Antideuterons search status

Marta Borchiellini – University of Groningen

04/11/24

Charge confusion study

Isotope identification

$$M = \frac{RZ\sqrt{1-\beta^2}}{\beta}$$

Background rejection

Study feature selection for RICH BDT

+

Build charge confusion estimator

Search for Antideuterons

CURRENT PROJECT

Aim:

build a charge confusion estimator (Data-driven BDT) to reject background composed by events whose rigidity sign is wrongly reconstructed.

Dataset

Data:
NAIA ISS Data v1.1.0/ISS.B1236/pass8
4 years (2015-2018) + one year (2023) of data

Preselection: antiproton like cut based selection

Features used: 3103 feature from Tracker and ToF detectors

Sample definition:
Signal Sample: events in mass proton range $0.75 < m < 1.25$ Gev/c^2
Background sample: events with low negative rigidity ($-20 < R < 0$) in the high-mass tail ($2.36 < m < 5$ Gev/c^2)

Antiproton-like selection

Z=1	**Z=1 TOF**	0.5<qtof<1.5 && qlowtof<2.0 && quptof<1.5
	Z=1 Tracker	0.5<q_inntr<1.5.
	God TOF Z	qup<1.5 && qdw<2.0
TOF	**Good TOF NCluster**	NBetaCluster == 4
	Good TOF chisq	chisqtn < 10 && chisqcn < 10
	Has Downgoing Track	Beta_tof>0.5
TRACKER	**Good Inner tracker chisq**	chisqInnerX_GBL< 10 && chisqnInnerY_GBL < 10
	Single track	ntrtrack == 1
	Tracker pattern	L2 && (L3 || L4) && (L5 || L6) && (L7 || L8)
	XY Hits	At least 3 XY hits
	Energy deposition	Less than 2.5 MeV deposited in Inner tracker (LayerEDep)
TRD	**Enough TRD hits**	NHitsOnTrack >10
	Likelihood e/p	Likelihood e/p >0.8
	Likelihood p/He	Likelihood p/He < 0.3
	Physics Trigger	IsPhysicsTrigger() == True
	Rigidity for isotope identification	|R_innner| < 20GV

Beta requirement RICH & TOF

TOF	
Good AGL beta	beta_tof < 0.9
RICH NAF	
Track in NAF	Track in NAF
NAF beta above threshold	Beta > 0.75 beta < 0.99
RICH AGL	
Track in AGL	Track in AGL
AGL beta above threshold	Beta_rich > 0.96 beta_rich < 0.997

Analysis workflow

Refine dataset

Skim ntuples to export relevant features in CSV

Balance the samples: apply undersampling to signal data

Final dataset: ~32000 events equally divided in signal and background

Analysis workflow

Refine dataset
Skim ntuples to export relevant features in CSV
Balance the samples: apply undersampling to signal data
Final dataset: ~32000 events equally divided in signal and background

4 years (2015-2018)		1 year (2023)
Feature Selection (6 months)	Classification (3.5 years)	Validation

Analysis workflow

Refine dataset

Skim ntuples to export relevant features in CSV

Balance the samples: apply undersampling to signal data

Final dataset: ~32000 events equally divided in signal and background

4 years (2015-2018)		1 year (2023)
Feature Selection (6 months)	Classification (3.5 years)	Validation

Feature selection

- Discard features dependent on Rigidity
- Apply different algorithms to select the most discriminant features
- Consider «Physics Driven» sets

Analysis workflow

Refine dataset

Skim ntuples to export relevant features in CSV

Balance the samples: apply undersampling to signal data

Final dataset: ~32000 events equally divided in signal and background

4 years (2015-2018)		1 year (2023)
Feature Selection (6 months)	Classification (3.5 years)	Validation

Feature selection

- Discard features dependent on Rigidity
- Apply different algorithms to select the most discriminant features
- Consider «Physics Driven» sets

Training + Test

- 2 BDt for each set of features (XGB vs AdaBoost)
- Use cross-validation to evaluate performance
- Test models and produce evaluation plots

Analysis workflow

Refine dataset
Skim ntuples to export Tracker + ToF features in CSV
Balance the samples: apply undersampling to signal data
Final dataset: ~32000 events equally divided in signal and background

4 years (2015-2018)		1 year (2023)
Feature Selection (6 months)	Classification (3.5 years)	Validation

Feature selection
- Discard features dependent on Rigidity
- Apply different algorithms to select the most discriminant features
- Consider «Physics Driven» sets

Training + Test
- 2 BDt for each set of features (XGB vs AdaBoost)
- Use cross-validation to evaluate performance
- Test models and produce evaluation plots

Test
Cross-check the models with a completely distinct dataset

Analysis workflow

Refine dataset

Skim ntuples to export Tracker + ToF features in CSV

Balance the samples: apply undersampling to signal data

Final dataset: ~32000 events equally divided in signal and background

4 years (2015-2018)		1 year (2023)
Feature Selection (6 months)	Classification (3.5 years)	Validation

Feature selection

- Discard features dependent on Rigidity
- Apply different algorithms to select the most discriminant features
- Consider «Physics Driven» sets

Training + Test

- 2 BDt for each set of features (XGB vs AdaBoost)
- Use cross-validation to evaluate performance
- Test models and produce evaluation plots

Test

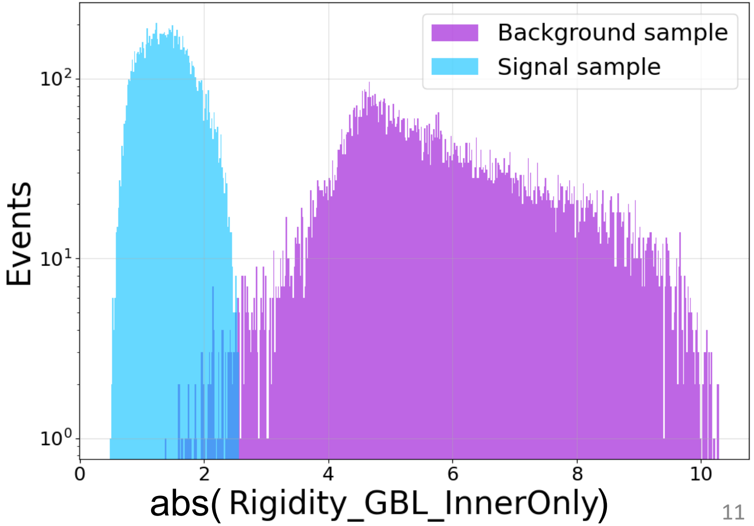
Cross-check the models with a completely distinct dataset

Depence on rigidity

Sample definition:
Signal Sample: events in mass proton range $0.75 < m < 1.25$ Gev/c^2
Background sample: events with low negative rigidity ($-20 < R < 0$) in the high-mass tail ($2.36 < m < 5$ Gev/c^2)

- The rigidity almost completely characterises the two samples (beta_tof < 0.9 cuts all signal events above ~2.5 GV)
- Leak of information if R or rigidity-dependent variables are used for training (BDTs perform perfectly)

→ Discard variables dependent on rigidity



Statistical tests

We need a way to test the rigidity dependence of ~3000 features

Performed different statistical tests on the complete set of features :

- Spearman correlation

test null hypothesis of no correlation

- Krukal-Wallis (KW)
- Kolmogorov Smirnov (KS)

tests if samples are drawn from the same distribution

We compute the p-value for each of the 3 statistical measures

We exclude features with p-value < 0.05.

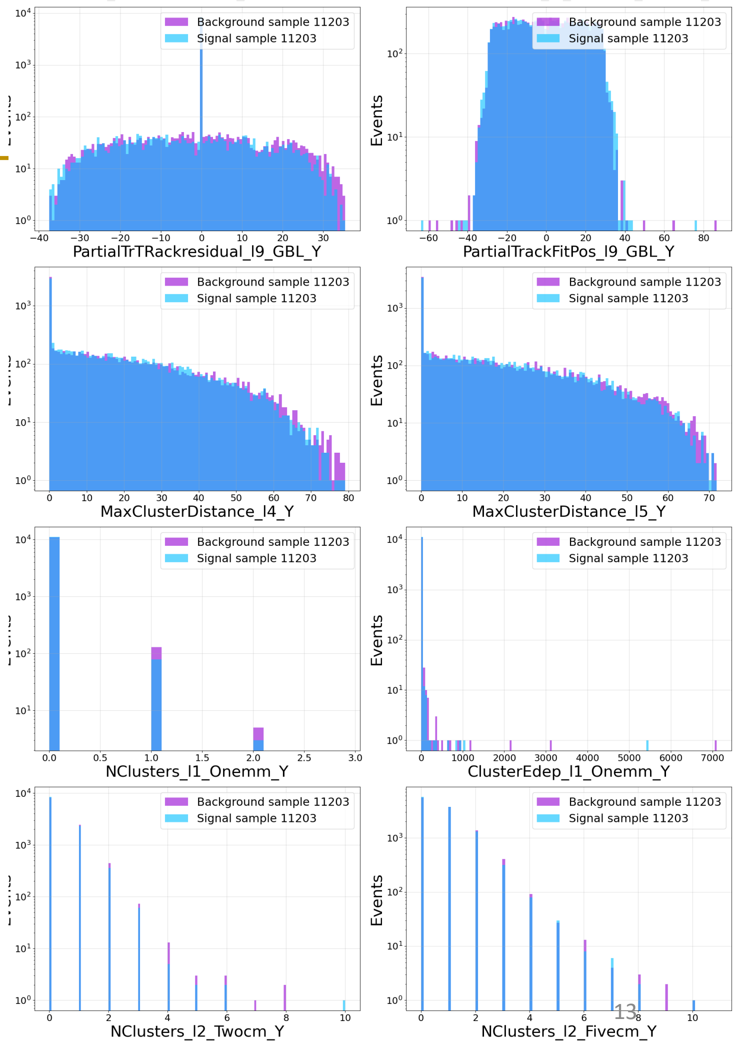
→ We reject features for which there is less than 5% probability to measure the value of the statistic obtained with the chosen test given the null hypothesis

Statistical tests - results

Number of features passing the tests:

- Spearman correlation: 679
- Krukal-Wallis: 661
- Kolmogorov Smirnov test: 187

- Kolmogorov Smirnov (KS) is still the more conservative
- The KS features seem to be independent of rigidity from visual inspection
- Features don't seem to be very discriminative



Feature selection with KS features

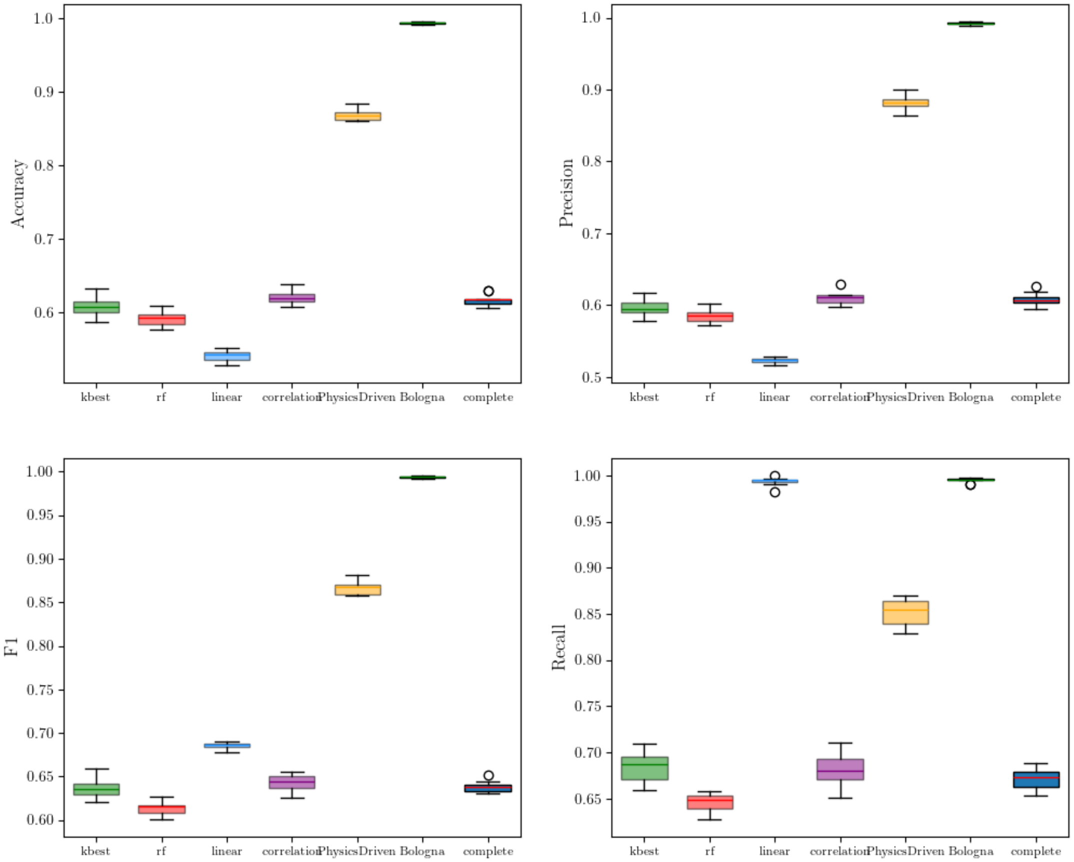
We input in the feature selection + classification pipeline the following features

- A ”complete set” of 187 features passing the KS test
- A Physics Driven set composed of 11 variables
- The set of 12 features used by the Bologna group

N.B. Physics Driven features and Bologna features do not pass the KS test with 0.05 threshold

Technique	Number of features selected
kbest	103
Random forest	53
Pearson’s correlation	160
Linear regression	1
Physics driven	11
Bologna	13

Cross-validation metrics (training phase)



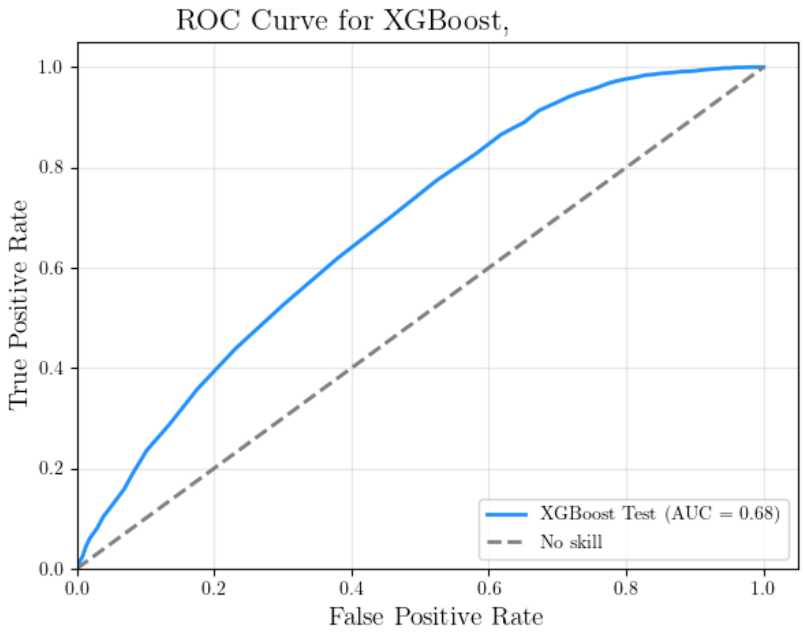
Performance metrics for XGBoost BDTs trained with the different feature sets

AdaBoost achieves similar performances

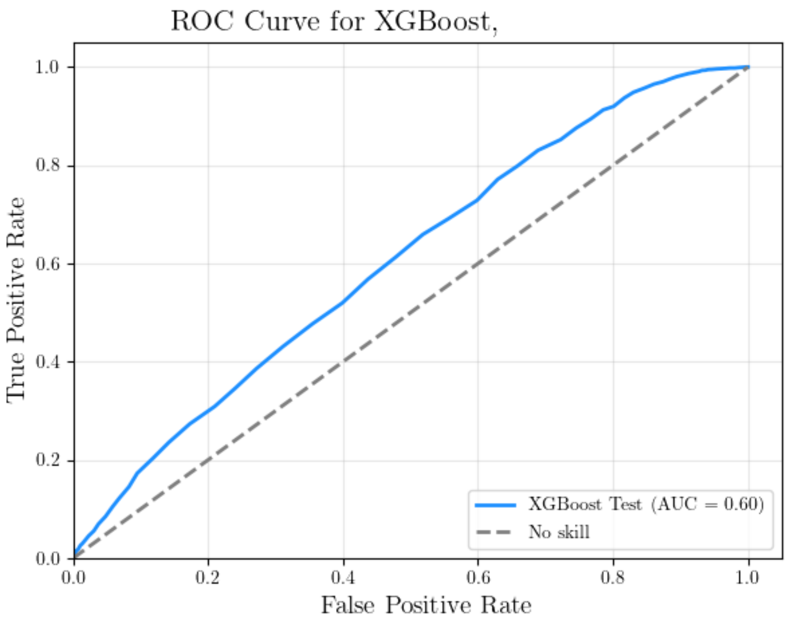
All the methods using subsets of the KS features perform poorly.

ROC curve – kbest

Validation 2015-2018 dataset

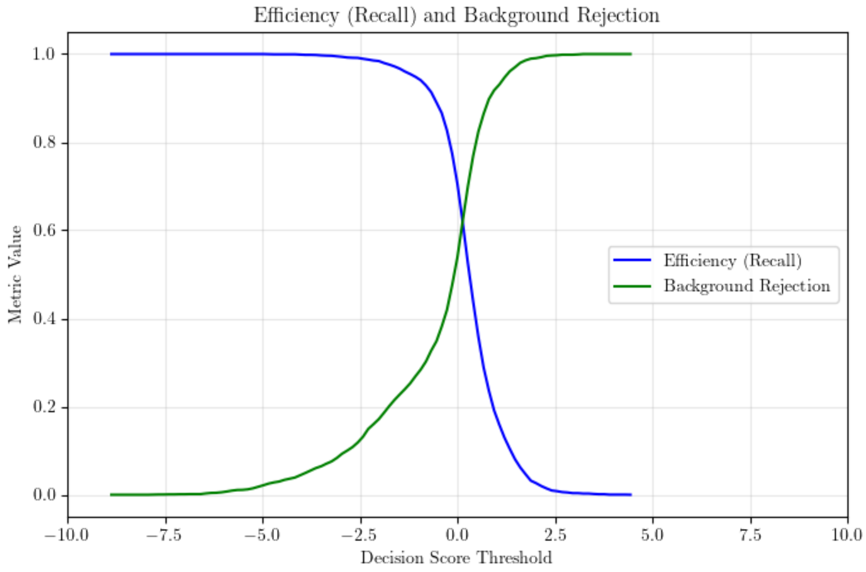


Validation 2023 dataset

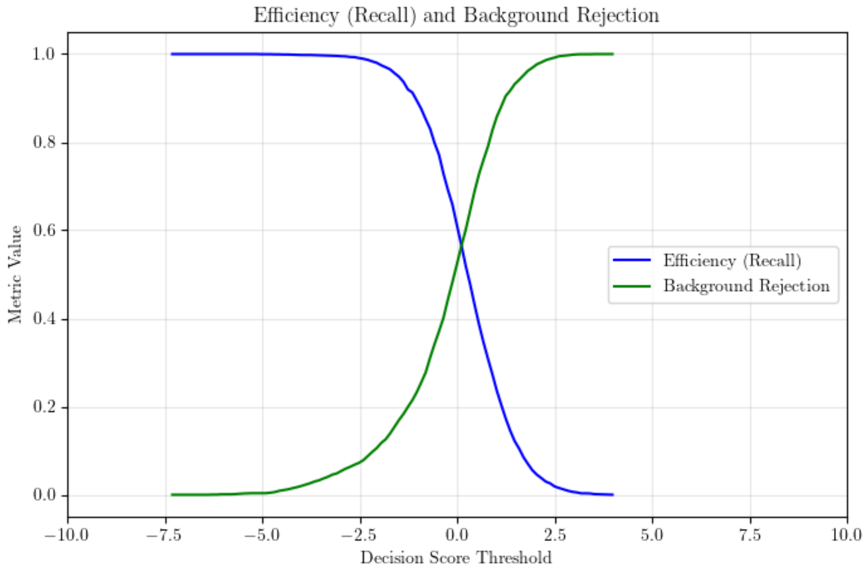


Efficiencies – kbest

Validation 2015-2018 dataset (XGB)

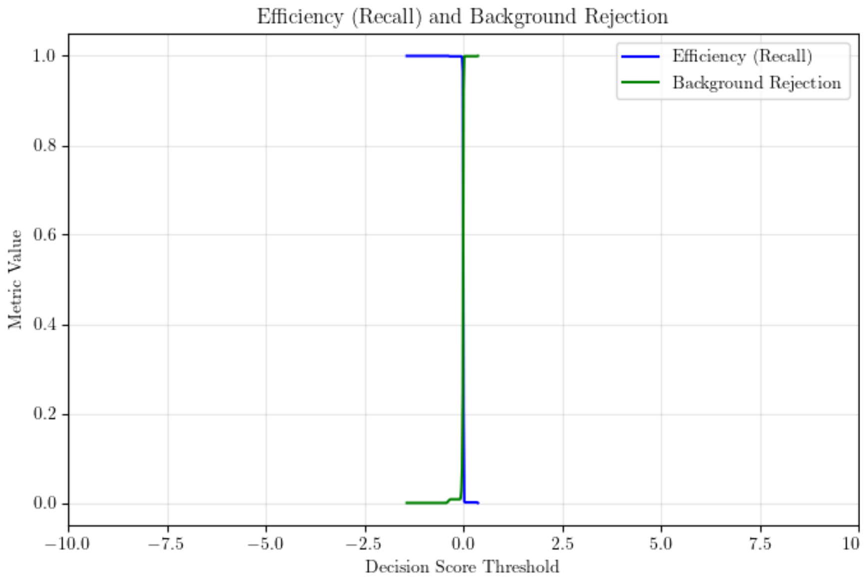


Validation 2023 dataset (XGB)

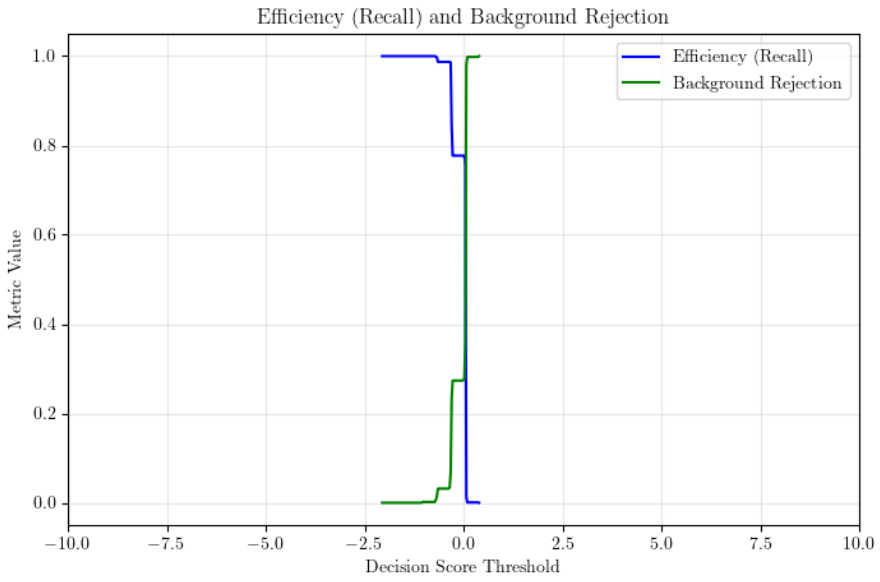


Efficiencies – kbest

Validation 2015-2018 dataset (Ada)



Validation 2023 dataset (Ada)



Conclusions

- All the BDT trained with features selected among the KS set perform poorly. XGBoost performs better than AdaBoost in terms of efficiencies

- The BDTs trained with Physics Driven features and Bologna features perform better. However, the features used do not pass the KS test with a threshold of 0.05

- All the models perform way worse on the 2023 dataset

- The definition of background and signal samples based on the mass associated with the current event selection limits the BDT performance as it requires excluding rigidity-dependent features , leaving just the less discriminative to be used.

Future steps

- Tune KS test threshold to know good variables
- Just select features in a physics-driven way
- Release cut on Beta_tod to have more signal events at higher rigidities?
- Use MC events?
- …?

BACK UP

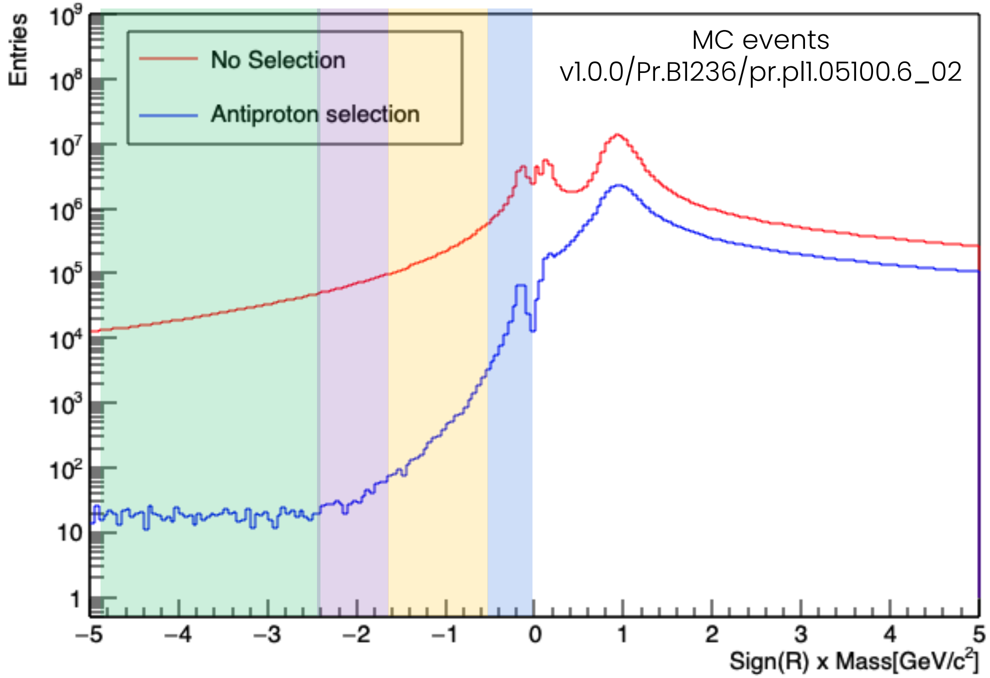
Background events

Light particles

p->pbar

p->dbar

High mass tail



Enlarged dataset

Previous dataset:

1 year (2015) + 2months (2023)

Undersampling was applied to the sign sample to even the number of events in the background and signal sample

~ 7400 events in total (3200 Feature selection, 3200 Classification, 1000 Validation test)

Dataset was enlarged to:

4 years (2015-2018) + one year (2023)

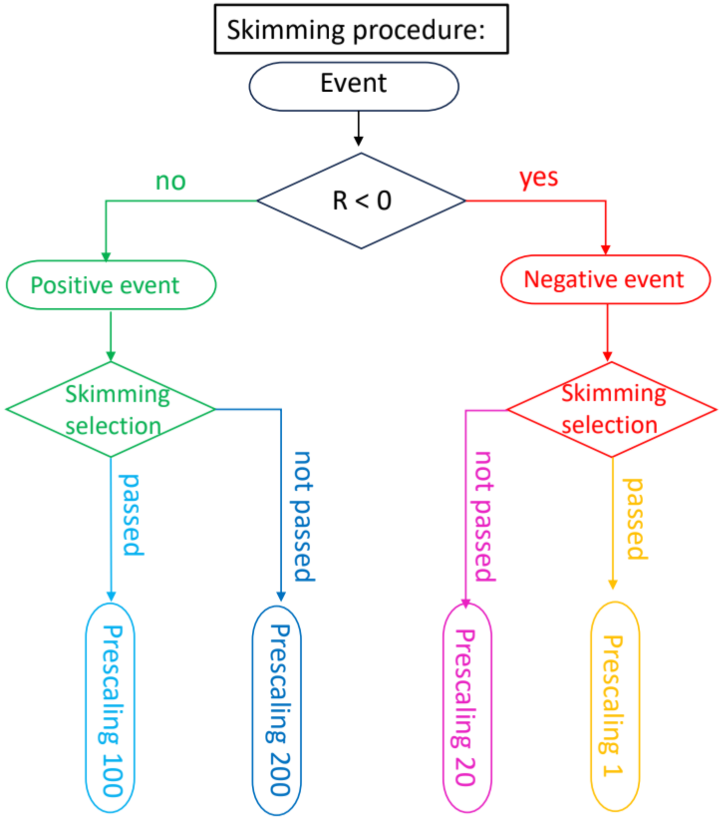
Undersampling applied as before

~ 32000 events in total (~ 3200 for Feature Selection , ~ 22600 for Classification, ~ 6000 for Validation test)

Skimming strategy

Goal: reduce the ntuple size otherwise the conversion to csv would take long time an disk space

Size:
~60Mb/ntuple ~115 Gb/month
~1,3Tb/year
→"skimming selection" == antiproton-like selection



Taken from F. D'angelo slides

Feature Selection

Machine Learning (ML) feature selection methods used:

- Kbest
- Random Forest
- Linear Regression
- Pearson's Correlation

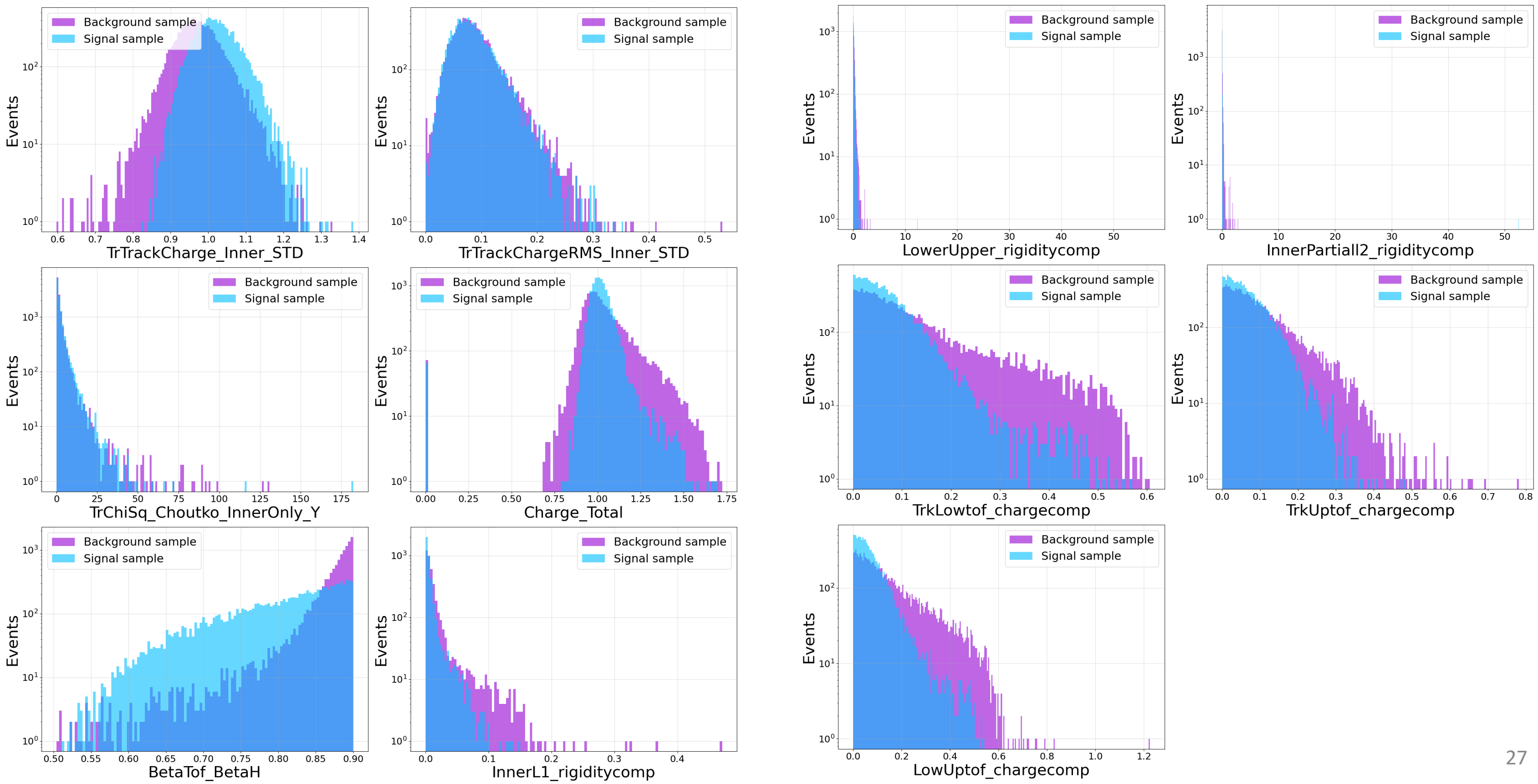
Feature dataset used for the analysis

We input in the feature selection + classification pipeline the following features

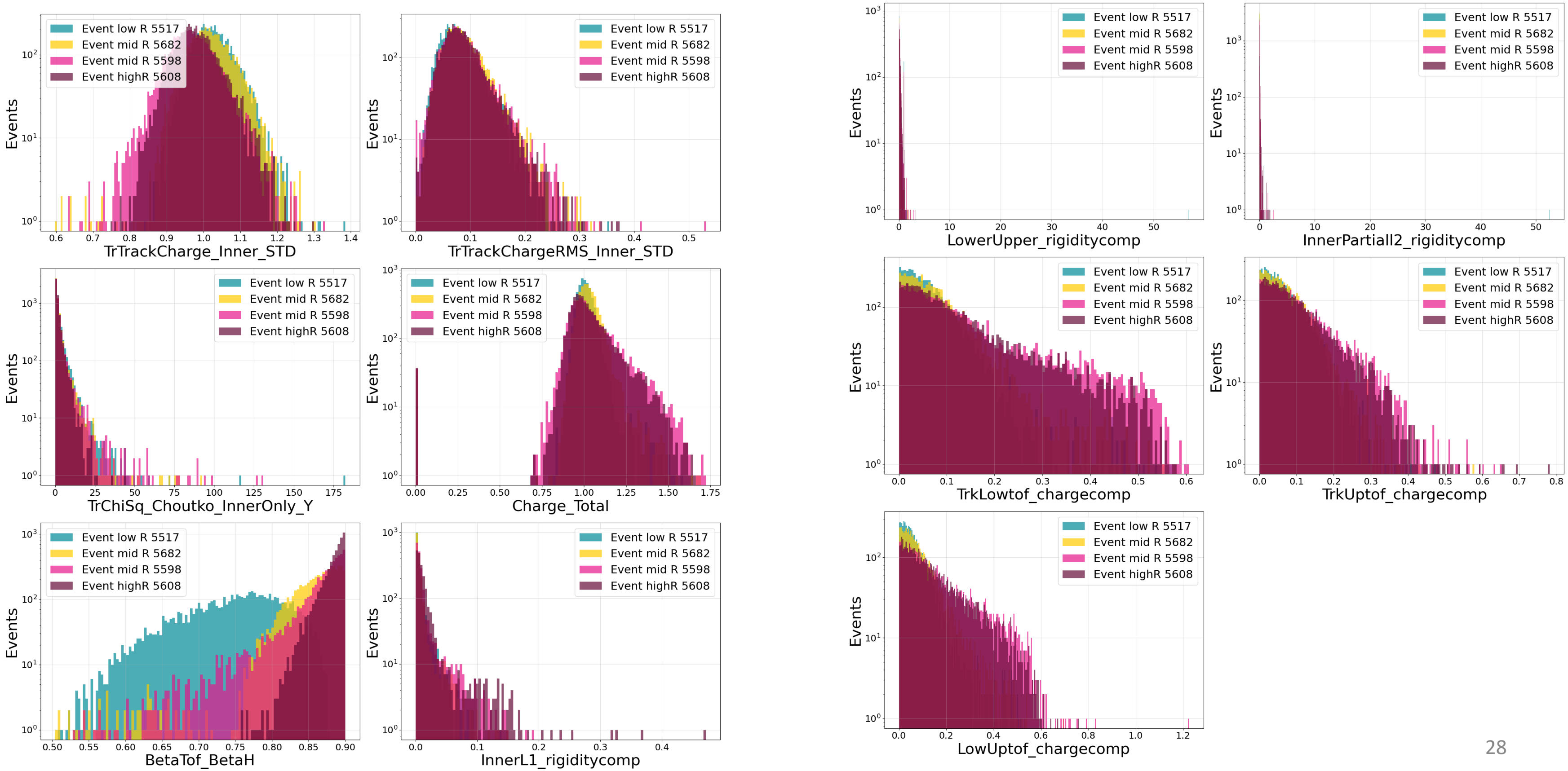
- A ”complete set” of 187 features passing the KS test
- A Physics Driven set composed of 11 variables
- The set of 12 features used by the Bologna group

N.B. Physics Driven features and Bologna features do not pass the KS test with 0.05 threshold

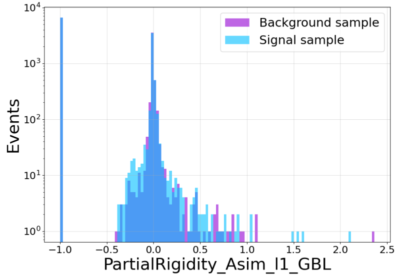
Physics Driven

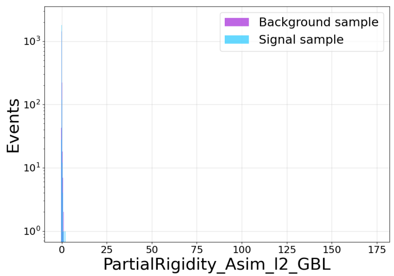


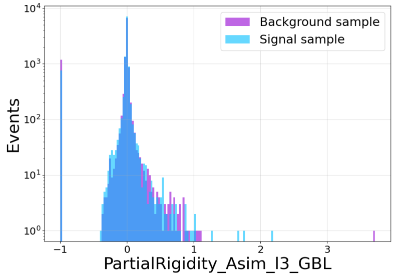
Physics Driven

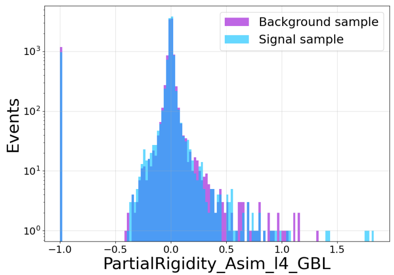


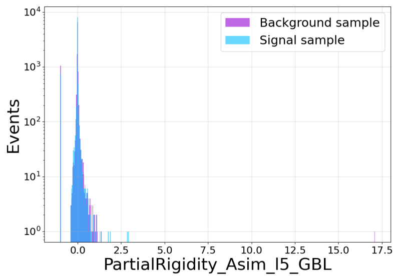
Bologna features

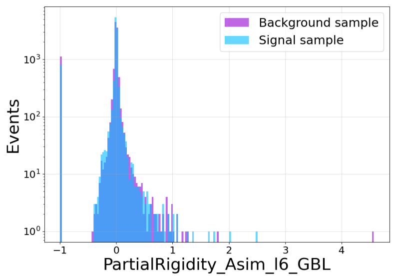


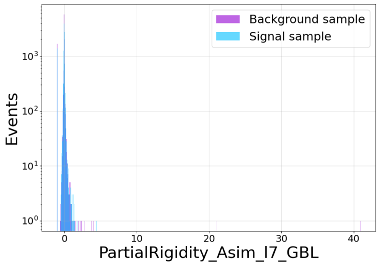


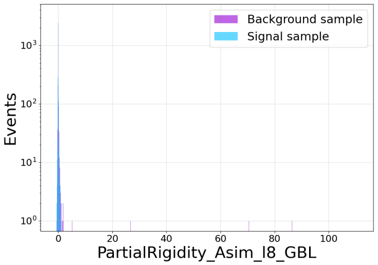


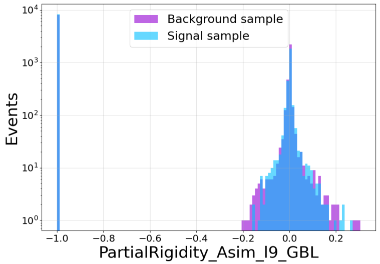


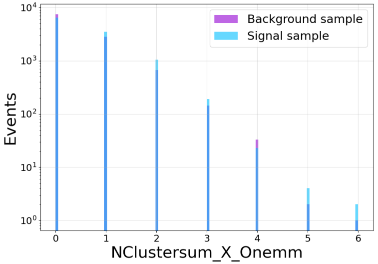


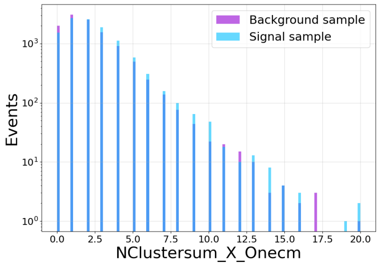


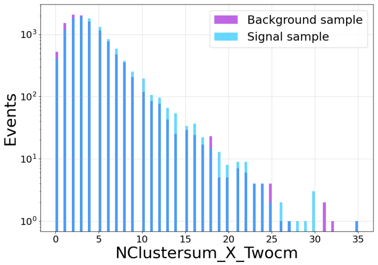


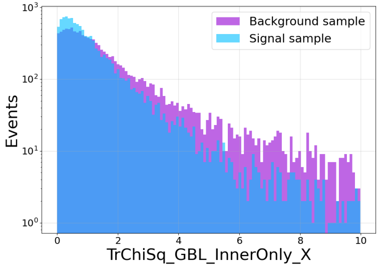




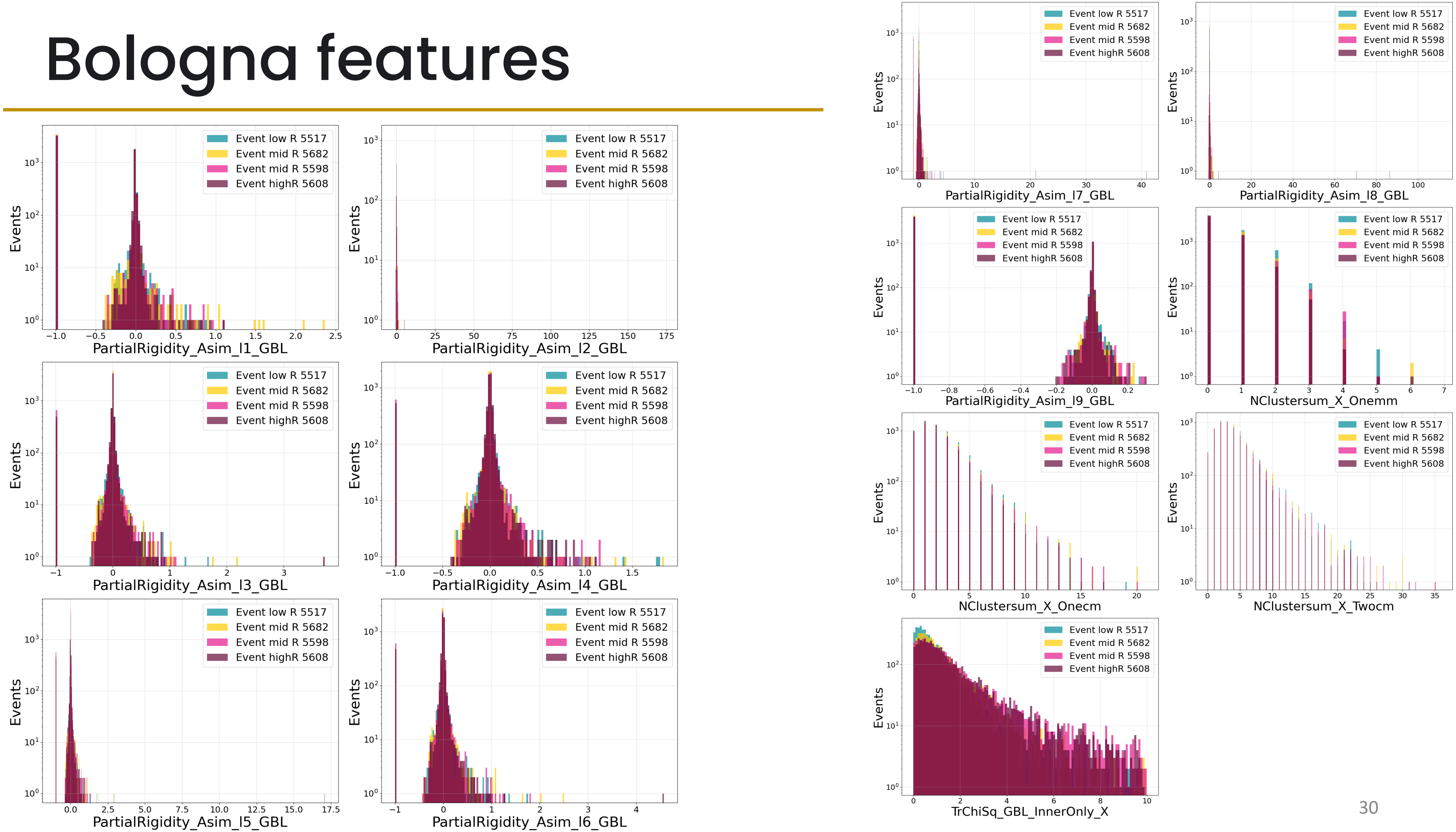








Bologna features

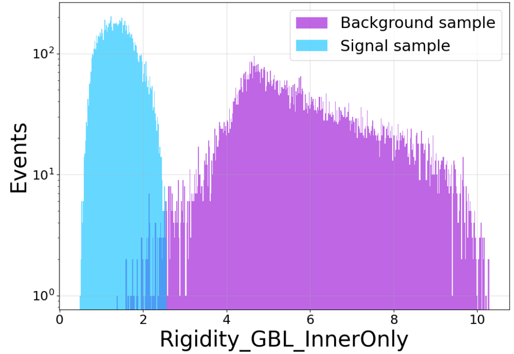


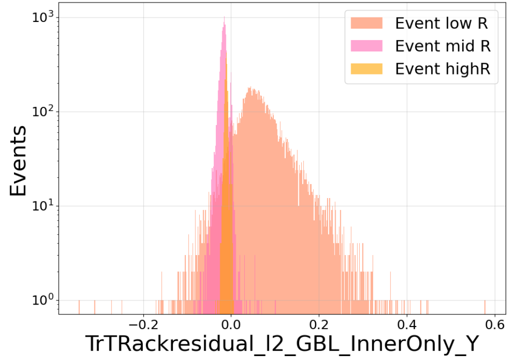
RECAP

The high performances of the BDTs are caused by a data leak due to rigidity and rigidity-dependent variables that bias their training.

2 parallel options:

A. Find a quick way to scan all the variables used (Tracker + ToF) to find the ones non-rigidity-dependent. (plotting all of them is not feasible)
B. Release the cut on beta_tof and check if this attenuates the relation between samples definition and rigidity





Spearman correlation

- Key assumptions:
  * Variables are at least ordinal (can be ranked)
  * Relationship is monotonic but not necessarily linear
  * No assumption about the distribution of variables

- Interpretation:
  * Range: -1 to +1
  * 0 indicates no monotonic relationship

We exclude features for which we have less than a 5% probability of obtaining the value measured if we assume that the distribution come from the same pop/are independent

Kruskal Wallis

- Key assumptions:
  * Ordinal or continuous data
  * Independent observations
  * No normal distribution assumption

- Interpretation:
  * Tests if samples come from the same distribution

Kolmogorov Smirnov

- Key assumptions:

  * Continuous data
  * No distributional assumptions (non-parametric)
  * Independent observations

- Interpretation:

  * Tests if two samples come from the same distribution
  * Sensitive to differences in both shape and location
  * More sensitive to differences in the center and less at the tails of distribution

- Test statistic:

  * Maximum distance between cumulative distribution functions
  * Range: 0 to 1. Larger values indicate greater differences

P-value meaning

We exclude features with p-value < 0.05

p-value: probability to measure the value of the statistic obtained with the chosen test, given the null hypothesis

Spearman → probability to measure that value of the statistic if sampling data from independent distributions (feature and rigidity)

Kw and ks → probability to measure that value of the statistic if the given distributions are coming from the same population

We reject features for which there is less than 5% probability to measure the computed statistic if the distribution were independent/coming from the same pop

P-value

p-value: probability to measure the value of the statistic obtained with the chosen test, given the null hypothesis

We compute the p-value for each of the 3 statistical measures
We exclude features with p-value < 0.05.

→ We reject features for which there is less than 5% probability to measure the value of the statistic obtained with the chosen test given the null hypothesis

Statistical tests - results

Number of features passing the tests:

- Spearman correlation: 710
- Krukal-Wallis: 670
- Kolmogorov Smirnov test: 265

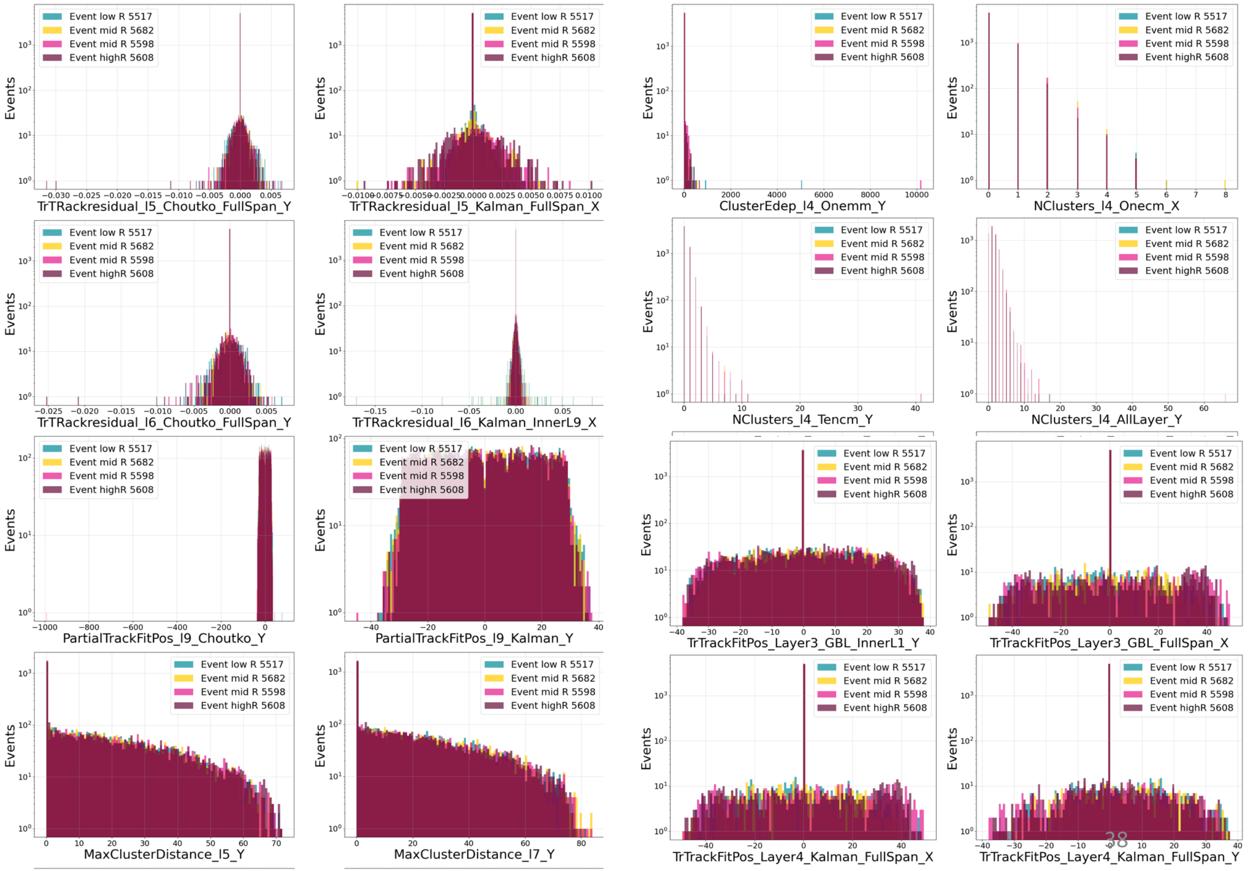
→ Choose the Kolmogorov Smirnov (KS) as it is the more conservative

→ Plot feature selected by KS as a cross-check: they seem to be rigidity-independent

KS features distributions

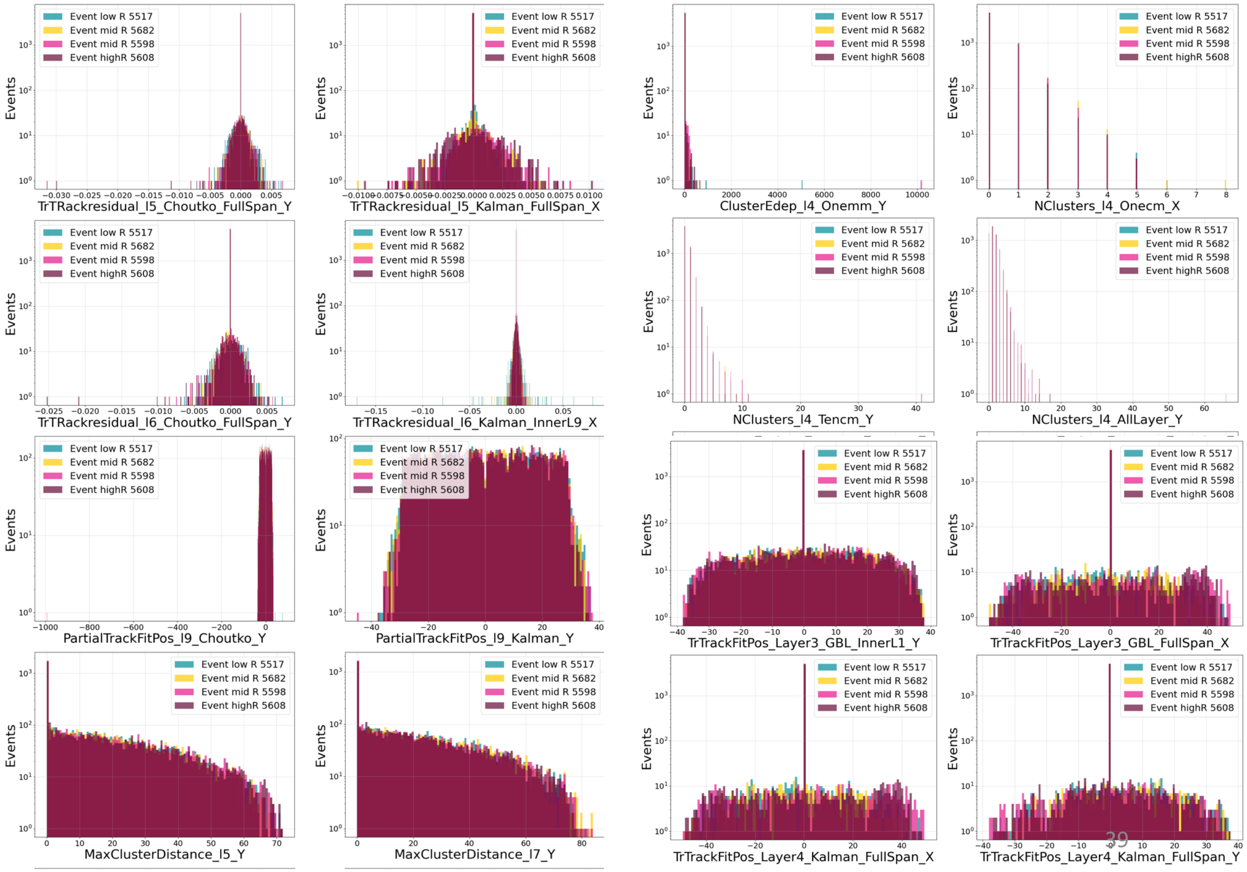
Some of the features selected by KS

The distributions for the different rigidity ranges are plotted in different colors.



KS features distributions

Some features have a peak at zero due most probably to non-controlled background events for fits and spans different from the one use din the mass definition (GBL, InnerOnly): investigate if this signature bias the test
→ Cut events at 0 and rerun the tests



Statistical tests – results no events at 0

Number of features passing the tests:

- Spearman correlation: 679
- Krukal-Wallis: 661
- Kolmogorov Smirnov test: 187

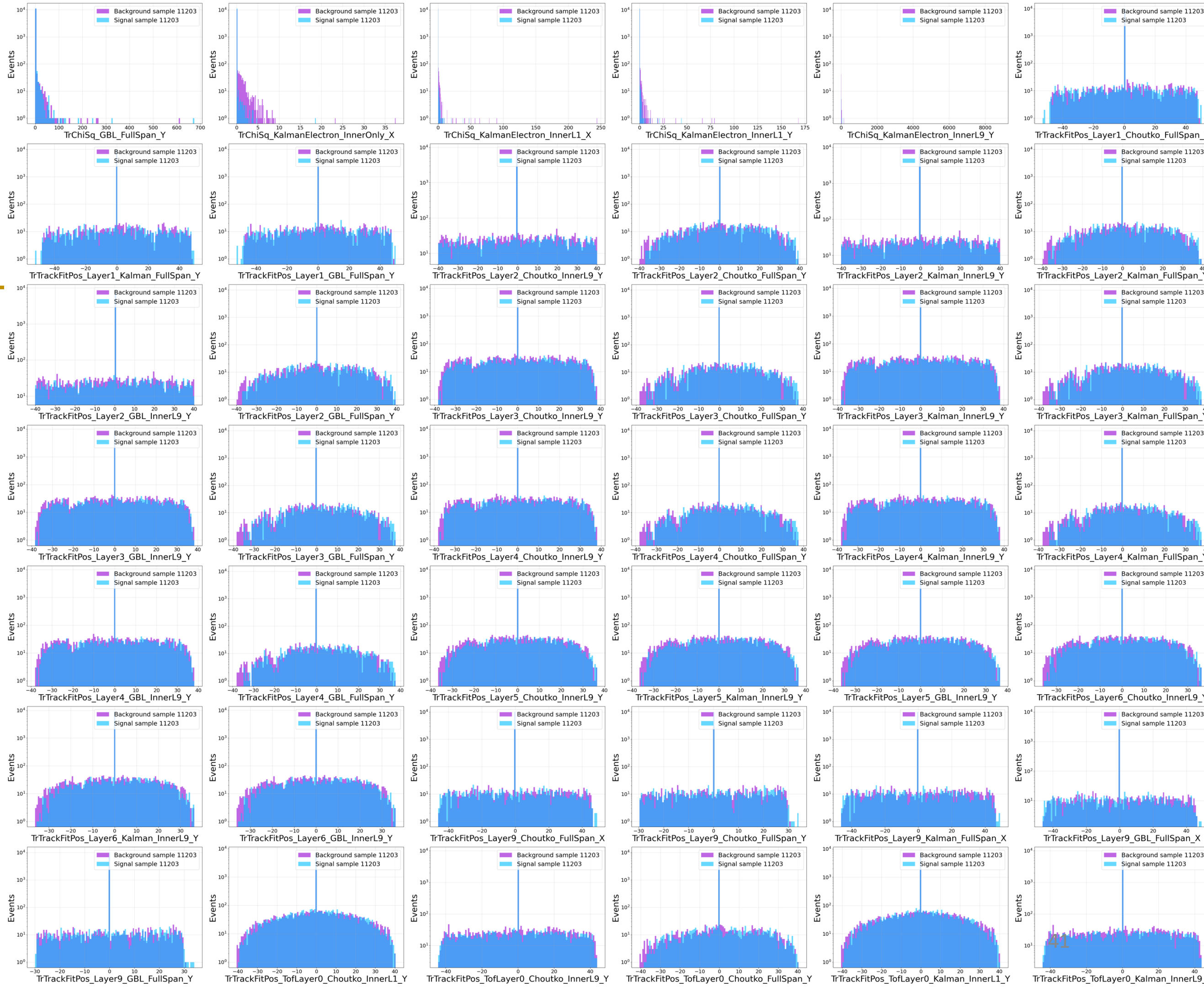
- Kolmogorov Smirnov (KS) is still the more conservative
- The features selected this time are fewer and, in part, different from the previous ones
- They seem to be independent from rigidity after visual inspection
- It is already clear that they are not going to discriminate well between background and signal → see plots

KS features

The features selected by KS

Bkg events in violet

Signal events in blue

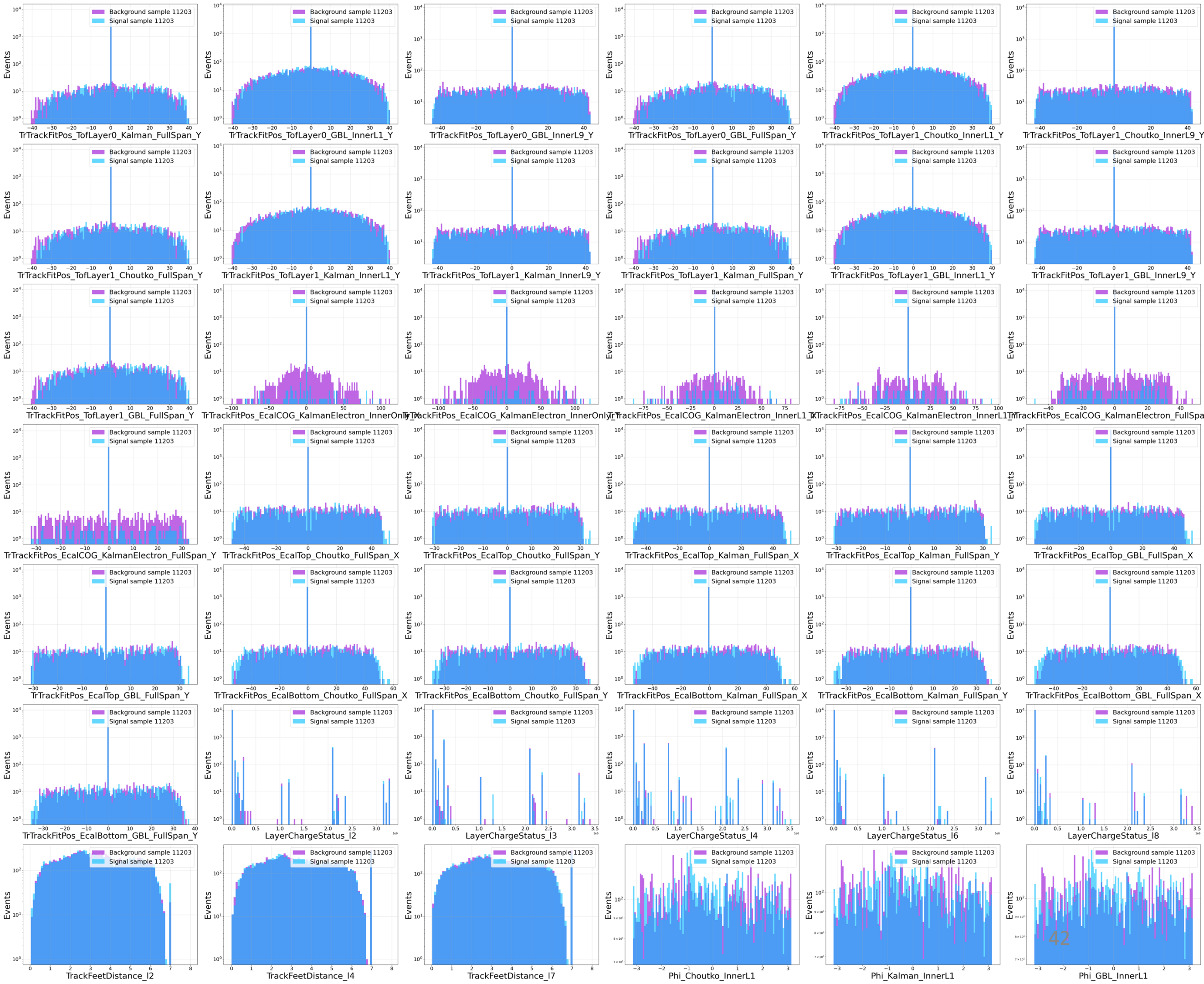


KS features

The features selected by KS

Bkg events in violet

Signal events in blue

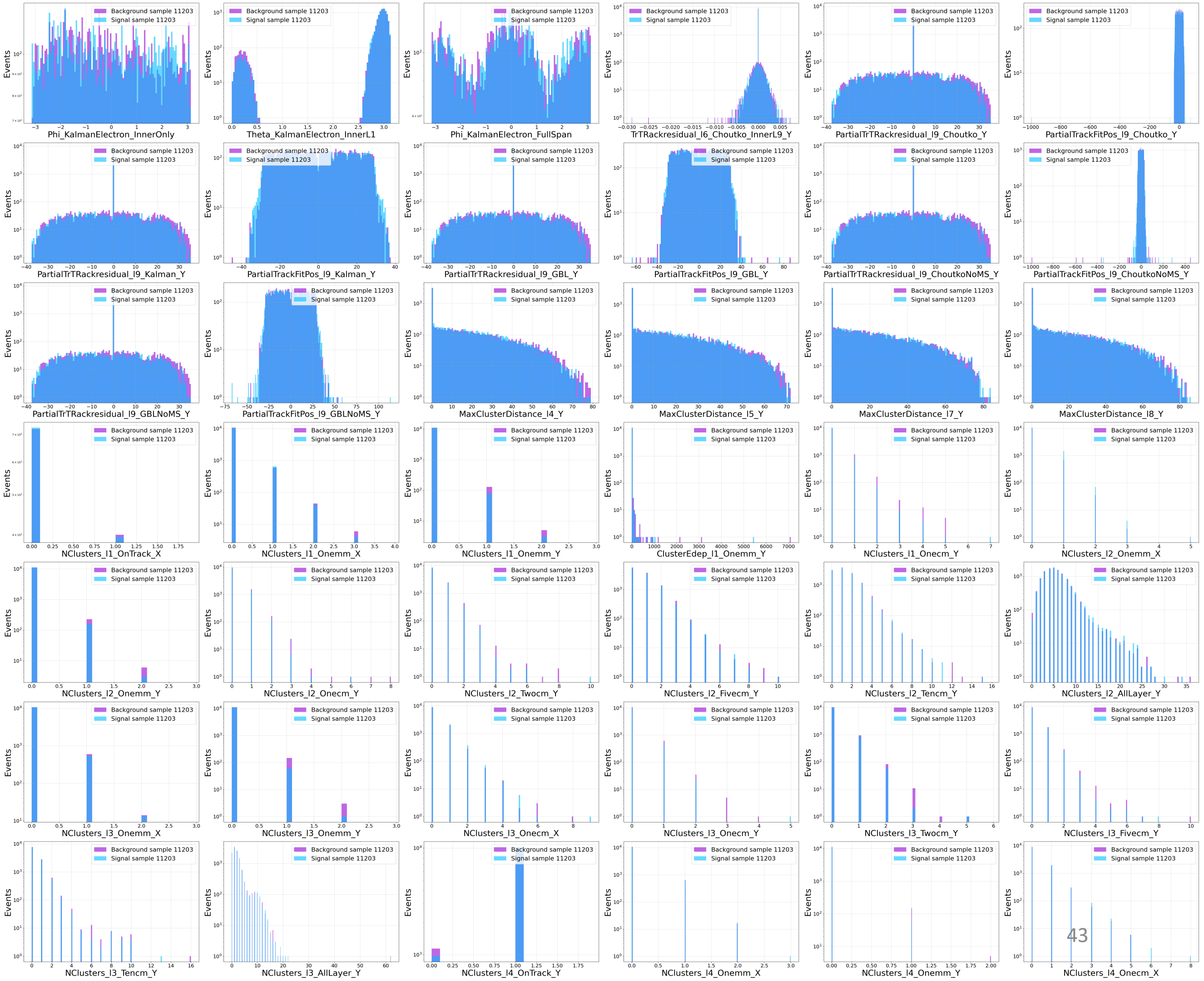


KS features

The features selected by KS

Bkg events in violet

Signal events in blue

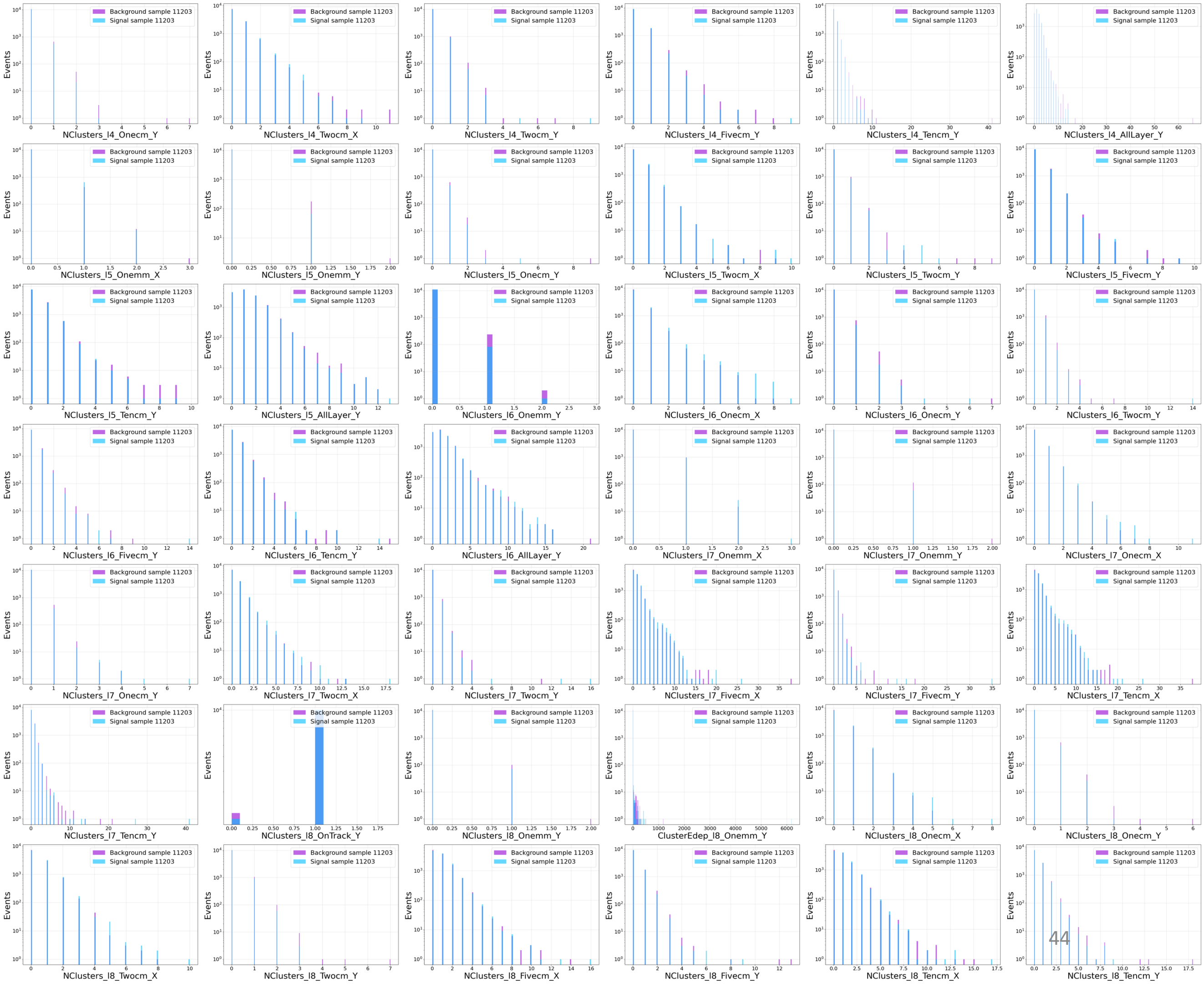


KS features

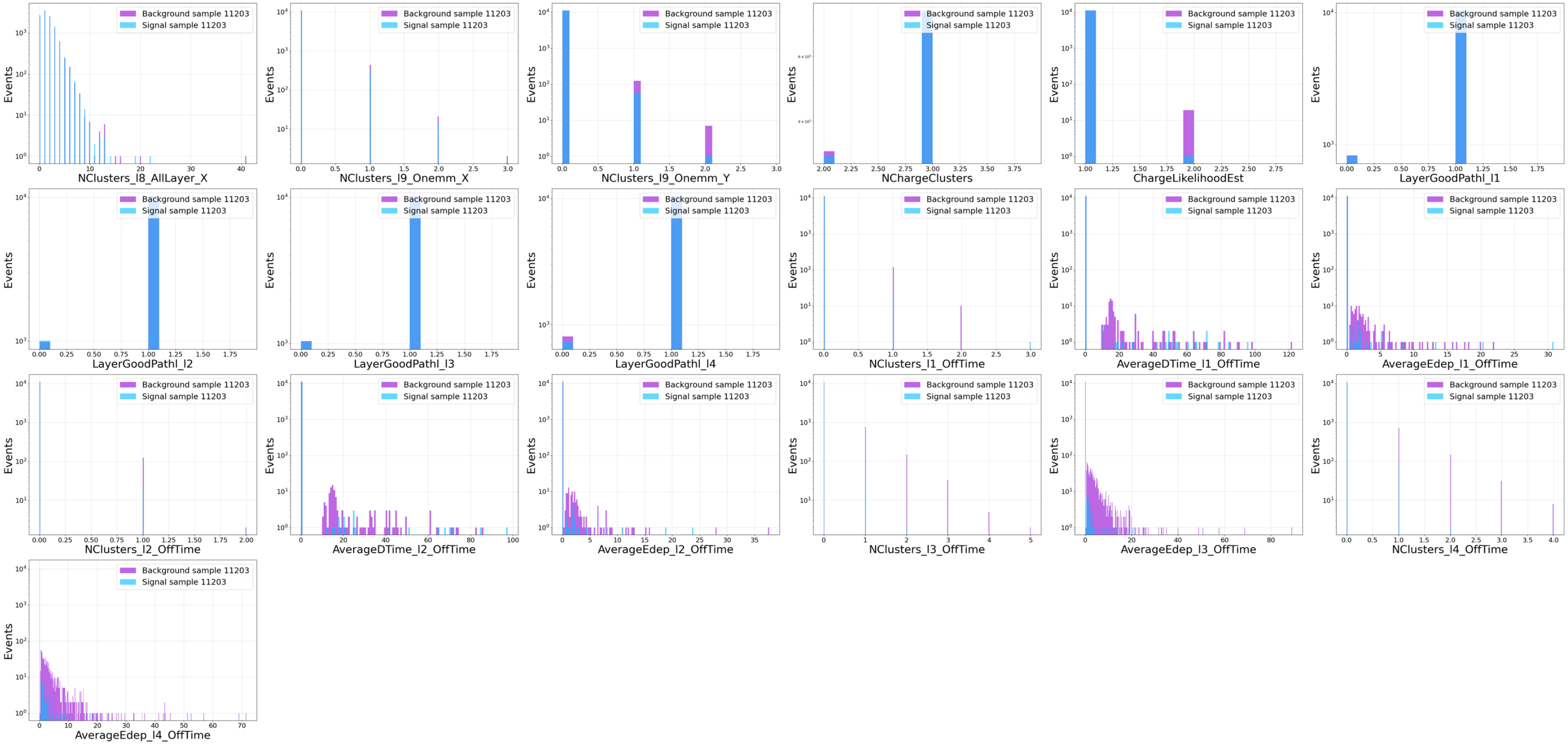
The features selected by KS

Bkg events in violet

Signal events in blue



KS features



Background sample 11203
Signal sample 11203
Events
NClusters_l8_AllLayer_X
NClusters_l9_Onemm_X
NClusters_l9_Onemm_Y
NChargeClusters
ChargeLikelihoodEst
LayerGoodPathl_l1
LayerGoodPathl_l2
LayerGoodPathl_l3
LayerGoodPathl_l4
NClusters_l1_OffTime
AverageDTime_l1_OffTime
AverageEdep_l1_OffTime
NClusters_l2_OffTime
AverageDTime_l2_OffTime
AverageEdep_l2_OffTime
NClusters_l3_OffTime
AverageEdep_l3_OffTime
NClusters_l4_OffTime
AverageEdep_l4_OffTime

Cross-validation metrics (training phase)

- Accuracy: this metric provides a general measure of the model's ability to correctly predict classes.
  Accuracy = (True Positives + True Negatives) / Total Examples

- Precision:. This metric focuses on the quality of the model's positive predictions, to avoid the erroneous classification of negative examples as positive
  Precision = True Positives / (True Positives + False Positives)

- Recall (signal efficiency): focuses on the model's ability to identify positive cases while effectively minimising false negatives.
  Recall = True Positives / (True Positives + False Negatives)

- F1-score: this metric combines the precision and recall metrics to provide a balanced measure of model performance. It is particularly relevant when the balance between accurately identifying positive cases and minimising false positives and false negatives is essential.
  F1-score = 2 X (Precision × Recall)/Precision + Recall

Performance metrics

Table 1. Training metrics XGBoost

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.6173 ± 0.0144	0.5913 ± 0.0188	0.541 ± 0.0146	0.6085 ± 0.0242	0.6202 ± 0.0178	0.8682 ± 0.0148	0.9934 ± 0.0028
Precision	0.6083 ± 0.0164	0.5848 ± 0.0184	0.5229 ± 0.0078	0.5964 ± 0.0216	0.6095 ± 0.0172	0.8822 ± 0.0182	0.9924 ± 0.0036
Recall	0.6718 ± 0.0224	0.6457 ± 0.0194	0.9935 ± 0.0088	0.685 ± 0.0332	0.6808 ± 0.0338	0.8518 ± 0.0286	0.9946 ± 0.005
F1-Score	0.6383 ± 0.0118	0.6137 ± 0.016	0.6852 ± 0.0082	0.6375 ± 0.023	0.6431 ± 0.019	0.8666 ± 0.0162	0.9935 ± 0.0028
ROC-AUC	0.6726 ± 0.0172	0.6355 ± 0.0212	0.5385 ± 0.0148	0.6609 ± 0.0222	0.6766 ± 0.0186	0.9412 ± 0.0126	0.9992 ± 0.001

Table 2. Validation metrics XGBoost None

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.624	0.593	0.536	0.621	0.624	0.869	0.993
Precision	0.626	0.595	0.727	0.625	0.627	0.869	0.993
Recall	0.624	0.593	0.536	0.621	0.624	0.869	0.993
F1-Score	0.622	0.592	0.418	0.618	0.623	0.869	0.993
ROC-AUC	0.625	0.594	0.542	0.622	0.625	0.869	0.993

Performance metrics

Table 3. Training metrics AdaBoost

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.6274 ± 0.0218	0.5998 ± 0.0284	0.541 ± 0.0146	0.6276 ± 0.0184	0.6318 ± 0.0256	0.8543 ± 0.0186	0.9891 ± 0.0028
Precision	0.6039 ± 0.0216	0.5824 ± 0.0242	0.5229 ± 0.0078	0.6007 ± 0.0164	0.6074 ± 0.0228	0.8646 ± 0.0202	0.9912 ± 0.0044
Recall	0.7526 ± 0.0296	0.7222 ± 0.0262	0.9935 ± 0.0088	0.7743 ± 0.0256	0.7581 ± 0.0316	0.8422 ± 0.0218	0.9871 ± 0.005
F1-Score	0.6696 ± 0.017	0.6448 ± 0.023	0.6852 ± 0.0082	0.6764 ± 0.0146	0.6743 ± 0.0212	0.8532 ± 0.019	0.9891 ± 0.0028
ROC-AUC	0.6836 ± 0.024	0.6444 ± 0.0344	0.5385 ± 0.0148	0.6815 ± 0.0124	0.6853 ± 0.0256	0.9287 ± 0.013	0.9982 ± 0.0016

Table 4. Validation metrics AdaBoost None

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.641	0.596	0.536	0.632	0.638	0.856	0.989
Precision	0.654	0.604	0.727	0.646	0.650	0.857	0.989
Recall	0.641	0.596	0.536	0.632	0.638	0.856	0.989
F1-Score	0.635	0.590	0.418	0.624	0.631	0.856	0.989
ROC-AUC	0.643	0.598	0.542	0.633	0.639	0.856	0.989

Performance metrics

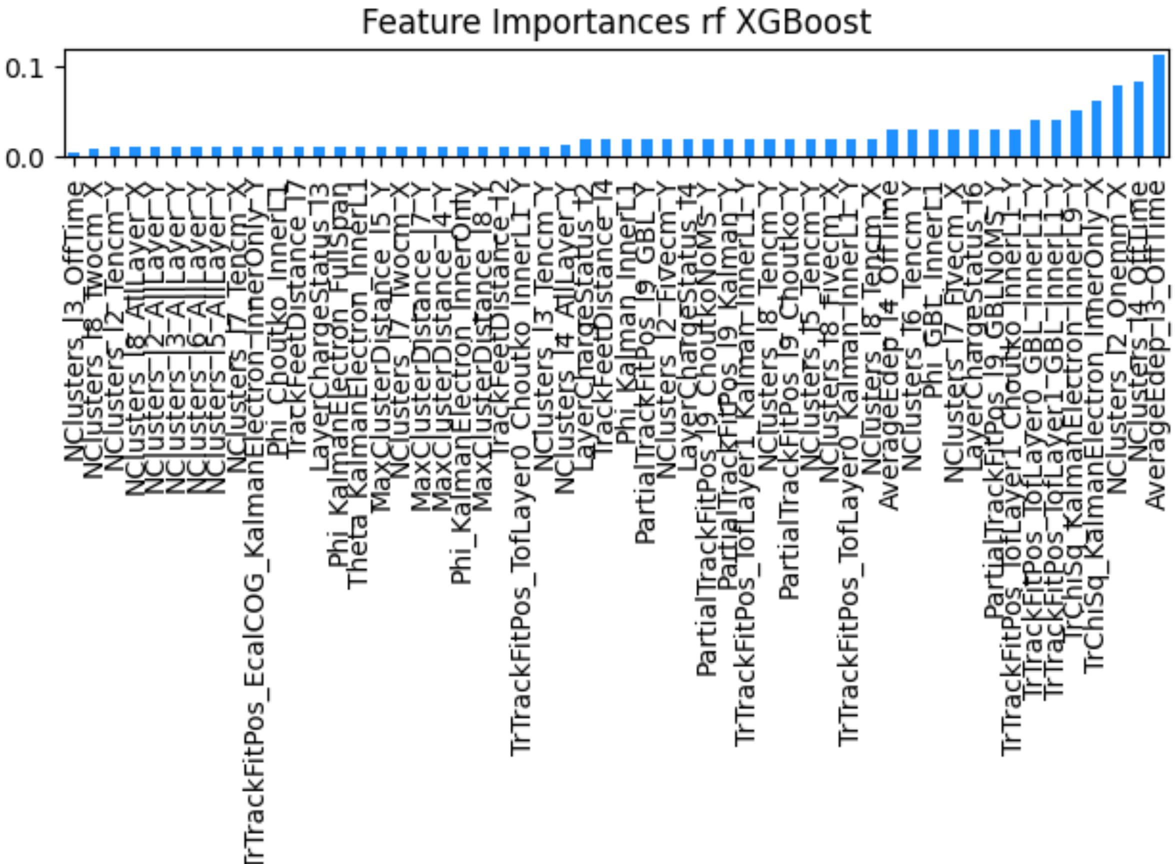
Table 5. Validation metrics XGBoost 2023

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.526	0.543	0.535	0.569	0.502	0.500	0.512
Precision	0.550	0.566	0.713	0.569	0.558	0.750	0.618
Recall	0.526	0.543	0.535	0.569	0.502	0.500	0.512
F1-Score	0.463	0.497	0.412	0.568	0.344	0.334	0.371
ROC-AUC	0.526	0.543	0.535	0.569	0.502	0.500	0.512

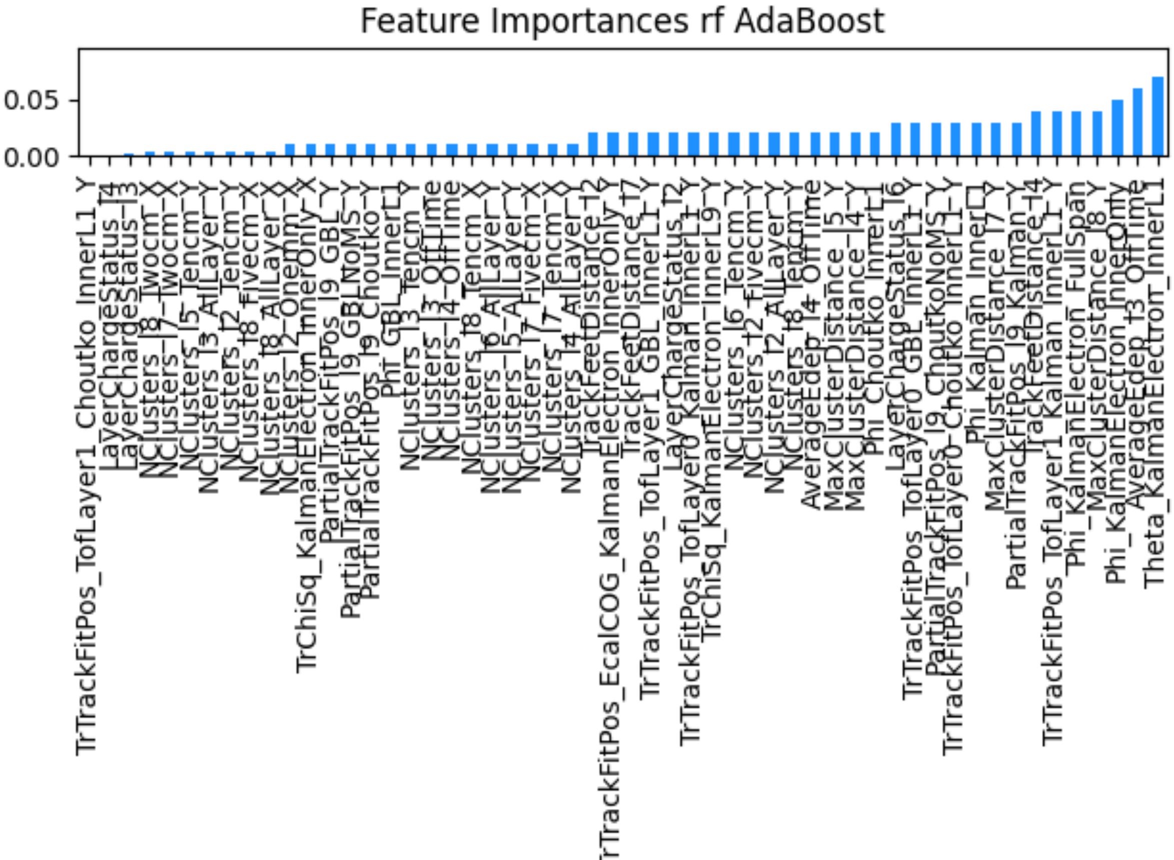
Table 6. Validation metrics AdaBoost 2023

	complete	rf	linear	kbest	correlation	PhysicsDriven	Bologna
Accuracy	0.501	0.515	0.535	0.526	0.529	0.501	0.513
Precision	0.576	0.668	0.713	0.534	0.579	0.667	0.706
Recall	0.501	0.515	0.535	0.526	0.529	0.501	0.513
F1-Score	0.339	0.371	0.412	0.494	0.442	0.335	0.365
ROC-AUC	0.501	0.515	0.535	0.526	0.529	0.501	0.513

Features importance – Random Forest

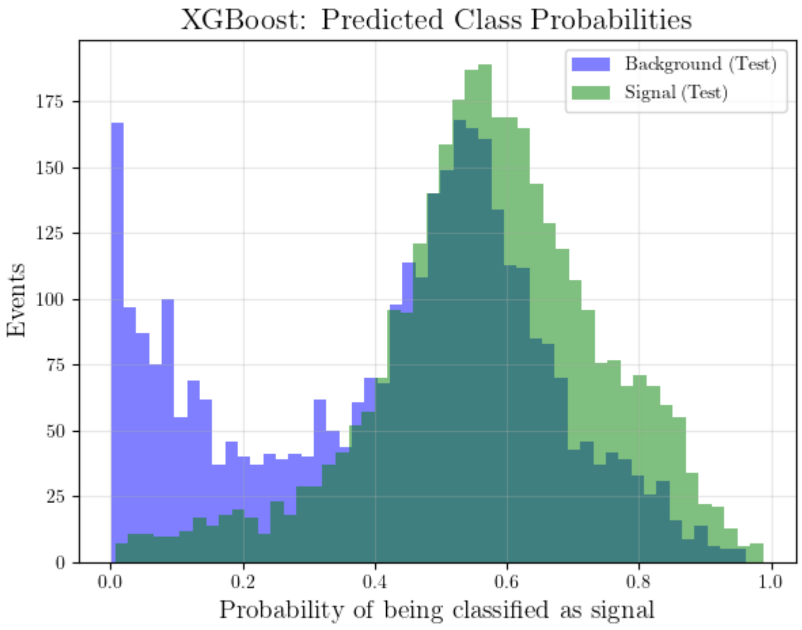


Features importance – Random Forest

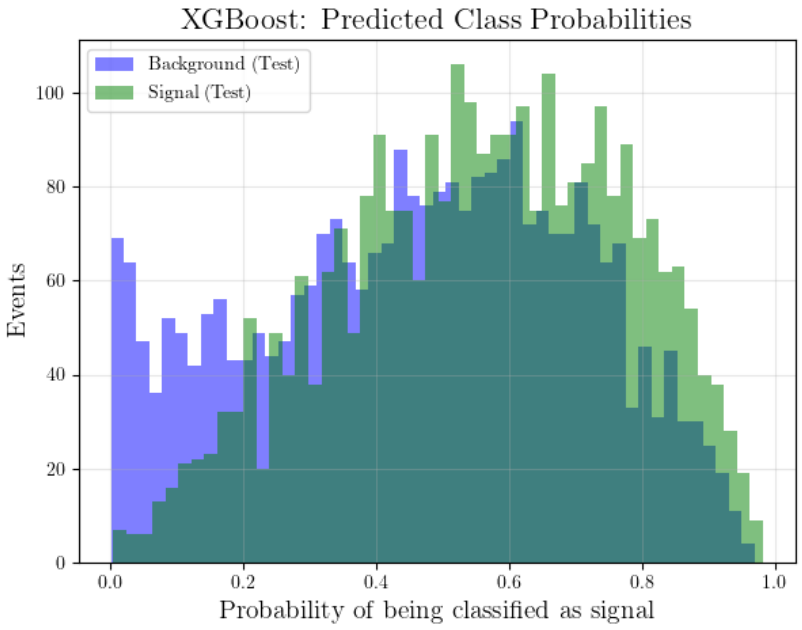


Probability of being classified as signal- kbest

Validation 2015-2018 dataset

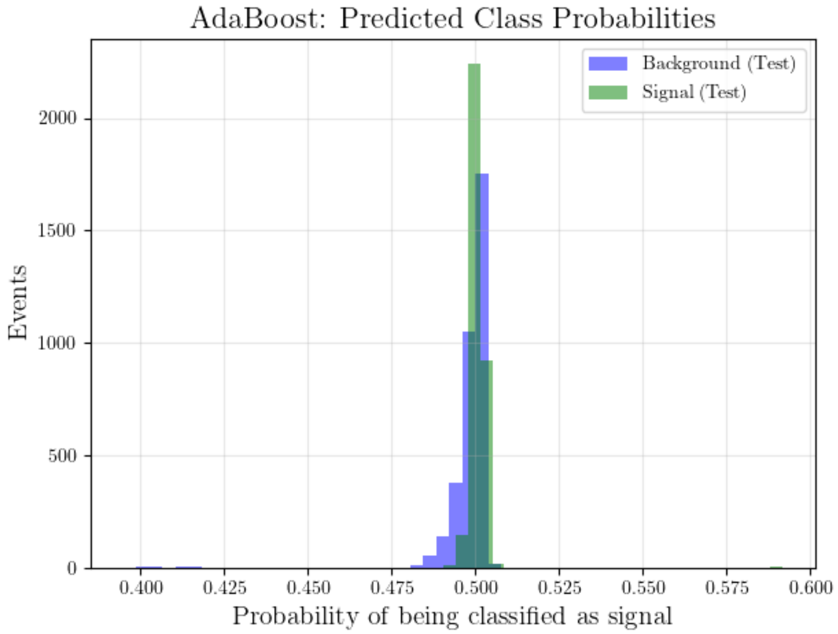


Validation 2023 dataset

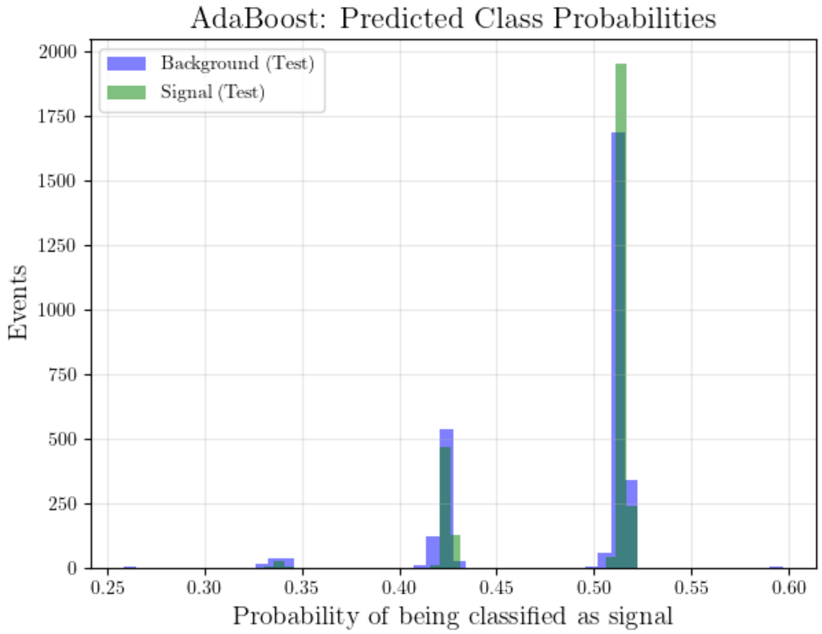


Probability of being classified as signal- kbest

Validation 2015-2018 dataset

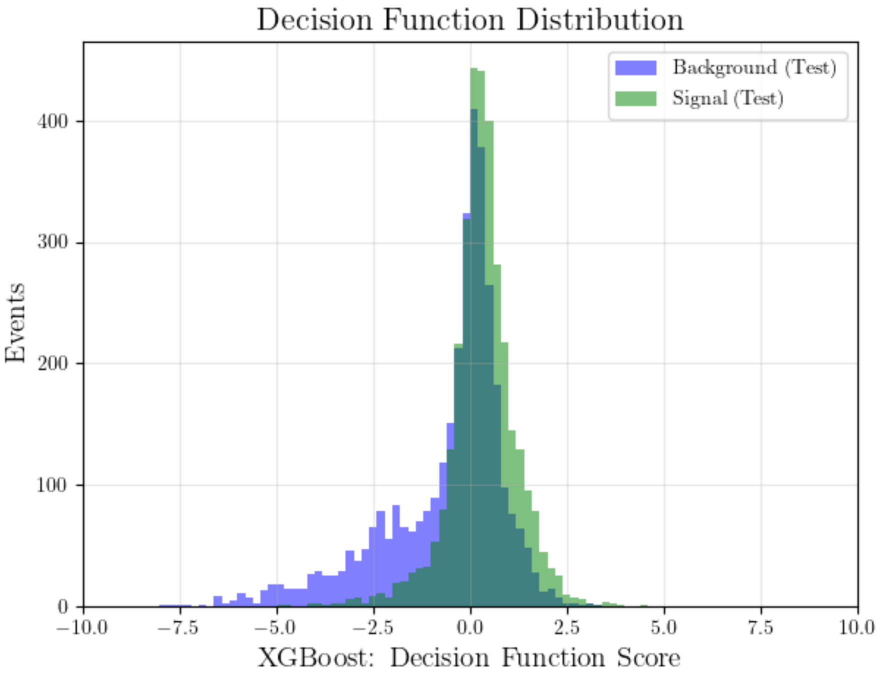


Validation 2023 dataset

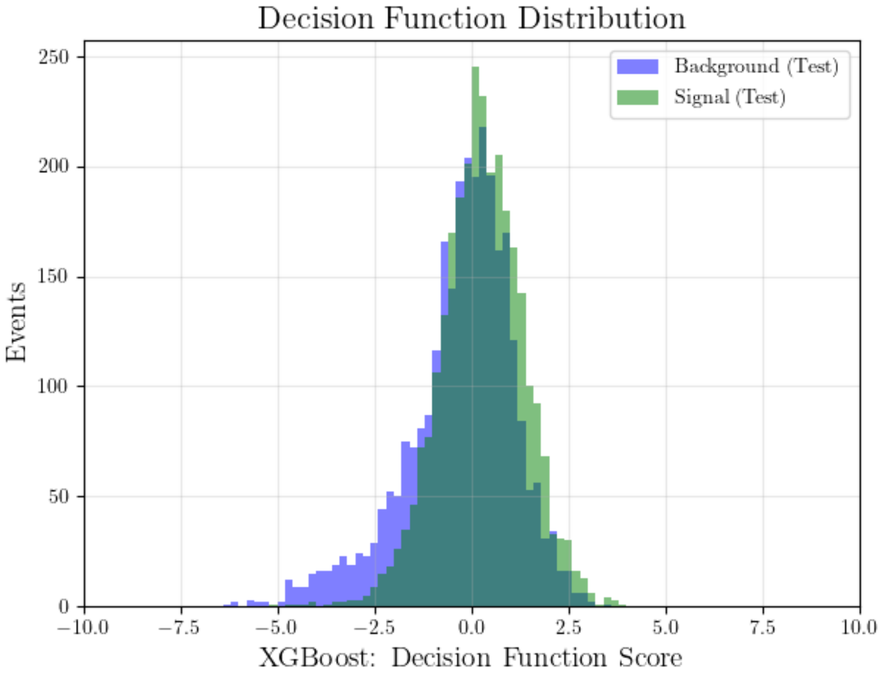


Decision score functions – kbest

Validation 2015-2018 dataset (XGB)

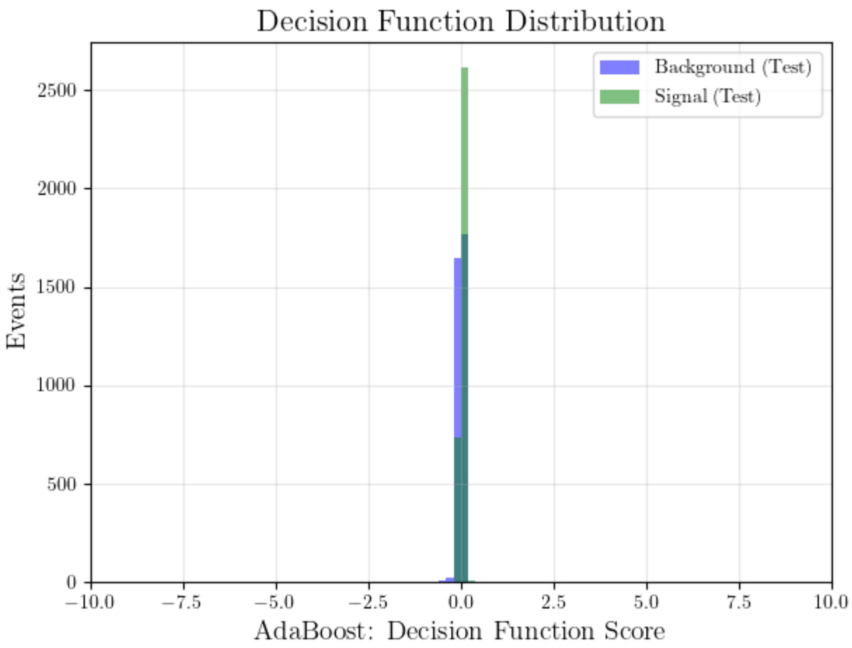


Validation 2023 dataset (XGB)

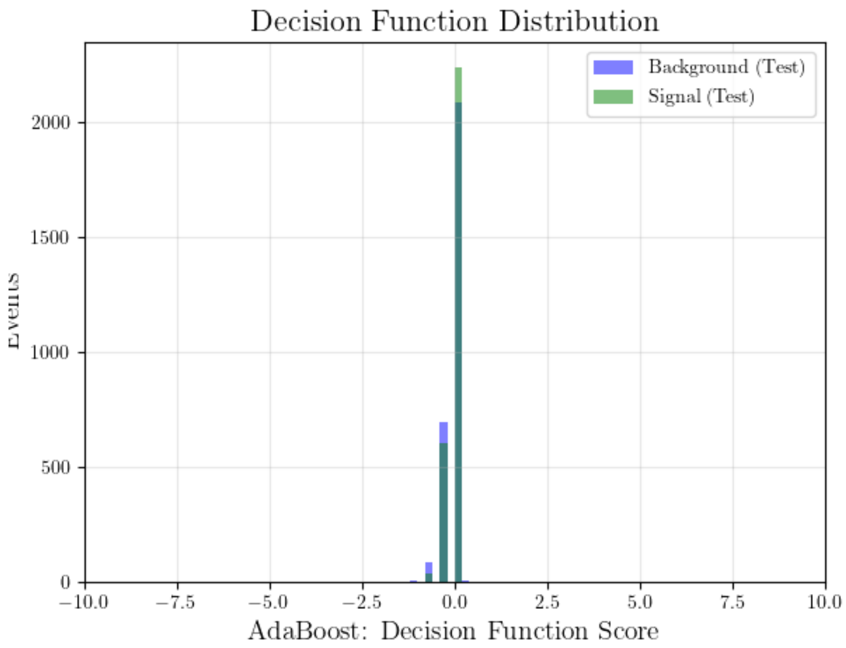


Decision score functions – kbest

Validation 2015-2018 dataset (Ada)

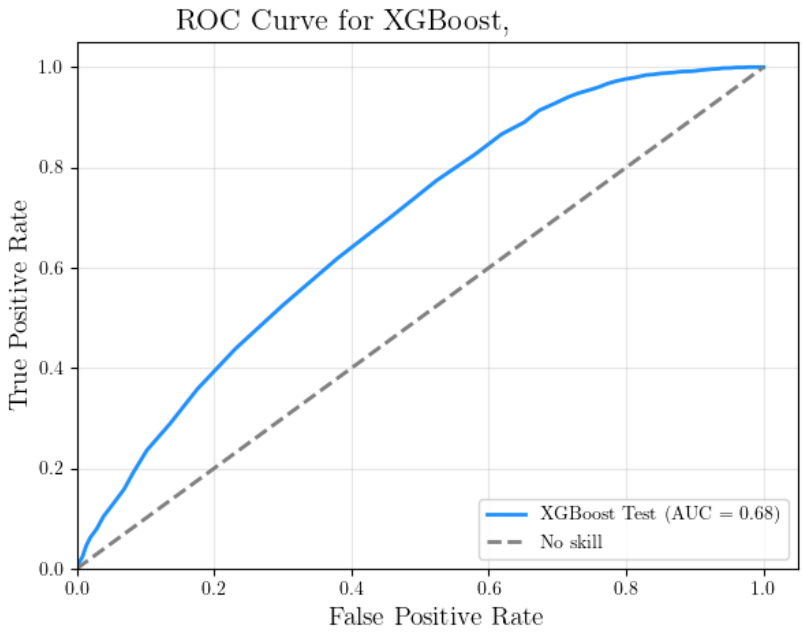


Validation 2023 dataset (Ada)

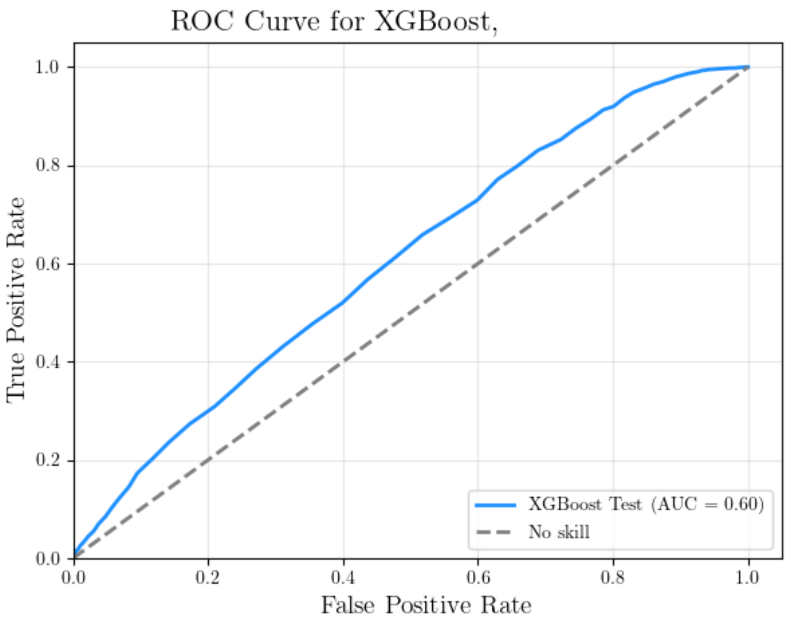


ROC curve – kbest

Validation 2015-2018 dataset

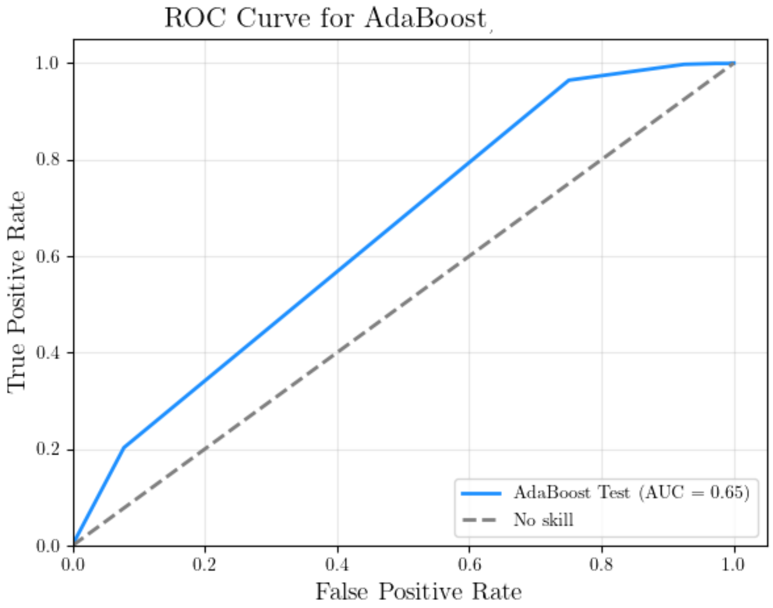


Validation 2023 dataset

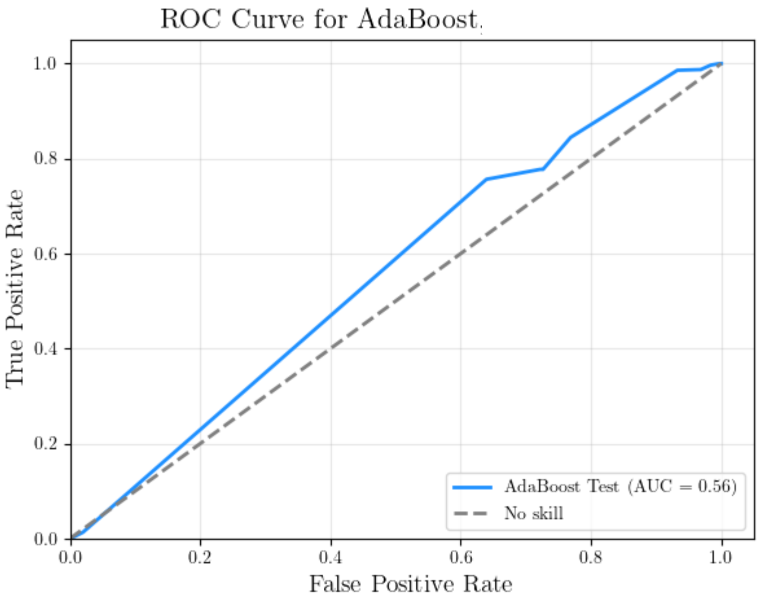


ROC curve – kbest

Validation 2015-2018 dataset

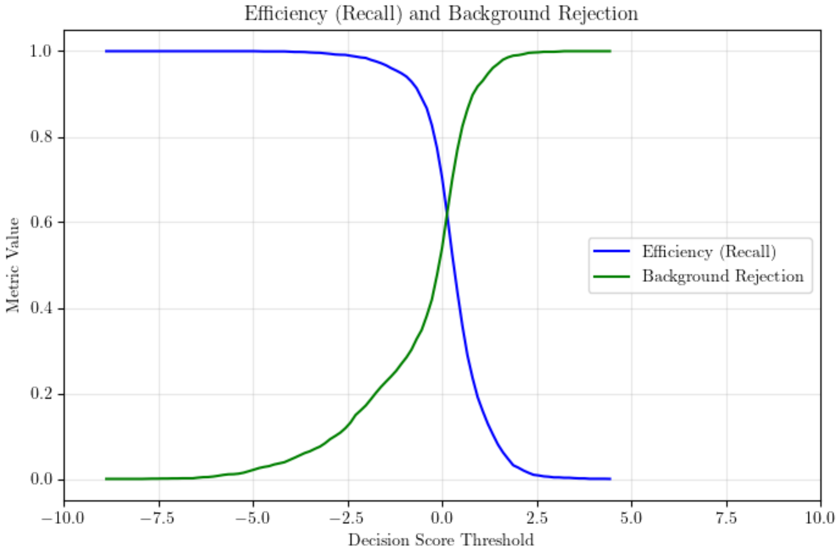


Validation 2023 dataset

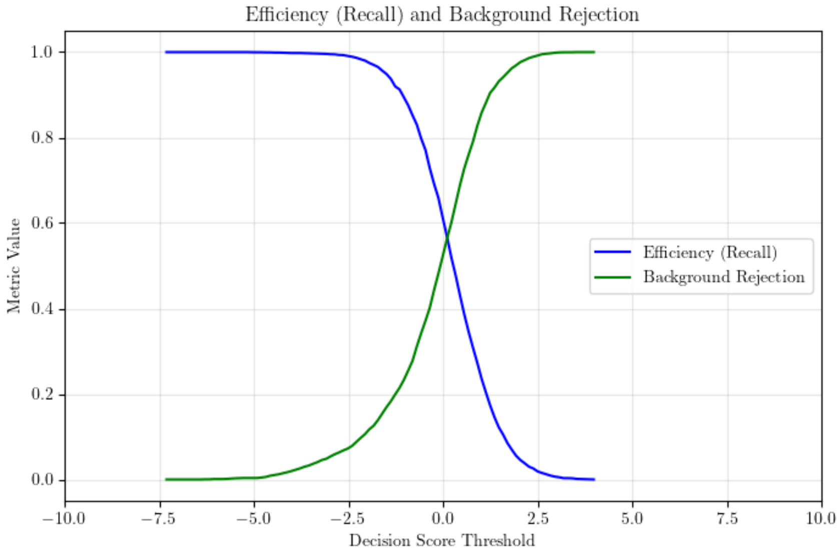


Efficiencies – kbest

Validation 2015-2018 dataset (XGB)

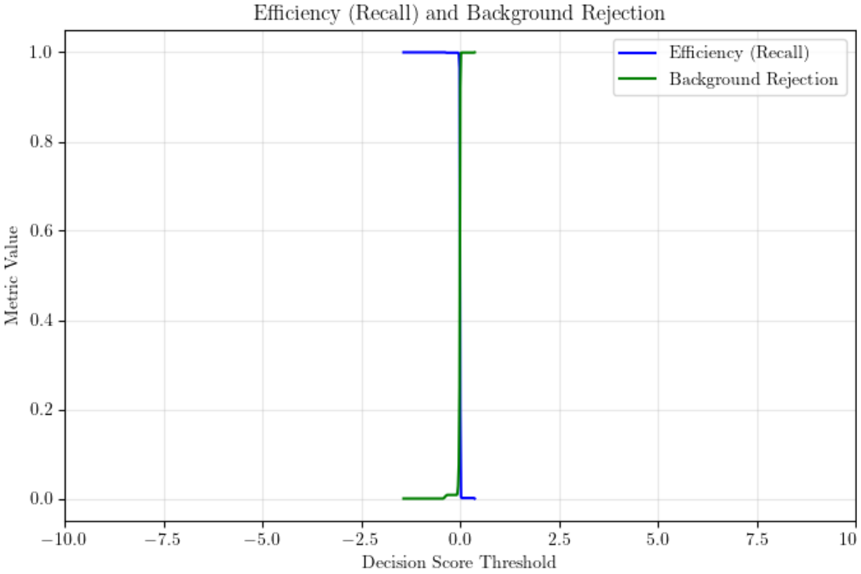


Validation 2023 dataset (XGB)

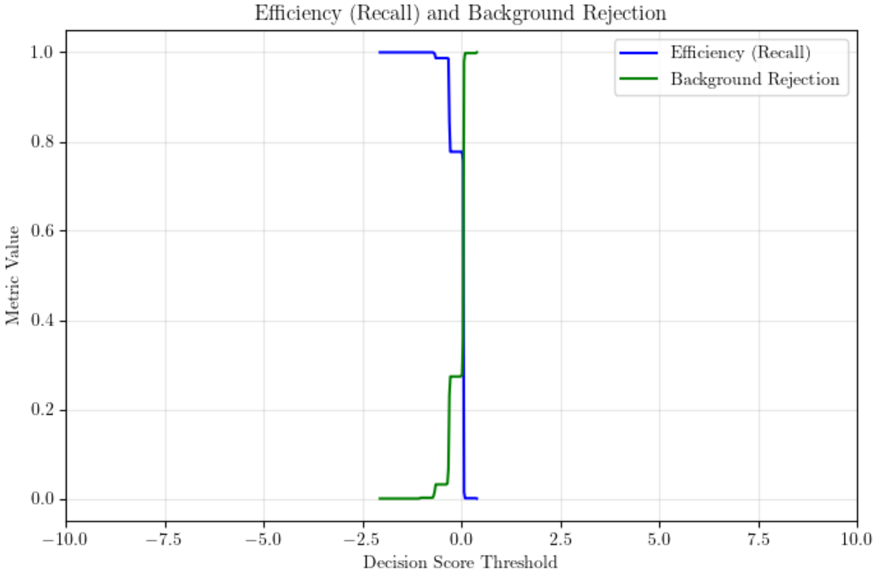


Efficiencies – kbest

Validation 2015-2018 dataset (Ada)

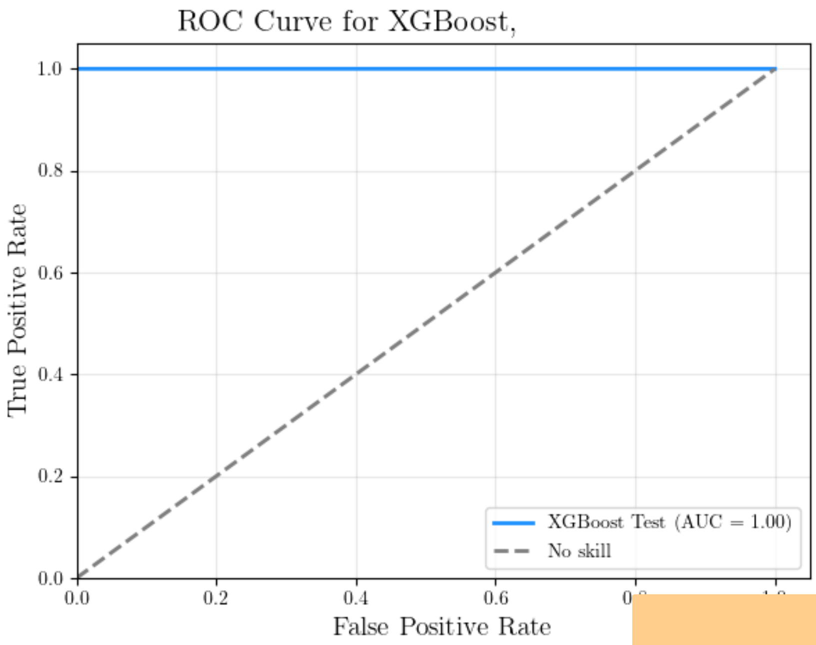


Validation 2023 dataset (Ada)

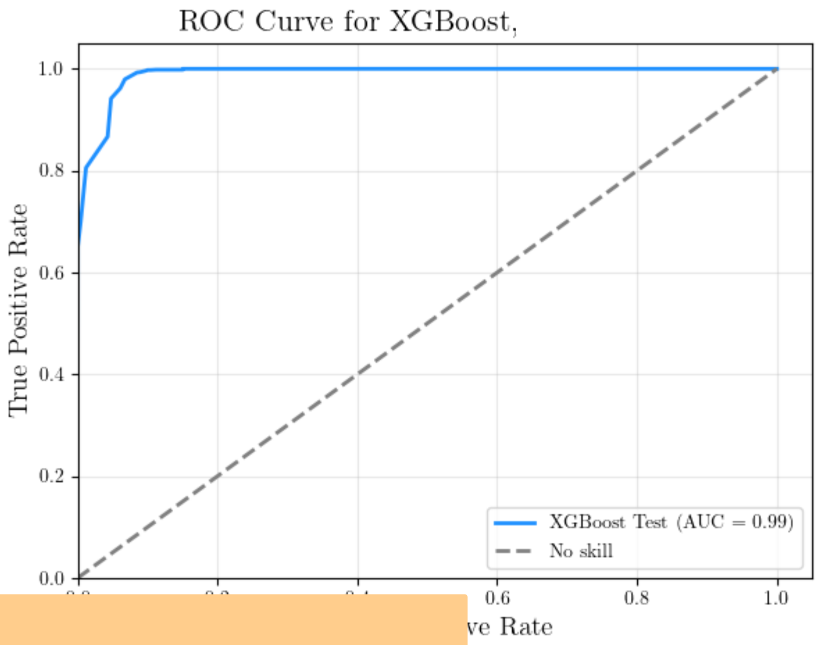


ROC curve – Random Forest

Validation 2015-2018 dataset



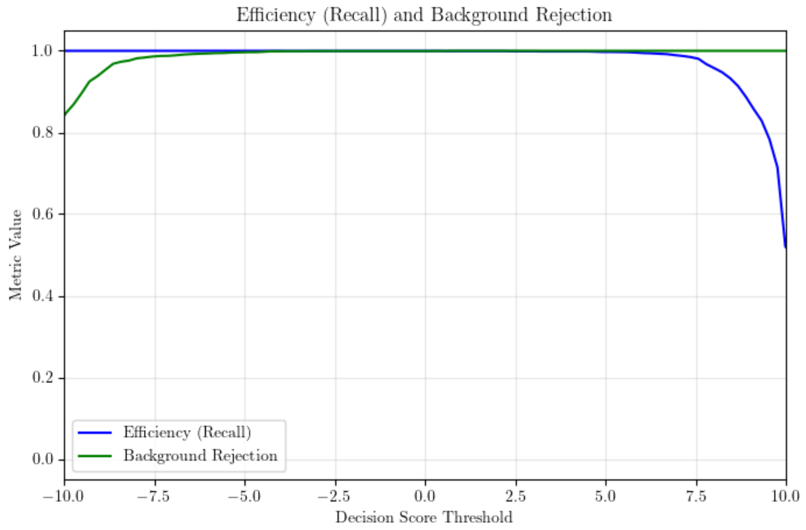
Validation 2023 dataset



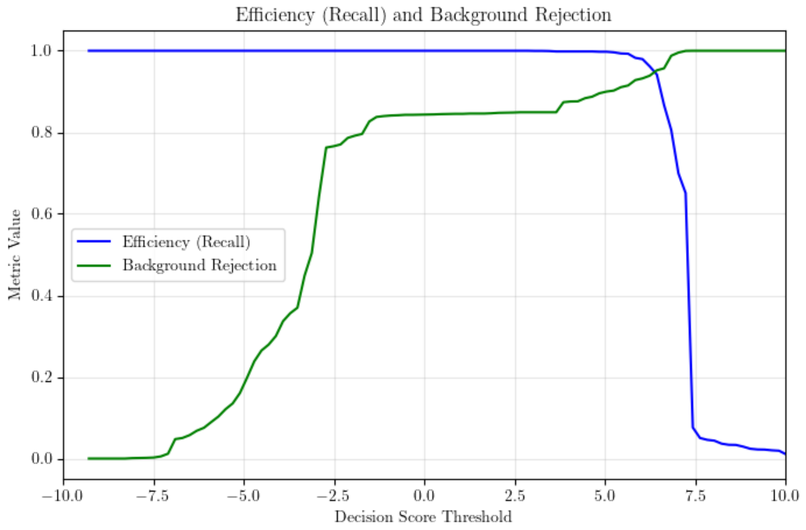
All features

Efficiencies – Random Forest

Validation 2015-2018 dataset



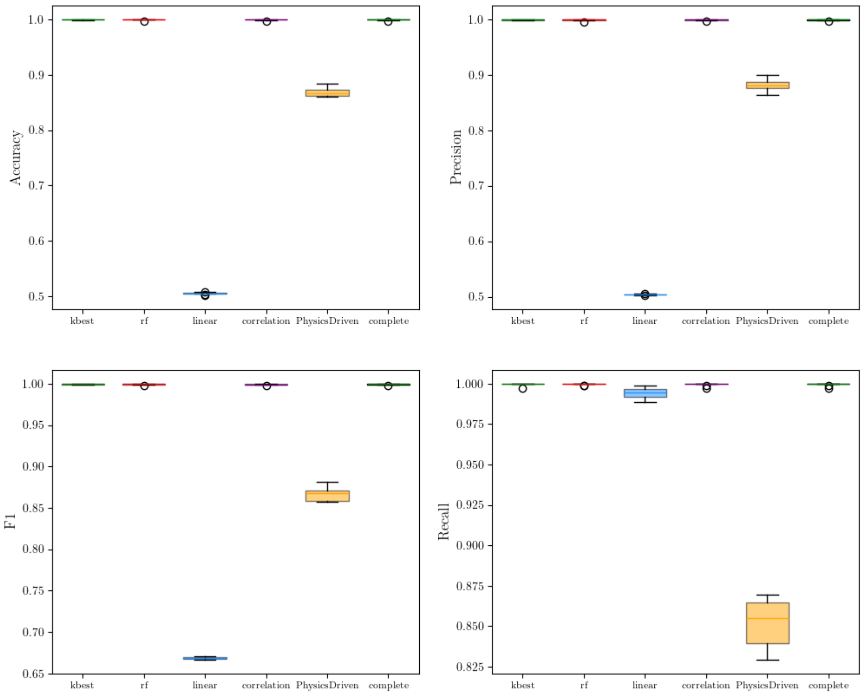
Validation 2023 dataset



All features

Boxplots (training metrics)

XGBoost



AdaBoost achieves similar performances (as seen in the previous slides), but the feature importances are very different from the two algorithms

All features

Features distributions

Rigidities

Plotted rigidities distribution for the two samples for different fits (span = Inner Tracker)

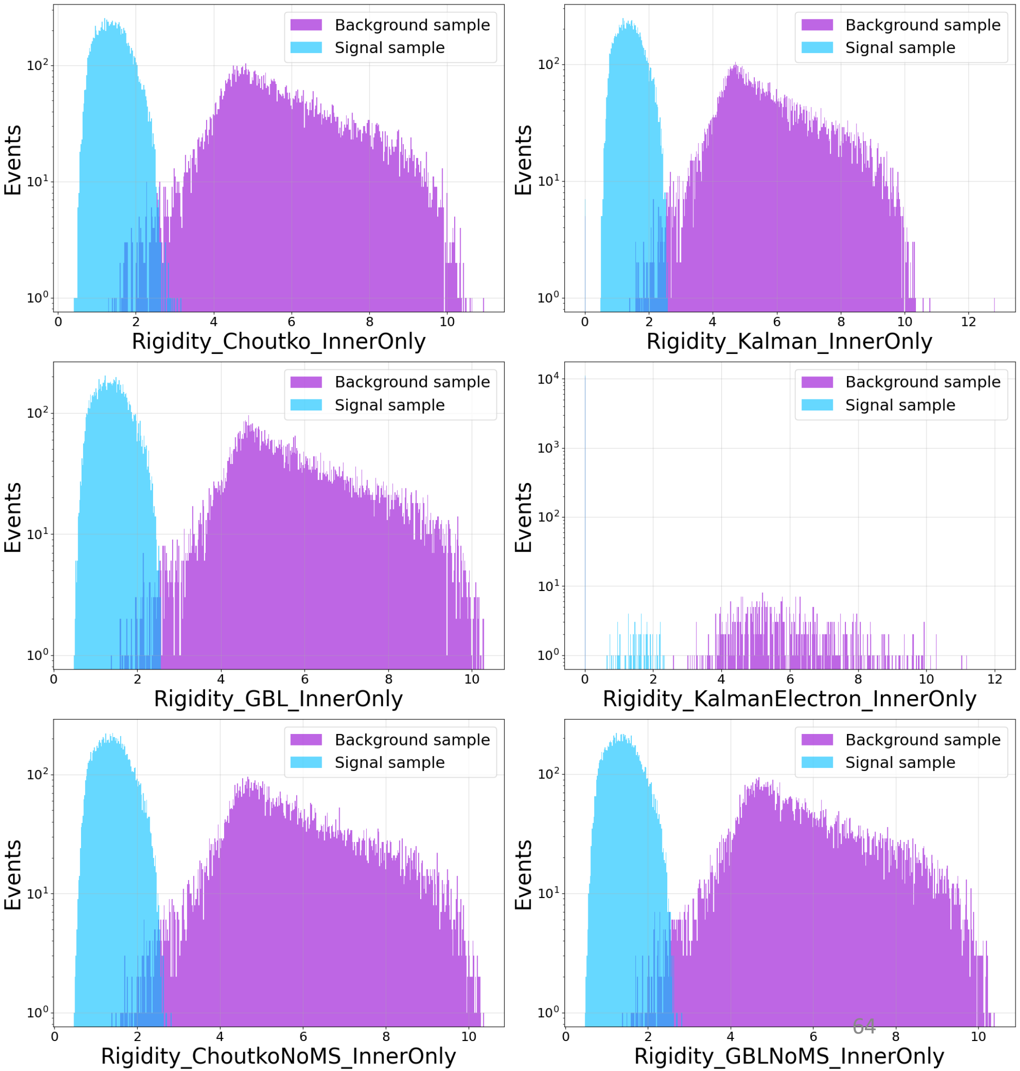
→ The rigidity almost completely characterizes the event of the two samples

i.e.

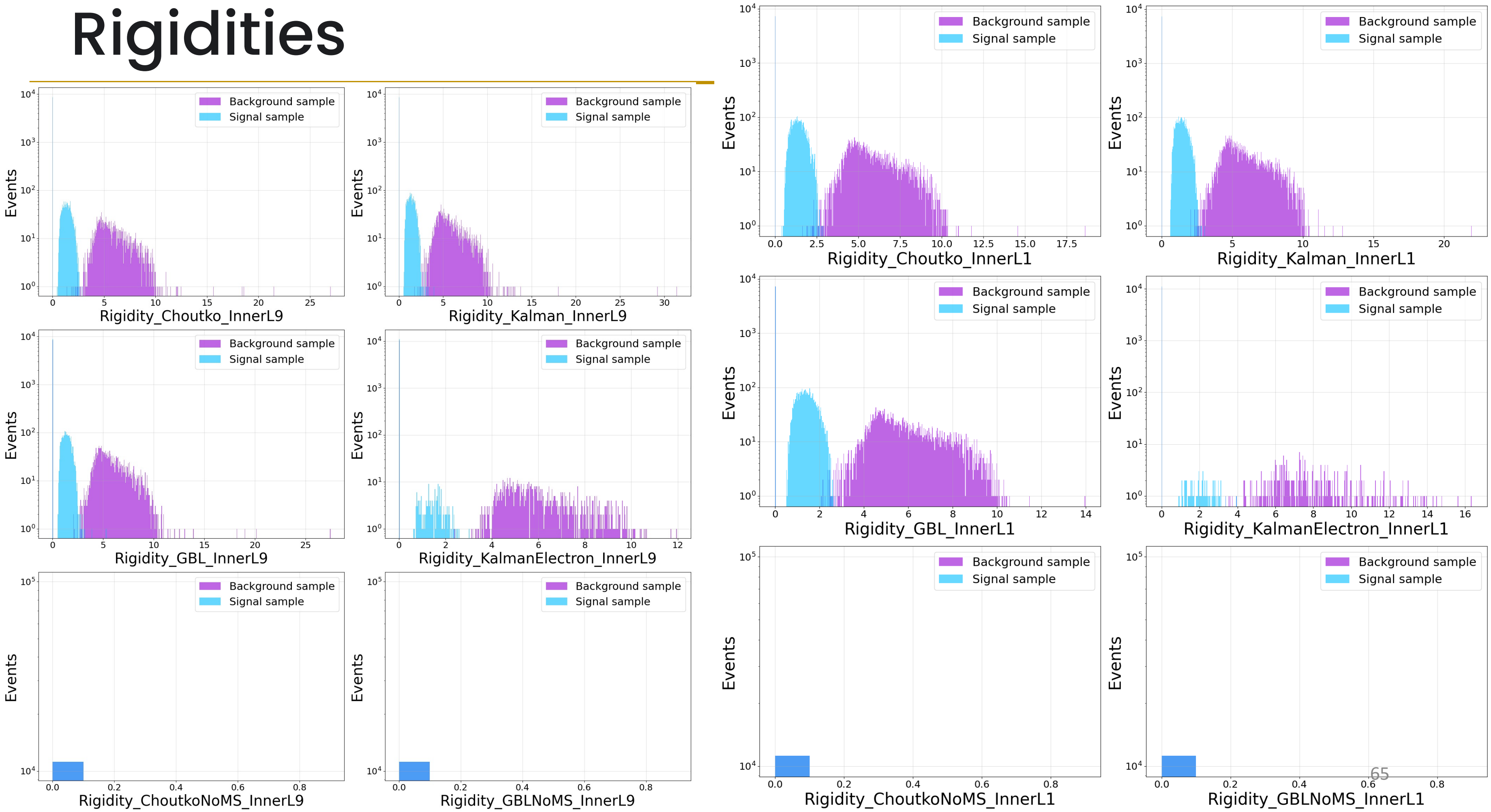
low R (R < 2.5 GV) == signal events

high R (2.5 GV < R < 20 GV) == bkg events

No true for every span

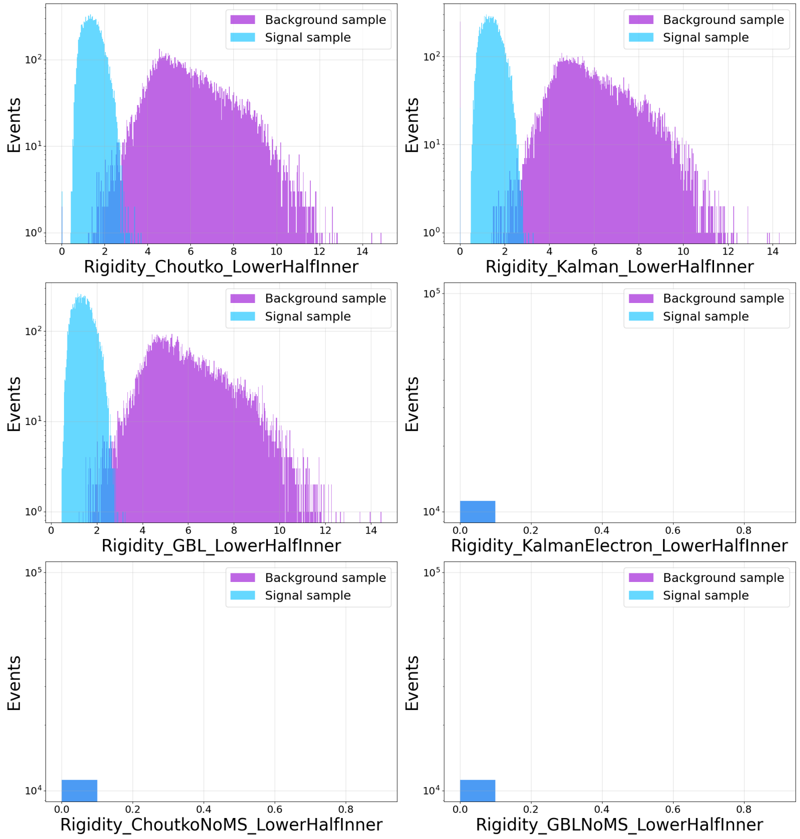


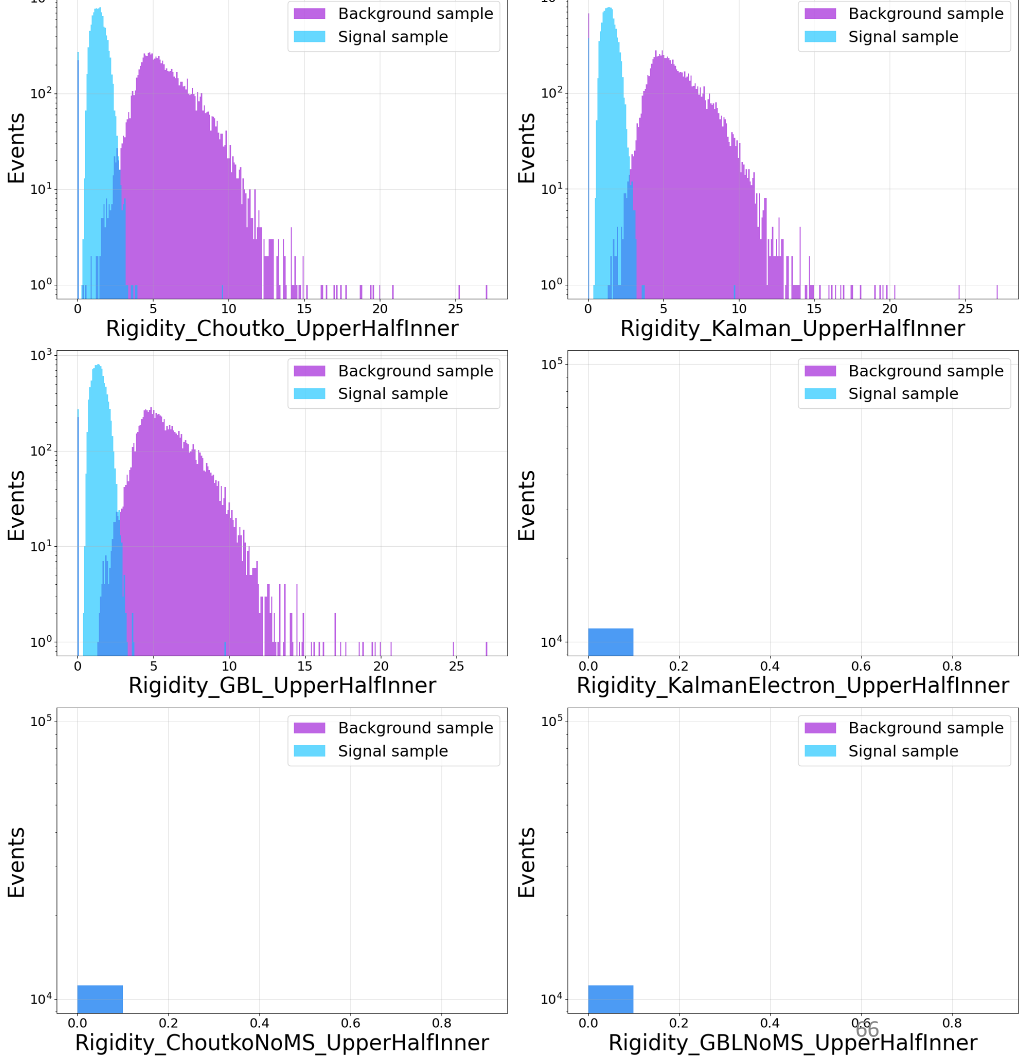
Rigidities



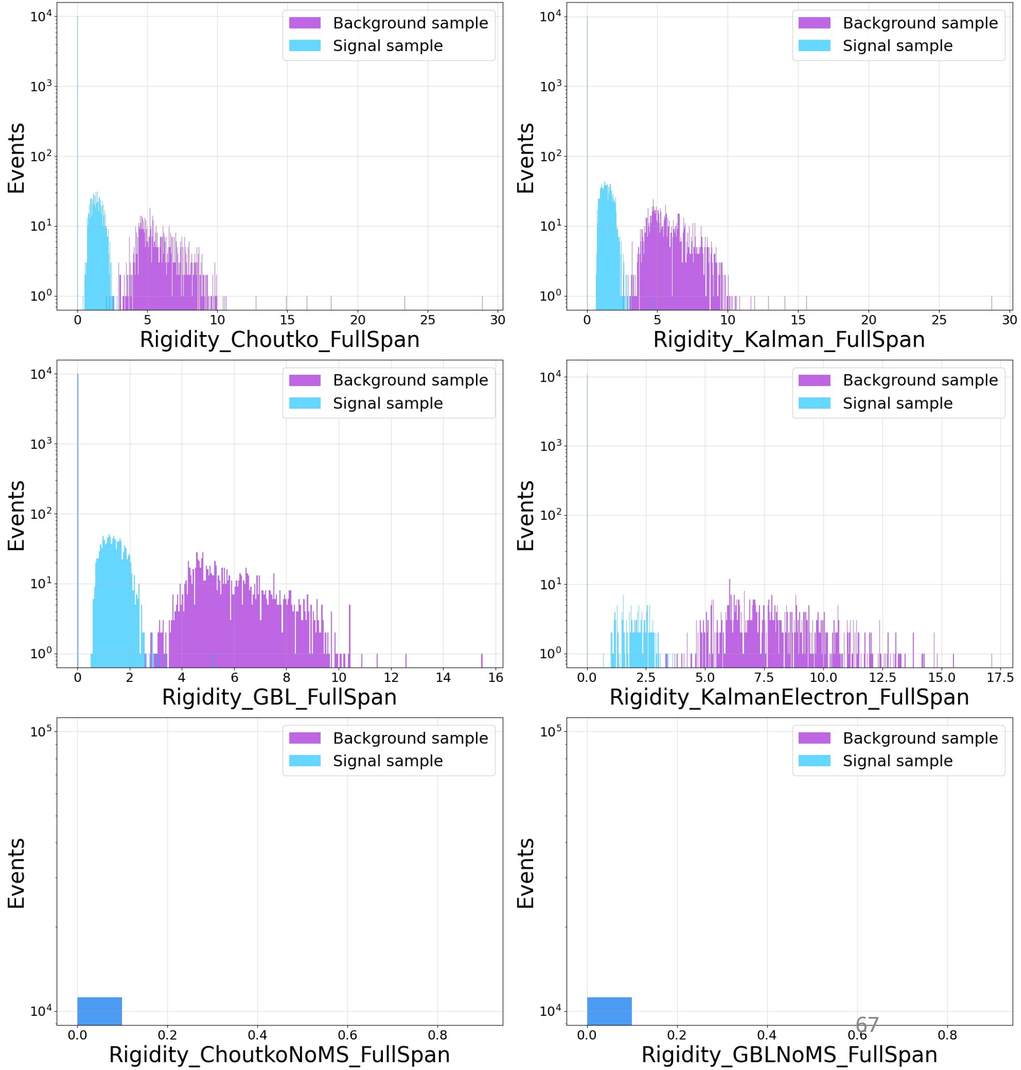
Background sample
Signal sample
Events
Rigidity_Choutko_InnerL9
Rigidity_Kalman_InnerL9
Rigidity_GBL_InnerL9
Rigidity_KalmanElectron_InnerL9
Rigidity_ChoutkoNoMS_InnerL9
Rigidity_GBLNoMS_InnerL9
Rigidity_Choutko_InnerL1
Rigidity_Kalman_InnerL1
Rigidity_GBL_InnerL1
Rigidity_KalmanElectron_InnerL1
Rigidity_ChoutkoNoMS_InnerL1
Rigidity_GBLNoMS_InnerL1

Rigidities


Background sample
Signal sample
Events
Rigidity_Choutko_LowerHalfInner
Rigidity_Kalman_LowerHalfInner
Rigidity_GBL_LowerHalfInner
Rigidity_KalmanElectron_LowerHalfInner
Rigidity_ChoutkoNoMS_LowerHalfInner
Rigidity_GBLNoMS_LowerHalfInner


Background sample
Signal sample
Events
Rigidity_Choutko_UpperHalfInner
Rigidity_Kalman_UpperHalfInner
Rigidity_GBL_UpperHalfInner
Rigidity_KalmanElectron_UpperHalfInner
Rigidity_ChoutkoNoMS_UpperHalfInner
Rigidity_GBLNoMS_UpperHalfInner

Rigidities



Background sample
Signal sample
Events
Rigidity_Choutko_FullSpan
Rigidity_Kalman_FullSpan
Rigidity_GBL_FullSpan
Rigidity_KalmanElectron_FullSpan
Rigidity_ChoutkoNoMS_FullSpan
Rigidity_GBLNoMS_FullSpan

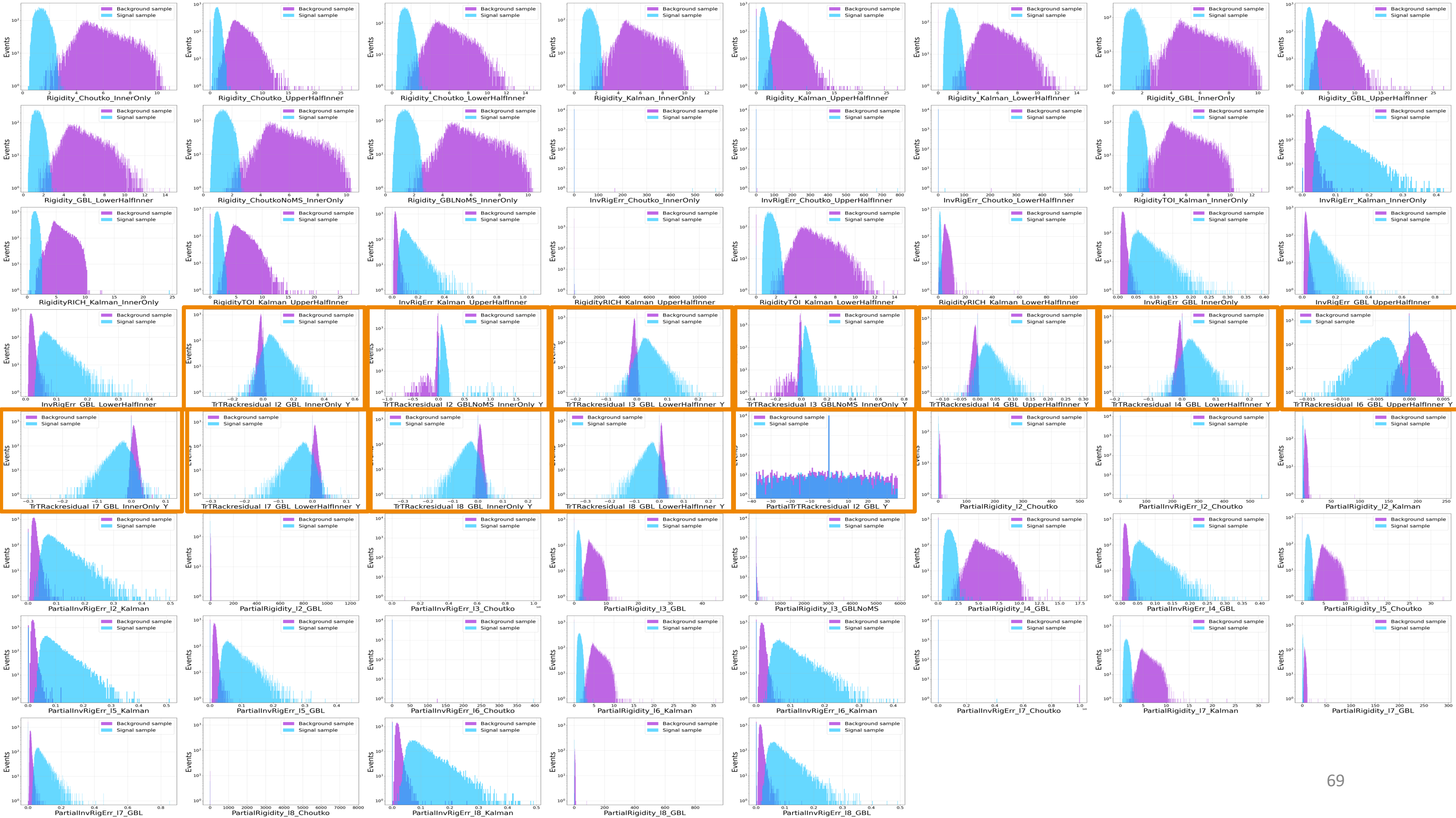
Other features: dependence on rigidity?

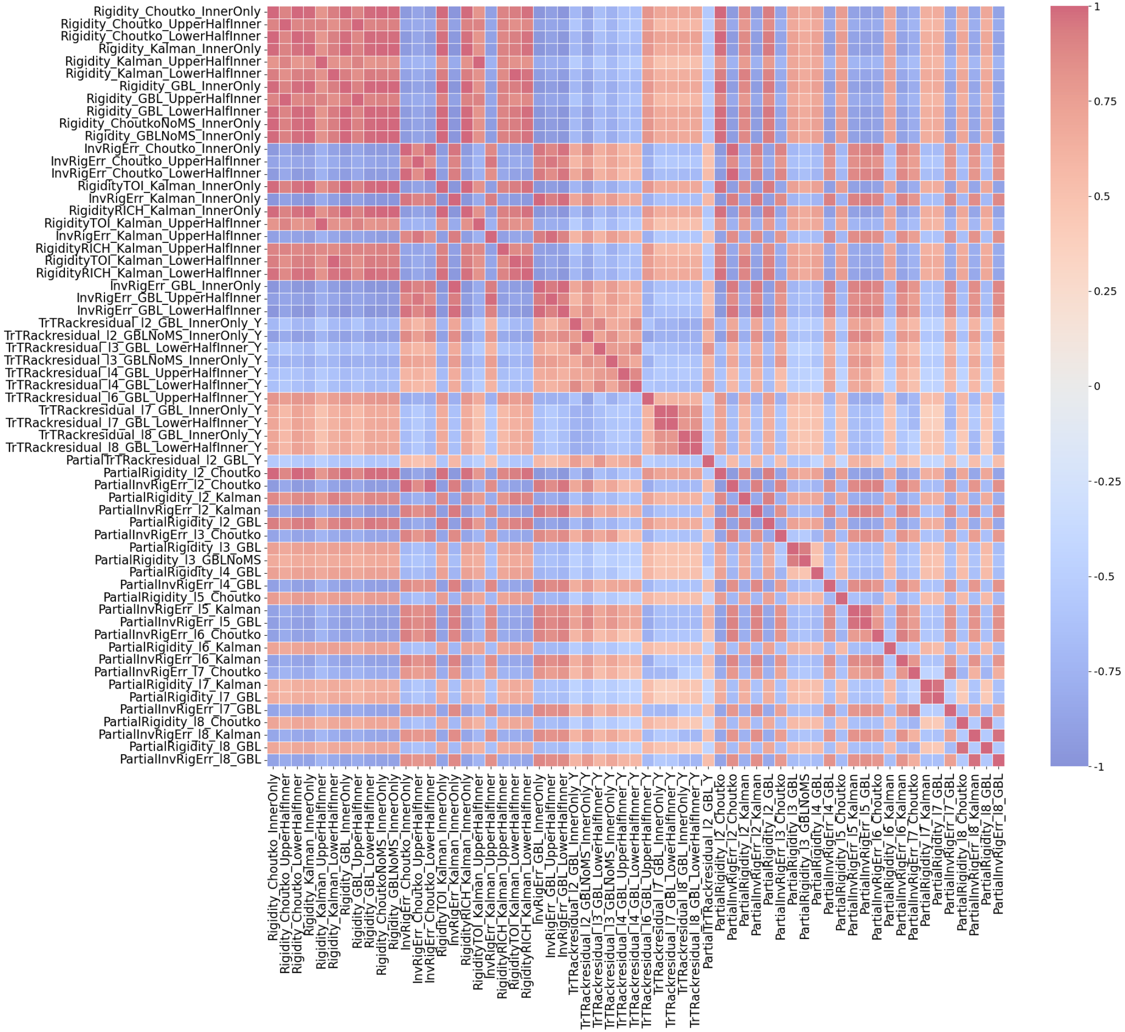
Need to check also for rigidity dependence on the other features

Start to check rigidity dependence in features selected by RF

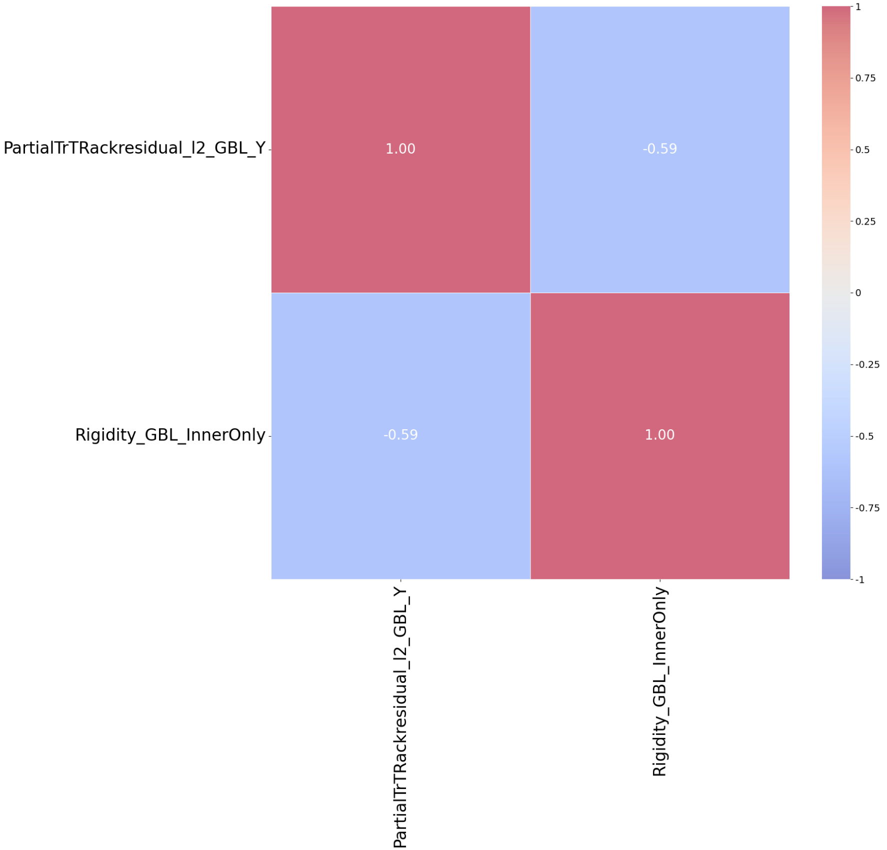
- Plot correlation matrix of rigidity variables plus other features to see if there is a correlation

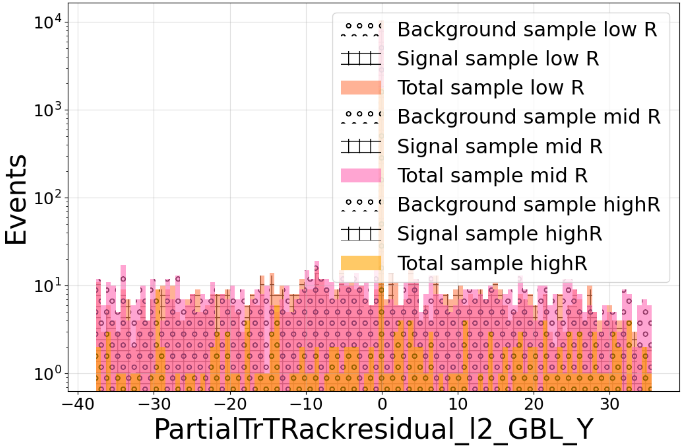
- Plot feature distribution for different rigidity bins


Background sample
Signal sample
Events
Rigidity_Choutko_InnerOnly
Rigidity_Choutko_UpperHalfInner
Rigidity_Choutko_LowerHalfInner
Rigidity_Kalman_InnerOnly
Rigidity_Kalman_UpperHalfInner
Rigidity_Kalman_LowerHalfInner
Rigidity_GBL_InnerOnly
Rigidity_GBL_UpperHalfInner
Rigidity_GBL_LowerHalfInner
Rigidity_ChoutkoNoMS_InnerOnly
Rigidity_GBLNoMS_InnerOnly
InvRigErr_Choutko_InnerOnly
InvRigErr_Choutko_UpperHalfInner
InvRigErr_Choutko_LowerHalfInner
RigidityTOI_Kalman_InnerOnly
InvRigErr_Kalman_InnerOnly
RigidityRICH_Kalman_InnerOnly
RigidityTOI_Kalman_UpperHalfInner
InvRigErr_Kalman_UpperHalfInner
RigidityRICH_Kalman_UpperHalfInner
RigidityTOI_Kalman_LowerHalfInner
RigidityRICH_Kalman_LowerHalfInner
InvRigErr_GBL_InnerOnly
InvRigErr_GBL_UpperHalfInner
InvRigErr_GBL_LowerHalfInner
TrTRackresidual_l2_GBL_InnerOnly_Y
TrTRackresidual_l2_GBLNoMS_InnerOnly_Y
TrTRackresidual_l3_GBL_LowerHalfInner_Y
TrTRackresidual_l3_GBLNoMS_InnerOnly_Y
TrTRackresidual_l4_GBL_UpperHalfInner_Y
TrTRackresidual_l4_GBL_LowerHalfInner_Y
TrTRackresidual_l6_GBL_UpperHalfInner_Y
TrTRackresidual_l7_GBL_InnerOnly_Y
TrTRackresidual_l7_GBL_LowerHalfInner_Y
TrTRackresidual_l8_GBL_InnerOnly_Y
TrTRackresidual_l8_GBL_LowerHalfInner_Y
PartialTrTRackresidual_l2_GBL_Y
PartialRigidity_l2_Choutko
PartialInvRigErr_l2_Choutko
PartialRigidity_l2_Kalman
PartialInvRigErr_l2_Kalman
PartialRigidity_l2_GBL
PartialInvRigErr_l3_Choutko
PartialRigidity_l3_GBL
PartialRigidity_l3_GBLNoMS
PartialRigidity_l4_GBL
PartialInvRigErr_l4_GBL
PartialRigidity_l5_Choutko
PartialInvRigErr_l5_Kalman
PartialInvRigErr_l5_GBL
PartialInvRigErr_l6_Choutko
PartialRigidity_l6_Kalman
PartialInvRigErr_l6_Kalman
PartialInvRigErr_l7_Choutko
PartialRigidity_l7_Kalman
PartialRigidity_l7_GBL
PartialInvRigErr_l7_GBL
PartialRigidity_l8_Choutko
PartialInvRigErr_l8_Kalman
PartialRigidity_l8_GBL
PartialInvRigErr_l8_GBL

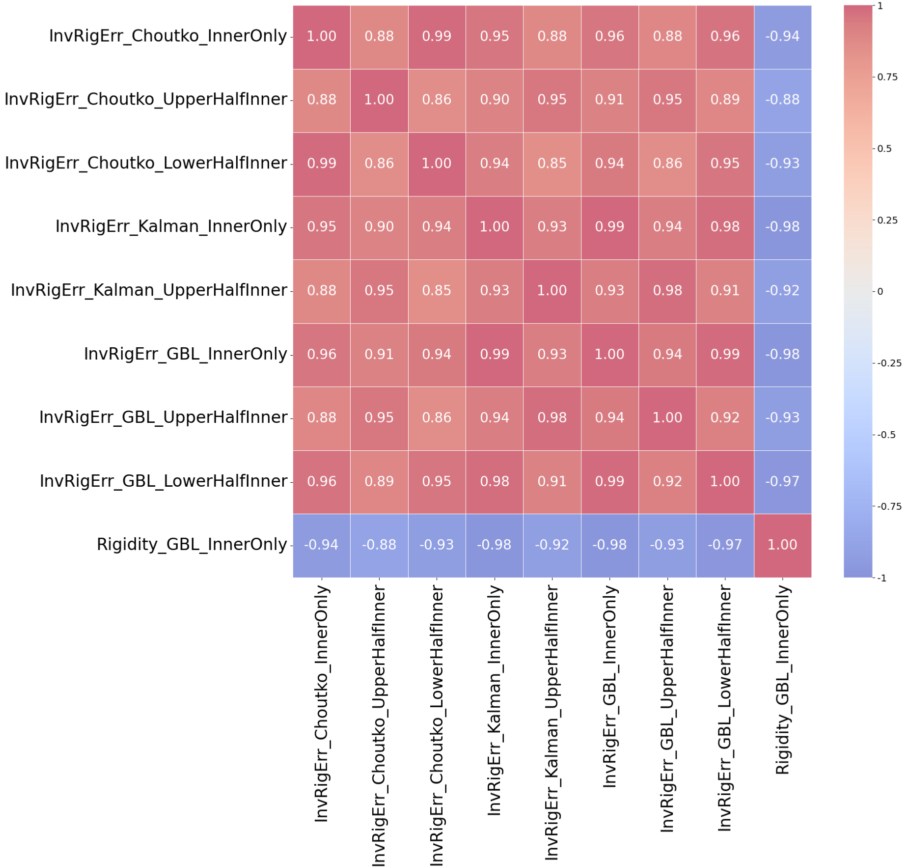

Rigidity_Choutko_InnerOnly
Rigidity_Choutko_UpperHalfInner
Rigidity_Choutko_LowerHalfInner
Rigidity_Kalman_InnerOnly
Rigidity_Kalman_UpperHalfInner
Rigidity_Kalman_LowerHalfInner
Rigidity_GBL_InnerOnly
Rigidity_GBL_UpperHalfInner
Rigidity_GBL_LowerHalfInner
Rigidity_ChoutkoNoMS_InnerOnly
Rigidity_GBLNoMS_InnerOnly
InvRigErr_Choutko_InnerOnly
InvRigErr_Choutko_UpperHalfInner
InvRigErr_Choutko_LowerHalfInner
RigidityTOI_Kalman_InnerOnly
InvRigErr_Kalman_InnerOnly
RigidityRICH_Kalman_InnerOnly
RigidityTOI_Kalman_UpperHalfInner
InvRigErr_Kalman_UpperHalfInner
RigidityRICH_Kalman_UpperHalfInner
RigidityTOI_Kalman_LowerHalfInner
RigidityRICH_Kalman_LowerHalfInner
InvRigErr_GBL_InnerOnly
InvRigErr_GBL_UpperHalfInner
InvRigErr_GBL_LowerHalfInner
TrTRackresidual_l2_GBL_InnerOnly_Y
TrTRackresidual_l2_GBLNoMS_InnerOnly_Y
TrTRackresidual_l3_GBL_LowerHalfInner_Y
TrTRackresidual_l3_GBLNoMS_InnerOnly_Y
TrTRackresidual_l4_GBL_UpperHalfInner_Y
TrTRackresidual_l4_GBL_LowerHalfInner_Y
TrTRackresidual_l6_GBL_UpperHalfInner_Y
TrTRackresidual_l7_GBL_InnerOnly_Y
TrTRackresidual_l7_GBL_LowerHalfInner_Y
TrTRackresidual_l8_GBL_InnerOnly_Y
TrTRackresidual_l8_GBL_LowerHalfInner_Y
PartialTrTRackresidual_l2_GBL_Y
PartialRigidity_l2_Choutko
PartialInvRigErr_l2_Choutko
PartialRigidity_l2_Kalman
PartialInvRigErr_l2_Kalman
PartialRigidity_l2_GBL
PartialInvRigErr_l3_Choutko
PartialRigidity_l3_GBL
PartialRigidity_l3_GBLNoMS
PartialRigidity_l4_GBL
PartialInvRigErr_l4_GBL
PartialRigidity_l5_Choutko
PartialInvRigErr_l5_Kalman
PartialInvRigErr_l5_GBL
PartialInvRigErr_l6_Choutko
PartialRigidity_l6_Kalman
PartialInvRigErr_l6_Kalman
PartialInvRigErr_l7_Choutko
PartialRigidity_l7_Kalman
PartialRigidity_l7_GBL
PartialInvRigErr_l7_GBL
PartialRigidity_l8_Choutko
PartialInvRigErr_l8_Kalman
PartialRigidity_l8_GBL
PartialInvRigErr_l8_GBL
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1

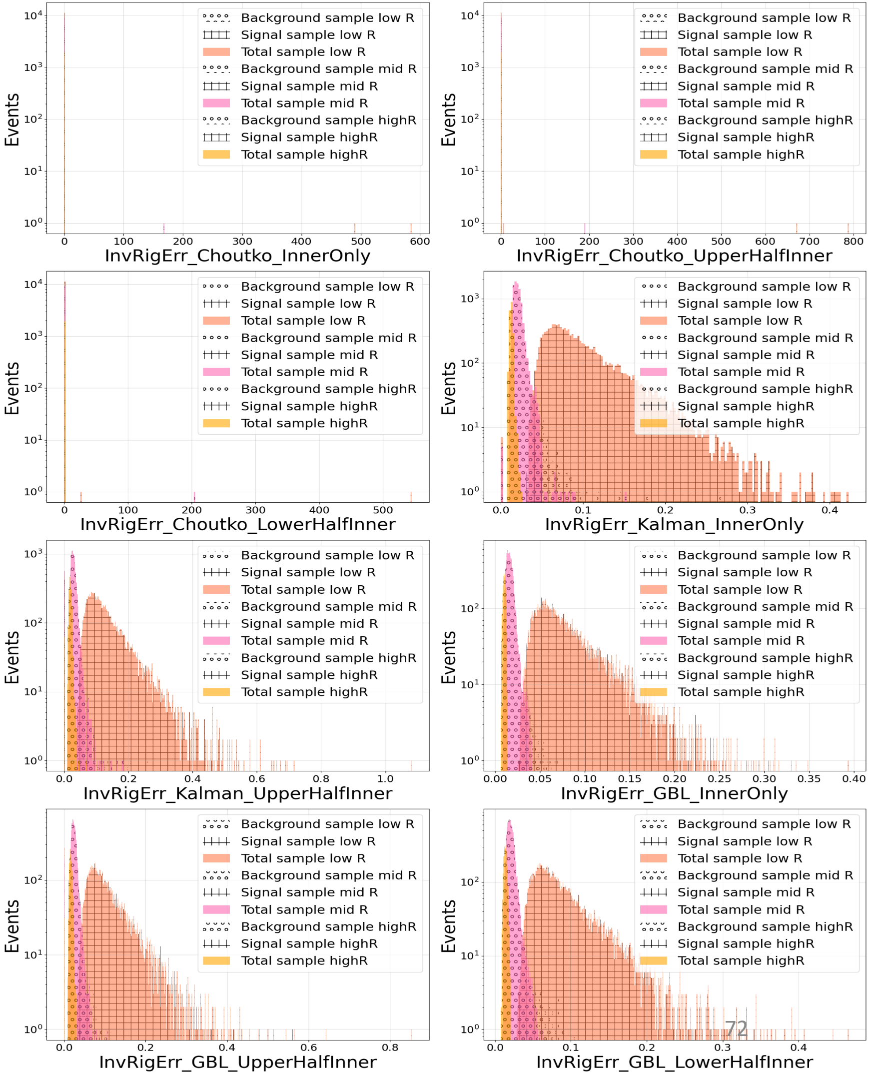
Partial TrTrackResiduals



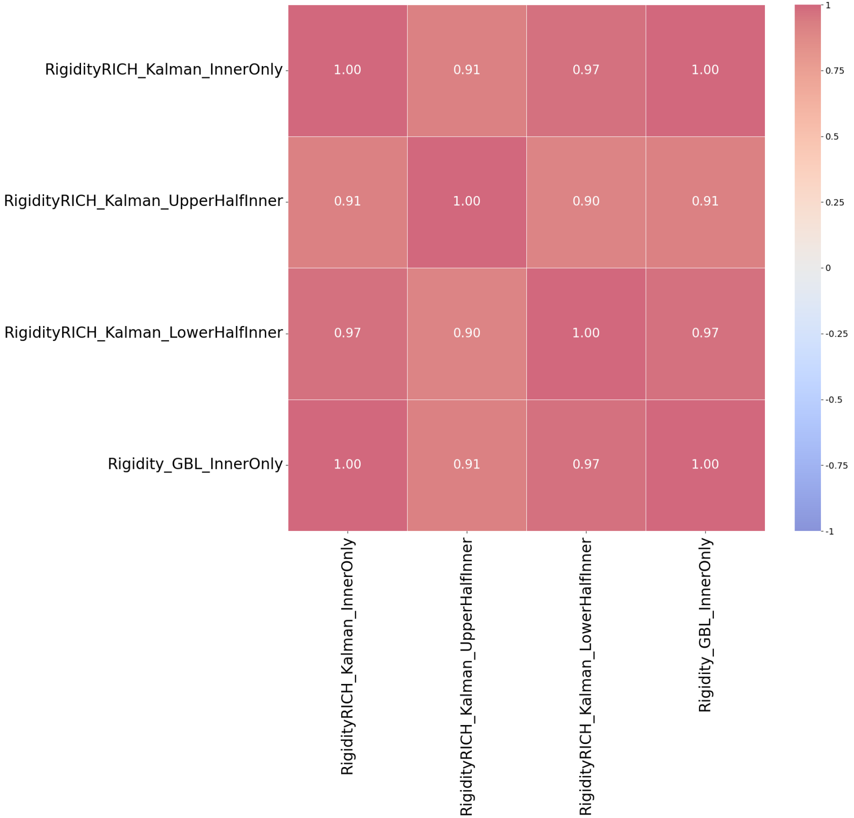


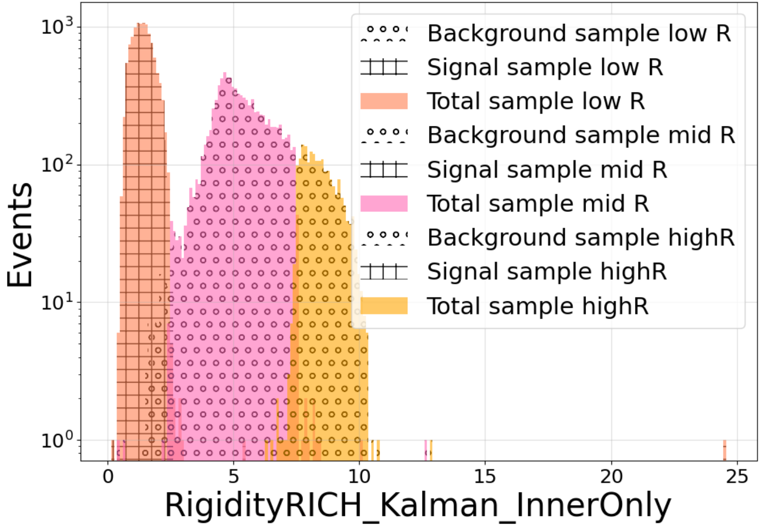
Inv Rig Err

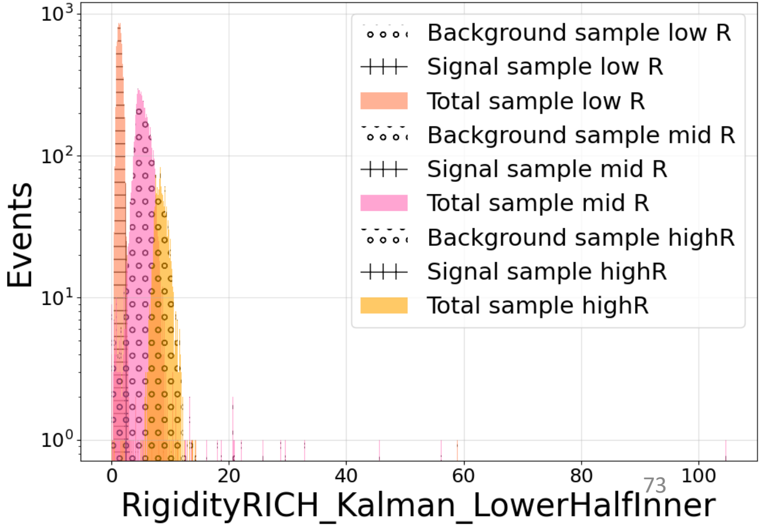

InvRigErr_Choutko_InnerOnly 1.00 0.88 0.99 0.95 0.88 0.96 0.88 0.96 -0.94
InvRigErr_Choutko_UpperHalfInner 0.88 1.00 0.86 0.90 0.95 0.91 0.95 0.89 -0.88
InvRigErr_Choutko_LowerHalfInner 0.99 0.86 1.00 0.94 0.85 0.94 0.86 0.95 -0.93
InvRigErr_Kalman_InnerOnly 0.95 0.90 0.94 1.00 0.93 0.99 0.94 0.98 -0.98
InvRigErr_Kalman_UpperHalfInner 0.88 0.95 0.85 0.93 1.00 0.93 0.98 0.91 -0.92
InvRigErr_GBL_InnerOnly 0.96 0.91 0.94 0.99 0.93 1.00 0.94 0.99 -0.98
InvRigErr_GBL_UpperHalfInner 0.88 0.95 0.86 0.94 0.98 0.94 1.00 0.92 -0.93
InvRigErr_GBL_LowerHalfInner 0.96 0.89 0.95 0.98 0.91 0.99 0.92 1.00 -0.97
Rigidity_GBL_InnerOnly -0.94 -0.88 -0.93 -0.98 -0.92 -0.98 -0.93 -0.97 1.00
InvRigErr_Choutko_InnerOnly
InvRigErr_Choutko_UpperHalfInner
InvRigErr_Choutko_LowerHalfInner
InvRigErr_Kalman_InnerOnly
InvRigErr_Kalman_UpperHalfInner
InvRigErr_GBL_InnerOnly
InvRigErr_GBL_UpperHalfInner
InvRigErr_GBL_LowerHalfInner
Rigidity_GBL_InnerOnly
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1


Background sample low R
Signal sample low R
Total sample low R
Background sample mid R
Signal sample mid R
Total sample mid R
Background sample highR
Signal sample highR
Total sample highR
Events
InvRigErr_Choutko_InnerOnly
InvRigErr_Choutko_UpperHalfInner
InvRigErr_Choutko_LowerHalfInner
InvRigErr_Kalman_InnerOnly
InvRigErr_Kalman_UpperHalfInner
InvRigErr_GBL_InnerOnly
InvRigErr_GBL_UpperHalfInner
InvRigErr_GBL_LowerHalfInner

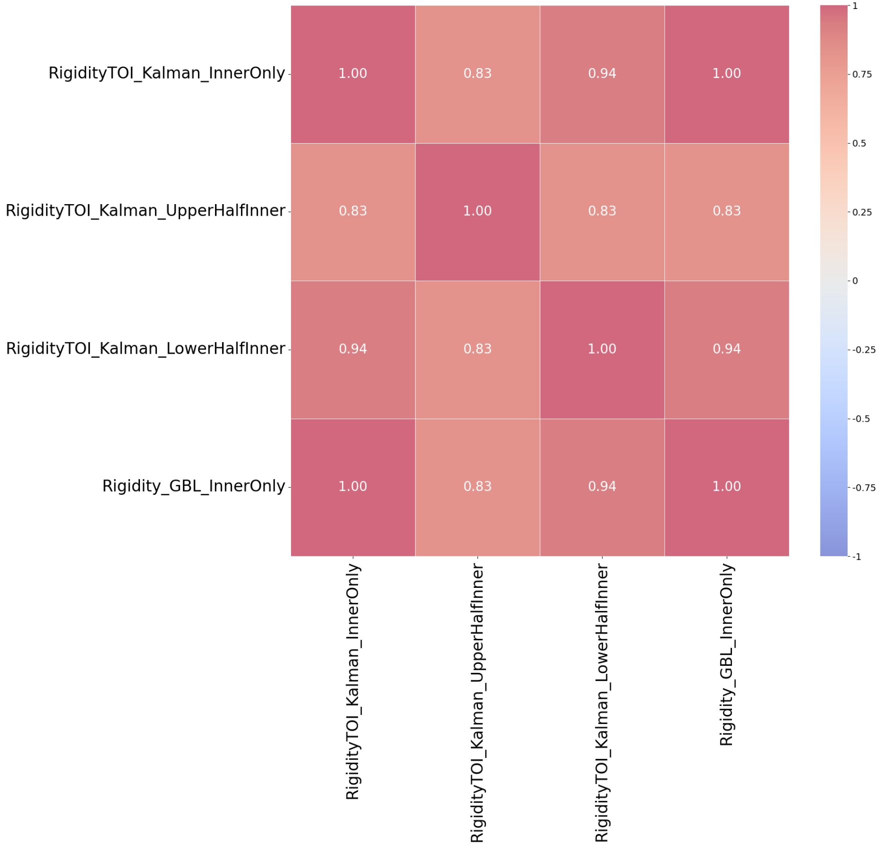
Rigidity RICH

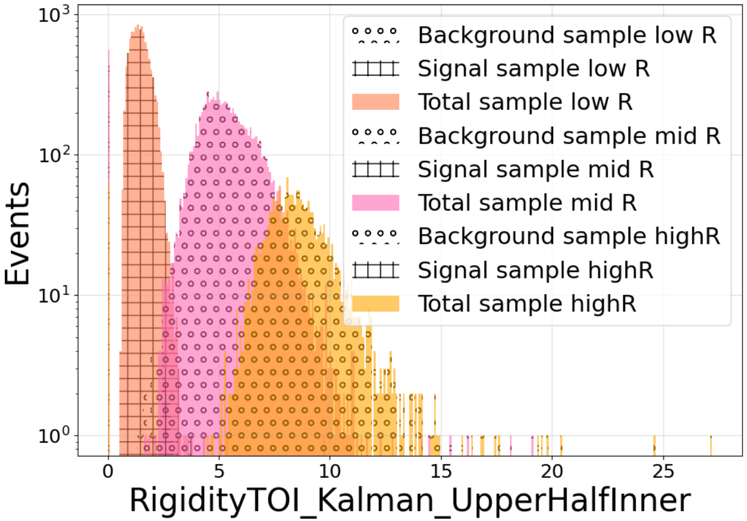


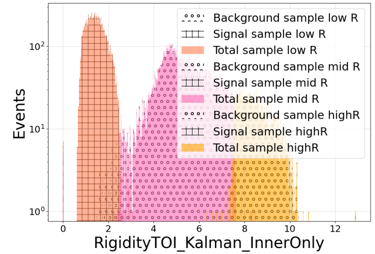


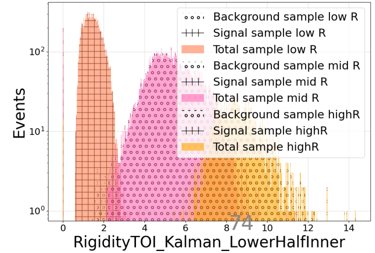


Rigidity TOI



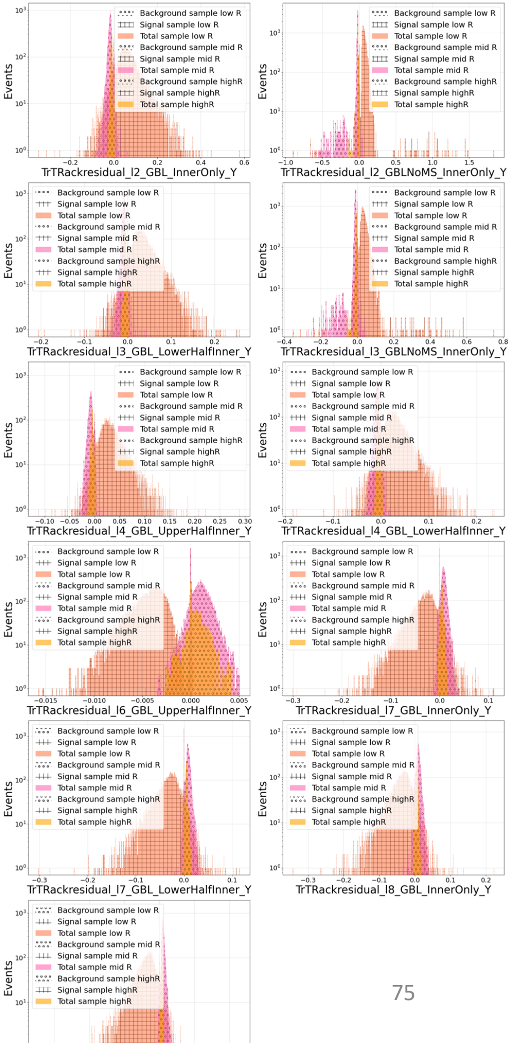




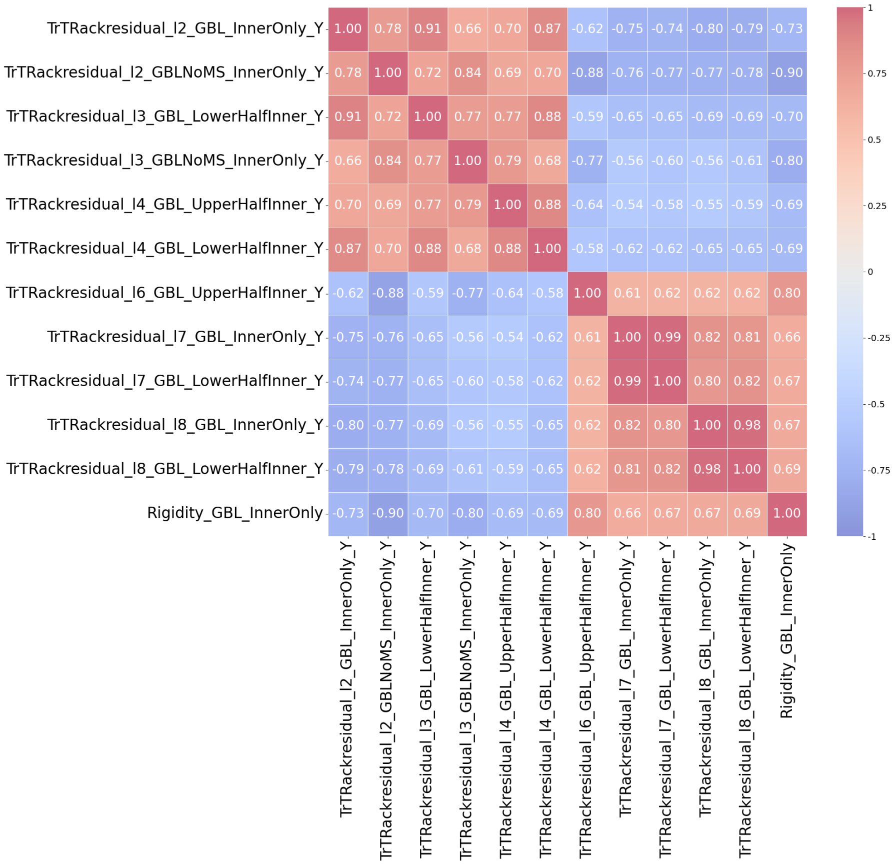


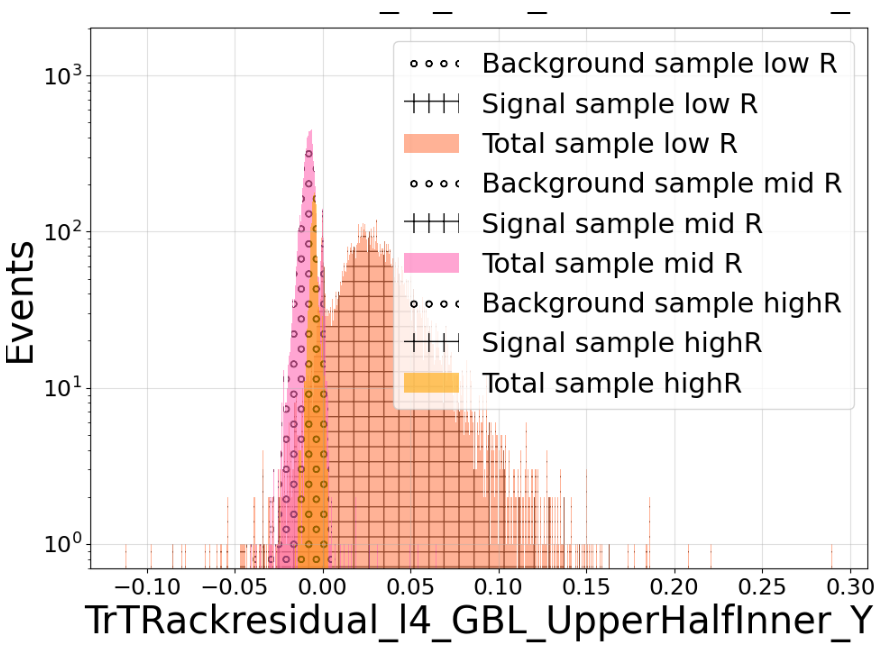
TrTrackResiduals

	TrTRackresidual_l2_GBL_InnerOnly_Y	TrTRackresidual_l2_GBLNoMS_InnerOnly_Y	TrTRackresidual_l3_GBL_LowerHalfInner_Y	TrTRackresidual_l3_GBLNoMS_InnerOnly_Y	TrTRackresidual_l4_GBL_UpperHalfInner_Y	TrTRackresidual_l4_GBL_LowerHalfInner_Y	TrTRackresidual_l6_GBL_UpperHalfInner_Y	TrTRackresidual_l7_GBL_InnerOnly_Y	TrTRackresidual_l7_GBL_LowerHalfInner_Y	TrTRackresidual_l8_GBL_InnerOnly_Y	TrTRackresidual_l8_GBL_LowerHalfInner_Y	Rigidity_GBL_InnerOnly
TrTRackresidual_l2_GBL_InnerOnly_Y	1.00	0.78	0.91	0.66	0.70	0.87	-0.62	-0.75	-0.74	-0.80	-0.79	-0.73
TrTRackresidual_l2_GBLNoMS_InnerOnly_Y	0.78	1.00	0.72	0.84	0.69	0.70	-0.88	-0.76	-0.77	-0.77	-0.78	-0.90
TrTRackresidual_l3_GBL_LowerHalfInner_Y	0.91	0.72	1.00	0.77	0.77	0.88	-0.59	-0.65	-0.65	-0.69	-0.69	-0.70
TrTRackresidual_l3_GBLNoMS_InnerOnly_Y	0.66	0.84	0.77	1.00	0.79	0.68	-0.77	-0.56	-0.60	-0.56	-0.61	-0.80
TrTRackresidual_l4_GBL_UpperHalfInner_Y	0.70	0.69	0.77	0.79	1.00	0.88	-0.64	-0.54	-0.58	-0.55	-0.59	-0.69
TrTRackresidual_l4_GBL_LowerHalfInner_Y	0.87	0.70	0.88	0.68	0.88	1.00	-0.58	-0.62	-0.62	-0.65	-0.65	-0.69
TrTRackresidual_l6_GBL_UpperHalfInner_Y	-0.62	-0.88	-0.59	-0.77	-0.64	-0.58	1.00	0.61	0.62	0.62	0.62	0.80
TrTRackresidual_l7_GBL_InnerOnly_Y	-0.75	-0.76	-0.65	-0.56	-0.54	-0.62	0.61	1.00	0.99	0.82	0.81	0.66
TrTRackresidual_l7_GBL_LowerHalfInner_Y	-0.74	-0.77	-0.65	-0.60	-0.58	-0.62	0.62	0.99	1.00	0.80	0.82	0.67
TrTRackresidual_l8_GBL_InnerOnly_Y	-0.80	-0.77	-0.69	-0.56	-0.55	-0.65	0.62	0.82	0.80	1.00	0.98	0.67
TrTRackresidual_l8_GBL_LowerHalfInner_Y	-0.79	-0.78	-0.69	-0.61	-0.59	-0.65	0.62	0.81	0.82	0.98	1.00	0.69
Rigidity_GBL_InnerOnly	-0.73	-0.90	-0.70	-0.80	-0.69	-0.69	0.80	0.66	0.67	0.67	0.69	1.00

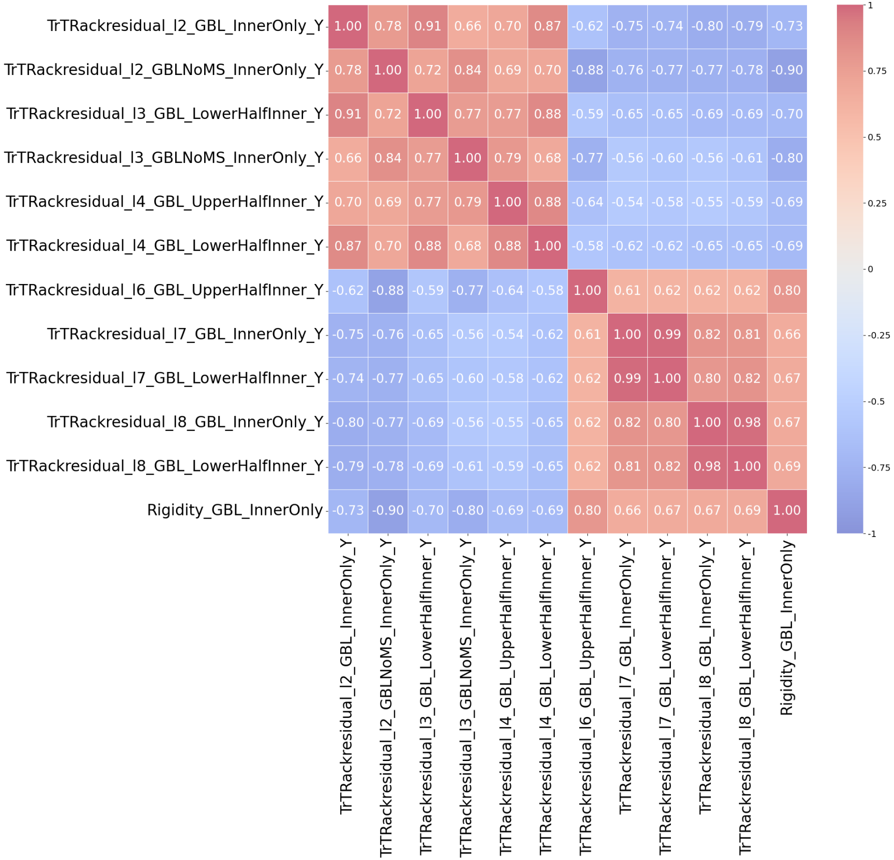


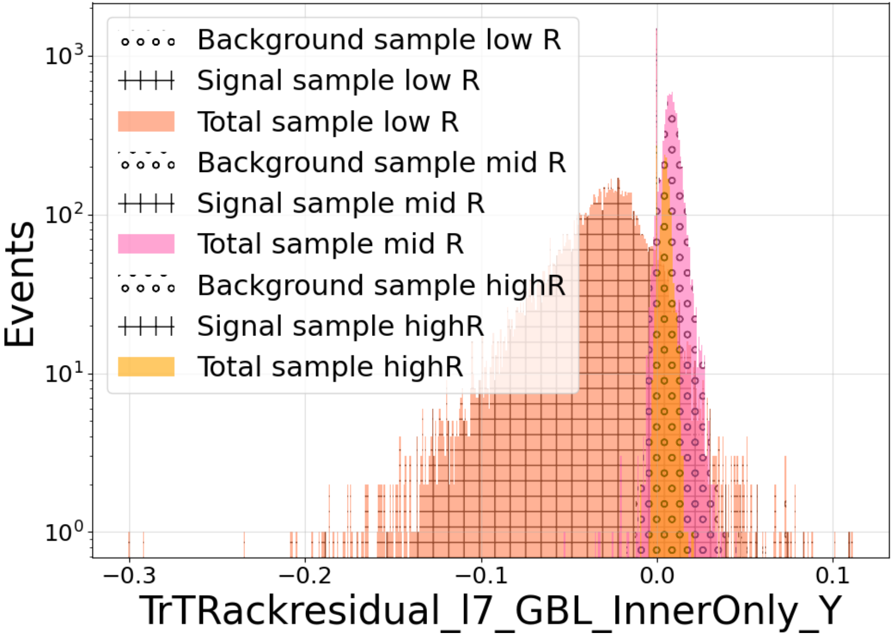
TrTrackResiduals





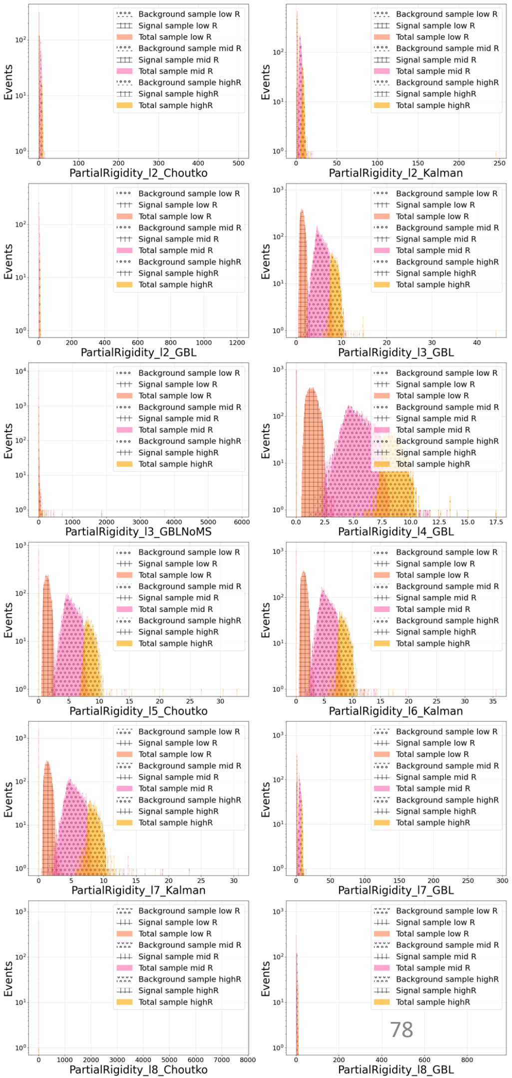
Partial TrTrackResiduals





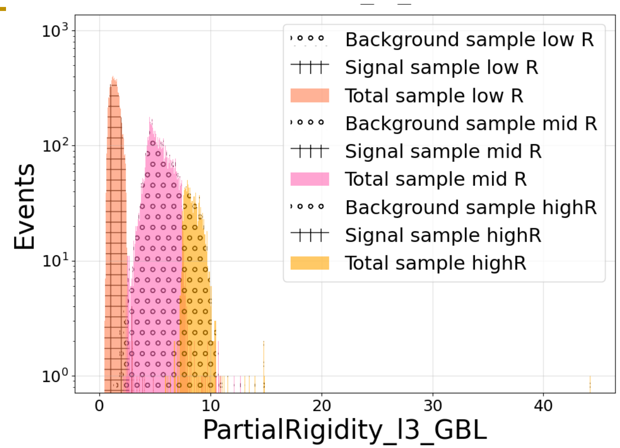
Partial Rigidity

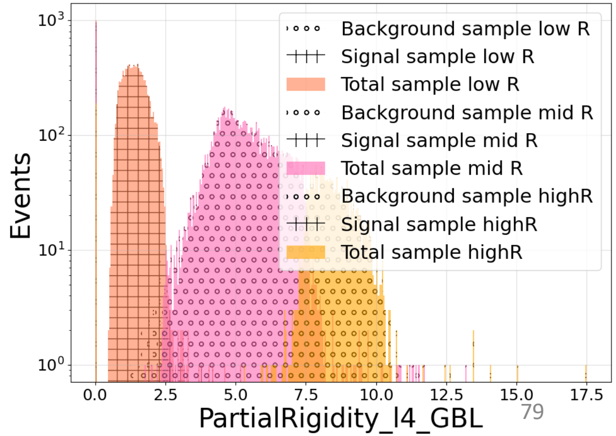
	PartialRigidity_l2_Choutko	PartialRigidity_l2_Kalman	PartialRigidity_l2_GBL	PartialRigidity_l3_GBL	PartialRigidity_l3_GBLNoMS	PartialRigidity_l4_GBL	PartialRigidity_l5_Choutko	PartialRigidity_l6_Kalman	PartialRigidity_l7_Kalman	PartialRigidity_l7_GBL	PartialRigidity_l8_Choutko	PartialRigidity_l8_GBL	Rigidity_GBL_InnerOnly
PartialRigidity_l2_Choutko	1.00	0.88	0.94	0.73	0.67	0.73	0.77	0.72	0.64	0.66	0.71	0.70	0.98
PartialRigidity_l2_Kalman	0.88	1.00	0.93	0.71	0.67	0.70	0.72	0.73	0.65	0.64	0.65	0.67	0.88
PartialRigidity_l2_GBL	0.94	0.93	1.00	0.73	0.67	0.72	0.74	0.73	0.64	0.65	0.68	0.69	0.94
PartialRigidity_l3_GBL	0.73	0.71	0.73	1.00	0.94	0.52	0.60	0.61	0.54	0.53	0.56	0.56	0.74
PartialRigidity_l3_GBLNoMS	0.67	0.67	0.67	0.94	1.00	0.47	0.55	0.56	0.50	0.49	0.51	0.52	0.67
PartialRigidity_l4_GBL	0.73	0.70	0.72	0.52	0.47	1.00	0.59	0.58	0.52	0.52	0.54	0.55	0.74
PartialRigidity_l5_Choutko	0.77	0.72	0.74	0.60	0.55	0.59	1.00	0.56	0.54	0.55	0.57	0.57	0.77
PartialRigidity_l6_Kalman	0.72	0.73	0.73	0.61	0.56	0.58	0.56	1.00	0.58	0.55	0.57	0.57	0.73
PartialRigidity_l7_Kalman	0.64	0.65	0.64	0.54	0.50	0.52	0.54	0.58	1.00	0.97	0.44	0.45	0.65
PartialRigidity_l7_GBL	0.66	0.64	0.65	0.53	0.49	0.52	0.55	0.55	0.97	1.00	0.42	0.43	0.66
PartialRigidity_l8_Choutko	0.71	0.65	0.68	0.56	0.51	0.54	0.57	0.57	0.44	0.42	1.00	0.98	0.70
PartialRigidity_l8_GBL	0.70	0.67	0.69	0.56	0.52	0.55	0.57	0.57	0.45	0.43	0.98	1.00	0.70
Rigidity_GBL_InnerOnly	0.98	0.88	0.94	0.74	0.67	0.74	0.77	0.73	0.65	0.66	0.70	0.70	1.00



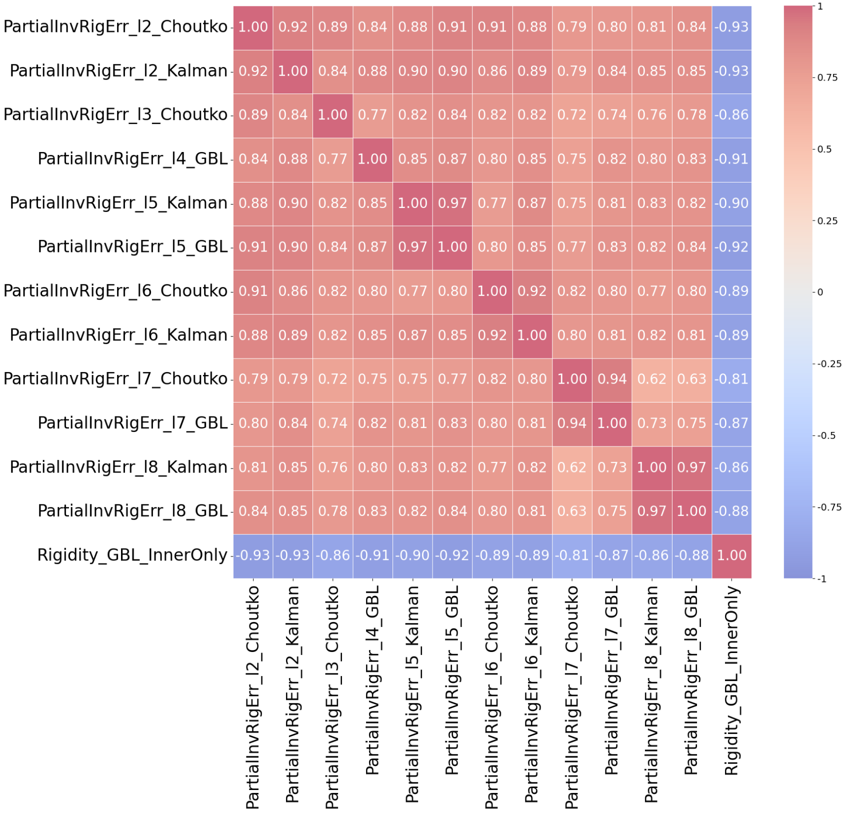
Partial Rigidity

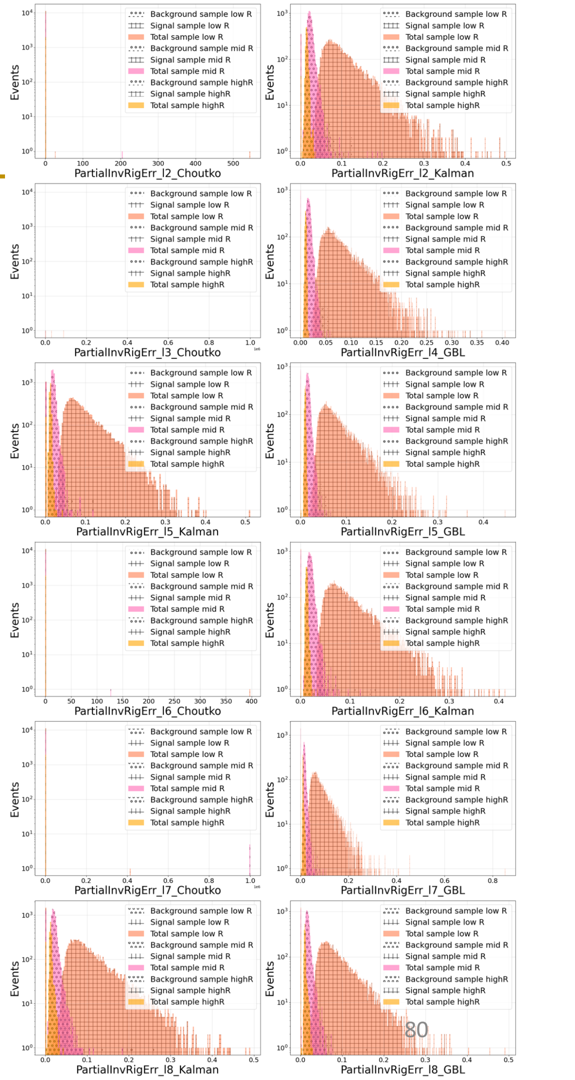
	PartialRigidity_l2_Choutko	PartialRigidity_l2_Kalman	PartialRigidity_l2_GBL	PartialRigidity_l3_GBL	PartialRigidity_l3_GBLNoMS	PartialRigidity_l4_GBL	PartialRigidity_l5_Choutko	PartialRigidity_l6_Kalman	PartialRigidity_l7_Kalman	PartialRigidity_l7_GBL	PartialRigidity_l8_Choutko	PartialRigidity_l8_GBL	Rigidity_GBL_InnerOnly
PartialRigidity_l2_Choutko	1.00	0.88	0.94	0.73	0.67	0.73	0.77	0.72	0.64	0.66	0.71	0.70	0.98
PartialRigidity_l2_Kalman	0.88	1.00	0.93	0.71	0.67	0.70	0.72	0.73	0.65	0.64	0.65	0.67	0.88
PartialRigidity_l2_GBL	0.94	0.93	1.00	0.73	0.67	0.72	0.74	0.73	0.64	0.65	0.68	0.69	0.94
PartialRigidity_l3_GBL	0.73	0.71	0.73	1.00	0.94	0.52	0.60	0.61	0.54	0.53	0.56	0.56	0.74
PartialRigidity_l3_GBLNoMS	0.67	0.67	0.67	0.94	1.00	0.47	0.55	0.56	0.50	0.49	0.51	0.52	0.67
PartialRigidity_l4_GBL	0.73	0.70	0.72	0.52	0.47	1.00	0.59	0.58	0.52	0.52	0.54	0.55	0.74
PartialRigidity_l5_Choutko	0.77	0.72	0.74	0.60	0.55	0.59	1.00	0.56	0.54	0.55	0.57	0.57	0.77
PartialRigidity_l6_Kalman	0.72	0.73	0.73	0.61	0.56	0.58	0.56	1.00	0.58	0.55	0.57	0.57	0.73
PartialRigidity_l7_Kalman	0.64	0.65	0.64	0.54	0.50	0.52	0.54	0.58	1.00	0.97	0.44	0.45	0.65
PartialRigidity_l7_GBL	0.66	0.64	0.65	0.53	0.49	0.52	0.55	0.55	0.97	1.00	0.42	0.43	0.66
PartialRigidity_l8_Choutko	0.71	0.65	0.68	0.56	0.51	0.54	0.57	0.57	0.44	0.42	1.00	0.98	0.70
PartialRigidity_l8_GBL	0.70	0.67	0.69	0.56	0.52	0.55	0.57	0.57	0.45	0.43	0.98	1.00	0.70
Rigidity_GBL_InnerOnly	0.98	0.88	0.94	0.74	0.67	0.74	0.77	0.73	0.65	0.66	0.70	0.70	1.00



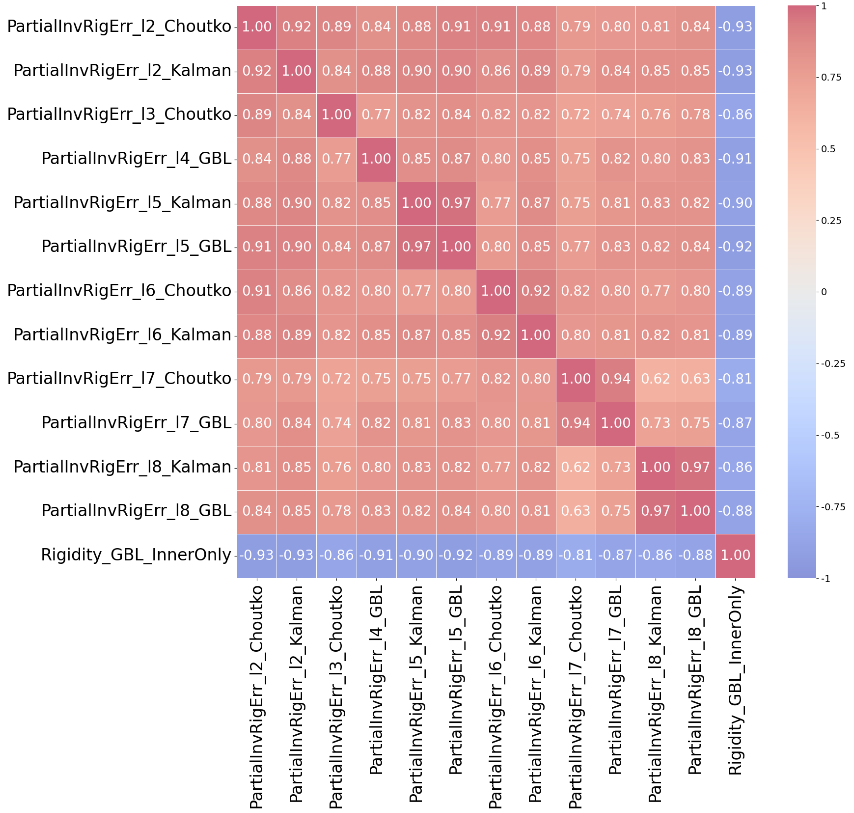


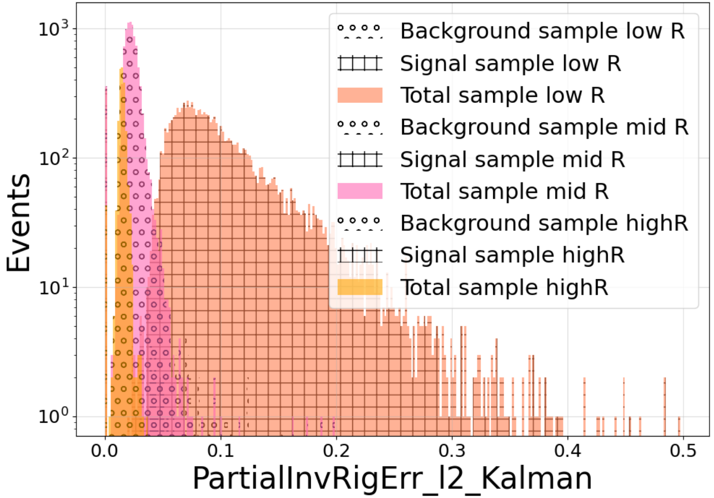
Partial Inv Rig Err

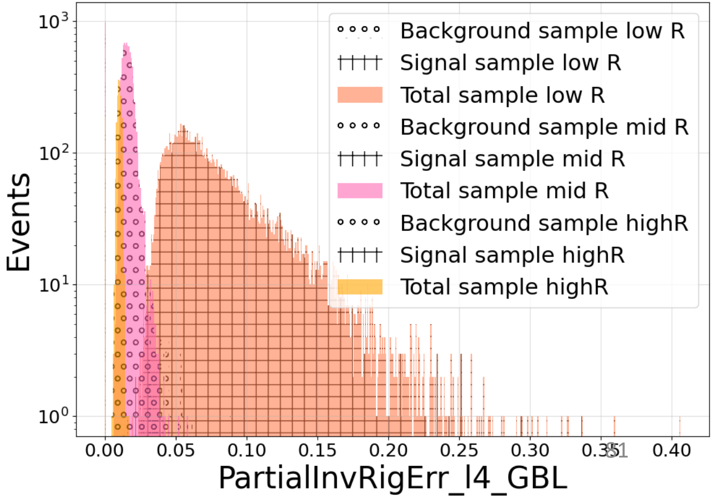




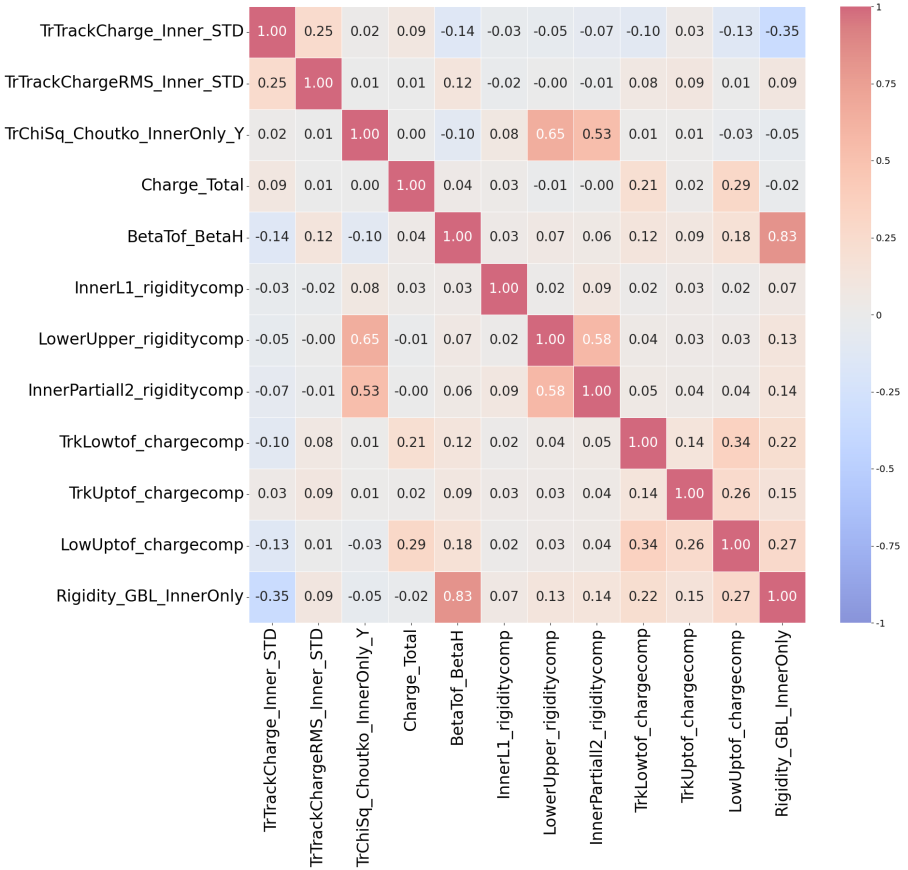
Partial Inv Rig Err



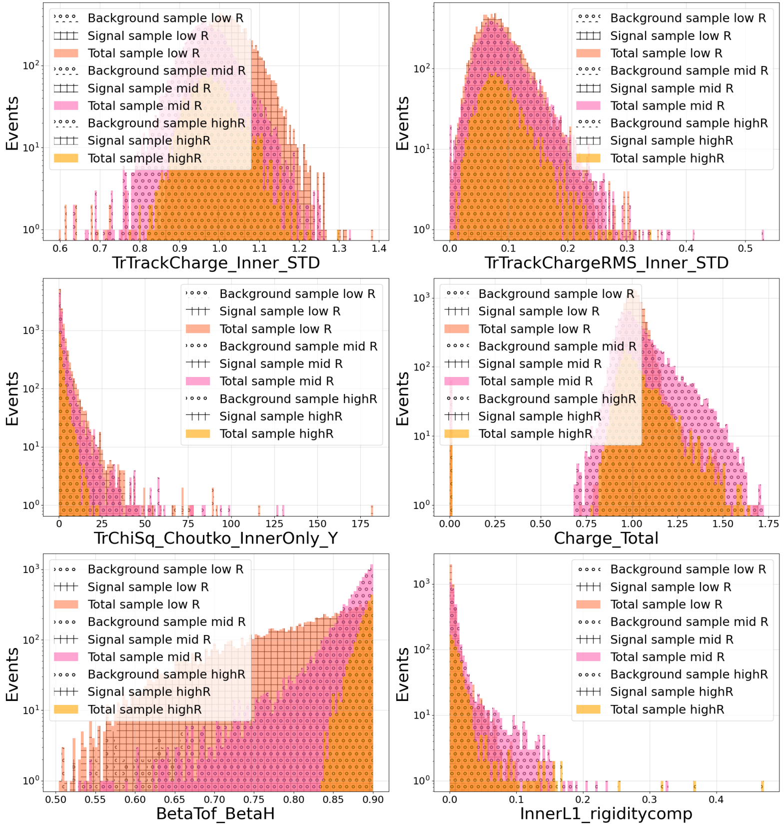


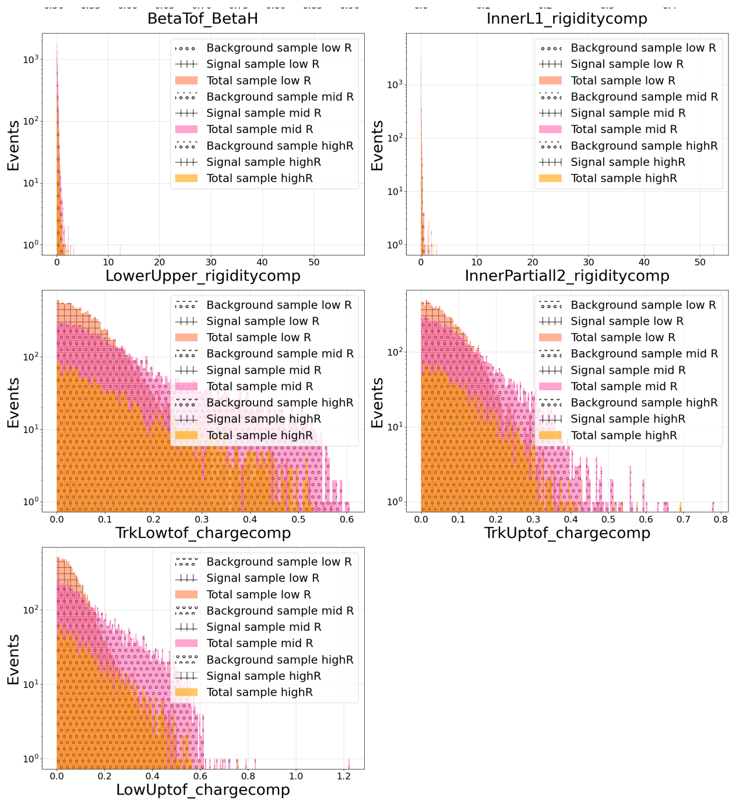


Physics Driven Features



Physics Driven Features



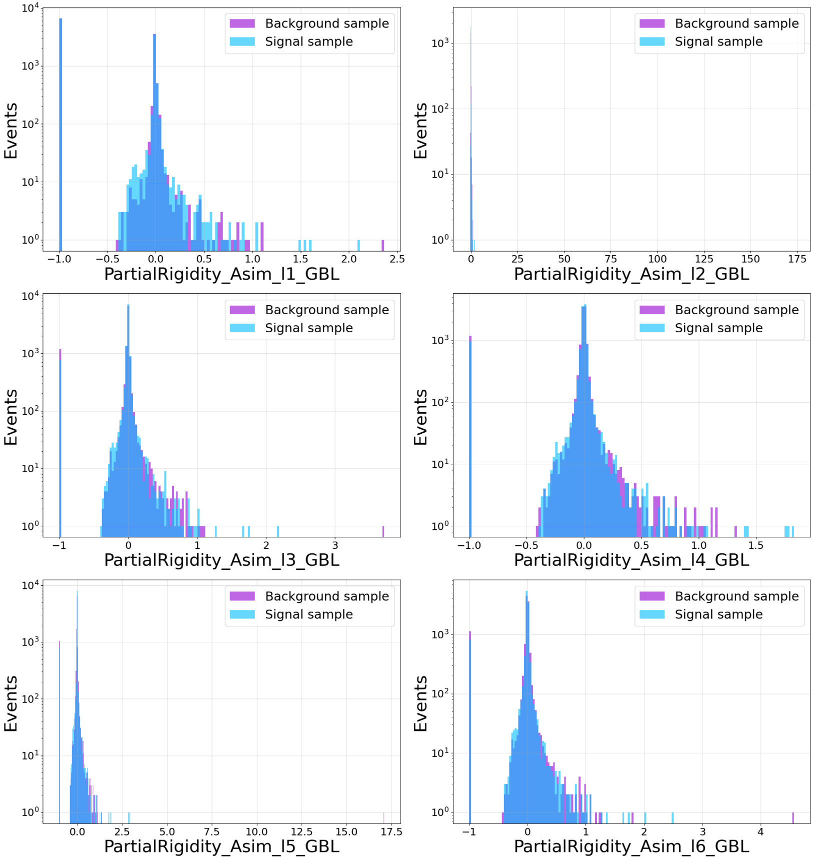


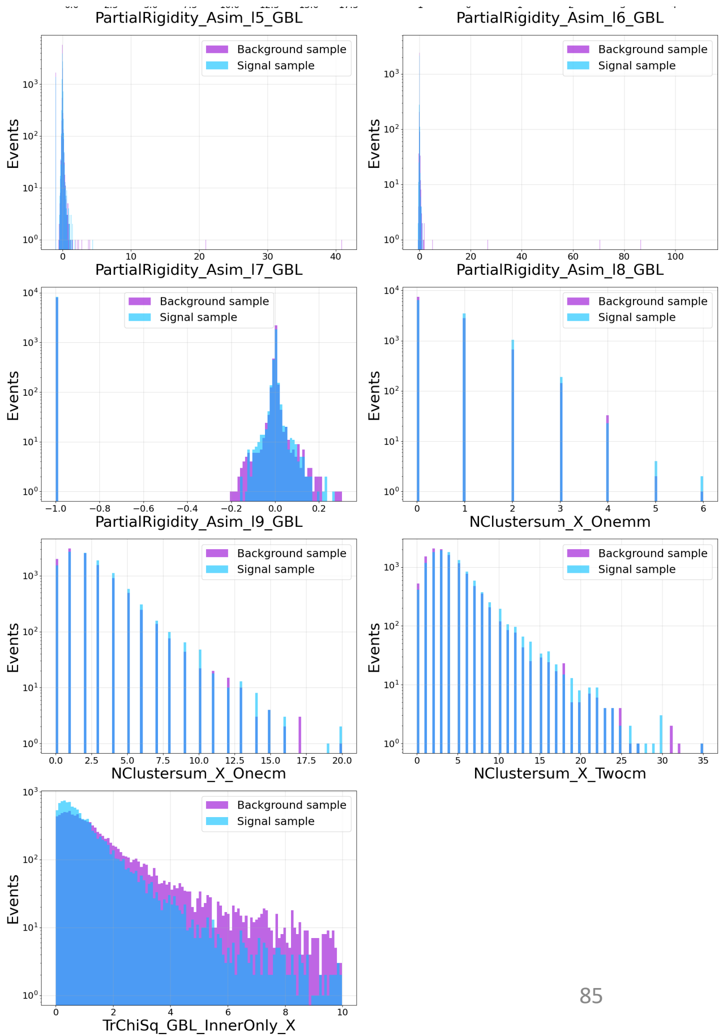
Bologna group features

	PartialRigidity_Asim_l1_GBL	PartialRigidity_Asim_l2_GBL	PartialRigidity_Asim_l3_GBL	PartialRigidity_Asim_l4_GBL	PartialRigidity_Asim_l5_GBL	PartialRigidity_Asim_l6_GBL	PartialRigidity_Asim_l7_GBL	PartialRigidity_Asim_l8_GBL	PartialRigidity_Asim_l9_GBL	NClustersum_X_Onemm	NClustersum_X_Onecm	NClustersum_X_Twocm	TrChiSq_GBL_InnerOnly_X	Rigidity_GBL_InnerOnly
PartialRigidity_Asim_l1_GBL	1.00	0.15	0.22	0.21	0.20	0.17	0.26	0.27	0.14	-0.06	-0.03	-0.02	0.06	0.02
PartialRigidity_Asim_l2_GBL	0.15	1.00	0.46	0.46	0.45	0.41	0.47	0.51	0.01	0.02	0.01	0.01	-0.05	-0.15
PartialRigidity_Asim_l3_GBL	0.22	0.46	1.00	0.46	0.49	0.45	0.61	0.64	0.02	-0.07	-0.05	-0.03	-0.00	-0.05
PartialRigidity_Asim_l4_GBL	0.21	0.46	0.46	1.00	0.50	0.44	0.64	0.67	0.02	-0.09	-0.05	-0.03	0.00	-0.04
PartialRigidity_Asim_l5_GBL	0.20	0.45	0.49	0.50	1.00	0.35	0.57	0.60	0.03	-0.07	-0.05	-0.03	-0.01	-0.04
PartialRigidity_Asim_l6_GBL	0.17	0.41	0.45	0.44	0.35	1.00	0.50	0.51	0.04	-0.07	-0.04	-0.04	-0.00	-0.04
PartialRigidity_Asim_l7_GBL	0.26	0.47	0.61	0.64	0.57	0.50	1.00	0.88	0.00	-0.06	-0.03	-0.03	-0.01	-0.05
PartialRigidity_Asim_l8_GBL	0.27	0.51	0.64	0.67	0.60	0.51	0.88	1.00	0.03	-0.02	-0.01	-0.01	-0.02	-0.05
PartialRigidity_Asim_l9_GBL	0.14	0.01	0.02	0.02	0.03	0.04	0.00	0.03	1.00	-0.02	-0.01	-0.01	0.03	0.04
NClustersum_X_Onemm	-0.06	0.02	-0.07	-0.09	-0.07	-0.07	-0.06	-0.02	-0.02	1.00	0.50	0.38	-0.01	-0.14
NClustersum_X_Onecm	-0.03	0.01	-0.05	-0.05	-0.05	-0.04	-0.03	-0.01	-0.01	0.50	1.00	0.84	-0.00	-0.09
NClustersum_X_Twocm	-0.02	0.01	-0.03	-0.03	-0.03	-0.04	-0.03	-0.01	-0.01	0.38	0.84	1.00	-0.01	-0.07
TrChiSq_GBL_InnerOnly_X	0.06	-0.05	-0.00	0.00	-0.01	-0.00	-0.01	-0.02	0.03	-0.01	-0.00	-0.01	1.00	0.20
Rigidity_GBL_InnerOnly	0.02	-0.15	-0.05	-0.04	-0.04	-0.04	-0.05	-0.05	0.04	-0.14	-0.09	-0.07	0.20	1.00

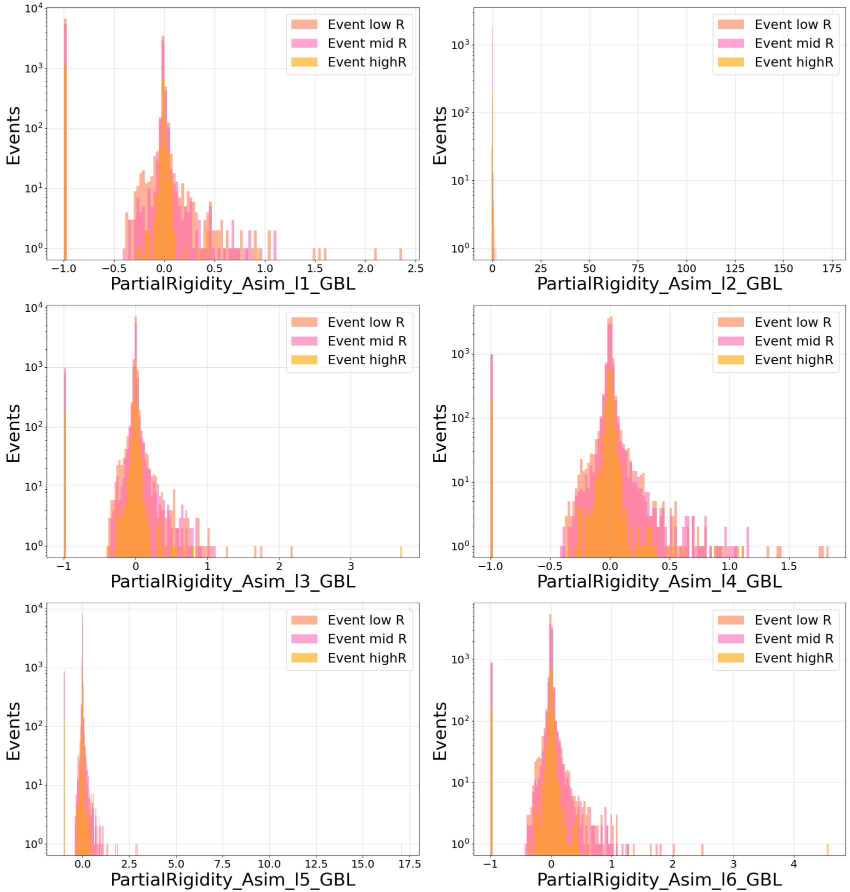
$$PartialRigidity\_Asim\_l_i\_GBL = \frac{(PartialRigidityl_i\_GBL - reconstructedR)}{reconstructedR}$$

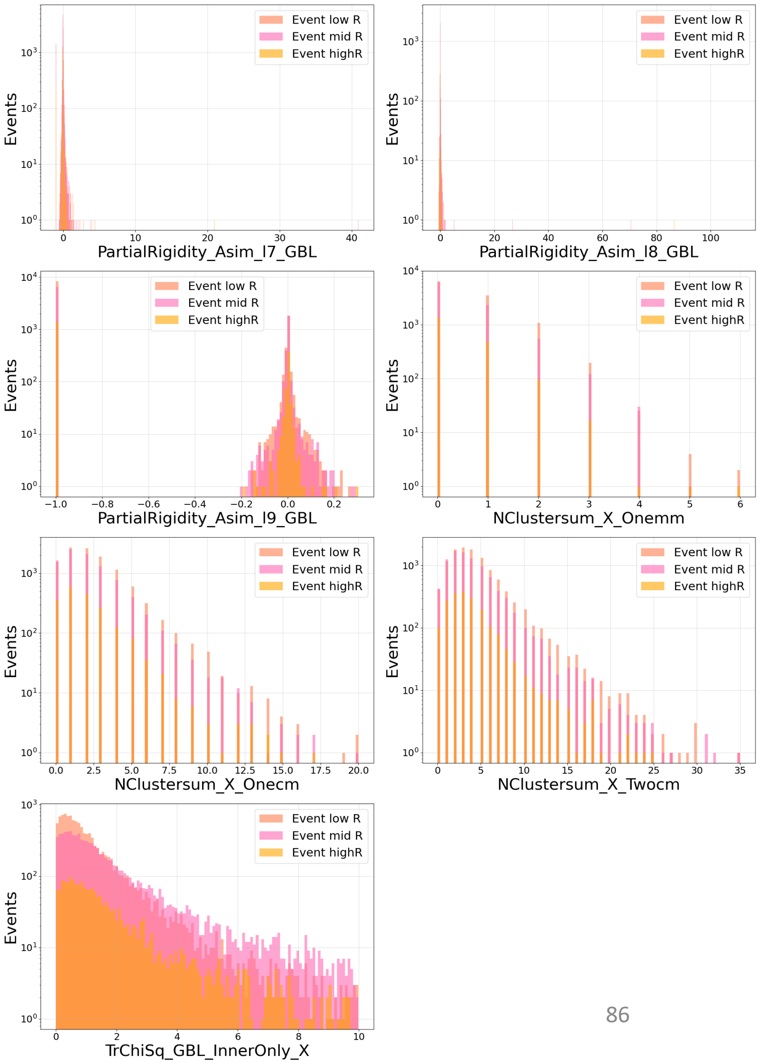
Bologna group features





Bologna group features





Bologna group features