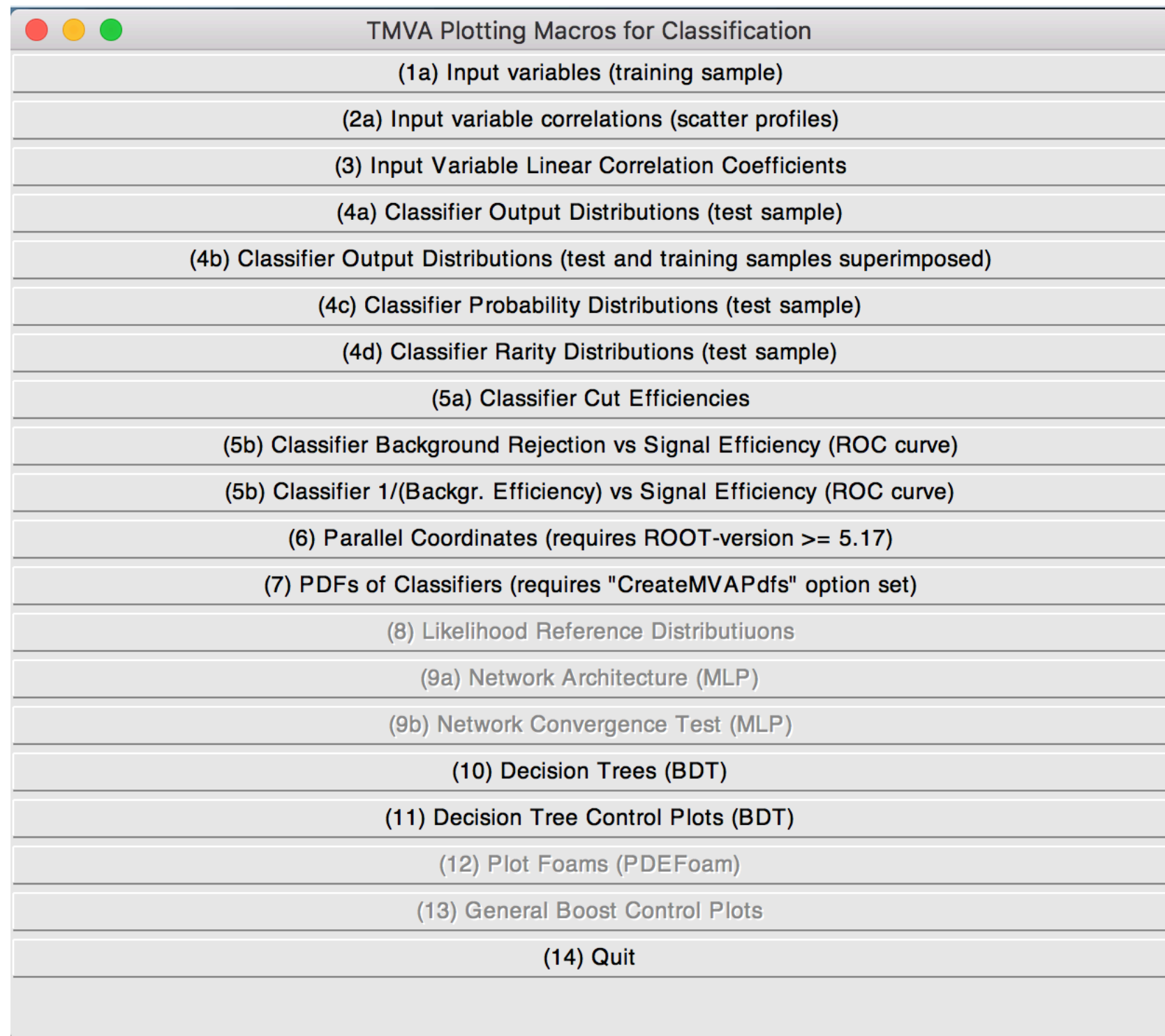# TMVA: BDT trained on RecoTk1IP from simulated ttbar samples
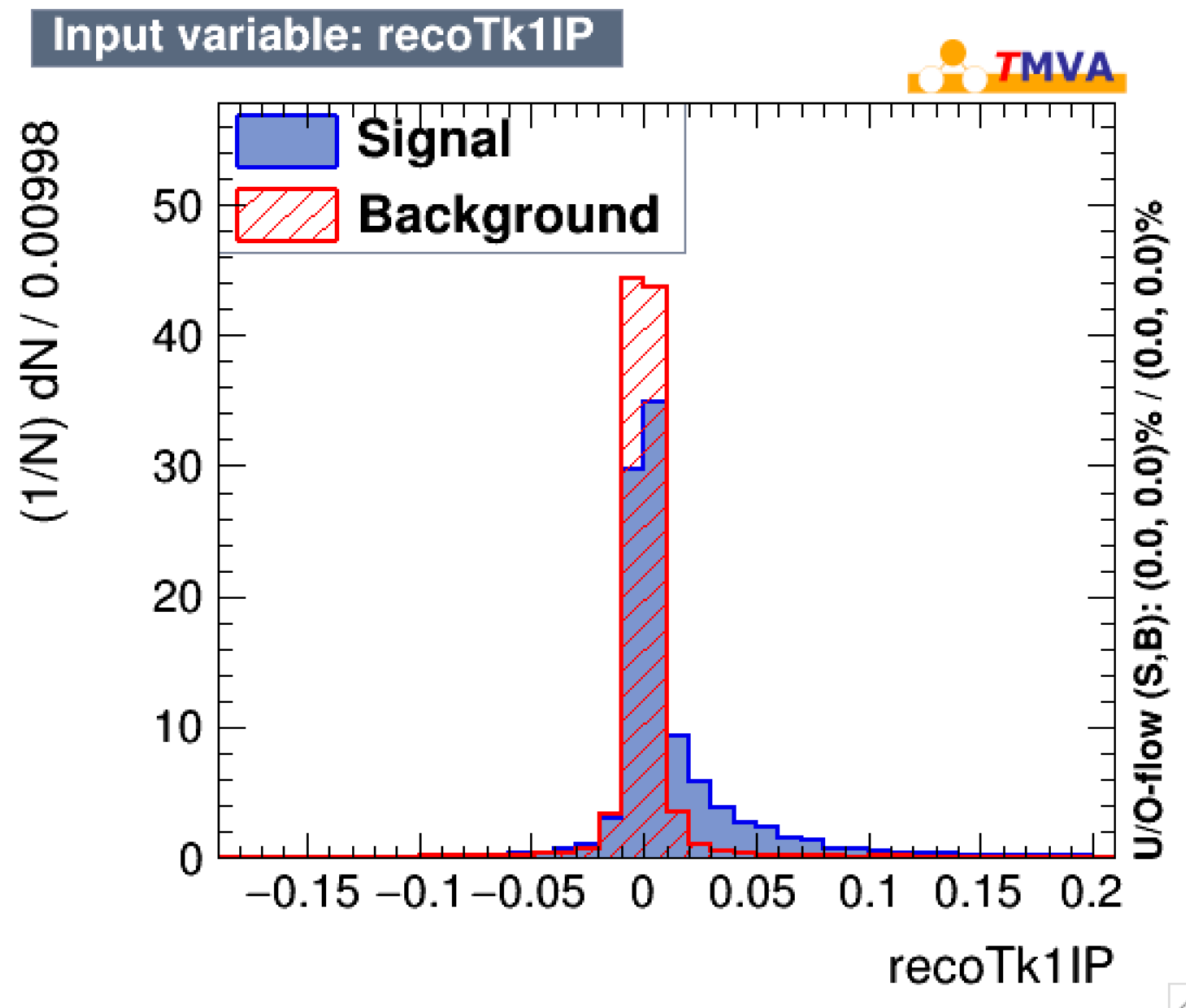
Aug 3, 2018

All plots were generated using the TMVA GUI (one click of a button!)

This slideshow just demonstrates how some of these plots look, for a BDT trained only on Track 1 2D IP.
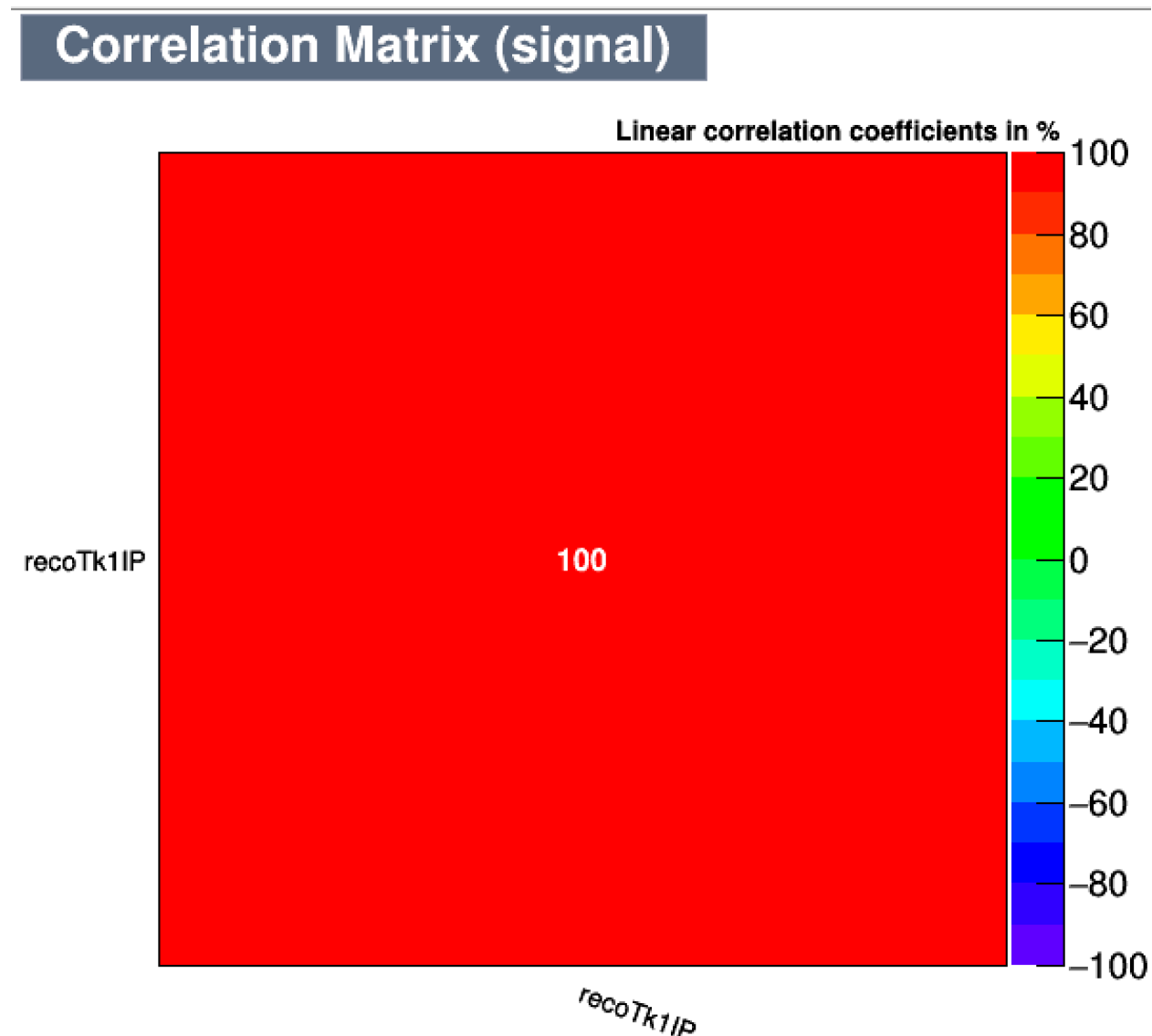
TMVA Plotting Macros for Classification

- (1a) Input variables (training sample)
- (2a) Input variable correlations (scatter profiles)
- (3) Input Variable Linear Correlation Coefficients
- (4a) Classifier Output Distributions (test sample)
- (4b) Classifier Output Distributions (test and training samples superimposed)
- (4c) Classifier Probability Distributions (test sample)
- (4d) Classifier Rarity Distributions (test sample)
- (5a) Classifier Cut Efficiencies
- (5b) Classifier Background Rejection vs Signal Efficiency (ROC curve)
- (5b) Classifier 1/(Backgr. Efficiency) vs Signal Efficiency (ROC curve)
- (6) Parallel Coordinates (requires ROOT-version >= 5.17)
- (7) PDFs of Classifiers (requires "CreateMVAPdfs" option set)
- (8) Likelihood Reference Distributiuons
- (9a) Network Architecture (MLP)
- (9b) Network Convergence Test (MLP)
- (10) Decision Trees (BDT)
- (11) Decision Tree Control Plots (BDT)
- (12) Plot Foams (PDEFoam)
- (13) General Boost Control Plots
- (14) Quit

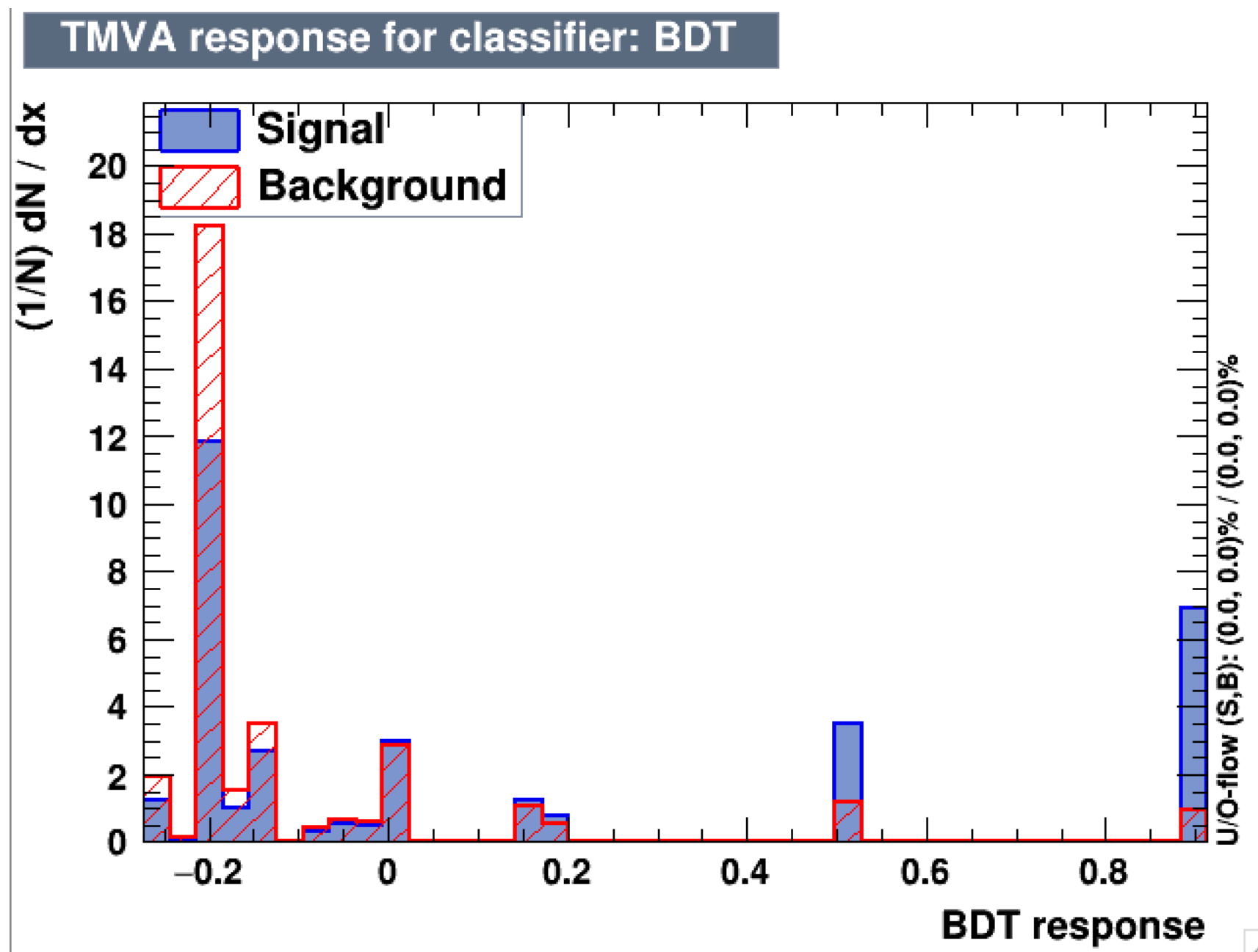# 1(a) Input variables (training sample)

(3) Input Variable Linear Correlation Coefficients

- Only one input variable: correlation coefficients are N/A
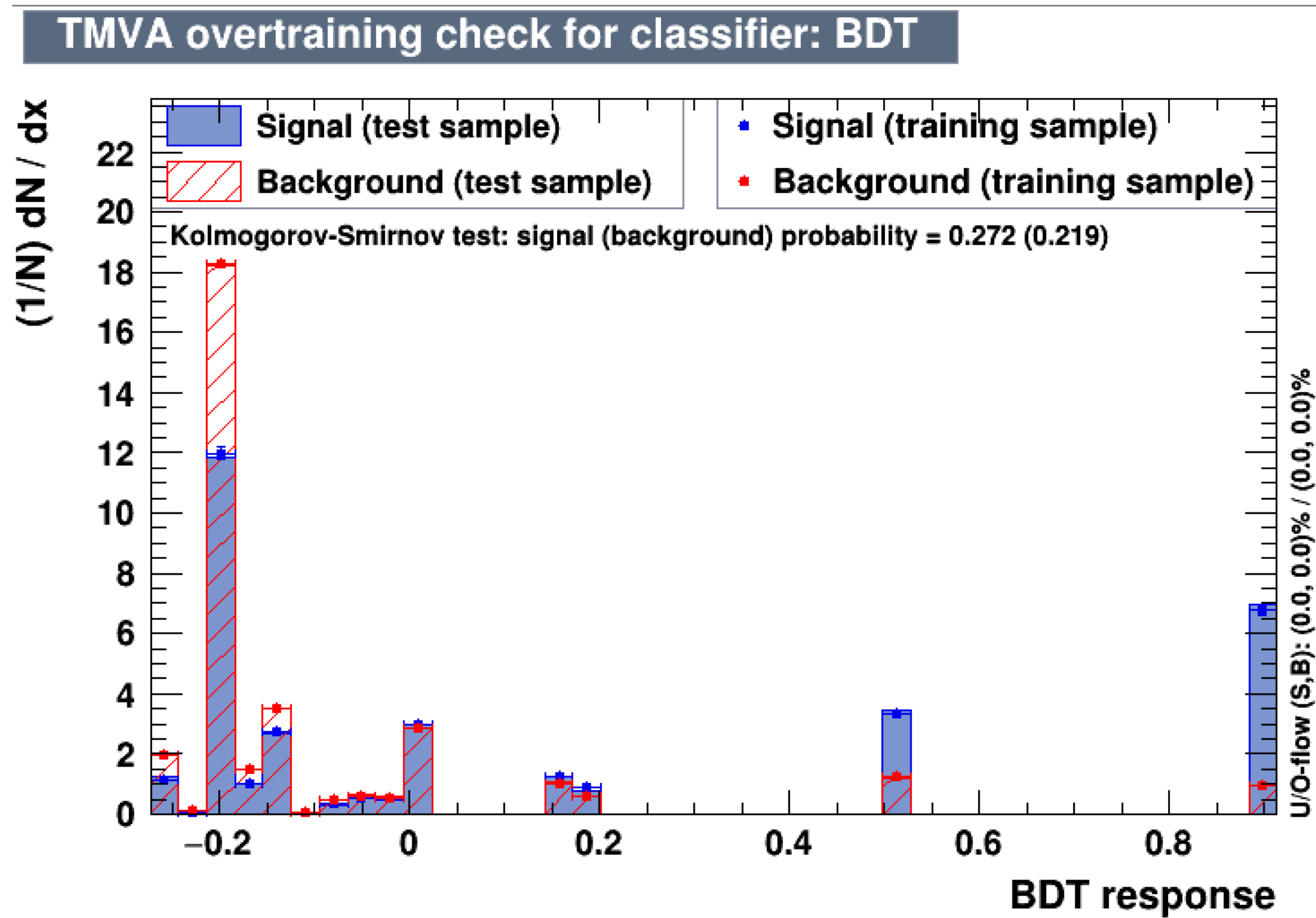
(4a) Classifier Output Distributions (test sample)

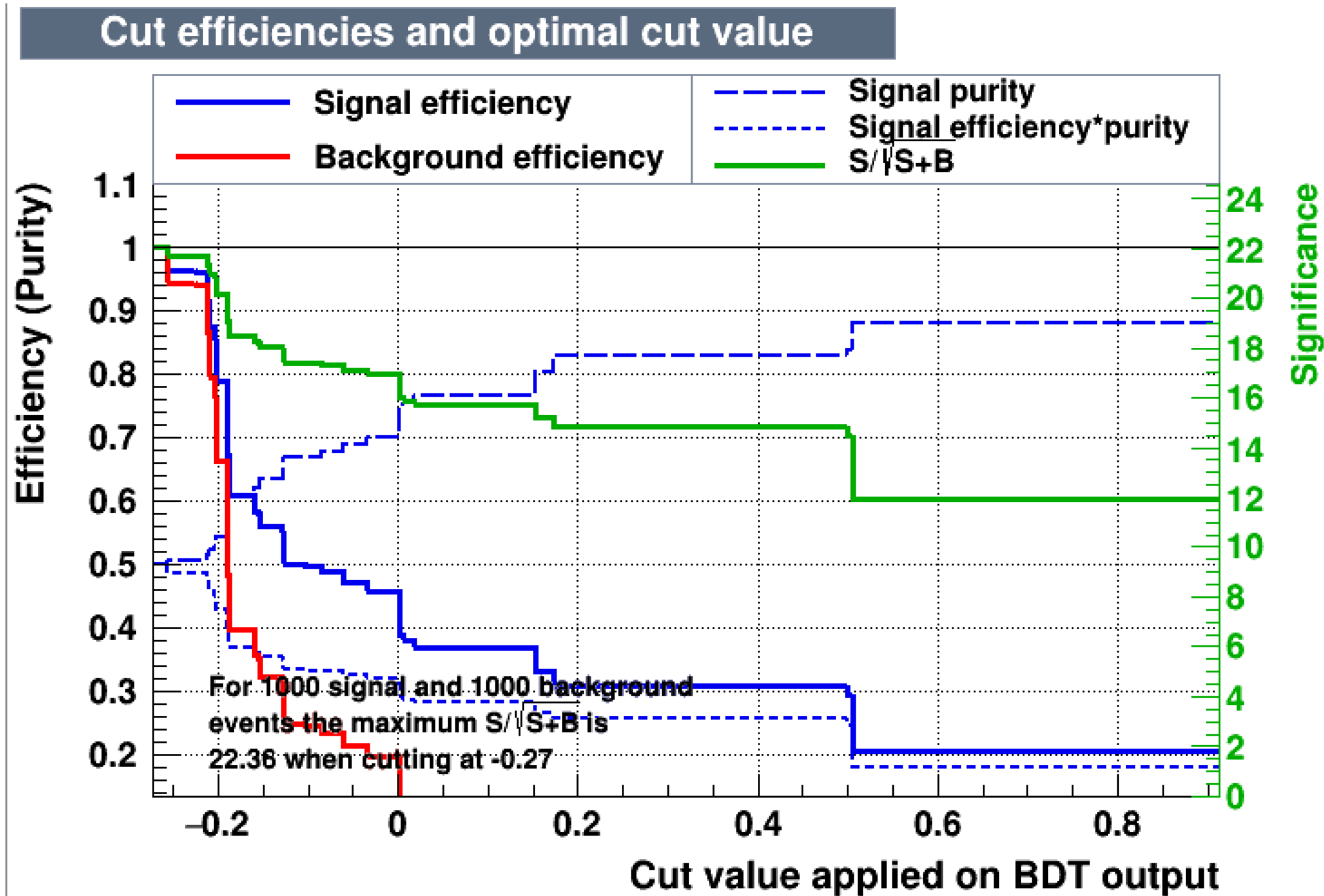- More signal events have a higher BDT classifier value as expected

(4b) Classifier Output Distributions (test and training samples superimposed)

- A warning sign of overtraining is when the training/test BDT responses don't match up
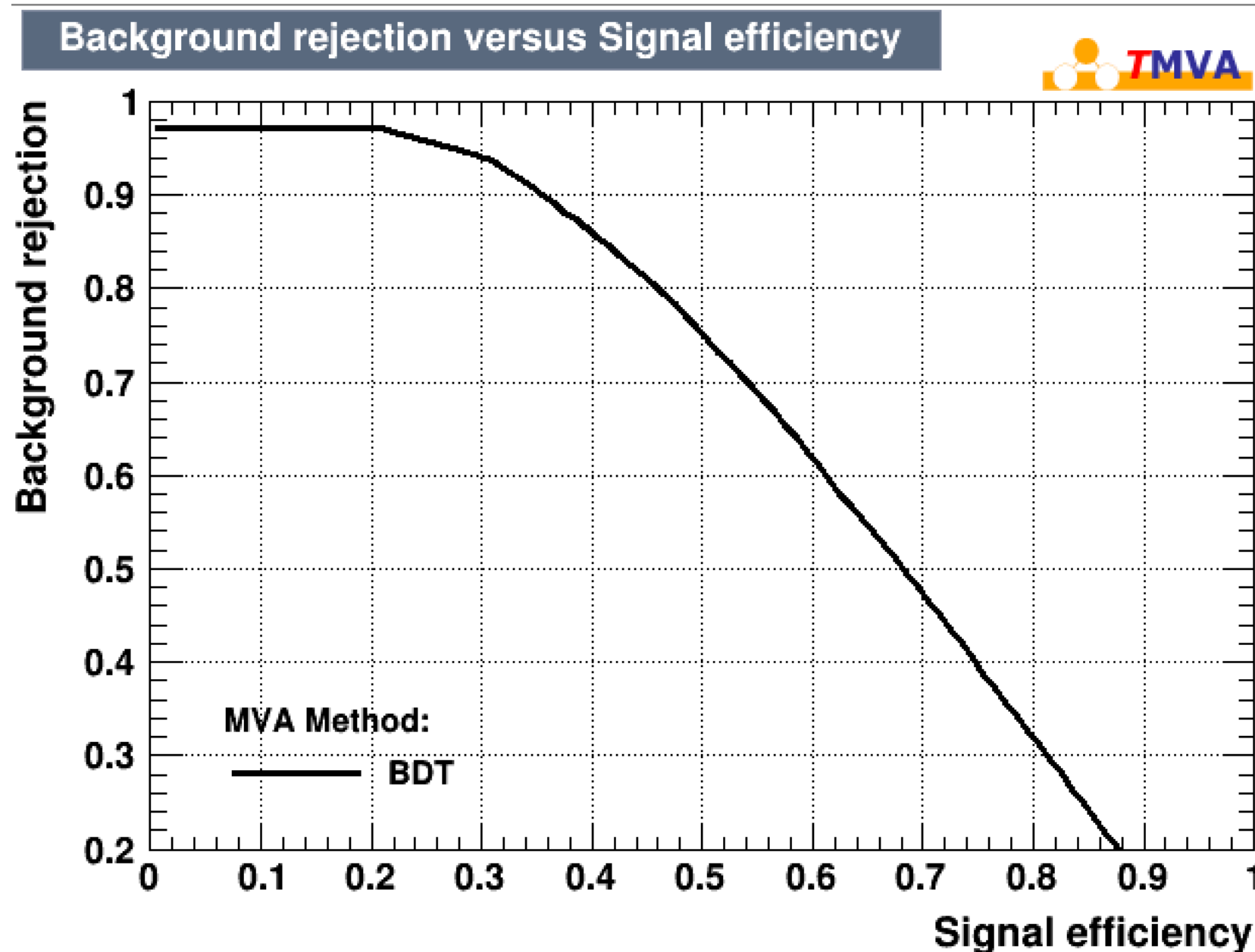
(5a) Classifier Cut Efficiencies

Cut efficiencies and optimal cut value

Legend:
- Signal efficiency
- Background efficiency
- Signal purity
- Signal efficiency*purity
- $S/\sqrt{S+B}$

For 1000 signal and 1000 background events the maximum $S/\sqrt{S+B}$ is 22.36 when cutting at -0.27

Efficiency (Purity) — vertical axis (left)
Significance — vertical axis (right)
Cut value applied on BDT output — horizontal axis

(5b) Classifier Background Rejection vs Signal Efficiency (ROC curve)
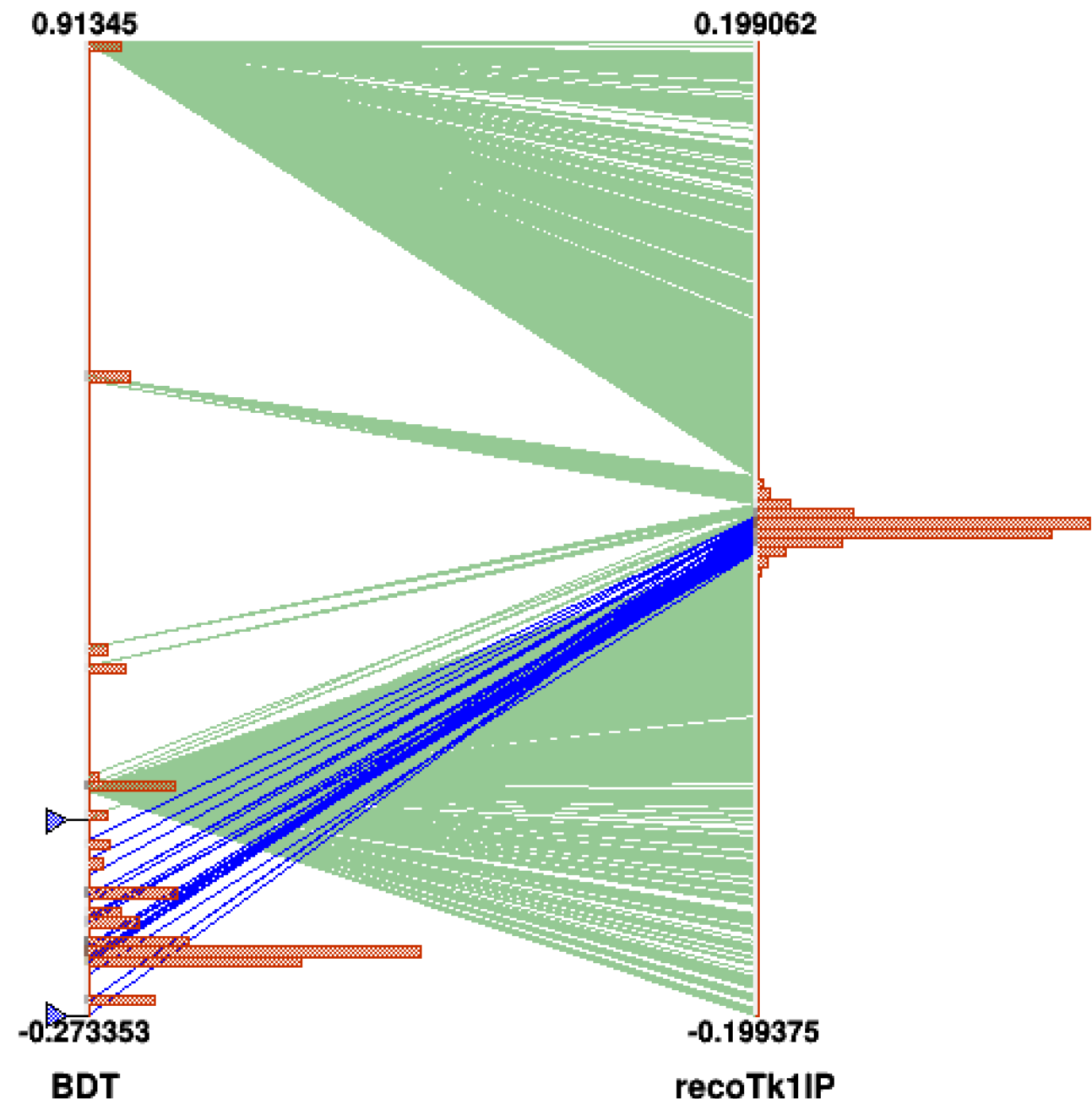
- More detailed metrics are in the logRun file

(6) Parallel Coordinate representation

BDT:recoTk1IP{classID==1}

- Mapping of recoTk1IP values to BDT discriminator

- This gave me a segfault when I clicked on it :(

(7) PDFs of Classifiers (requires "CreateMVAPdfs" option set)

- (Still trying to set "CreateMVAPDFs" option)

(10) Decision Trees (BDT)

(11) Decision Tree Control Plots (BDT)