

# Automatic MVA Evaluation

Thomas Keck  
Moritz Gelb  
Nils Braun

August 9, 2019

## **Abstract**

Evaluation plots

# Contents

<b>1</b>	<b>Classifiers</b>	<b>3</b>
1.1	mva_cs_id_3.xml . . . . .	3
<b>2</b>	<b>Variables</b>	<b>3</b>
2.1	Importance . . . . .	4
2.2	Correlation . . . . .	4
2.3	CleoConeCS(3) . . . . .	5
2.4	KSFVariables(mm2) . . . . .	5
2.5	KSFVariables(et) . . . . .	6
2.6	cosTBz . . . . .	6
2.7	KSFVariables(hso22) . . . . .	7
2.8	thrustOm . . . . .	7
2.9	KSFVariables(hoo2) . . . . .	8
2.10	R2 . . . . .	8
2.11	KSFVariables(hso12) . . . . .	9
2.12	KSFVariables(hso02) . . . . .	9
2.13	thrustBm . . . . .	10
2.14	cosTBTO . . . . .	10
<b>3</b>	<b>Classifier Plot</b>	<b>10</b>
<b>4</b>	<b>ROC Plot</b>	<b>11</b>
<b>5</b>	<b>Classification Results</b>	<b>12</b>
5.1	mva_c . . . . .	12
<b>6</b>	<b>Diagonal Plot</b>	<b>12</b>
6.1	mva_c . . . . .	12
6.2	Overtraining Plot . . . . .	13
<b>7</b>	<b>Spectators</b>	<b>13</b>
7.1	mbc . . . . .	14
7.1.1	mbc with classifier mva_cs_id_3.xml . . . . .	14

# 1 Classifiers

This section contains the GeneralOptions and SpecificOptions of all classifiers represented by an XML tree. The same information can be retrieved using the basf2\_mva\_info tool.

Table 1: Abbreviations of identifiers

Identifier	Abbreviation
mva_cs_id_3.xml	mva_c

## 1.1 mva\_cs\_id\_3.xml

```
<?xml version="1.0" encoding="utf-8"?>
<method>FastBDT</method>
<weightfile>mva_cs_id_3.xml</weightfile>
<treename>tree</treename>
<target_variable>isSignal</target_variable>
<weight_variable/>
<signal_class>1</signal_class>
<max_events>0</max_events>
<number_feature_variables>12</number_feature_variables>
<variable0>R2</variable0>
<variable1>thrustBm</variable1>
<variable2>thrustOm</variable2>
<variable3>cosTBT0</variable3>
<variable4>cosTBz</variable4>
<variable5>KSFWVariables(et)</variable5>
<variable6>KSFWVariables(mm2)</variable6>
<variable7>KSFWVariables(hso02)</variable7>
<variable8>KSFWVariables(hso12)</variable8>
<variable9>KSFWVariables(hso22)</variable9>
<variable10>KSFWVariables(hoo2)</variable10>
<variable11>CleoConeCS(3)</variable11>
<number_spectator_variables>1</number_spectator_variables>
<spectator0>mbc</spectator0>
<number_data_files>1</number_data_files>
<datafile0>../eganiev/analysis/phase3/train.root</datafile0>
<FastBDT_version>2</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>
<FastBDT_flatnessLoss>-1</FastBDT_flatnessLoss>
<FastBDT_sPlot>false</FastBDT_sPlot>
<FastBDT_number_individual_nCuts>0</FastBDT_number_individual_nCuts>
<FastBDT_purityTransformation>false</FastBDT_purityTransformation>
<FastBDT_number_individualPurityTransformation>0</FastBDT_number_individualPurityTransformation>
```

## 2 Variables

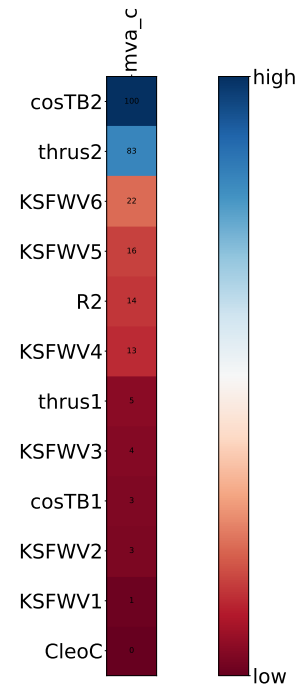
This section contains an overview of the importance and correlation of the variables used by the classifiers. And distribution plots of the variables on the independent dataset. The distributions are normed for signal and background separately, and only the region  $\pm 3$  sigma around the mean is shown.

Table 2: Abbreviations of variables

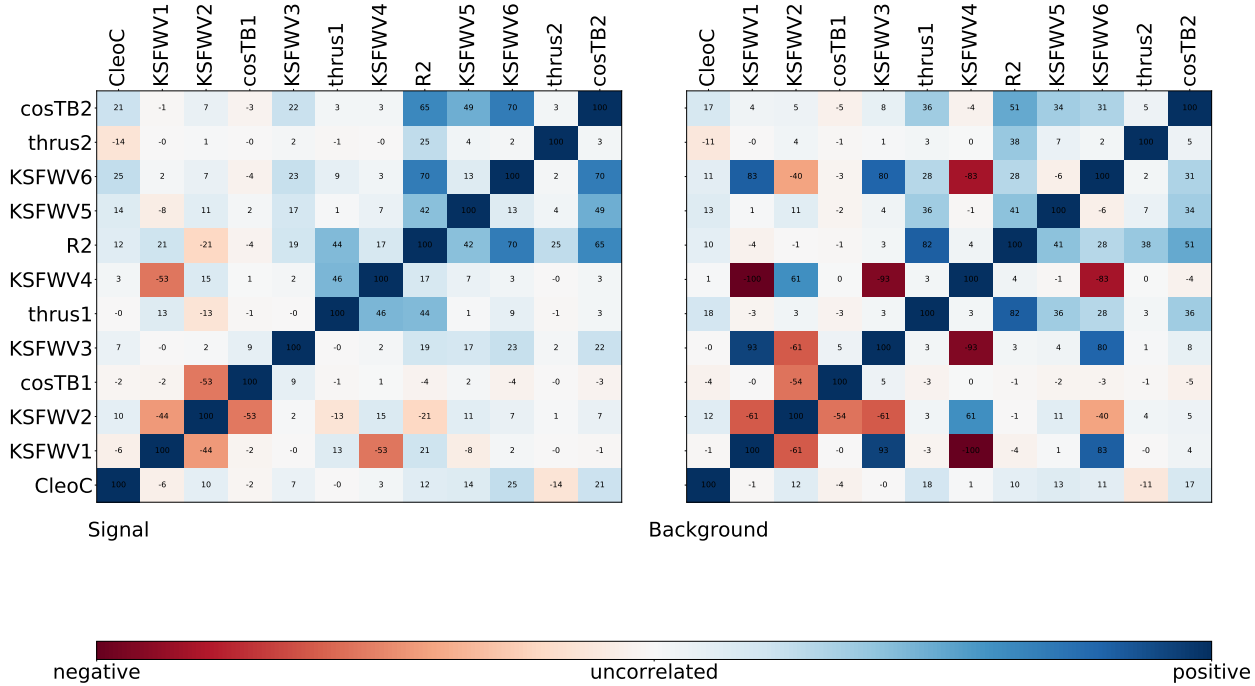
Variable	Abbreviation
CleoConeCS(3)	CleoC
KSFWVariables(mm2)	KSFWV1
KSFWVariables(et)	KSFWV2
cosTBz	cosTB1
KSFWVariables(hso22)	KSFWV3
thrustOm	thrus1
KSFWVariables(hoo2)	KSFWV4
R2	R2
KSFWVariables(hso12)	KSFWV5
KSFWVariables(hso02)	KSFWV6

thrus2  
cosTB2

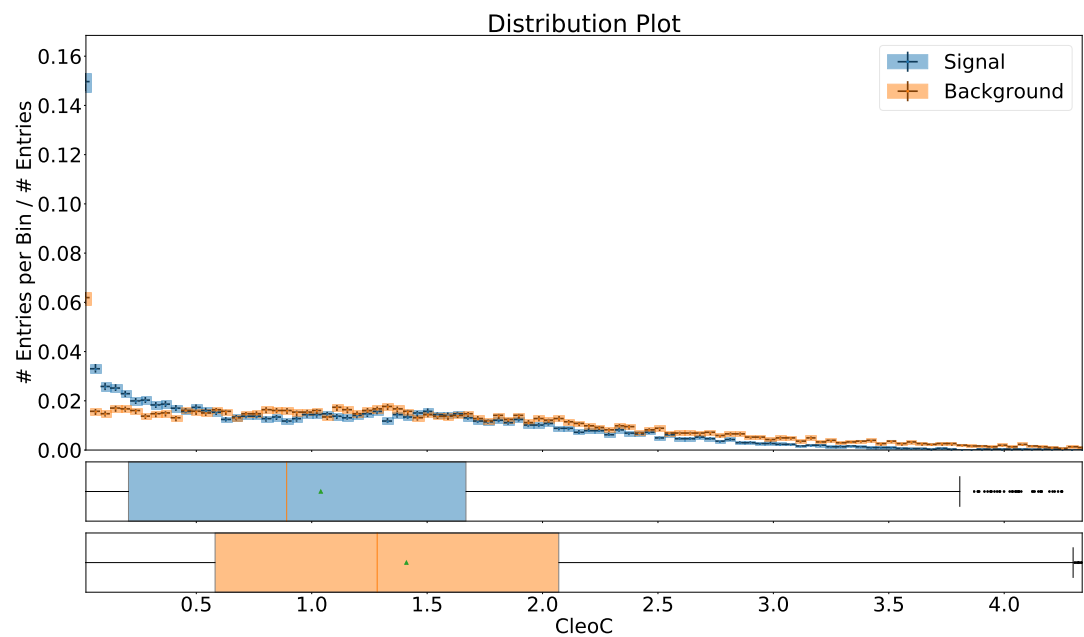
## 2.1 Importance



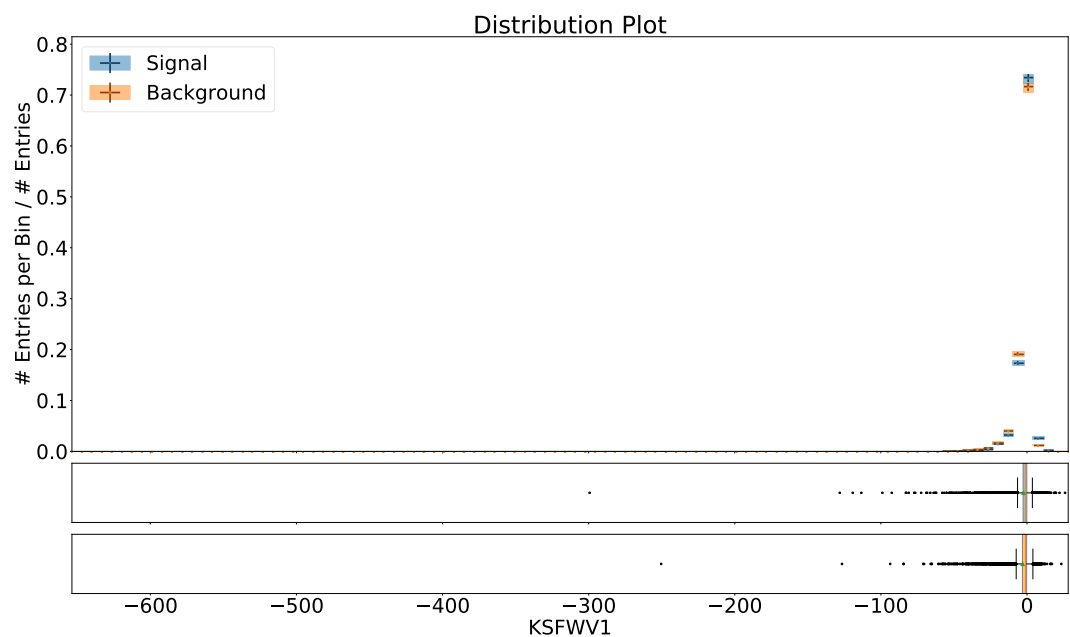
## 2.2 Correlation



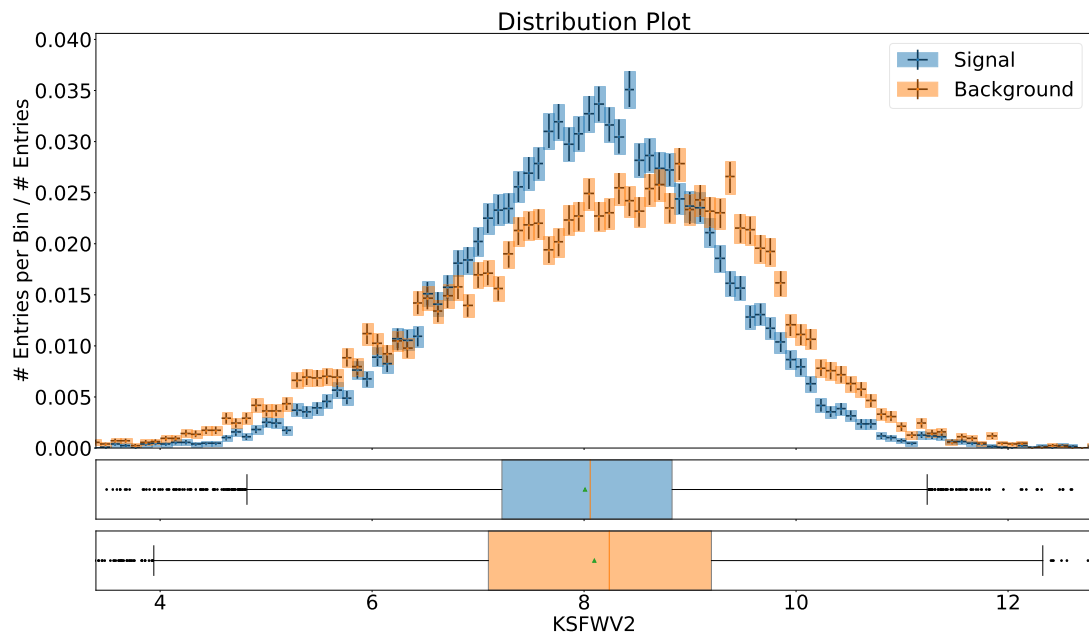
2.3 CleoConeCS(3)



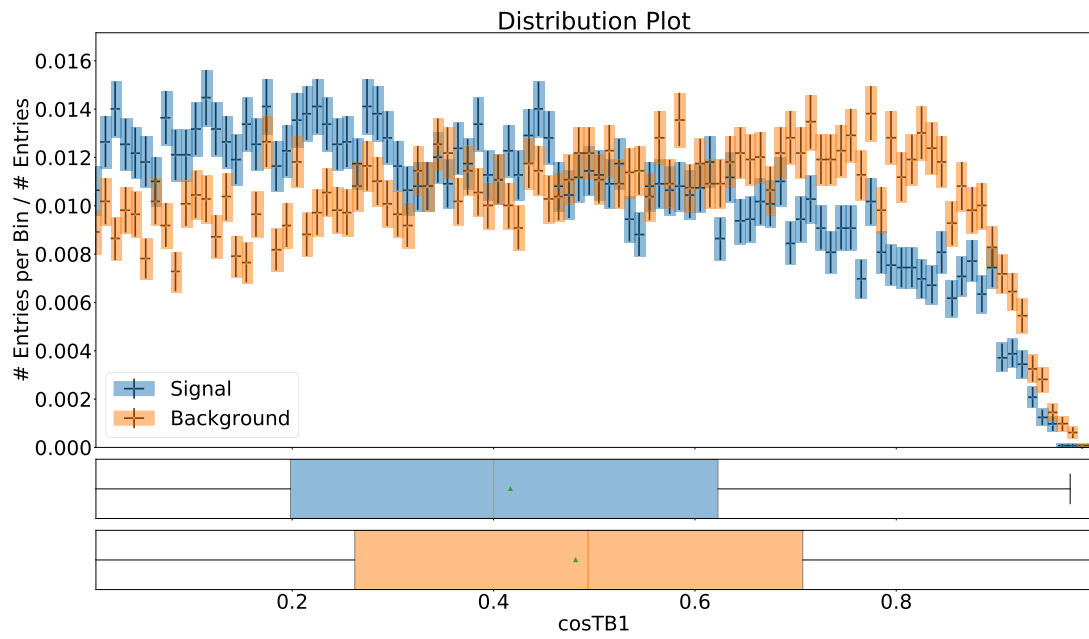
2.4 KSFVVariables(mm2)



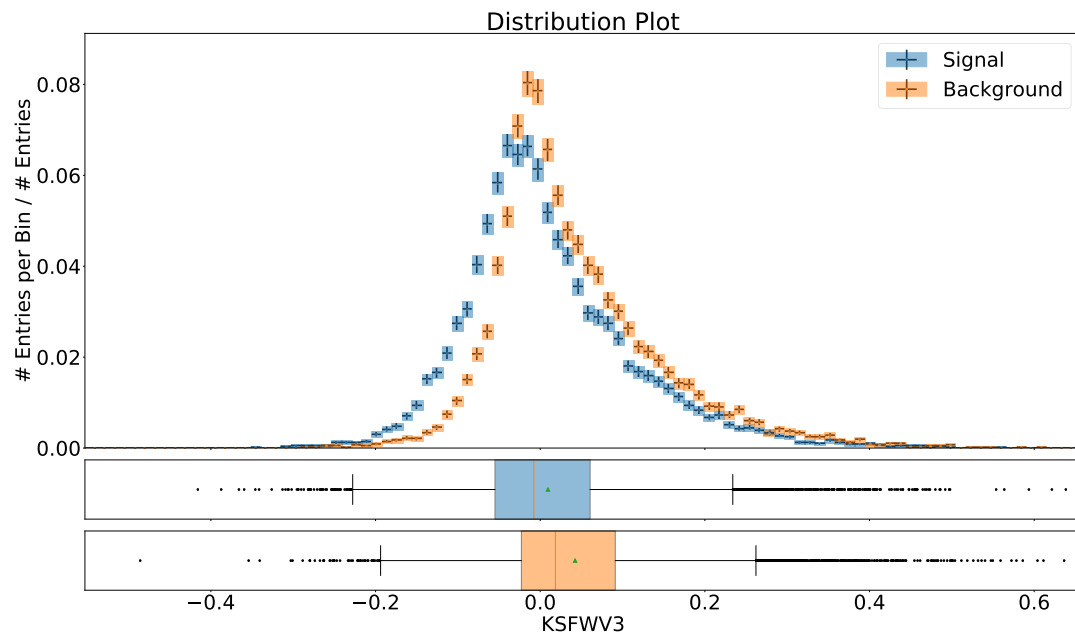
## 2.5 KSFWVariables(et)



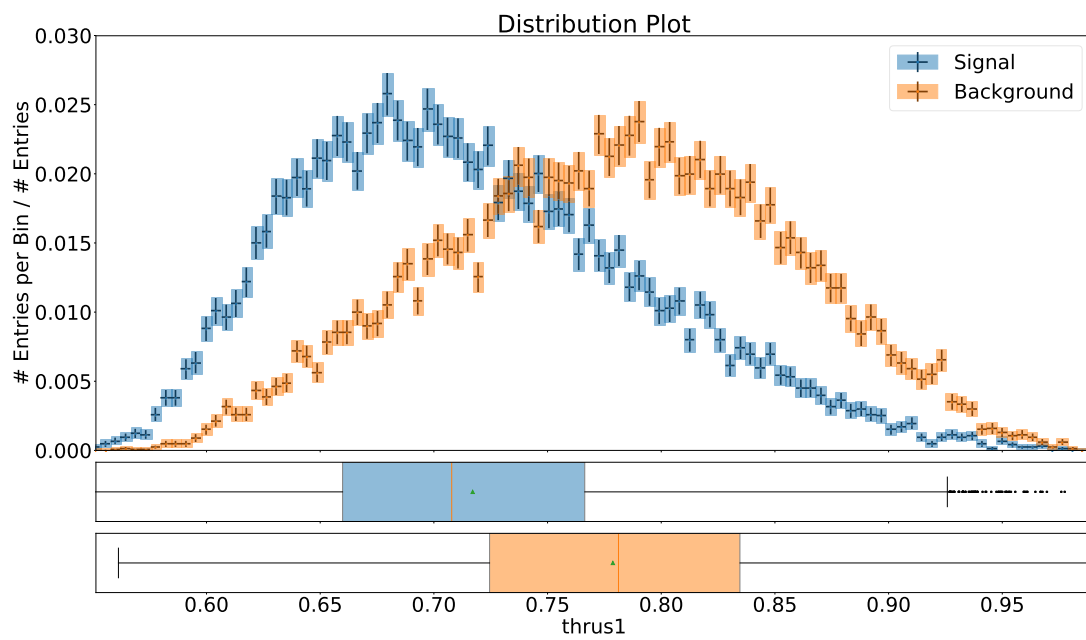
## 2.6 $\cos\text{TBz}$



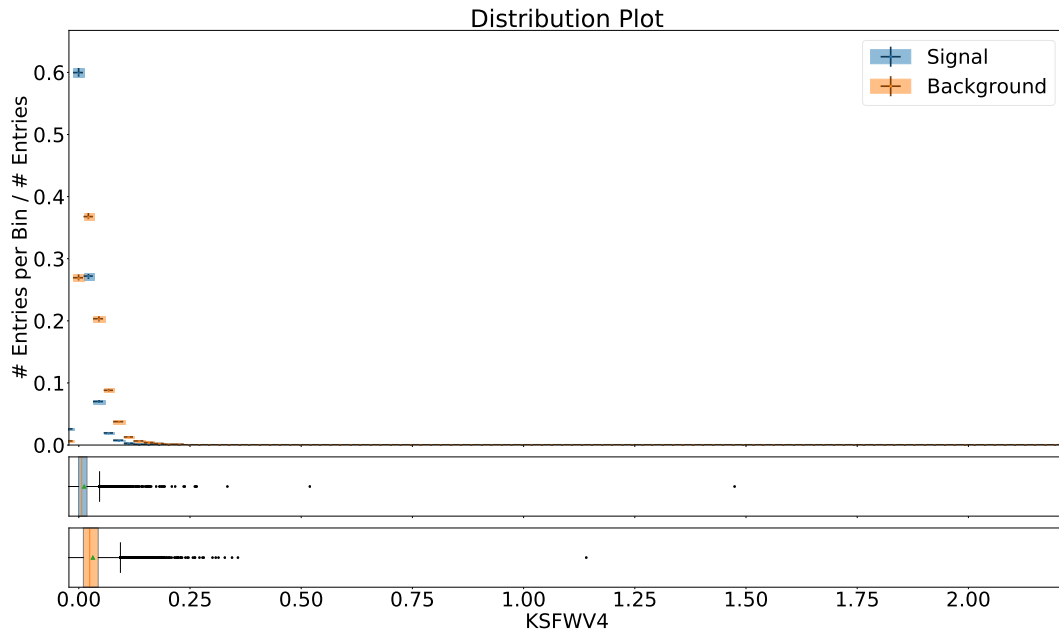
## 2.7 KSFVVariables(hso22)



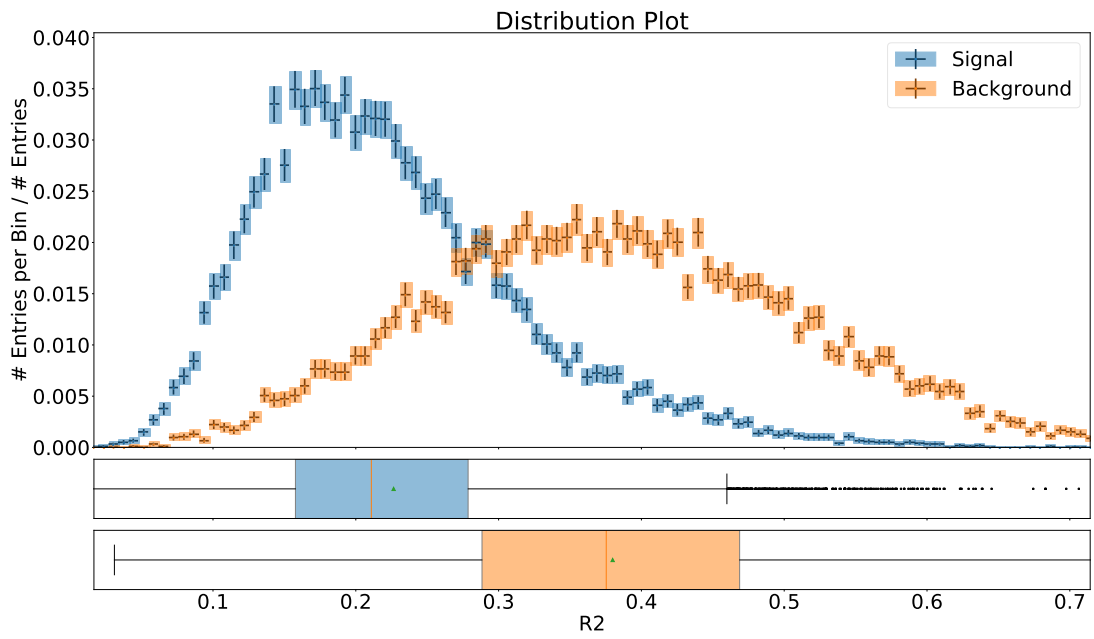
## 2.8 thrustOm



## 2.9 KSFVVariables(hoo2)

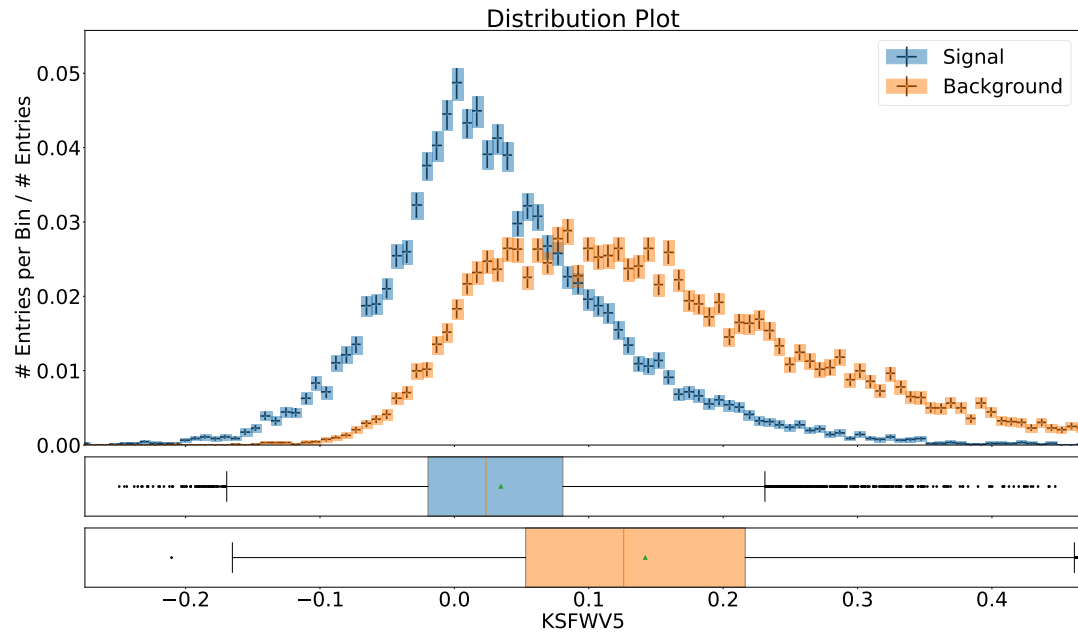


## 2.10 R2

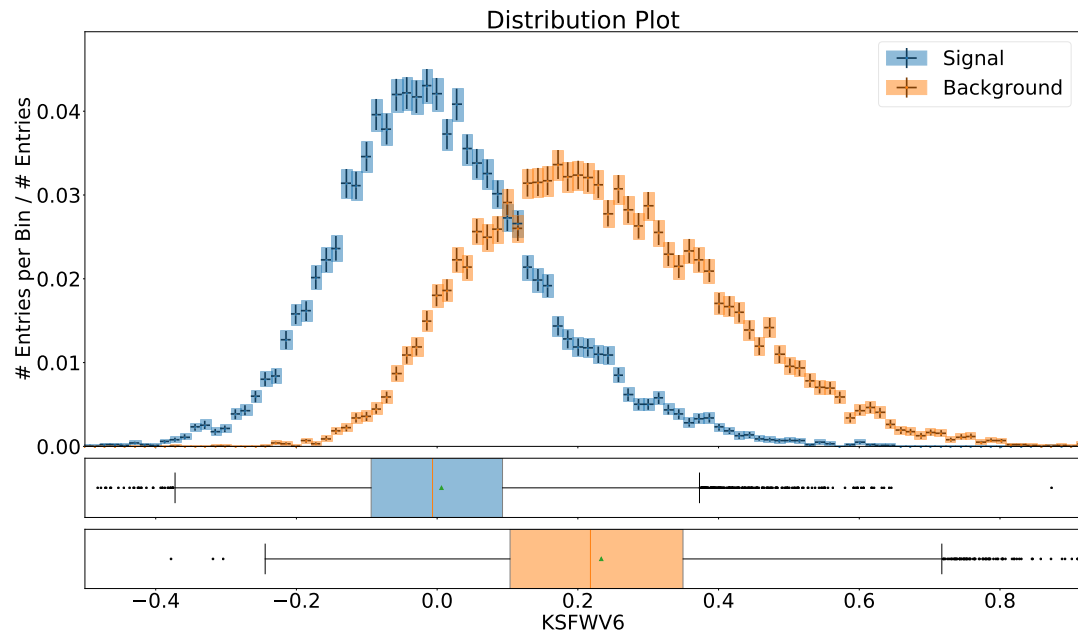




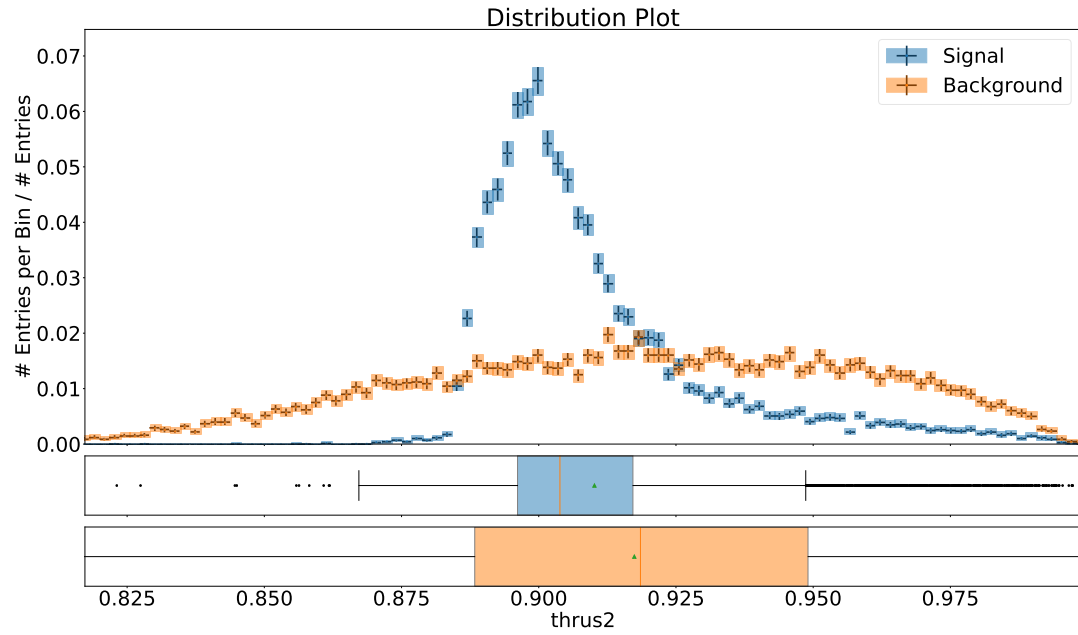
## 2.11 KSFVVariables(hso12)



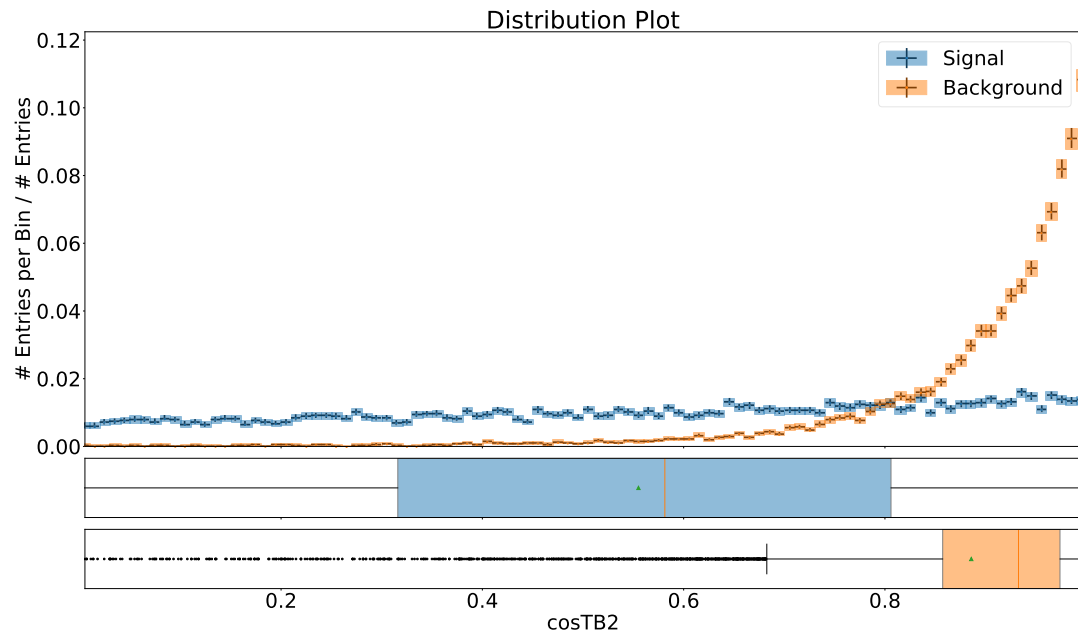
## 2.12 KSFVVariables(hso02)



### 2.13 thrustBm



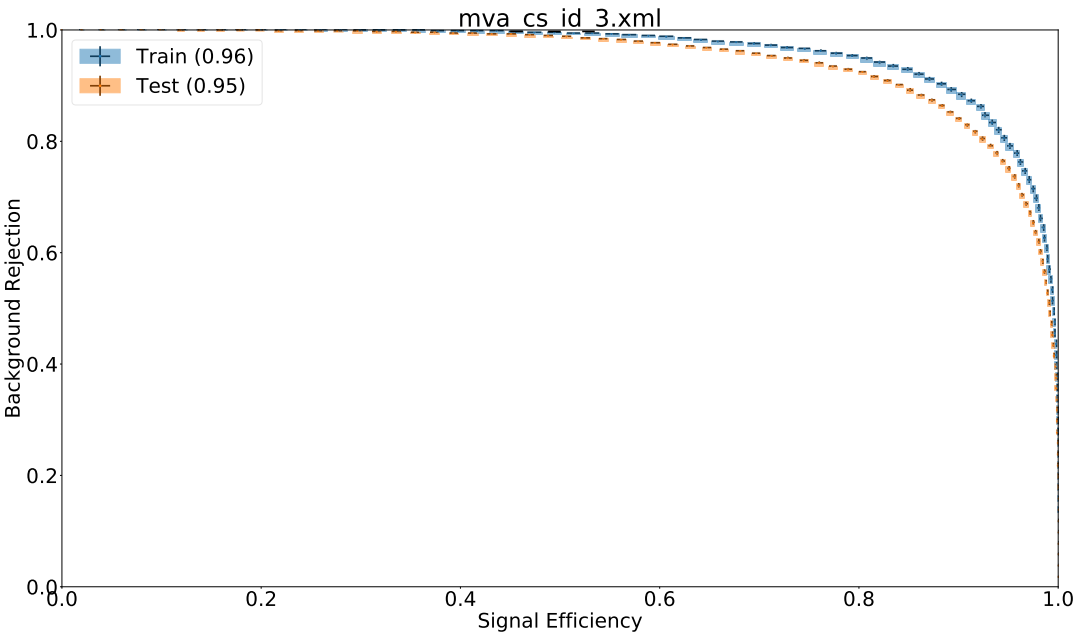
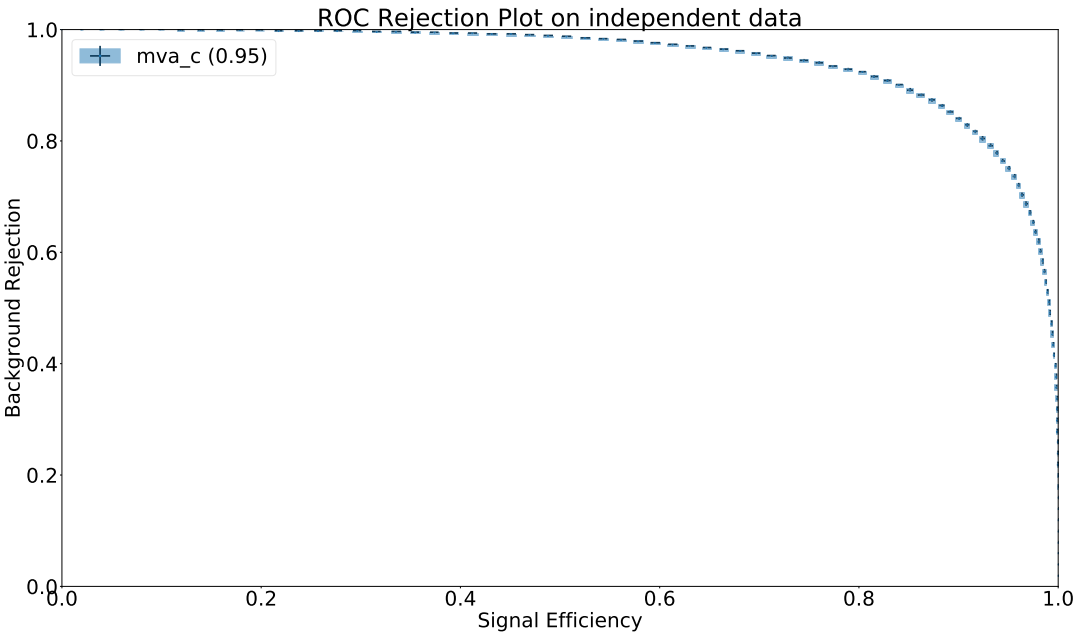
### 2.14 cosTBTO



## 3 Classifier Plot

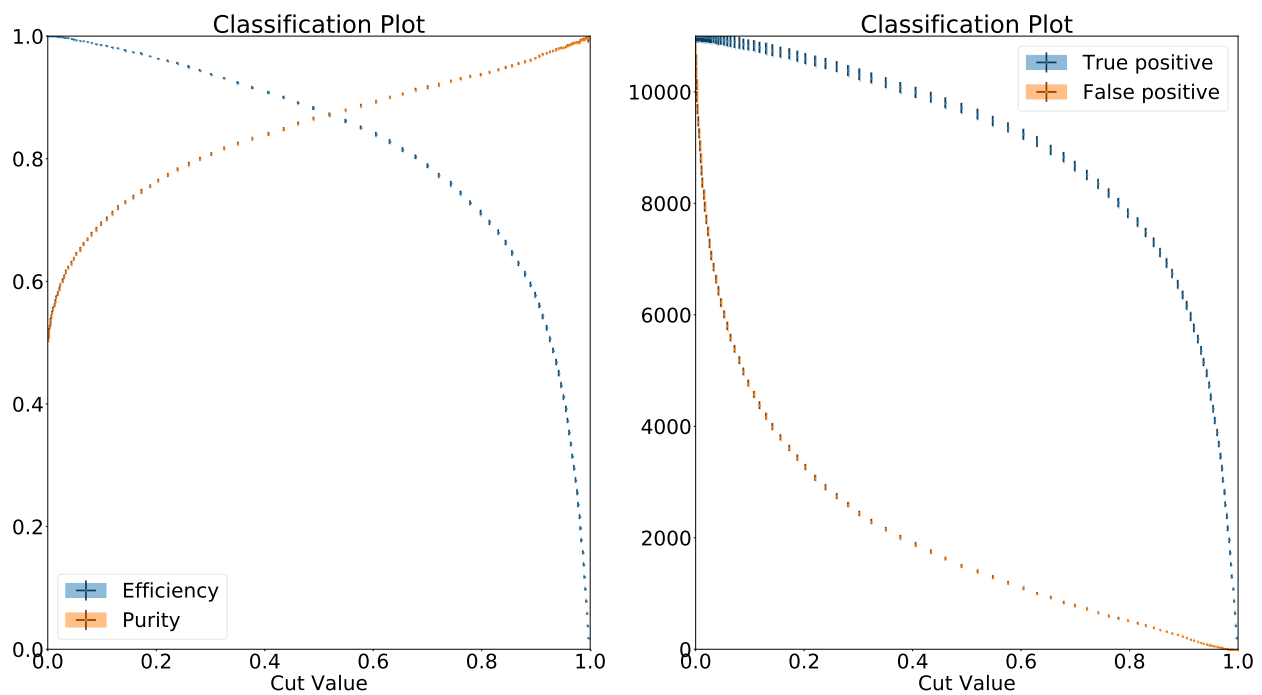
This section contains the receiver operating characteristics (ROC), purity projection, ...of the classifiers on training and independent data. The legend of each plot contains the shortened identifier and the area under the ROC curve in parenthesis.

# 4 ROC Plot



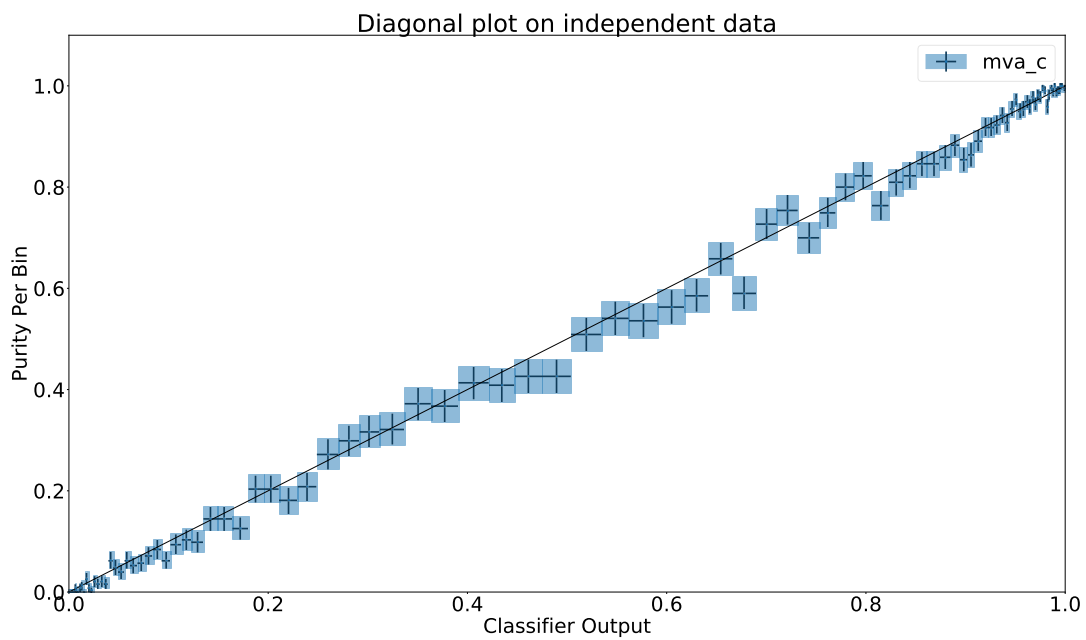
## 5 Classification Results

### 5.1 mva\_c

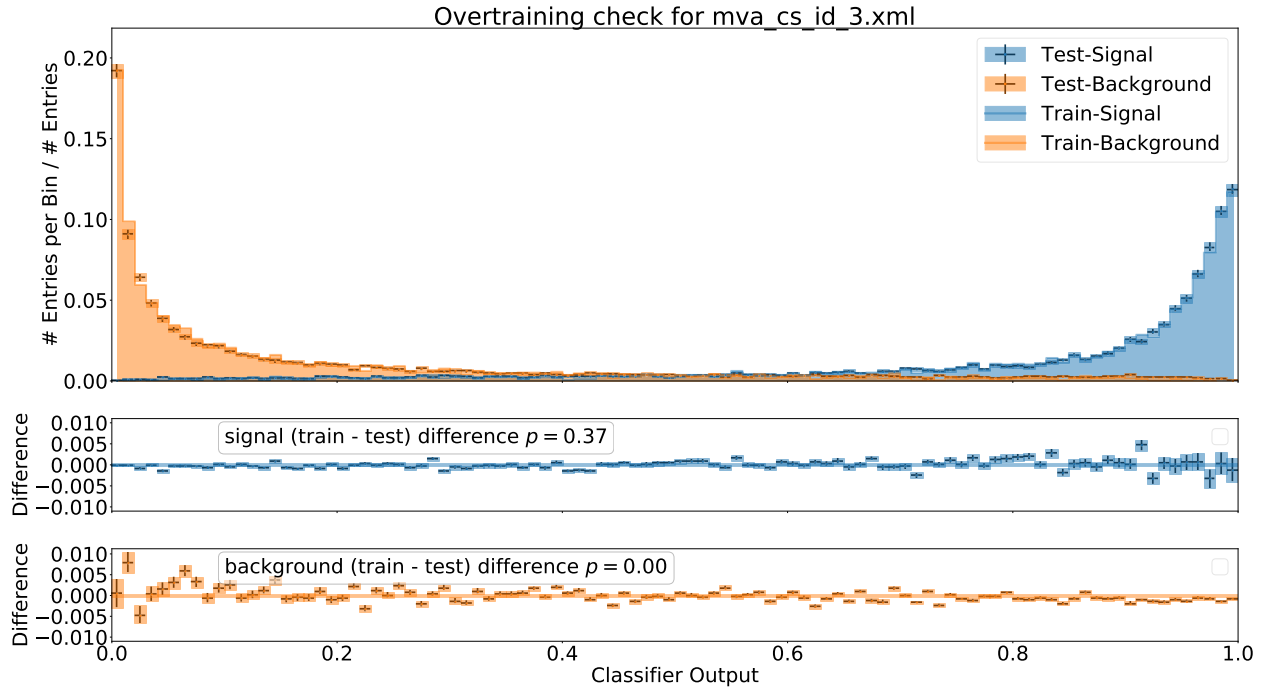


## 6 Diagonal Plot

### 6.1 mva\_c



## 6.2 Overtraining Plot



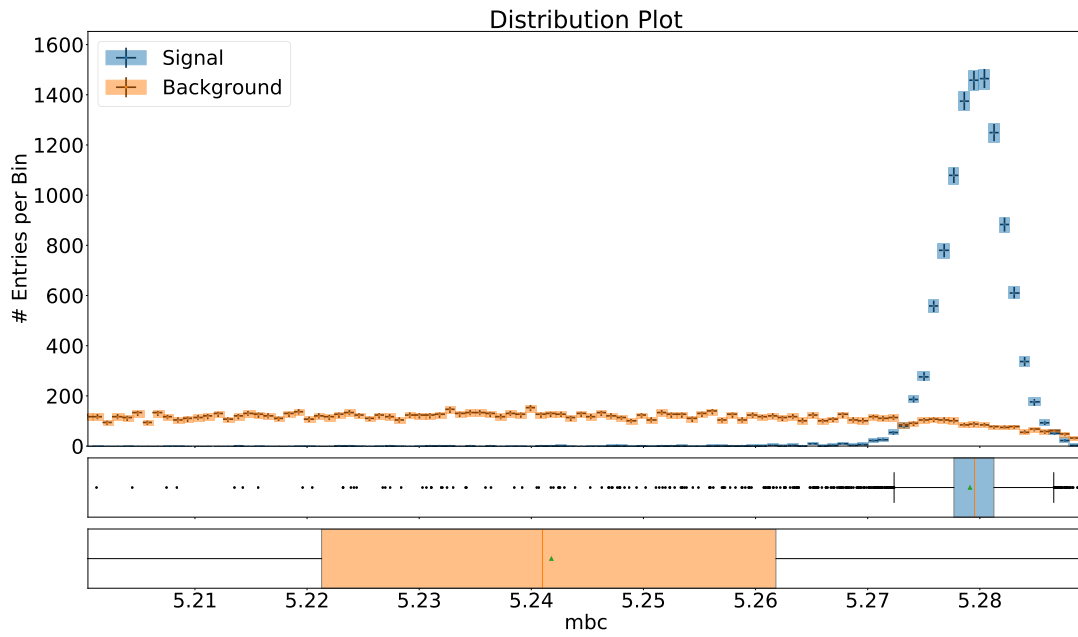
## 7 Spectators

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

Table 3: Abbreviations of spectators

Spectator	Abbreviation
mbc	mbc

## 7.1 mbc



### 7.1.1 mbc with classifier mva\_cs\_id\_3.xml

