$B_s^0 \to \mu^+\mu^-$ Cross Check Analysis

Purdue, Pisa

April 12, 2013

Abstract

The aim of this analysis note is to provide supporting material for the cross check analysis for the search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays. It follows closely the documentation provided in [1] and includes direct comparison to the results in that note.

Contents

1 Introduction **3**

Part I (HCP analysis) **3**

2 Datasets **3**

3 Selection **3**
3.1 pre-selection . . . 3
3.2 muon identification . . . 3
3.3 variable distributions . . . 3
3.4 variable ranking and correlations . . . 3

4 Boosted Decision Tree **16**

5 Cut and count analysis **18**
5.1 optimization and blinded results . . . 18
5.2 unblind . . . 18

6 BDT analysis **18**
6.1 training and overtraining . . . 18
6.2 blinded results . . . 18
6.3 unblind . . . 18

7 Normalization channel **24**
7.1 datasets . . . 24
7.2 selection . . . 24
7.3 BDT . . . 24
7.4 yields . . . 24

Part II (full updated analysis) **25**

8 Full dataset **25**

9 Selection **25**
9.1 datasets . . . 25
9.2 muon identification . . . 25
9.3 variable distributions, correlations, ranking . . . 25
9.4 TMVA training . . . 25
9.4.1 MLP . . . 25
9.4.2 BDT . . . 25
9.5 Normalization channel . . . 25
9.5.1 MLP . . . 25
9.5.2 BDT . . . 25
9.6 Limits . . . 25

10 Summary **25**

Official MC datasets
BsToMuMu_BsFilter_8TeV-pythia6-evtgen/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM
Data /MuOnia/Run2012A-13Jul2012-v1/AOD /MuOnia/Run2012A-recover-06Aug2012-v1/AOD /MuOnia/Run2012B-13Jul2012-v1/AOD /MuOnia/Run2012C-24Aug2012-v1/AOD /MuOnia/Run2012C-PromptReco-v2/AOD /MuOnia/Run2012C-EcalRecover_11Dec2012-v1/AOD /MuOnia/Run2012D-PromptReco-v1/AOD

1 Introduction

Goals and how the note is organized.

2 Datasets

The datasets used are shown in table **??**.

The events are selected using the same triggers as described in the reference [1].

The analysis is based on a boosted decision tree (BDT) algorithm to select $B_s^0 \to \mu^+\mu^-$ events from the background. A preselection is applied to the data and MC samples to reduce the size and remove outliers that might confuse the BDT training. The signal is taken from the $B_s^0 \to \mu^+\mu^-$ MC sample, while the background is taken from the sidebands in the data sample.
Despite the availability of the full 2012 dataset, in order to allow a direct comparison with the results in [1] this study is limited to runs <= 203002. The dataset is split in three different categories according the remainder of "event number%3".

3 Selection

3.1 pre-selection

The same preselection as in [1] is used.

3.2 muon identification

The muons are chosen to pass global muon prompt tight selection (GM_PT). [to be updated]

3.3 variable distributions

The distributions of the variables used for the BDT training are shown in Figures 1 and 2.

3.4 variable ranking and correlations

Tables **??** and **??** show the ranking of variables before the BDT training.

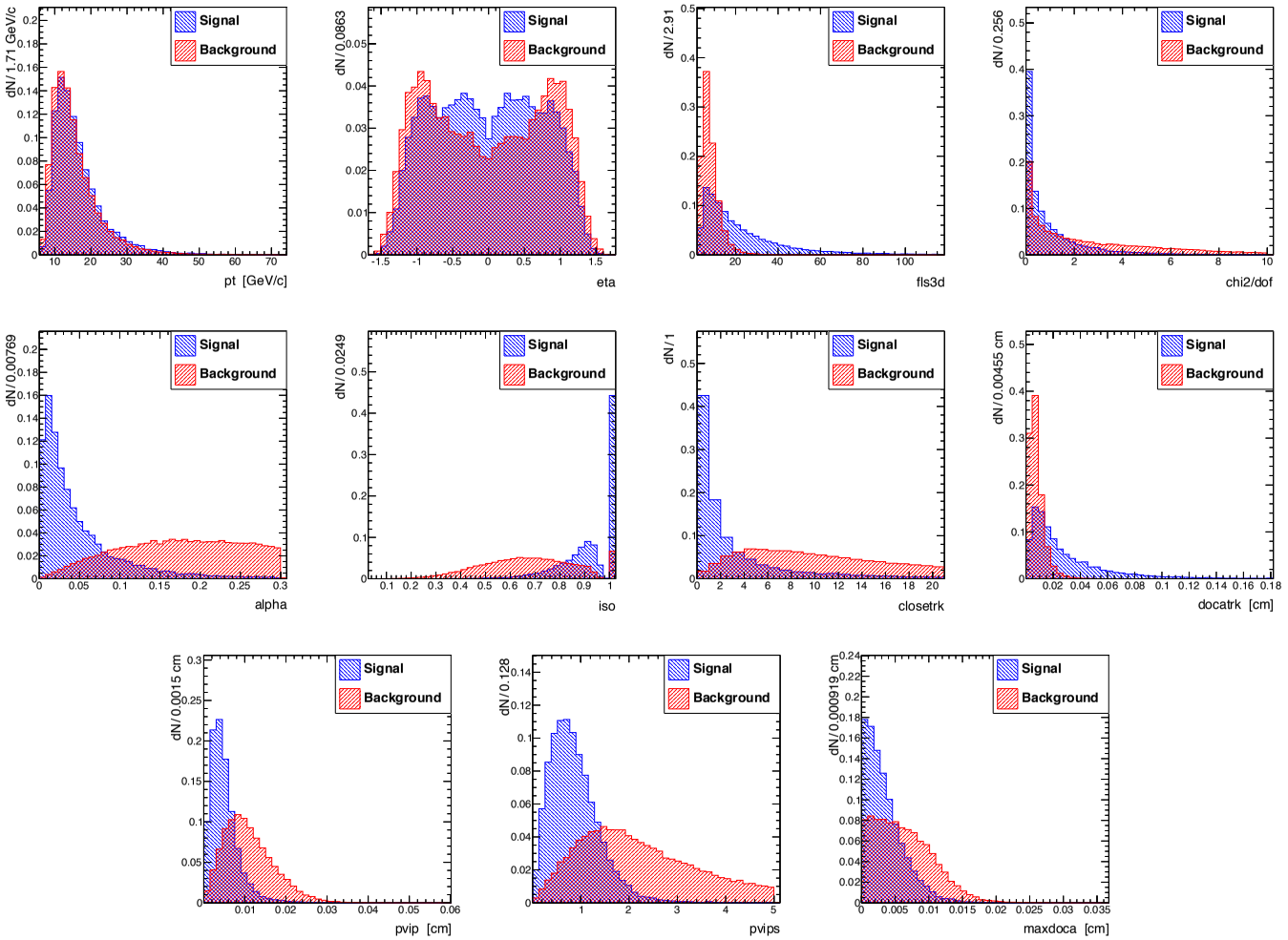

Figure 1: Standard TMVA plot of the input variables for the barrel BDT for signal (blue) and background (red). The background is extracted from data dimuon sidebands.

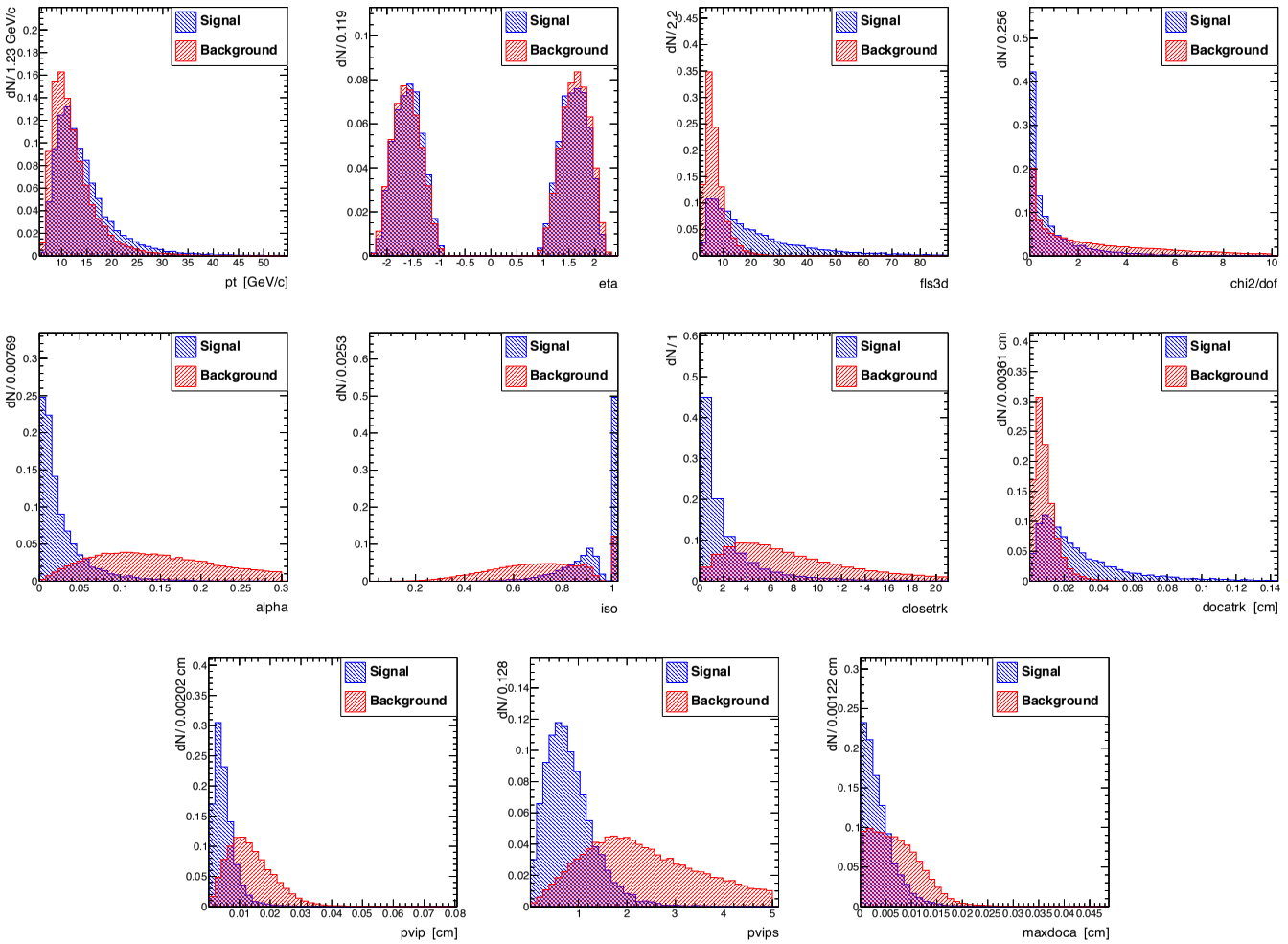

Figure 2: Standard TMVA plot of the input variables for the endcaps BDT for signal (blue) and background (red). The background is extracted from data dimuon sidebands.

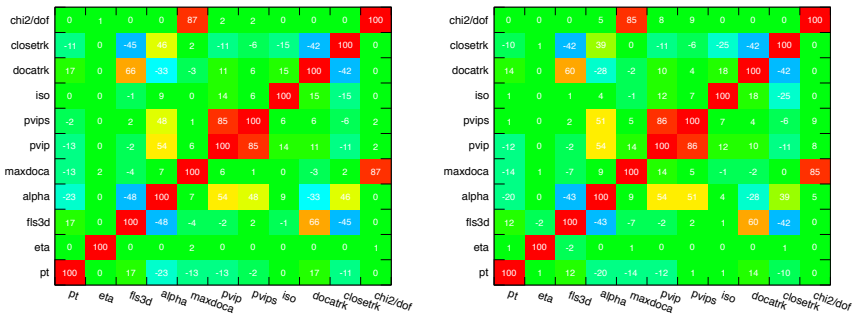

Figure 3: Correlation matrix for signal events in the barrel (left) and the endcap (right).

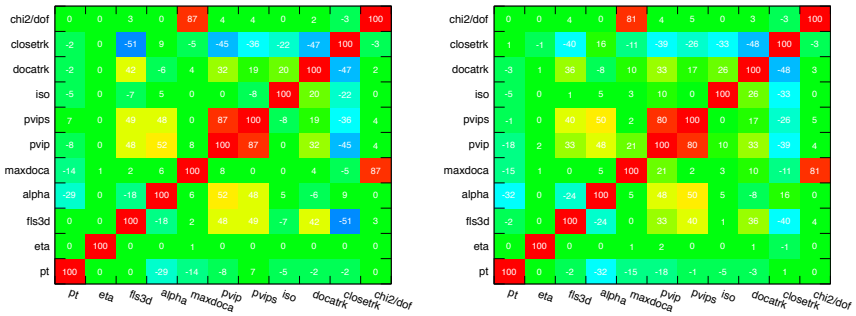

Figure 4: Correlation matrix for background events in the barrel (left) and the endcap (right).

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	alpha	5.319e-01	alpha	5.275e-01	alpha	5.359e-01
2	closetrk	4.440e-01	iso	4.433e-01	iso	4.375e-01
3	iso	4.378e-01	closetrk	4.407e-01	closetrk	4.332e-01
4	pvips	3.723e-01	pvips	3.684e-01	pvips	3.722e-01
5	docatrk	3.333e-01	docatrk	3.299e-01	docatrk	3.285e-01
6	fls3d	3.243e-01	fls3d	3.262e-01	fls3d	3.260e-01
7	pvip	2.974e-01	pvip	2.897e-01	pvip	3.028e-01
8	maxdoca	1.377e-01	maxdoca	1.380e-01	maxdoca	1.409e-01
9	chi2/dof	1.307e-01	chi2/dof	1.320e-01	chi2/dof	1.371e-01
10	eta	9.476e-03	eta	9.818e-03	eta	1.016e-02
11	pt	6.318e-03	pt	6.104e-03	pt	6.269e-03

Table 1: Variable ranking for events of the three different event samples in the barrel before BDT training.

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	alpha	5.601e-01	alpha	5.558e-01	alpha	5.597e-01
2	closetrk	4.273e-01	closetrk	4.290e-01	closetrk	4.352e-01
3	fls3d	4.076e-01	fls3d	4.152e-01	fls3d	4.150e-01
4	iso	3.439e-01	iso	3.464e-01	iso	3.494e-01
5	pvips	3.233e-01	pvips	3.159e-01	pvips	3.239e-01
6	docatrk	3.205e-01	docatrk	3.155e-01	docatrk	3.234e-01
7	pvip	2.786e-01	pvip	2.715e-01	pvip	2.754e-01
8	maxdoca	1.511e-01	maxdoca	1.547e-01	maxdoca	1.549e-01
9	chi2/dof	1.399e-01	chi2/dof	1.401e-01	chi2/dof	1.426e-01
10	pt	3.847e-02	pt	4.095e-02	pt	3.861e-02
11	eta	3.436e-03	eta	2.581e-03	eta	2.434e-03

Table 2: Variable ranking for events of the three different event samples in the endcaps before BDT training.

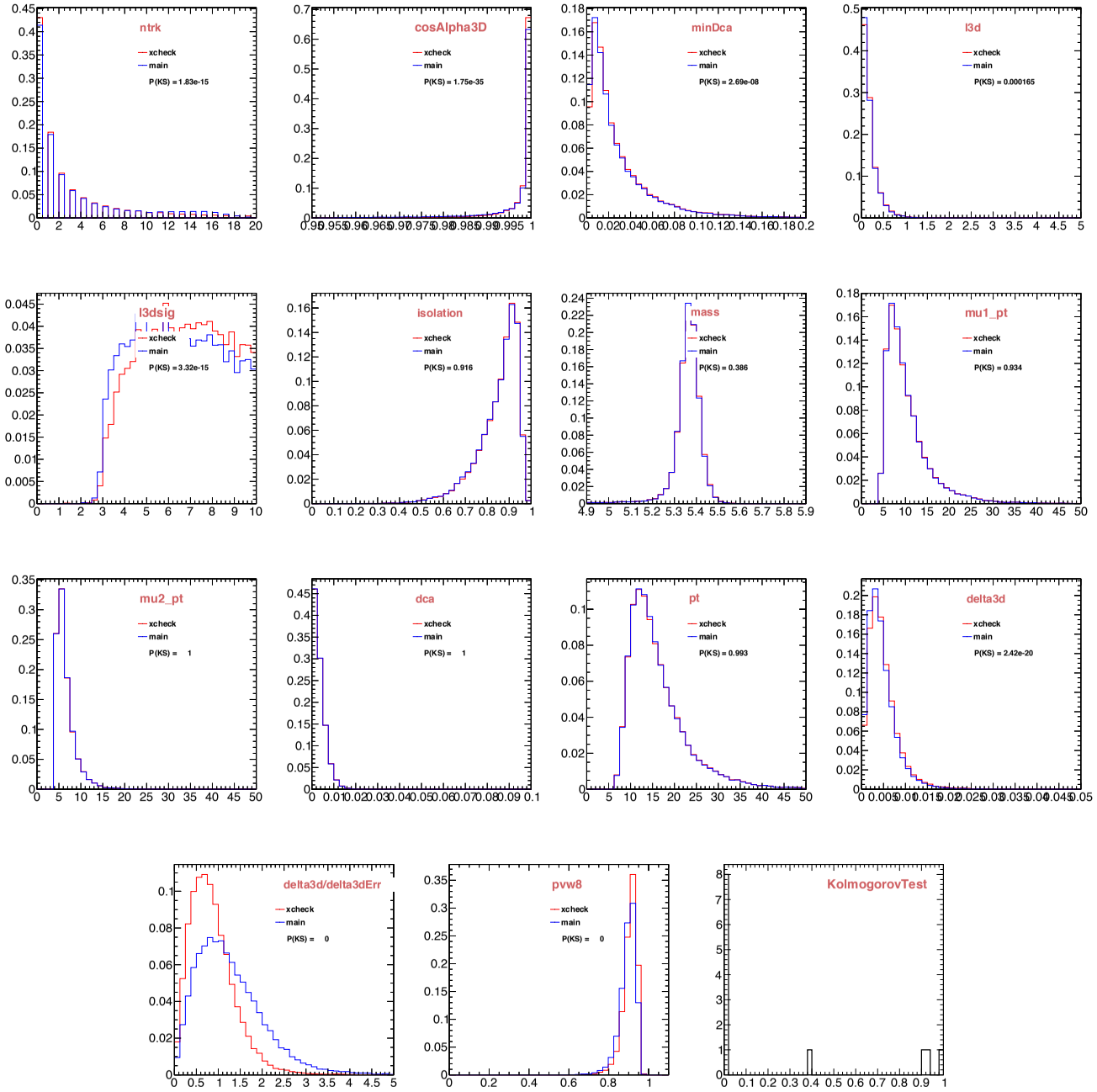

Figure 5: Variable comparisons between the main analysis and the cross-check analysis. Part I: MC barrel.

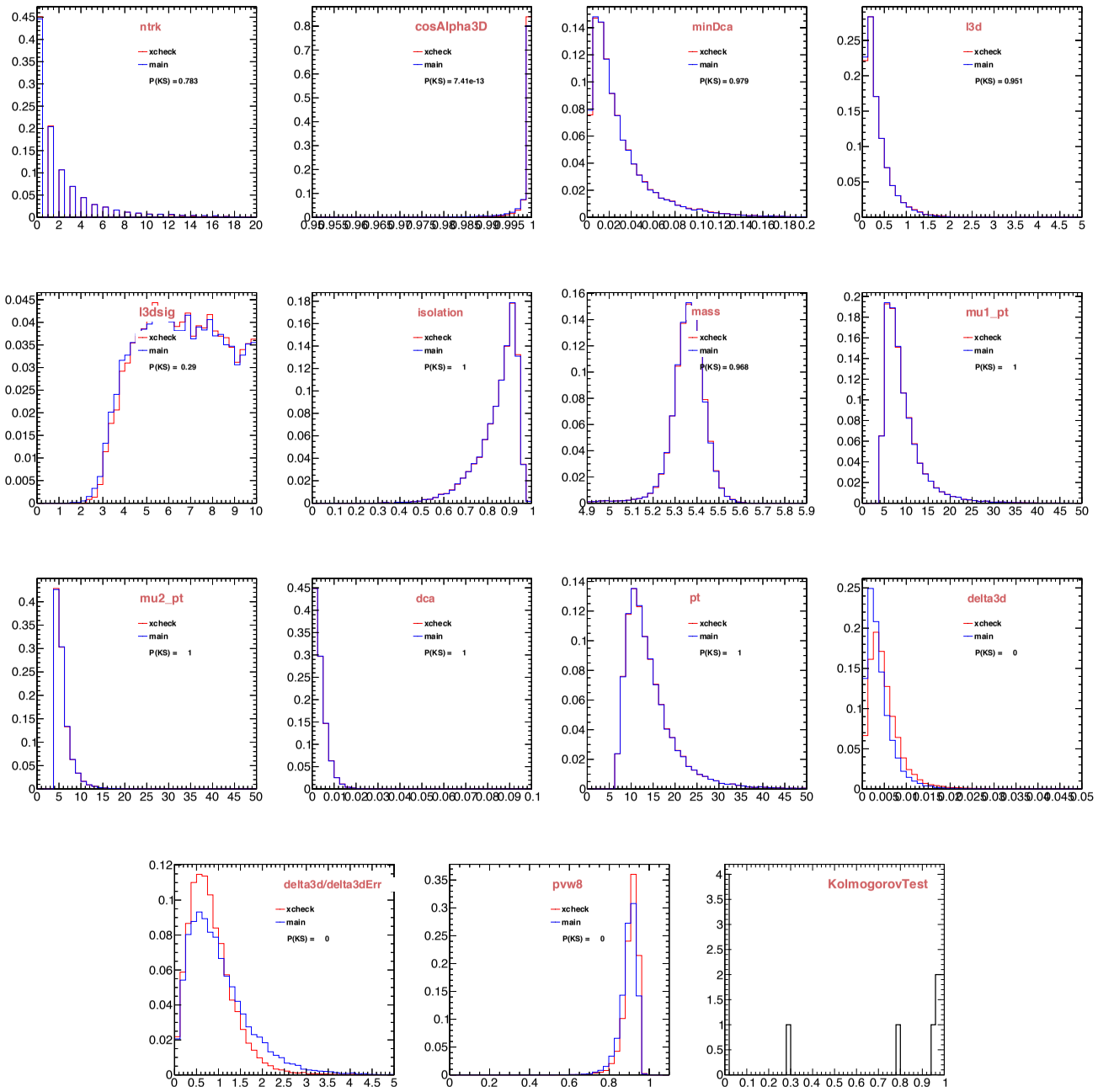

Figure 6: Variable comparisons between the main analysis and the cross-check analysis. Part I: MC endcaps.

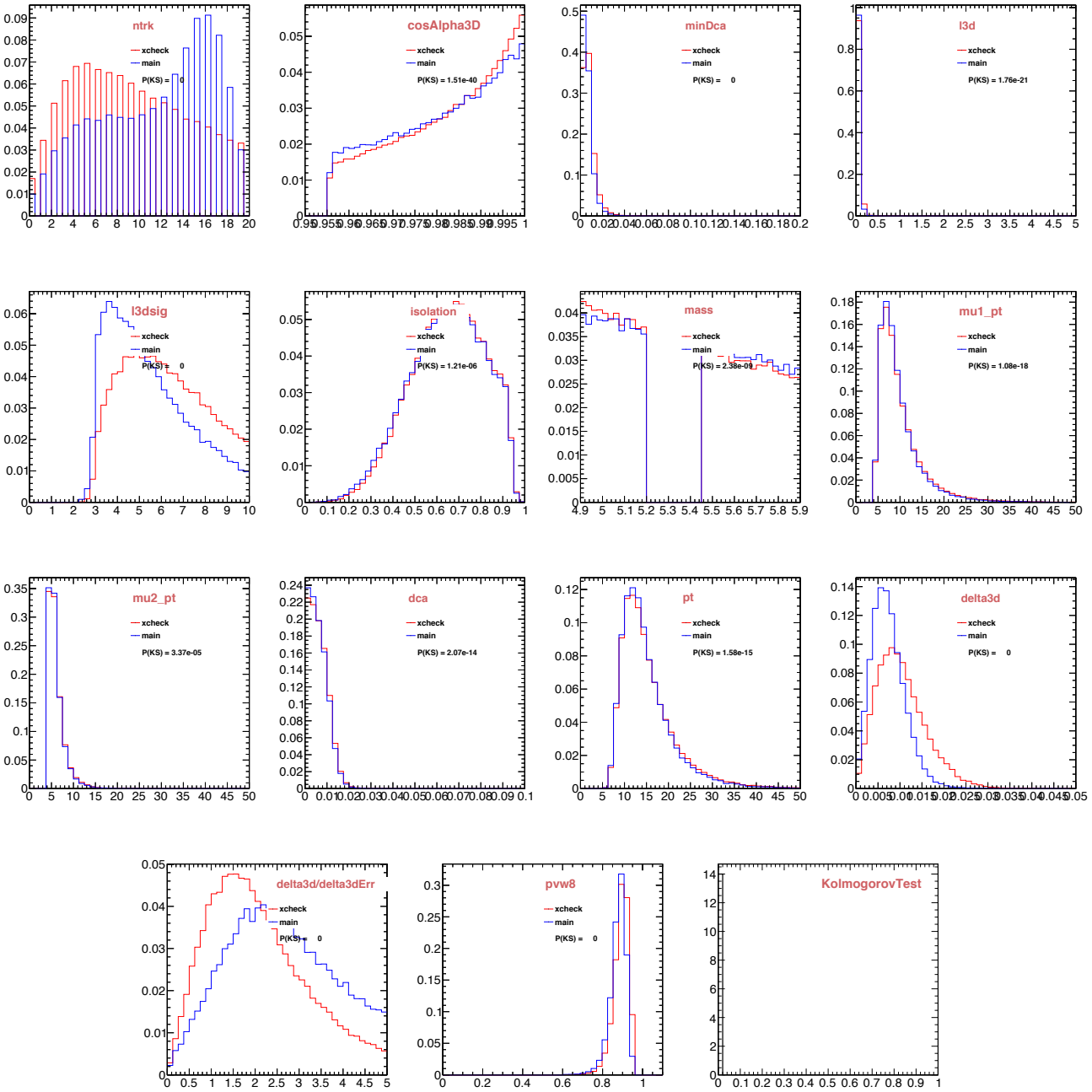

Figure 7: Variable comparisons between the main analysis and the cross-check analysis. Part I: Data barrel.

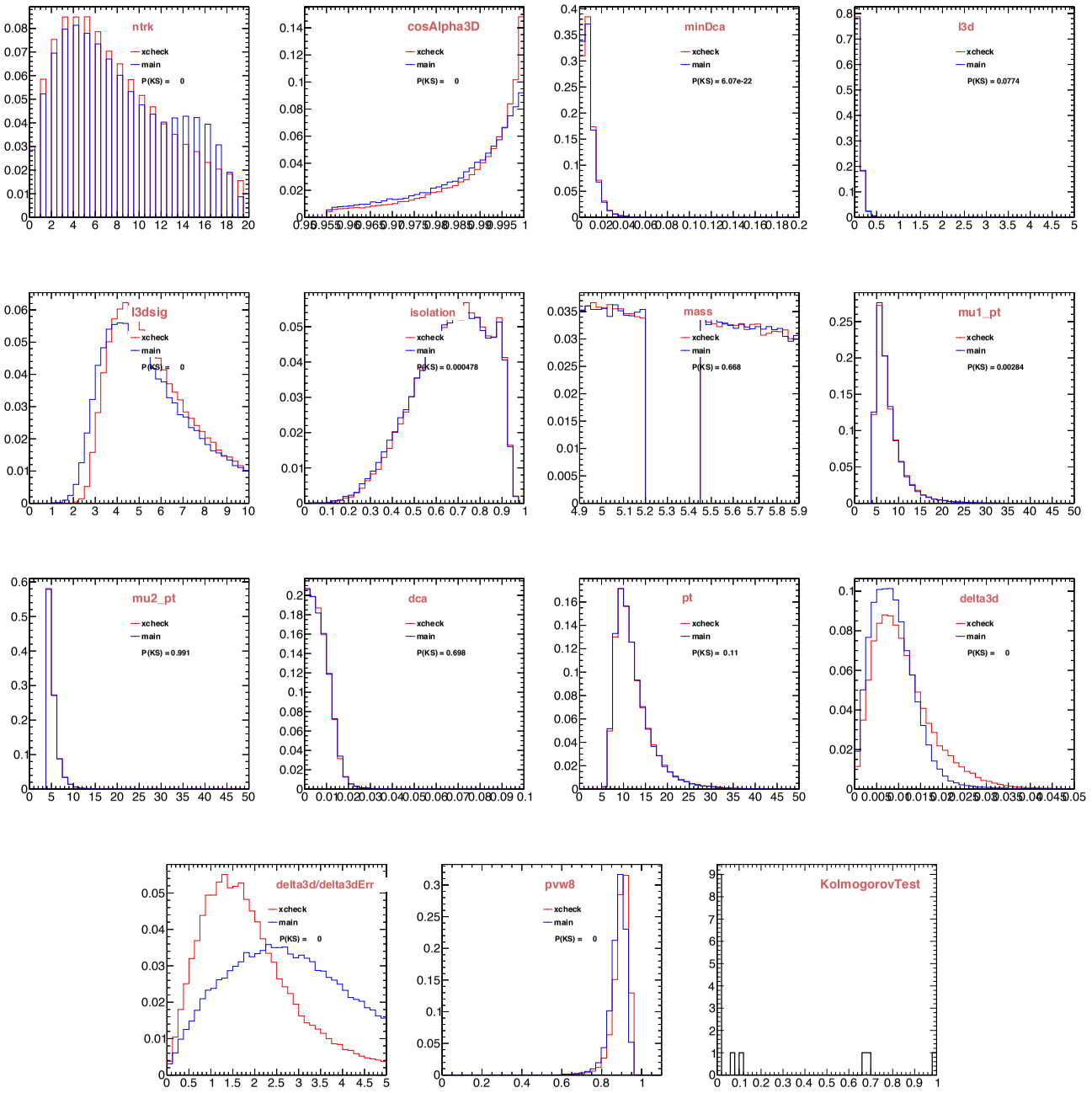

Figure 8: Variable comparisons between the main analysis and the cross-check analysis. Part I: Data endcaps.

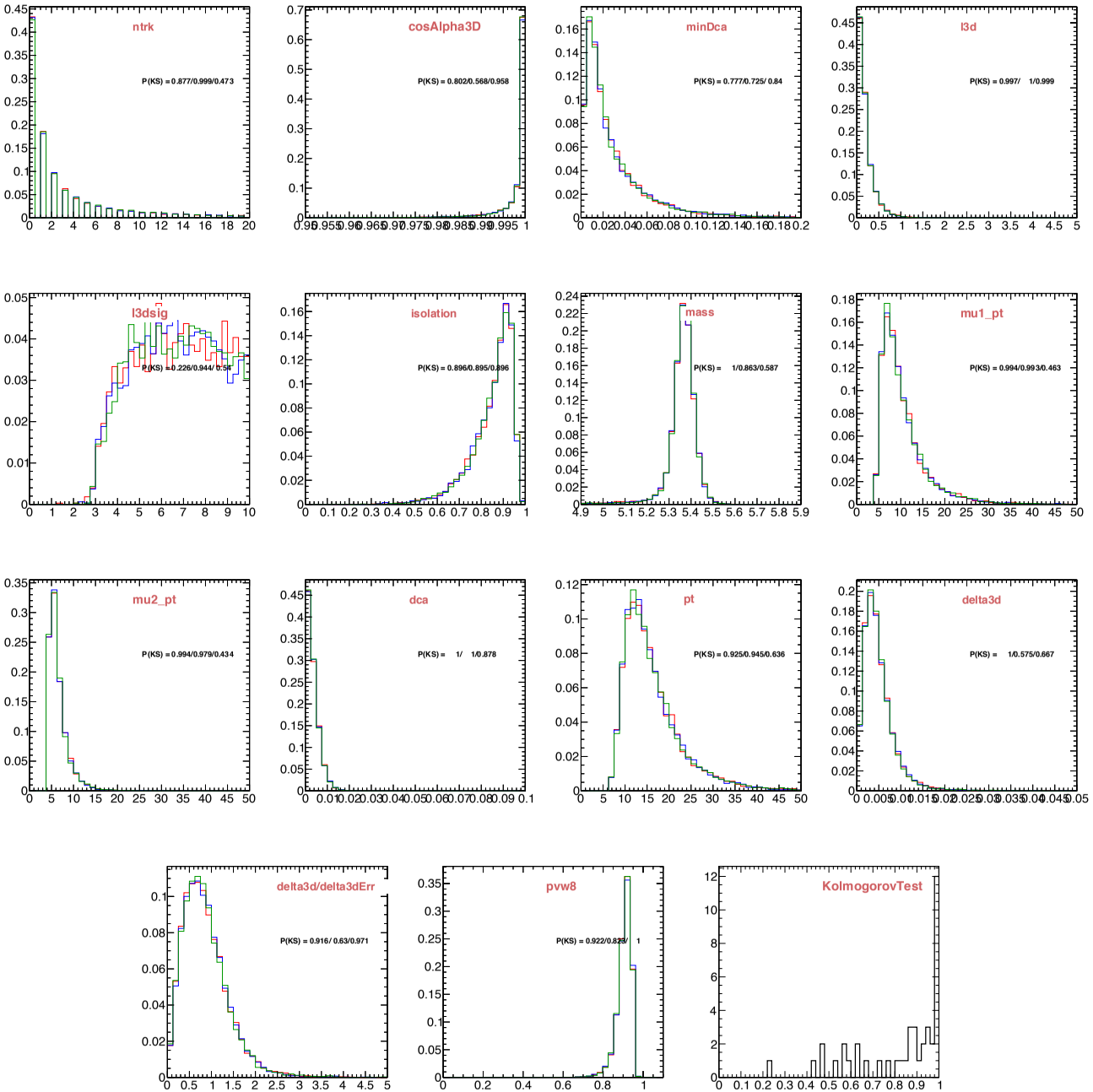

Figure 9: Overlay of BDT training variable distributions in Signal MC for events of the three subsets in the barrel. The plot on the bottom right summarizes all KS probablities.

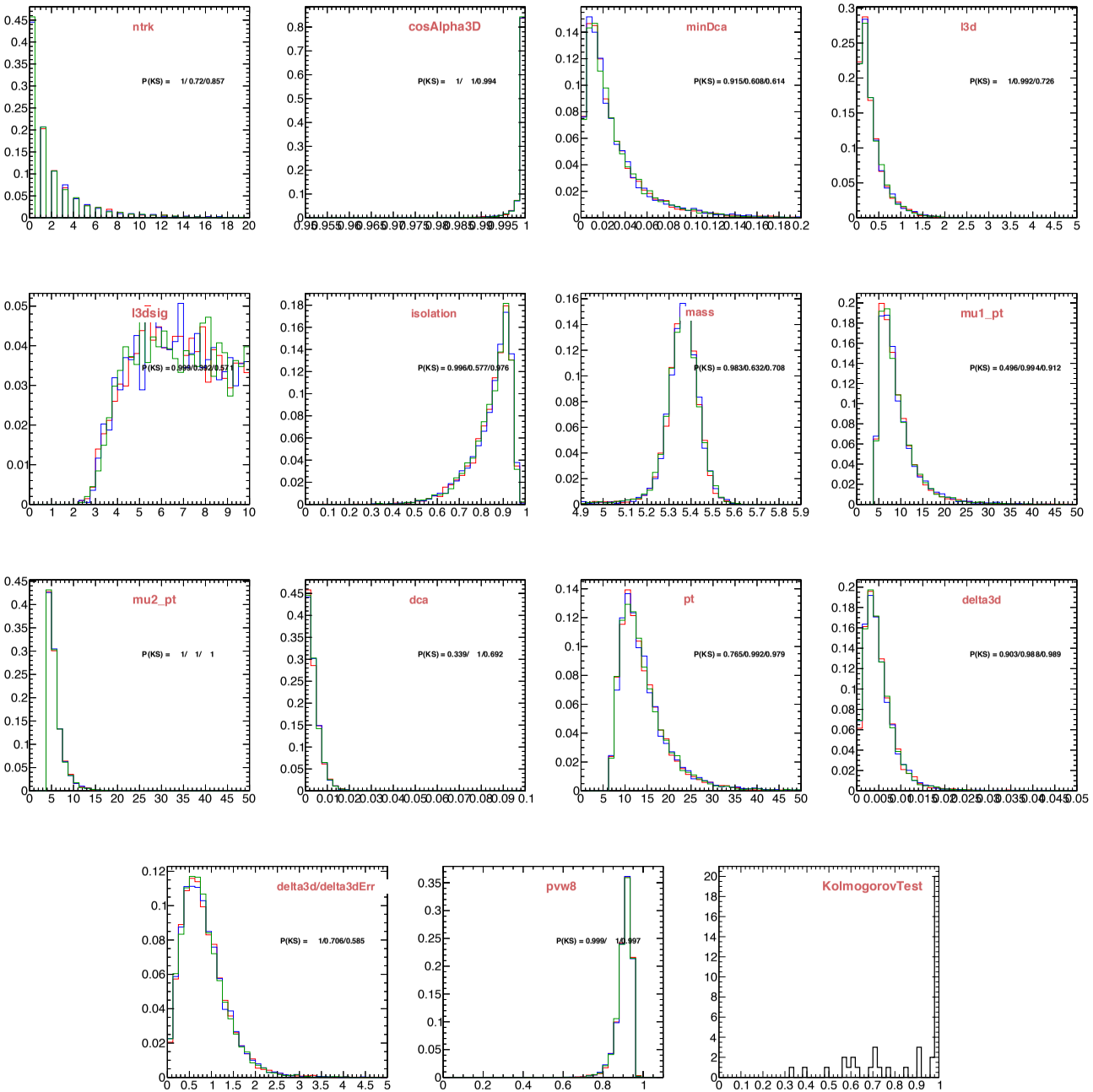

Figure 10: Overlay of BDT training variable distributions in Signal MC for events of the three subsets in the endcap. The plot on the bottom right summarizes all KS probablities.

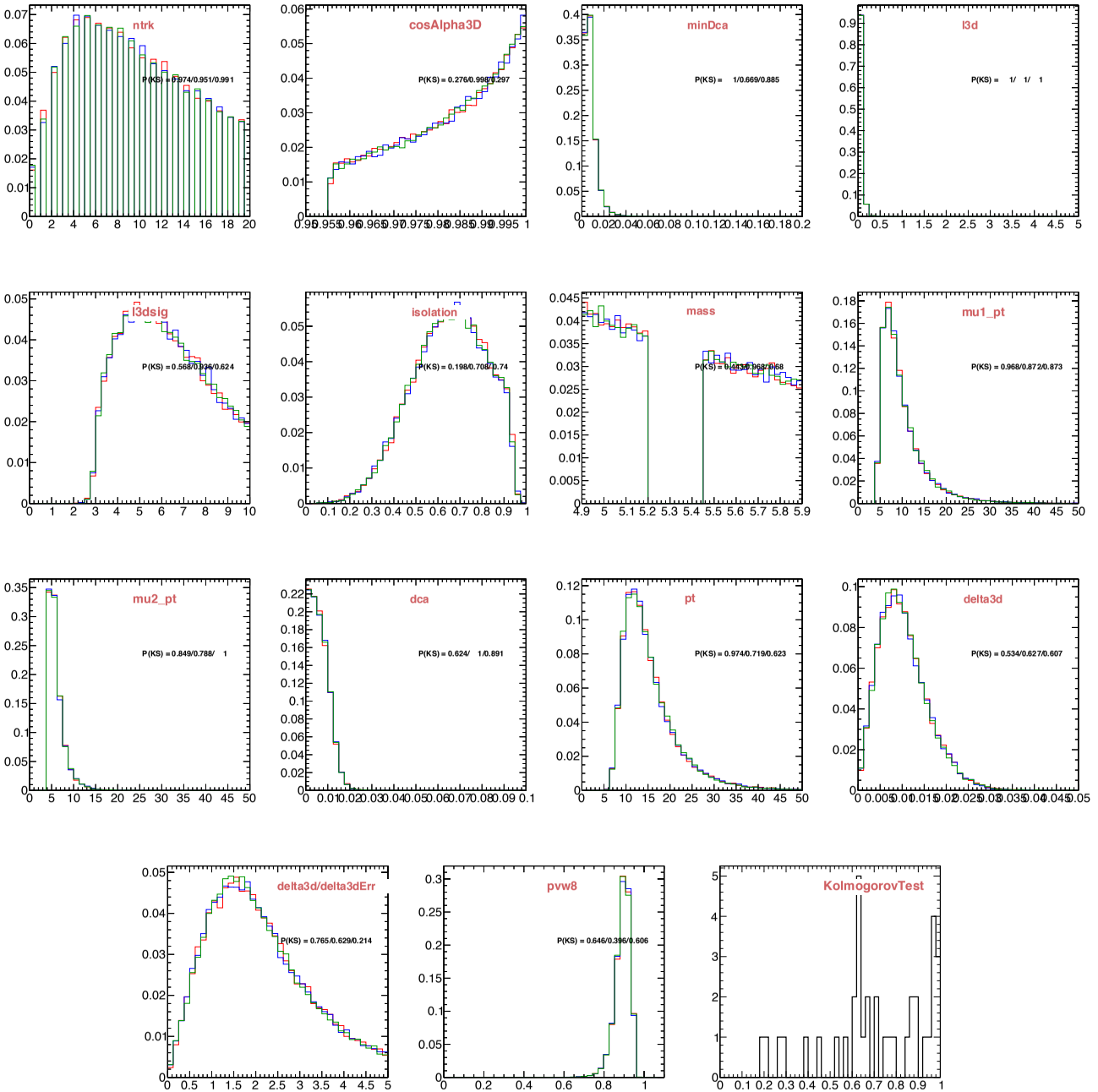

Figure 11: Overlay of BDT training variable distributions in data sideband background for events of the three subsets in the barrel. The plot on the bottom right summarizes all KS probablities.

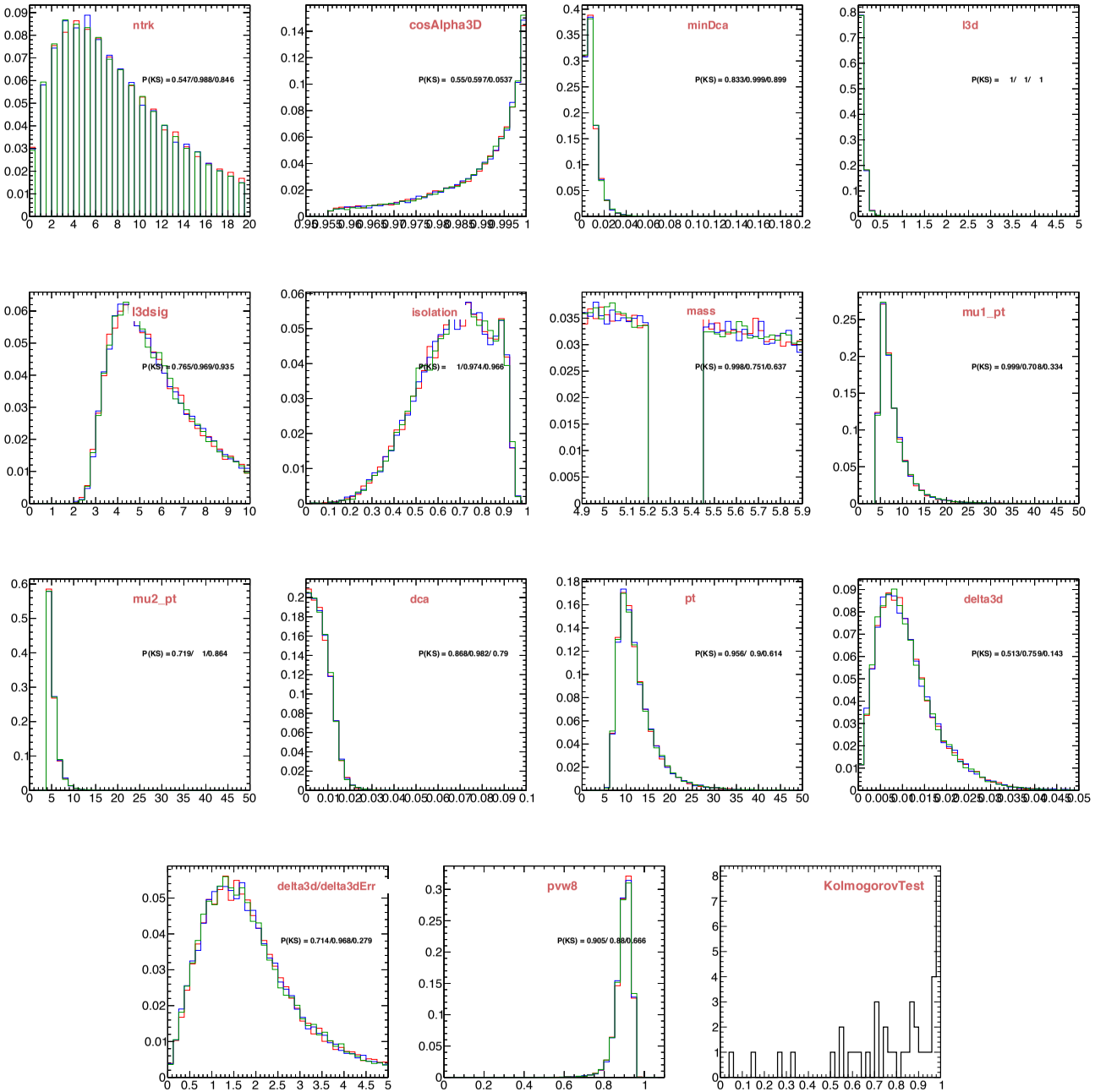

Figure 12: Overlay of BDT training variable distributions in data sideband background for events of the three subsets in the endcap. The plot on the bottom right summarizes all KS probablities.

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	pvips	1.577e-01	iso	1.548e-01	iso	1.400e-01
2	iso	1.572e-01	pvips	1.257e-01	pvips	1.319e-01
3	alpha	1.396e-01	alpha	1.197e-01	alpha	1.278e-01
4	fls3d	8.819e-02	pt	9.154e-02	fls3d	9.723e-02
5	chi2dof	8.550e-02	fls3d	8.968e-02	pt	9.553e-02
6	closetrk	7.571e-02	chi2dof	7.928e-02	closetrk	8.801e-02
7	pt	7.090e-02	closetrk	7.494e-02	chi2dof	7.864e-02
8	docatrk	6.949e-02	docatrk	7.439e-02	maxdoca	7.295e-02
9	maxdoca	5.860e-02	maxdoca	7.032e-02	eta	5.824e-02
10	eta	5.850e-02	pvip	6.783e-02	pvip	5.773e-02
11	pvip	3.857e-02	eta	5.181e-02	docatrk	5.196e-02

Table 3: Variable ranking for events of the three different event samples in the barrel after BDT training.

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	iso	1.389e-01	alpha	1.746e-01	pvips	1.483e-01
2	alpha	1.377e-01	iso	1.684e-01	iso	1.424e-01
3	pvips	1.195e-01	pvips	1.460e-01	alpha	1.387e-01
4	fls3d	9.320e-02	fls3d	9.906e-02	maxdoca	8.872e-02
5	chi2dof	9.047e-02	chi2dof	9.223e-02	pvip	8.600e-02
6	closetrk	8.517e-02	closetrk	7.533e-02	fls3d	8.021e-02
7	pvip	7.938e-02	pt	6.066e-02	chi2dof	7.470e-02
8	maxdoca	7.763e-02	docatrk	5.926e-02	docatrk	7.352e-02
9	pt	6.702e-02	pvip	4.886e-02	closetrk	7.205e-02
10	docatrk	5.680e-02	maxdoca	4.086e-02	pt	5.839e-02
11	eta	5.421e-02	eta	3.474e-02	eta	3.708e-02

Table 4: Variable ranking for events of the three different event samples in the endcaps after BDT training.

4 Boosted Decision Tree

The inclusive samples are split in three different subsamples according to the rule $index = eventNumber\%3$. These samples are then used as follows:

- events of type 0: analyzed by BDT0, trained on type-1 events, tested on type-2 events
- events of type 1: analyzed by BDT1, trained on type-2 events, tested on type-0 events
- events of type 2: analyzed by BDT2, trained on type-0 events, tested on type-1 events

for the training and testing.

Tables **??** and **??** show the ranking of variables before the BDT training.

Sample	Type 0	Type 1	Type 2
Signal barrel	11235	11074	11042
Signal endcaps	6686	6589	6622
Background barrel	27512	27381	27432
Background endcaps	20702	20615	20642

Table 5: Number of events per type for signal and background events in the barrel and endcap.

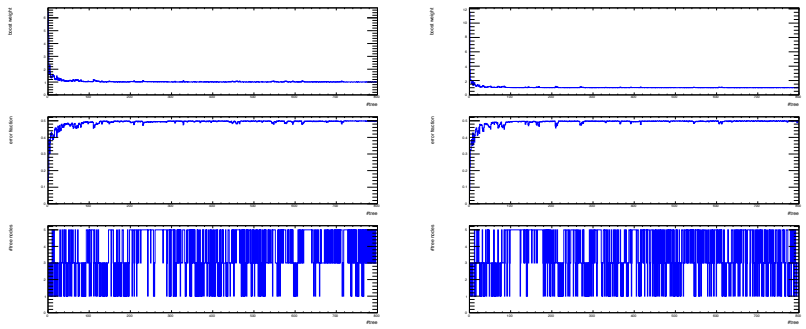

Figure 13: TMVA BDT charaterization plots for the barrel (left) and the endcap (right). Shown versus the tree number is the boost weight (top) and the event misclassification rate (middle), and the number of nodes before pruning (bottom).

Tables **??** shows the number of candidates in each of the subsamples. The events are after all preselections including muon-id (tight muon).

Figure 13 shows control plots for the BDT training.

5 Cut and count analysis

5.1 optimization and blinded results

5.2 unblind

6 BDT analysis

6.1 training and overtraining

6.2 blinded results

6.3 unblind

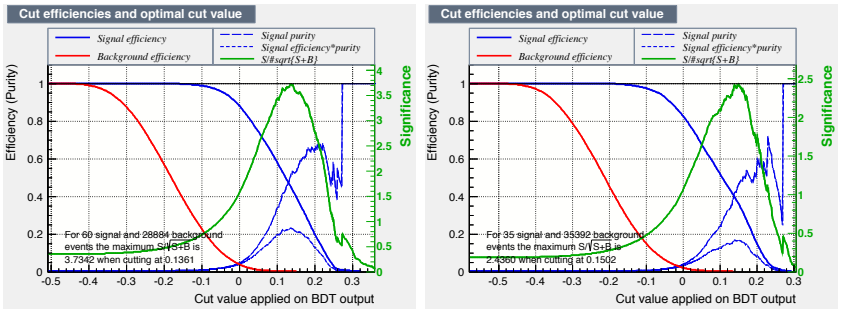

Figure 14: Optimal BDT cut value for barrel (left) and endcaps (right).

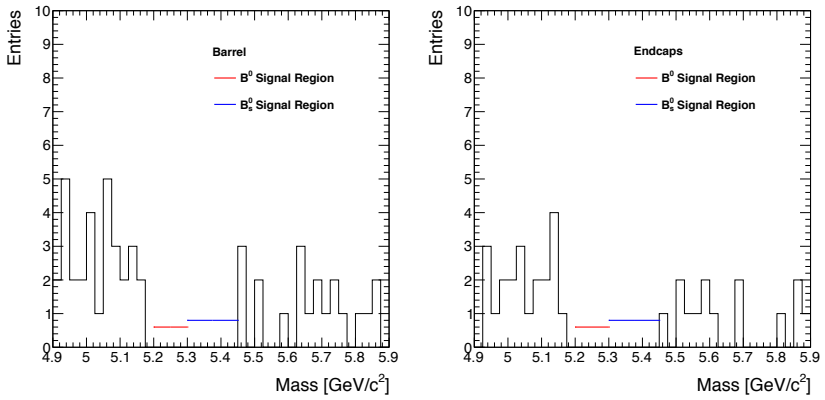

Figure 15: Blinded invariant mass distribution for all candidates passing the BDT selection for barrel (left) and endcaps (right).

Figures 14 show the optimal value of the BDT cut for the estimated numbers of signal and background events.

Figures 15 show the blinded invariant mass distribution after applying the BDT selection.

Figures 16, 17, show variable distributions after the application of the BDT selection.

In figure **??** the standard control plots of TMVA with a linear scale are shown that check against overtraining of the BDT in the barrel and the endcap.

Figures **??** and **??** show the output of the BDT application on different samples for data and $B_s^0 \to \mu^+\mu^-$ simulation.

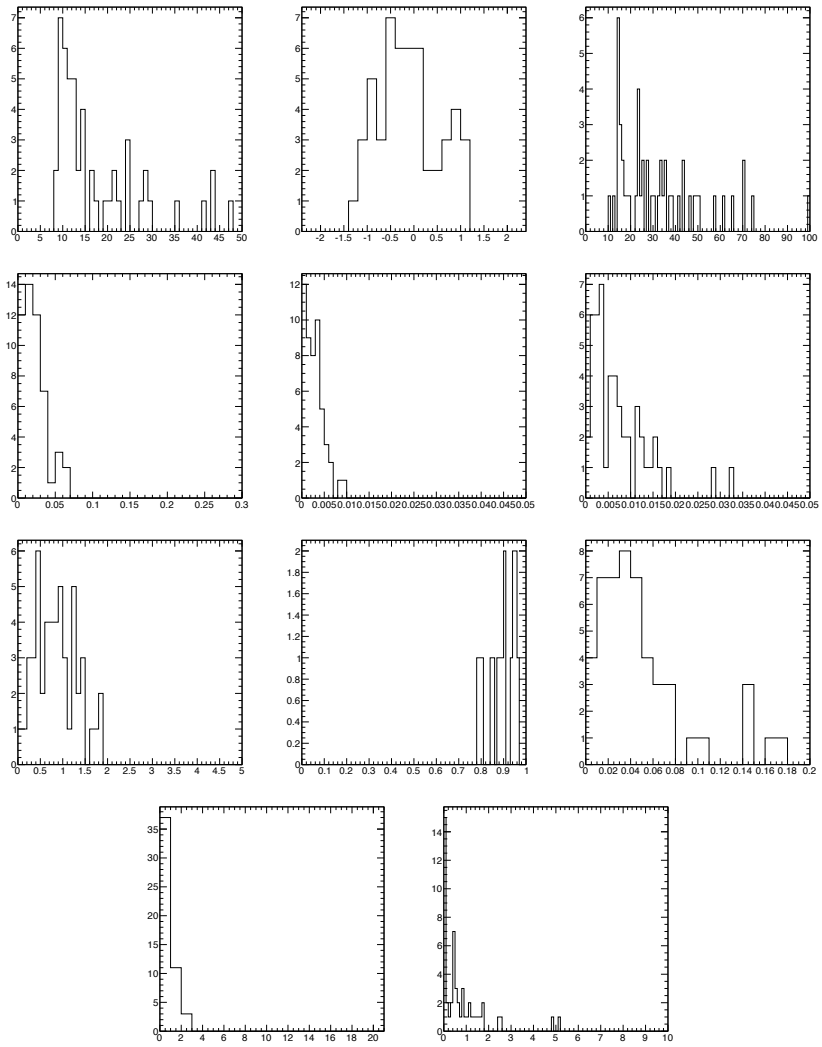

Figure 16: Variable distributions for the barrel after BDT cuts on blinded data.

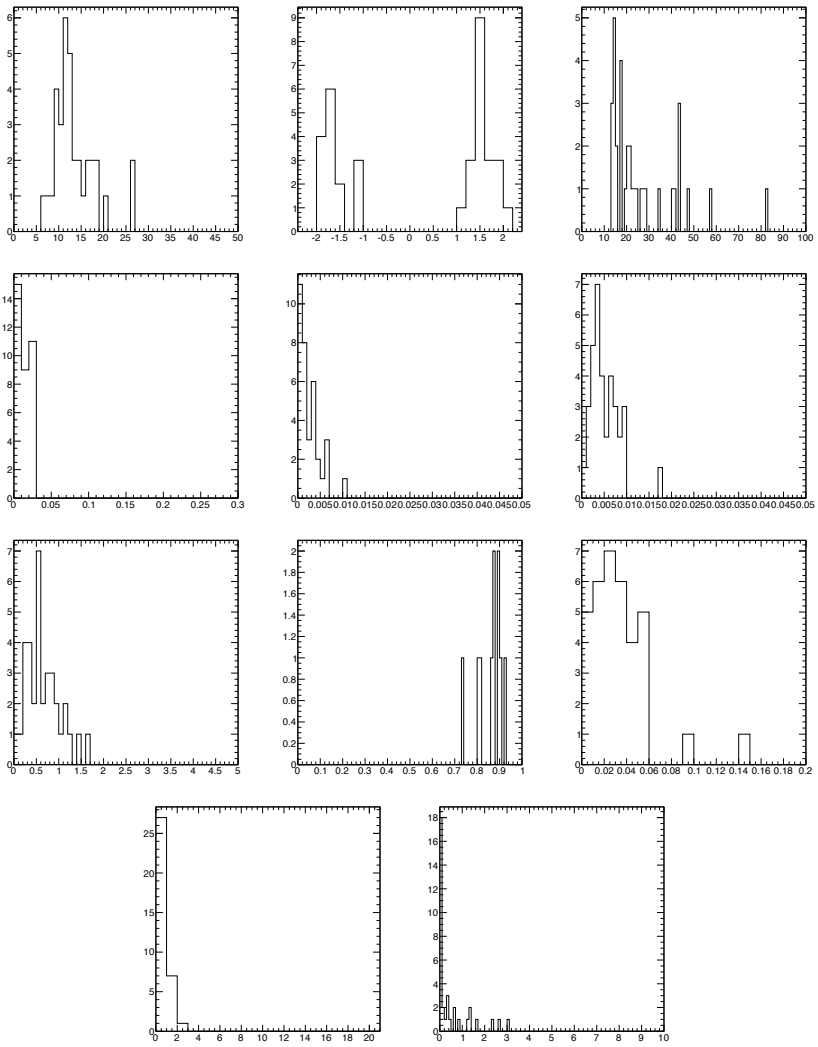

Figure 17: Variable distributions for the endcaps after BDT cuts on blinded data.

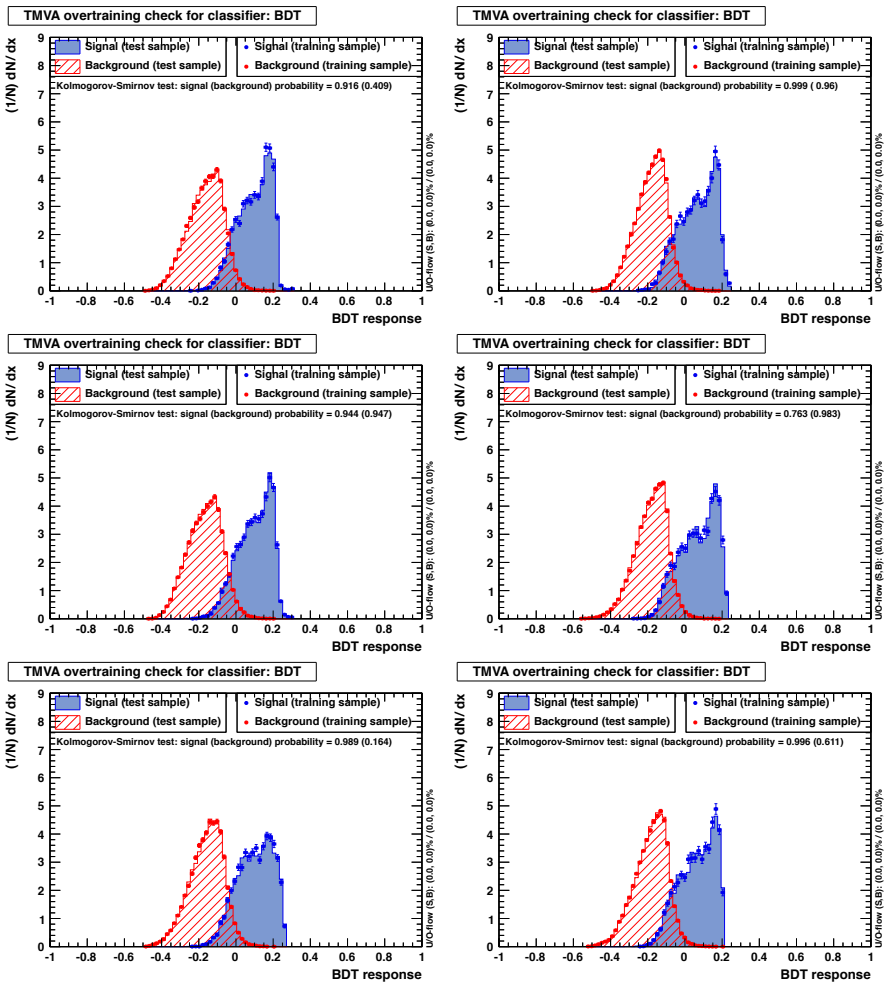

Figure 18: TMVA overtraining control plot for the barrel(left) and the end-cap(right), for events of type 0...2 from the top to the bottom.

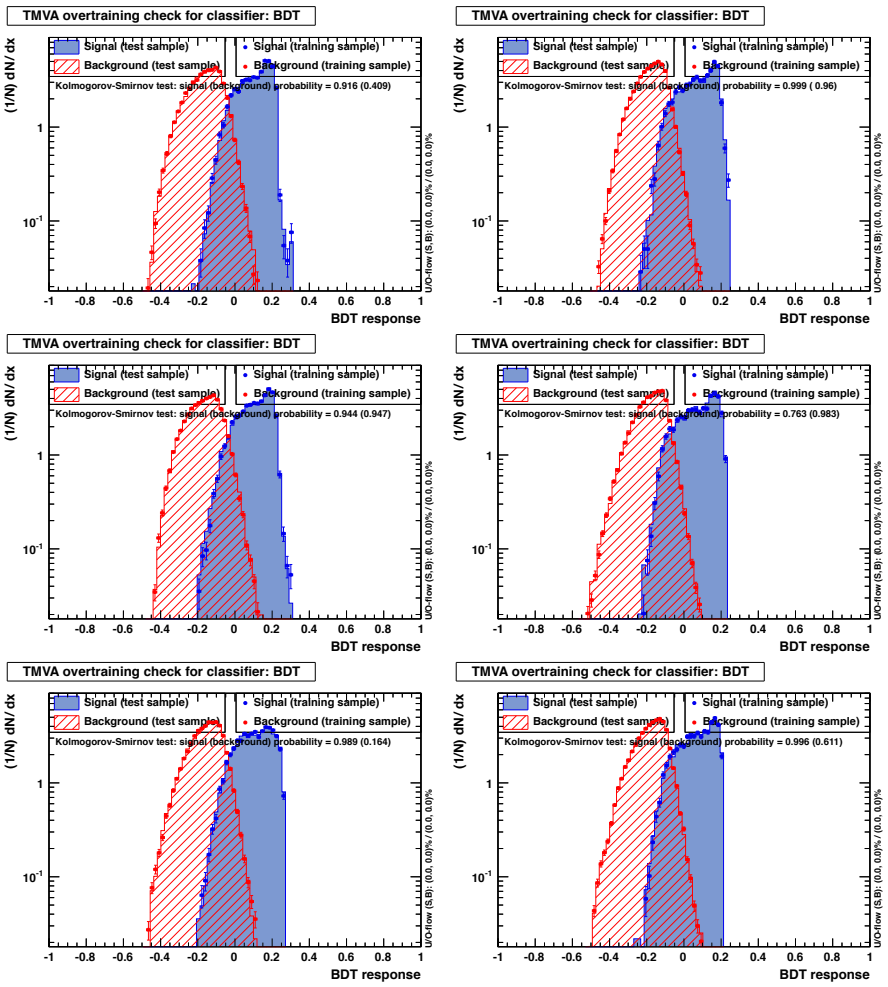

Figure 19: TMVA overtraining control plot for the barrel(left) and the endcap(right), for events of type 0...2 from the top to the bottom. This Figure shows on a logarithmic scale.

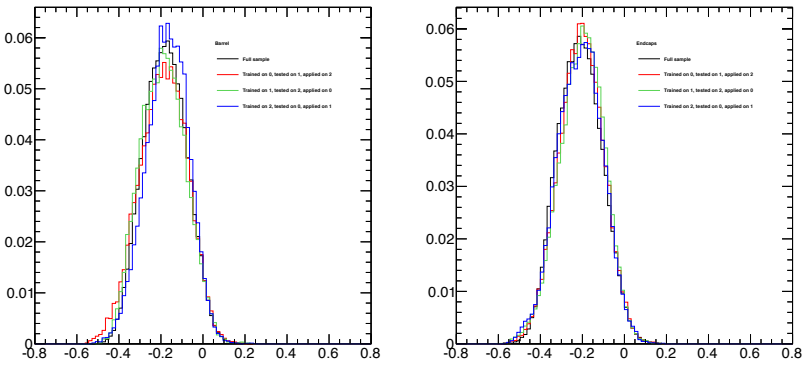

Figure 20: BDT output distribution for the application on data for different control samples and the full sample for (left) and endcaps (right).

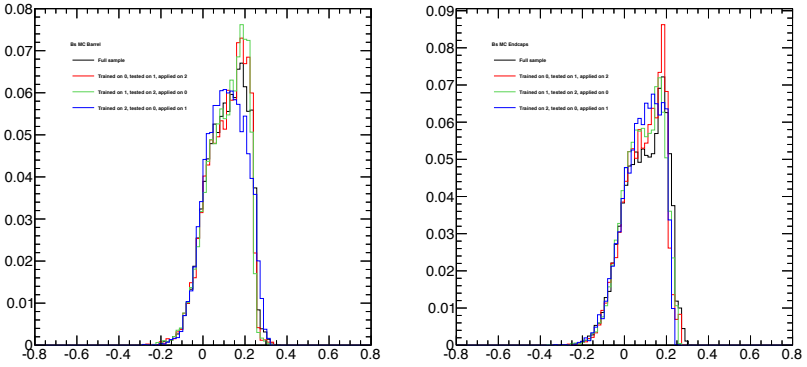

Figure 21: BDT output distribution for the application on $B_s^0 \to \mu^+\mu^-$ signal MC different control samples and the full sample for (left) and endcaps (right).

7 Normalization channel

7.1 datasets

7.2 selection

7.3 BDT

7.4 yields

8 Full dataset

9 Selection

9.1 datasets

9.2 muon identification

9.3 variable distributions, correlations, ranking

9.4 TMVA training

9.4.1 MLP

9.4.2 BDT

9.5 Normalization channel

9.5.1 MLP

9.5.2 BDT

9.6 Limits

10 Summary

References

[1] U. Langenegger et al. AN-12-358: "Search for $B_s \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ with the 2011 and 2012 data". Technical report, CMS Collaboration, 2012.