

Automatic MVA Evaluation

Thomas Keck
Moritz Gelb
Nils Braun

April 8, 2023

Abstract

Evaluation plots

Contents

1	Classifiers	3
1.1	/home/belle2/ssana/MC15ri_cs1/cs/test/MVAFastBDT.root	3
2	Variables	3
2.1	Importance	4
2.2	Correlation	5
2.3	KSFVariables(hso01)	5
2.4	KSFVariables(hso20)	6
2.5	CleoConeCS(2)	6
2.6	KSFVariables(hoo0)	7
2.7	thrustOm	7
2.8	KSFVariables(et)	8
2.9	CleoConeCS(1)	8
2.10	cosTBz	9
2.11	KSFVariables(hso10)	9
2.12	KSFVariables(hso02)	10
2.13	KSFVariables(hso12)	10
2.14	CMS_cosTheta	11
2.15	thrustBm	11
2.16	cosTBTO	12
2.17	abs_qr	12
2.18	DeltaZ	13
2.19	R2	13
3	Classifier Plot	13
4	ROC Plot	14
5	Classification Results	15
5.1	/home	15
6	Diagonal Plot	15
6.1	/home	15
6.2	Overtraining Plot	16
7	Spectators	16

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
/home/belle2/ssana/MC15ri_cs1/cs/test/MVAFastBDT.root	/home

1.1 /home/belle2/ssana/MC15ri_cs1/cs/test/MVAFastBDT.root

```
<?xml version="1.0" encoding="utf-8"?>
<method>FastBDT</method>
<weightfile>/home/belle2/ssana/MC15ri_cs1/cs/test/MVAFastBDT.root</weightfile>
<treename>tree</treename>
<target_variable>isSignal</target_variable>
<weight_variable>__weight__</weight_variable>
<signal_class>1</signal_class>
<max_events>0</max_events>
<number_feature_variables>17</number_feature_variables>
<variable0>abs_qr</variable0>
<variable1>DeltaZ</variable1>
<variable2>R2</variable2>
<variable3>thrustBm</variable3>
<variable4>thrustOm</variable4>
<variable5>cosTBT0</variable5>
<variable6>cosTBz</variable6>
<variable7>CMS_cosTheta</variable7>
<variable8>KSFWVariables(et)</variable8>
<variable9>KSFWVariables(hso01)</variable9>
<variable10>KSFWVariables(hso02)</variable10>
<variable11>KSFWVariables(hso10)</variable11>
<variable12>KSFWVariables(hso12)</variable12>
<variable13>KSFWVariables(hso20)</variable13>
<variable14>KSFWVariables(hoo0)</variable14>
<variable15>CleoConeCS(1)</variable15>
<variable16>CleoConeCS(2)</variable16>
<number_spectator_variables>0</number_spectator_variables>
<number_data_files>1</number_data_files>
<datafile0>/home/belle2/ssana/MC15ri_cs1/cs/train/signal_scaled/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

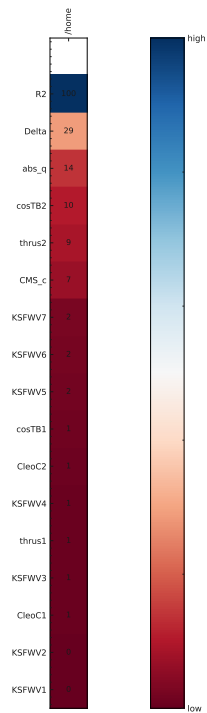
This section contains an overview of the importance and correlation of the variables used by the classifiers. And distribution plots of the variables on the independent dataset. The distributions are normed for signal and background separately, and only the region ± 3 sigma around the mean is shown.

Table 2: Abbreviations of variables

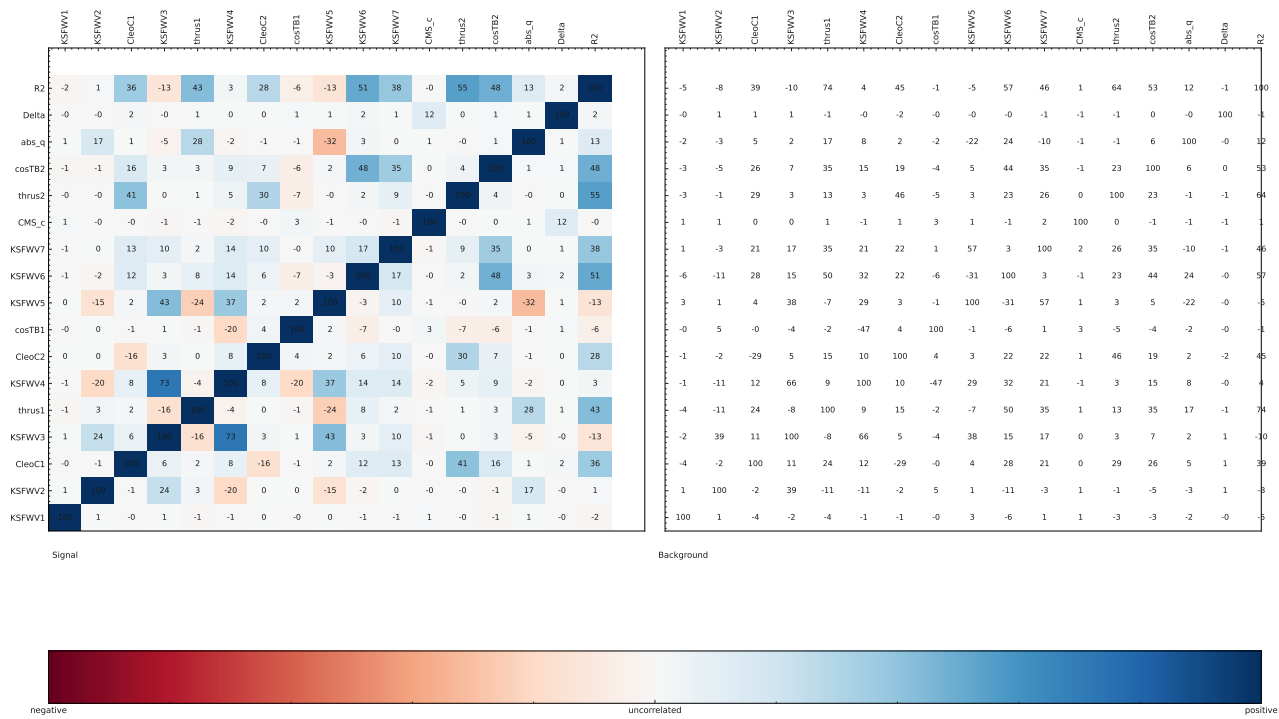
Variable	Abbreviation
KSFWVariables(hso01)	KSFWV1
KSFWVariables(hso20)	KSFWV2
CleoConeCS(2)	CleoC1
KSFWVariables(hoo0)	KSFWV3
thrustOm	thrus1
KSFWVariables(et)	KSFWV4

CleoConeCS(1)	CleoC2
cosTBz	cosTB1
KSFVVariables(hso10)	KSFVW5
KSFVVariables(hso02)	KSFVW6
KSFVVariables(hso12)	KSFVW7
CMS_cosTheta	CMS_c
thrustBm	thrus2
cosTBTO	cosTB2
abs_qr	abs_q
DeltaZ	Delta
R2	R2

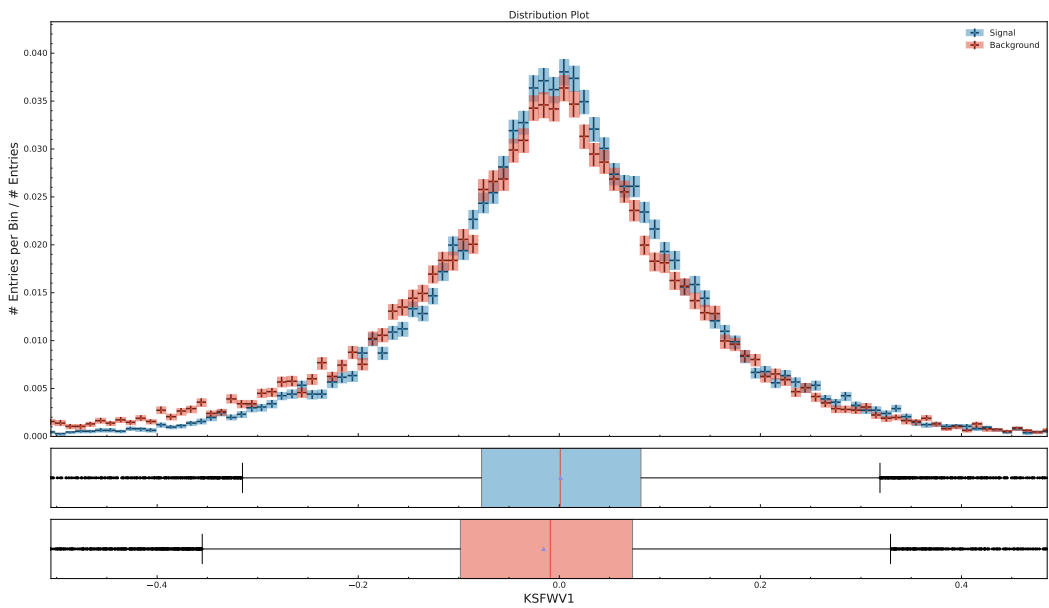
2.1 Importance



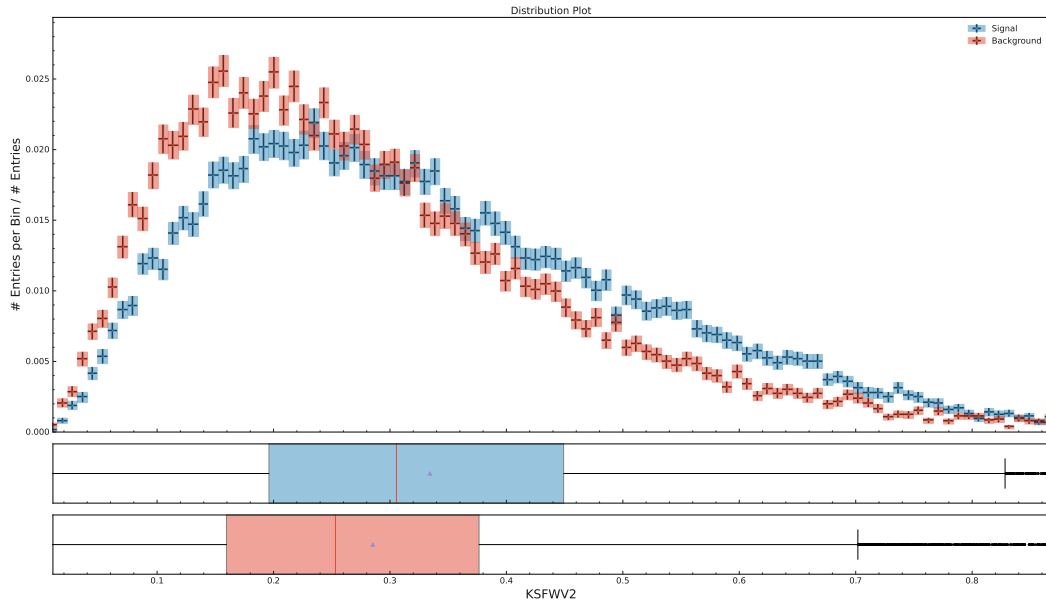
2.2 Correlation



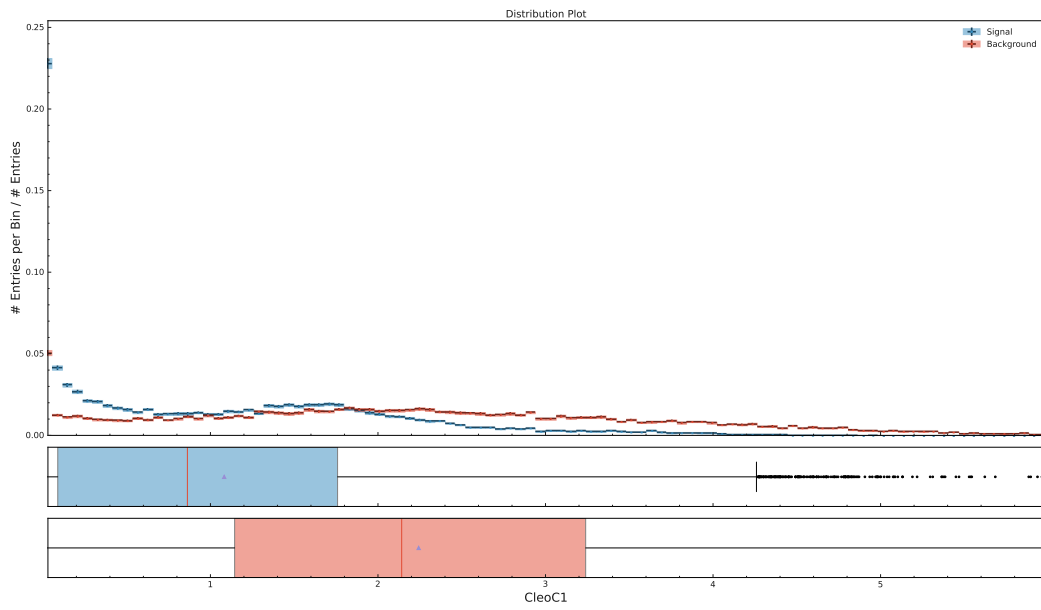
2.3 KSFVVariables(hso01)



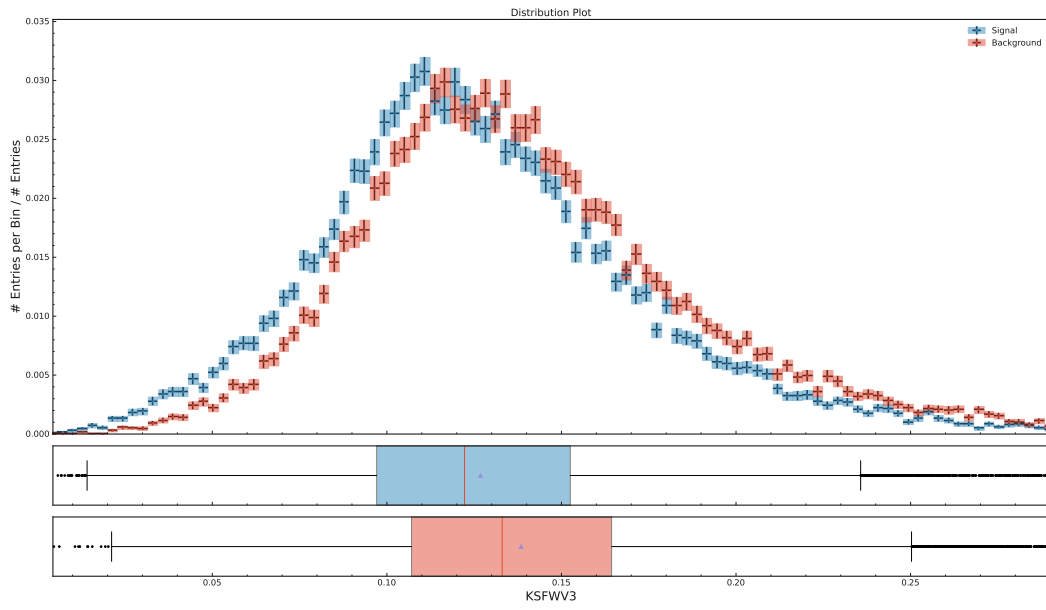
2.4 KSFVVariables(hso20)



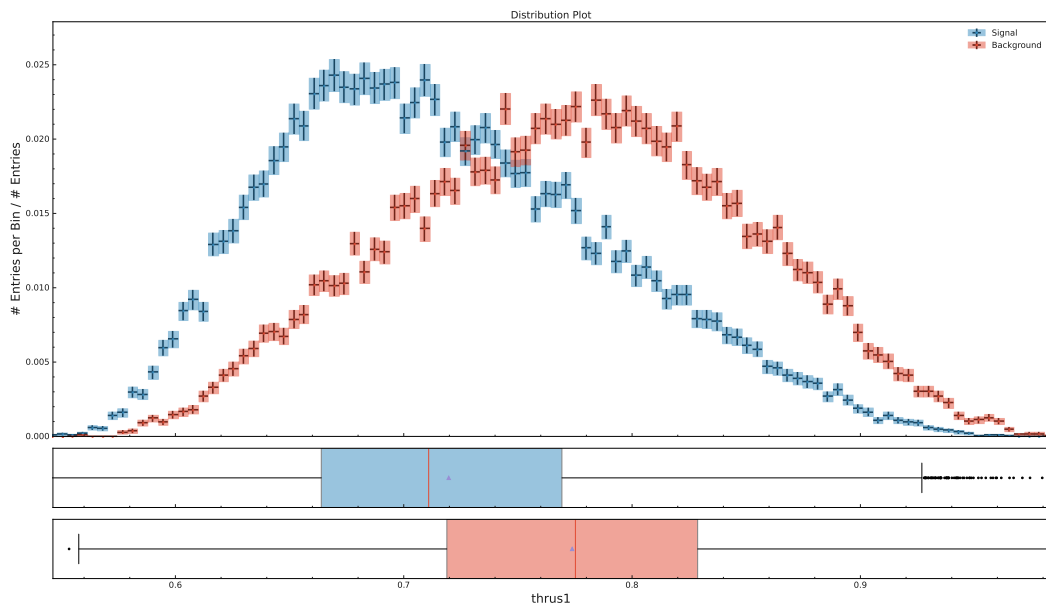
2.5 CleoConeCS(2)



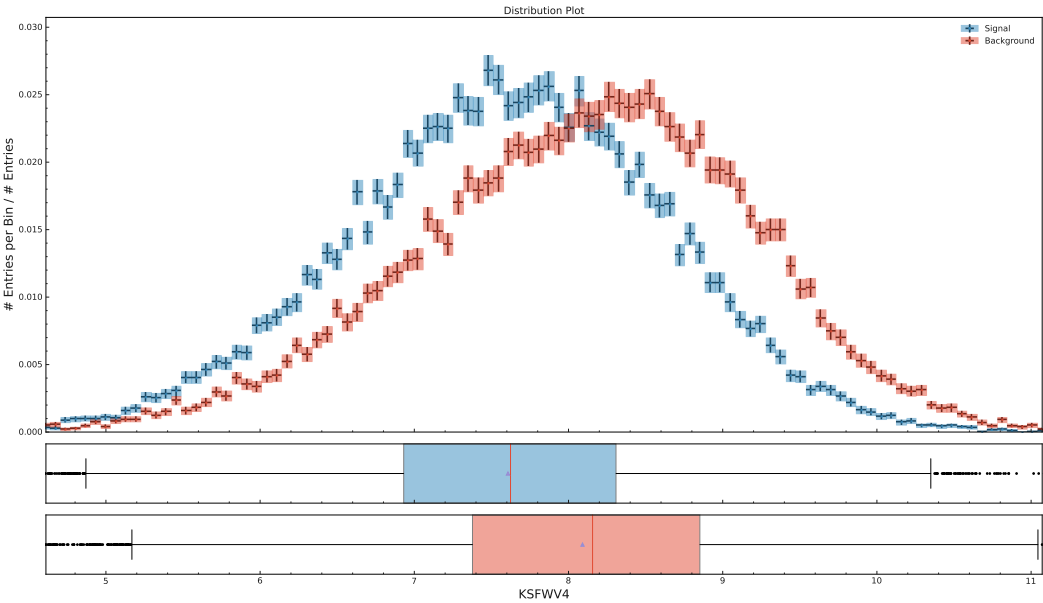
2.6 KSFVVariables(hoo0)



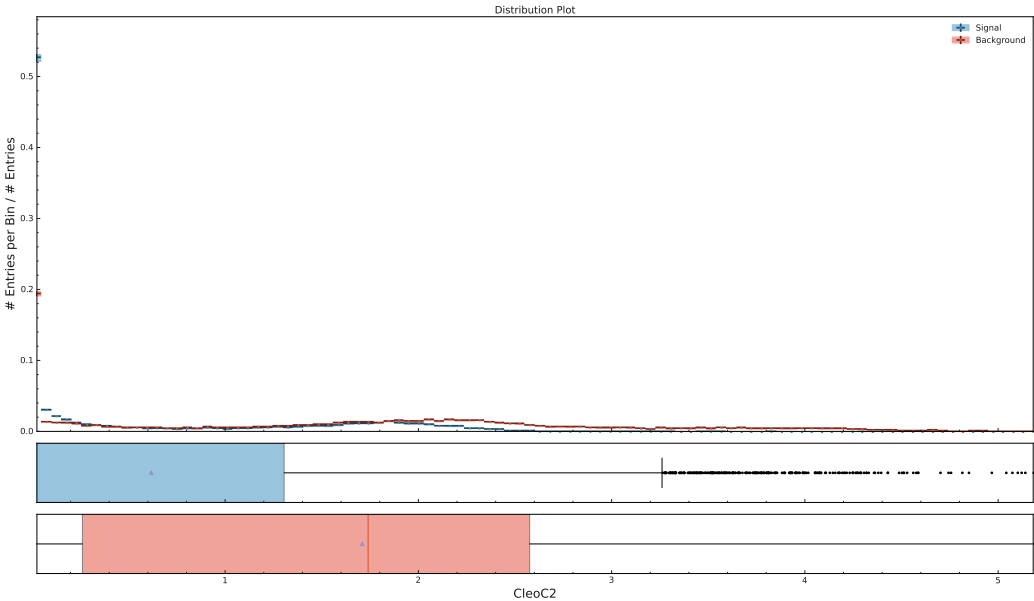
2.7 thrustOm



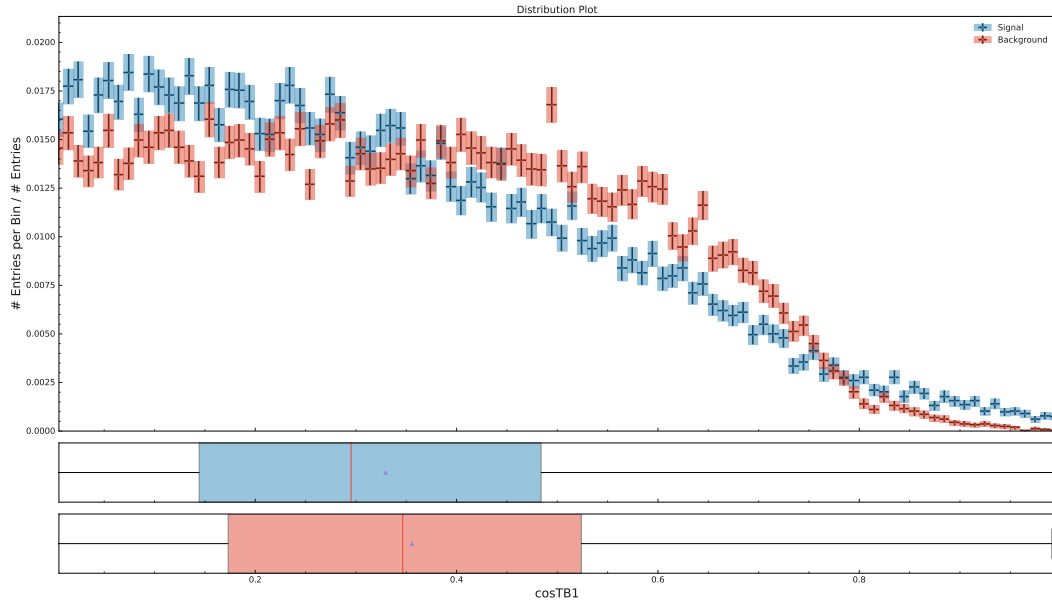
2.8 KSFVVariables(et)



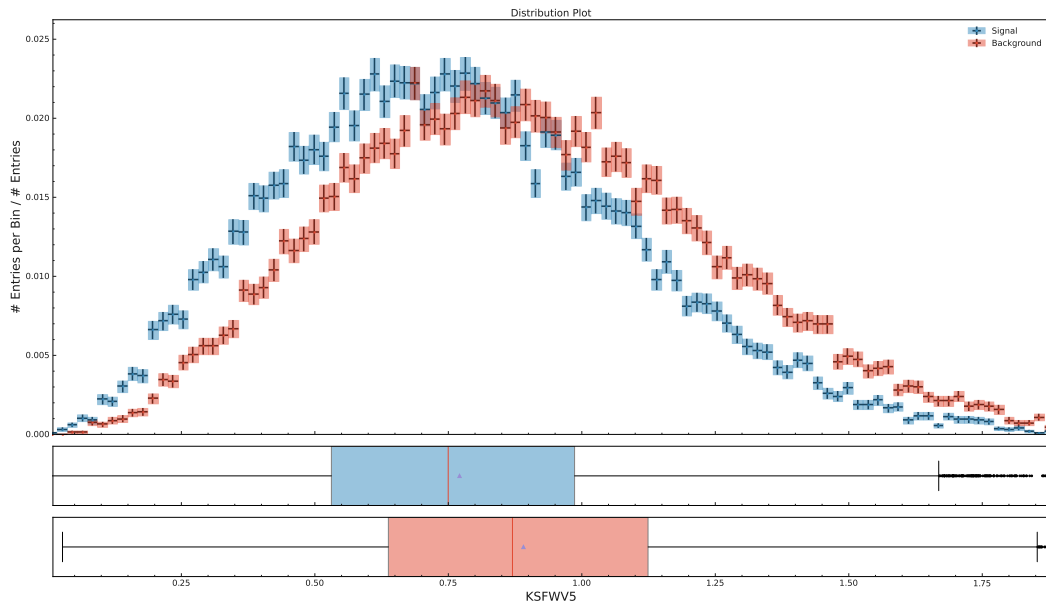
2.9 CleoConeCS(1)



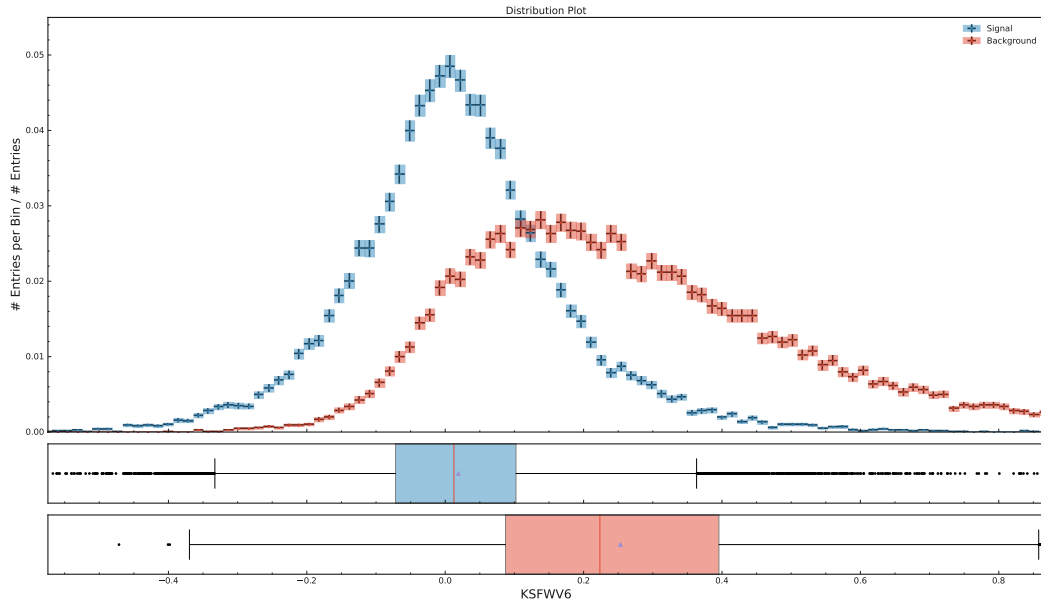
2.10 cosTBz



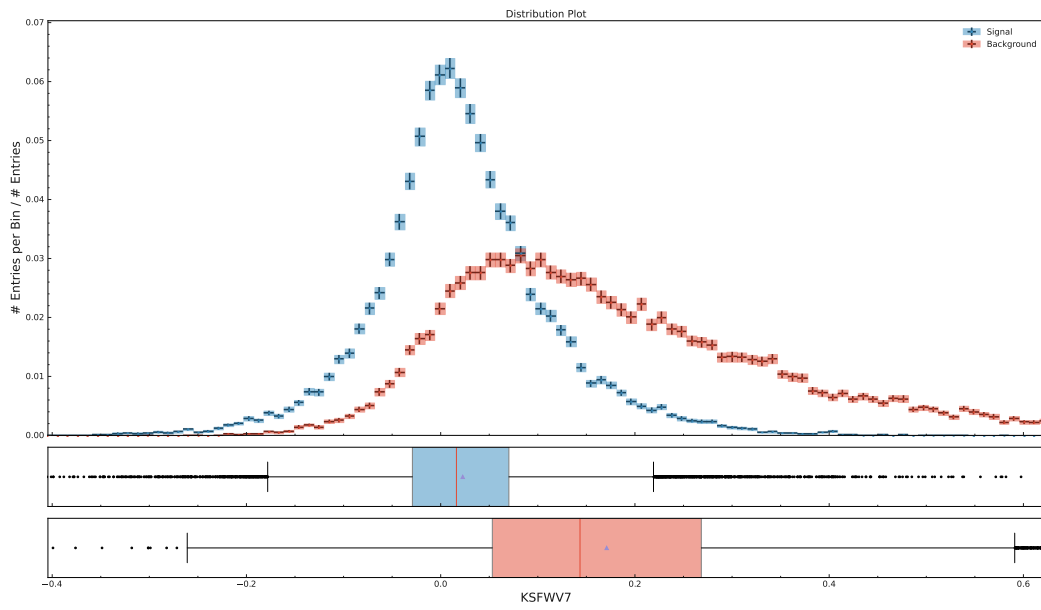
2.11 KSFWVariables(hso10)



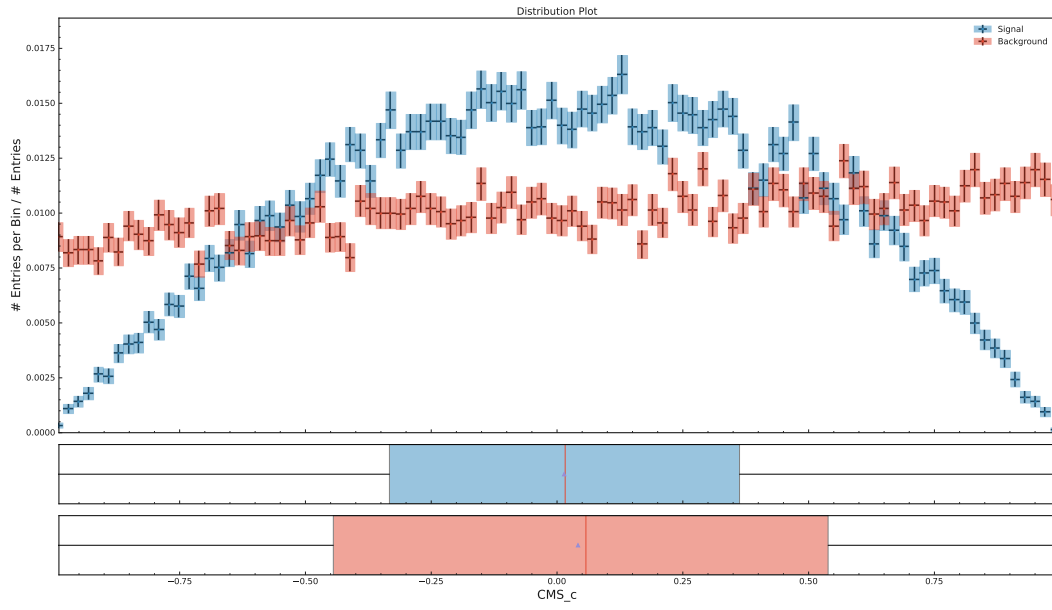
2.12 KSFVVariables(hso02)



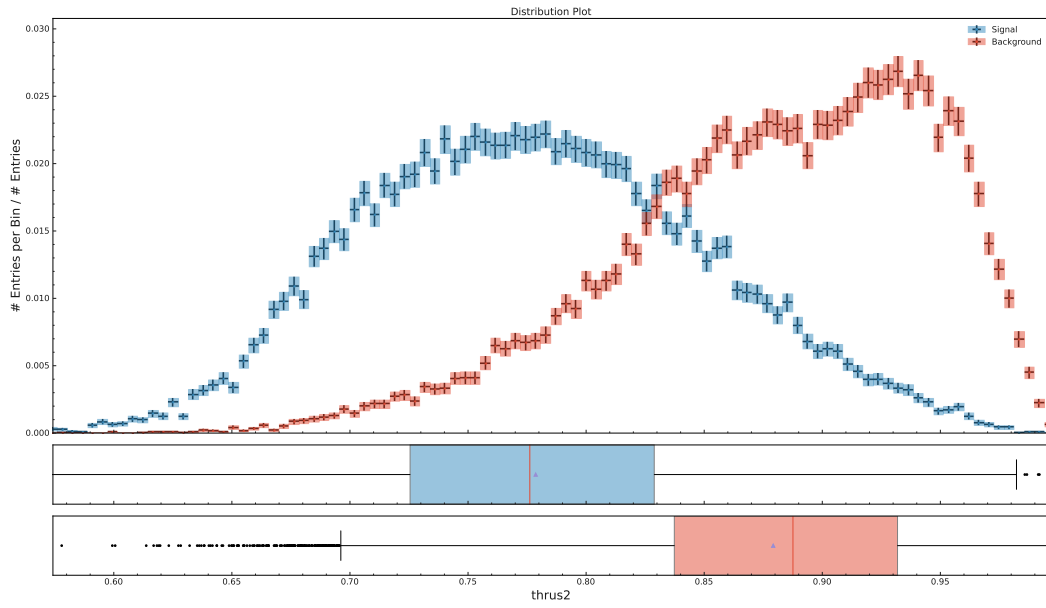
2.13 KSFVVariables(hso12)



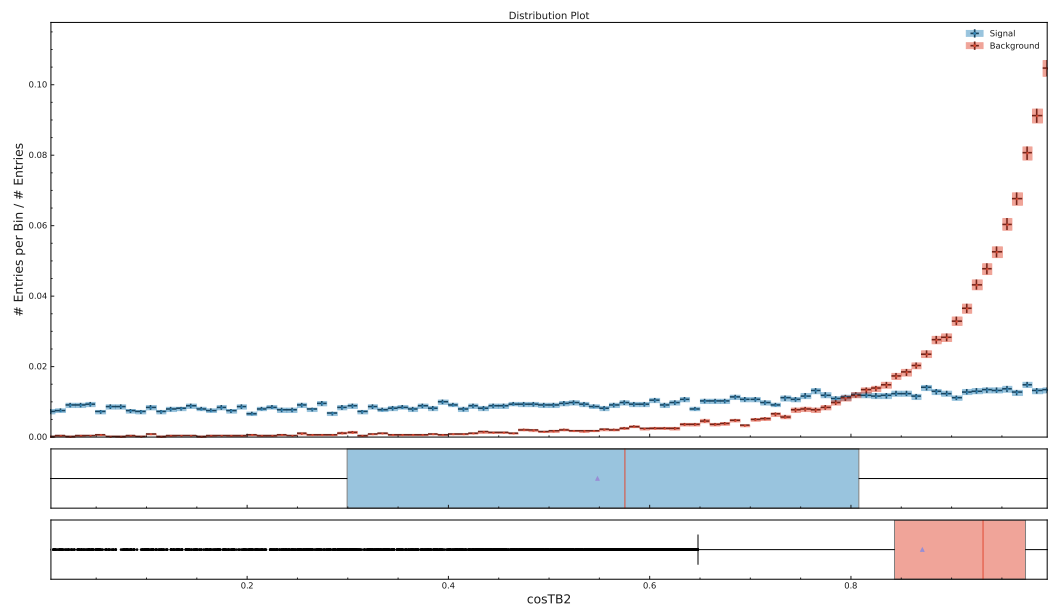
2.14 CMS_cosTheta



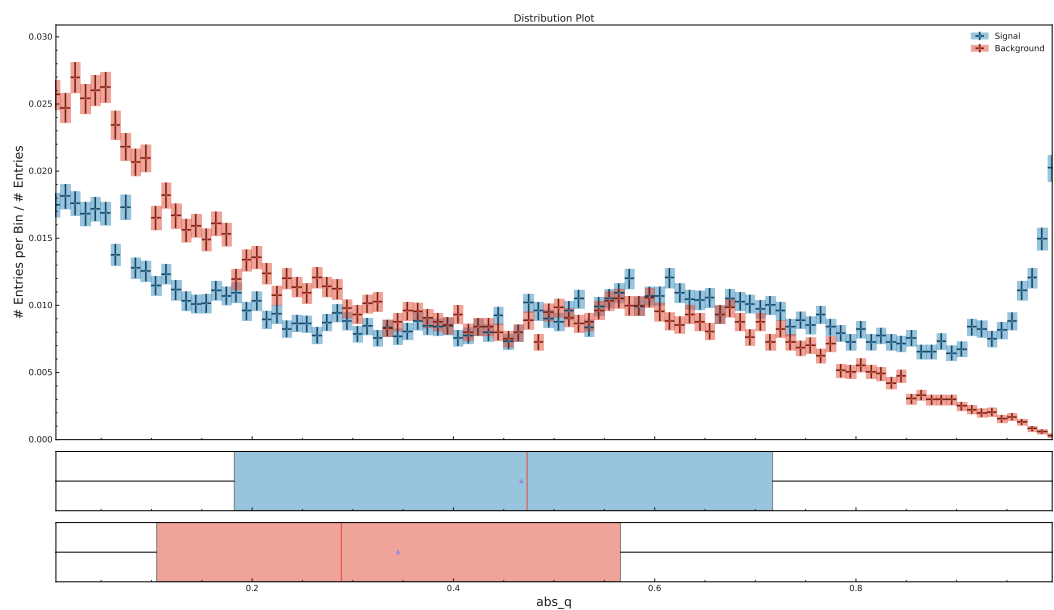
2.15 thrustBm



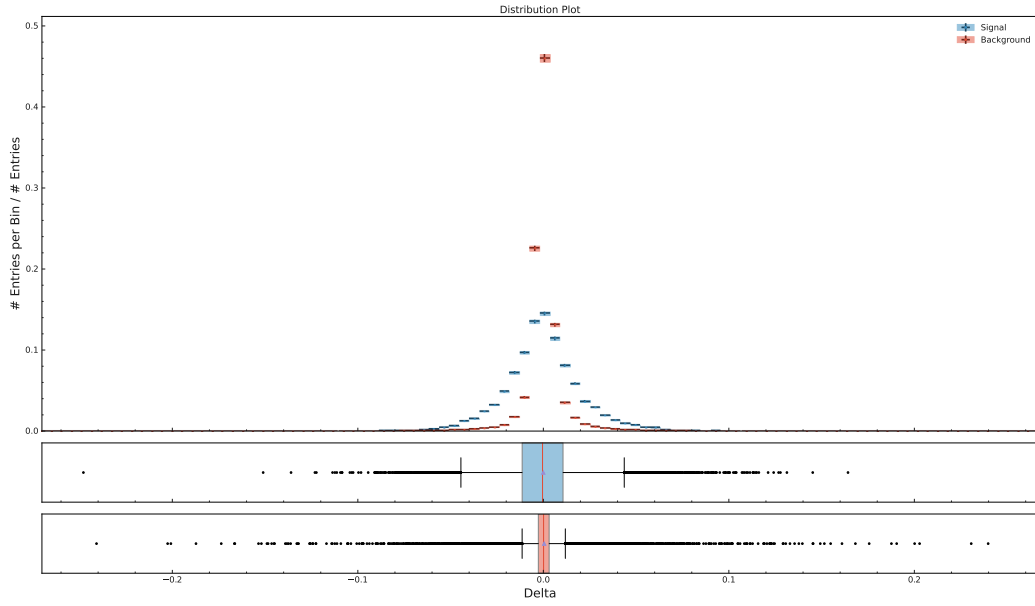
2.16 cosTBTO



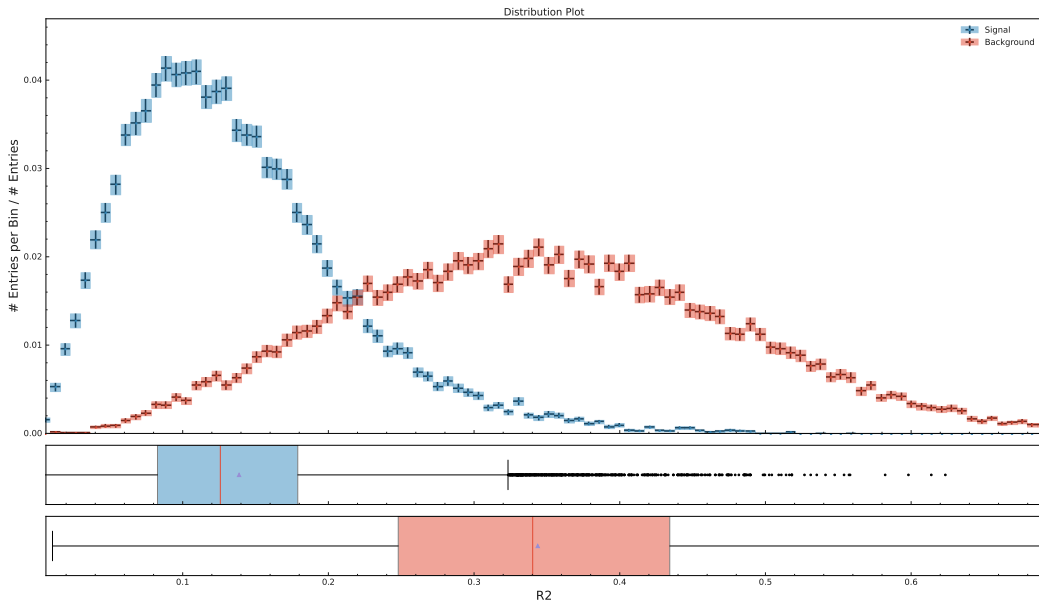
2.17 abs_qr



2.18 DeltaZ



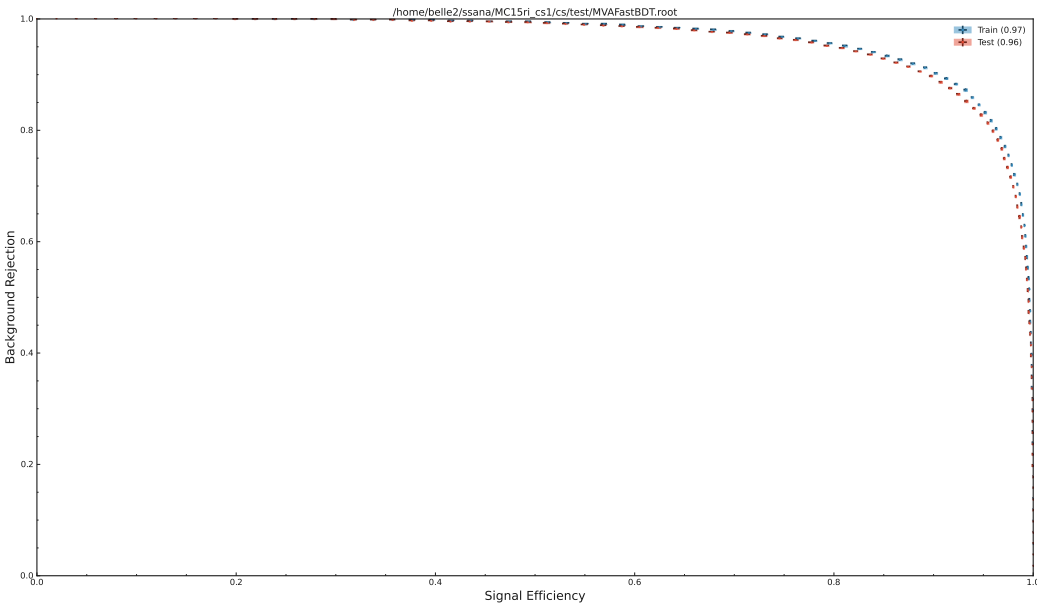
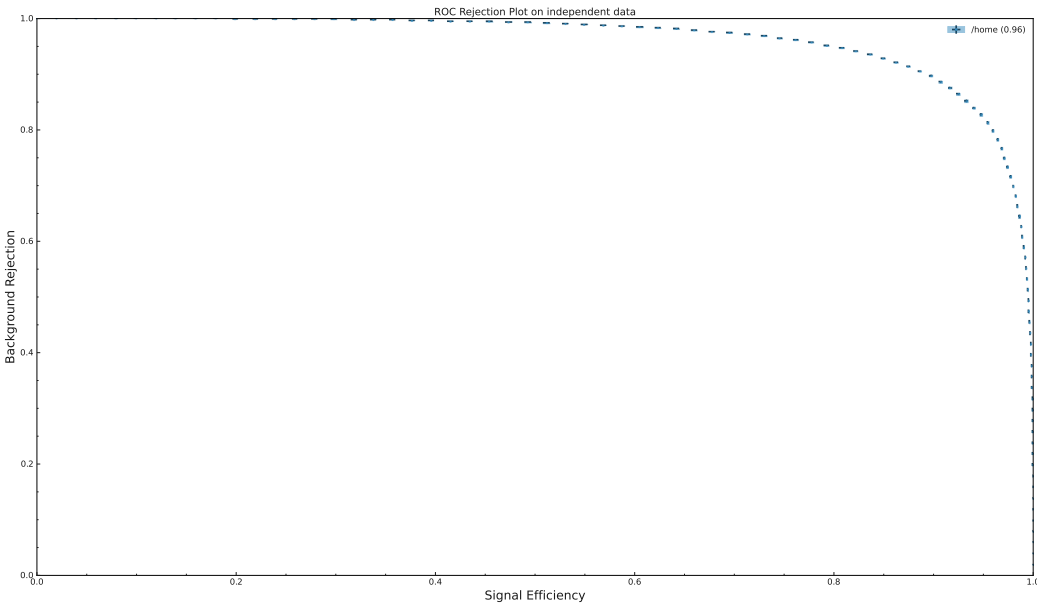
2.19 R2



3 Classifier Plot

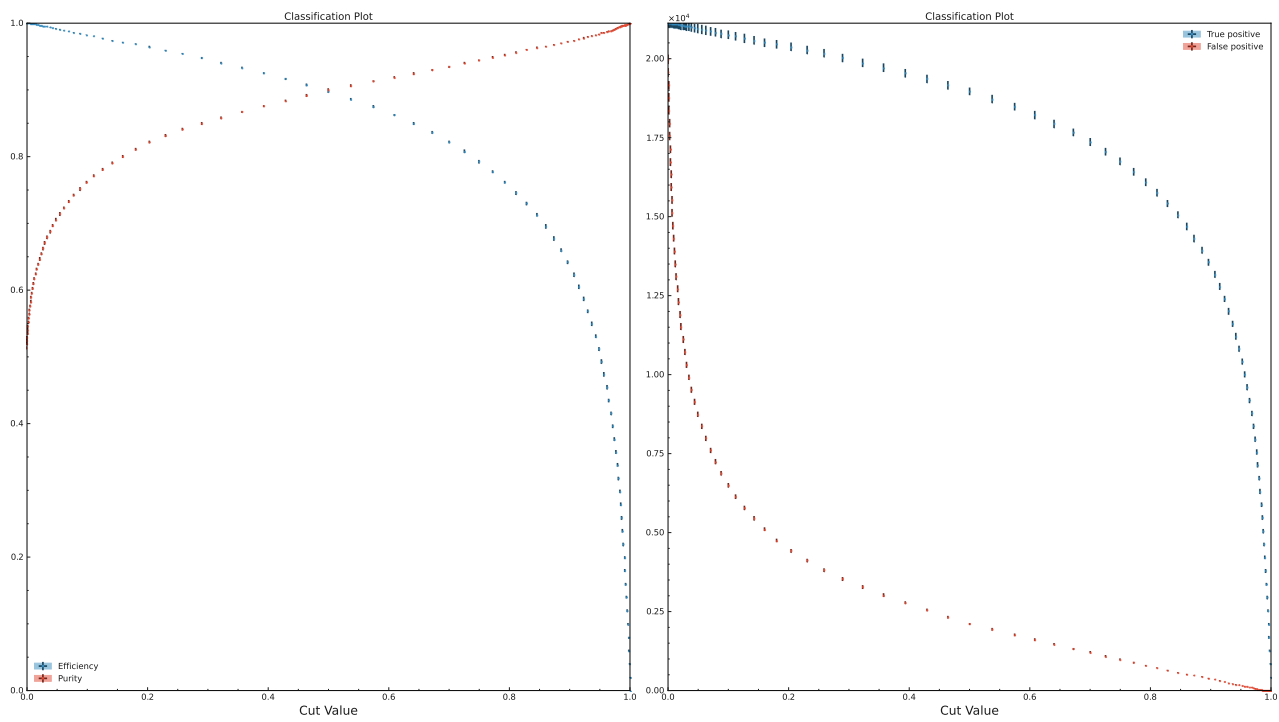
This section contains the receiver operating characteristics (ROC), purity projection, ...of the classifiers on training and independent data. The legend of each plot contains the shortened identifier and the area under the ROC curve in parenthesis.

4 ROC Plot



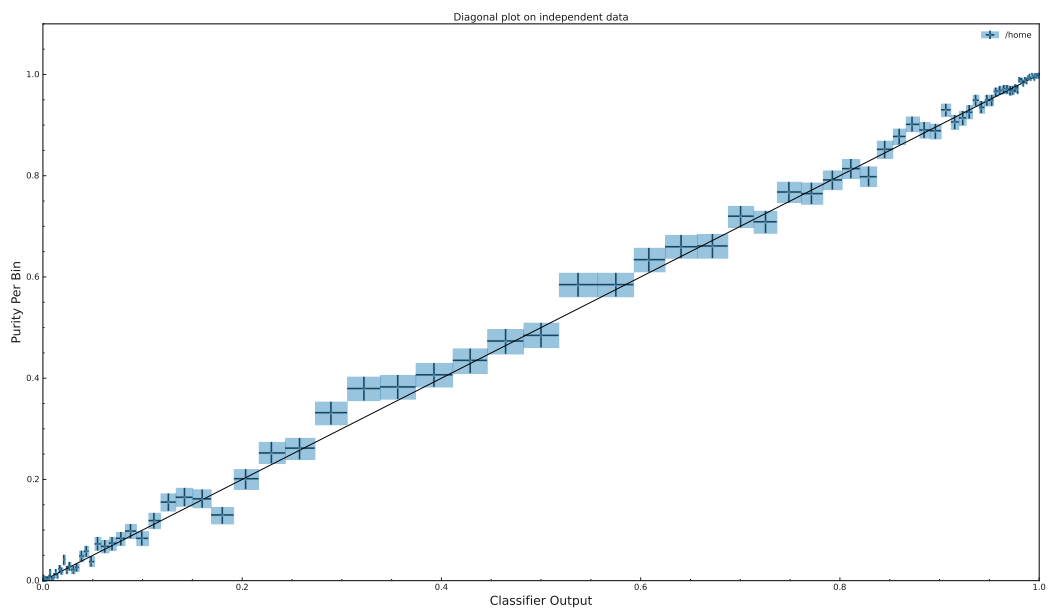
5 Classification Results

5.1 /home

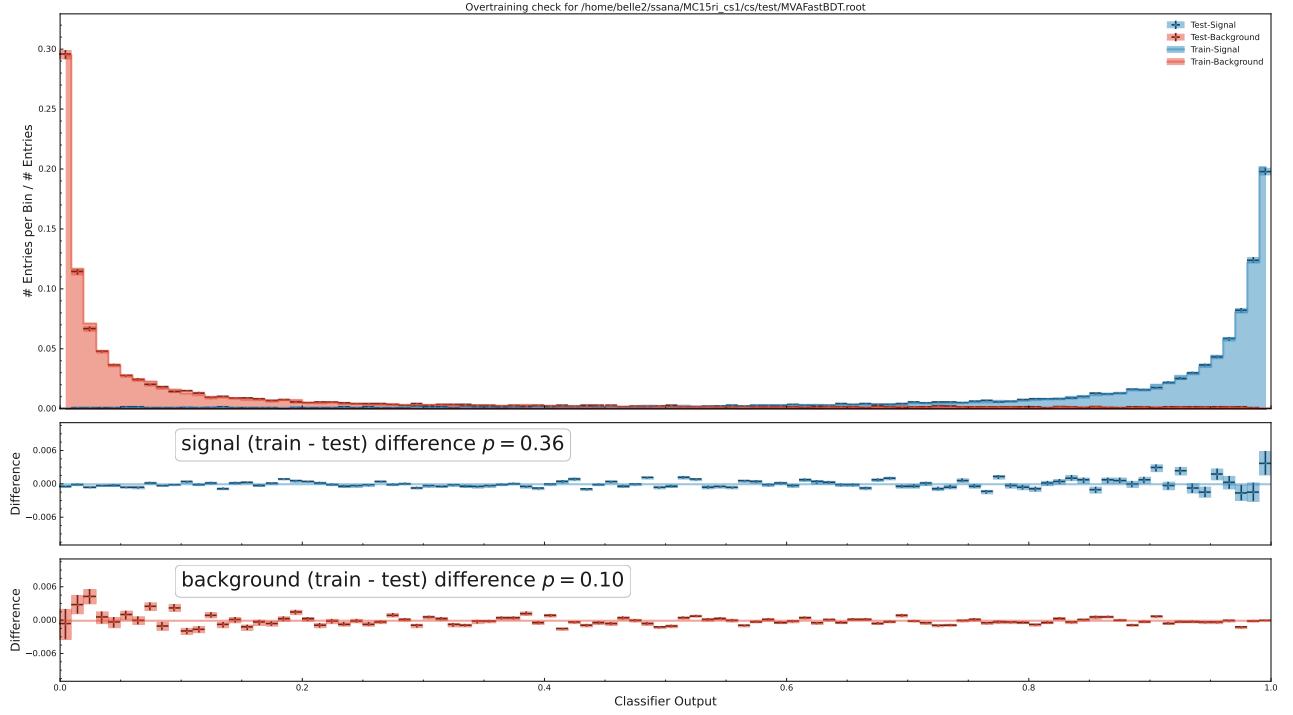


6 Diagonal Plot

6.1 /home



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
-----------	--------------