Automatic MVA Evaluation

Thomas Keck
Moritz Gelb
Nils Braun

April 8, 2023

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 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>35</number_feature_variables>
<variable0>abs_qr</variable0>
<variable1>DeltaZ</variable1>
<variable2>R2</variable2>
<variable3>thrustBm</variable3>
<variable4>thrustOm</variable4>
<variable5>cosTBTO</variable5>
<variable6>cosTBz</variable6>
<variable7>CMS_cosTheta</variable7>
<variable8>KSFWVariables(et)</variable8>
<variable9>KSFWVariables(mm2)</variable9>
<variable10>KSFWVariables(hso00)</variable10>
<variable11>KSFWVariables(hso01)</variable11>
<variable12>KSFWVariables(hso02)</variable12>
<variable13>KSFWVariables(hso03)</variable13>
<variable14>KSFWVariables(hso04)</variable14>
<variable15>KSFWVariables(hso10)</variable15>
<variable16>KSFWVariables(hso12)</variable16>
<variable17>KSFWVariables(hso14)</variable17>
<variable18>KSFWVariables(hso20)</variable18>
<variable19>KSFWVariables(hso22)</variable19>
<variable20>KSFWVariables(hso24)</variable20>
<variable21>KSFWVariables(hoo0)</variable21>
<variable22>KSFWVariables(hoo1)</variable22>
<variable23>KSFWVariables(hoo2)</variable23>
<variable24>KSFWVariables(hoo3)</variable24>
<variable25>KSFWVariables(hoo4)</variable25>
<variable26>CleoConeCS(1)</variable26>
<variable27>CleoConeCS(2)</variable27>
<variable28>CleoConeCS(3)</variable28>
<variable29>CleoConeCS(4)</variable29>
<variable30>CleoConeCS(5)</variable30>
<variable31>CleoConeCS(6)</variable31>
<variable32>CleoConeCS(7)</variable32>
<variable33>CleoConeCS(8)</variable33>
<variable34>CleoConeCS(9)</variable34>
<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>
```

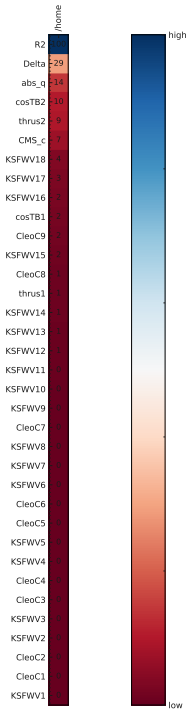

2 Variables

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.

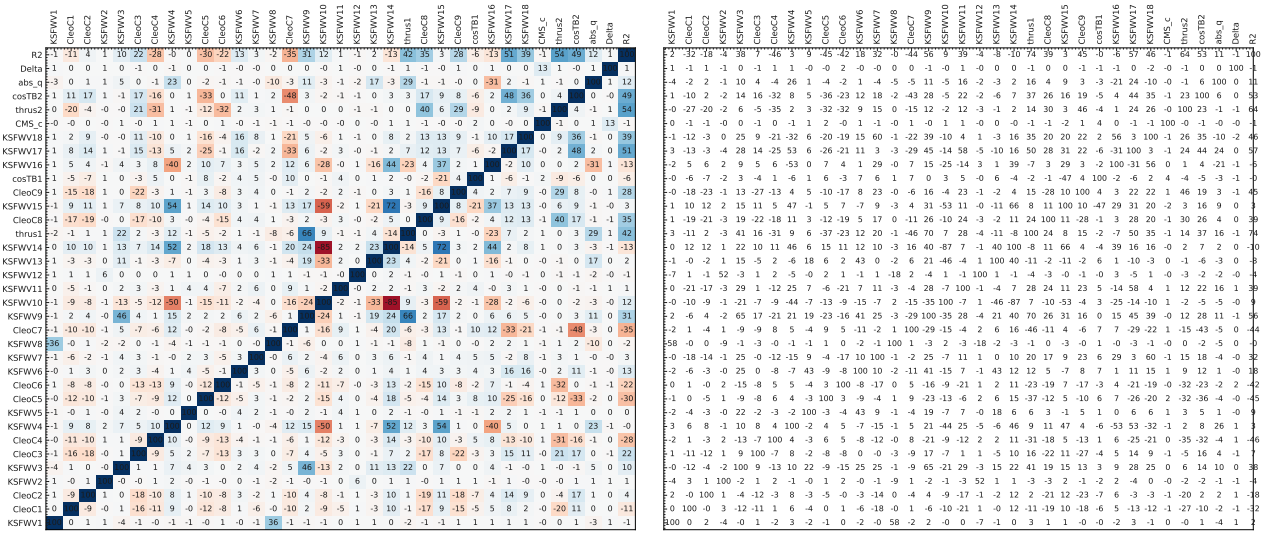
Table 2: Abbreviations of variables

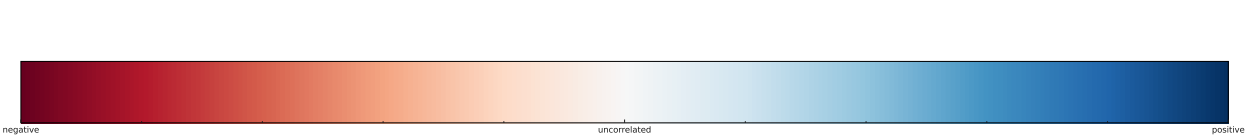
Variable	Abbreviation
KSFWVariables(hoo3)	KSFWV1
CleoConeCS(5)	CleoC1
CleoConeCS(4)	CleoC2
KSFWVariables(hso03)	KSFWV2
KSFWVariables(hoo4)	KSFWV3
CleoConeCS(3)	CleoC3
CleoConeCS(7)	CleoC4
KSFWVariables(hso00)	KSFWV4
KSFWVariables(hso24)	KSFWV5
CleoConeCS(8)	CleoC5
CleoConeCS(6)	CleoC6
KSFWVariables(hso22)	KSFWV6
KSFWVariables(hso14)	KSFWV7
KSFWVariables(hoo1)	KSFWV8
CleoConeCS(9)	CleoC7
KSFWVariables(hoo2)	KSFWV9
KSFWVariables(mm2)	KSFWV10
KSFWVariables(hso04)	KSFWV11
KSFWVariables(hso01)	KSFWV12
KSFWVariables(hso20)	KSFWV13
KSFWVariables(hoo0)	KSFWV14
thrustOm	thrus1
CleoConeCS(2)	CleoC8
KSFWVariables(et)	KSFWV15
CleoConeCS(1)	CleoC9
cosTBz	cosTB1
KSFWVariables(hso10)	KSFWV16
KSFWVariables(hso02)	KSFWV17
KSFWVariables(hso12)	KSFWV18
CMS_cosTheta	CMS_c
thrustBm	thrus2
cosTBTO	cosTB2
abs_qr	abs_q
DeltaZ	Delta
R2	R2

2.1 Importance

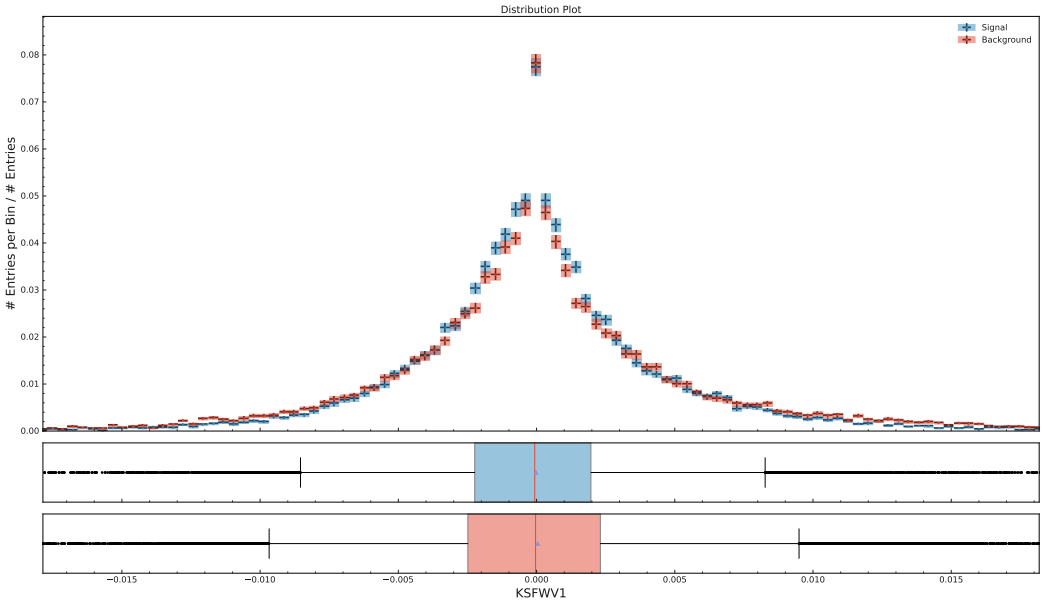


2.2 Correlation

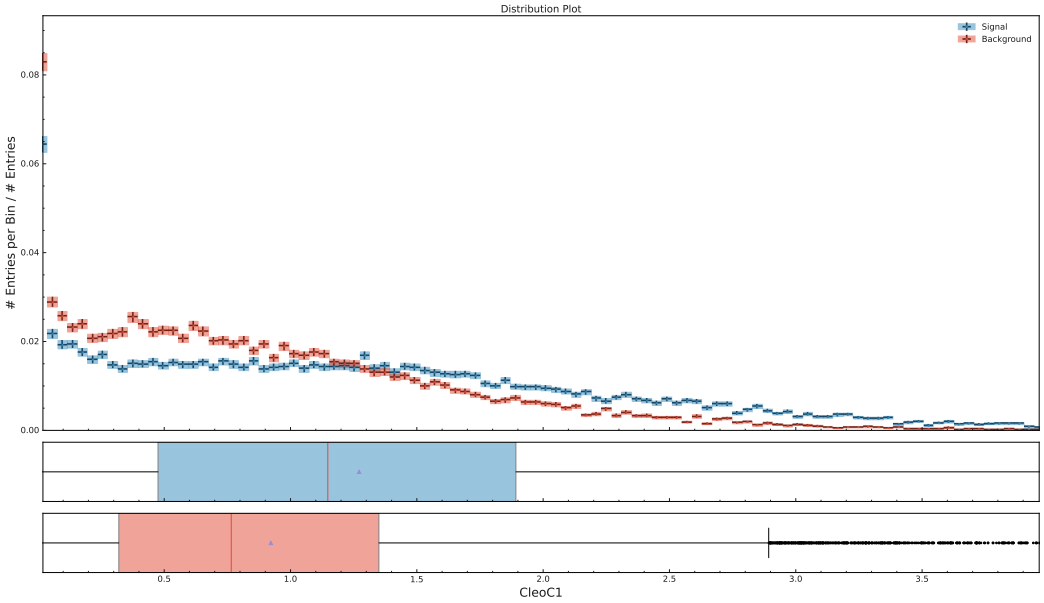




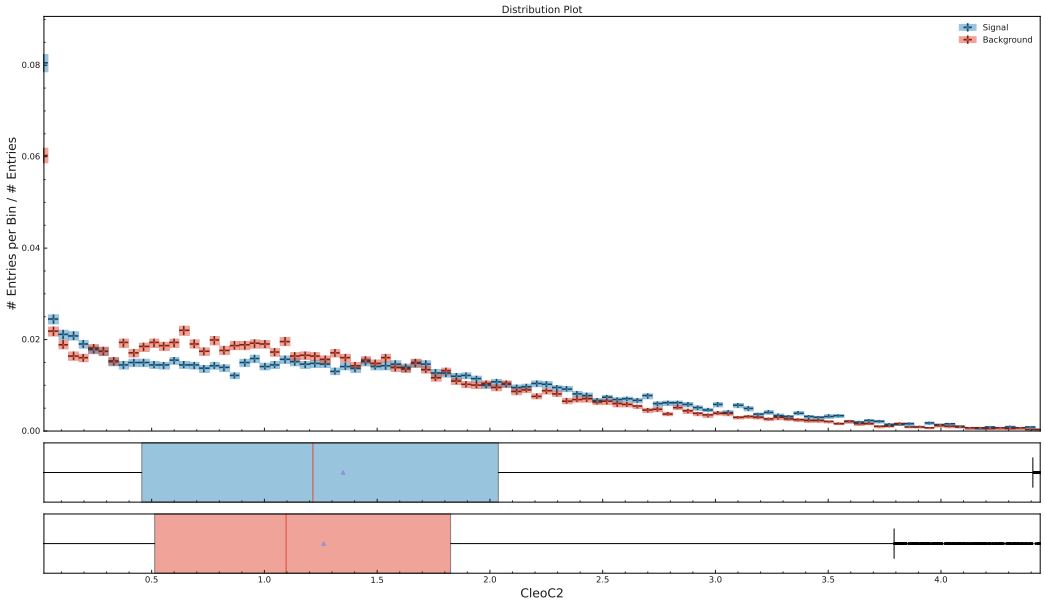
2.3 KSFWVariables(hoo3)



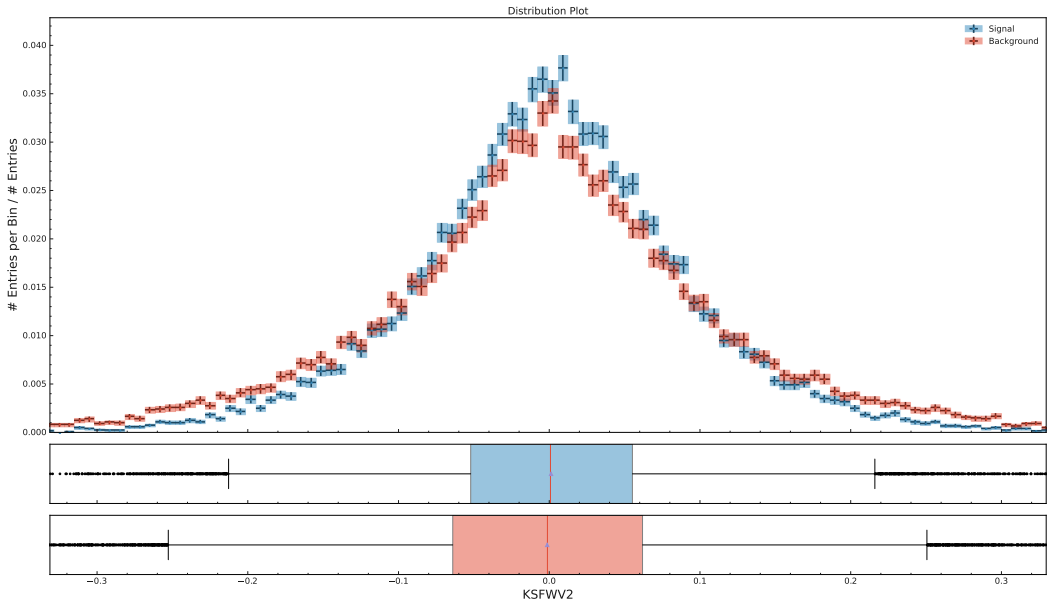
2.4 CleoConeCS(5)



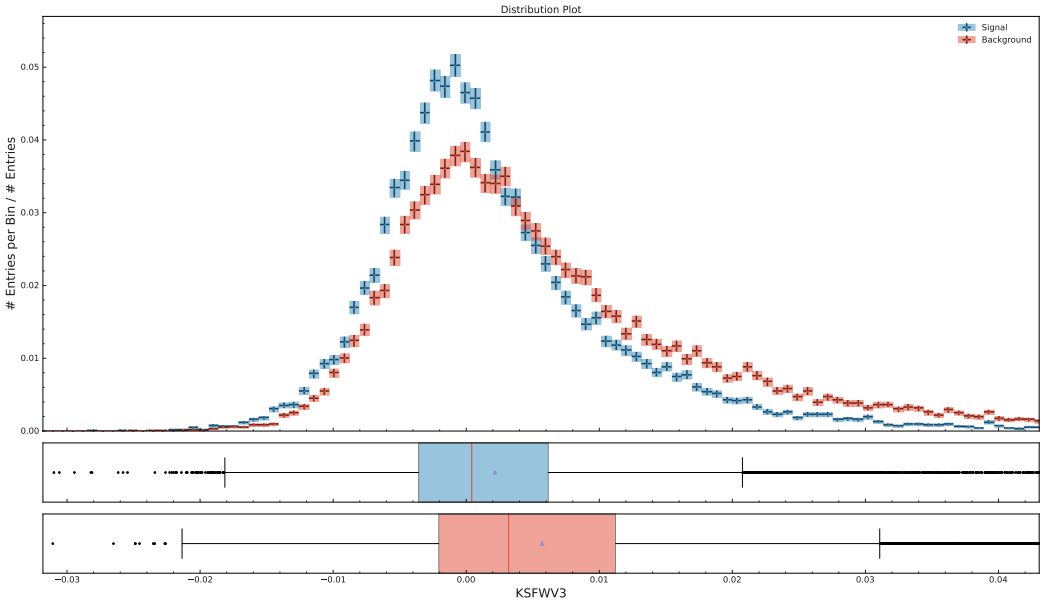
2.5 CleoConeCS(4)



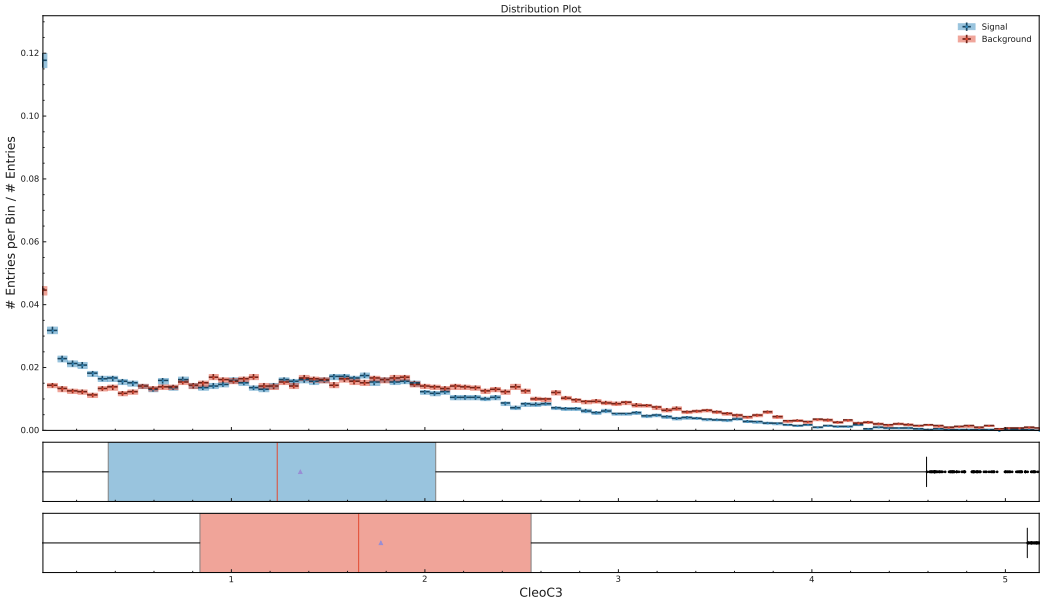
2.6 KSFWVariables(hso03)



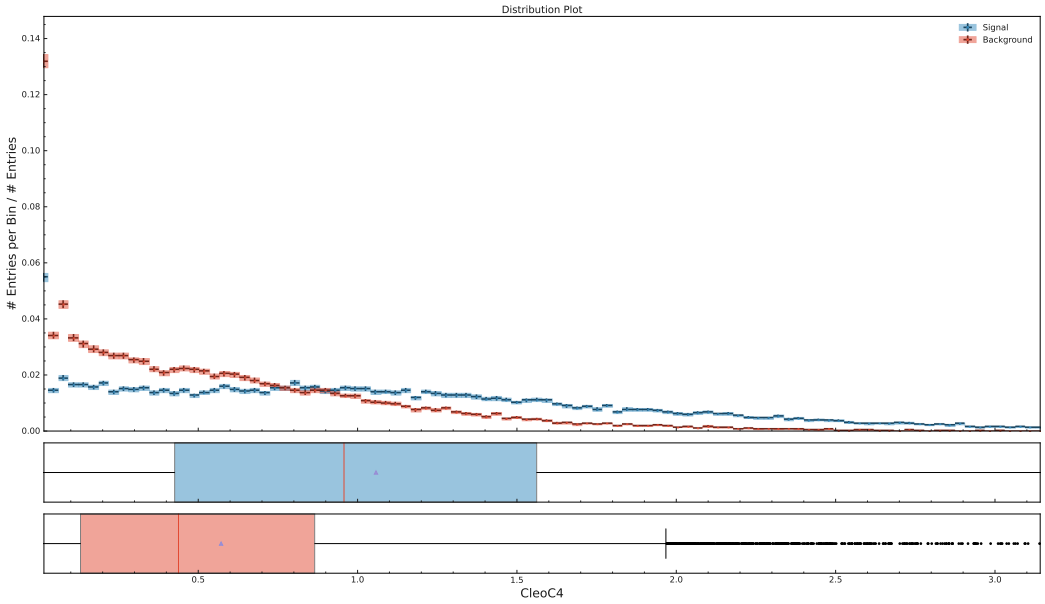
2.7 KSFWVariables(hoo4)



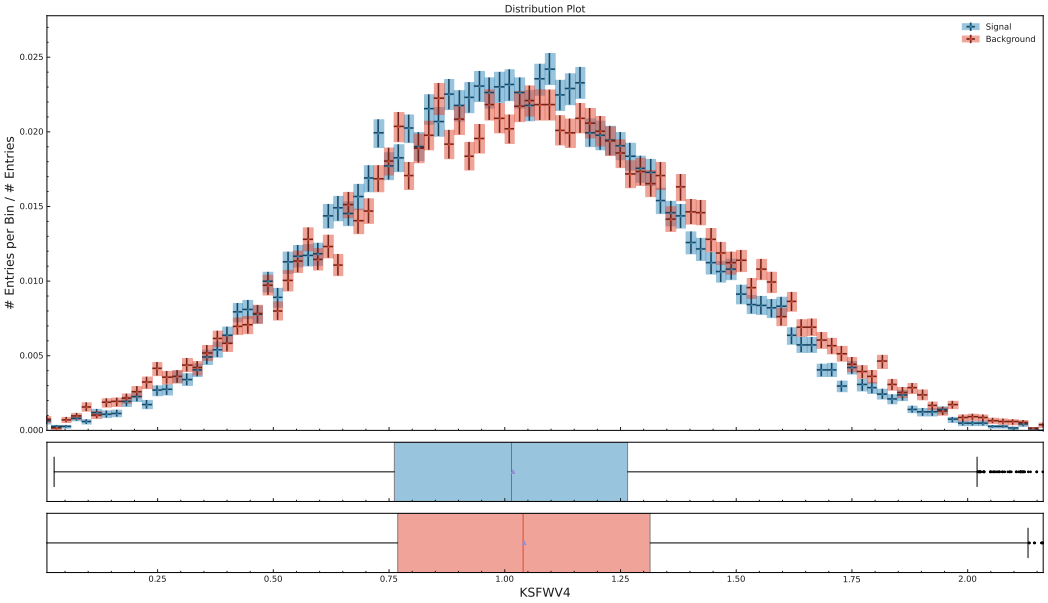
2.8 CleoConeCS(3)



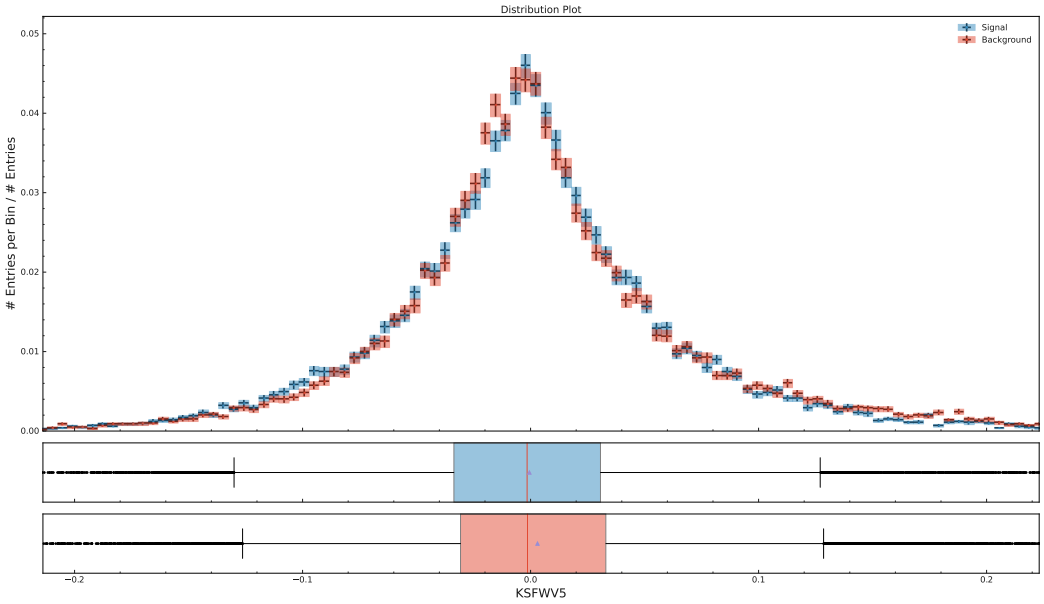
2.9 CleoConeCS(7)



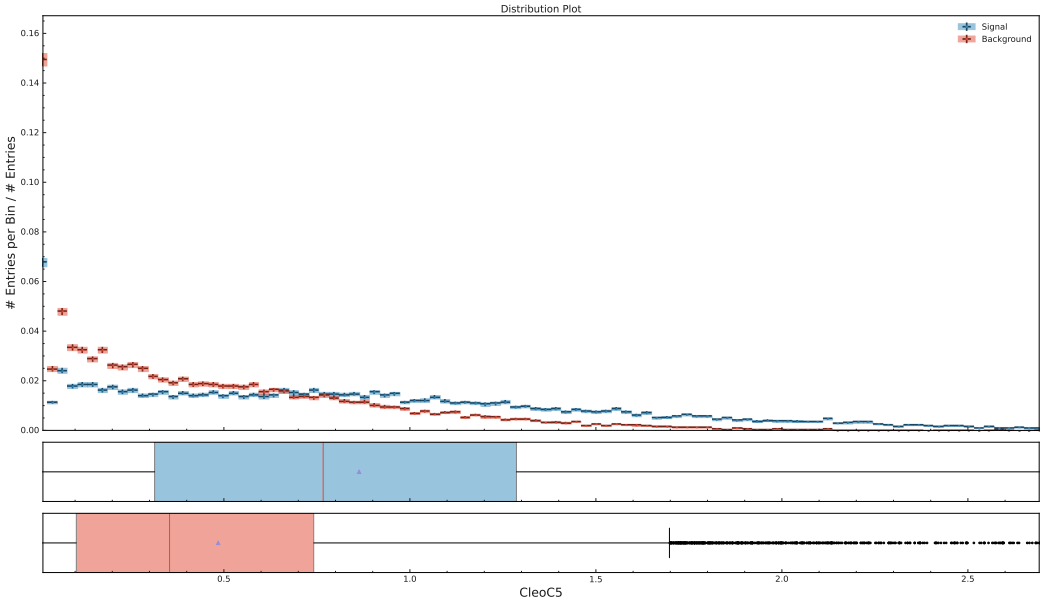
2.10 KSFWVariables(hso00)



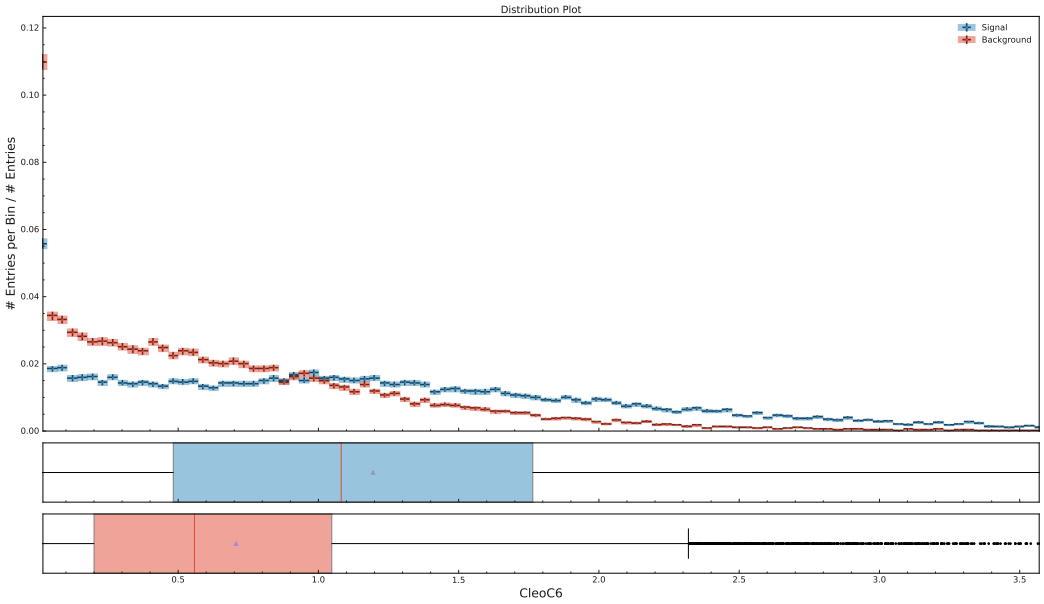
2.11 KSFWVariables(hso24)



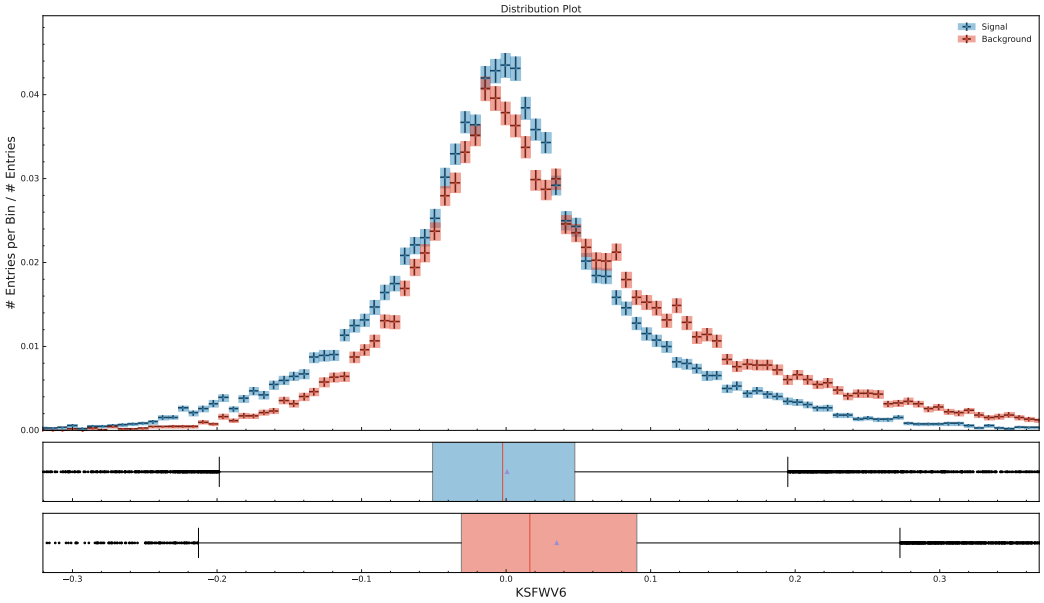
2.12 CleoConeCS(8)



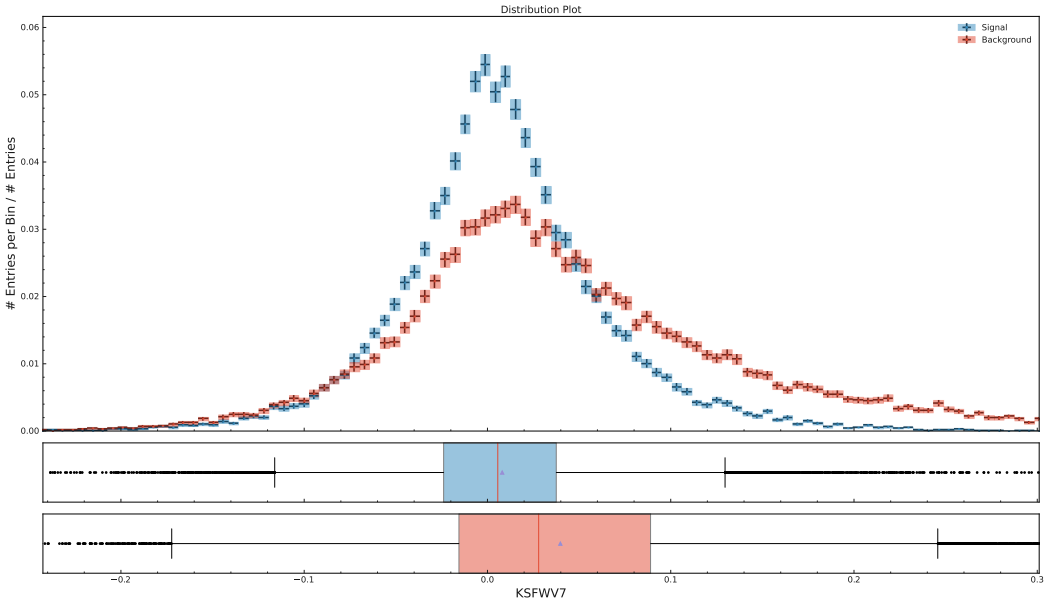
2.13 CleoConeCS(6)



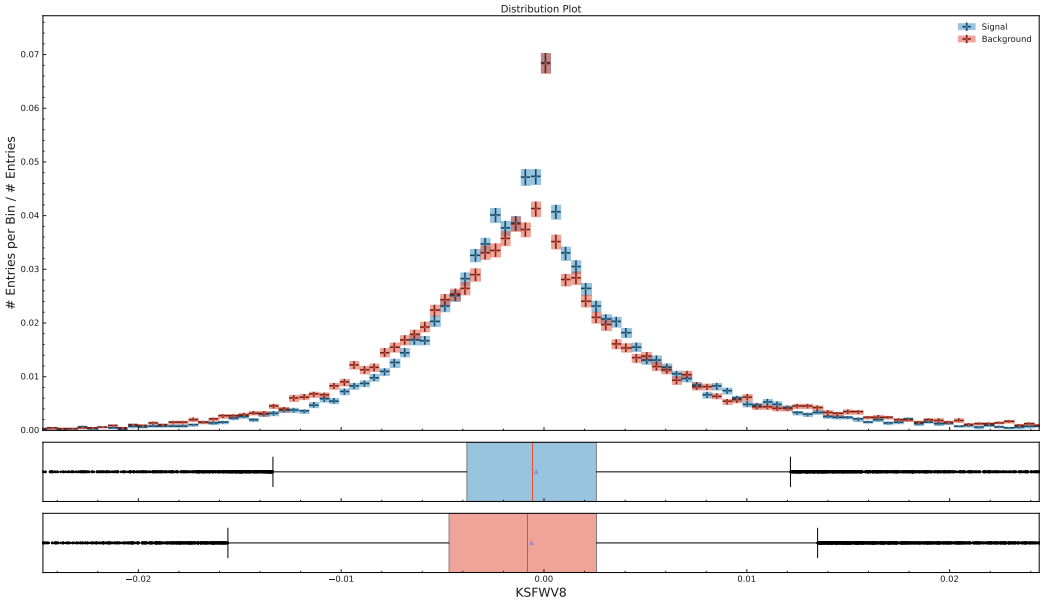
2.14 KSFWVariables(hso22)



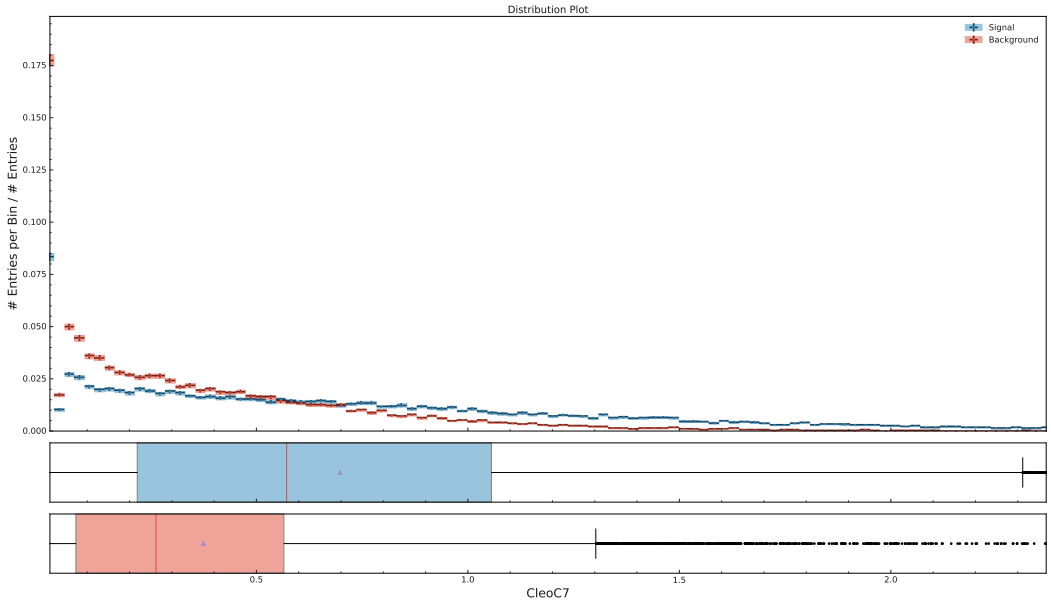
2.15 KSFWVariables(hso14)



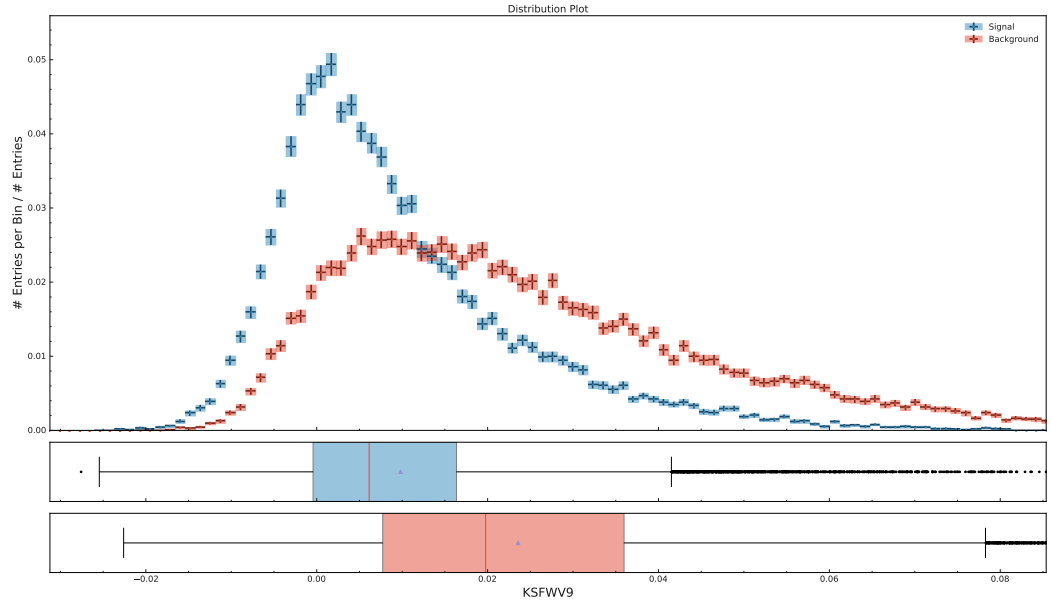
2.16 KSFWVariables(hoo1)



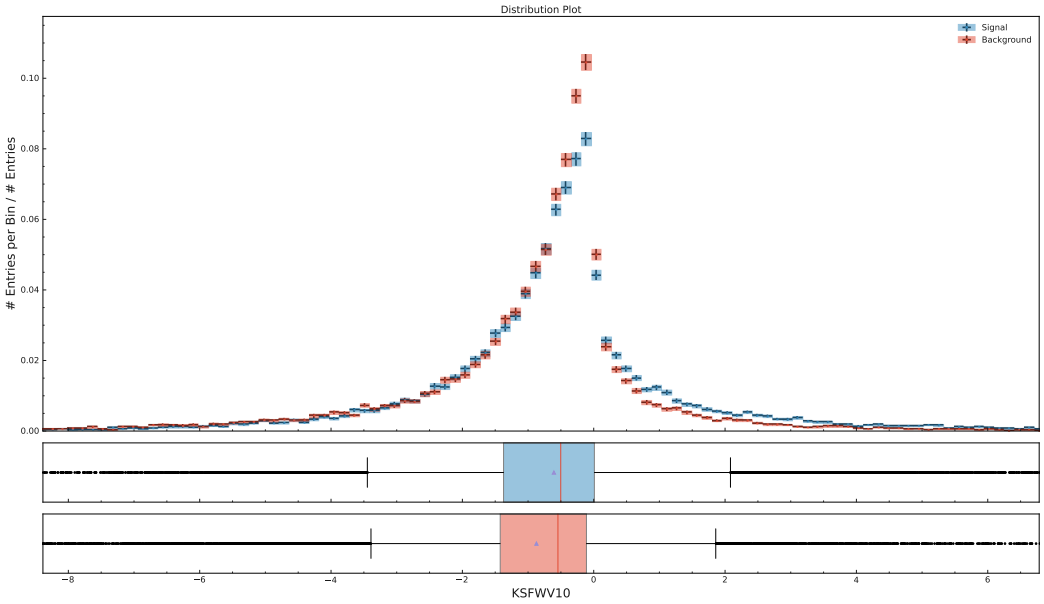
2.17 CleoConeCS(9)



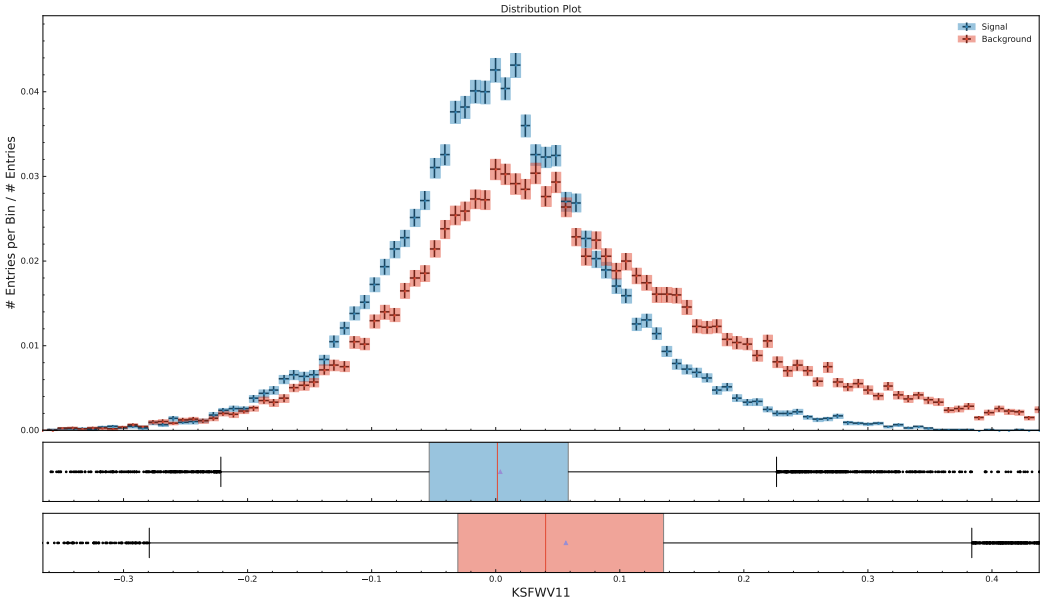
2.18 KSFWVariables(hoo2)



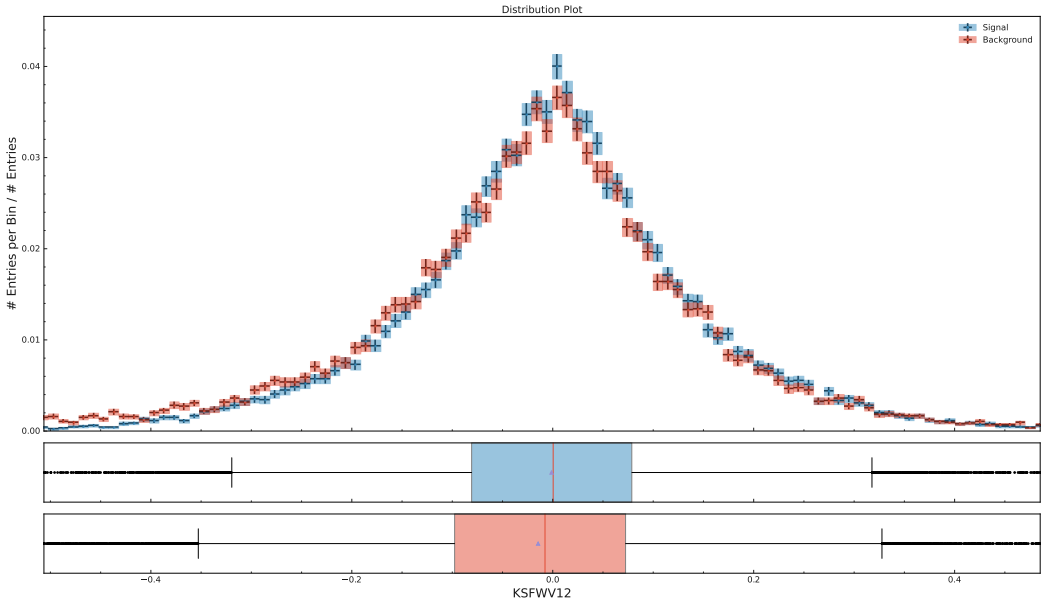
2.19 KSFWVariables(mm2)



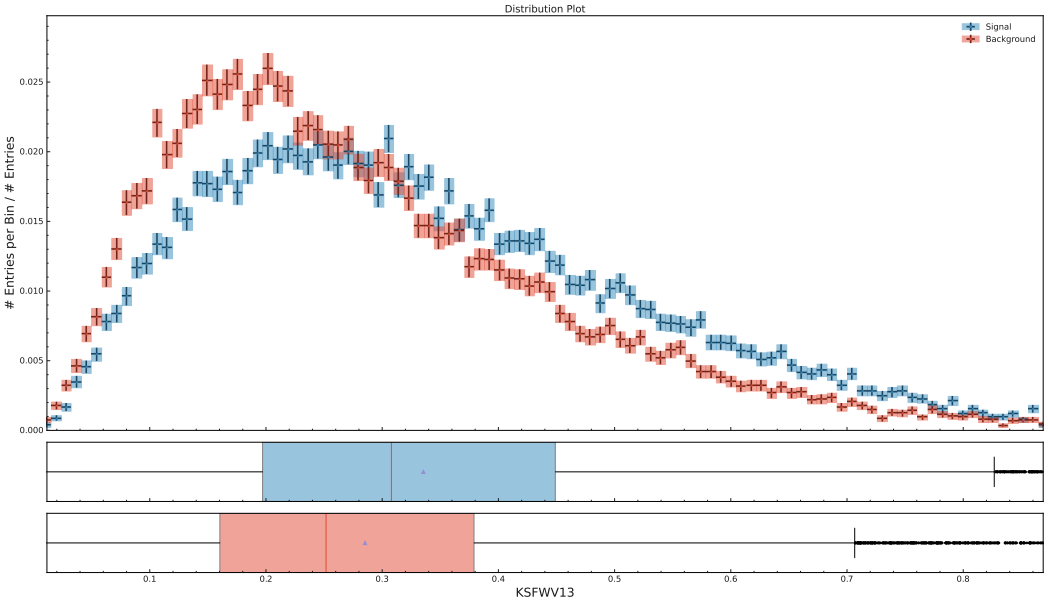
2.20 KSFWVariables(hso04)



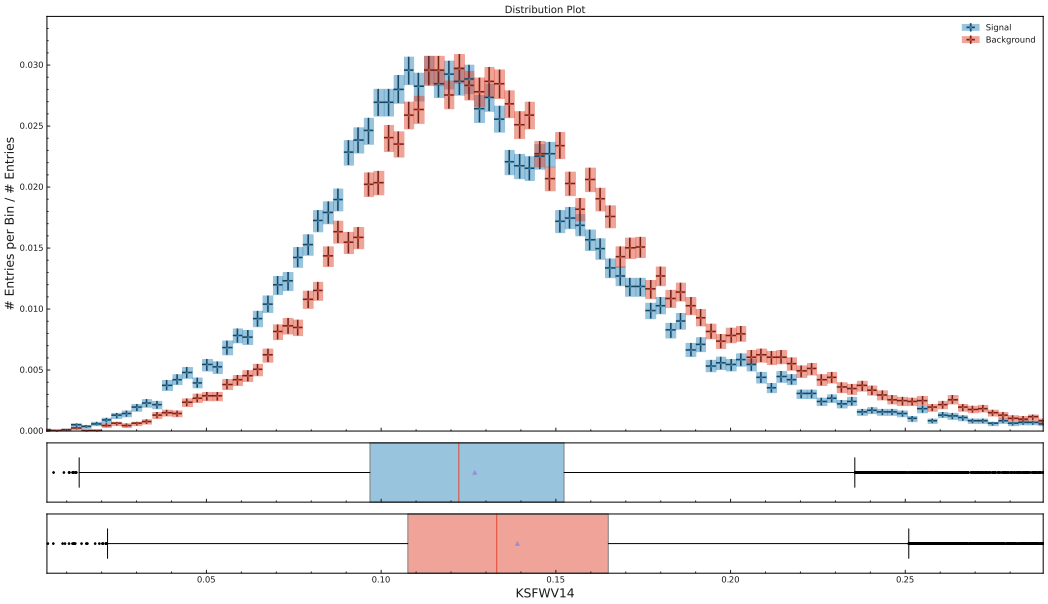
2.21 KSFWVariables(hso01)



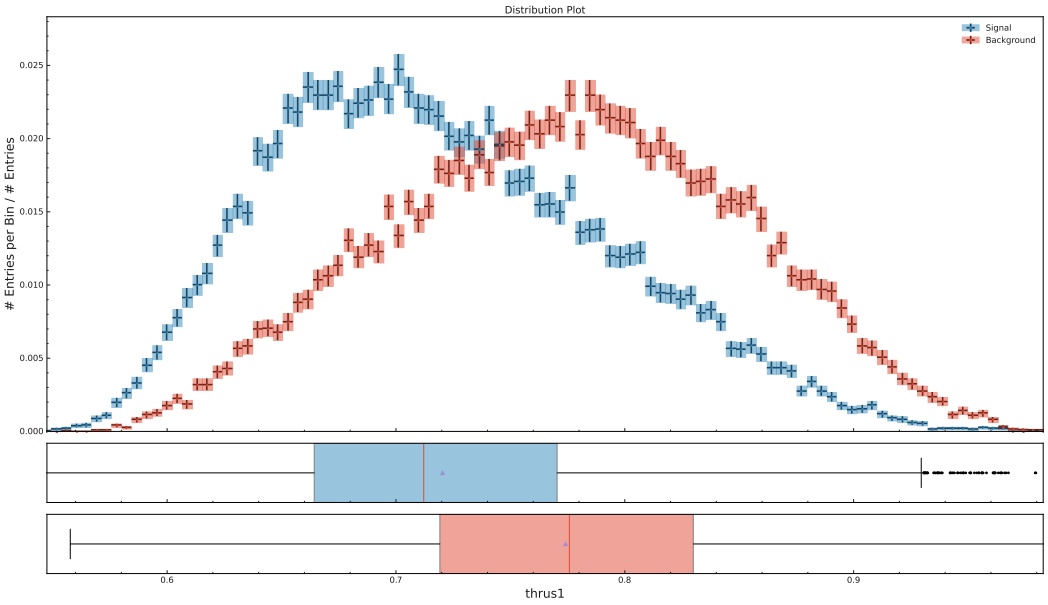
2.22 KSFWVariables(hso20)



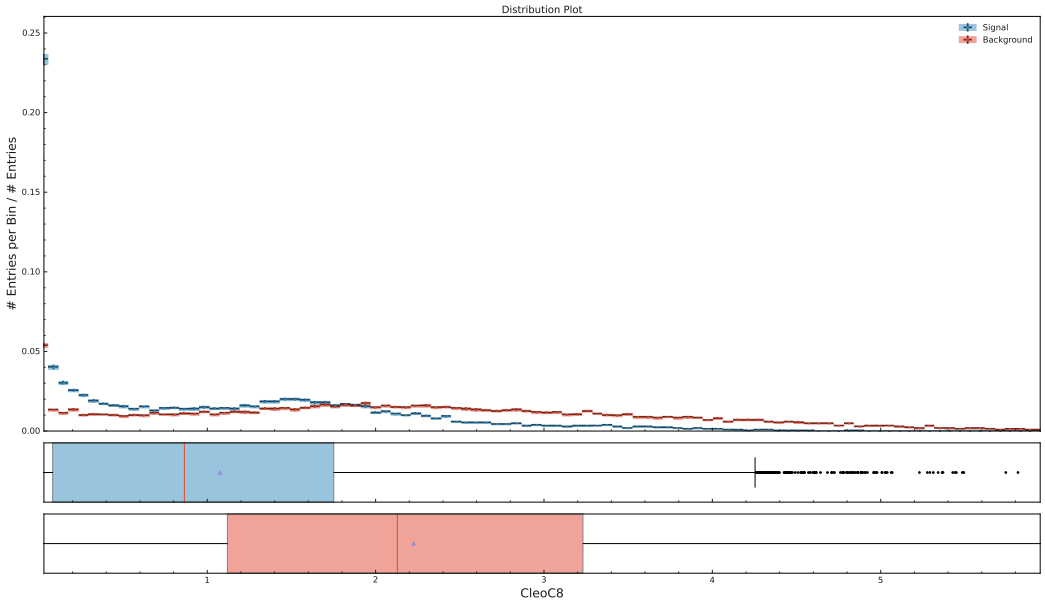
2.23 KSFWVariables(hoo0)



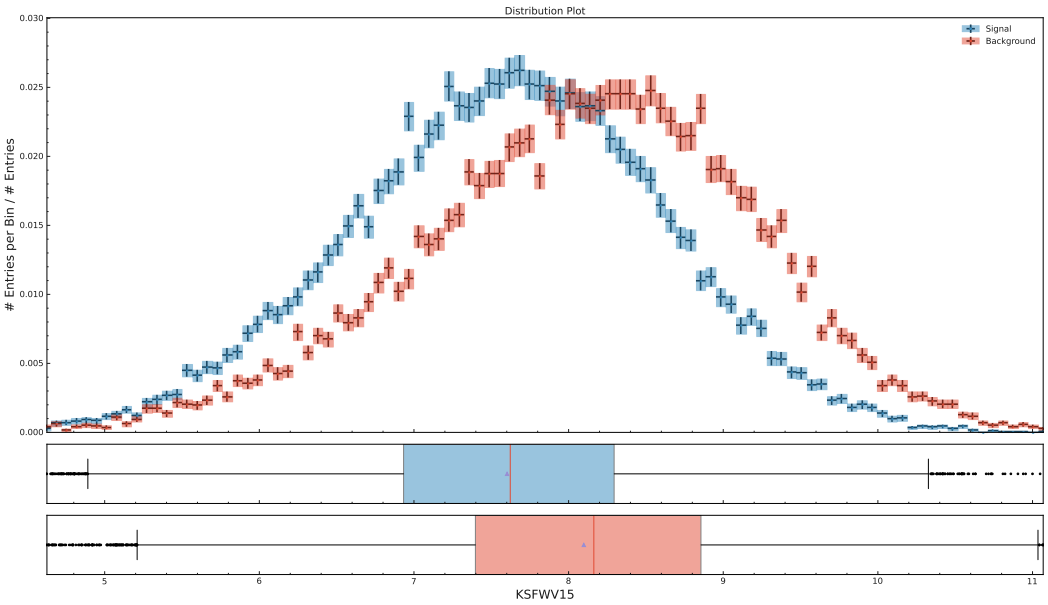
2.24 thrustOm



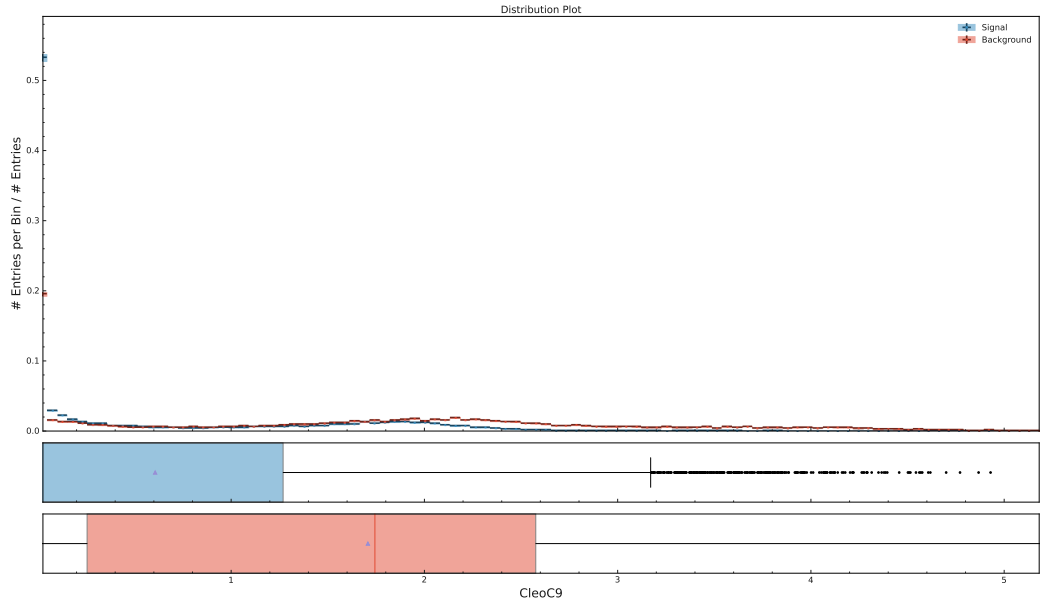
2.25 CleoConeCS(2)



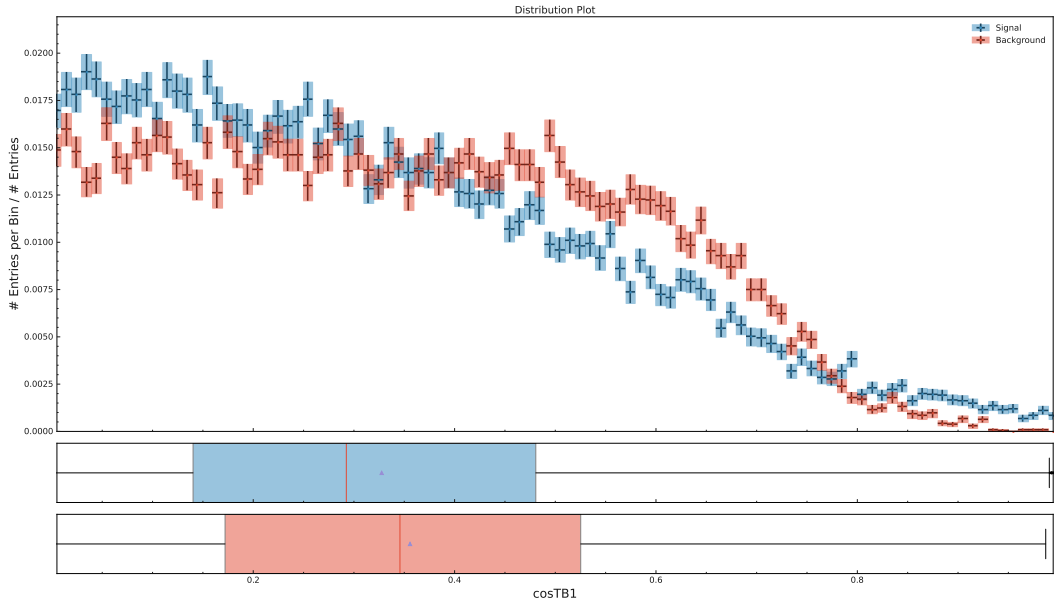
2.26 KSFWVariables(et)



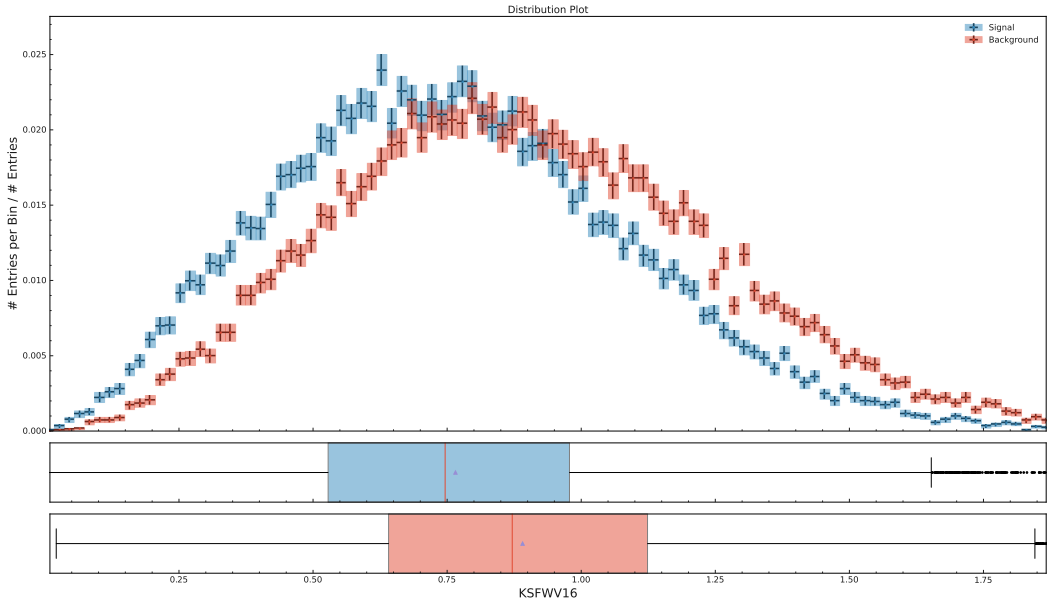
2.27 CleoConeCS(1)



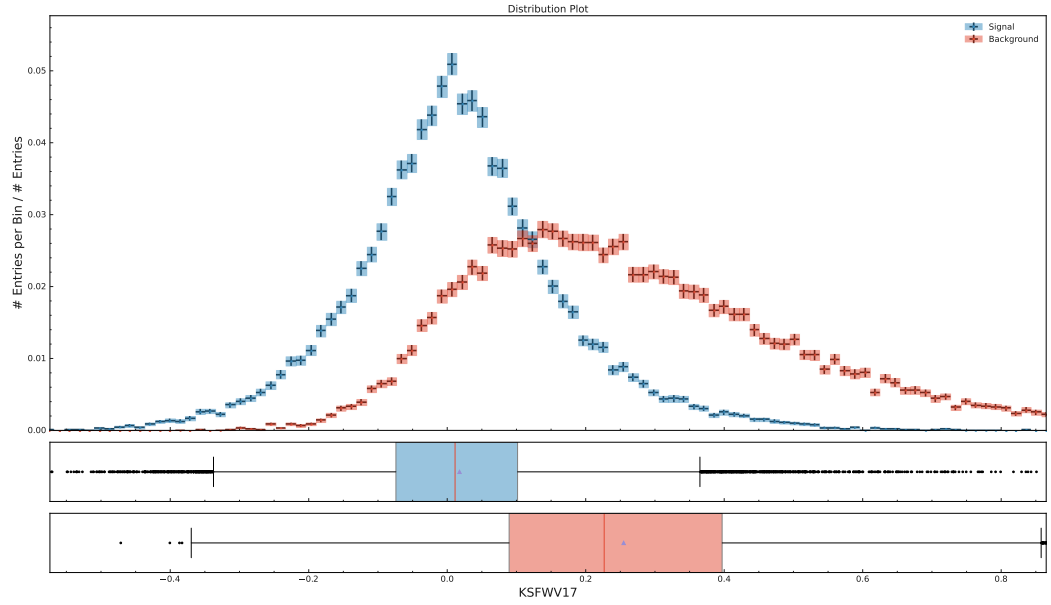
2.28 cosTBz



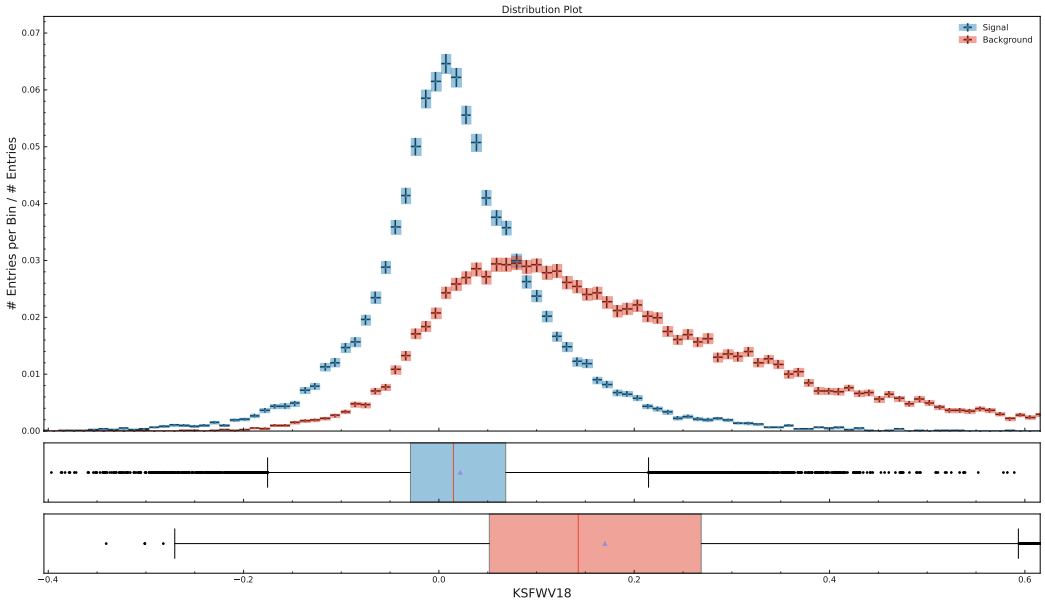
2.29 KSFWVariables(hso10)



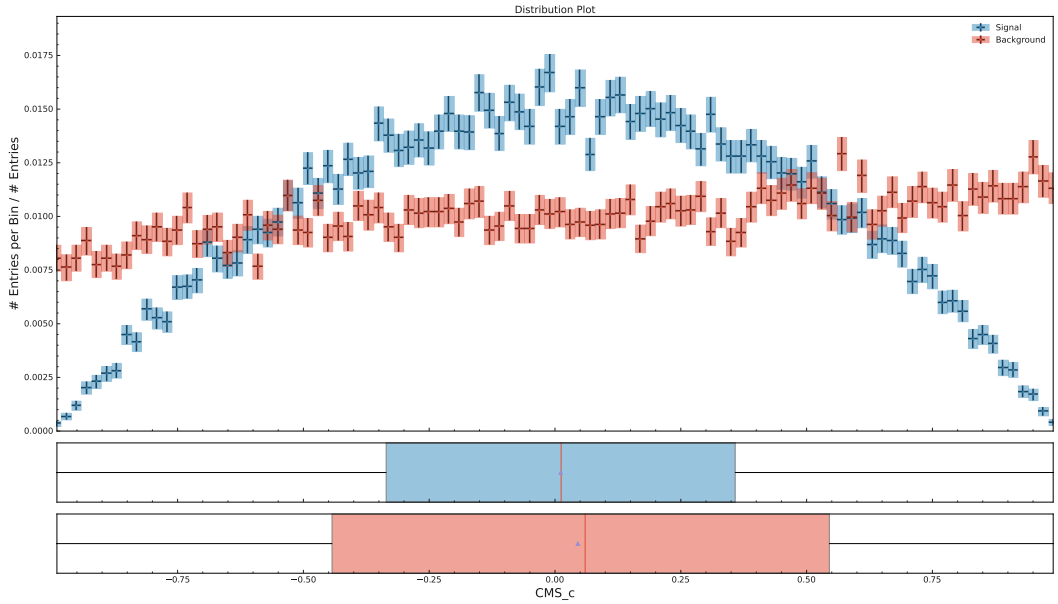
2.30 KSFWVariables(hso02)



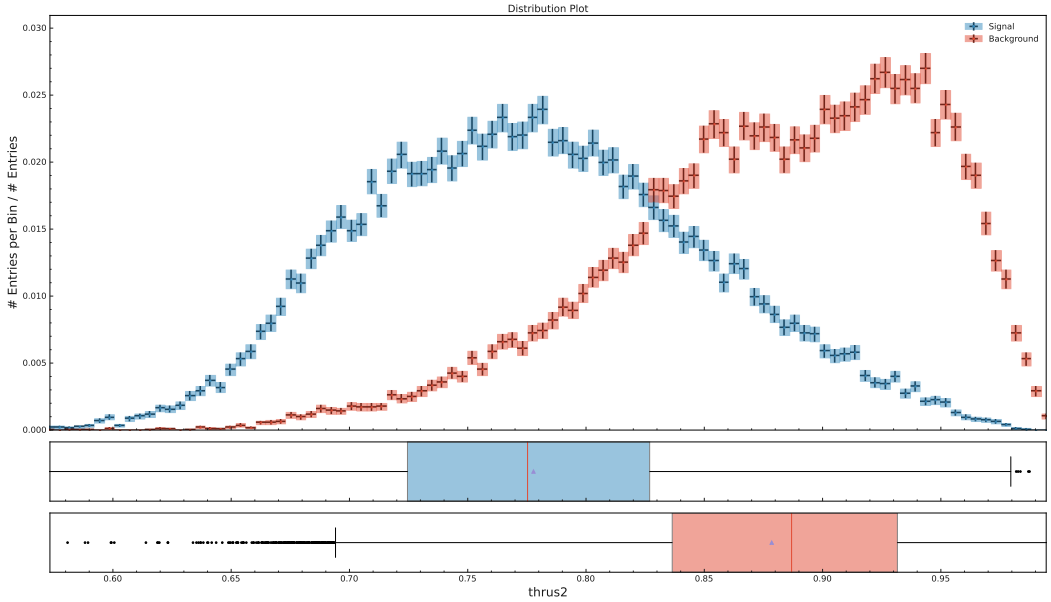
2.31 KSFWVariables(hso12)



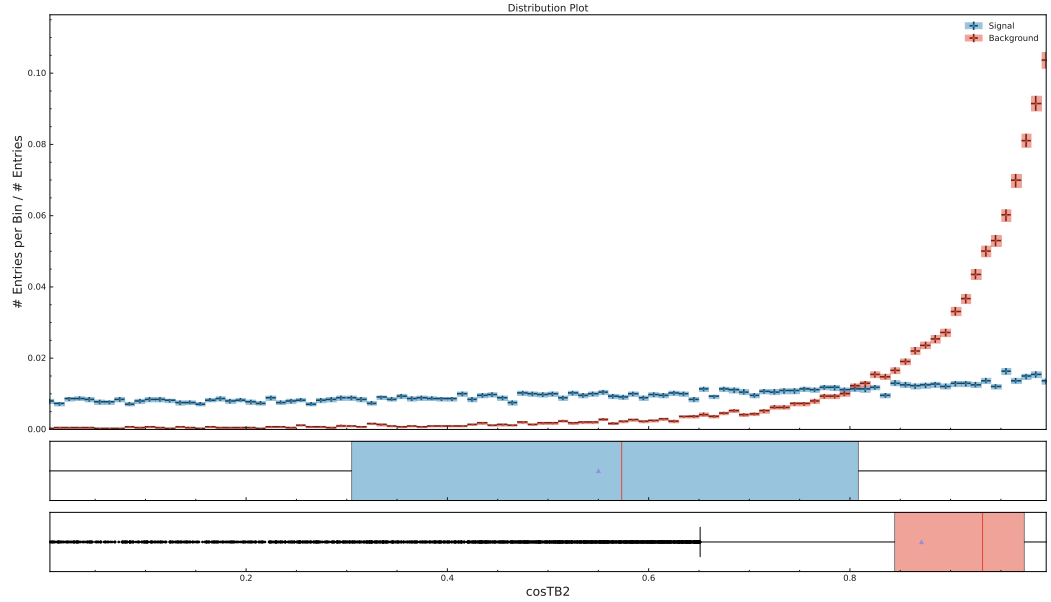
2.32 CMS_cosTheta



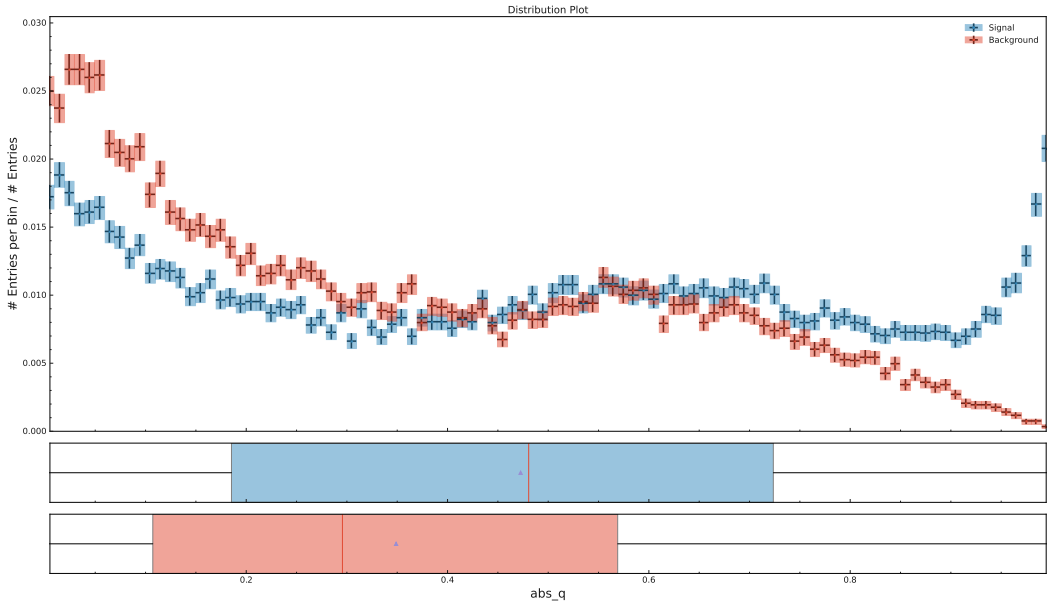
2.33 thrustBm



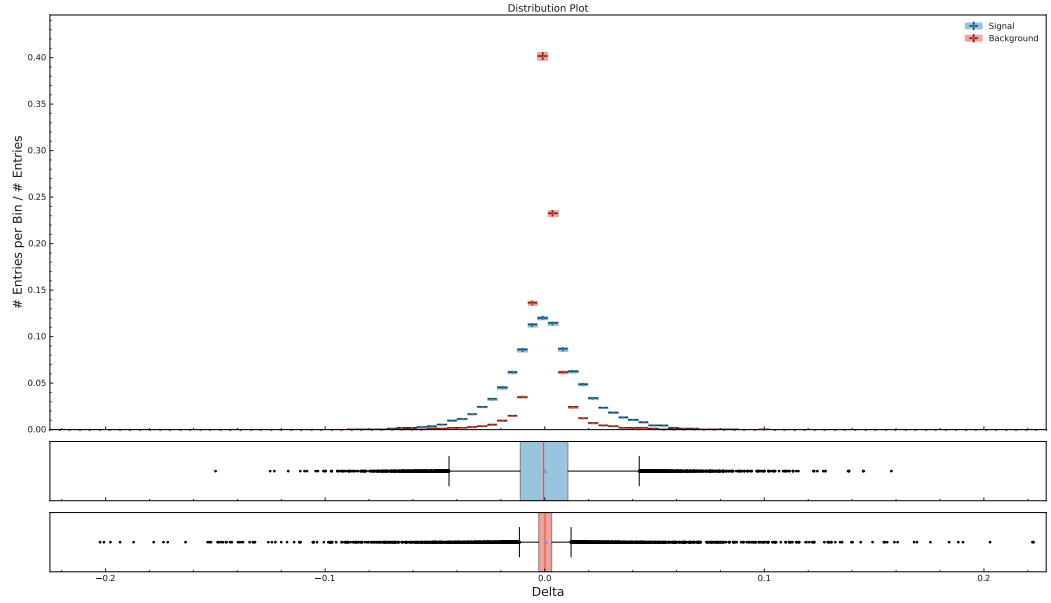
2.34 cosTBTO



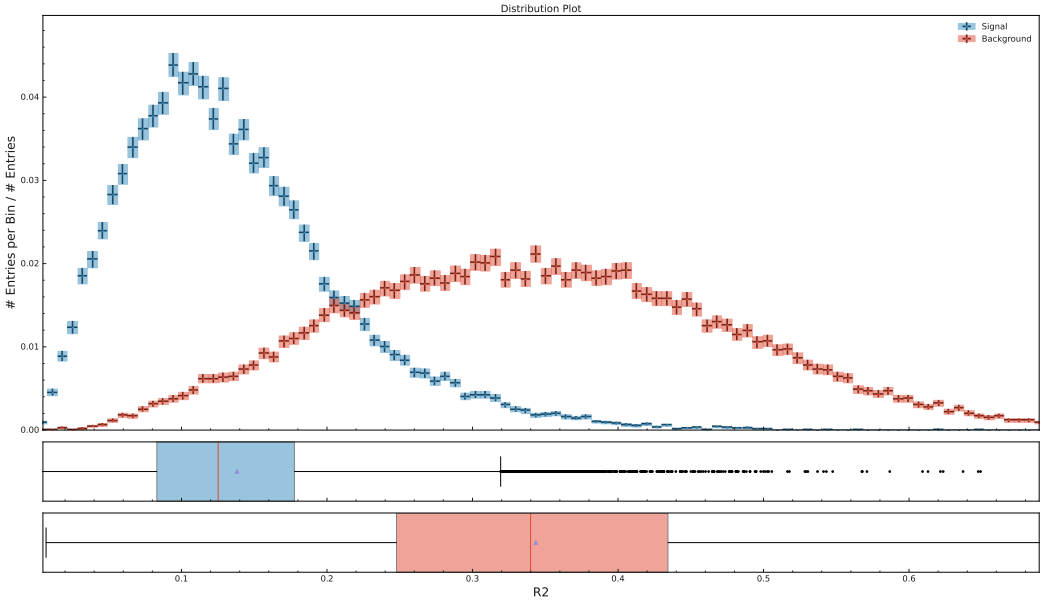
2.35 abs_qr



2.36 DeltaZ



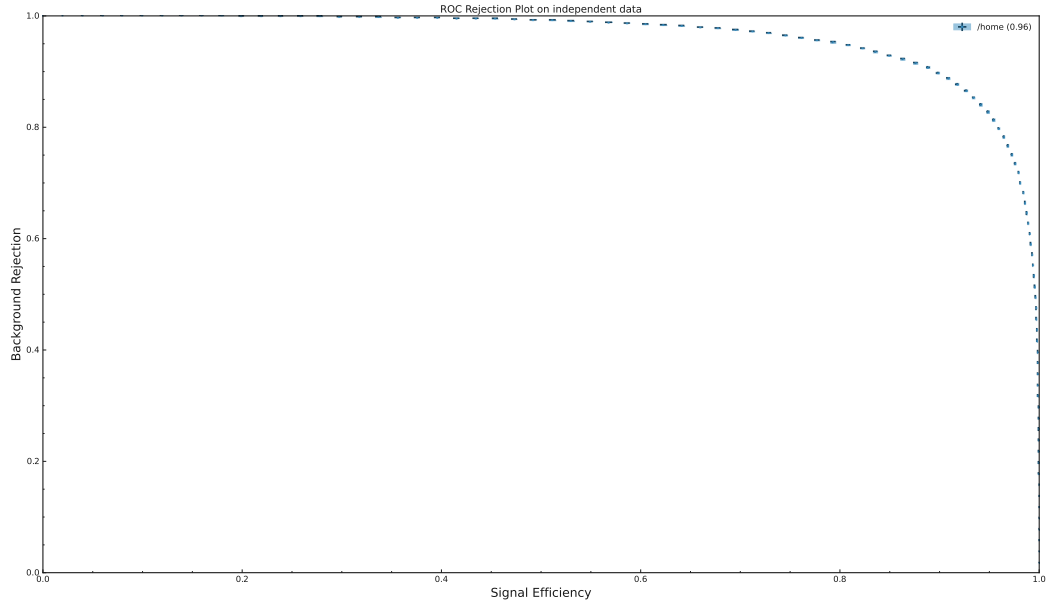
2.37 R2

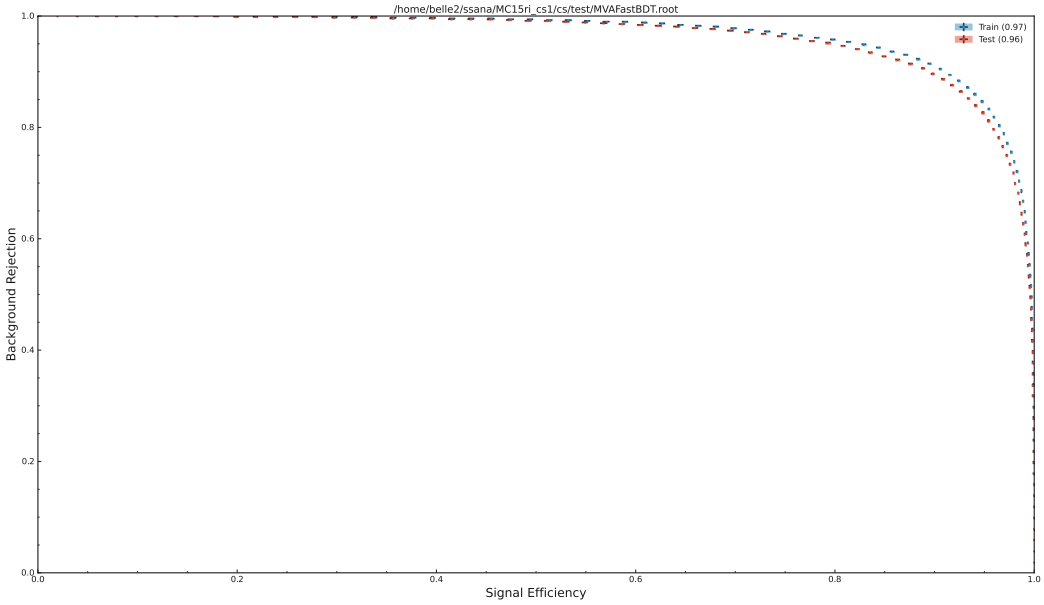


3 Classifier Plot

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 curvein parenthesis.

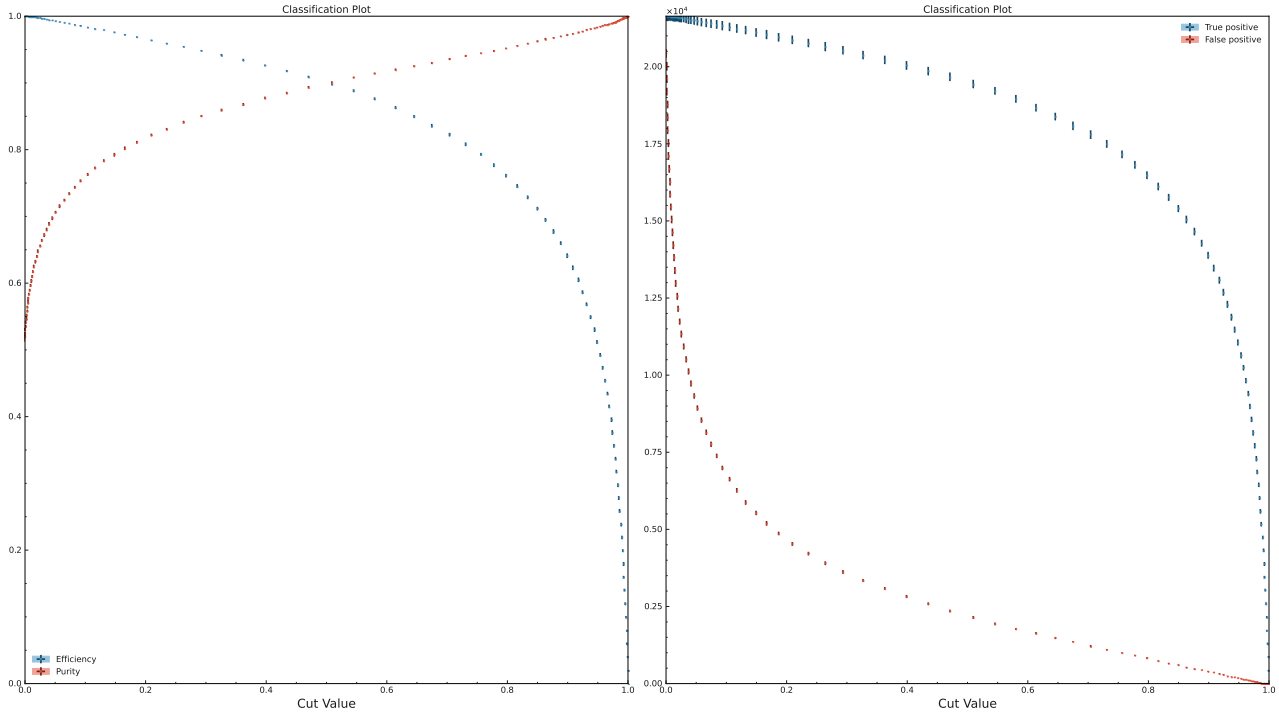
4 ROC Plot





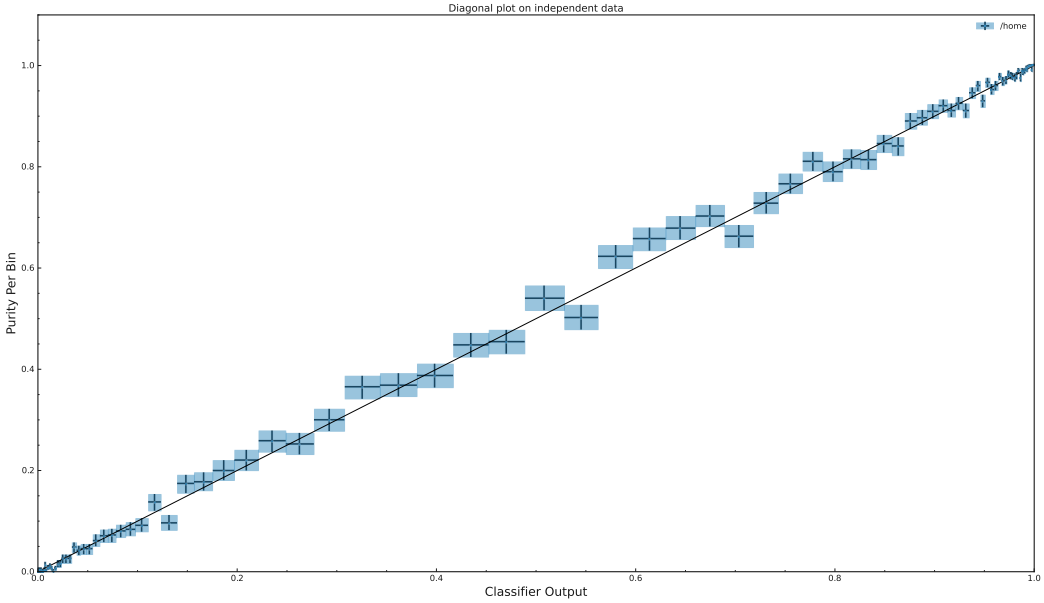
5 Classification Results

5.1 /home

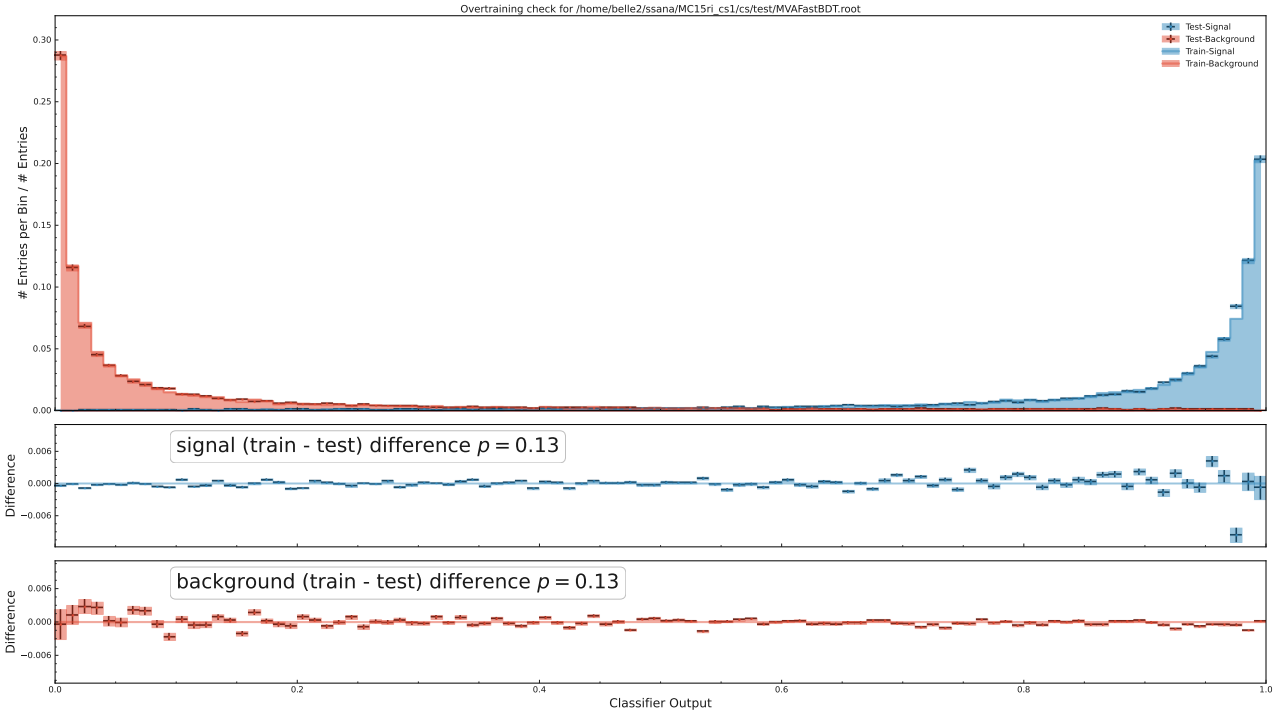


6 Diagonal Plot

6.1 /home



6.2 Overtraining Plot



7 Spectators

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

Table 3: Abbreviations of spectators

Spectator	Abbreviation