VBF Higgs → ZZ → mumuJetJet

- Event selection:
  - Z → mumu (at least 2 leptons, check what is the cut on the Z mass if any)
  - N jets (3 or 4 depends on the energy of one of the jets)
  - Tag jets identified with a BDT (no cut)
- This selection was applied to all signal and background sample in the next slide
  - The BDT trained in the corresponding signal mass point was used in the backgrounds too

MC samples

- Signal:
  - VBFHiggs with M = 500 GeV
- Background:
  - Z+Jets in two energy bins:
    - 10 < M < 50
    - 50 < M < 10000
  - Ttbar
  - ZZ
  - WZ
  - VQQ

About the TagJet BDT

- Variables used in the training:
  - HighEtaJJ
  - DeltaEtaJJ
  - DeltaPhiJJ
  - EtaXetaJJ
  - PzXpzJJ
  - HighPtJJ
  - LowPtJJ
  - SumetaJJ
  - DeltaPtJJ
  - SumPtJJ
  - PtJJsys
  - MassJJsys
  - DeltaEtaJJsysZlep
  - scalProdJJsysZlep

Possible variables for HadZ BDT

Had Z identification variables

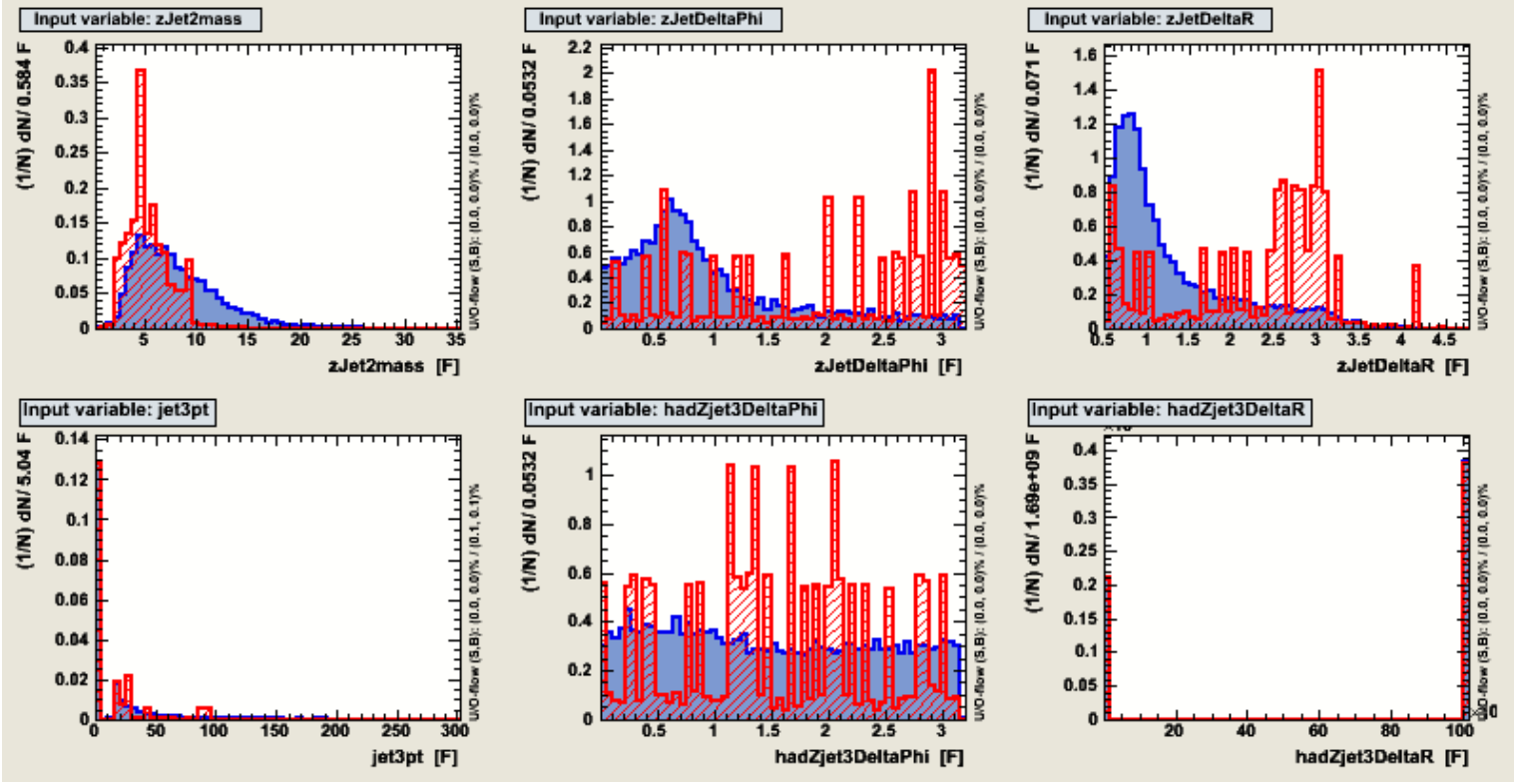
- TagBDToutput
- Zjet1pt
- Zjet2pt
- Zjet1eta
- Zjet2eta
- Zjet1mass
- Zjet2mass
- ZjetsDeltaPhi
- ZjetsDeltaR
- Jet3pt
- Jet3eta
- HadZjet3DeltaPhi
- HadZjet3DeltaR
- LepZjet3DeltaPhi
- LepZjet3DeltaR
- HadZmass
- HadZpt
- ZzDeltaPhi
- ZzDeltaR
- DeltaEtaJJsysZhad
- ScalProdJJsysZhad

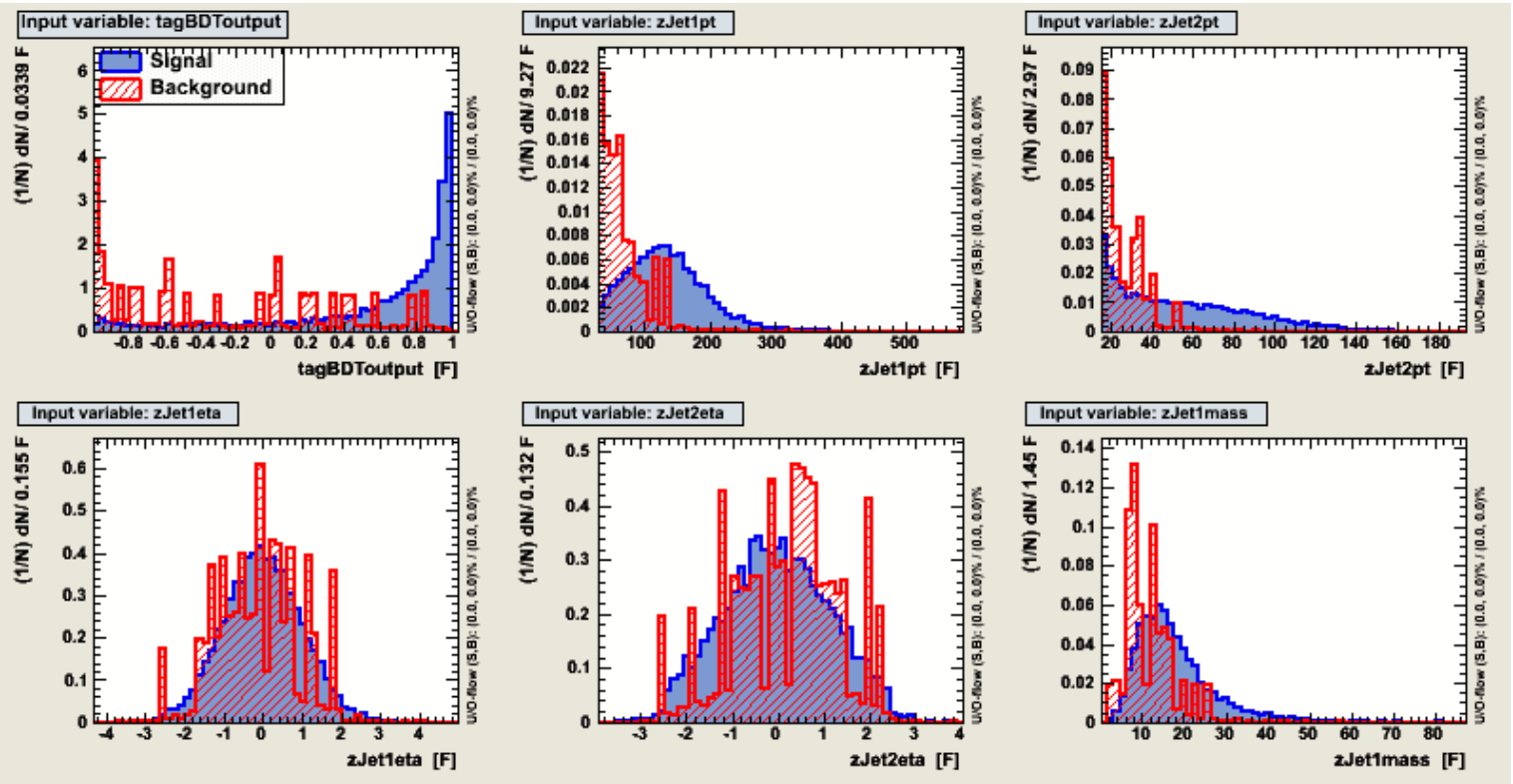
Other variables for the full system

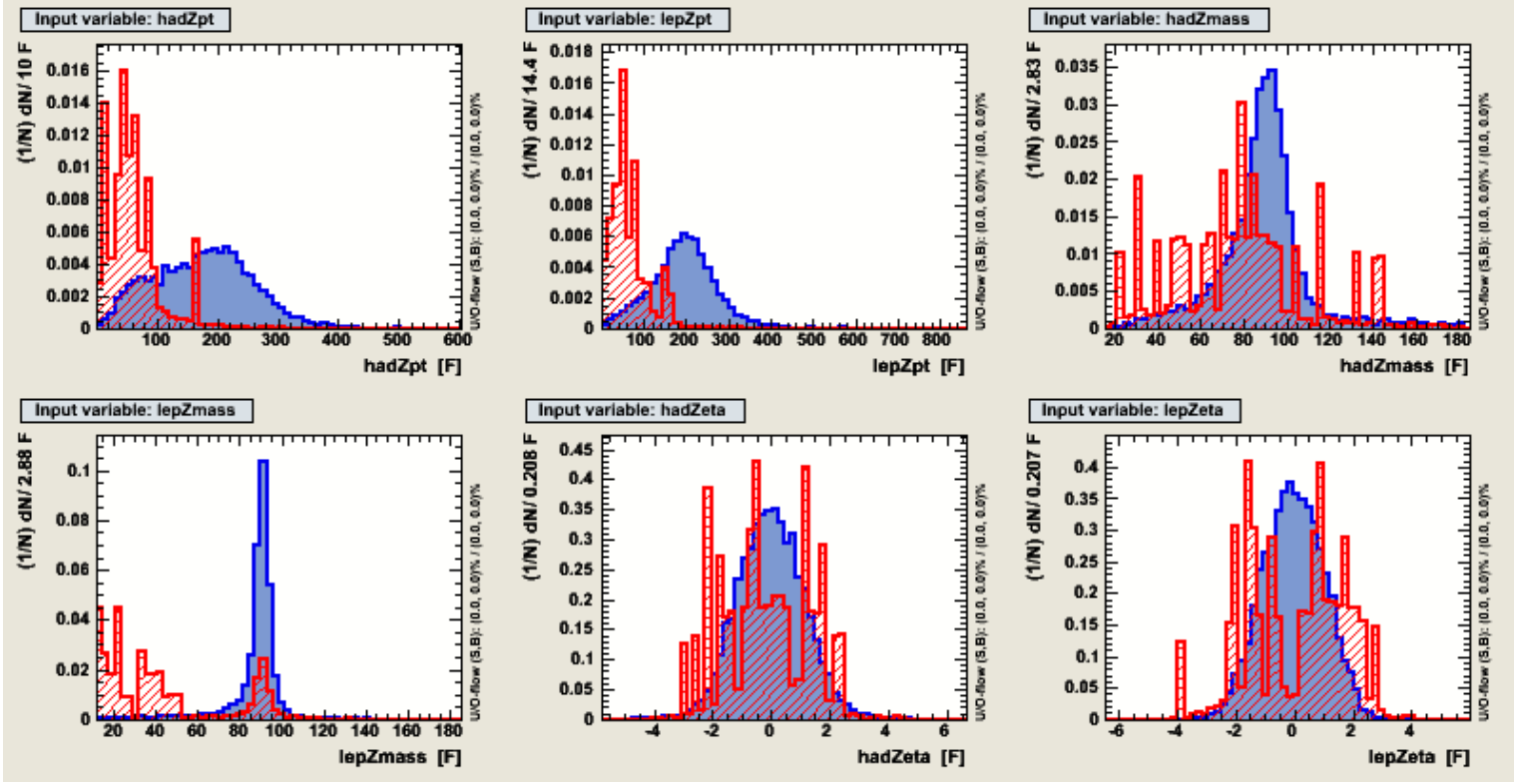
- Zmuon1pt
- Zmuon2pt
- Zmuon1eta
- Zmuon2eta
- LepZpt
- LepZmass
- LepZeta
- HadZeta
- PtBalance
- Met
- SumET

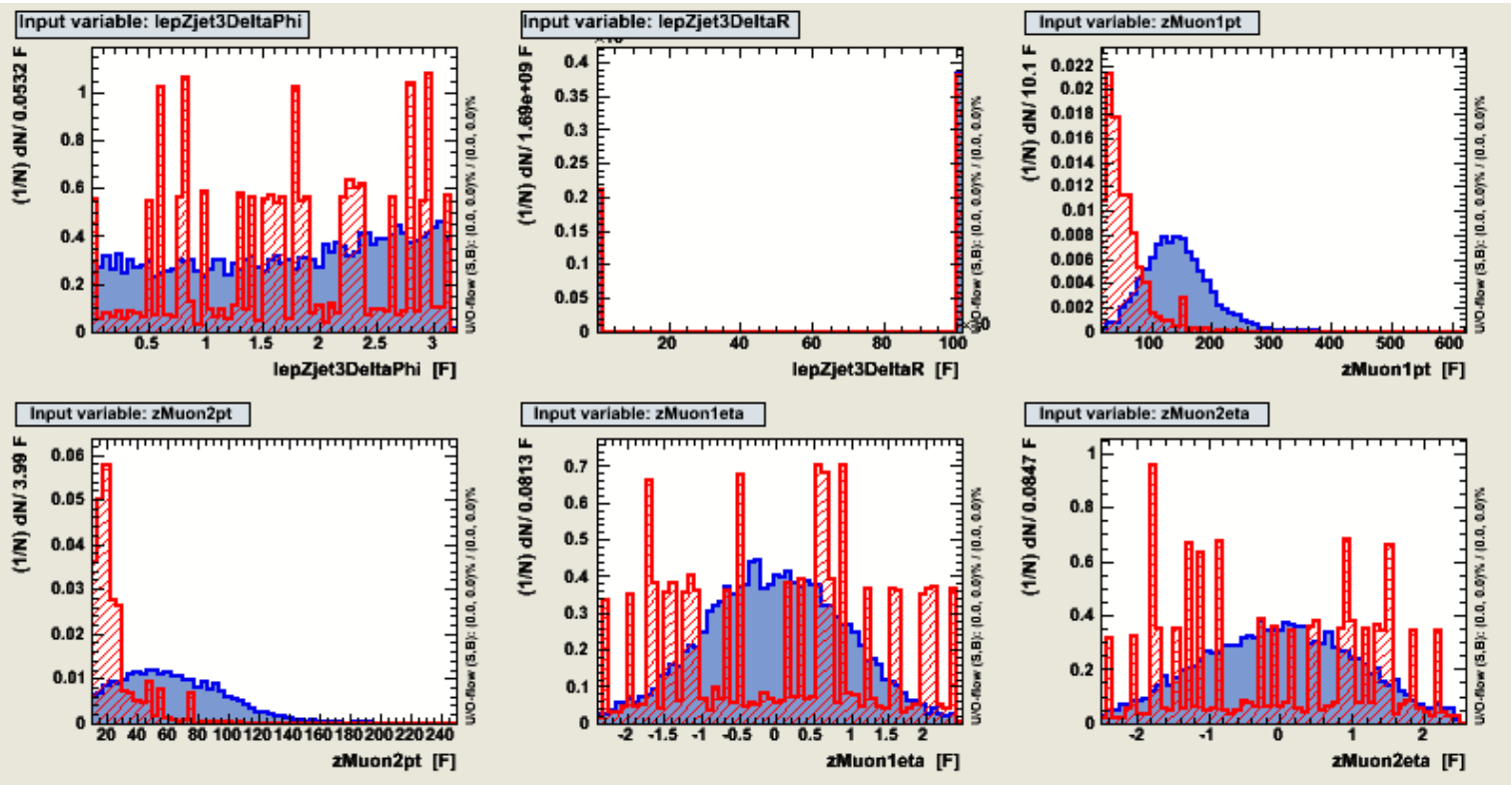
Using all samples

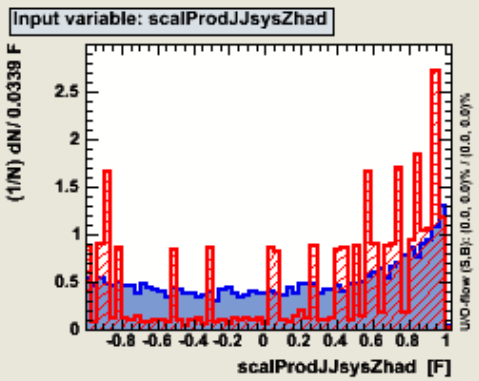
half events training half test

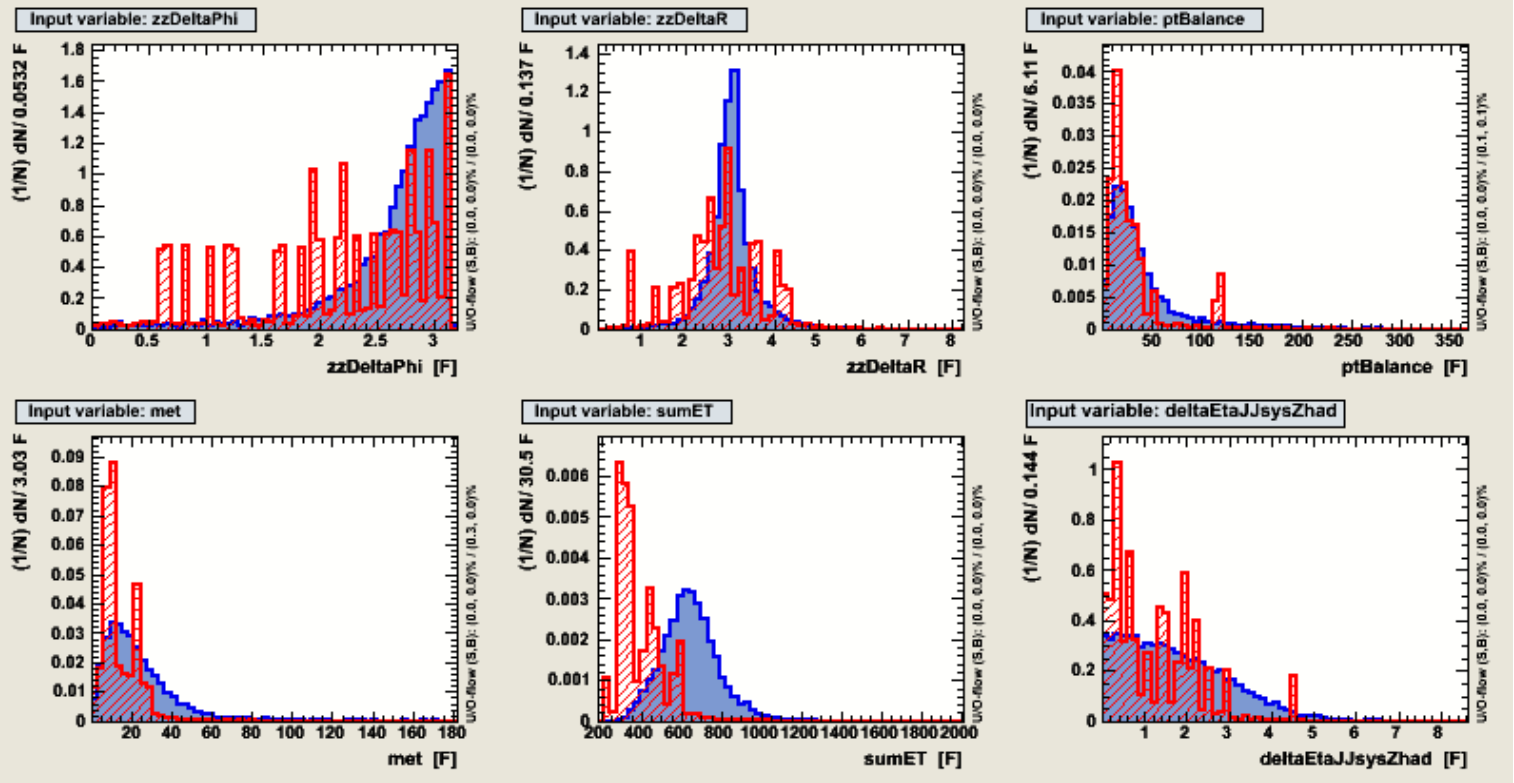

Input variable: zJet2mass
(1/N) dN/ 0.584 F
zJet2mass [F]
Input variable: zJetDeltaPhi
(1/N) dN/ 0.0532 F
zJetDeltaPhi [F]
Input variable: zJetDeltaR
(1/N) dN/ 0.071 F
zJetDeltaR [F]
Input variable: jet3pt
(1/N) dN/ 5.04 F
jet3pt [F]
Input variable: hadZjet3DeltaPhi
(1/N) dN/ 0.0532 F
hadZjet3DeltaPhi [F]
Input variable: hadZjet3DeltaR
(1/N) dN/ 1.69e+09 F
hadZjet3DeltaR [F]


Input variable: tagBDToutput
Signal
Background
(1/N) dN/ 0.0339 F
tagBDToutput [F]
Input variable: zJet1pt
(1/N) dN/ 9.27 F
zJet1pt [F]
Input variable: zJet2pt
(1/N) dN/ 2.97 F
zJet2pt [F]
Input variable: zJet1eta
(1/N) dN/ 0.155 F
zJet1eta [F]
Input variable: zJet2eta
(1/N) dN/ 0.132 F
zJet2eta [F]
Input variable: zJet1mass
(1/N) dN/ 1.45 F
zJet1mass [F]


Input variable: hadZpt
(1/N) dN/ 10 F
hadZpt [F]
Input variable: lepZpt
(1/N) dN/ 14.4 F
lepZpt [F]
Input variable: hadZmass
(1/N) dN/ 2.83 F
hadZmass [F]
Input variable: lepZmass
(1/N) dN/ 2.88 F
lepZmass [F]
Input variable: hadZeta
(1/N) dN/ 0.208 F
hadZeta [F]
Input variable: lepZeta
(1/N) dN/ 0.207 F
lepZeta [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%

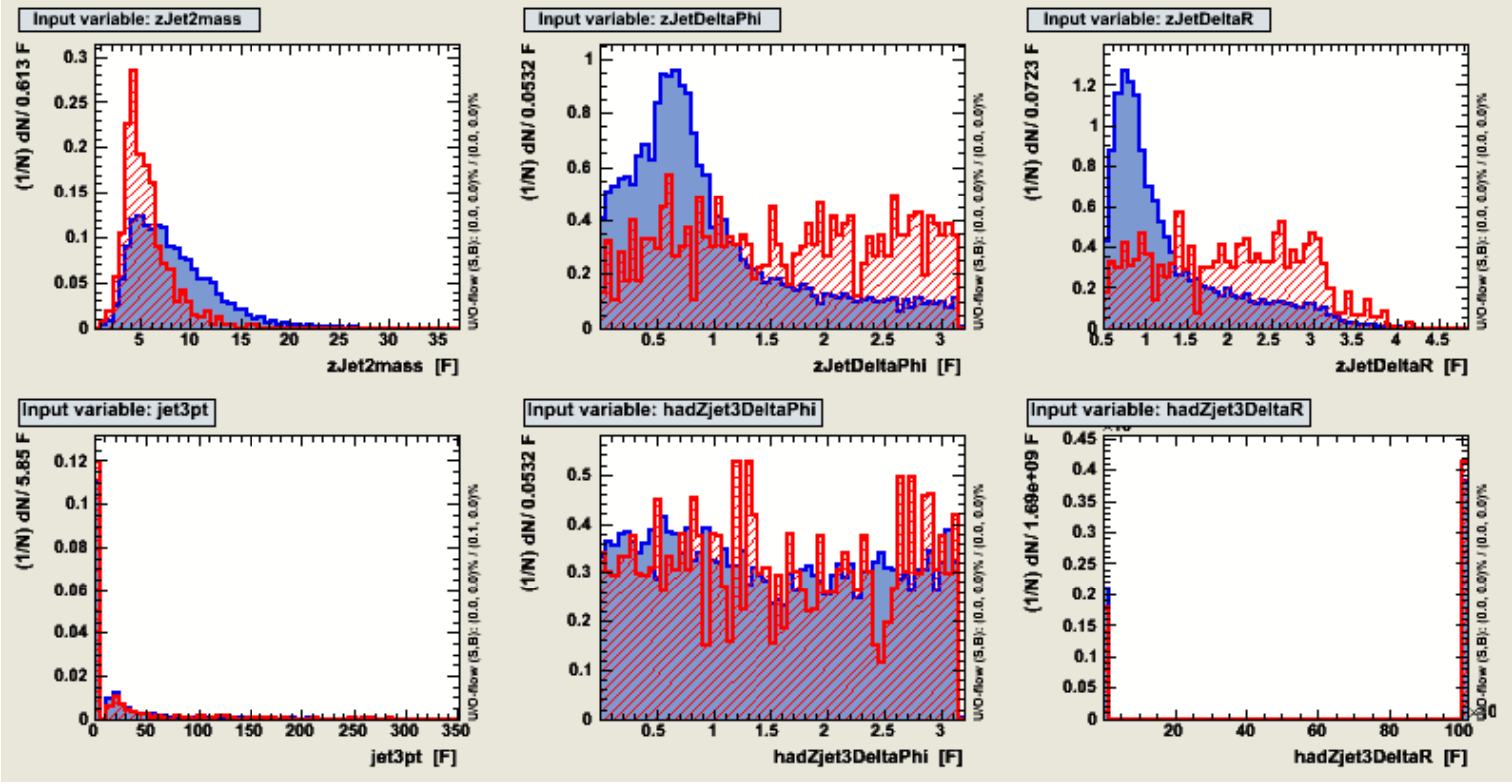

Input variable: lepZjet3DeltaPhi
(1/N) dN/ 0.0532 F
lepZjet3DeltaPhi [F]
Input variable: lepZjet3DeltaR
(1/N) dN/ 1.69e+09 F
lepZjet3DeltaR [F]
Input variable: zMuon1pt
(1/N) dN/ 10.1 F
zMuon1pt [F]
Input variable: zMuon2pt
(1/N) dN/ 3.99 F
zMuon2pt [F]
Input variable: zMuon1eta
(1/N) dN/ 0.0813 F
zMuon1eta [F]
Input variable: zMuon2eta
(1/N) dN/ 0.0847 F
zMuon2eta [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%

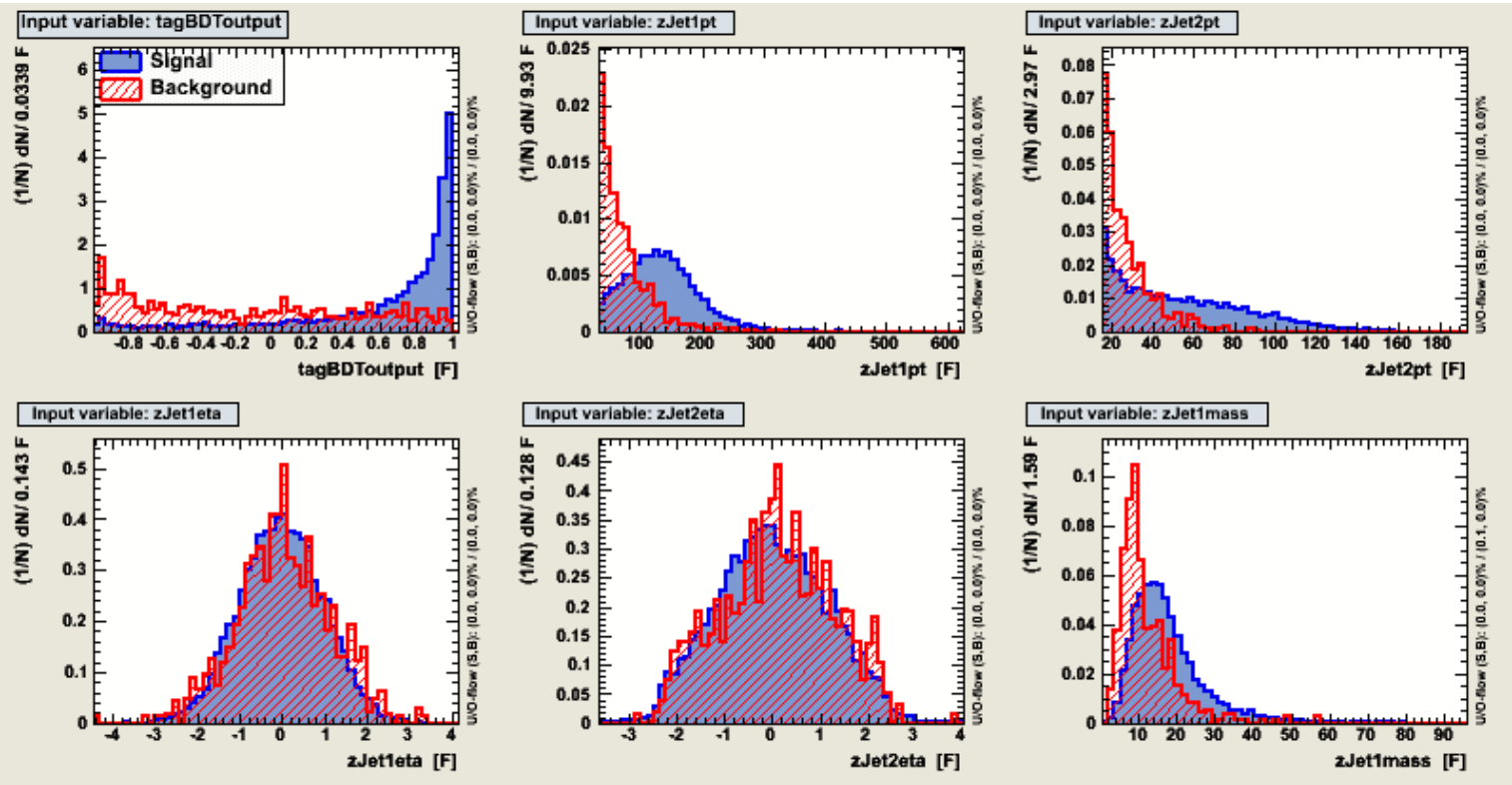

Input variable: scalProdJJsysZhad
(1/N) dN/ 0.0339 F
scalProdJJsysZhad [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%

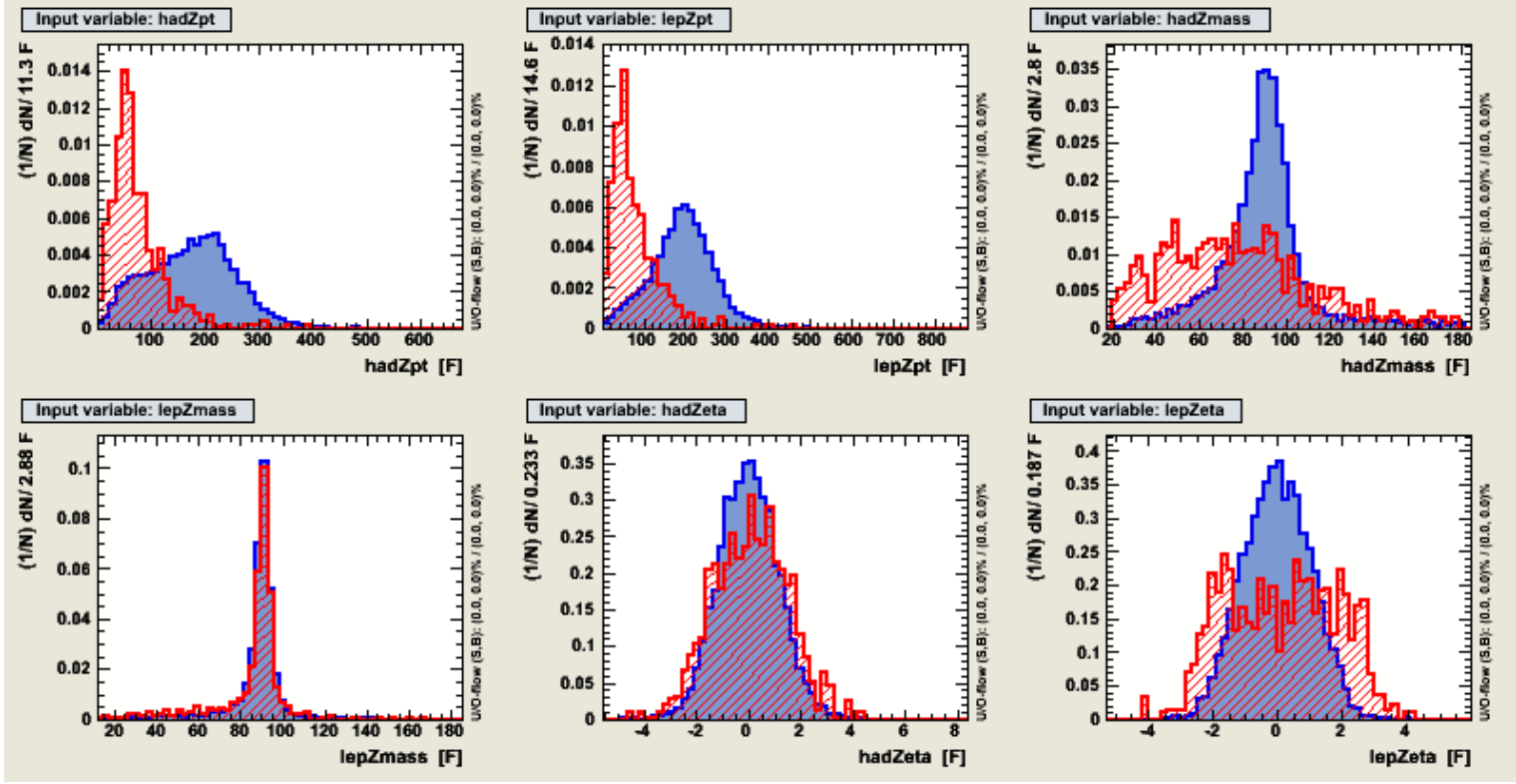

Input variable: zzDeltaPhi
(1/N) dN/ 0.0532 F
zzDeltaPhi [F]
Input variable: zzDeltaR
(1/N) dN/ 0.137 F
zzDeltaR [F]
Input variable: ptBalance
(1/N) dN/ 6.11 F
ptBalance [F]
Input variable: met
(1/N) dN/ 3.03 F
met [F]
Input variable: sumET
(1/N) dN/ 30.5 F
sumET [F]
Input variable: deltaEtaJJsysZhad
(1/N) dN/ 0.144 F
deltaEtaJJsysZhad [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%

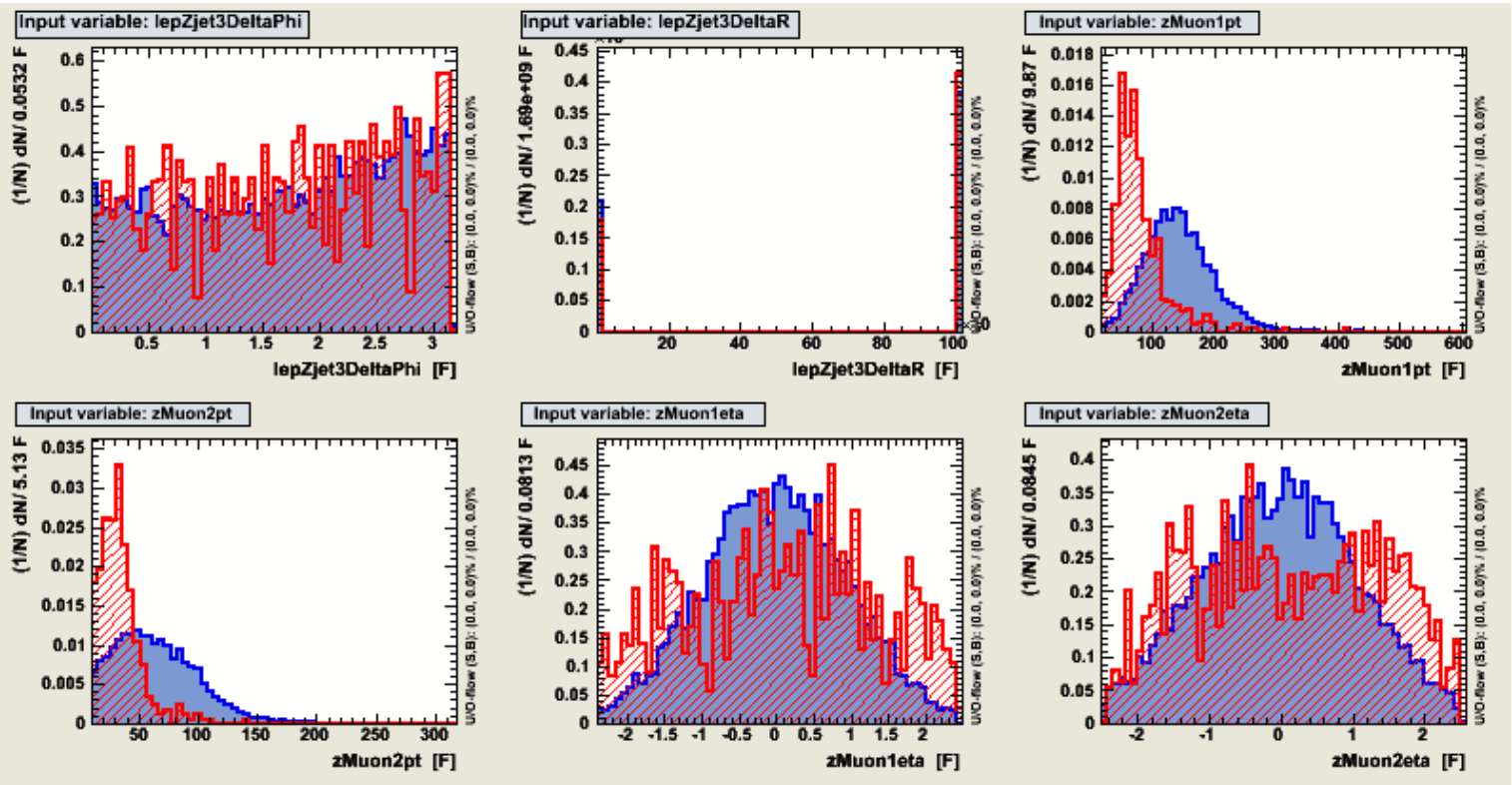
Removing the low mass Z+jet sample

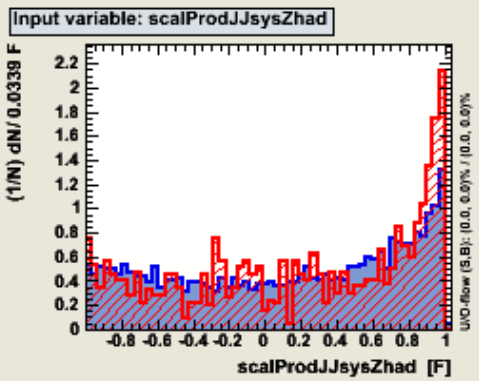
- It has very few events with a huge weight... It could distort the training
  - One possibility is to leave it out of the training and use it in the test phase to check that the events are still properly treated.
- Next slides show the same variables when the Z+jet with $10 < M < 50$ GeV sample is removed

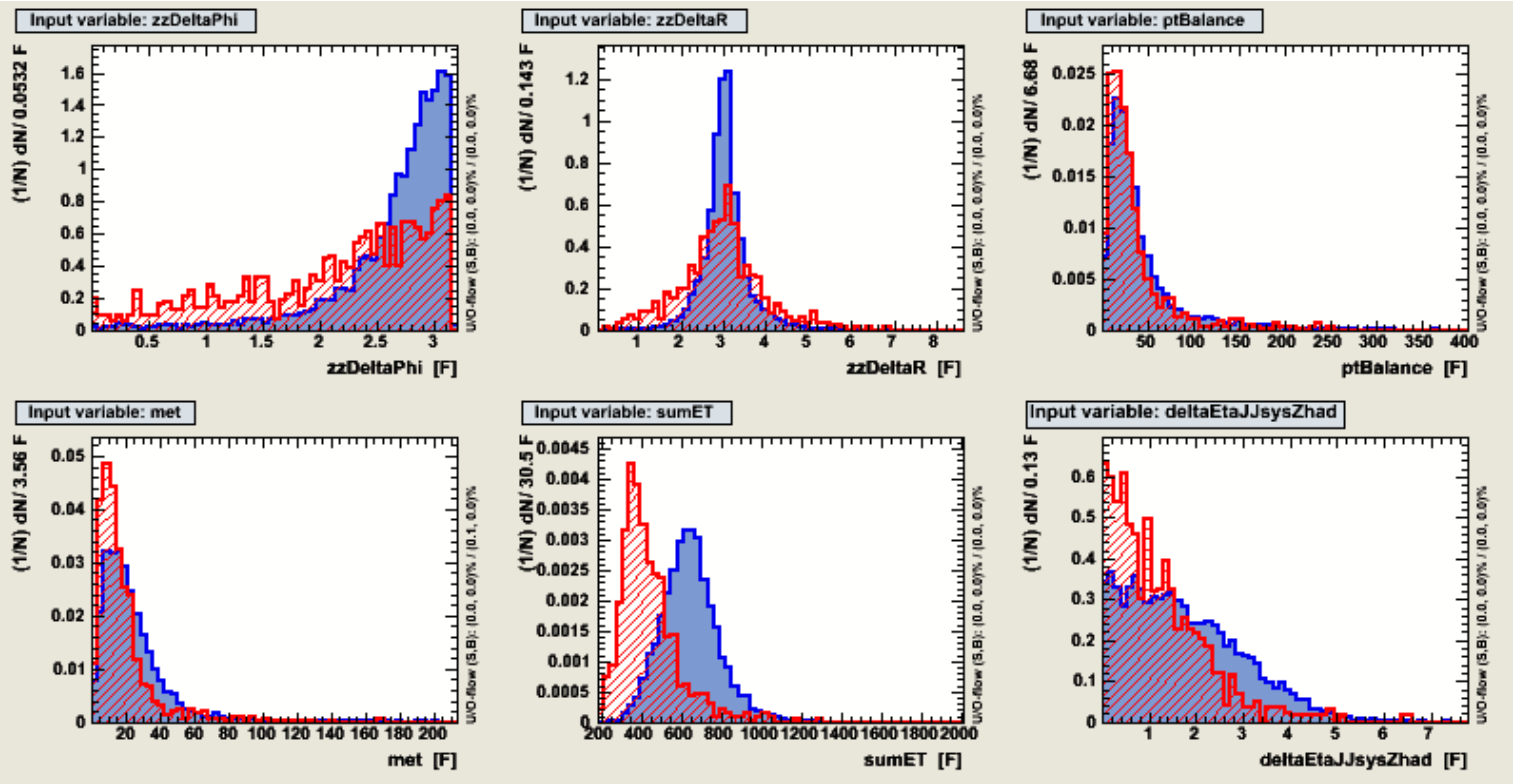

Input variable: zJet2mass
(1/N) dN/ 0.613 F
zJet2mass [F]
Input variable: zJetDeltaPhi
(1/N) dN/ 0.0532 F
zJetDeltaPhi [F]
Input variable: zJetDeltaR
(1/N) dN/ 0.0723 F
zJetDeltaR [F]
Input variable: jet3pt
(1/N) dN/ 5.85 F
jet3pt [F]
Input variable: hadZjet3DeltaPhi
(1/N) dN/ 0.0532 F
hadZjet3DeltaPhi [F]
Input variable: hadZjet3DeltaR
(1/N) dN/ 1.69e+09 F
hadZjet3DeltaR [F]


Input variable: tagBDToutput
Signal
Background
(1/N) dN/ 0.0339 F
tagBDToutput [F]
Input variable: zJet1pt
(1/N) dN/ 9.93 F
zJet1pt [F]
Input variable: zJet2pt
(1/N) dN/ 2.97 F
zJet2pt [F]
Input variable: zJet1eta
(1/N) dN/ 0.143 F
zJet1eta [F]
Input variable: zJet2eta
(1/N) dN/ 0.128 F
zJet2eta [F]
Input variable: zJet1mass
(1/N) dN/ 1.59 F
zJet1mass [F]


Input variable: hadZpt
(1/N) dN/ 11.3 F
hadZpt [F]
Input variable: lepZpt
(1/N) dN/ 14.6 F
lepZpt [F]
Input variable: hadZmass
(1/N) dN/ 2.8 F
hadZmass [F]
Input variable: lepZmass
(1/N) dN/ 2.88 F
lepZmass [F]
Input variable: hadZeta
(1/N) dN/ 0.233 F
hadZeta [F]
Input variable: lepZeta
(1/N) dN/ 0.187 F
lepZeta [F]
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


Input variable: lepZjet3DeltaPhi
(1/N) dN/ 0.0532 F
lepZjet3DeltaPhi [F]
Input variable: lepZjet3DeltaR
(1/N) dN/ 1.69e+09 F
lepZjet3DeltaR [F]
Input variable: zMuon1pt
(1/N) dN/ 9.87 F
zMuon1pt [F]
Input variable: zMuon2pt
(1/N) dN/ 5.13 F
zMuon2pt [F]
Input variable: zMuon1eta
(1/N) dN/ 0.0813 F
zMuon1eta [F]
Input variable: zMuon2eta
(1/N) dN/ 0.0845 F
zMuon2eta [F]
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


Input variable: scalProdJJsysZhad
(1/N) dN/ 0.0339 F
scalProdJJsysZhad [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%


Input variable: zzDeltaPhi
(1/N) dN/ 0.0532 F
zzDeltaPhi [F]
Input variable: zzDeltaR
(1/N) dN/ 0.143 F
zzDeltaR [F]
Input variable: ptBalance
(1/N) dN/ 6.68 F
ptBalance [F]
Input variable: met
(1/N) dN/ 3.56 F
met [F]
Input variable: sumET
(1/N) dN/ 30.5 F
sumET [F]
Input variable: deltaEtaJJsysZhad
(1/N) dN/ 0.13 F
deltaEtaJJsysZhad [F]
U/O-flow (S,B): |0.0, 0.0|% / |0.0, 0.0|%

A few considerations

- The variables shapes seem more or less the same when using the low mass Z+jet sample (despite the spikes...)
  - The only notable difference is the leptonic Z mass (of course).
  - This cut does not seem to be strong for the other backgrounds but should be used to remove this one, which has a very high cross section with respect to the others.

Had Z selection, most significant variables (still preliminary)

- For the hadroinc Z identification:
  - hadZmass
  - hadZpt
  - sumET
  - zJetsDeltaR
- For the full system
  - tagBDToutput
  - lepZmass
  - LepZpt

Other possible variables

- Consider using the value of the btagging variable for both hadZ jets
- Consider eventually the alpha_t variable (see the SUSY pas SUS-10-001) to remove backgrounds with real MEt