

Automatic MVA Evaluation

Thomas Keck
Moritz Gelb
Nils Braun

August 8, 2019

Abstract

Evaluation plots

Contents

1	Classifiers	3
1.1	mva_cs_id_1.xml	3
2	Variables	3
2.1	Importance	4
2.2	Correlation	5
2.3	KSFVariables(hoo4)	5
2.4	KSFVariables(hso00)	6
2.5	CleoConeCS(7)	6
2.6	KSFVariables(hoo3)	7
2.7	KSFVariables(hoo1)	7
2.8	KSFVariables(mm2)	8
2.9	CleoConeCS(8)	8
2.10	KSFVariables(hso22)	9
2.11	KSFVariables(hso20)	9
2.12	KSFVariables(hso14)	10
2.13	KSFVariables(hso10)	10
2.14	KSFVariables(hoo0)	11
2.15	KSFVariables(hso24)	11
2.16	KSFVariables(et)	12
2.17	cosTBz	12
2.18	KSFVariables(hso04)	13
2.19	thrustOm	13
2.20	R2	14
2.21	KSFVariables(hoo2)	14
2.22	KSFVariables(hso12)	15
2.23	KSFVariables(hso02)	15
2.24	thrustBm	16
2.25	cosTBTO	16
3	Classifier Plot	16
4	ROC Plot	17
5	Classification Results	18
5.1	mva_c	18
6	Diagonal Plot	18
6.1	mva_c	18
6.2	Overtraining Plot	19
7	Spectators	19
7.1	mbc	20
7.1.1	mbc with classifier mva_cs_id_1.xml	20

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_1.xml	mva_c

1.1 mva_cs_id_1.xml

```
<?xml version="1.0" encoding="utf-8"?>
<method>FastBDT</method>
<weightfile>mva_cs_id_1.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>23</number_feature_variables>
<variable0>R2</variable0>
<variable1>thrustBm</variable1>
<variable2>thrustOm</variable2>
<variable3>cosTBTO</variable3>
<variable4>cosTBz</variable4>
<variable5>KSFWVariables(et)</variable5>
<variable6>KSFWVariables(mm2)</variable6>
<variable7>KSFWVariables(hso00)</variable7>
<variable8>KSFWVariables(hso02)</variable8>
<variable9>KSFWVariables(hso04)</variable9>
<variable10>KSFWVariables(hso10)</variable10>
<variable11>KSFWVariables(hso12)</variable11>
<variable12>KSFWVariables(hso14)</variable12>
<variable13>KSFWVariables(hso20)</variable13>
<variable14>KSFWVariables(hso22)</variable14>
<variable15>KSFWVariables(hso24)</variable15>
<variable16>KSFWVariables(hoo0)</variable16>
<variable17>KSFWVariables(hoo1)</variable17>
<variable18>KSFWVariables(hoo2)</variable18>
<variable19>KSFWVariables(hoo3)</variable19>
<variable20>KSFWVariables(hoo4)</variable20>
<variable21>CleoConeCS(7)</variable21>
<variable22>CleoConeCS(8)</variable22>
<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>0.1000000000000001</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

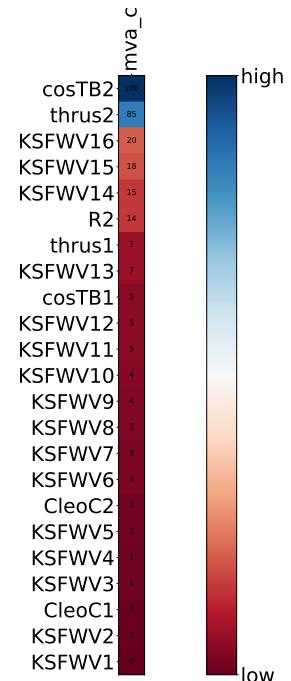
This section contains an overview of the importance and correlation of the variables used by the classifiers. And distribution plots of the variables on the independent dataset. The distributions are normed for signal and background separately, and only the region ± 3 sigma around the mean is shown.

Table 2: Abbreviations of variables

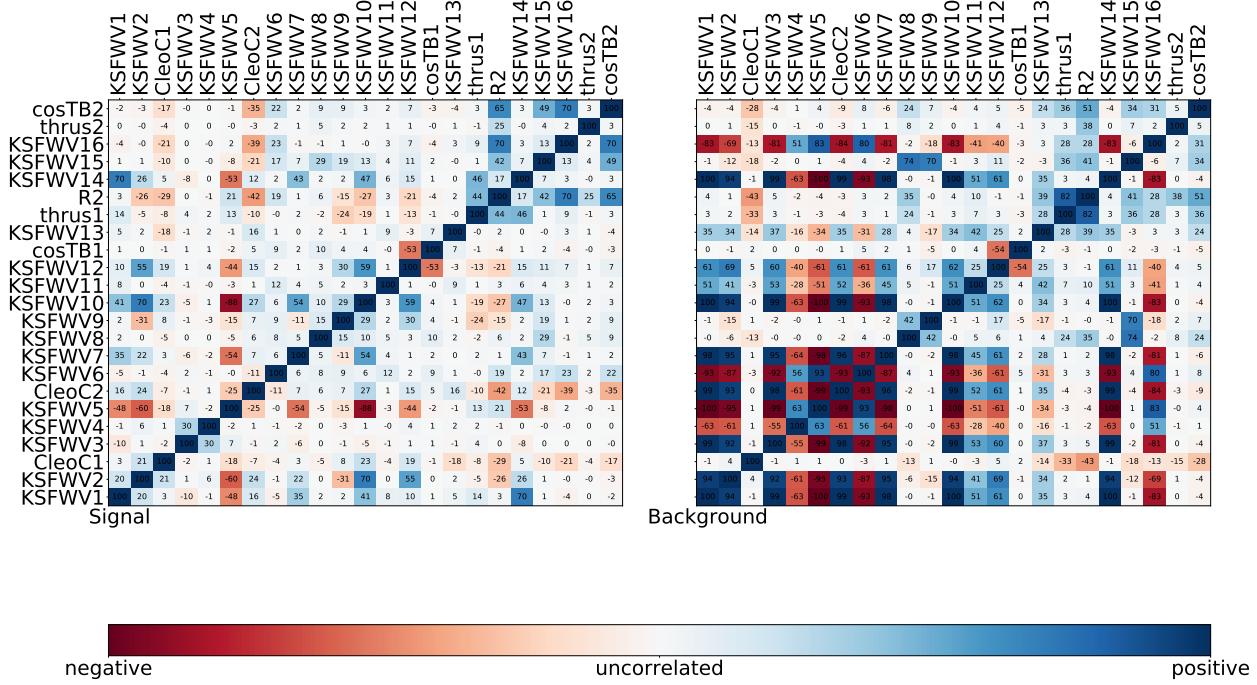
Variable	Abbreviation
KSFWVariables(hoo4)	KSFWV1
KSFWVariables(hso00)	KSFWV2
CleoConeCS(7)	CleoC1

KSFVVariables(hoo3)	KSFV3
KSFVVariables(hoo1)	KSFV4
KSFVVariables(mm2)	KSFV5
CleoConeCS(8)	CleoC2
KSFVVariables(hso22)	KSFV6
KSFVVariables(hso20)	KSFV7
KSFVVariables(hso14)	KSFV8
KSFVVariables(hso10)	KSFV9
KSFVVariables(hoo0)	KSFV10
KSFVVariables(hso24)	KSFV11
KSFVVariables(et)	KSFV12
cosTBz	cosTB1
KSFVVariables(hso04)	KSFV13
thrustOm	thrus1
R2	R2
KSFVVariables(hoo2)	KSFV14
KSFVVariables(hso12)	KSFV15
KSFVVariables(hso02)	KSFV16
thrustBm	thrus2
cosTBTO	cosTB2

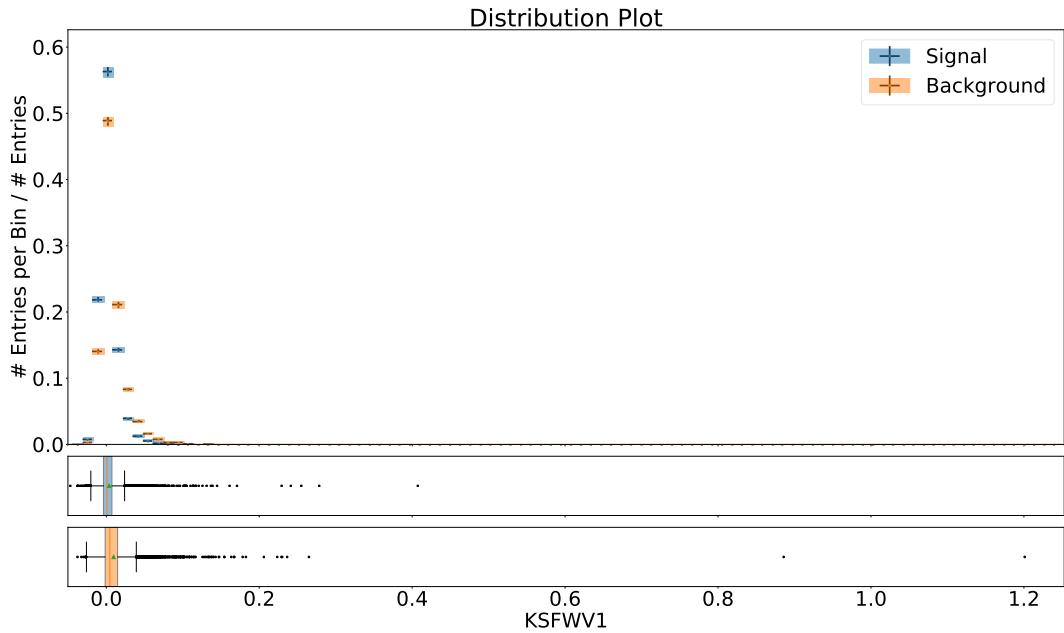
2.1 Importance



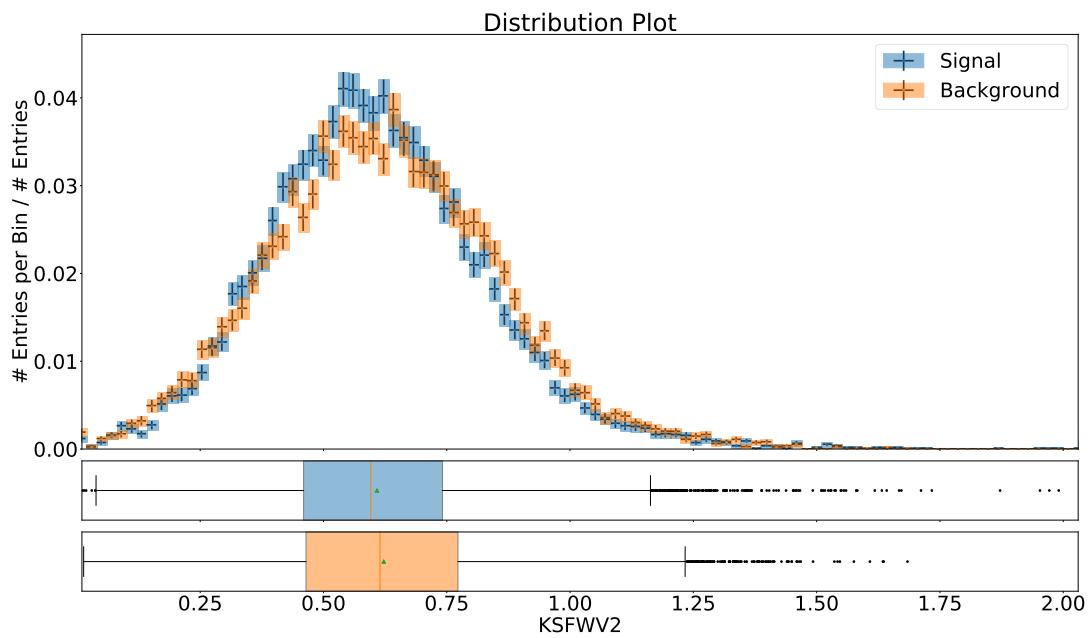
2.2 Correlation



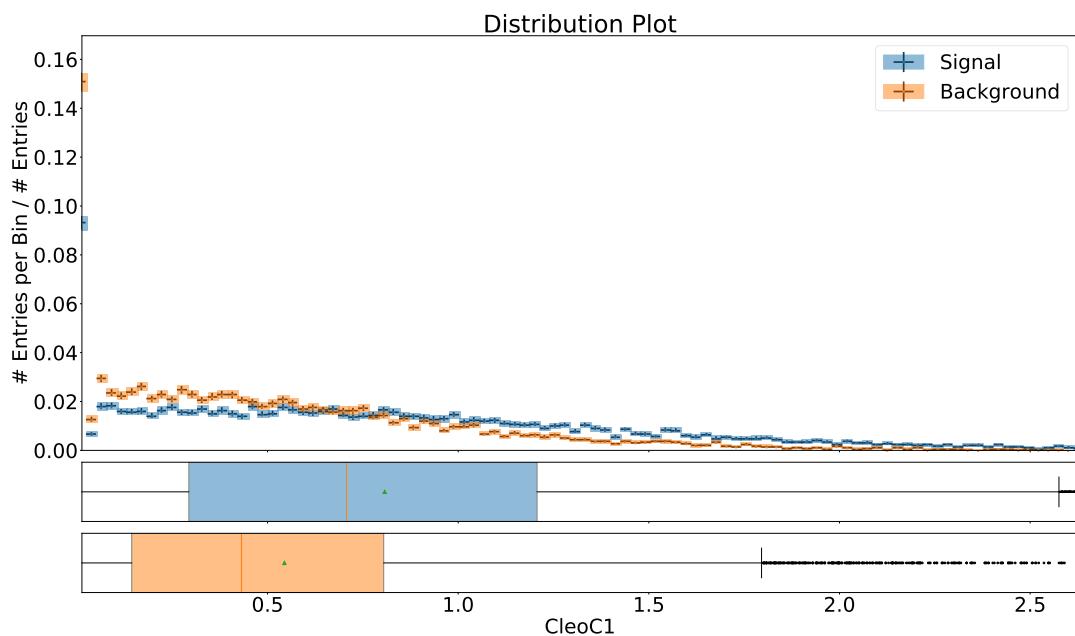
2.3 KSFVariables(hoo4)



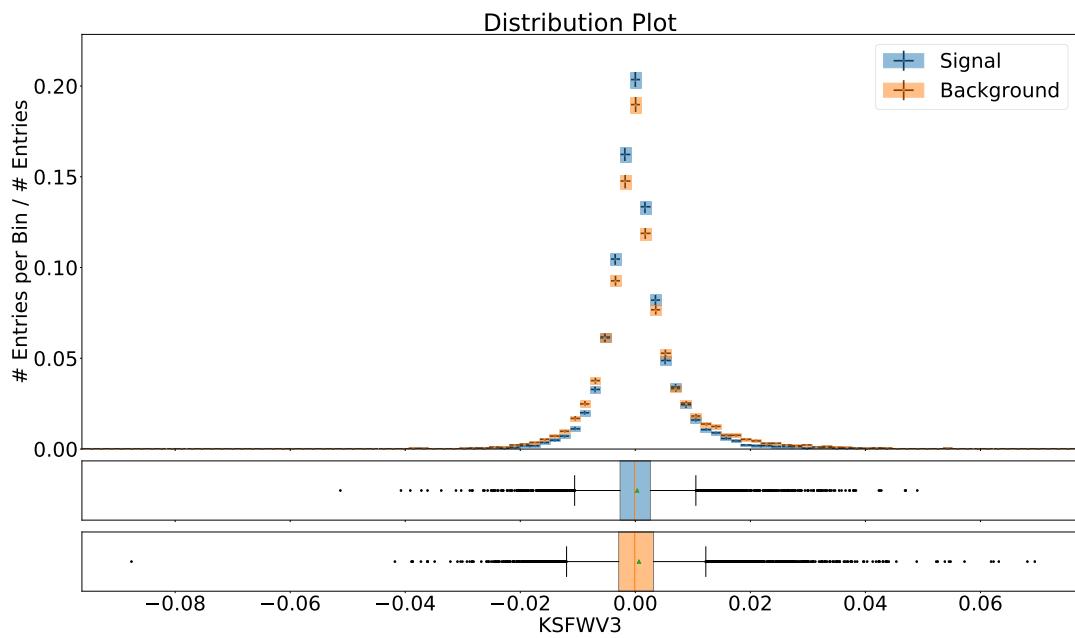
2.4 KSFVWVariables(hso00)



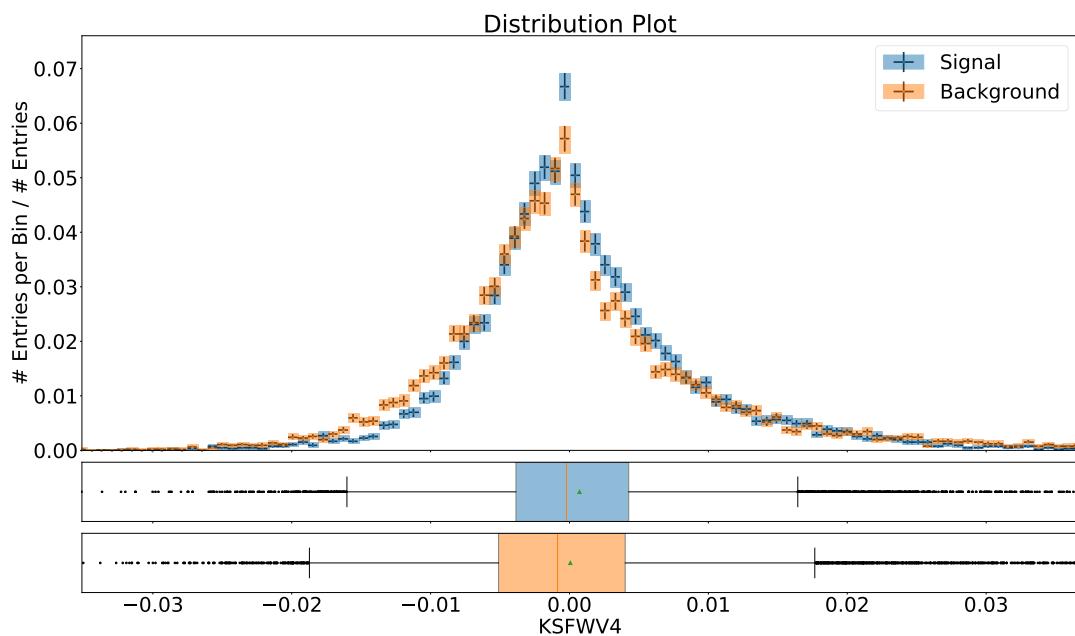
2.5 CleoConeCS(7)



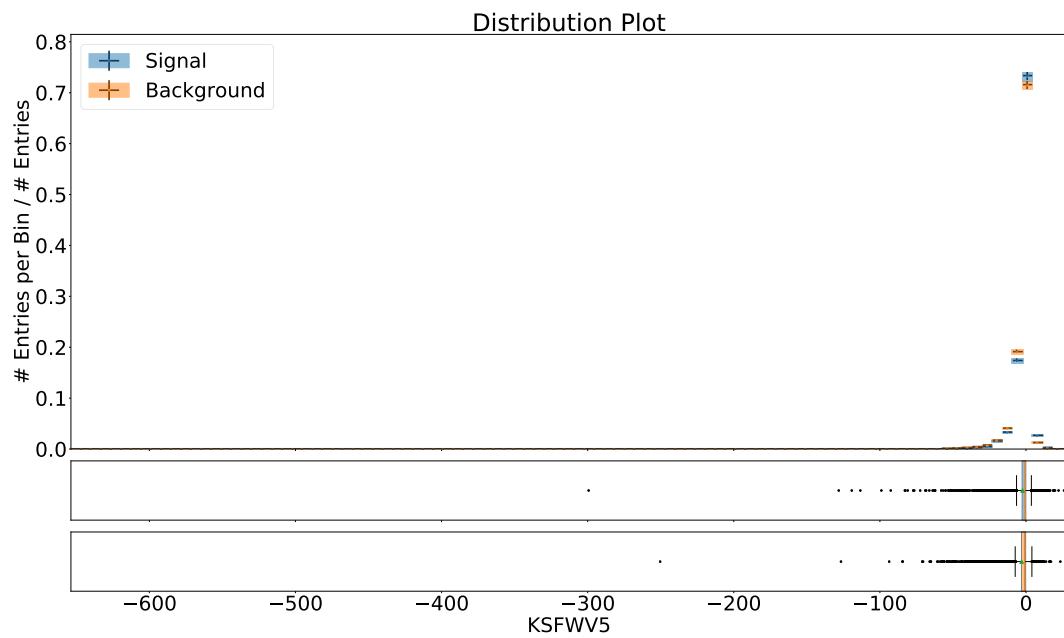
2.6 KSFVariables(hoo3)



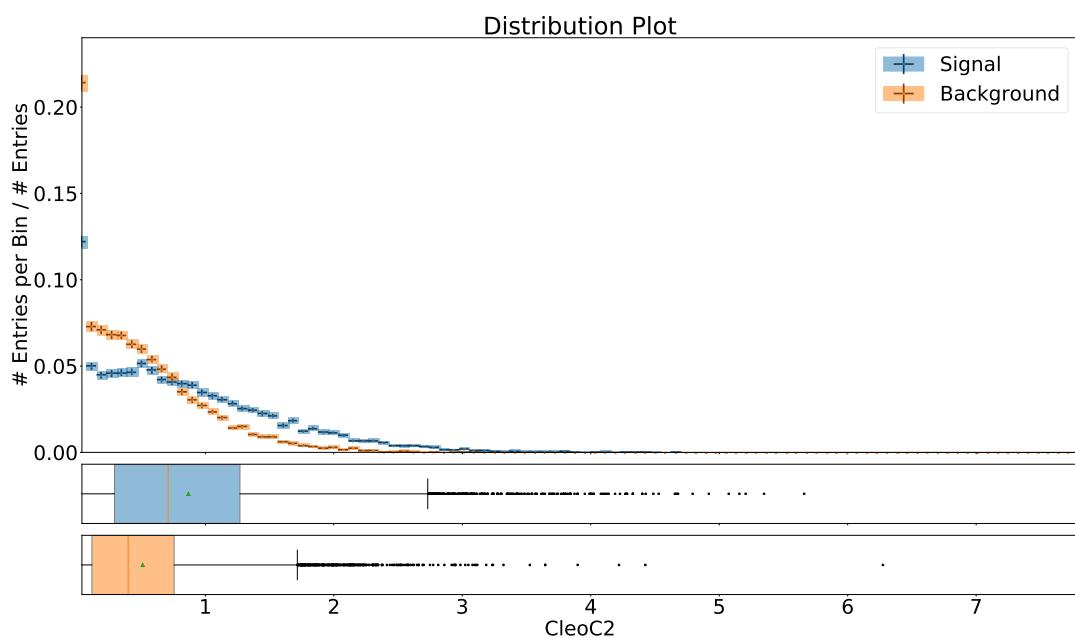
2.7 KSFVariables(hoo1)



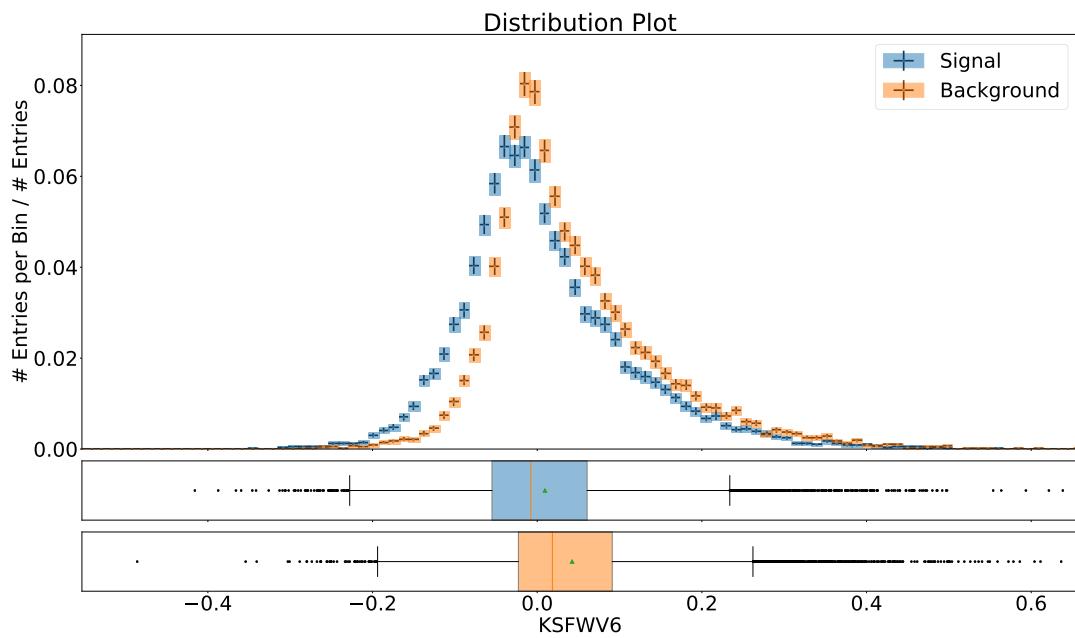
2.8 KSFVWVariables(mm2)



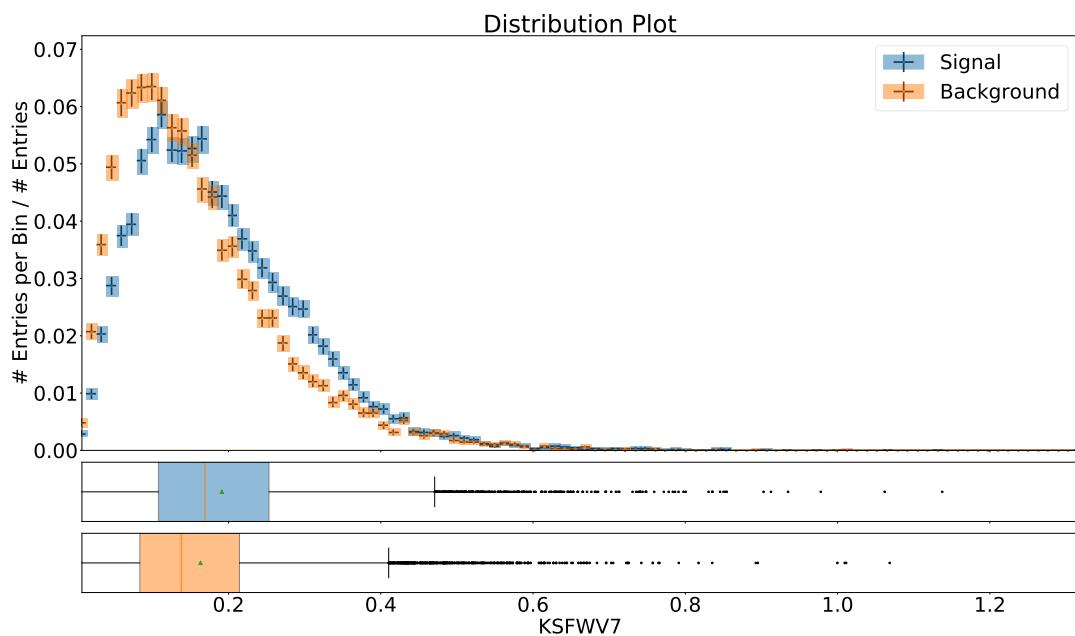
2.9 CleoConeCS(8)



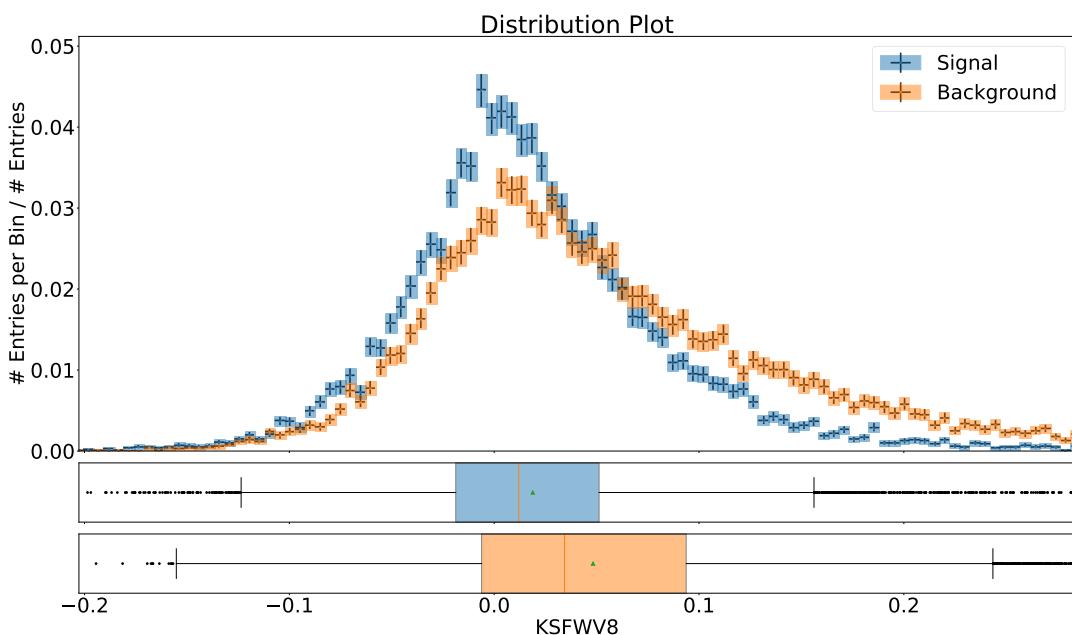
2.10 KSFVVariables(hso22)



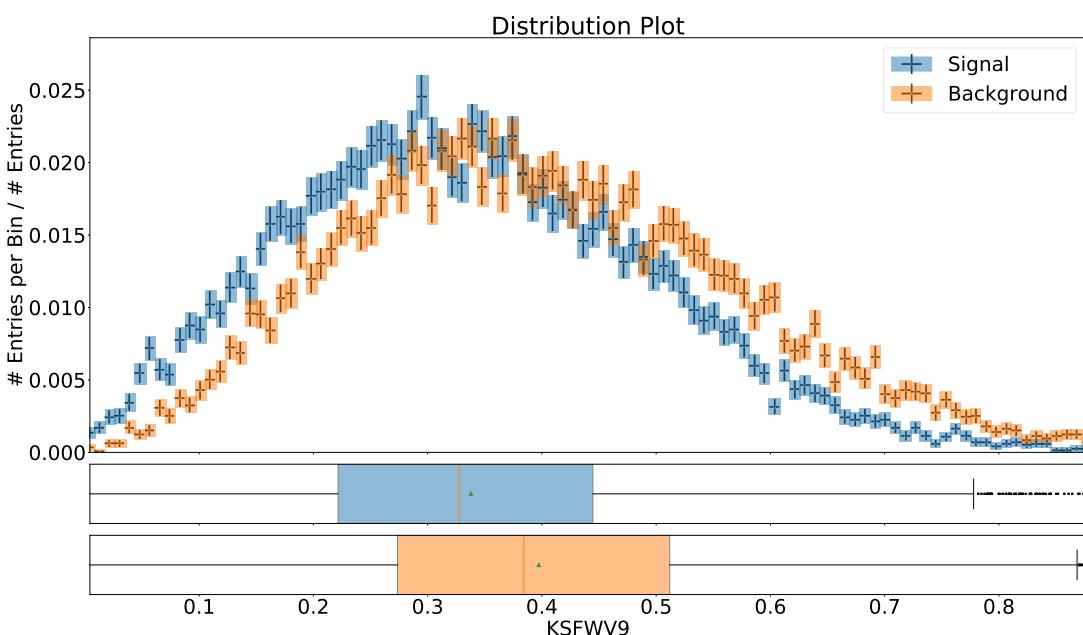
2.11 KSFVVariables(hso20)



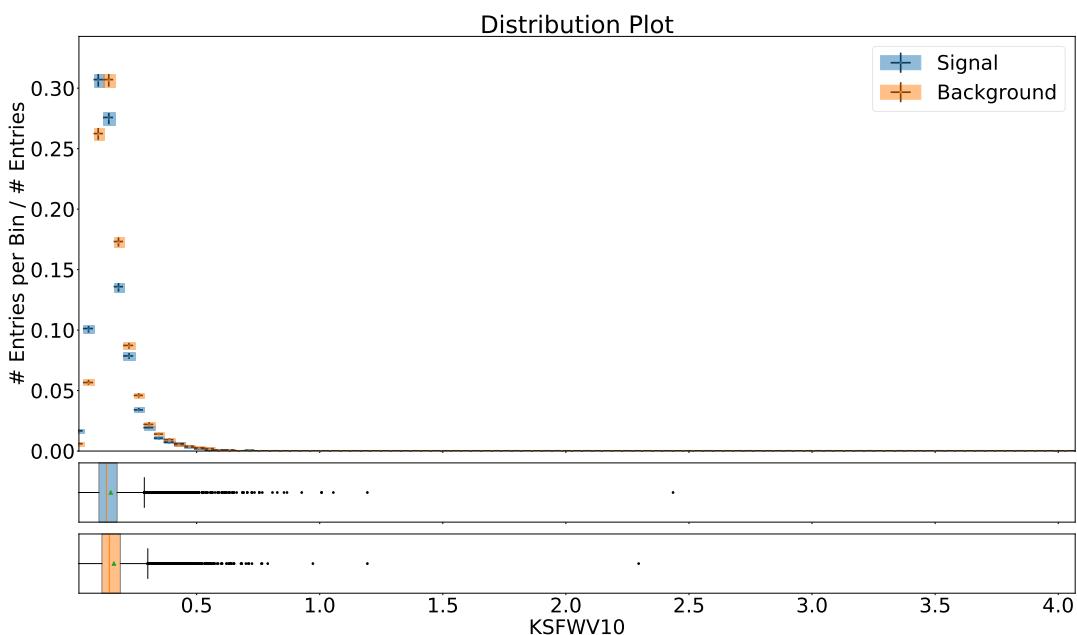
2.12 KSFVVariables(hso14)



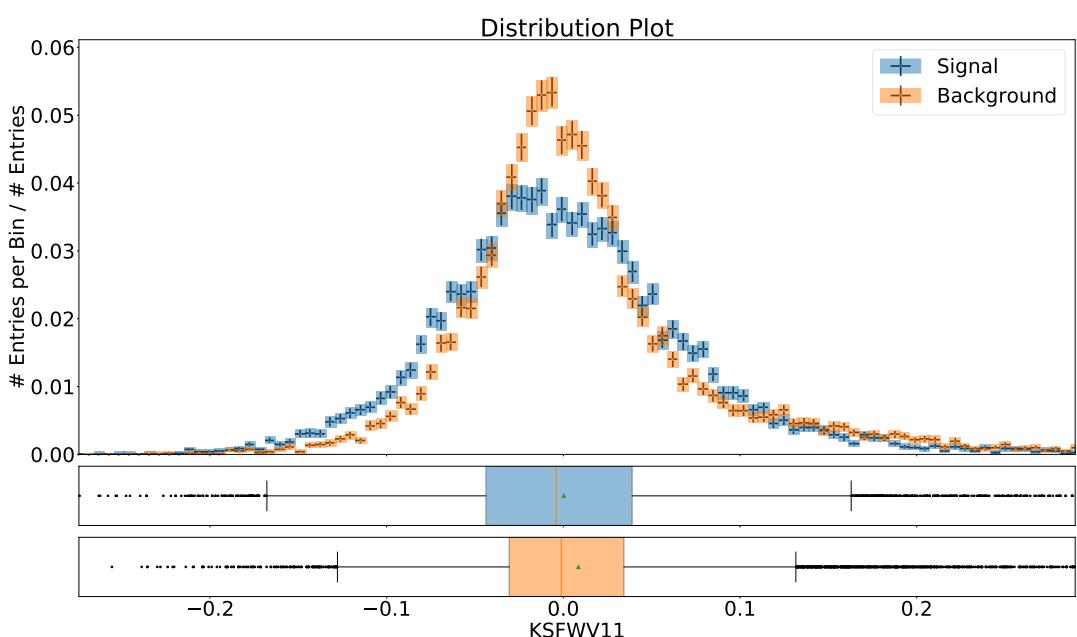
2.13 KSFVVariables(hso10)



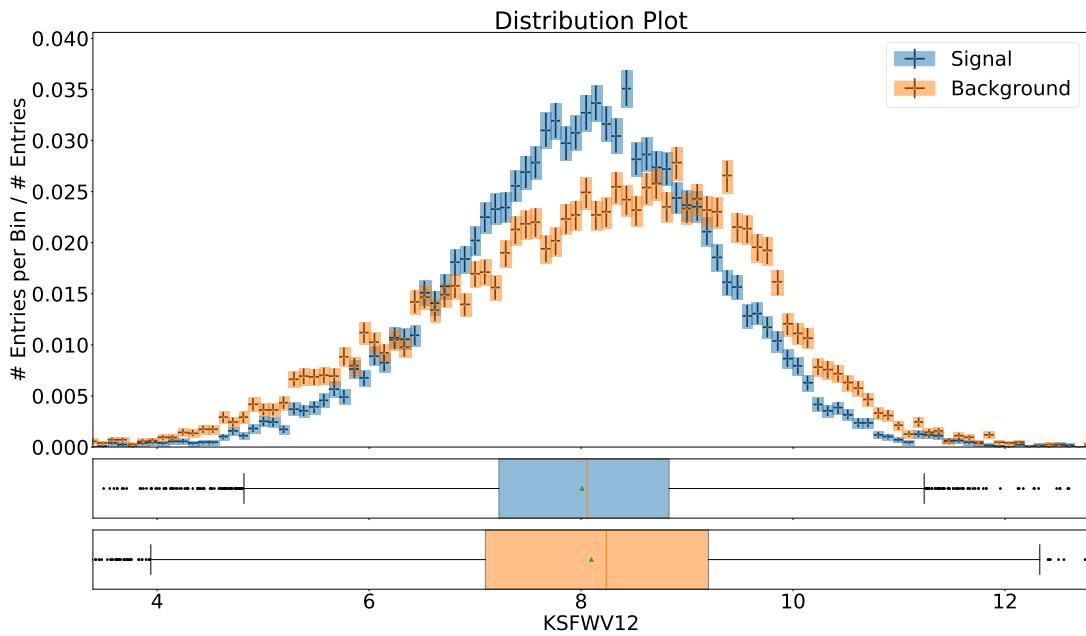
2.14 KSFVVariables(hoo0)



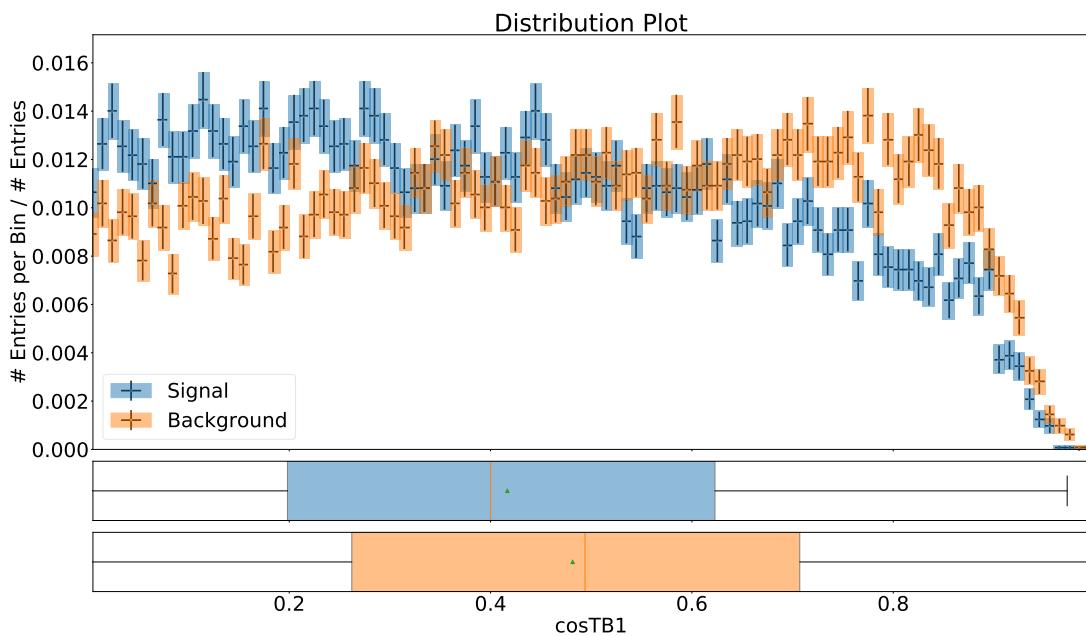
2.15 KSFVVariables(hso24)



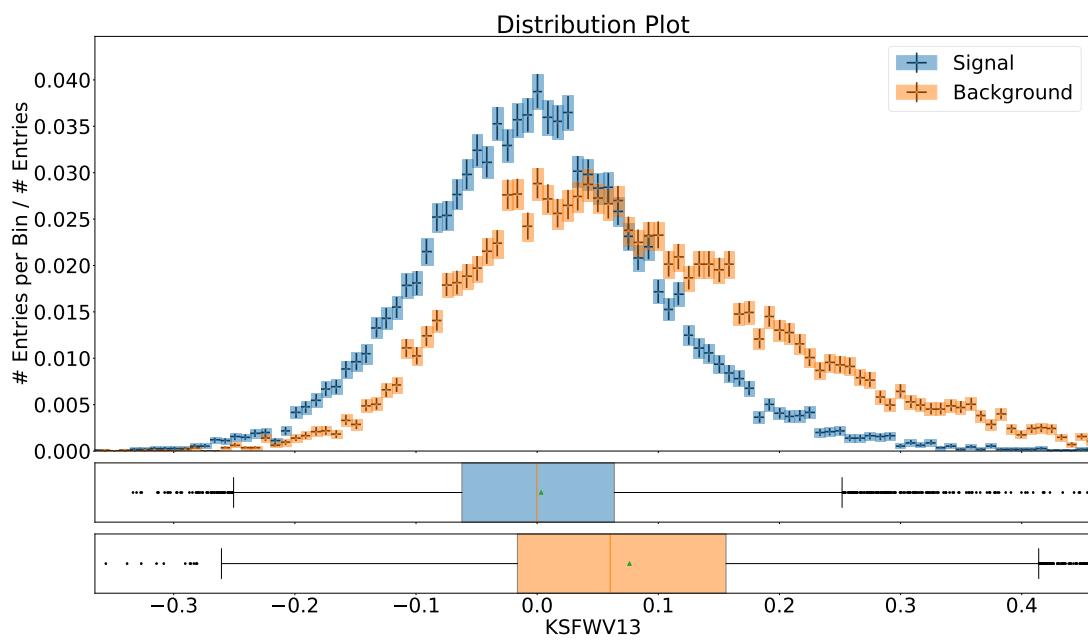
2.16 KSFVWVariables(et)



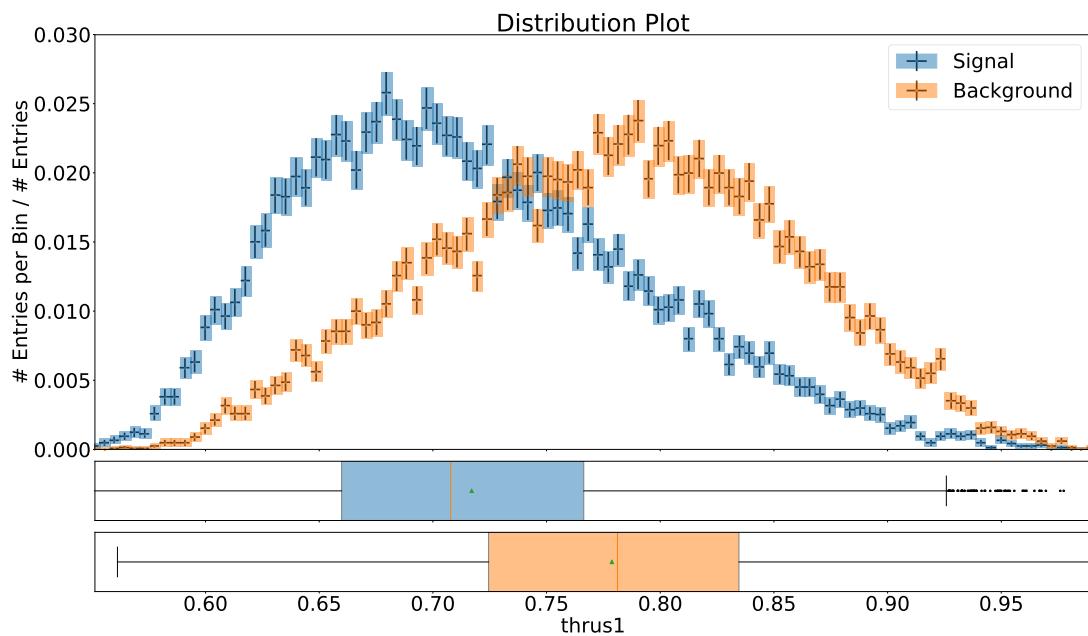
2.17 cosTBz



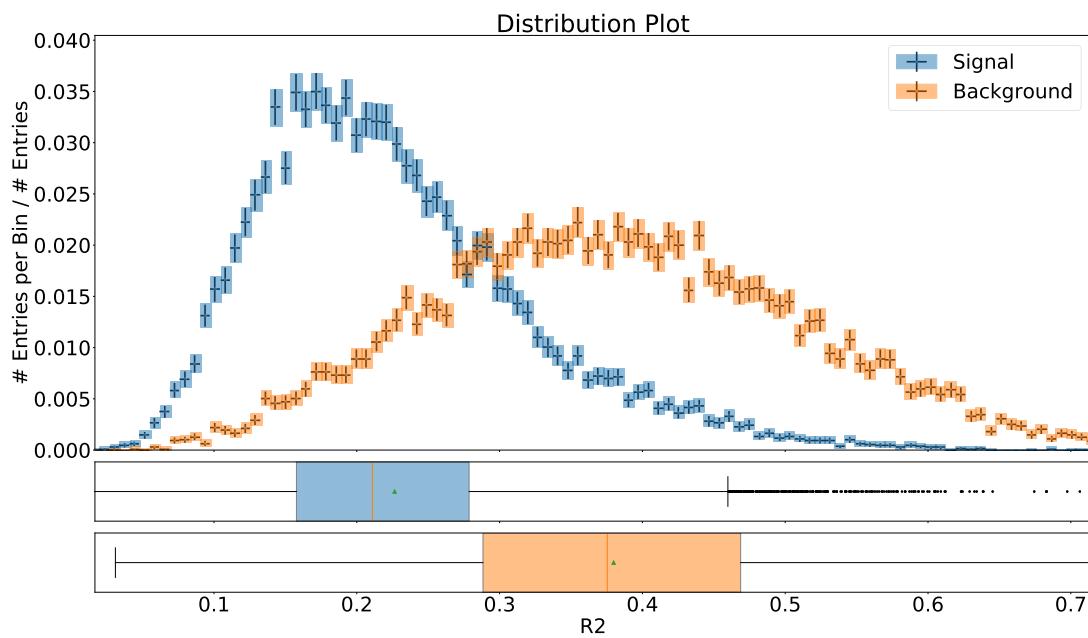
2.18 KSFVWVariables(hso04)



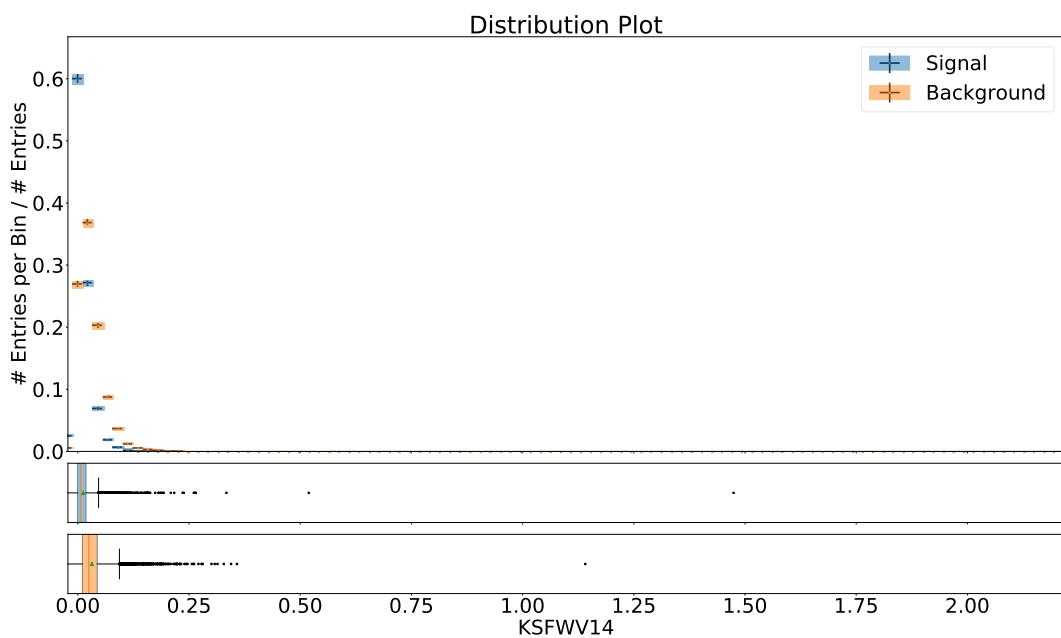
2.19 thrustOm



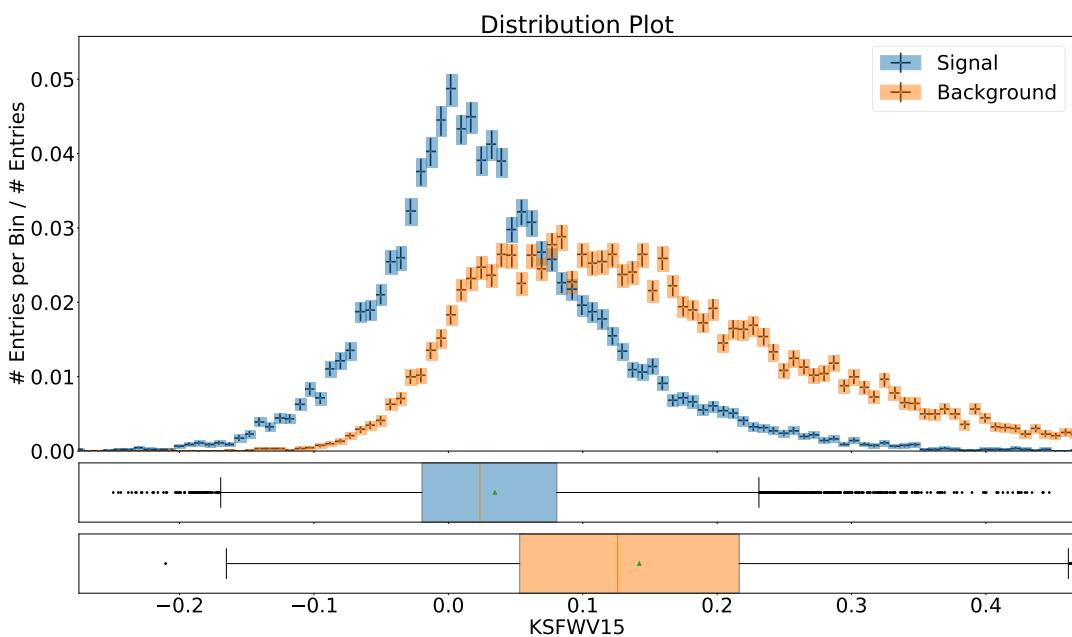
2.20 R2



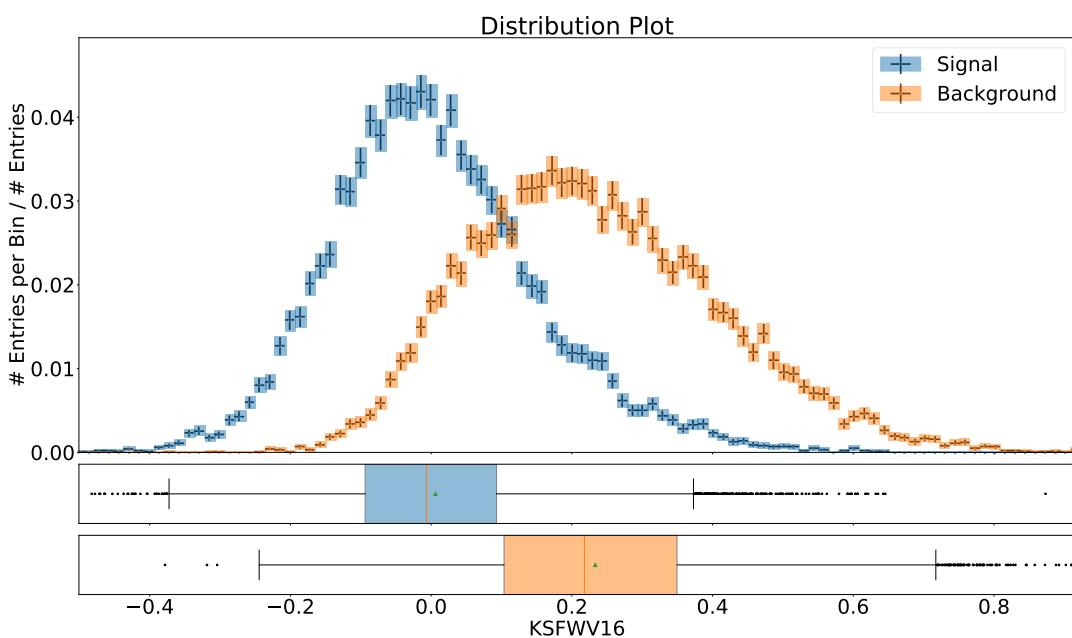
2.21 KSFVWVariables(hoo2)



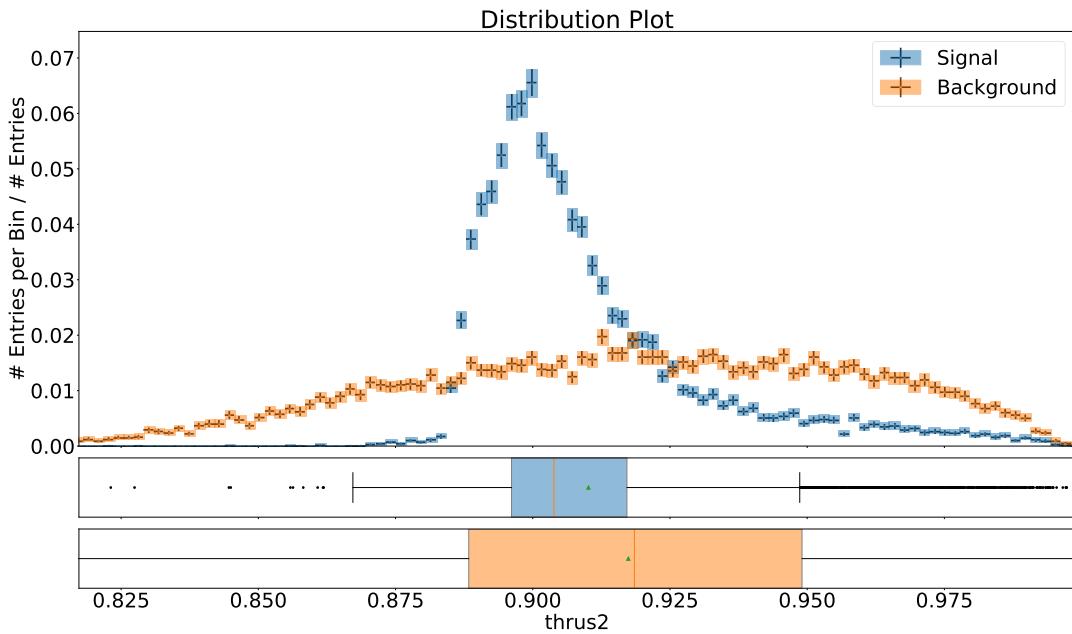
2.22 KSFVVariables(hso12)



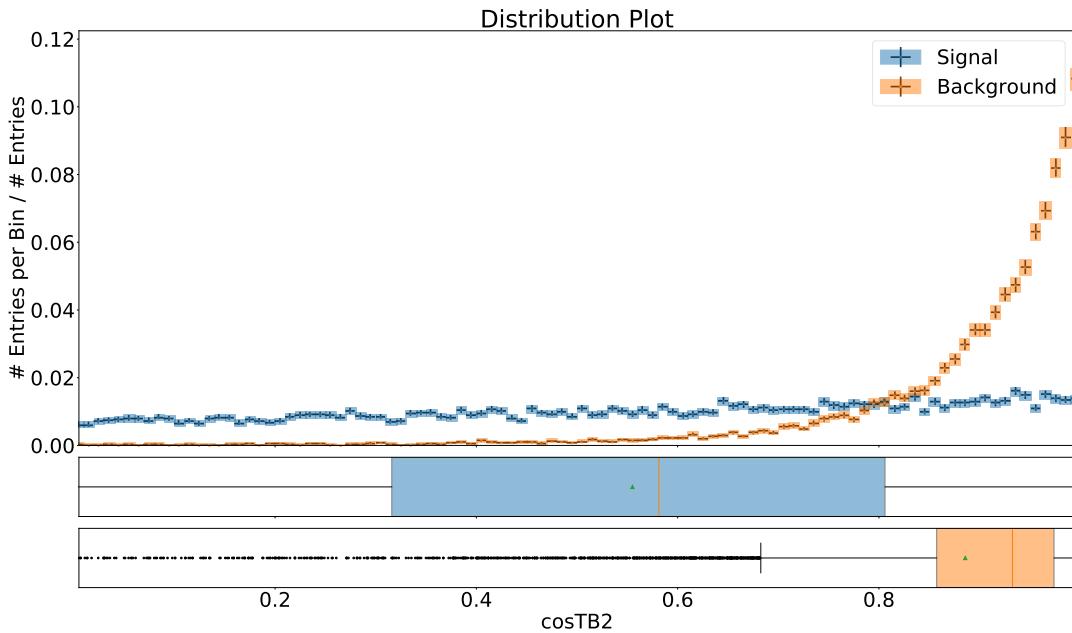
2.23 KSFVVariables(hso02)



2.24 thrustBm



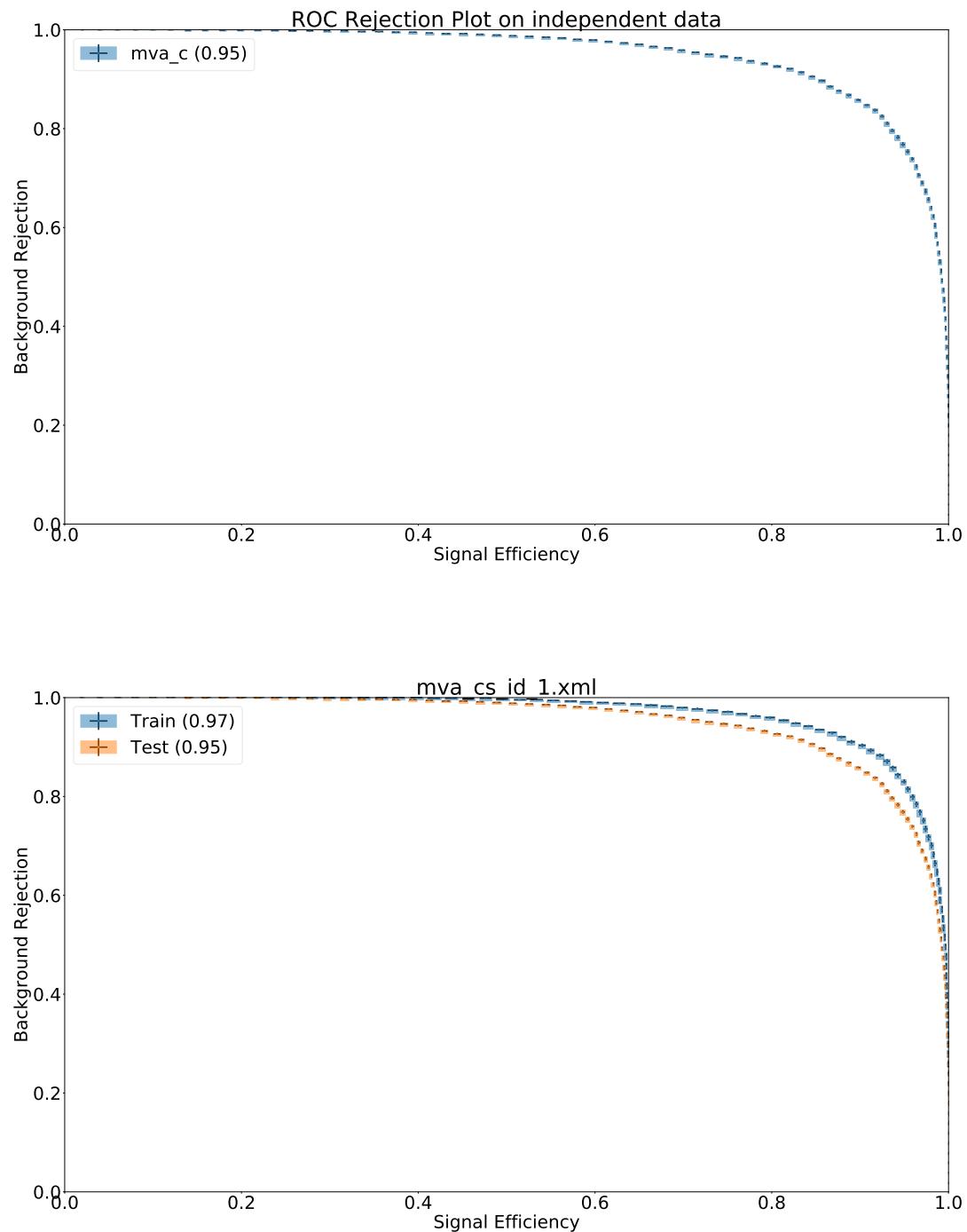
2.25 cosTBTO



3 Classifier Plot

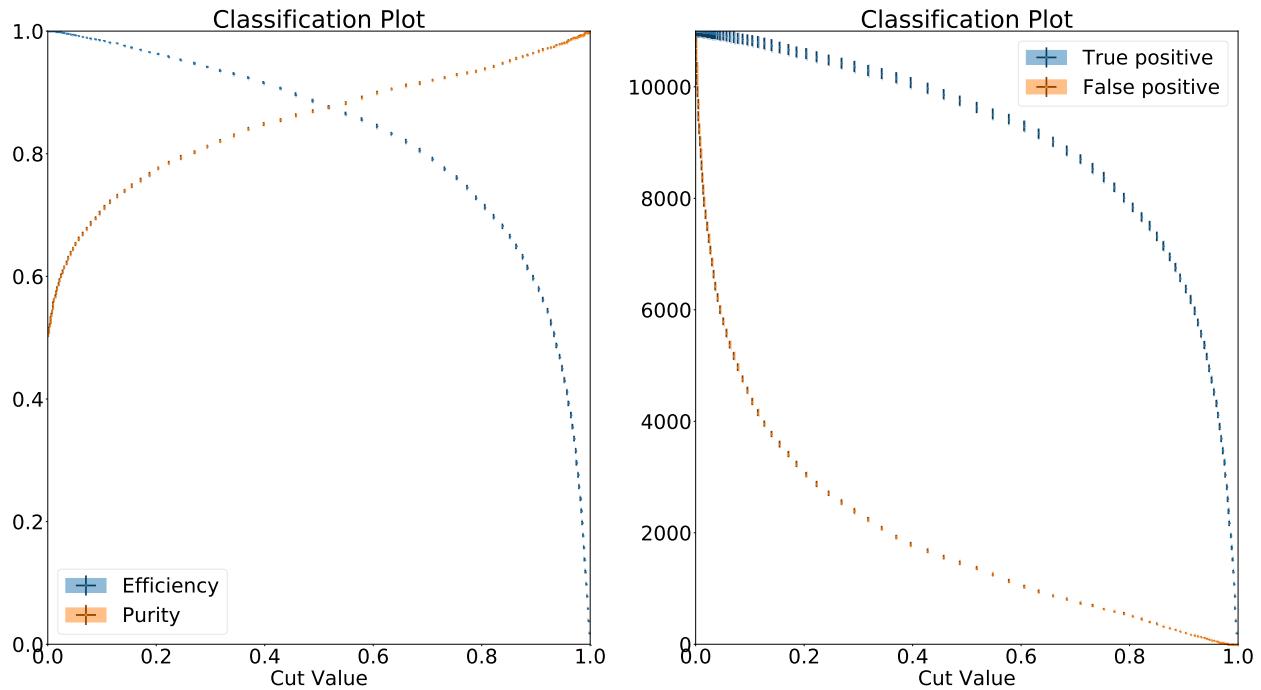
This section contains the receiver operating characteristics (ROC), purity projection, ...of the classifiers on training and independent data. The legend of each plot contains the shortened identifier and the area under the ROC curve in parenthesis.

4 ROC Plot



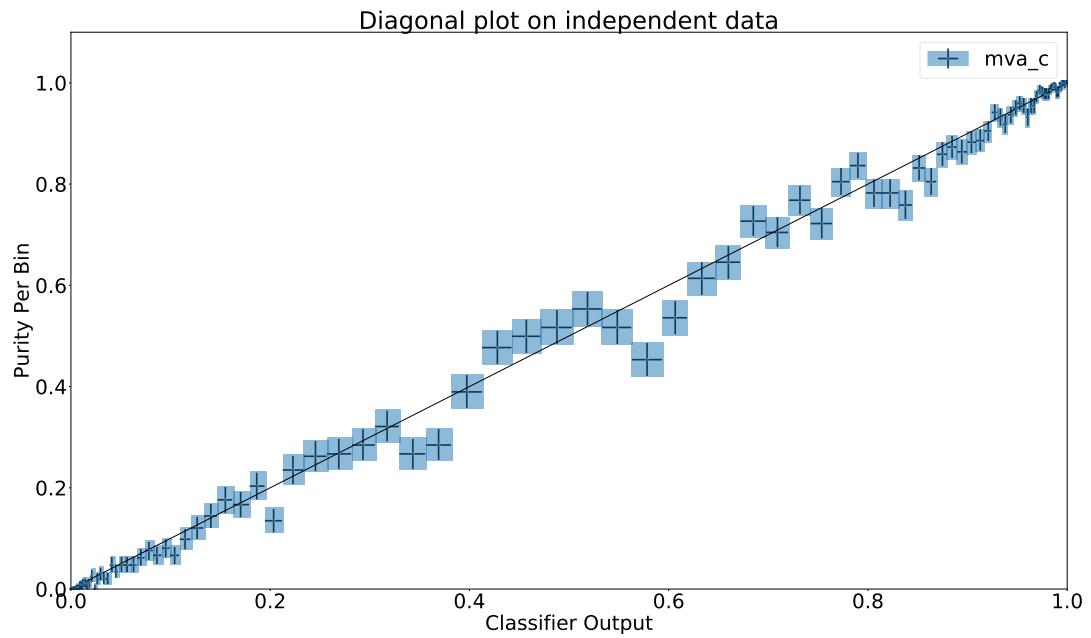
5 Classification Results

5.1 mva_c

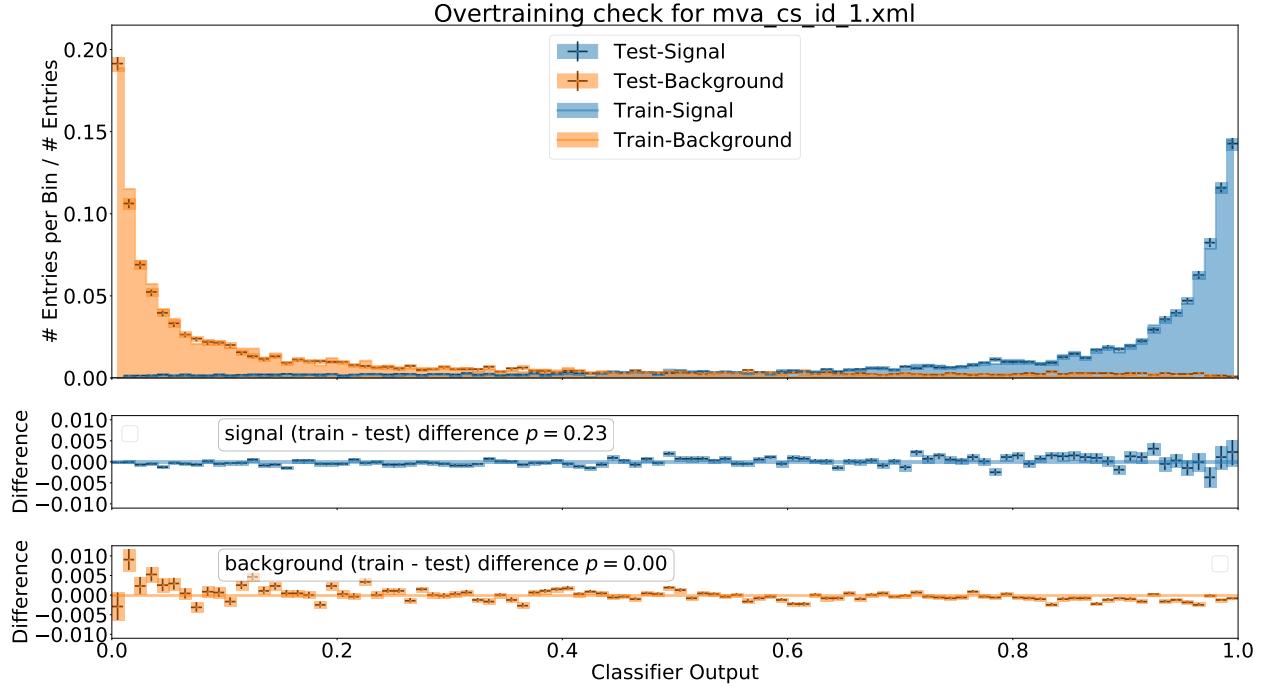


6 Diagonal Plot

6.1 mva_c



6.2 Overtraining Plot



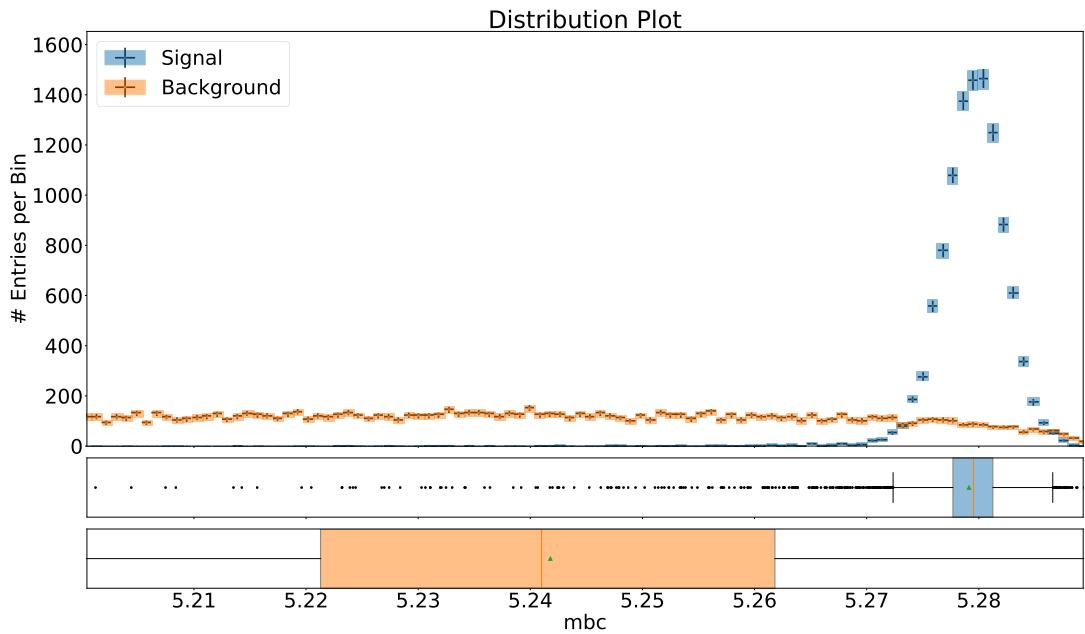
7 Spectators

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_1.xml

