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

Purdue, Pisa

March 28, 2013

Abstract

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

Contents

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

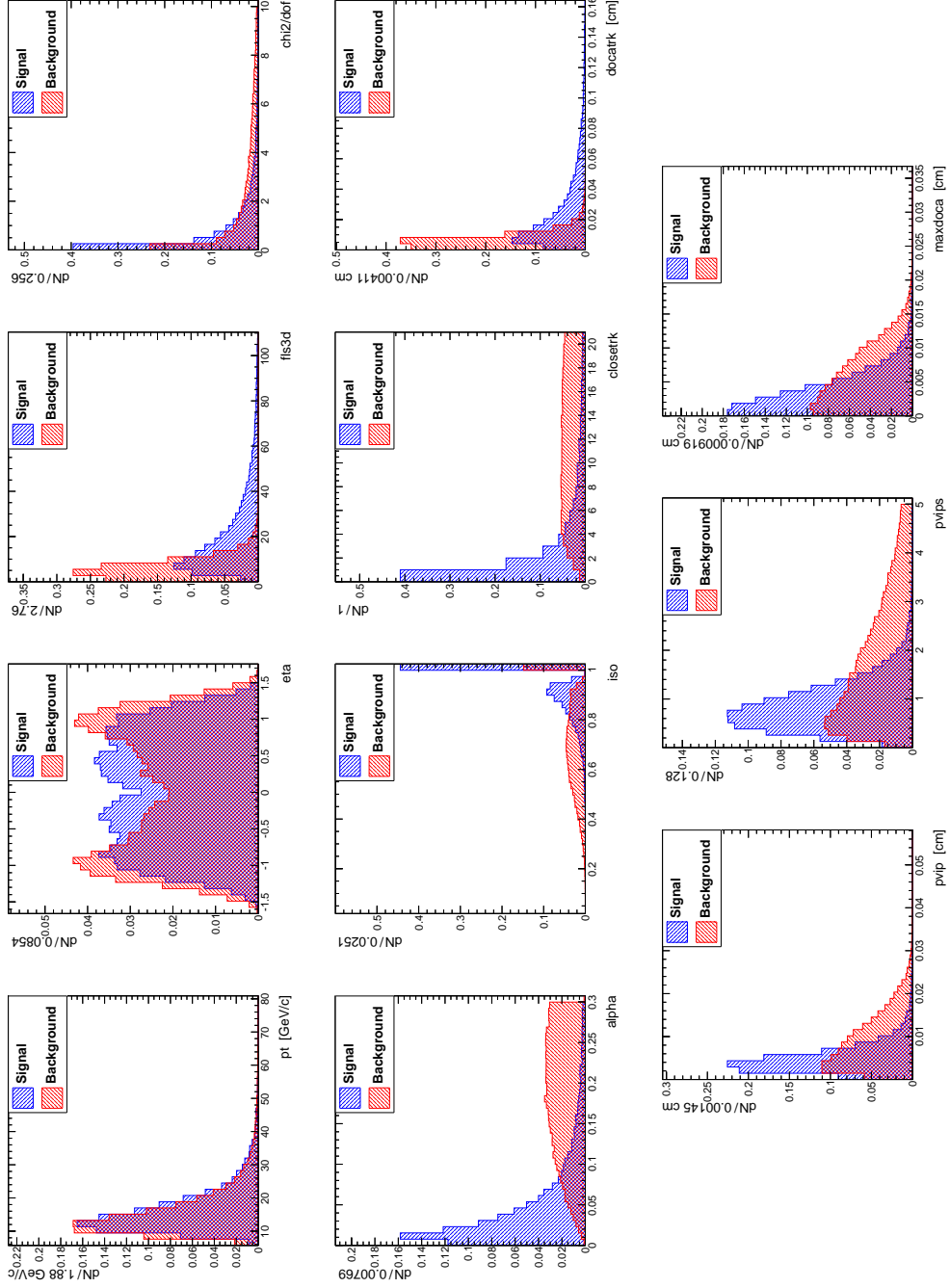Tables **??** and **??** show the ranking of variables before the BDT training.

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.
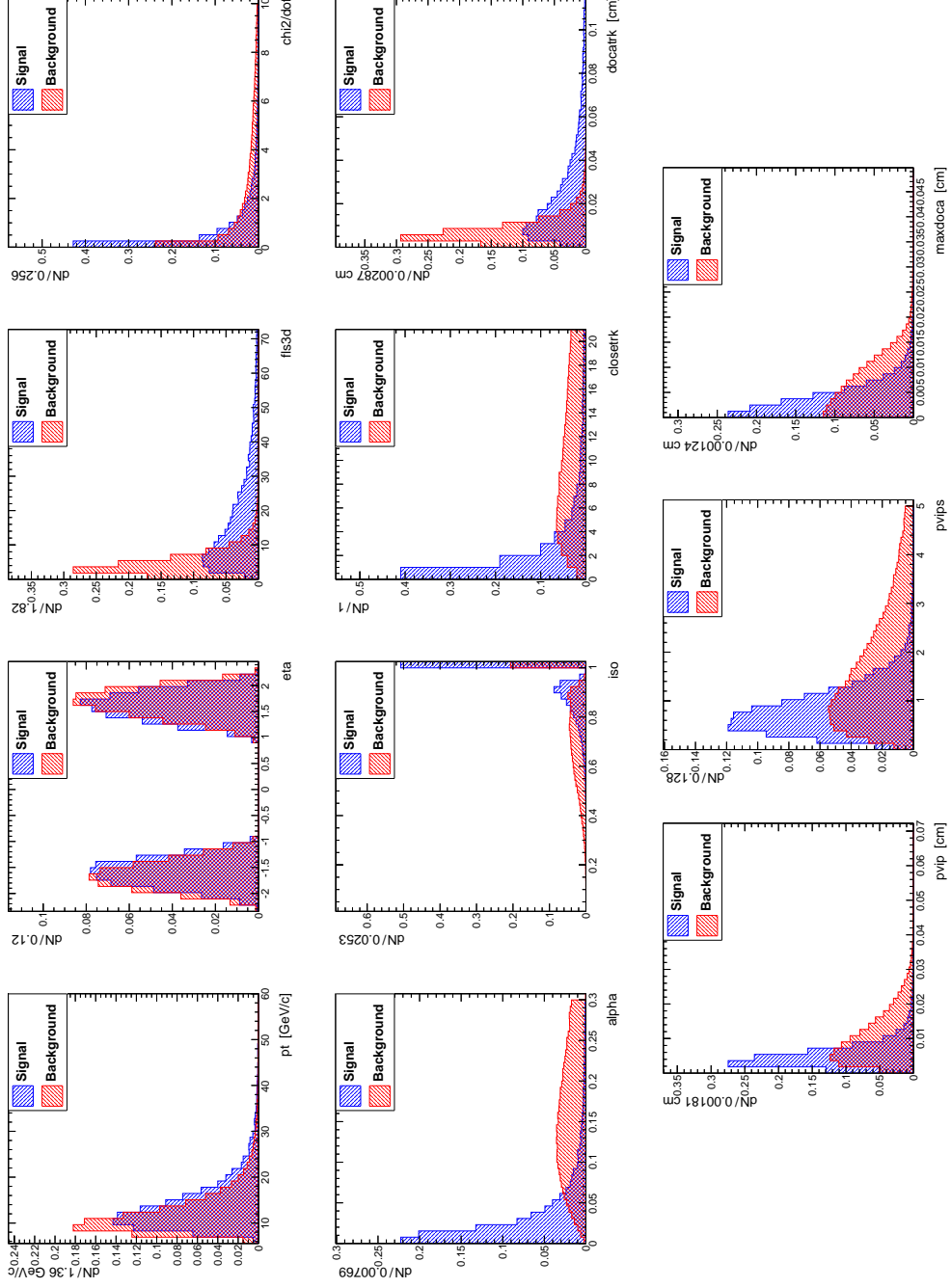
Figure 2: Standard TMVA plot of the input variables for the endcaps BDT for signal (blue) and background (red). The background is extracted from data dimuon sidebands.
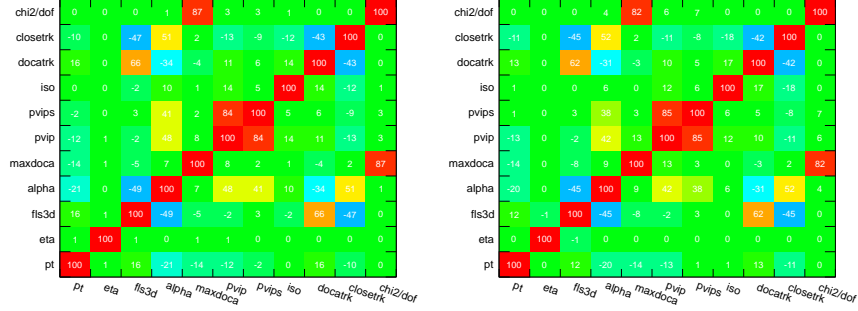
5

Figure 3: Correlation matrix for signal events in the barrel (left) and the endcap (right).
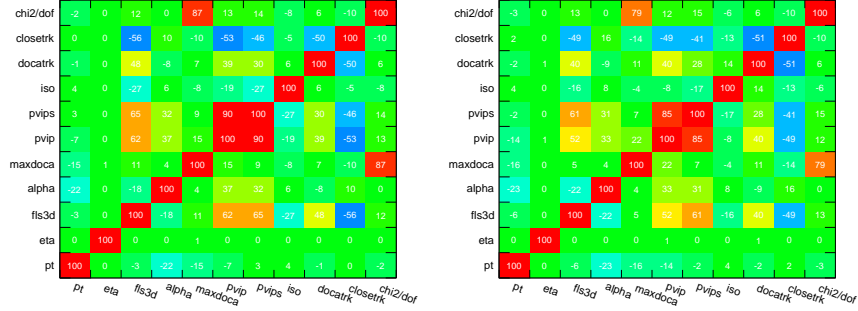
Figure 4: Correlation matrix for background events in the barrel (left) and the endcap (right).

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	alpha	5.152e-01	alpha	5.187e-01	alpha	5.231e-01
2	closetrk	4.468e-01	closetrk	4.501e-01	closetrk	4.390e-01
3	fls3d	3.449e-01	fls3d	3.506e-01	fls3d	3.476e-01
4	docatrk	3.433e-01	docatrk	3.429e-01	docatrk	3.403e-01
5	iso	3.186e-01	iso	3.197e-01	iso	3.166e-01
6	pvips	2.223e-01	pvips	2.212e-01	pvips	2.201e-01
7	pvip	1.617e-01	pvip	1.562e-01	pvip	1.635e-01
8	maxdoca	1.069e-01	maxdoca	1.075e-01	maxdoca	1.106e-01
9	chi2/dof	1.014e-01	chi2/dof	1.019e-01	chi2/dof	1.071e-01
10	eta	2.601e-02	eta	2.483e-02	eta	2.552e-02
11	pt	9.944e-03	pt	1.004e-02	pt	1.009e-02

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	5.692e-01	alpha	5.674e-01	alpha	5.632e-01
2	closetrk	4.172e-01	closetrk	4.115e-01	closetrk	4.166e-01
3	fls3d	3.764e-01	fls3d	3.747e-01	fls3d	3.728e-01
4	docatrk	3.063e-01	docatrk	3.002e-01	docatrk	3.094e-01
5	iso	2.525e-01	iso	2.534e-01	iso	2.564e-01
6	pvips	2.193e-01	pvips	2.179e-01	pvips	2.233e-01
7	pvip	2.022e-01	pvip	2.008e-01	pvip	2.043e-01
8	maxdoca	1.272e-01	maxdoca	1.306e-01	maxdoca	1.293e-01
9	chi2/dof	1.047e-01	chi2/dof	1.057e-01	chi2/dof	1.078e-01
10	pt	3.555e-02	pt	3.862e-02	pt	3.534e-02
11	eta	1.208e-02	eta	1.299e-02	eta	1.072e-02

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

	0		1		2	
rank	variable	separation	variable	separation	variable	separation
1	pvips	1.660e-01	iso	1.383e-01	pvips	1.485e-01
2	iso	1.295e-01	pvips	1.283e-01	iso	1.435e-01
3	alpha	1.249e-01	alpha	1.206e-01	alpha	1.388e-01
4	fls3d	9.997e-02	pt	1.026e-01	fls3d	1.070e-01
5	pt	8.929e-02	fls3d	1.017e-01	pt	9.982e-02
6	closetrk	8.100e-02	closetrk	7.681e-02	closetrk	9.170e-02
7	chi2dof	7.528e-02	chi2dof	7.167e-02	chi2dof	8.199e-02
8	docatrk	7.456e-02	docatrk	6.902e-02	eta	7.281e-02
9	eta	7.163e-02	eta	6.747e-02	maxdoca	5.140e-02
10	pvip	4.615e-02	maxdoca	6.372e-02	pvip	3.815e-02
11	maxdoca	4.177e-02	pvip	5.978e-02	docatrk	2.639e-02

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

4 Boosted Decision Tree

The inclusive samples 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.

rank	0 variable	0 separation	1 variable	1 separation	2 variable	2 separation
1	alpha	1.648e-01	alpha	1.641e-01	pvips	1.750e-01
2	pvips	1.461e-01	pvips	1.555e-01	alpha	1.690e-01
3	iso	1.265e-01	iso	1.304e-01	iso	1.256e-01
4	closetrk	1.077e-01	closetrk	1.000e-01	closetrk	9.363e-02
5	pt	9.134e-02	fls3d	8.571e-02	pt	7.921e-02
6	fls3d	7.871e-02	pt	8.055e-02	maxdoca	7.662e-02
7	maxdoca	6.374e-02	chi2dof	7.770e-02	chi2dof	6.989e-02
8	pvip	5.972e-02	maxdoca	6.384e-02	eta	6.810e-02
9	chi2dof	5.492e-02	pvip	5.309e-02	fls3d	6.457e-02
10	docatrk	5.421e-02	eta	5.159e-02	pvip	4.433e-02
11	eta	5.220e-02	docatrk	3.759e-02	docatrk	3.403e-02

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

Sample	Type 0	Type 1	Type 2
Signal barrel	11713	11481	11500
Signal endcaps	7392	7318	7390
Background barrel	38206	37947	38308
Background endcaps	68327	68233	68157

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

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

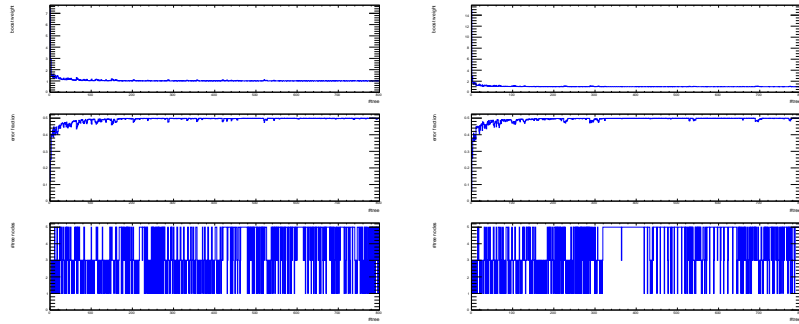Figure 5 shows control plots for the BDT training.

Figure 5: TMVA BDT charaterization plots for the barrel (left) and the end-cap (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).

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

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.