Automatic MVA Evaluation

Thomas Keck
Moritz Gelb
Nils Braun

November 16, 2017

Abstract

Evaluation plots

Contents

1 Classifiers

This section contains the GeneralOptions and SpecificOptions of all classifiers represented by an XML tree. The same information can be retreived using the basf2_mva_info tool.

Table 1: Abbreviations of identifiers

Identifier	Abbreviation
MyExample.xml	MyExa

1.1 MyExample.xml

```
<?xml version="1.0" encoding="utf-8"?>
<method>FastBDT</method>
<weightfile>MyExample.xml</weightfile>
<treename>B0Signal</treename>
<target_variable>B0_isSignal</target_variable>
<weight_variable/>
<signal_class>1</signal_class>
<max_events>0</max_events>
<number_feature_variables>18</number_feature_variables>
<variable0>B0_P</variable0>
<variable1>B0_K_10_P</variable1>
<variable2>B0_K_10_pi0_P</variable2>
<variable3>B0_K_10_pi1_P</variable3>
<variable4>B0_K_10_K_S0_P</variable4>
<variable5>B0_gamma_P</variable5>
<variable6>B0_cosTheta</variable6>
<variable7>B0_M</variable7>
<variable8>B0_ErrM</variable8>
<variable9>B0_SigM</variable9>
<variable10>B0_K_10_M</variable10>
<variable11>B0_K_10_ErrM</variable11>
<variable12>B0_K_10_SigM</variable12>
<variable13>B0_K_10_K_S0_M</variable13>
<variable14>B0_K_10_K_S0_Rho</variable14>
<variable15>B0_DeltaTErr</variable15>
<variable16>B0_deltae</variable16>
<variable17>B0_mbc</variable17>
<number_spectator_variables>0</number_spectator_variables>
<number_data_files>1</number_data_files>
<datafile0>/home/sbilokin/BelleII/test/signal-cs2/train.root</datafile0>
<FastBDT_version>1</FastBDT_version>
<FastBDT_nTrees>200</FastBDT_nTrees>
<FastBDT_nCuts>8</FastBDT_nCuts>
<FastBDT_nLevels>3</FastBDT_nLevels>
<FastBDT_shrinkage>0.10000000000000001</FastBDT_shrinkage>
<FastBDT_randRatio>0.5</FastBDT_randRatio>
```

2 Variables

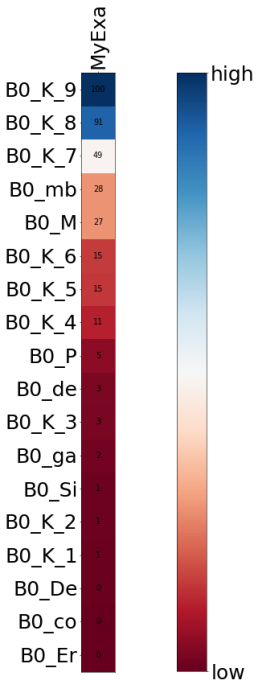
This section contains an overview of the importance and correlation of the variables used by the classifiers. And distribution plots of the variables on the independent dataset. The distributions are normed for signal and background separately, and only the region +- 3 sigma around the mean is shown.

Table 2: Abbreviations of variables

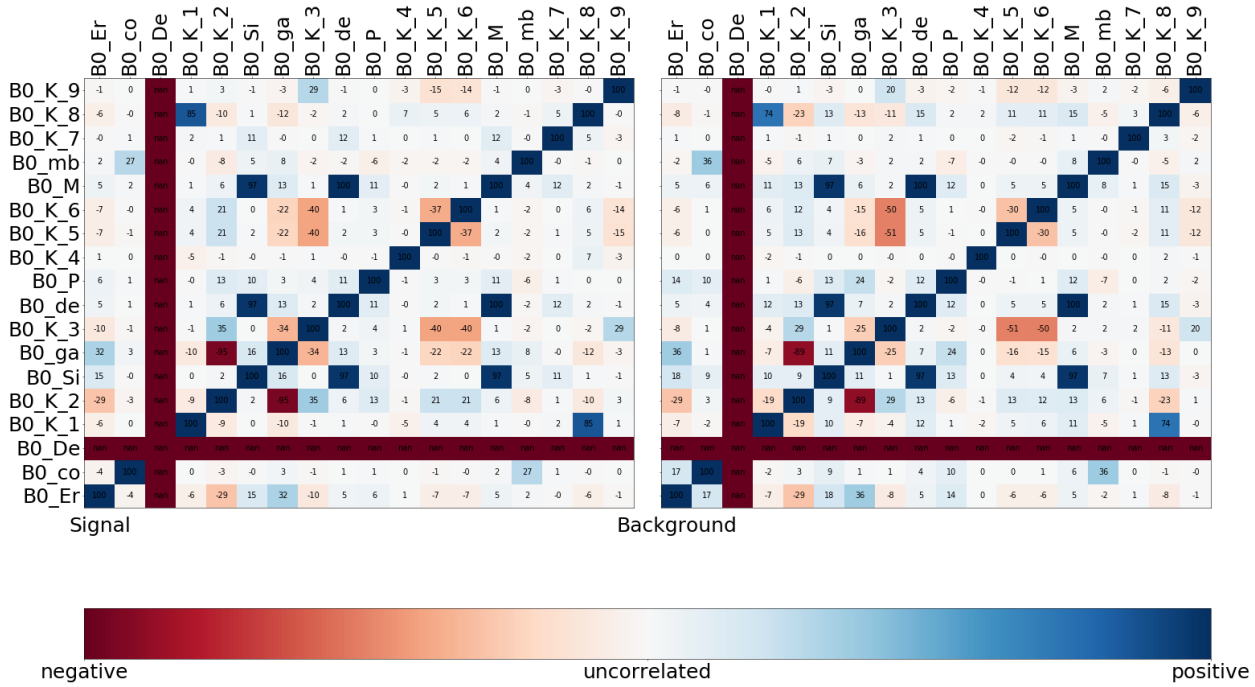
Variable	Abbreviation
B0_ErrM	B0_Er
B0_cosTheta	B0_co
B0_DeltaTErr	B0_De
B0_K_10_SigM	B0_K_1
B0_K_10_P	B0_K_2
B0_SigM	B0_Si
B0_gamma_P	B0_ga
B0_K_10_K_S0_P	B0_K_3
B0_deltae	B0_de
B0_P	B0_P

B0_K_10_ErrM	B0_K_4
B0_K_10_pi0_P	B0_K_5
B0_K_10_pi1_P	B0_K_6
B0_M	B0_M
B0_mbc	B0_mb
B0_K_10_K_S0_M	B0_K_7
B0_K_10_M	B0_K_8
B0_K_10_K_S0_Rho	B0_K_9

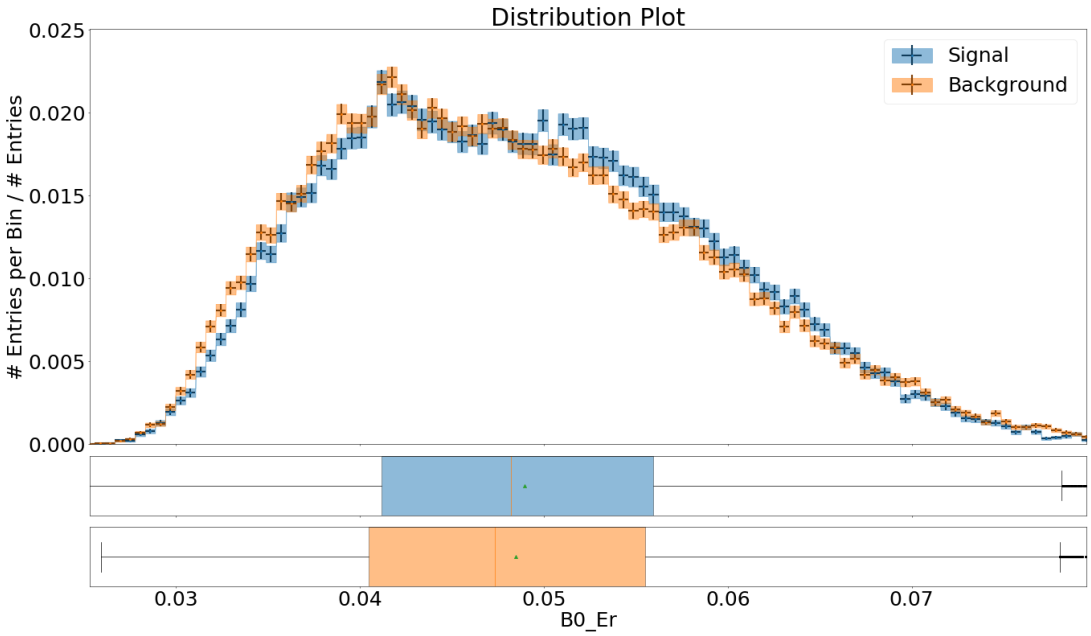
2.1 Importance



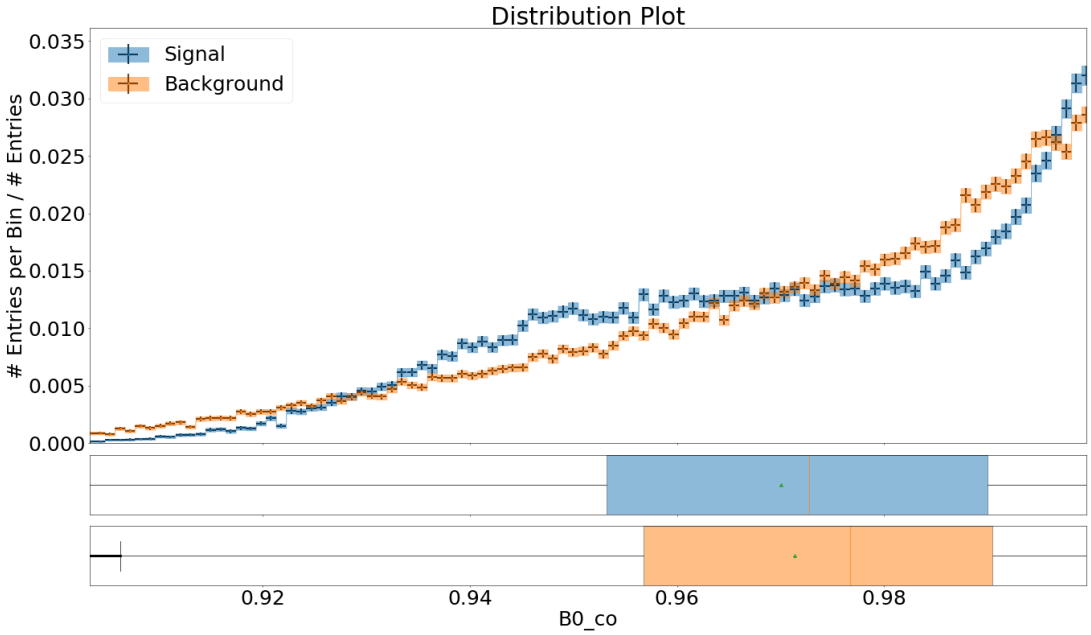
2.2 Correlation



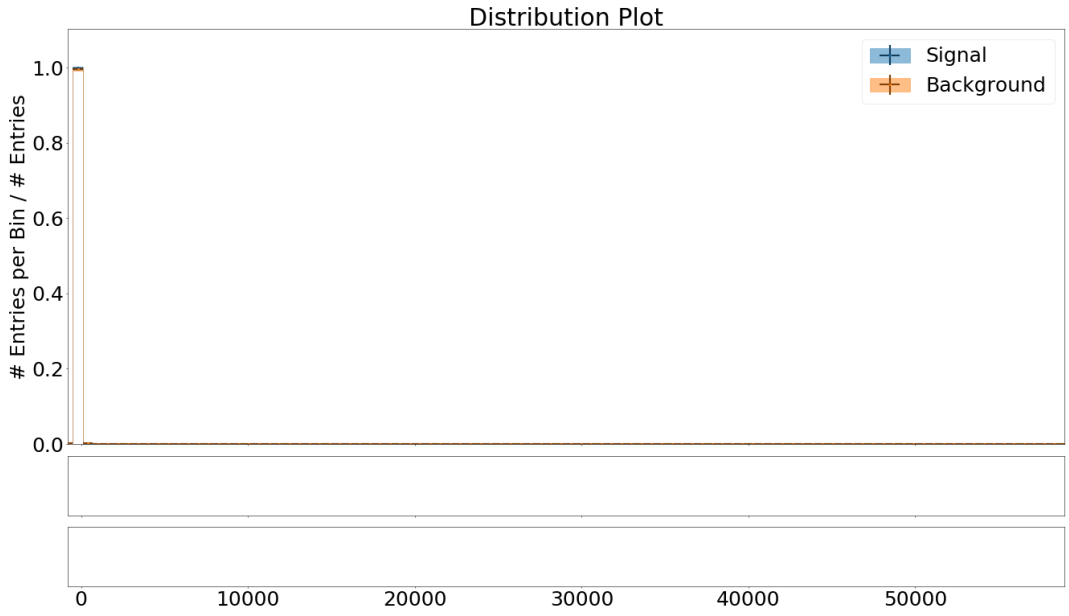
2.3 B0_ErrM



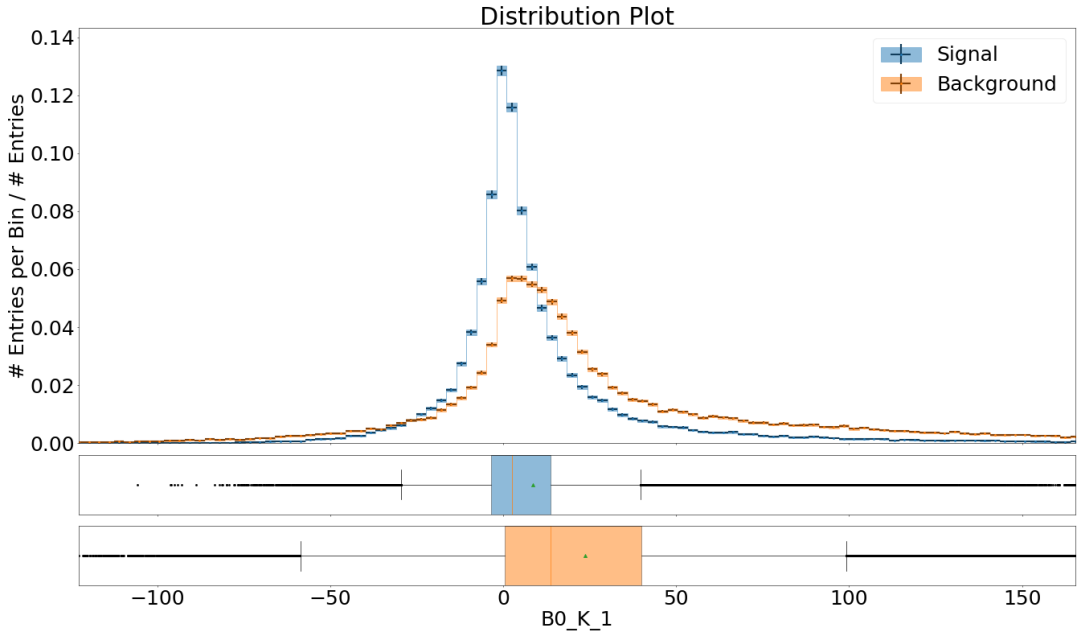
2.4 B0_cosTheta



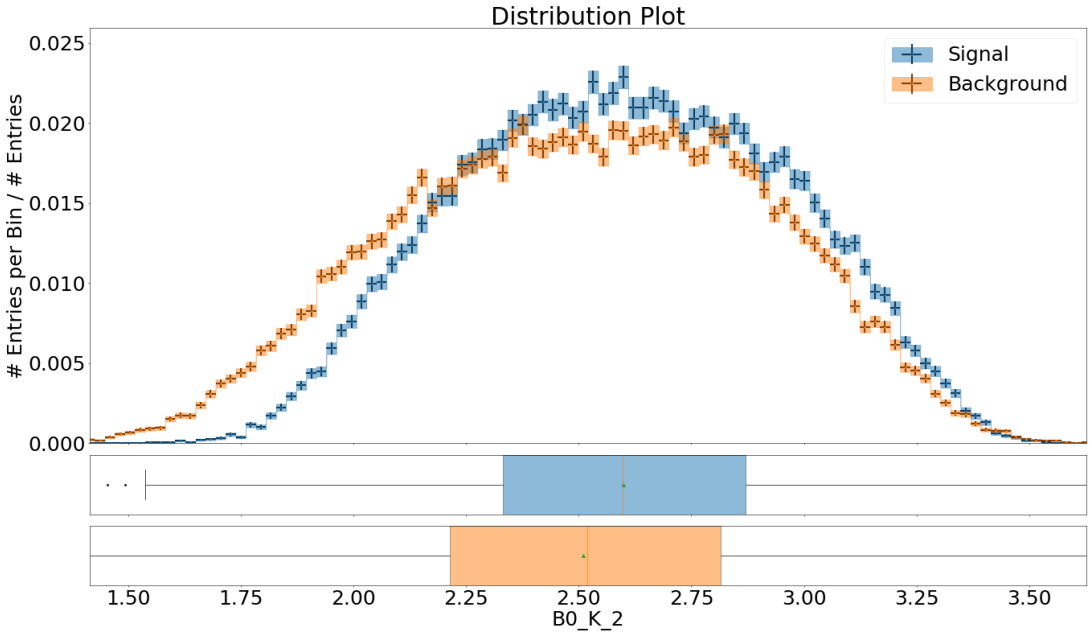
2.5 B0_DeltaTErr



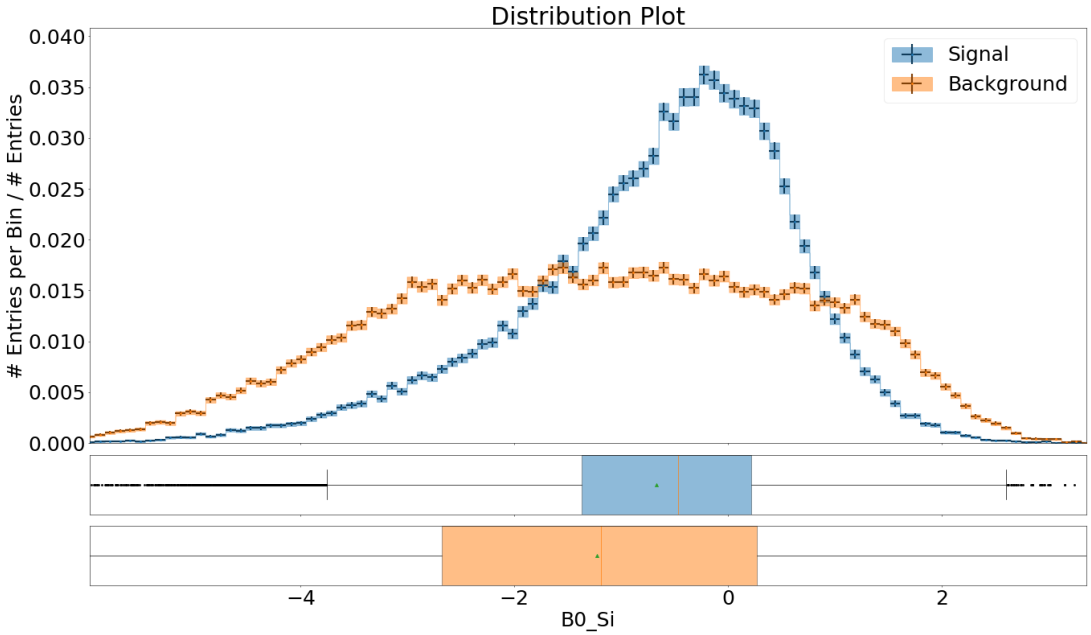
2.6 B0_K_10_SigM



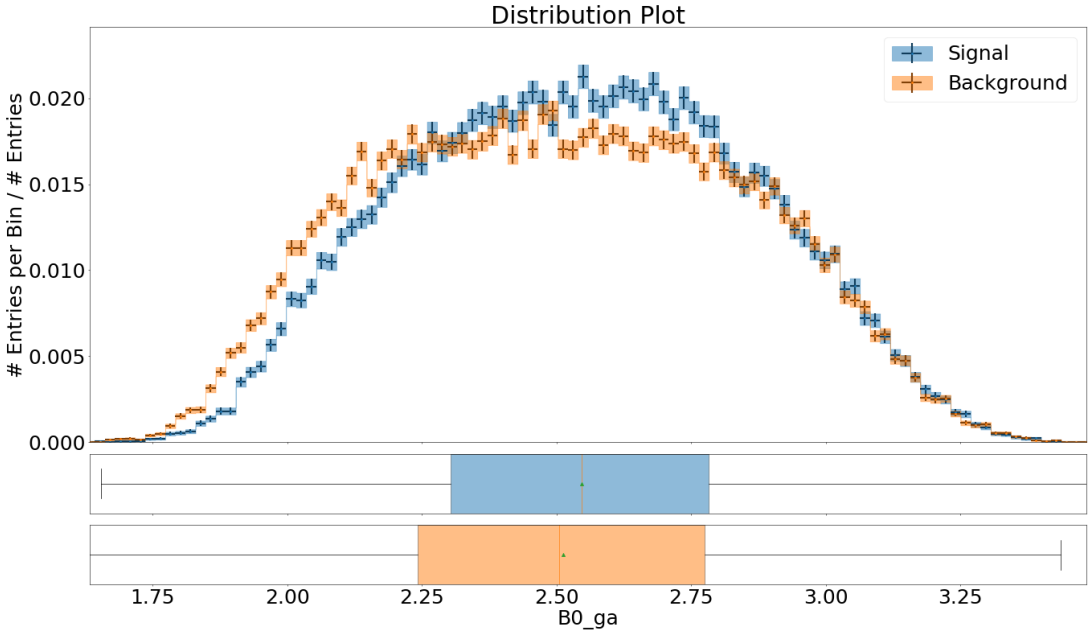
2.7 B0_K_10_P



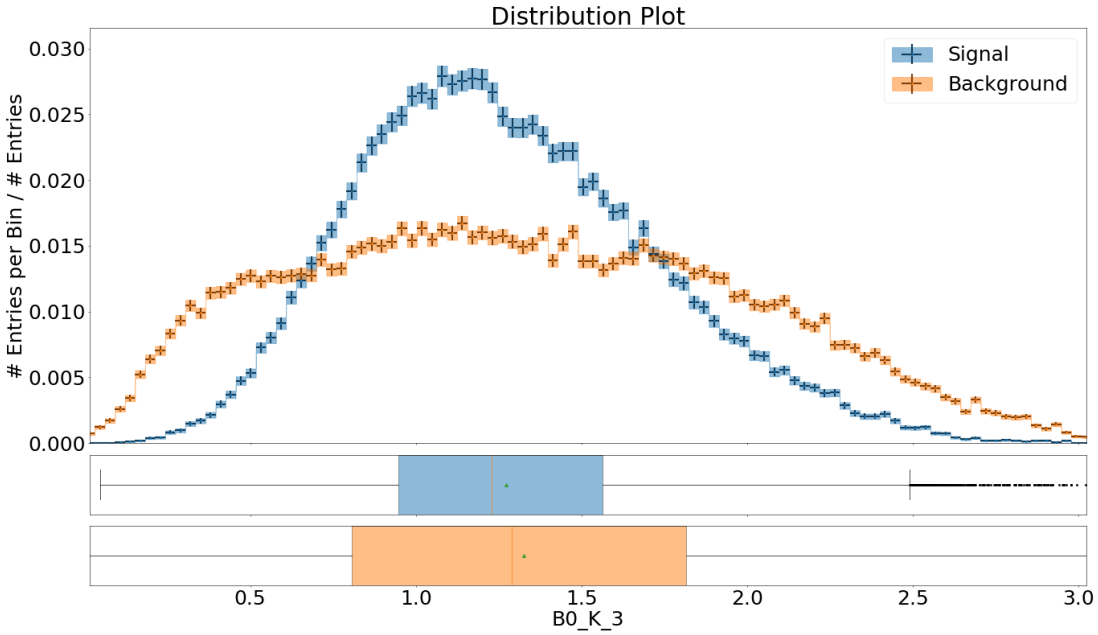
2.8 B0_SigM



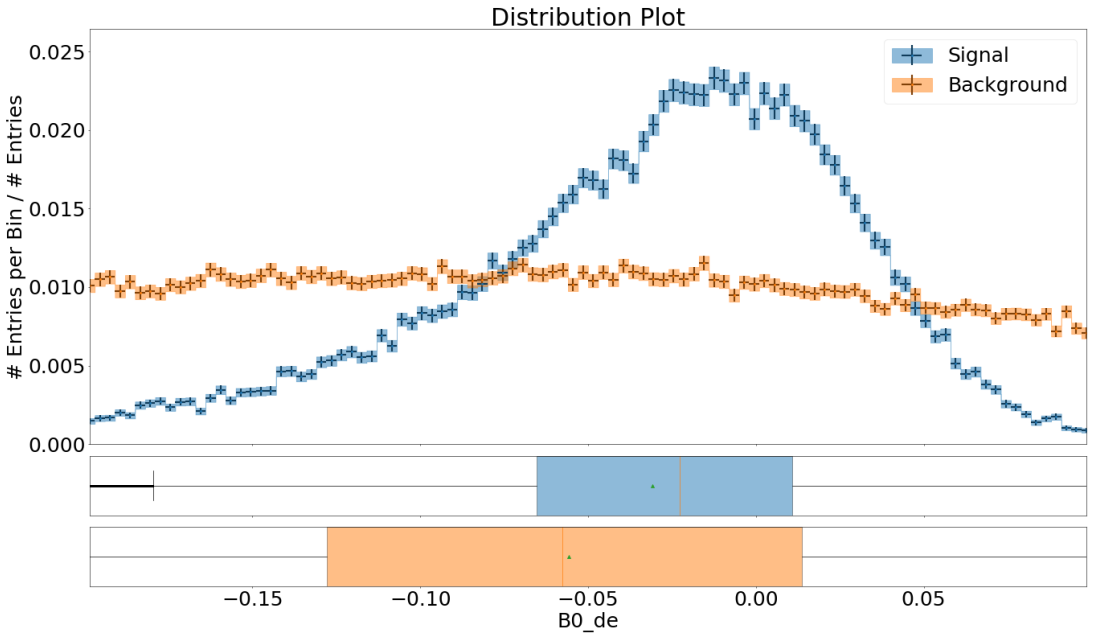
2.9 B0_gamma_P



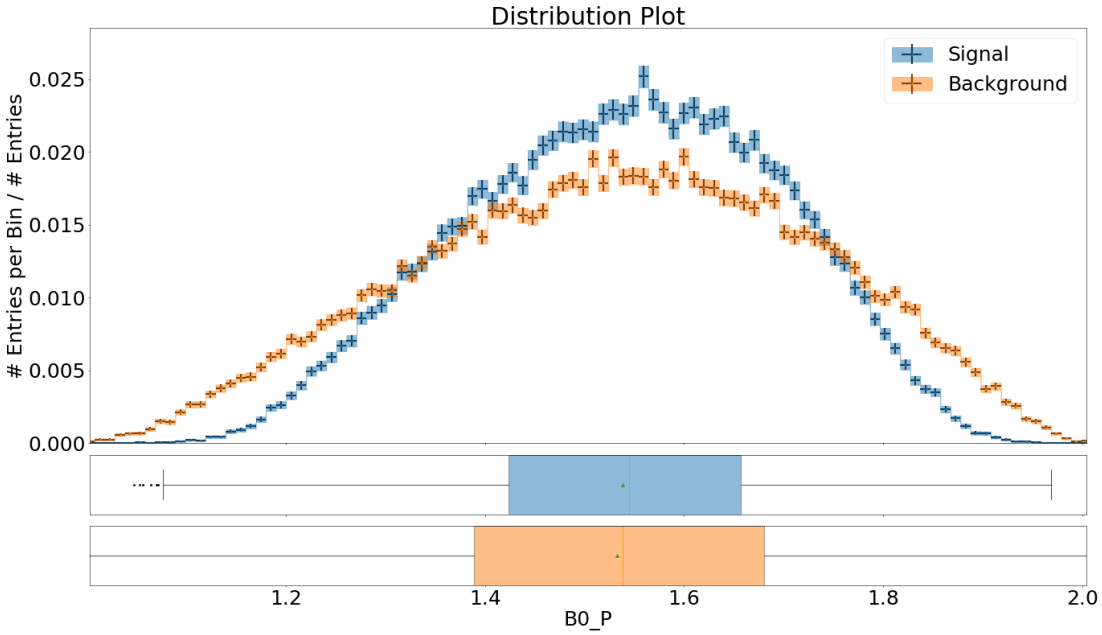
2.10 B0_K_10_K_S0_P



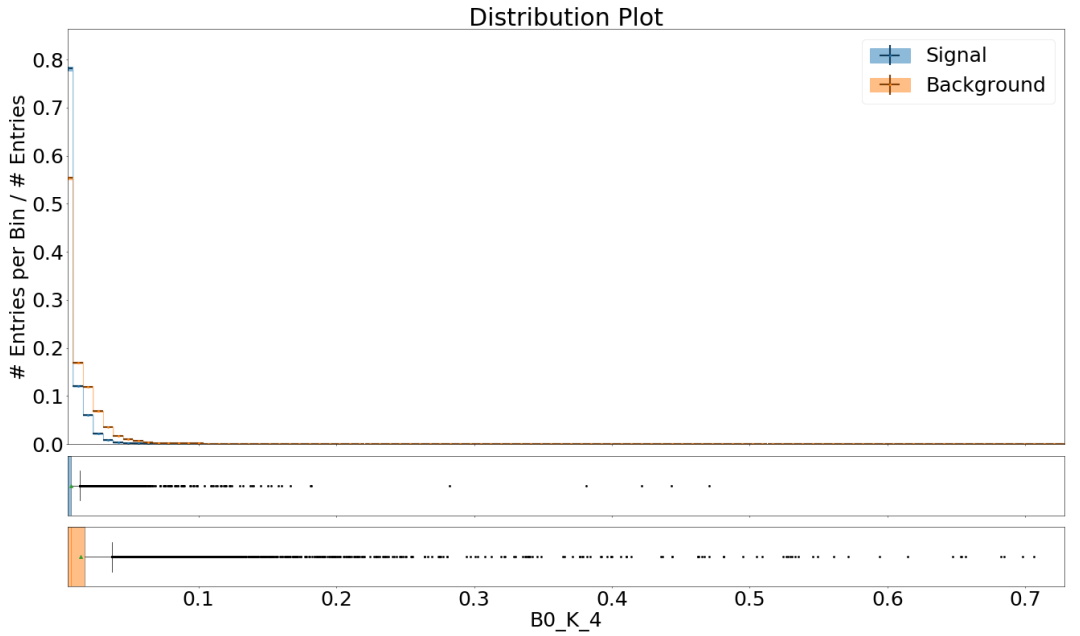
2.11 B0_deltae



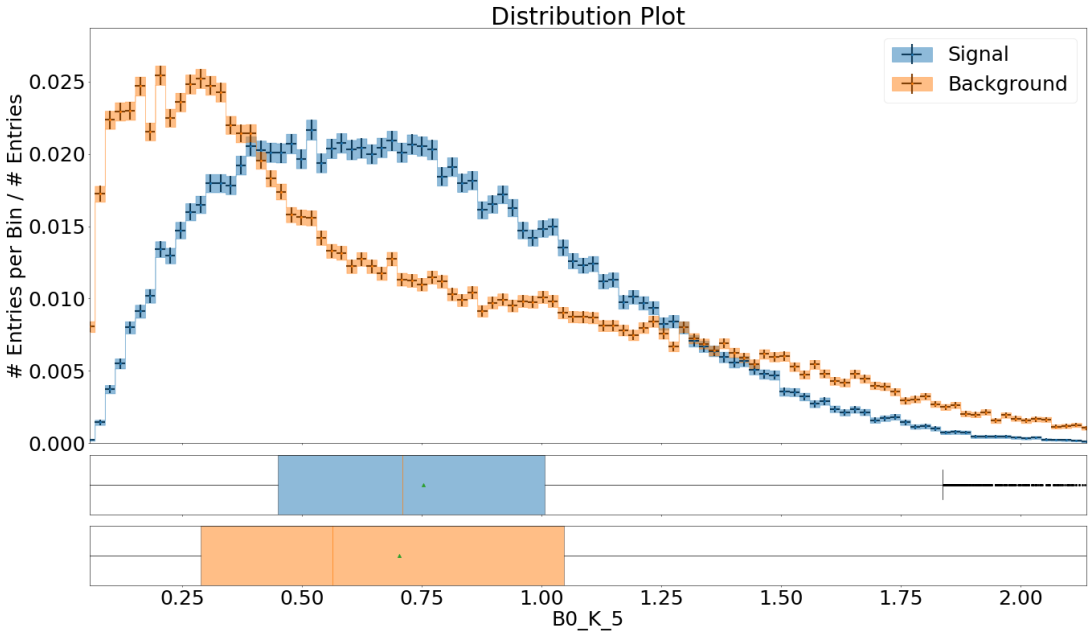
2.12 B0_P



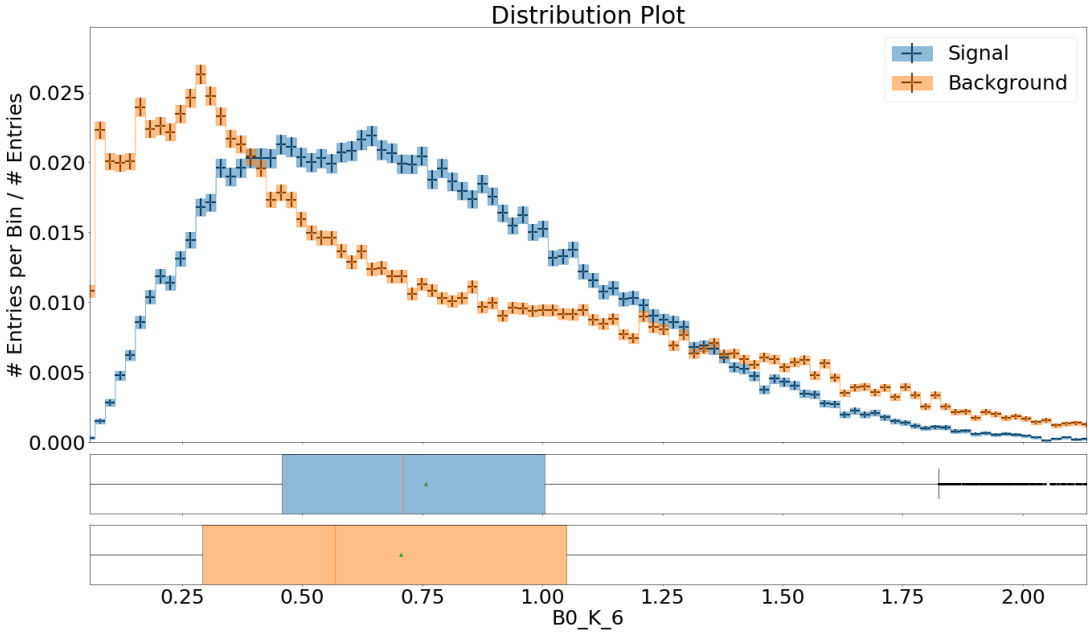
2.13 B0_K_10_ErrM



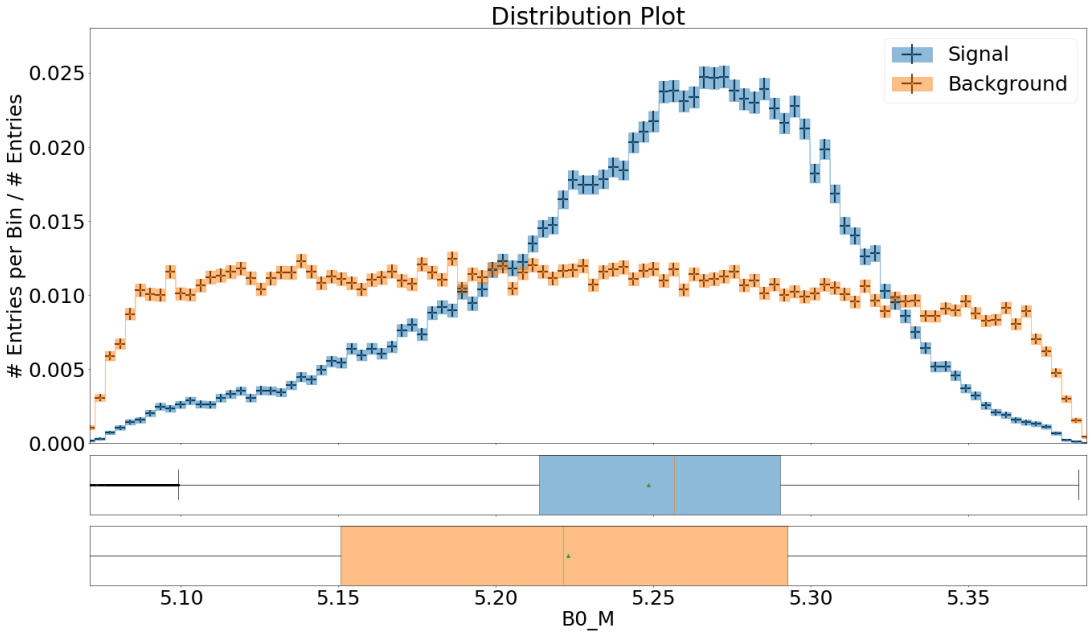
2.14 B0_K_10_pi0_P



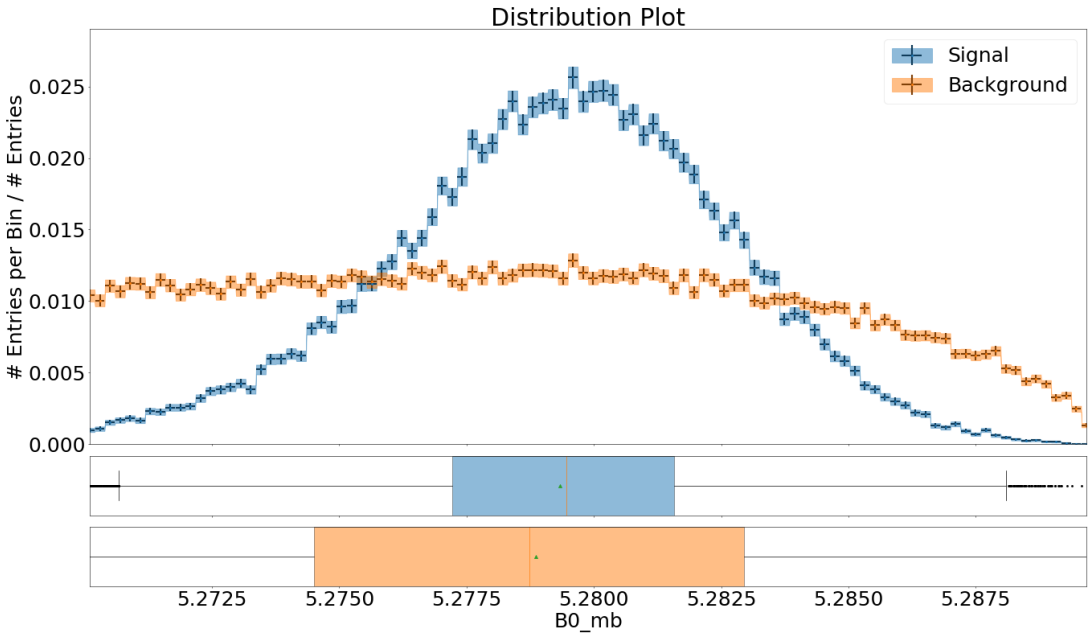
2.15 B0_K_10_pi1_P



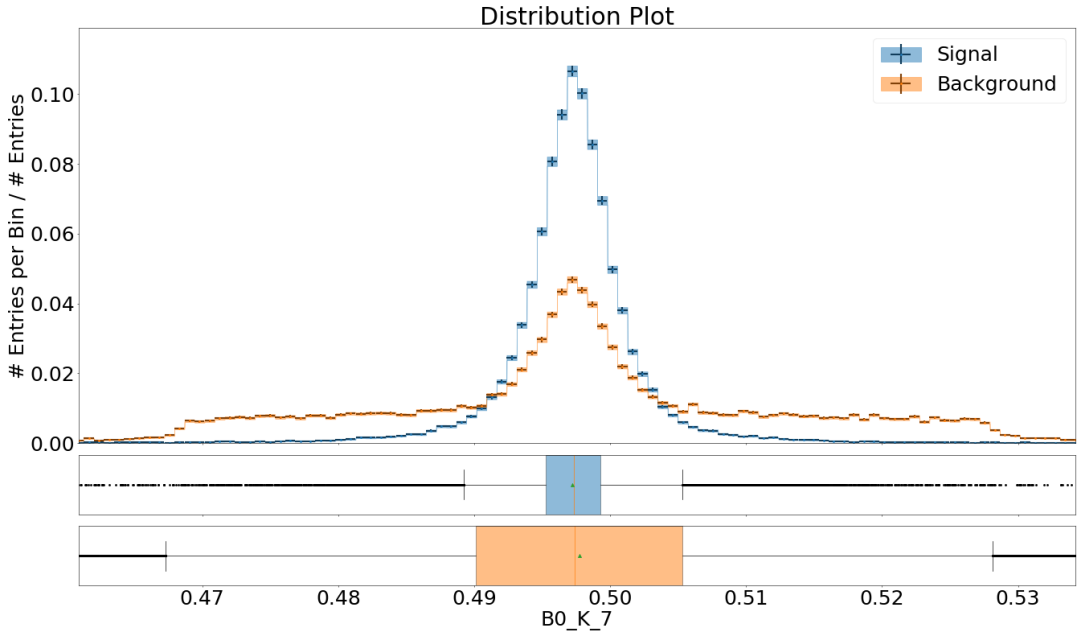
2.16 B0_M



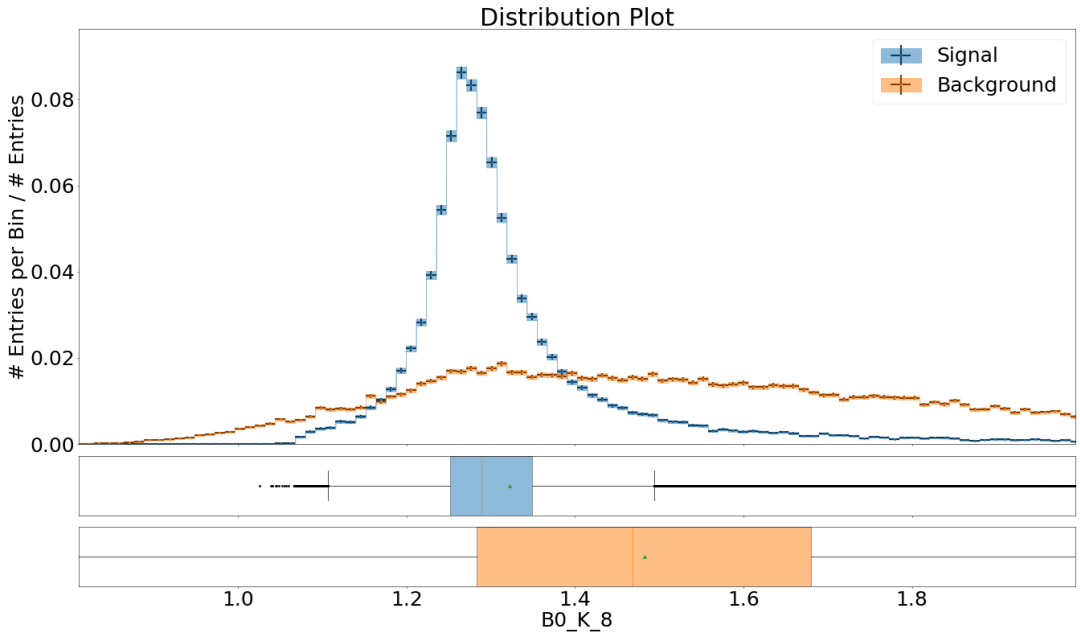
2.17 B0_mbc



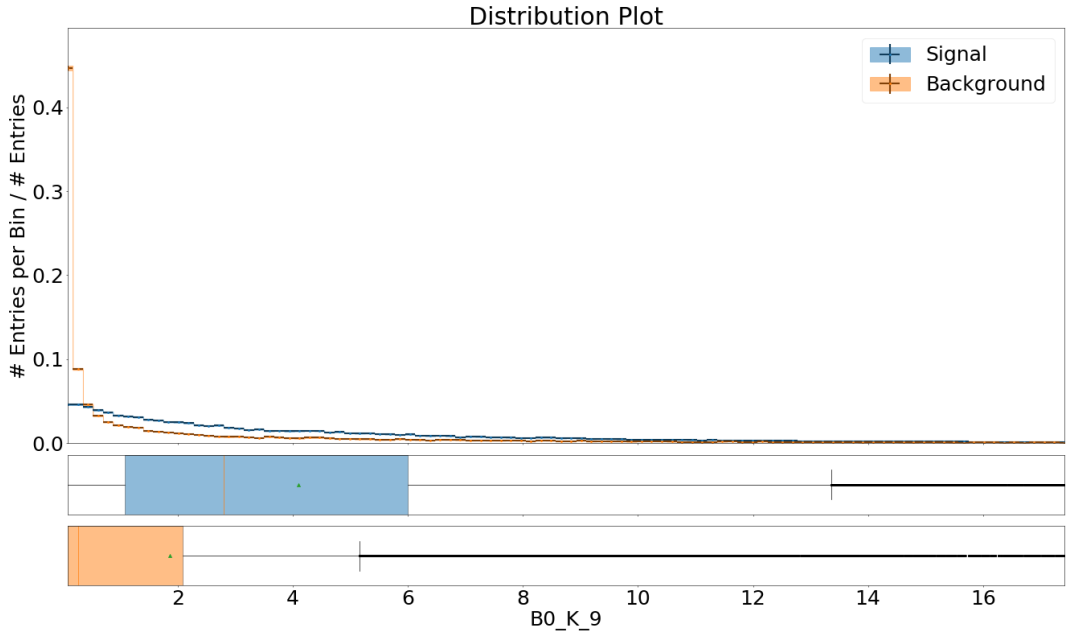
2.18 B0_K_10_K_S0_M



2.19 B0_K_10_M



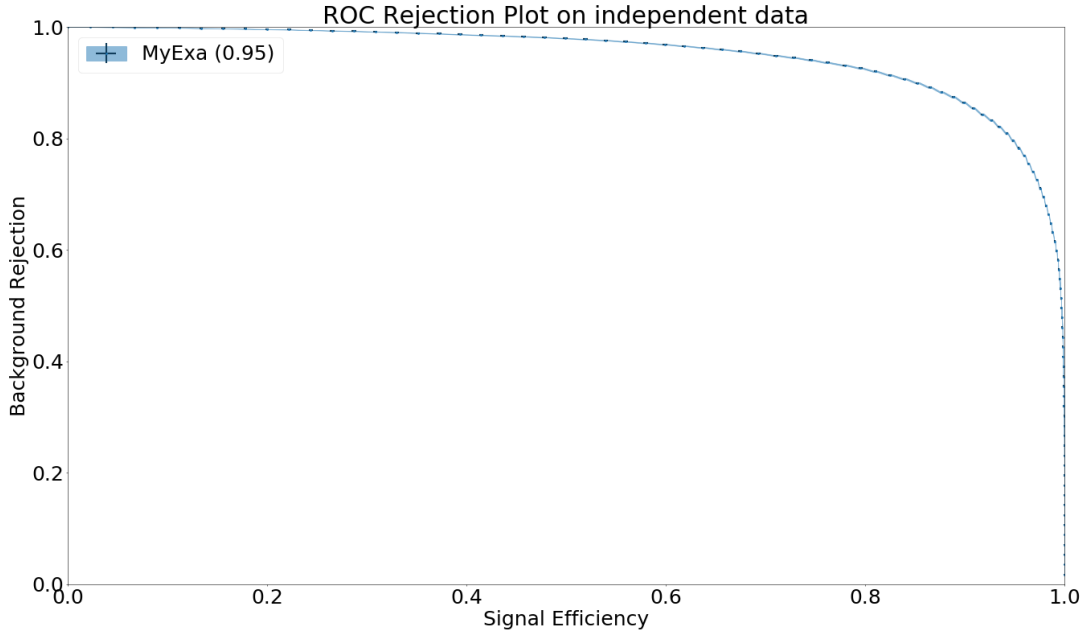
2.20 B0_K_10_K_S0_Rho

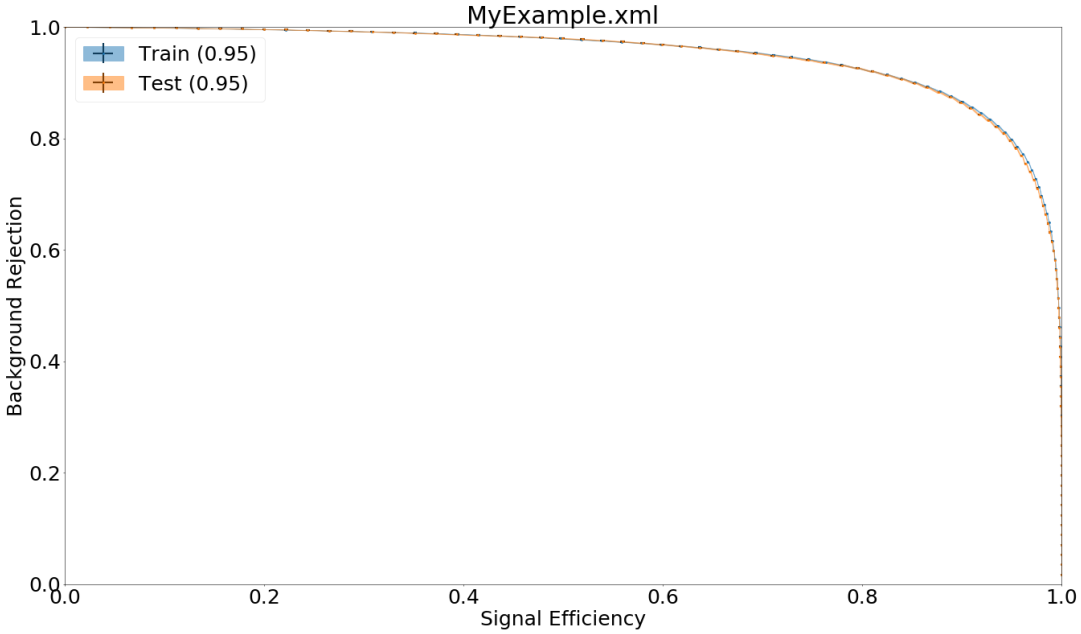


3 Classifier Plot

This section contains the receiver operating characteristics (ROC), purity projection, ...of the classifiers on training and independent data.The legend of each plot contains the shortened identifier and the area under the ROC curvein parenthesis.

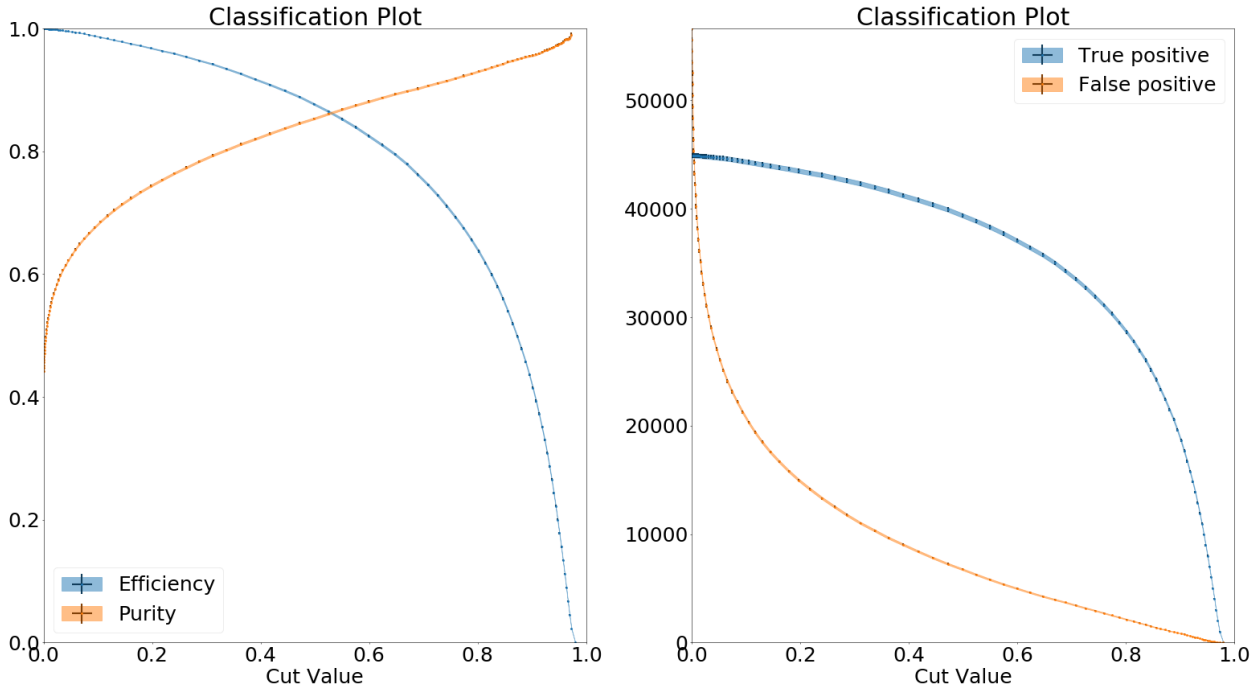
4 ROC Plot





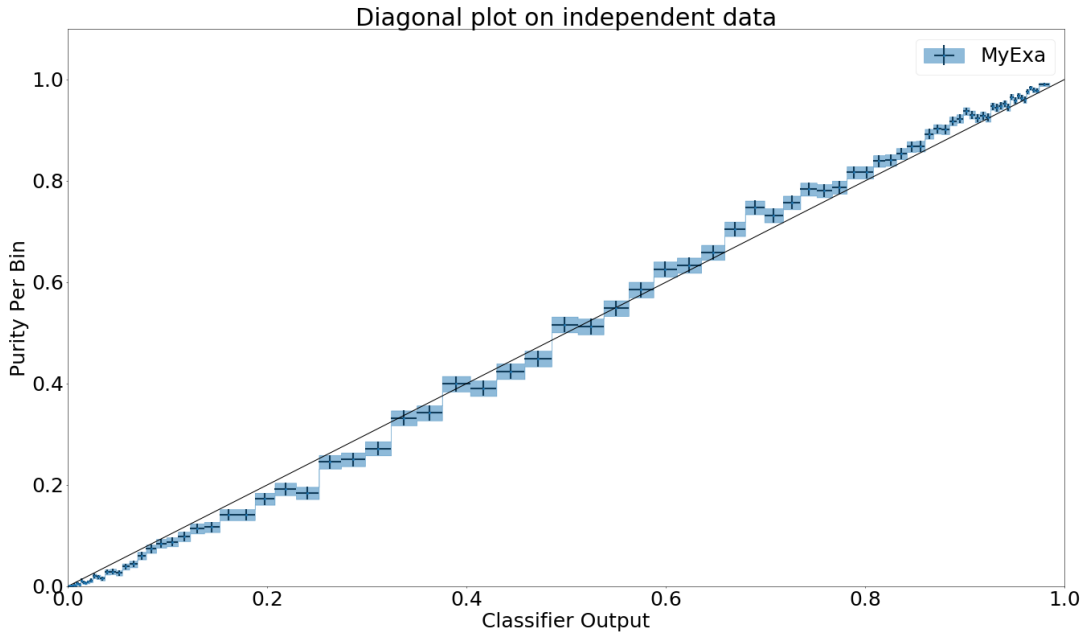
5 Classification Results

5.1 MyExa

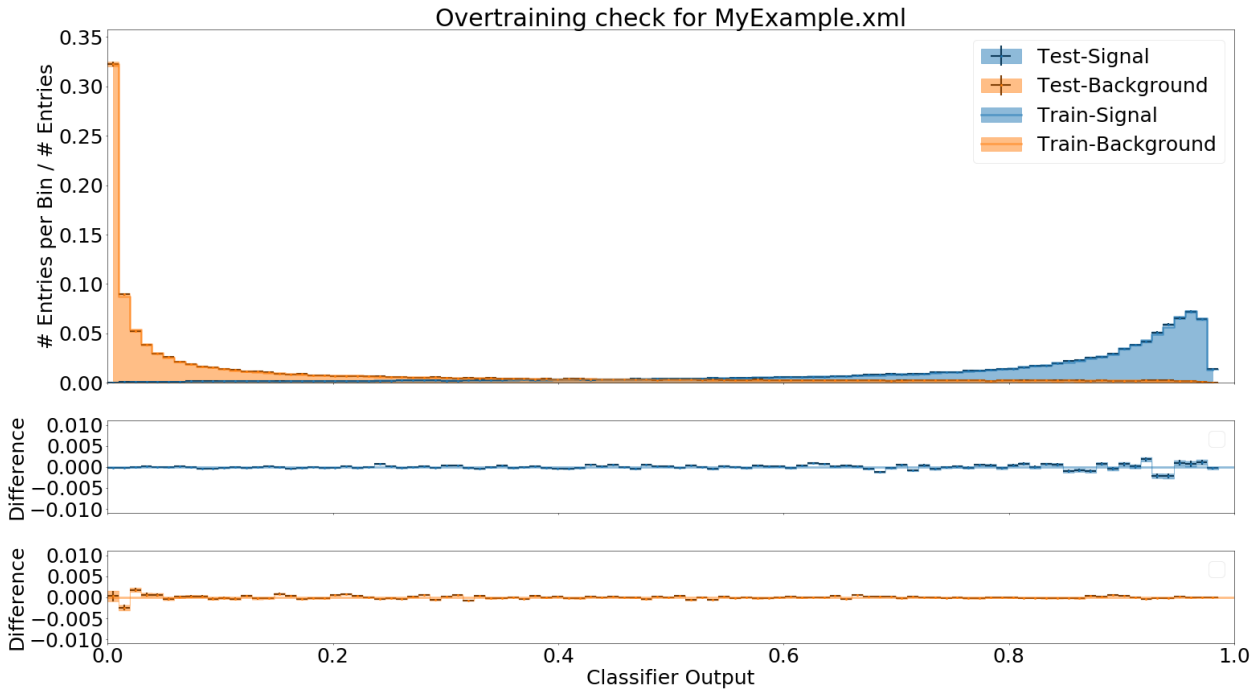


6 Diagonal Plot

6.1 MyExa



6.2 Overtraining Plot



7 Spectators

This section contains the distribution and dependence on theclassifier outputs of all spectator variables.

Table 3: Abbreviations of spectators

Spectator	Abbreviation