中国科学技术大学
University of Science and Technology of China

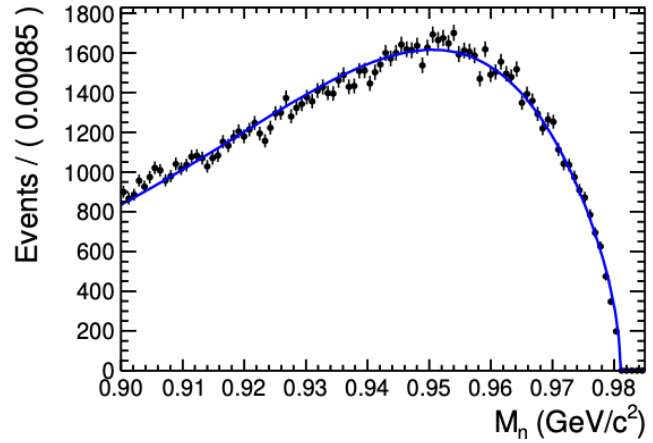
BDT test in Xi analysis

Liu Liang

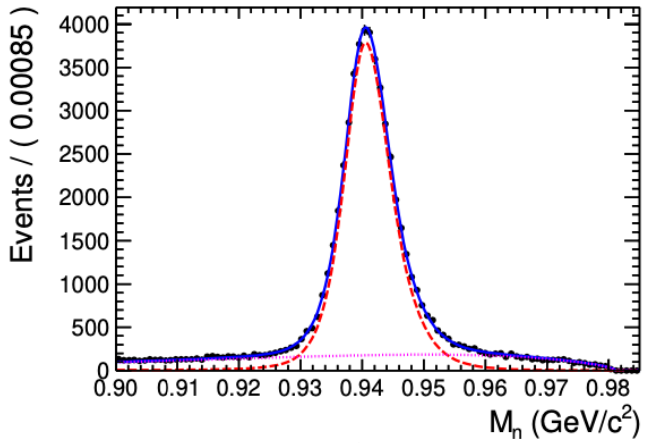
liangzy@mail.ustc.edu.cn

Motivation

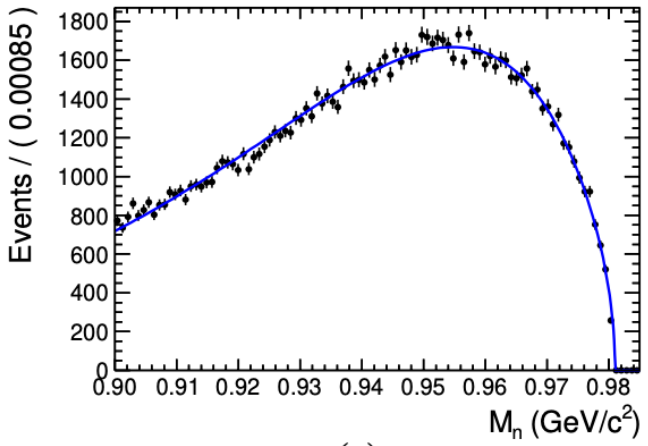
- High background level in the signal region of neutron anti-neutron invariant mass.
- It might have an opportunity to be improved by applying an MVA method.



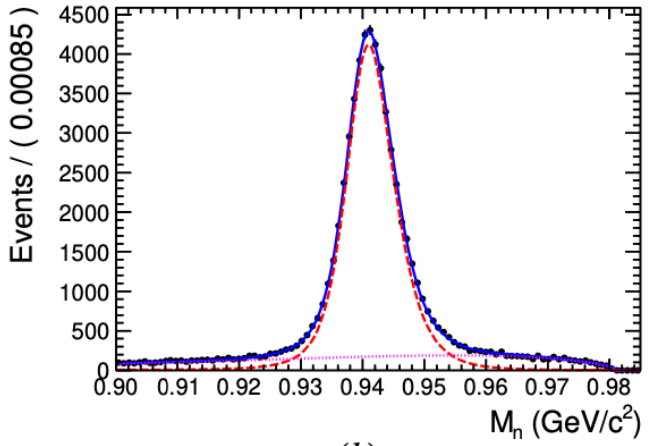
(a)



(b)



(a)



(b)

Motivation

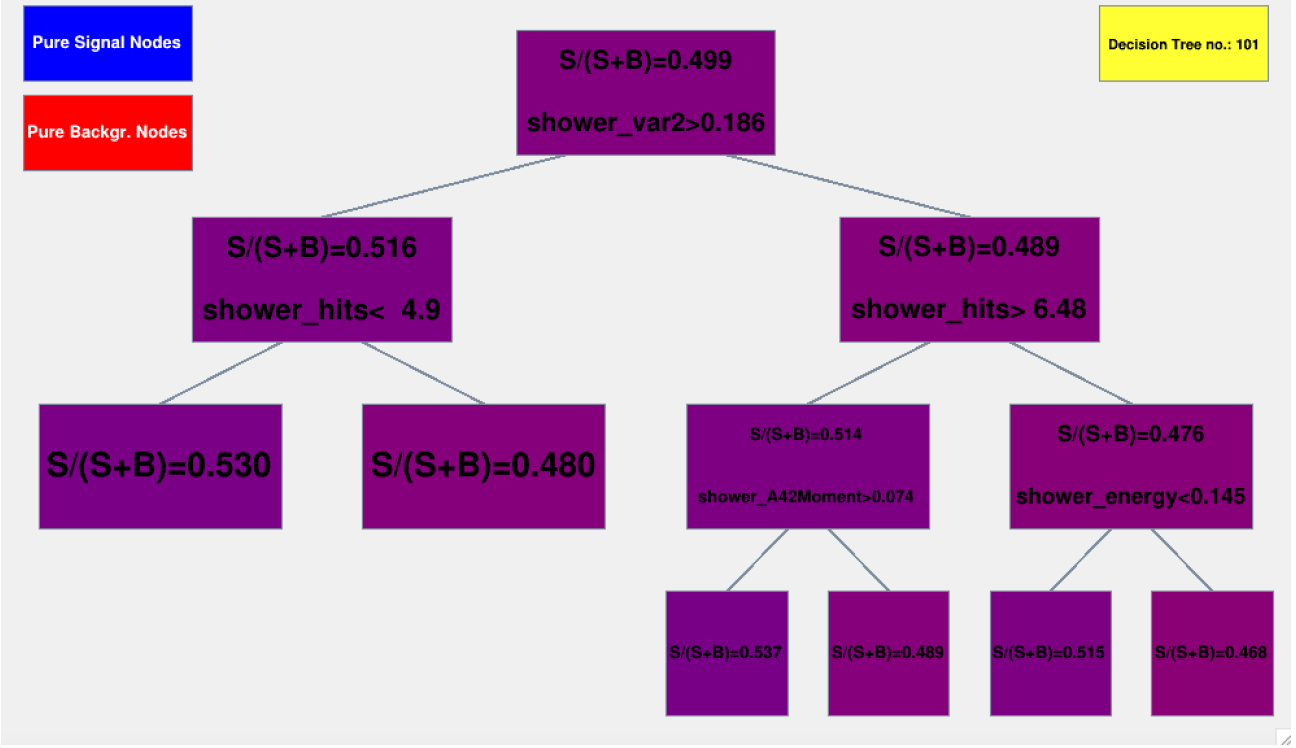
	Neutron Channel	Anti-neutron channel
2009	2461 ± 56	2946 ± 61
2012	12862 ± 127	13217 ± 127
2018	49162 ± 250	55638 ± 264
2019	46409 ± 246	51507 ± 252

	Neutron channel	Anti-neutron channel
2009	405±40	377±41
2012	2020±92	1754±86
2018	7991±182	7589±184
2019	7213±304	7246±178

The background level is about ~15%

Brief introduction of BDT

- Decision tree: a binary tree structured classifier (regressor).
- The phase space is split this way into many regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final leaf node.



Brief introduction of BDT

- The boosting of a decision tree extends this concept from one tree to several trees which form a forest.
- The boosted event classifi- cation yBoost(x) is then given by

$$y_{\mathrm{Boost}}(\mathbf{x}) = \frac{1}{N_{\mathrm{collection}}} \cdot \sum_{i}^{N_{\mathrm{collection}}} \ln(\alpha_i) \cdot h_i(\mathbf{x}) \,,$$

for signal and background as h(x) = +1 and − 1, respectively.
The boost weight is derived from the misclassification rate, err, of the previous tree

$$\alpha = \frac{1 - \mathrm{err}}{\mathrm{err}} \,.$$

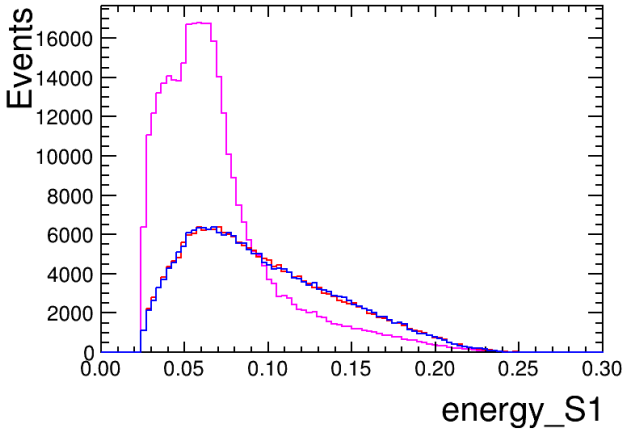
AdaBoost method, Gradient Boosting method.

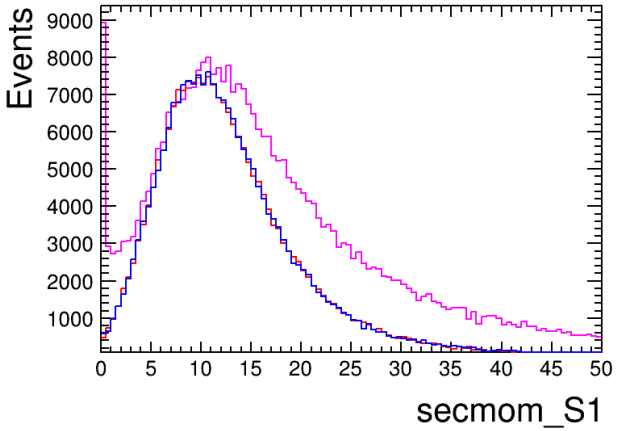
Variables of photon shower

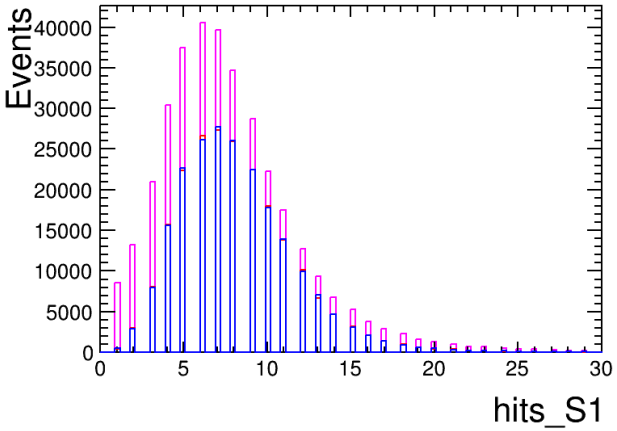
double	energy()	Shower energy after correction.
double	dE()	Error for shower energy.
double	eSeed()	Seed"s energy.
double	e3x3()	Energy sum of 3x3 crystal around seed.
double	e5x5()	Energy sum of 5x5 crystal around seed.
double	secondMoment()	Second moment for shower shape.
double	latMoment()	Lateral moment for shower shape.
double	A20Moment()	Zernike-20 moment for shower shape.
double	A42Moment()	Zernike-42 moment for shower shape.

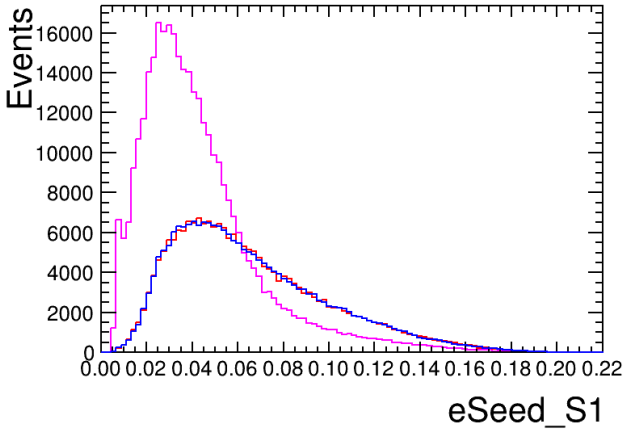
Energy, numHits, secondMoment,
eSeed, e3x3, e5x5,
latMoment, A20Moment, A42Moment

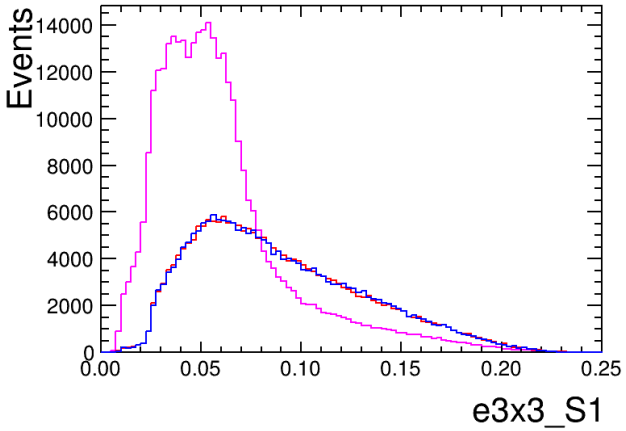
Variables of photon shower

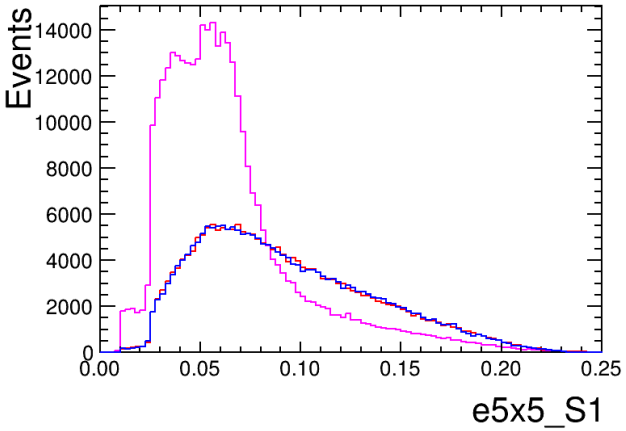




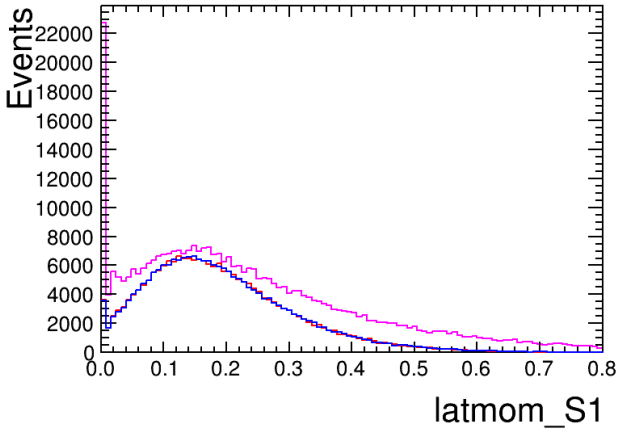


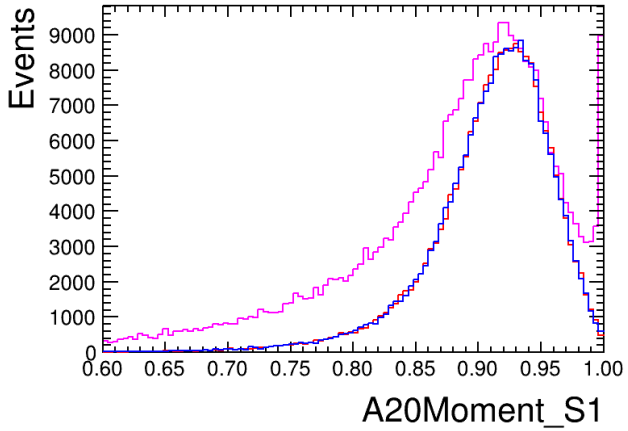


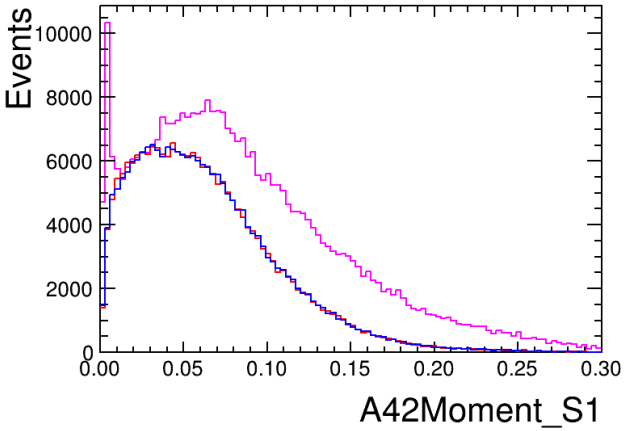


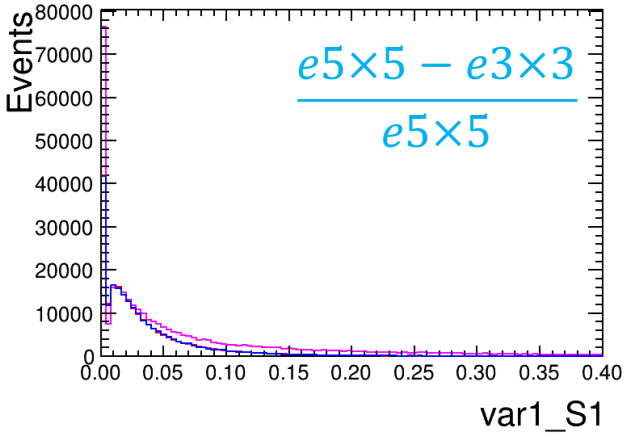


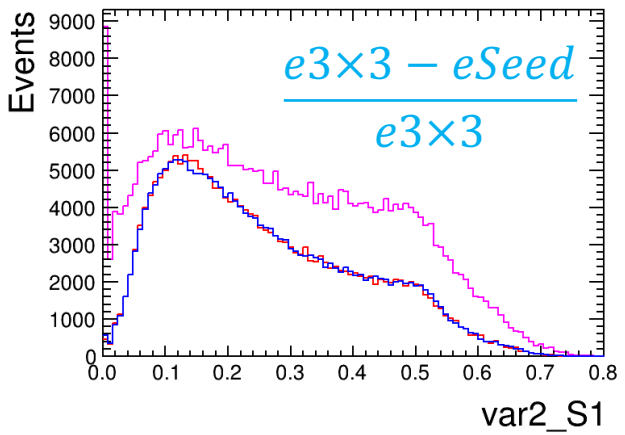
Variables of photon shower

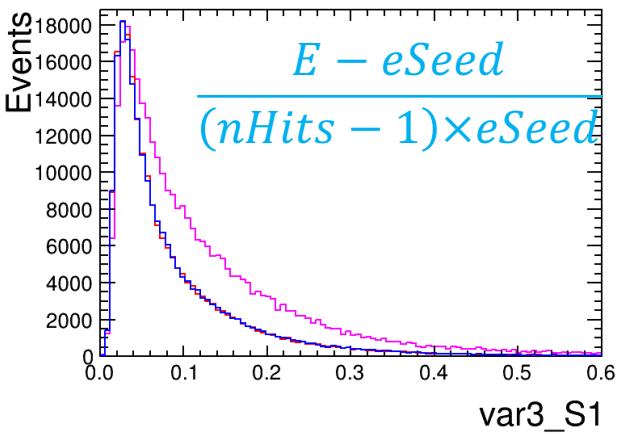




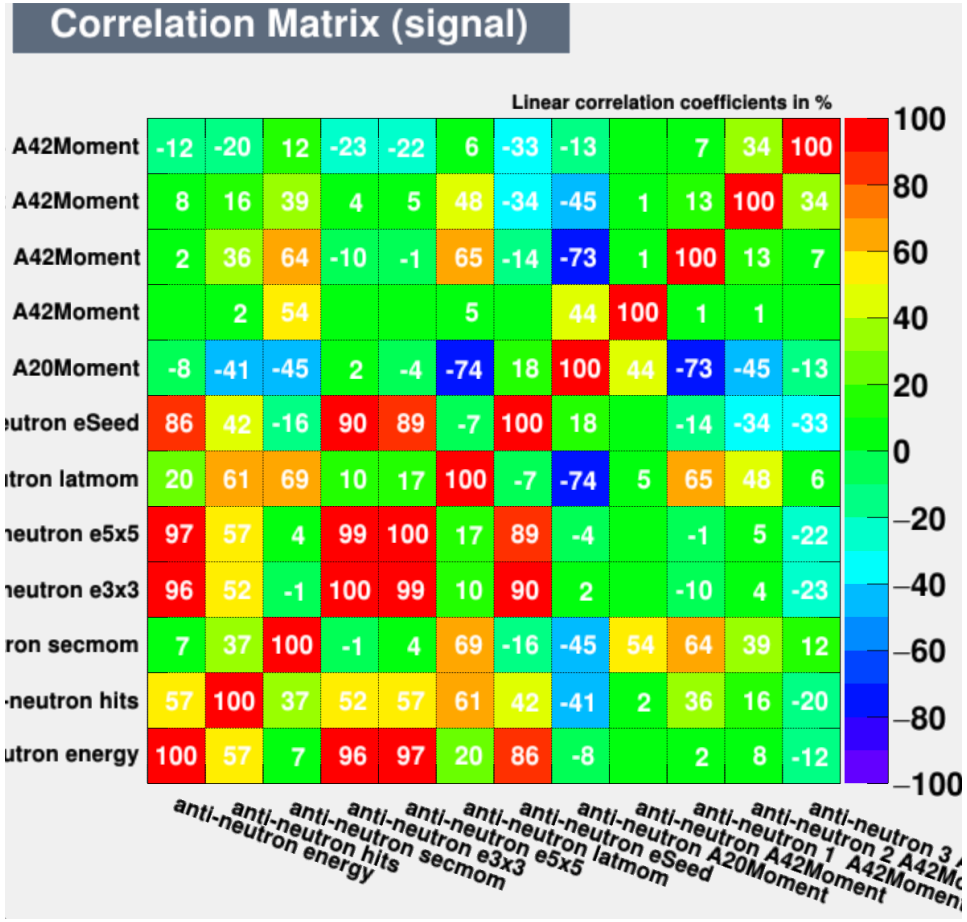


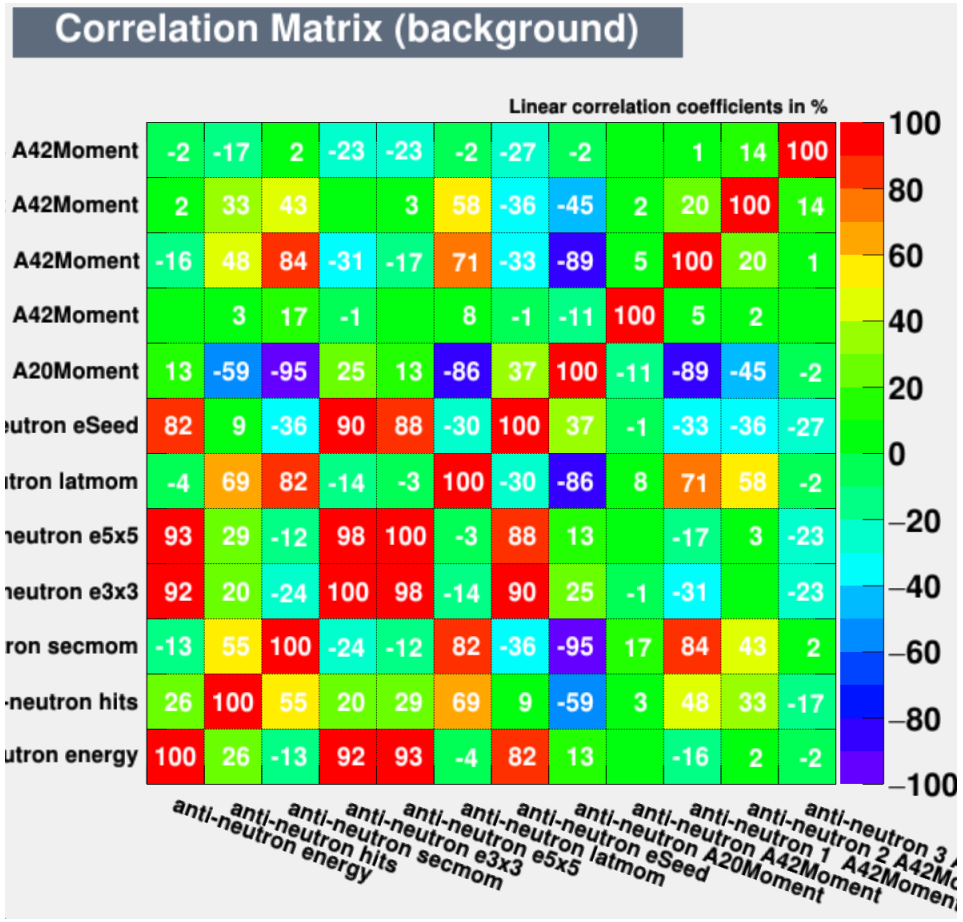




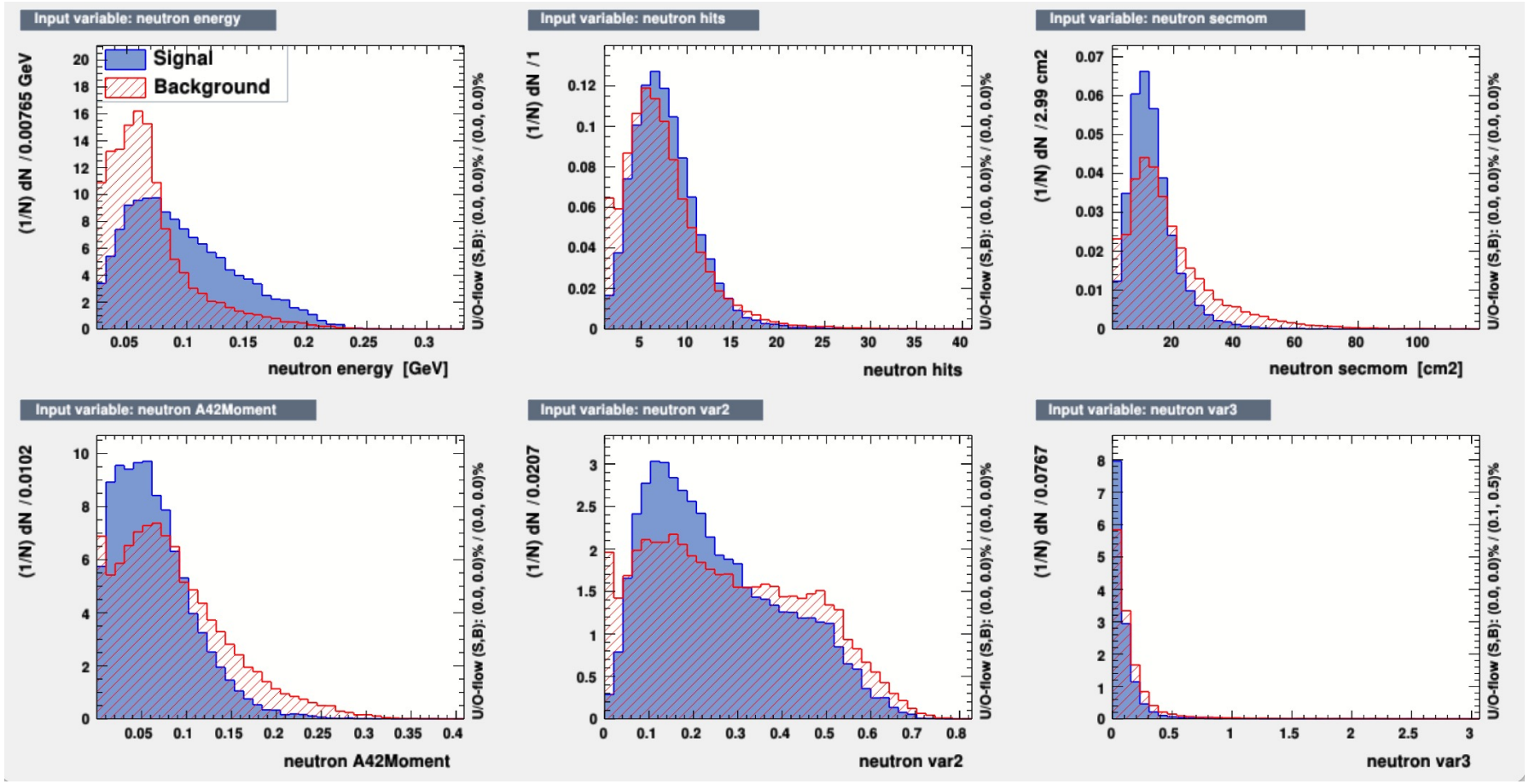


Correlations

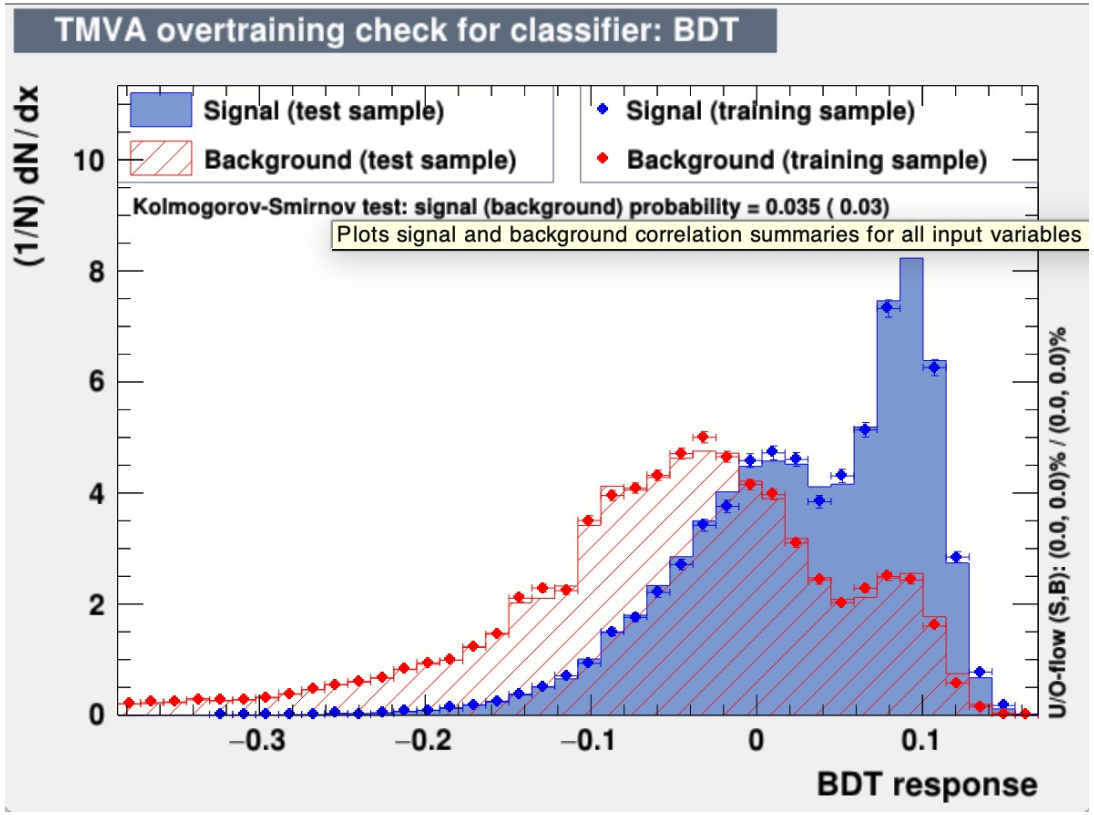


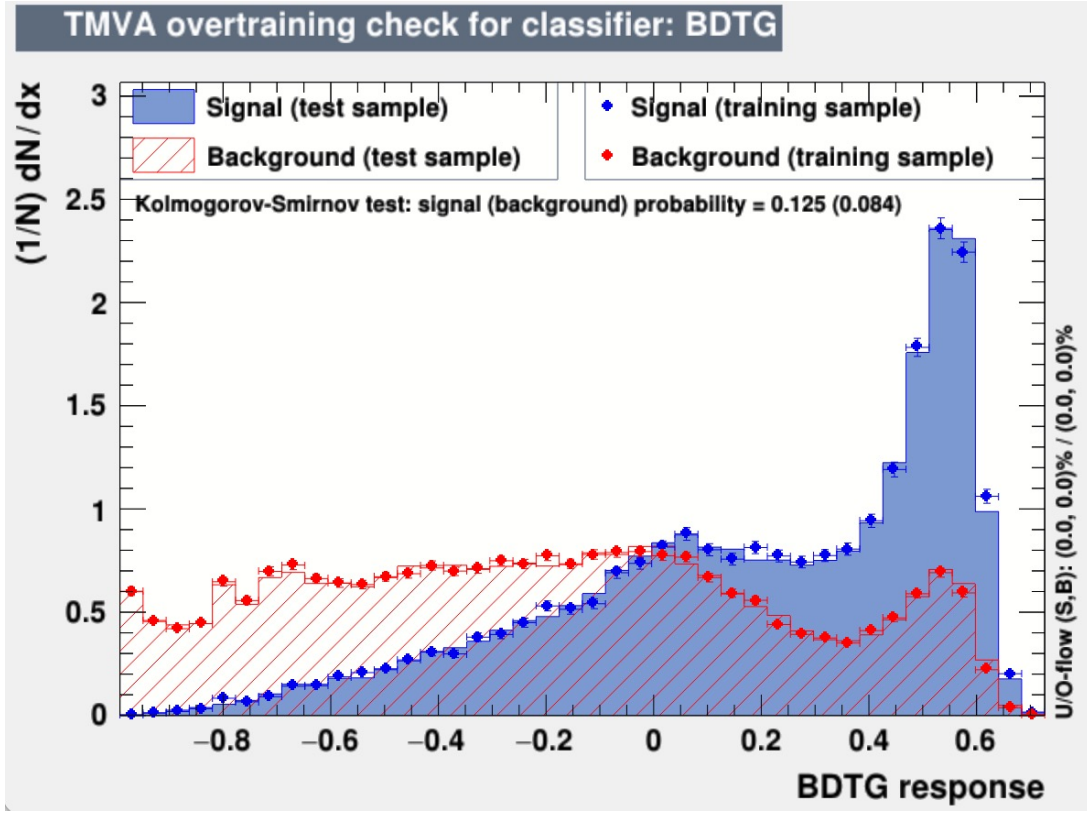


Neutron channel test

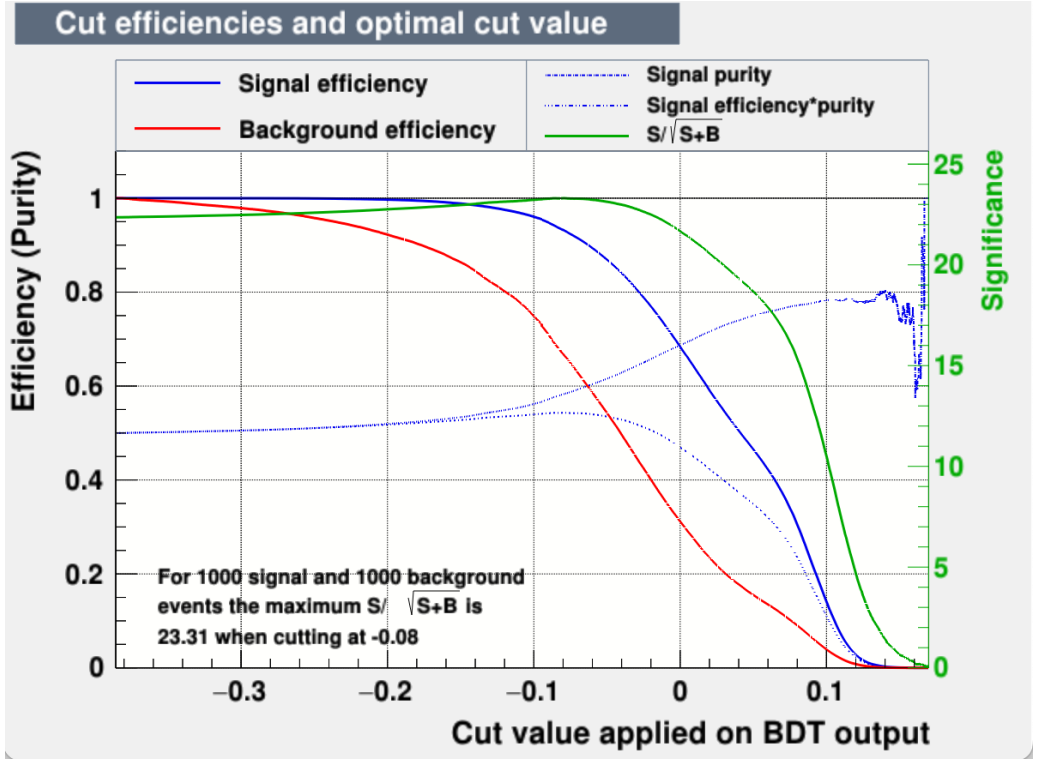


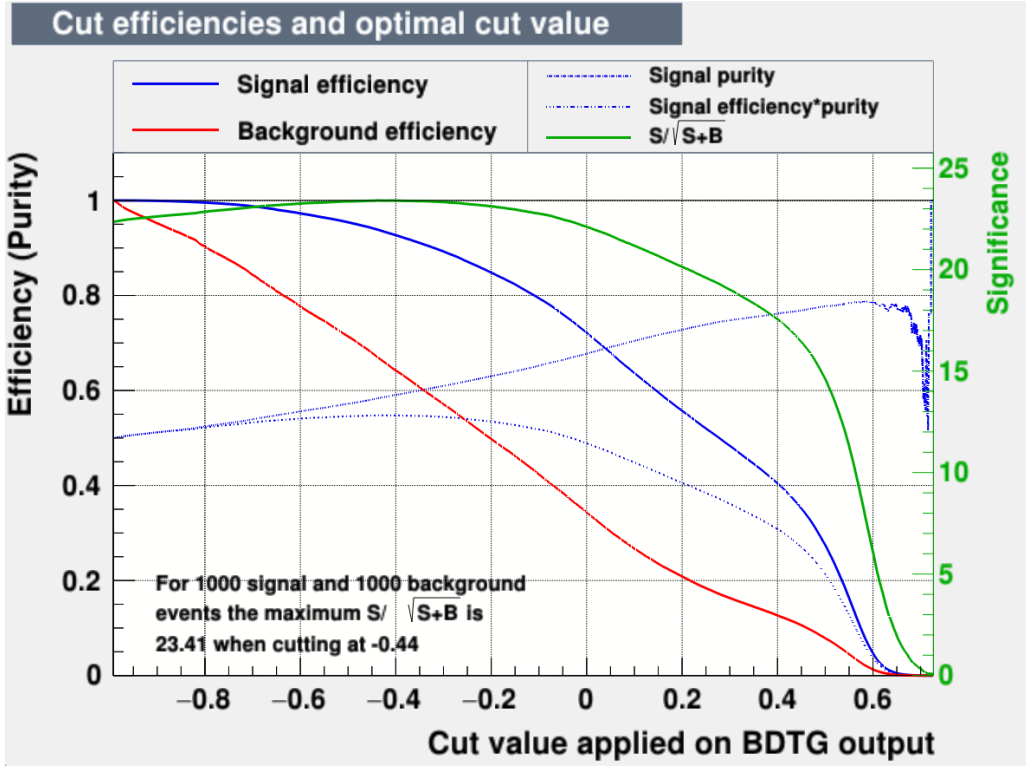
Neutron channel test





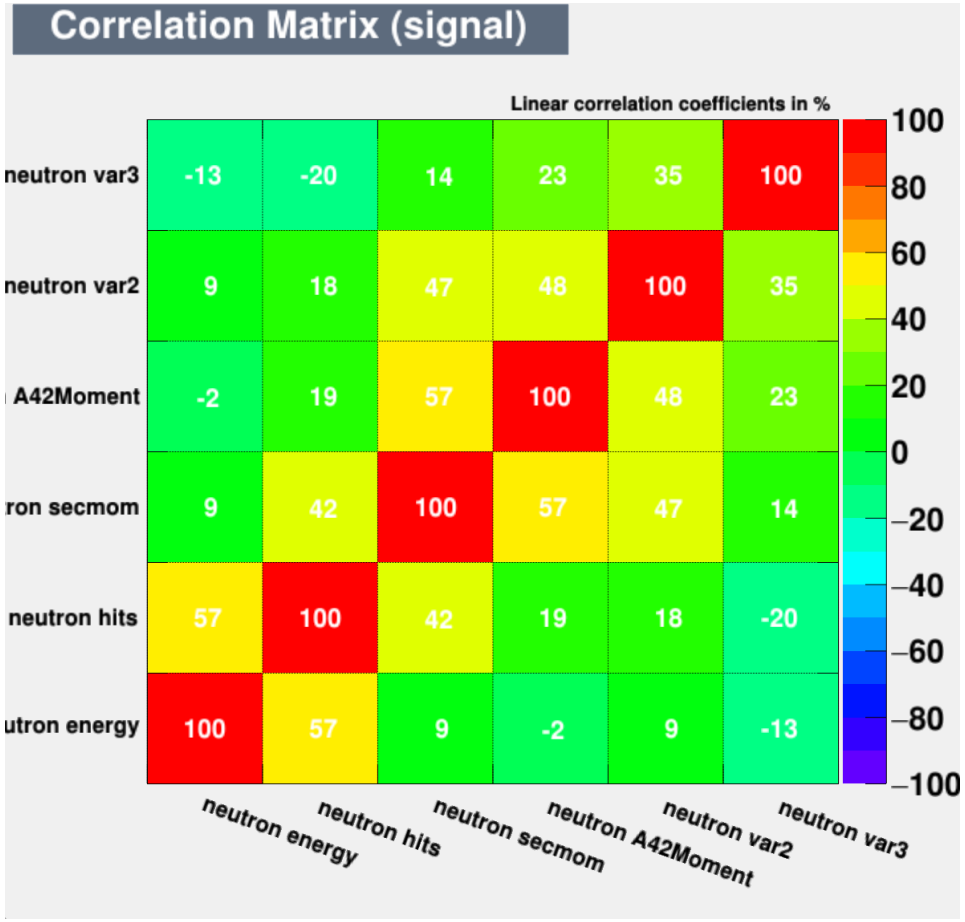
Neutron channel test

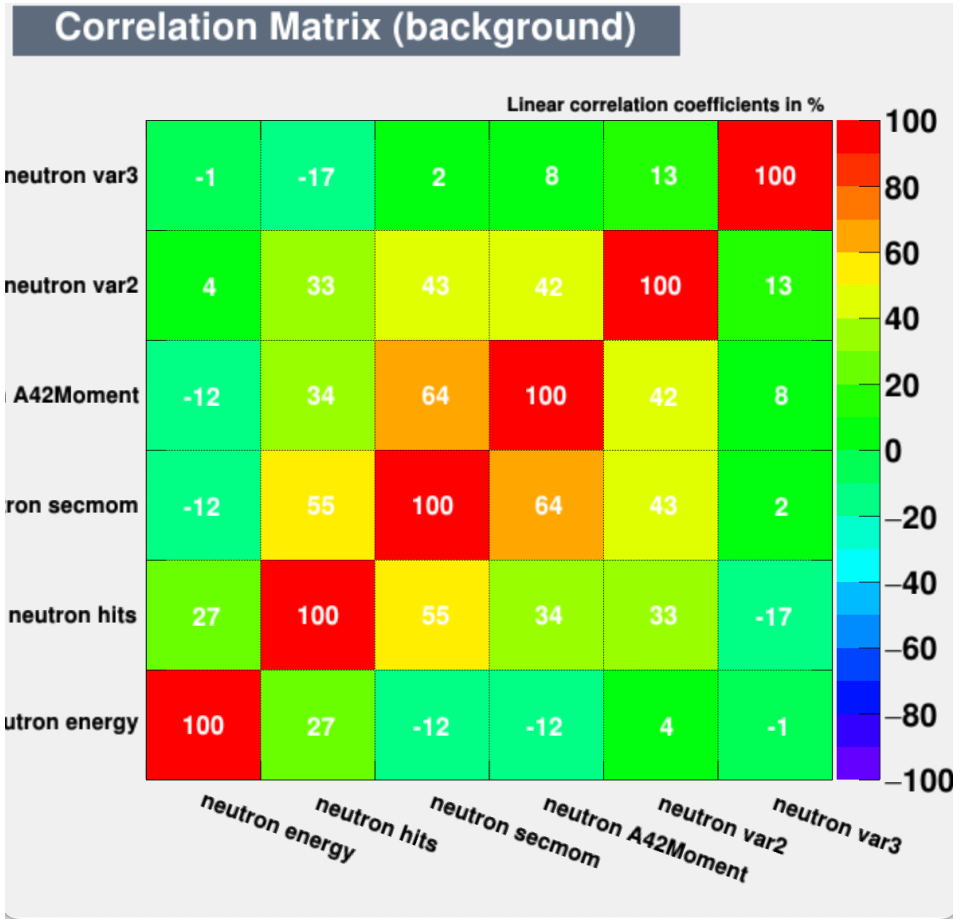




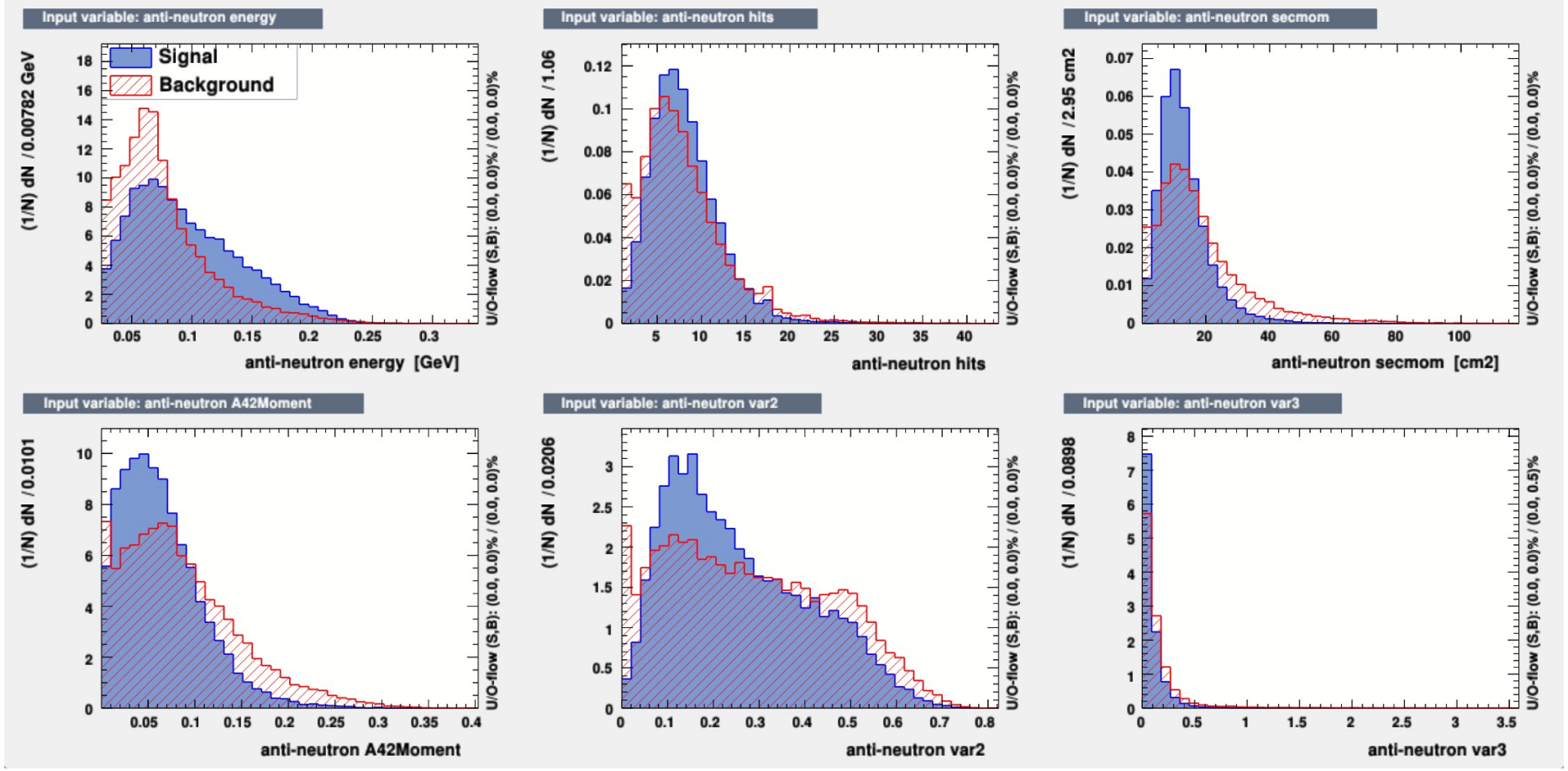

```
--- ============================================================================================
--- Classifier   (  #signal, #backgr.)  Optimal-cut  S/sqrt(S+B)      NSig      NBkg    EffSig    EffBkg
--- --------------------------------------------------------------------------------------------
---       BDTG:  (     1000,     1000)      -0.4373        23.408  938.5034  668.9654    0.9385     0.669
---        BDT:  (     1000,     1000)      -0.0794        23.309  931.6928  666.0246    0.9317     0.666
--- --------------------------------------------------------------------------------------------
```

Neutron channel test

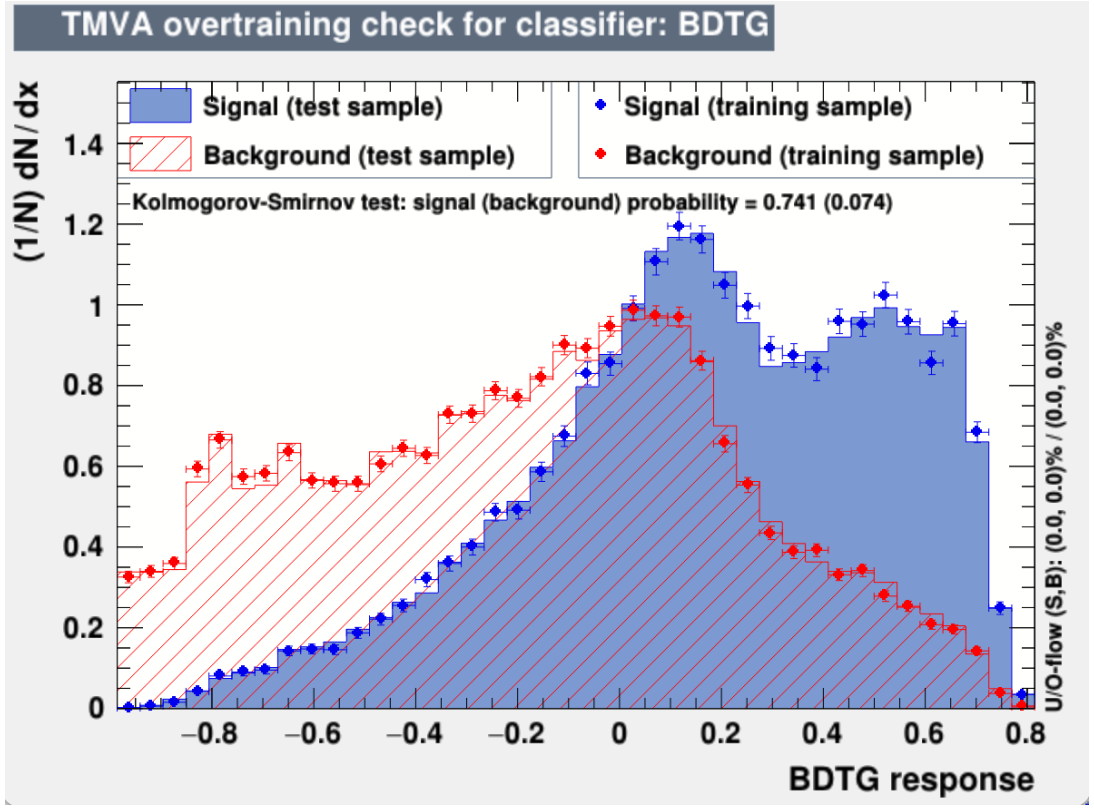


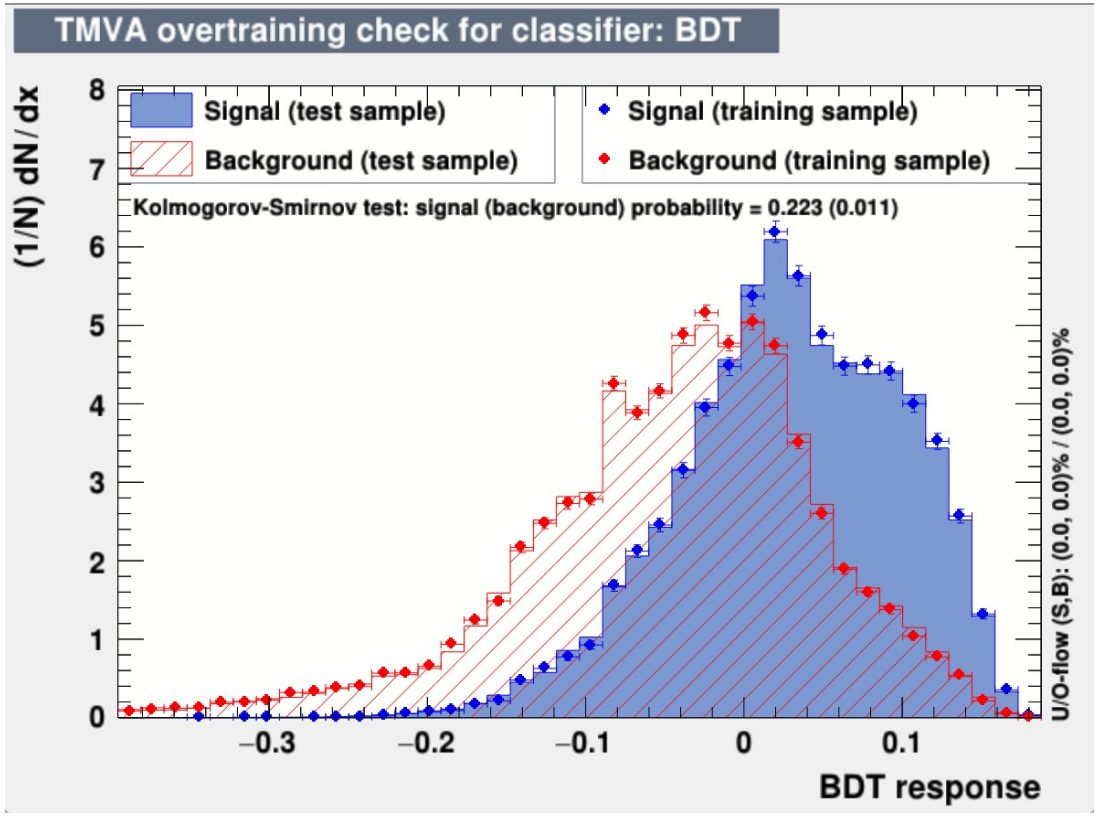


Anti-neutron channel

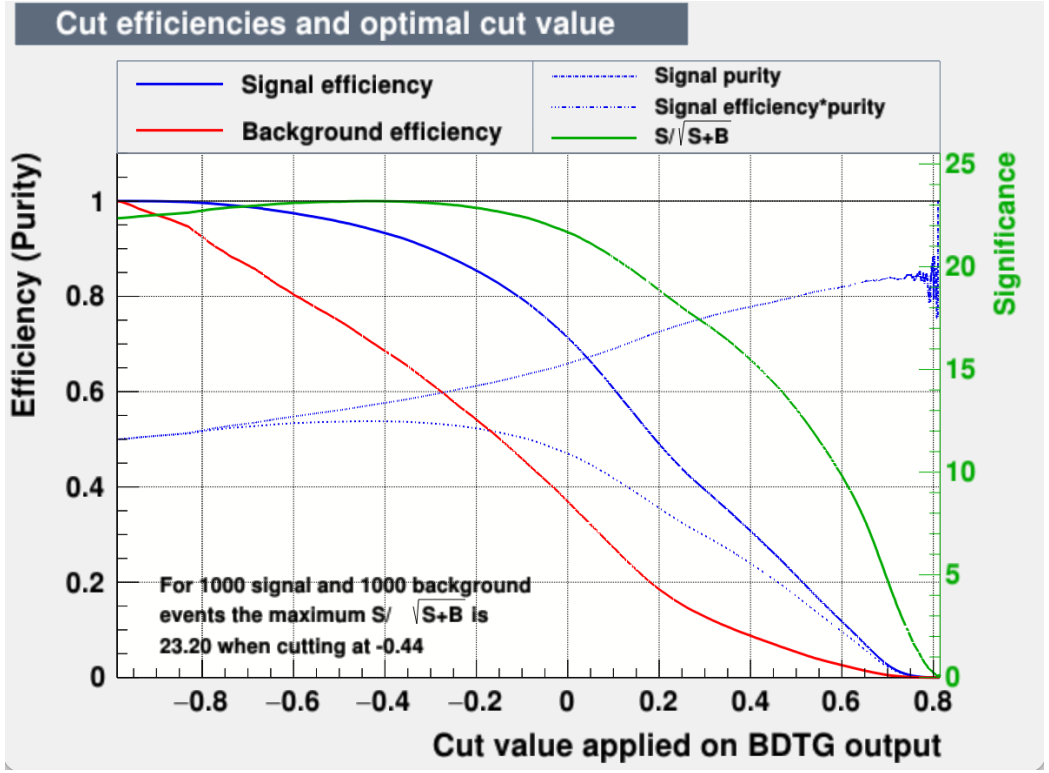


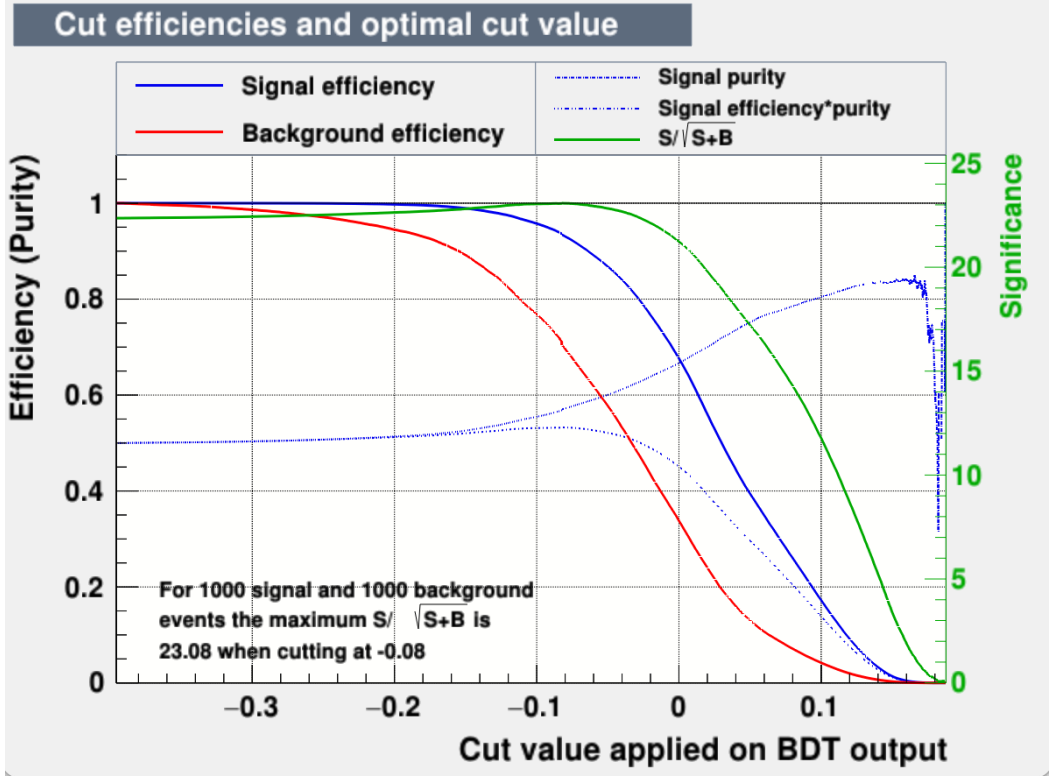
Anti-neutron channel





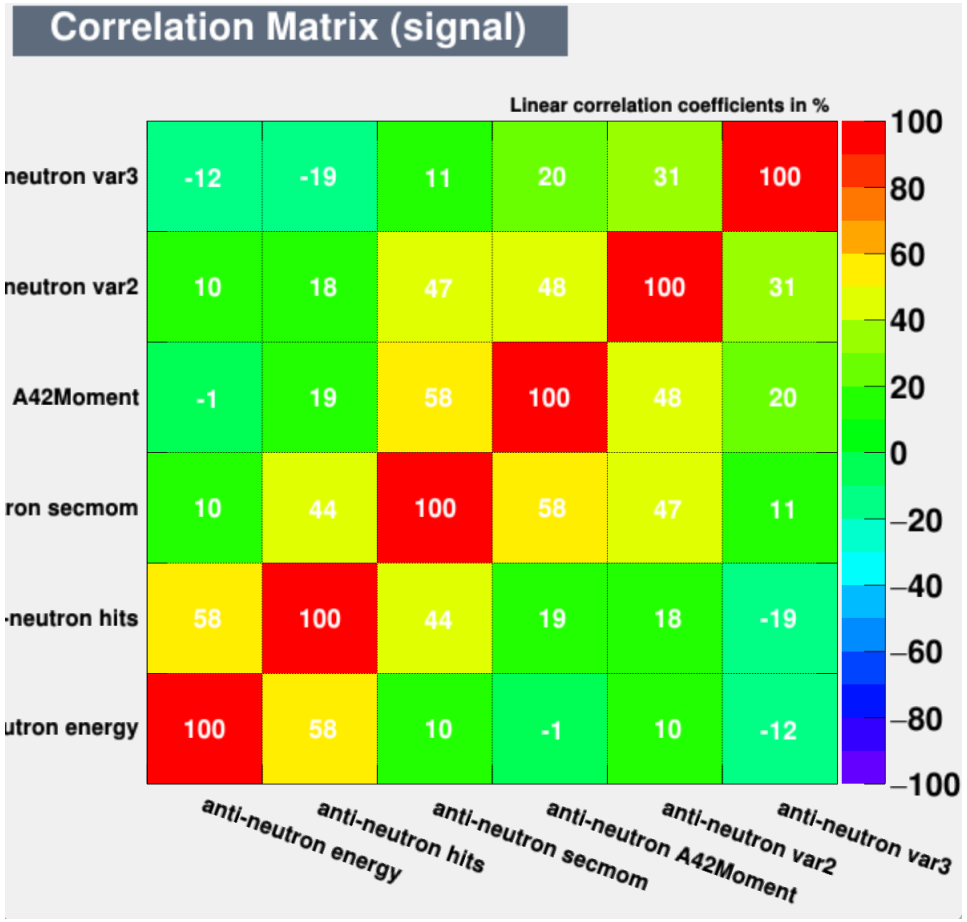
Anti-neutron channel

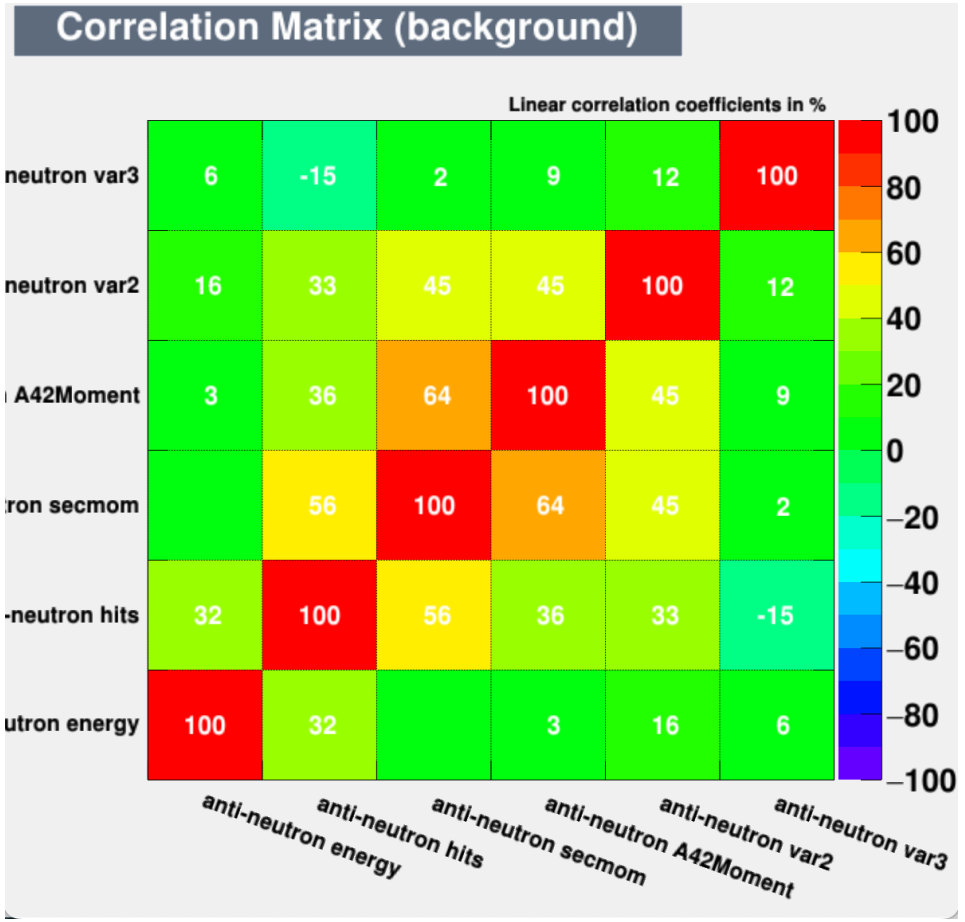





```
--- ==========================================================================================
--- Classifier   (  #signal, #backgr.)  Optimal-cut  S/sqrt(S+B)      NSig      NBkg    EffSig    EffBkg
--- ------------------------------------------------------------------------------------------
---       BDTG:  (     1000,     1000)      -0.4409       23.1972   943.664   711.202    0.9437    0.7112
---        BDT:  (     1000,     1000)      -0.0824        23.082  934.3693  704.2983    0.9344    0.7043
--- ------------------------------------------------------------------------------------------
```

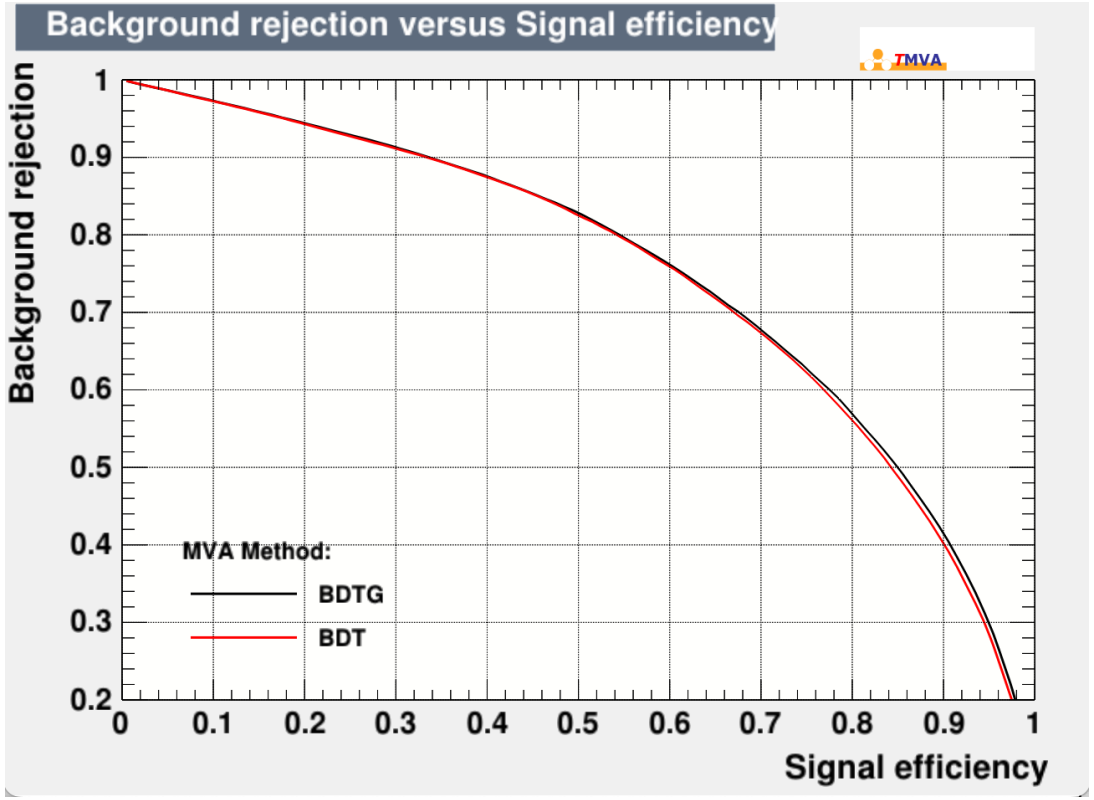
Anti-neutron channel



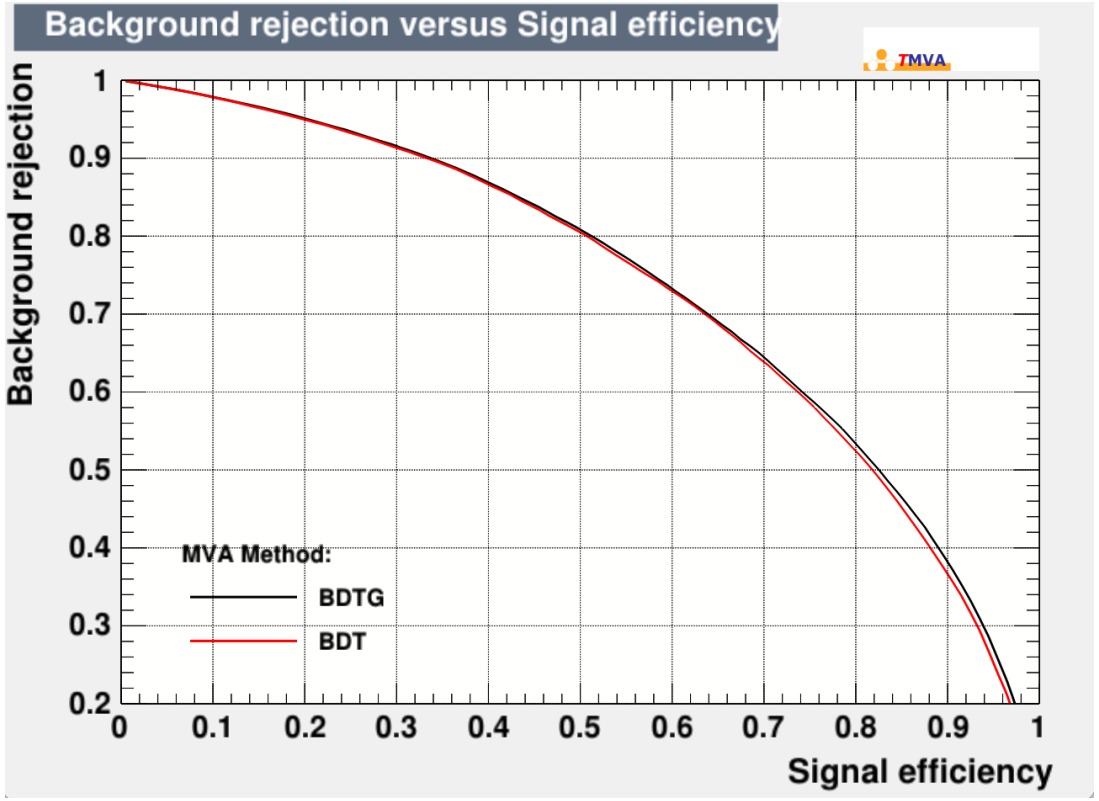


RejB vs EffS

Neutron channel



Anti-neutron channel



Backup

Neutron channel

```
                         .
                         : Ranking input variables (method specific)...
BDTG                     : Ranking result (top variable is best ranked)
                         : ----------------------------------------------
                         : Rank : Variable         : Variable Importance
                         : ----------------------------------------------
                         :    1 : shower_var2      : 2.149e-01
                         :    2 : shower_energy    : 1.981e-01
                         :    3 : shower_A42Moment : 1.753e-01
                         :    4 : shower_hits      : 1.649e-01
                         :    5 : shower_secmom    : 1.535e-01
                         :    6 : shower_var3      : 9.323e-02
                         : ----------------------------------------------
BDT                      : Ranking result (top variable is best ranked)
                         : ----------------------------------------------
                         : Rank : Variable         : Variable Importance
                         : ----------------------------------------------
                         :    1 : shower_energy    : 2.088e-01
                         :    2 : shower_var2      : 1.974e-01
                         :    3 : shower_secmom    : 1.921e-01
                         :    4 : shower_A42Moment : 1.821e-01
                         :    5 : shower_hits      : 1.571e-01
                         :    6 : shower_var3      : 6.249e-02
                         : ----------------------------------------------
```

Anti-neutron channel

```
                         :
                         : Ranking input variables (method specific)...
BDTG                     : Ranking result (top variable is best ranked)
                         : ----------------------------------------------
                         : Rank : Variable         : Variable Importance
                         : ----------------------------------------------
                         :    1 : shower_var2      : 2.149e-01
                         :    2 : shower_energy    : 1.944e-01
                         :    3 : shower_A42Moment : 1.840e-01
                         :    4 : shower_hits      : 1.642e-01
                         :    5 : shower_secmom    : 1.502e-01
                         :    6 : shower_var3      : 9.235e-02
                         : ----------------------------------------------
BDT                      : Ranking result (top variable is best ranked)
                         : ----------------------------------------------
                         : Rank : Variable         : Variable Importance
                         : ----------------------------------------------
                         :    1 : shower_energy    : 2.058e-01
                         :    2 : shower_var2      : 1.978e-01
                         :    3 : shower_A42Moment : 1.861e-01
                         :    4 : shower_secmom    : 1.835e-01
                         :    5 : shower_hits      : 1.668e-01
                         :    6 : shower_var3      : 6.002e-02
                         : ----------------------------------------------
```