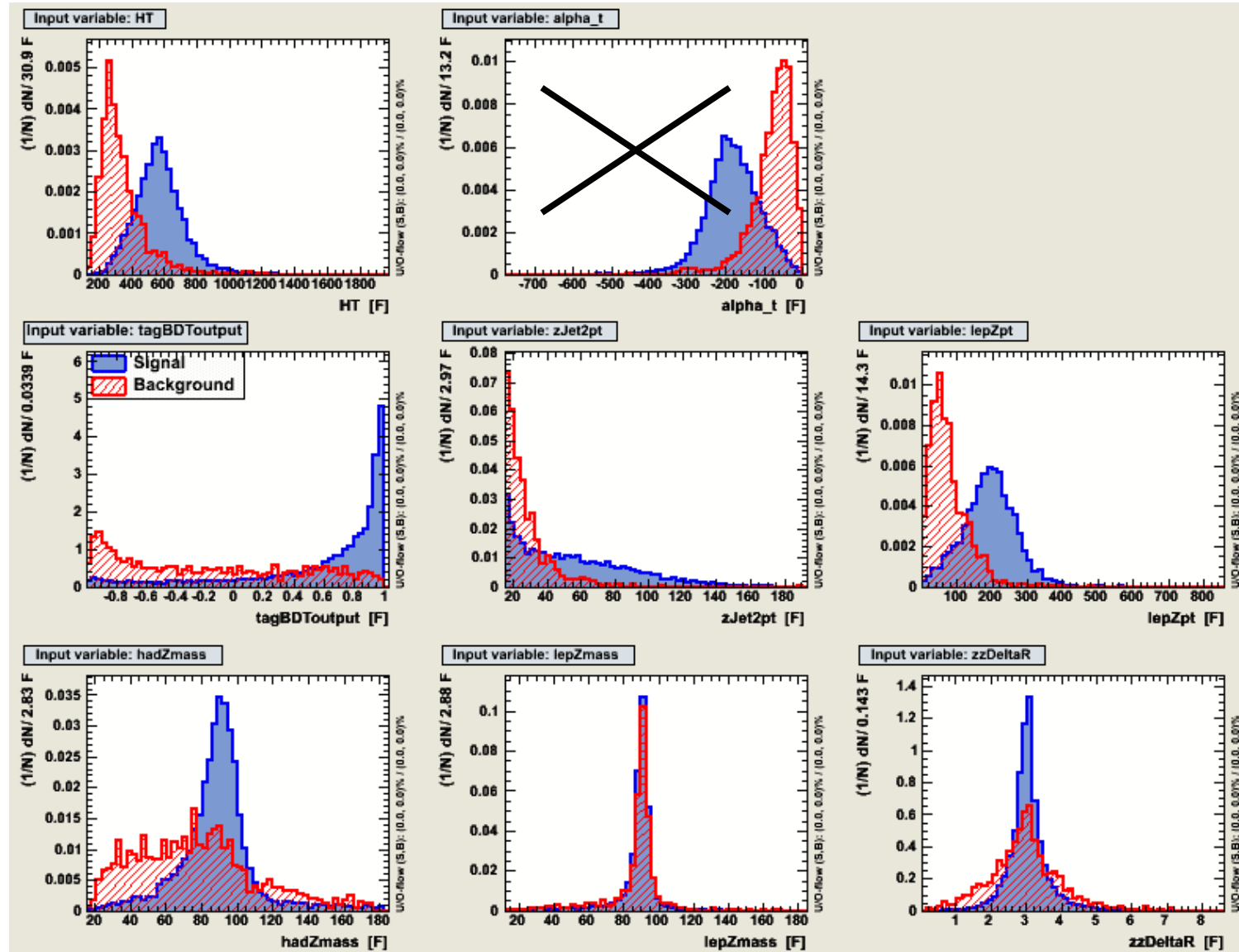# Choice of a subset of variables

- Variables were chosen according to classification and correlation
  - Test different sets of variables and look at the output significance returned by TMVA as a measure of goodness:
    - All: 2.569
    - Reduced to 12: 2.439
    - Reduced to 9: 2.468
    - Reduced to 7: 2.470
    - Reduced to 5: 2.039
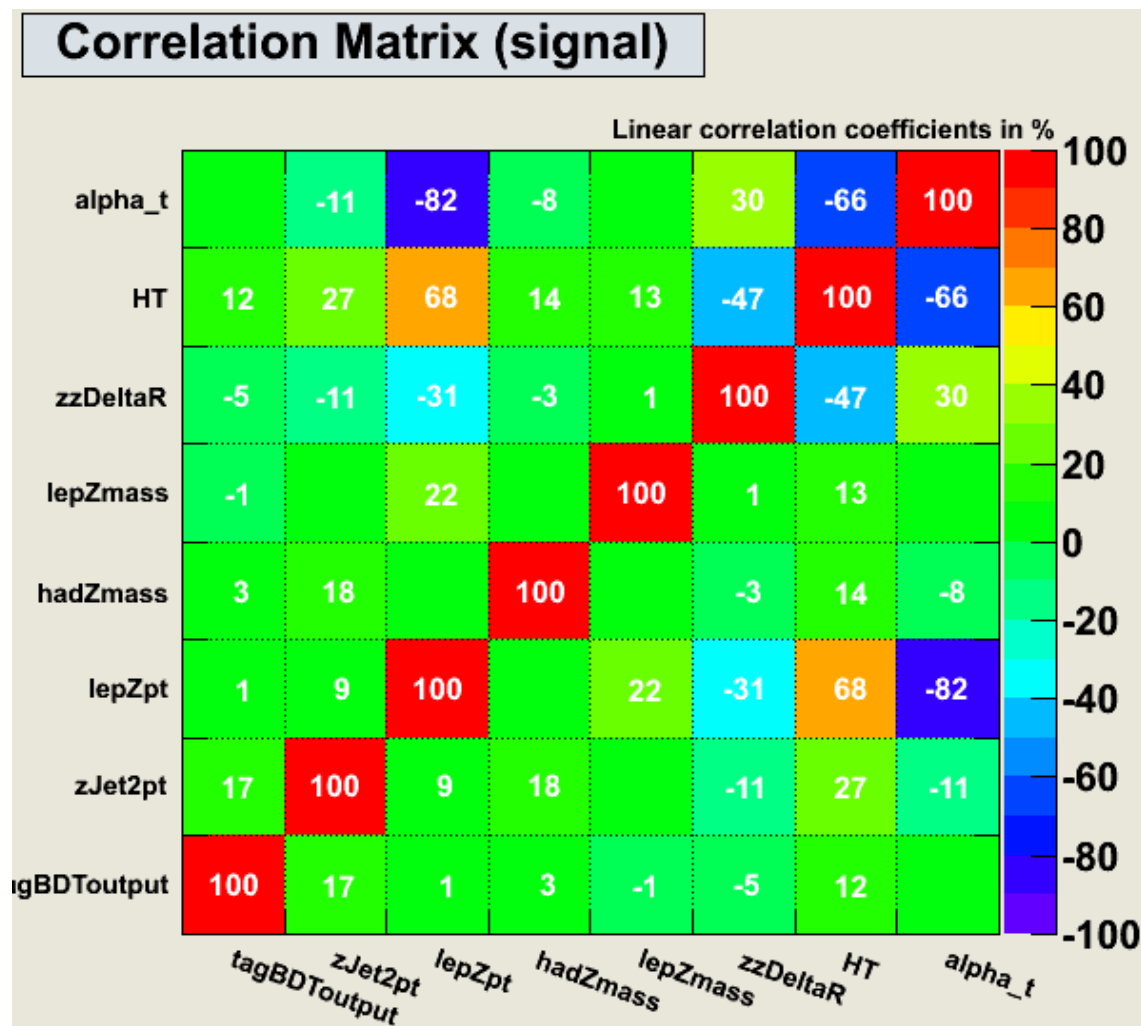 - The best compromise seems to be the set of 7 (see next slide)

Training for mH = 500 GeV

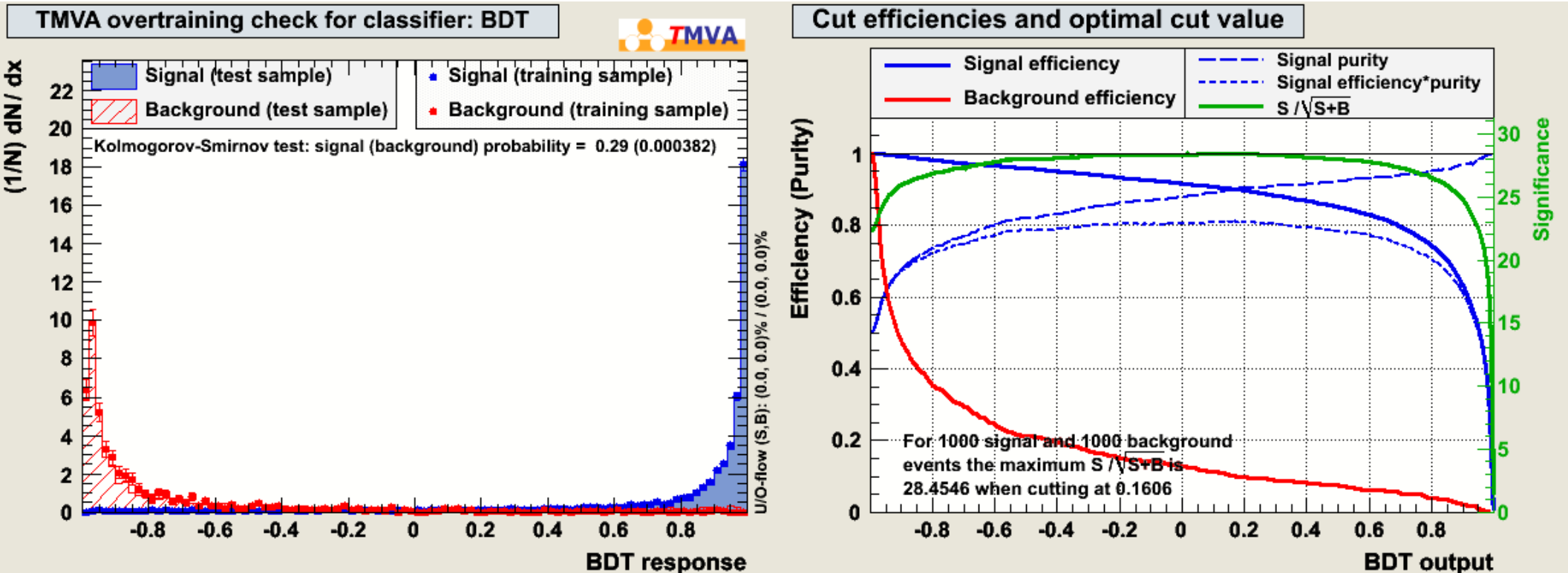- Selected variables:
  - alpha_t not included in the training

Linear correlation coefficients

- Alpha_t is very anticorrelated with HT and leptonic Z pt

  - Of course! It does not include the muons!

  - Coult try to redo it, maybe it will reduce the correlation and become more important...

Correlation Matrix (signal)
Linear correlation coefficients in %

Removed alpha_t, training result

- Good performance on test sample (half the statistics)

ROC

Background rejection versus Signal efficiency