 0 jet has almost no Fake/Wγ; the continuum background is WW and the WZ and ZZ are concentrated at the Z mass with about 30% from ZZ.
  - The Low S/B 0 jet has a significant Fake/Wγ contribution, especially below the Z mass. The data are systematically higher and hence this Fake/Wγ may be underestimted.
  - For High S/B and 1 jet, there is about 50% of the expected number of events. There is a low mass bump of pure DY.
  - Fir Low S/B and 1 jet, statistics are low but the agreement is fine, and there is also a low mass bump.

- For $0.1 < NN < 80\%$ cut, there is no evidence of a Z contribution.
  - 0 jet: The Fake/Wγ in at lower mass in High and at all masses below the Z for Low S/B.
  - 1 Jet: Very little Fake/Wγ in high S/B. More DY in high S/B. Also, WZ/ZZ contribute noticeable here in the 1 jet High S/B and less so in the other cases. The 1 jet low S/B again has Fake/Wγ that masks the W/Z contribution.
- for NN$> 80\%$ cut the kinematic coverage is the same as for the $0.1 < NN < 80\%$ which therefor provides a good control region.

Observations Concerning MetSig

- $NN < 0.1$
  - The DY is at low MetSig in all cases
  - There is signficant Fake/Wγ for Low S/B 0 jet.
  - The WZ and ZZ are uniformly dense over MetSig$>$2.5
  - Top has its main contirbution for MetSig$>$2.5 and High S/B 1 jet.
- $0.1 < NN <$ 80% Cut
  - The Fakes/Wγ contribution dominates in Low S/B 0 jet.
  - The DY contribution apprears stronger in the 1 jet and especially in High S/B. The W/Z contribution seen in High S/B 1 jet is also visible here in the MetSig$>$2.5 region.
- $NN >$ 80% Cut

- The kinematic coverage here compares well with that in the $0.1 < NN <$ 80% region. The composition and proportion of backgrounds is well matched in the 0 jet.
- The 1 jet background in the control region contains a DY contribution that is not expected in this signal region. Nonetheless WW dominates the High S/B and the mixture of WW and Fakes/Wγ in the Low S/B is reasonable, although statistics are poor.

LRHWW Observations

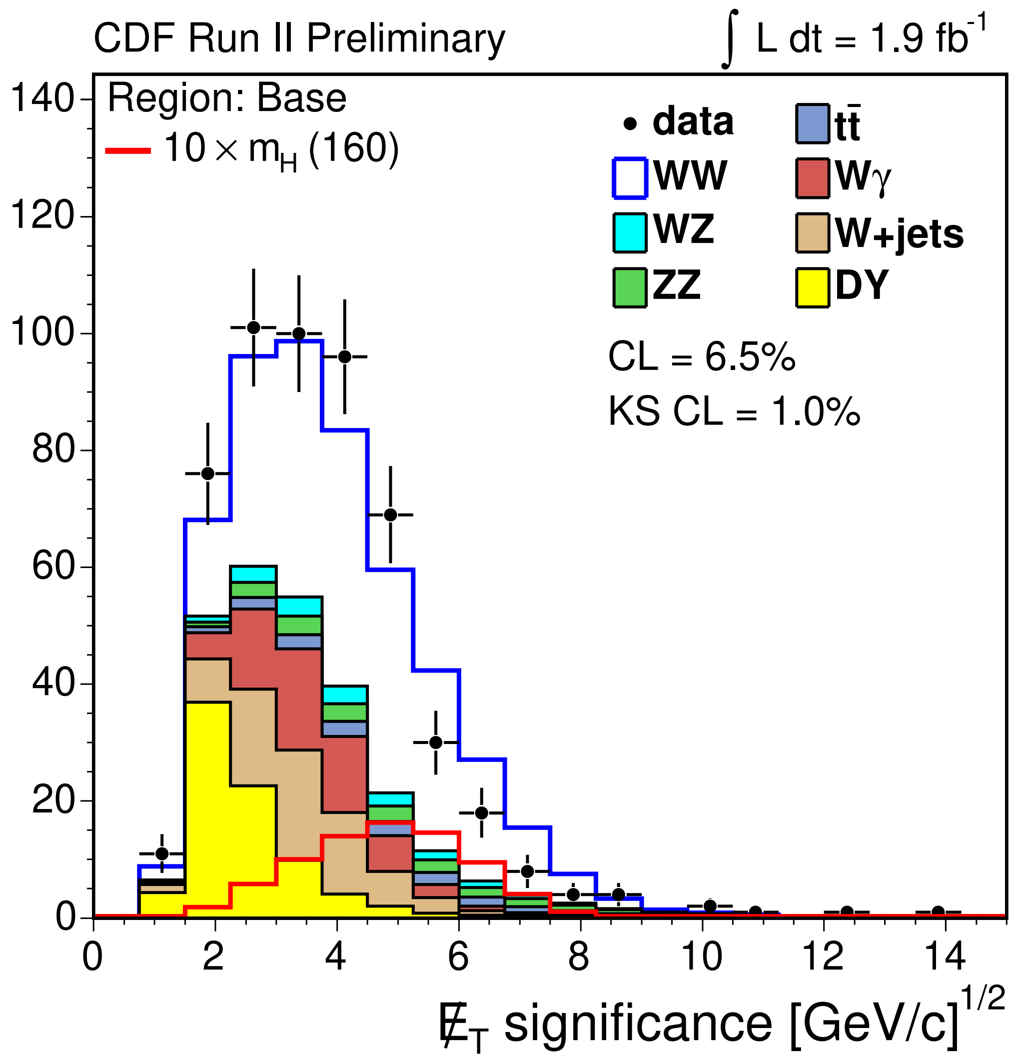
- The $NN < 0.1$ cut corresponds strongly to the peak at LRHWW=0.
- The intermediate NN score region corresponds to the intermediate LRHWW values as was expected from the high degree of correlation in the two estimators.
- The high score region has little statistics and again reflects the correlation between the estimators. There is clearly no control region in this case.

Reasons to Doubt MetSig

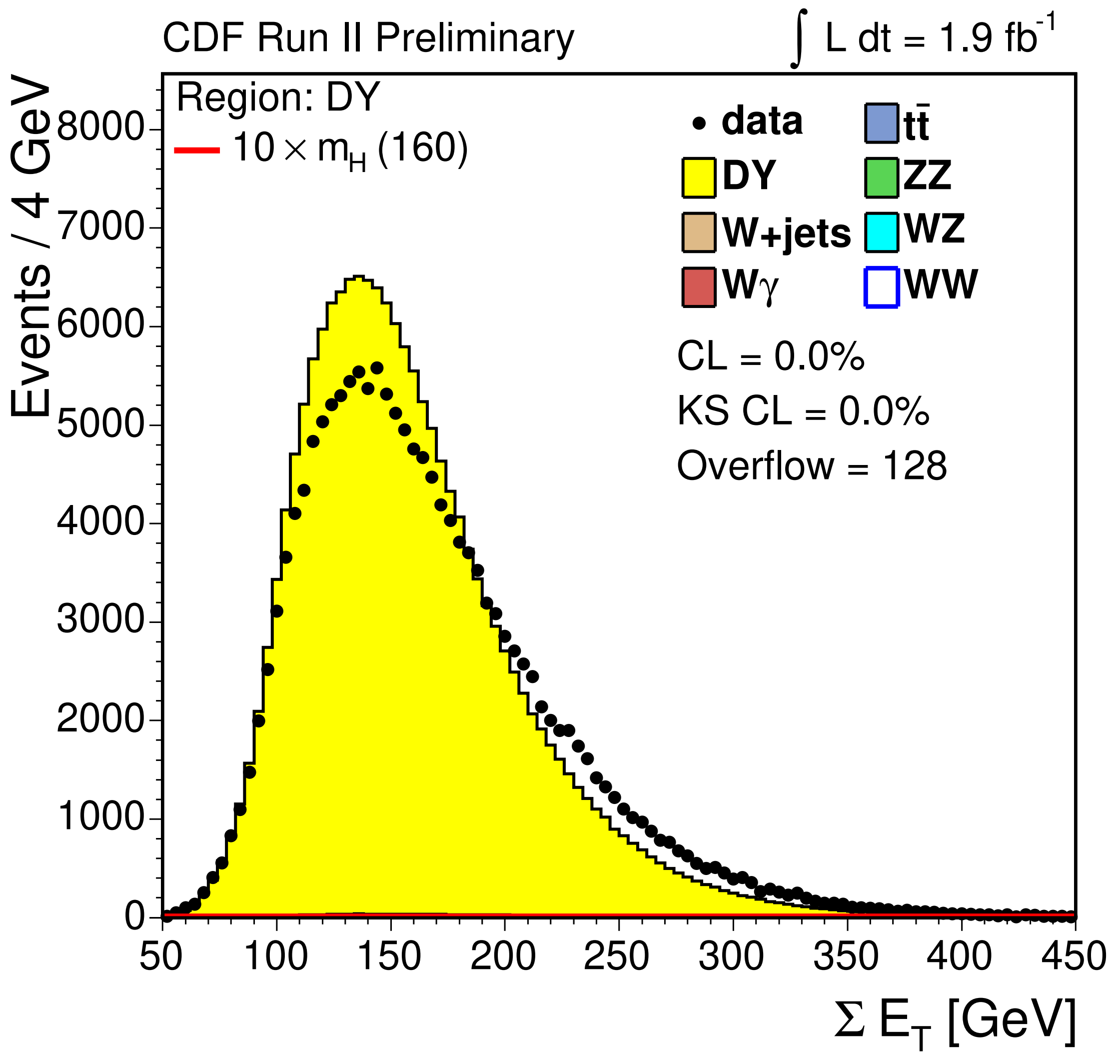
- The SumEt which is the denominator of MetSig is well known not to be well modeled by Pythia. This is one of the reasons that we don't use a MetSig cut in the ME analysis.
- The WW looks more like the culprit. WW is clearly overestimated in the large METsig region.
  - Note the bad KS tests, and it's clear why.
  - Also note the HWW shape. This suggests that this is the source of the signal excess when using MetSig.
- The DY SumEt plot showing with large statistics that mismodeling of SumEt is a generic Pythia issue, not specific to WW and hence not likely to be due to any background issues.
- In the WW cross-section measurement, we calculated a

systematic on the WW acceptance by reweighting the WW events by the ratio of DY data/MC as a function of SumEt. The effect was very small, but that may be partly because we only used a MetSig cut in the Z-peak region.

WW plot of MetSig

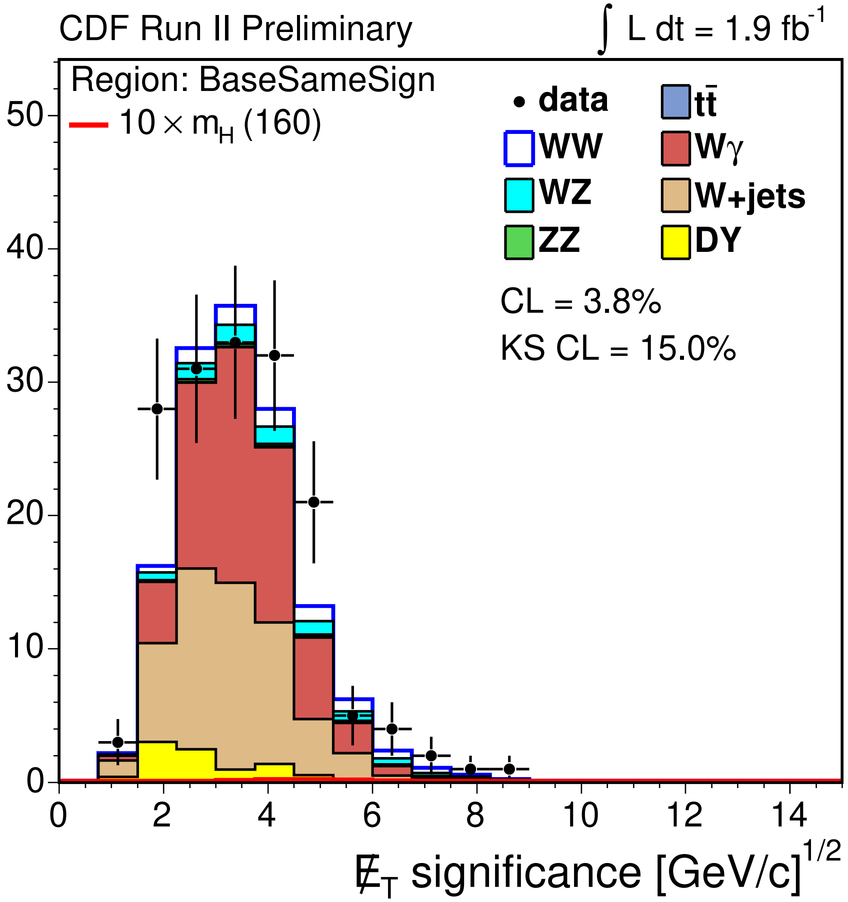


DY Plot of SumEt

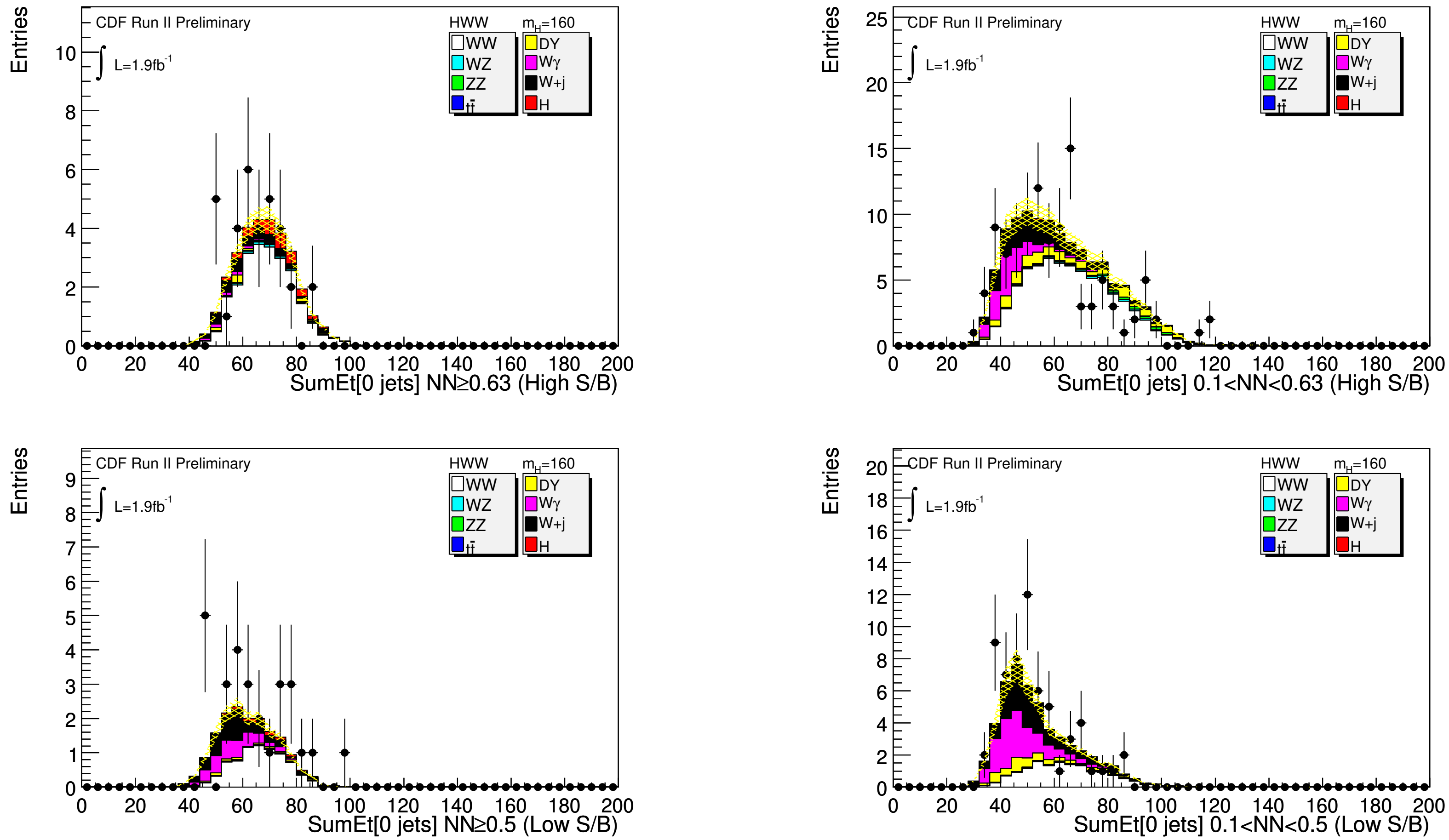


Reason to believe Fakes

Plot of (Base=llMET) Same Sign which shows the fakes are reasonably well modeled



Signal Coverage and DY

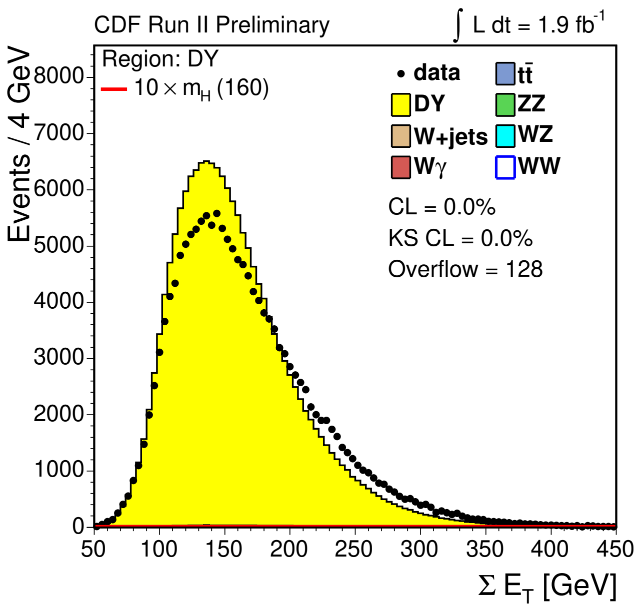


- Signal is in the region from 40-80.

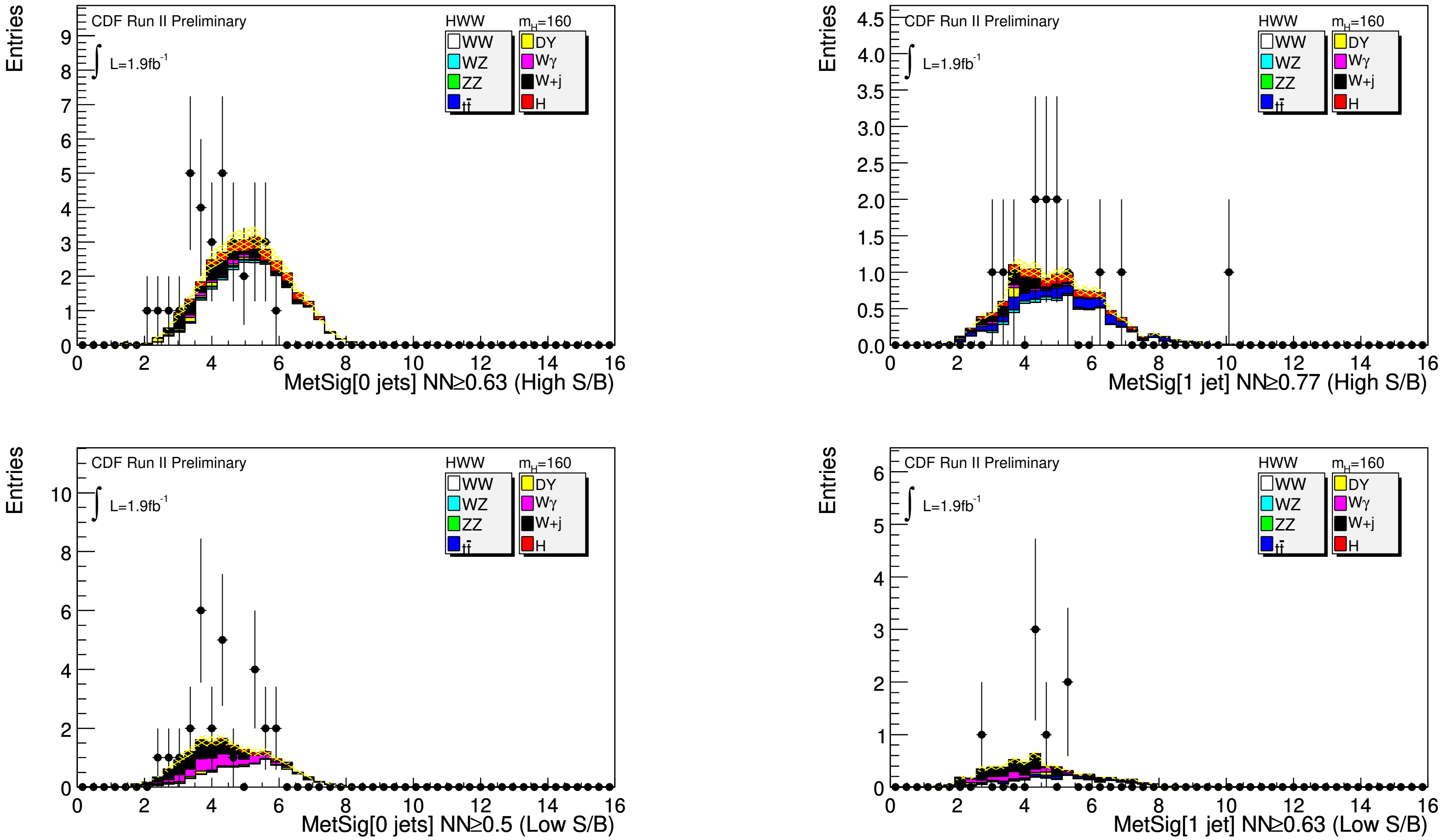
- High S/B: background is dominantly WW; bit of fakes that are smaller than the signal.
- Low S/B background is about 1/2 fake/WGamma and 1/2 WW.

DY Plot of SumEt

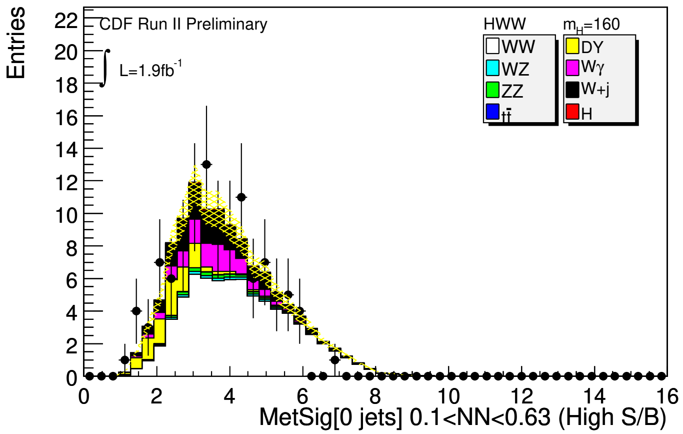
- SumEt for a Drell-Yan sample and what catches your eye is the disagreement in a region we dont care about. In fact the SumEt region from 40-80, (really can only see 50-80) is ok.
- Need to do Drell Yan in 0, 1 jet and high/low s/b.

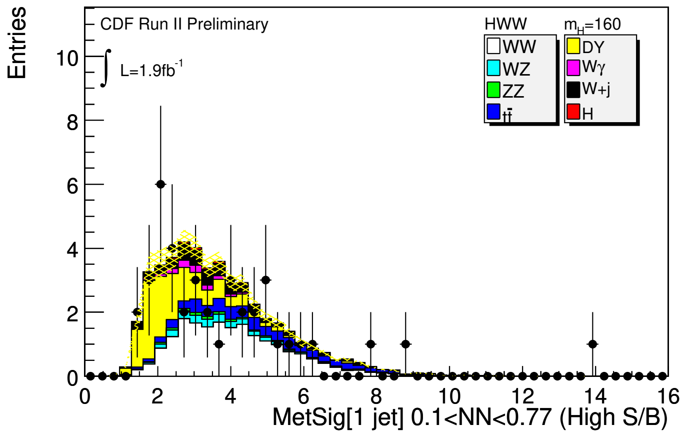


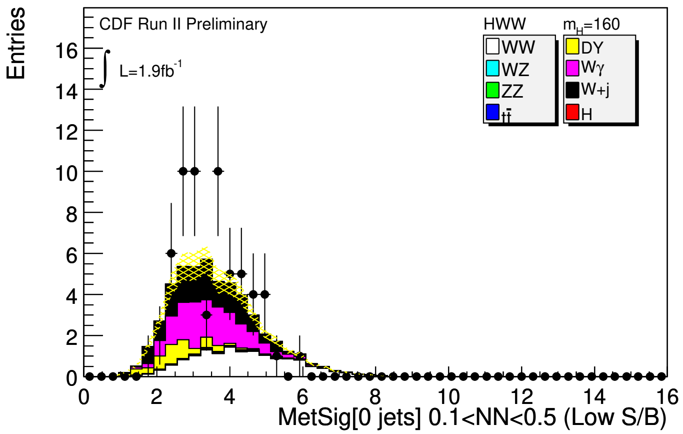
- MetSig composed of 0,1 and high/low S/B
- Good kinematic coverage;

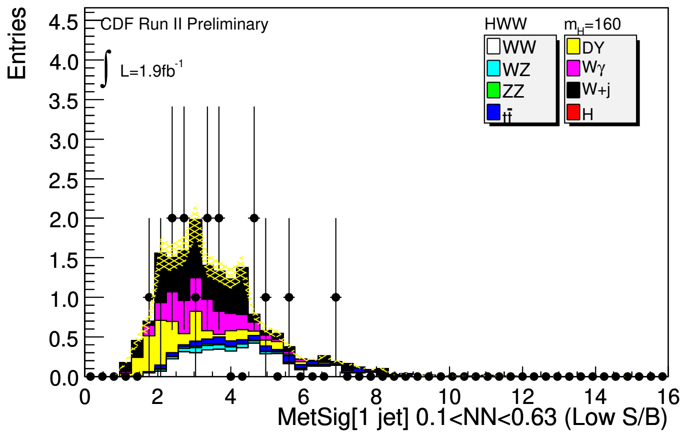


Low NN


CDF Run II Preliminary
$\int$ L=1.9fb^{-1}
HWW m_H=160
WW
WZ
ZZ
$t\bar{t}$
DY
Wγ
W+j
H
Entries
MetSig[0 jets] 0.1<NN<0.63 (High S/B)

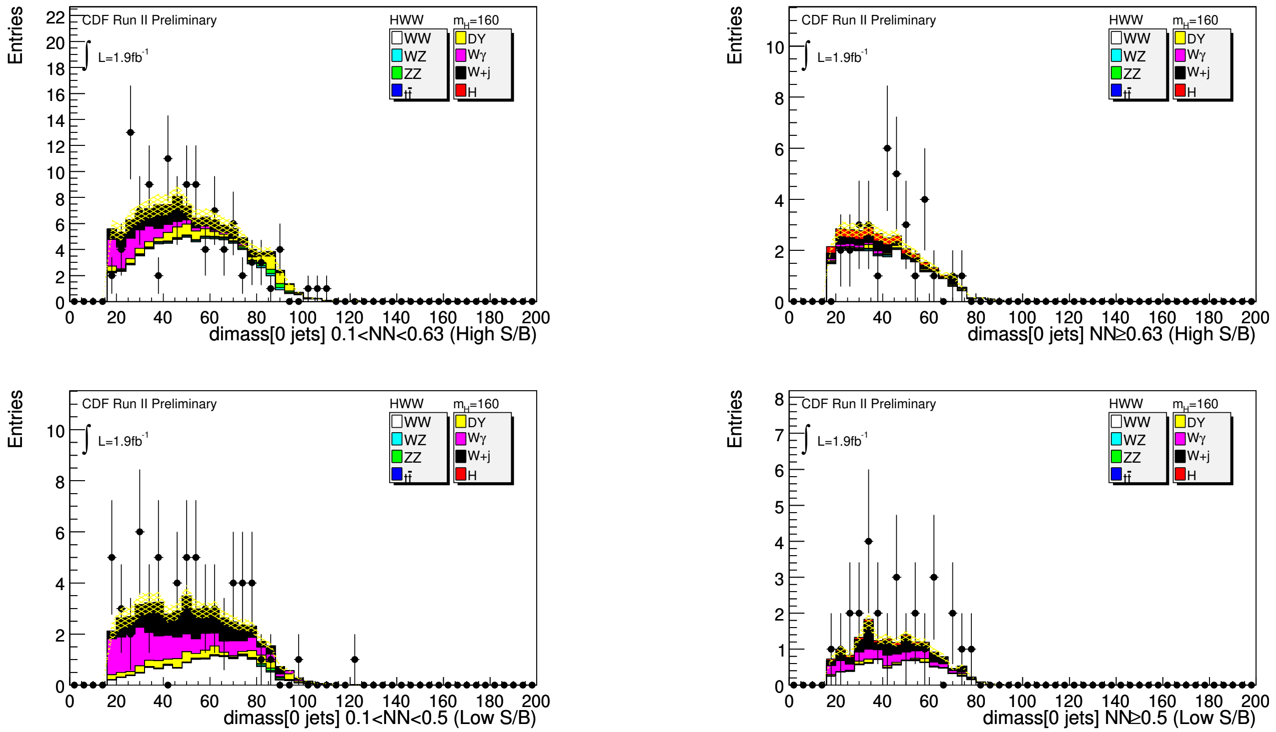

CDF Run II Preliminary
$\int$ L=1.9fb^{-1}
HWW m_H=160
WW
WZ
ZZ
$t\bar{t}$
DY
Wγ
W+j
H
Entries
MetSig[1 jet] 0.1<NN<0.77 (High S/B)


CDF Run II Preliminary
$\int$ L=1.9fb^{-1}
HWW m_H=160
WW
WZ
ZZ
$t\bar{t}$
DY
Wγ
W+j
H
Entries
MetSig[0 jets] 0.1<NN<0.5 (Low S/B)


CDF Run II Preliminary
$\int$ L=1.9fb^{-1}
HWW m_H=160
WW
WZ
ZZ
$t\bar{t}$
DY
Wγ
W+j
H
Entries
MetSig[1 jet] 0.1<NN<0.63 (Low S/B)

Fakes

- Shapes different for WW and Fakes/Wγ
- Low S/B shows difference.



DY and WW

- The DY and WW have the same reasons for SumEt and Met if one removes leptons!
- DY can be used for high stat test, divided into jet and high s/b bins. Difference in diagram where gluon emitted from propagator. Need to plot a new Met and MetSig with leptons removed to compare.
- Same sign Fakes need to be done by High/low S/B and 0 1 jet to see if consistent.
- Can fit fakes in Mll since shape is different and use that (for low s/b).

Summary

- For the discriminant variables that caused concerns regarding the agreement of data and Monte Carlo, we have been able to define control regions in which the agreement is good and remove regions that were not relevant from consideration: the $NN < 0.1$ region
- However, for this $NN < 0.1$ region, we do see a contribution of ZZ to the Z mass that is particularly clean in the High S/B 0 jet although an identically shaped DY contribution is present. A cut on MetSig clearly would remove this with high efficiency.
- For Mll, the isolation of the Z peak with $NN < 0.1$ is striking. This cut also brings out the lowest LRHWW values.
- The region from $0.1 < NN <$80% cut offers a good control

region with a good mixture of backgrounds distributed over the same kinematics as the signal region, $NN > 80\%$ cut. The exception is the additional DY background in the 1 jet High S/B control region that is not expected in the signal region.

- The control region shows that for the Low S/B there is some discrepancy in the data and background predictions for 0 jet. Statistics are too low in the 1 jet bin to say much. The dominant background is Fake/Wγ and this needs to be investigated. While this does not dominate the High S/B region it may be important since it does lie in the kinematic signal region.
- DY and Fakes-same-sign will be studied in 0,1 jet and high/low s/b. DY will be examined with leptons removed to allow comparison to WW.