BDT update

Background was over used in the last update

05/12/2024

Method

Approach 1

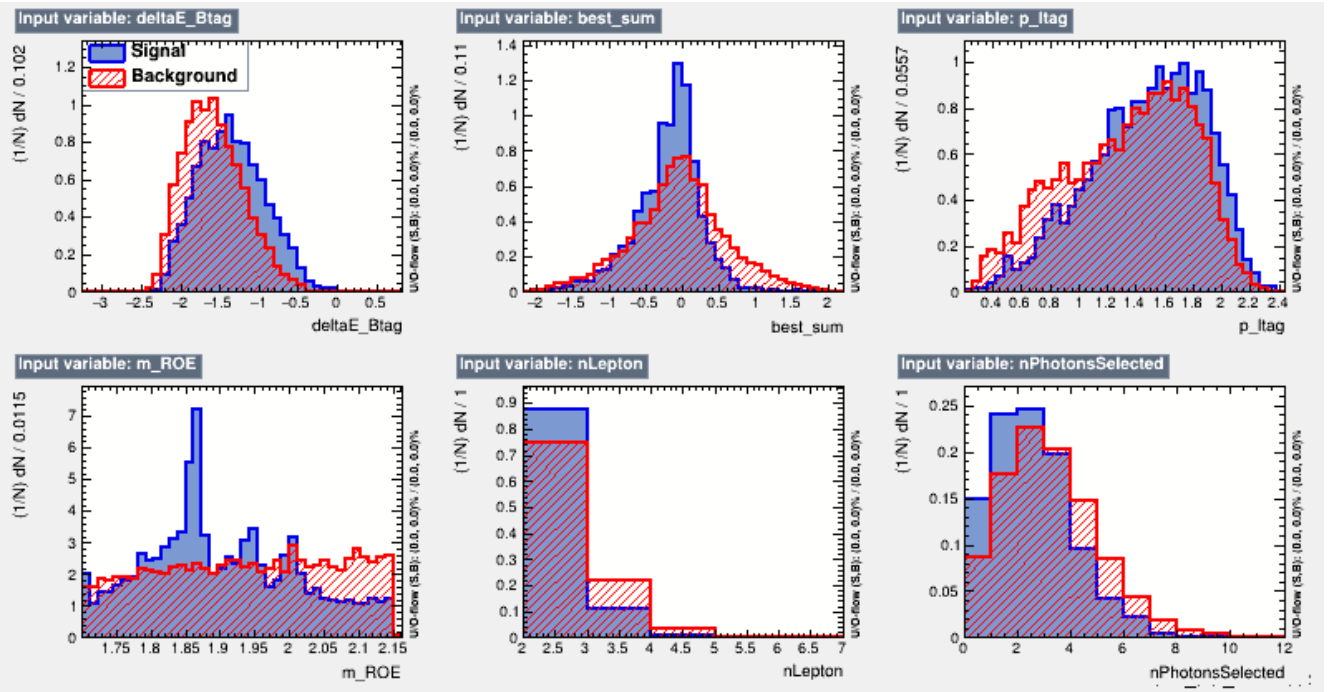
- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

Approach 1

- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.



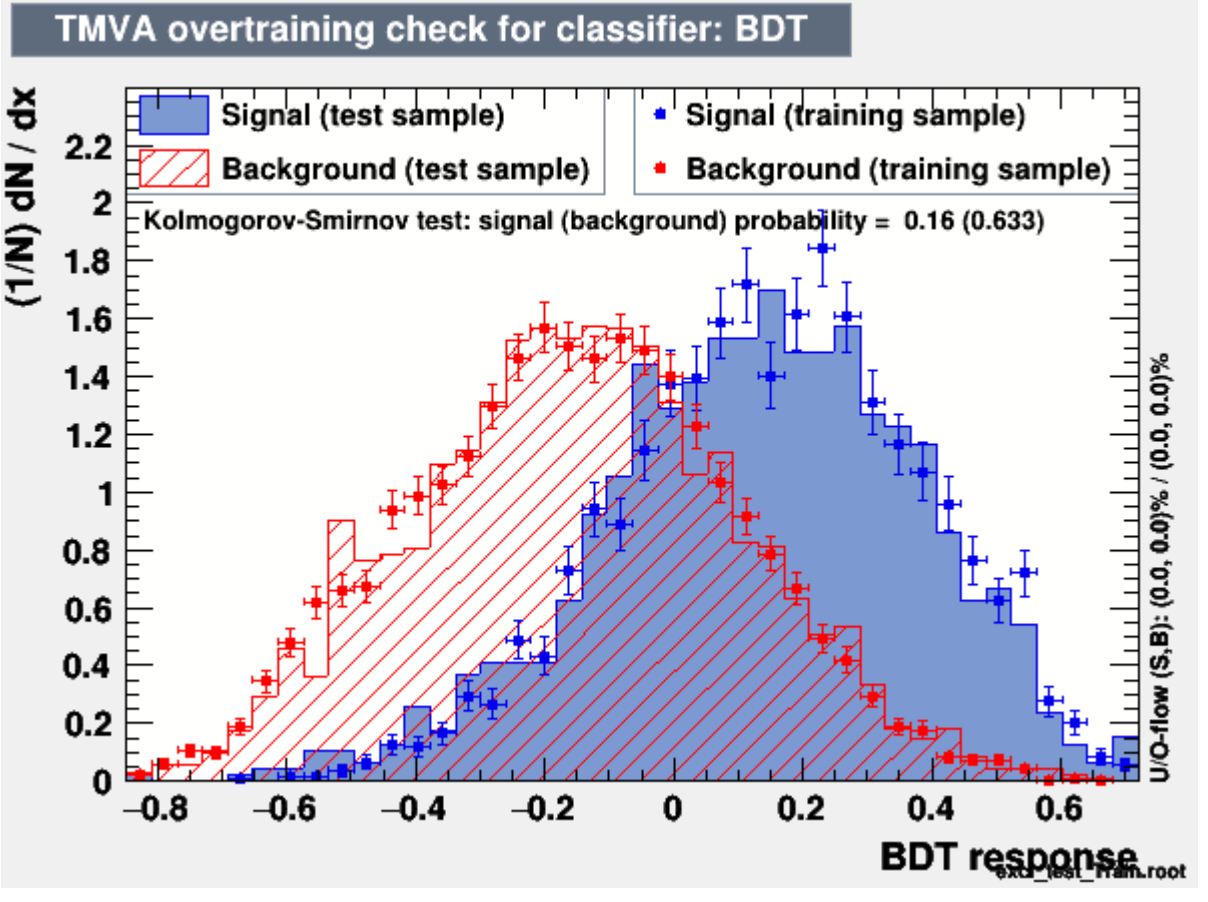
Var. importance

```
Ranking result (top variable is best ranked)
-------------------------------------------
Rank : Variable         : Variable Importance
-------------------------------------------
   1 : best_sum         : 2.390e-01
   2 : deltaE_Btag      : 2.055e-01
   3 : p_ltag           : 1.955e-01
   4 : m_ROE            : 1.817e-01
   5 : nPhotonsSelected : 1.019e-01
   6 : nLepton          : 7.626e-02
```

Approach 1

- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

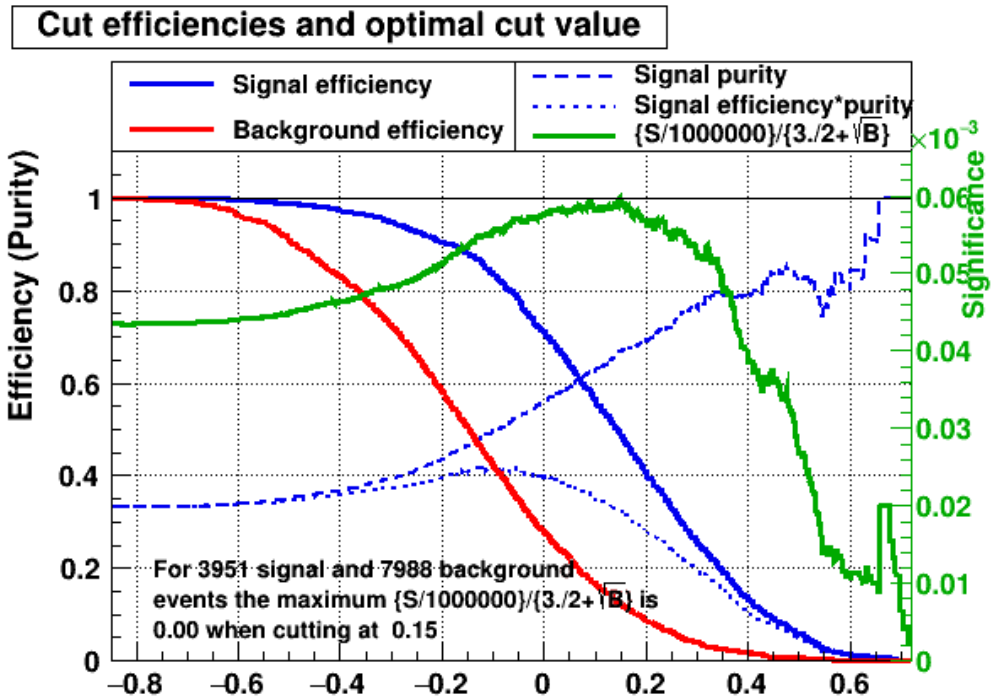
BDT response



Approach 1

- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

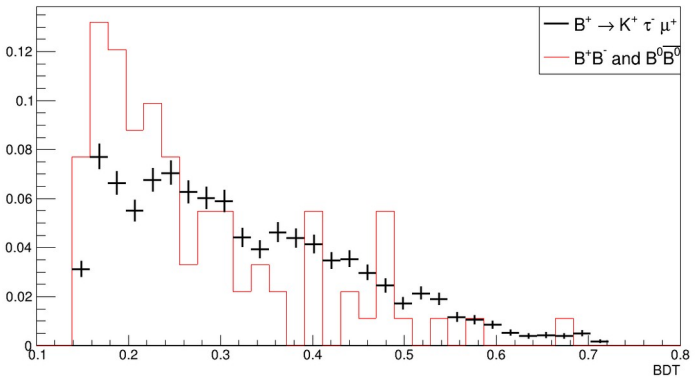
BDT > 0.15



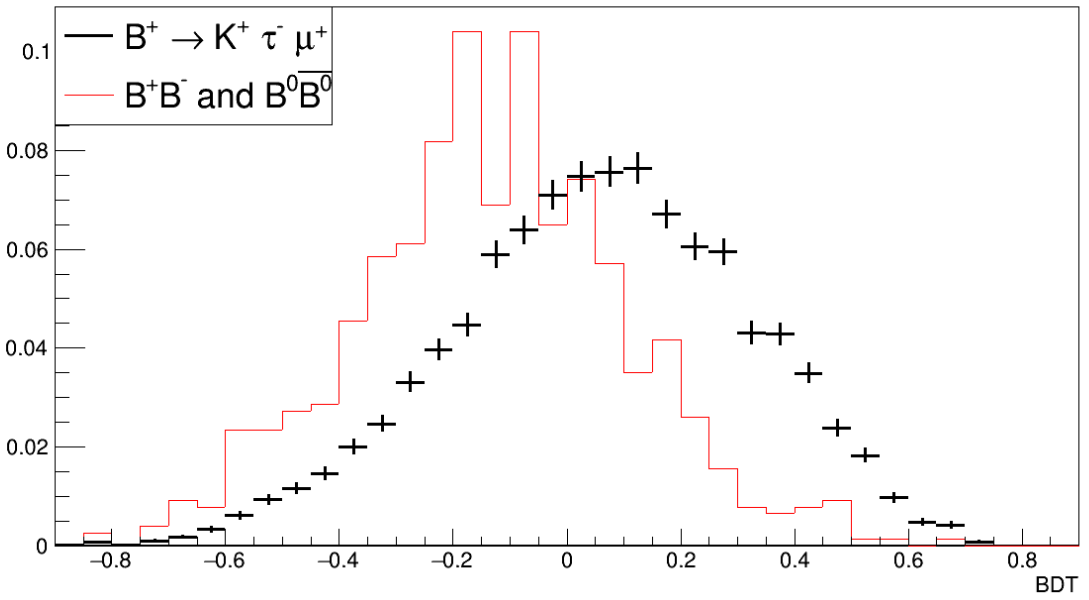
Classifier	(#signal, #backgr.)	Optimal-cut	(S/1000000)/(3./2+sqrt(B))	NSig	NBkg	EffSig	EffBkg
BDT:	(3951, 7988)	0.1489	5.97378e-05	1950.494	970.3792	0.4937	0.1215
BDTG:	(3951, 7988)	0.0804	5.77831e-05	2553.98	1823.238	0.6464	0.2282
Fisher:	(3951, 7988)	0.0329	5.25783e-05	2473.959	2075.067	0.6262	0.2598
MLP:	(3951, 7988)	0.6806	6.00227e-05	1880.476	889.794	0.4759	0.1114

Approach 1

- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.



BDT score

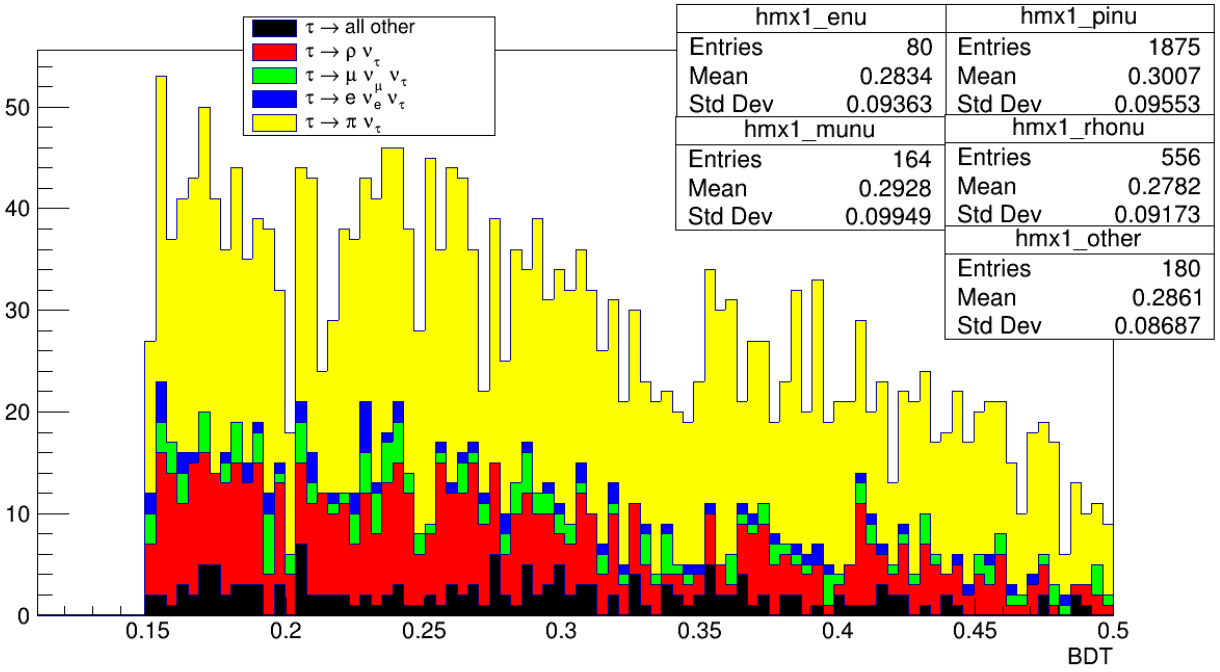


Approach 1

- Training (70%) and testing (30%) on the 1.0 M $\tau \rightarrow \pi$ sample.
- Applying on the 8.9 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

N_{pi}	1875	1
N_{e}	80	0.04
N_{mu}	164	0.08
N_{rho}	556	0.30
N_{others}	180	0.09

BDT score



For 5×10^{-5} BF

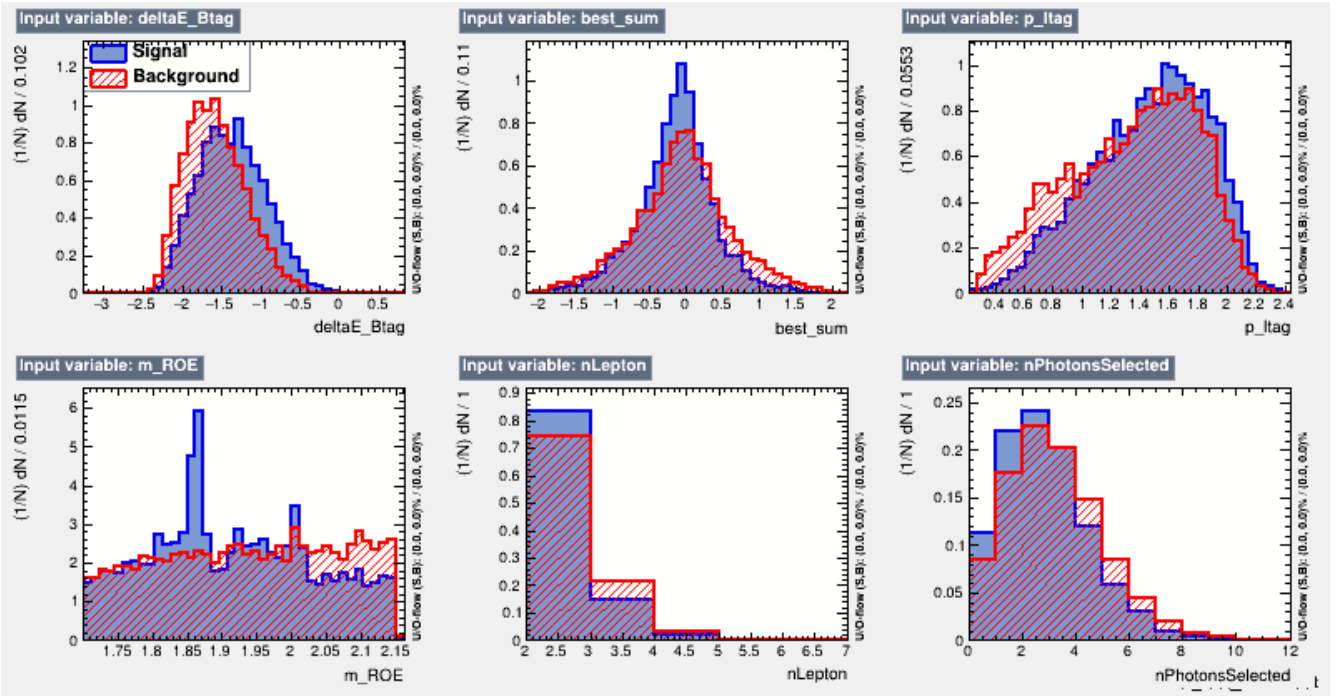
$N_{sig} = 12$ & $N_{sig} = 8$ (for only pi mode)

$N_{B+B-} = 63$

$N_{B0B0} = 28$

Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.



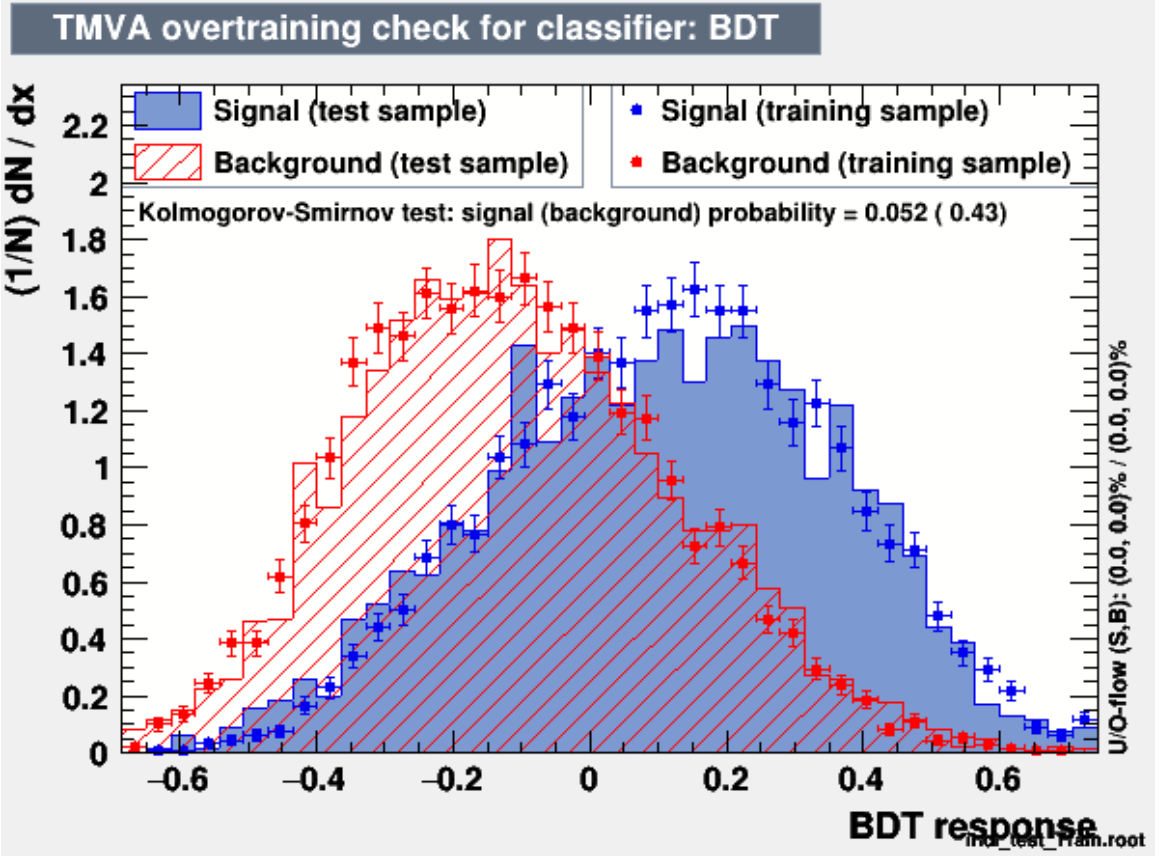
Var. importance

```
Ranking result (top variable is best ranked)
-----------------------------------------------
Rank : Variable         : Variable Importance
-----------------------------------------------
   1 : deltaE_Btag      : 2.542e-01
   2 : m_ROE            : 2.287e-01
   3 : best_sum         : 1.811e-01
   4 : p_ltag           : 1.595e-01
   5 : nPhotonsSelected : 1.005e-01
   6 : nLepton          : 7.592e-02
```

Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

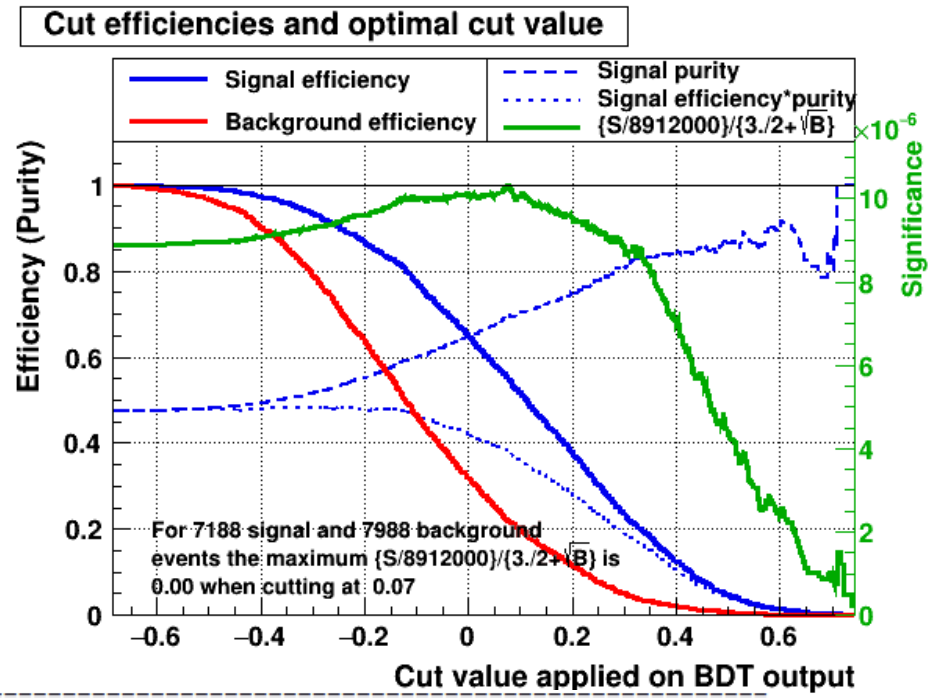
BDT response



Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

Ponzi FOM



BDT optimal cut

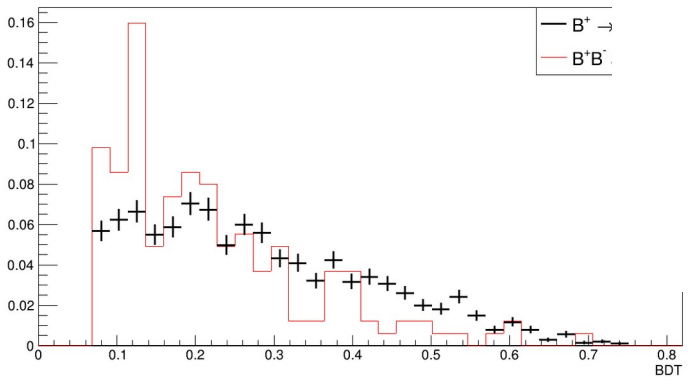
BDT > 0.07

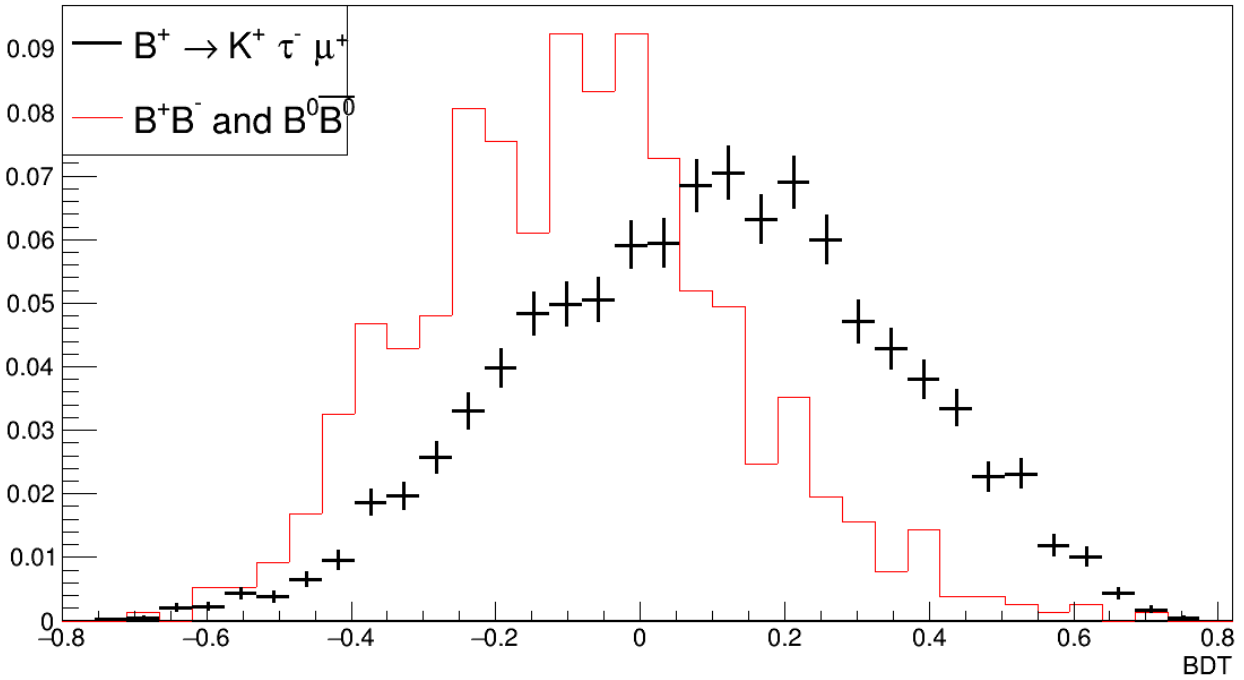
```
================================================================================================================
Classifier   (  #signal, #backgr.)  Optimal-cut  (S/8912000)/(3./2+sqrt(B))      NSig      NBkg    EffSig    EffBkg
----------------------------------------------------------------------------------------------------------------
       BDT:  (     7188,     7988)       0.0745   1.0326e-05   3990.74  1752.726   0.5552    0.2194
      BDTG:  (     7188,     7988)       0.2591   9.76914e-06  3317.282  1339.728   0.4615    0.1677
    Fisher:  (     7188,     7988)      -0.0244   9.51159e-06  4737.545  2958.145   0.6591    0.3703
       MLP:  (     7188,     7988)       0.4911   1.02653e-05  4887.573  2696.244     0.68    0.3375
----------------------------------------------------------------------------------------------------------------
```

Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

BDT score



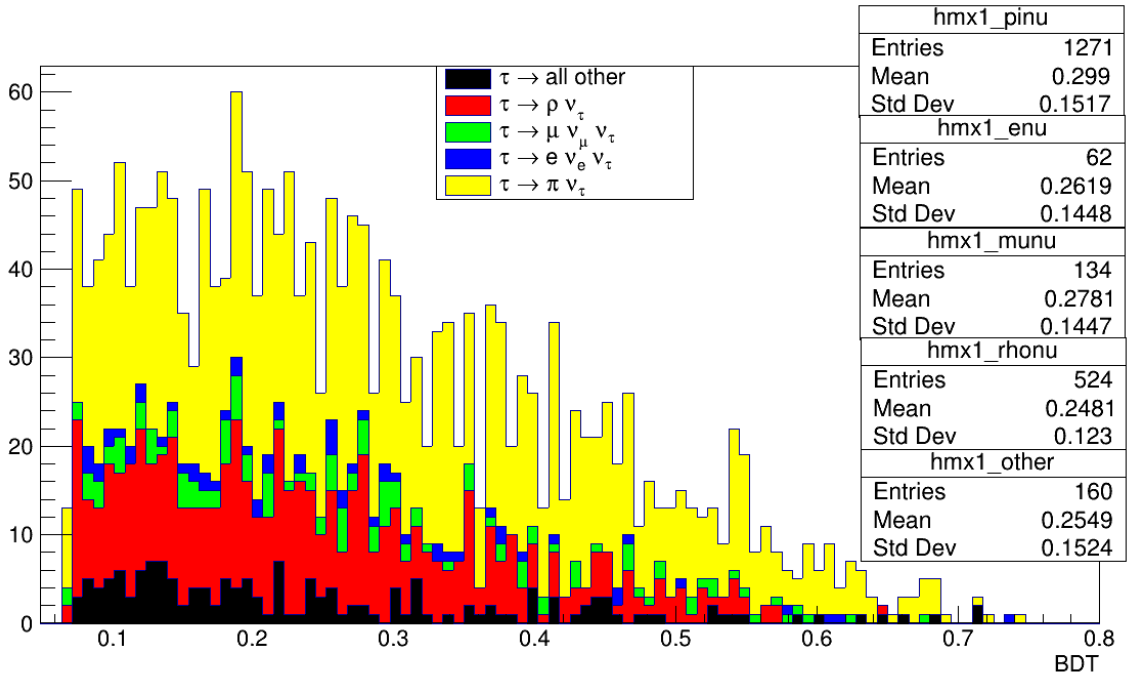


Approach 2

- Training (70%) and testing (30%) on 8.9M $\tau \rightarrow$ generic sample.
- Applying on the 4.4 M $\tau \rightarrow$ generic signal and ten streams (07 training, 03 testing) of MC.
- Estimated the N_{sig} and N_{bg}.

N_{pi}	1271	1
N_e	62	0.05
N_{mu}	134	0.11
N_{rho}	524	0.41
N_{others}	160	0.13

BDT score



For 5×10^{-5} BF

N_{sig} = 19 & N_{sig} = 11 (for only pi mode)

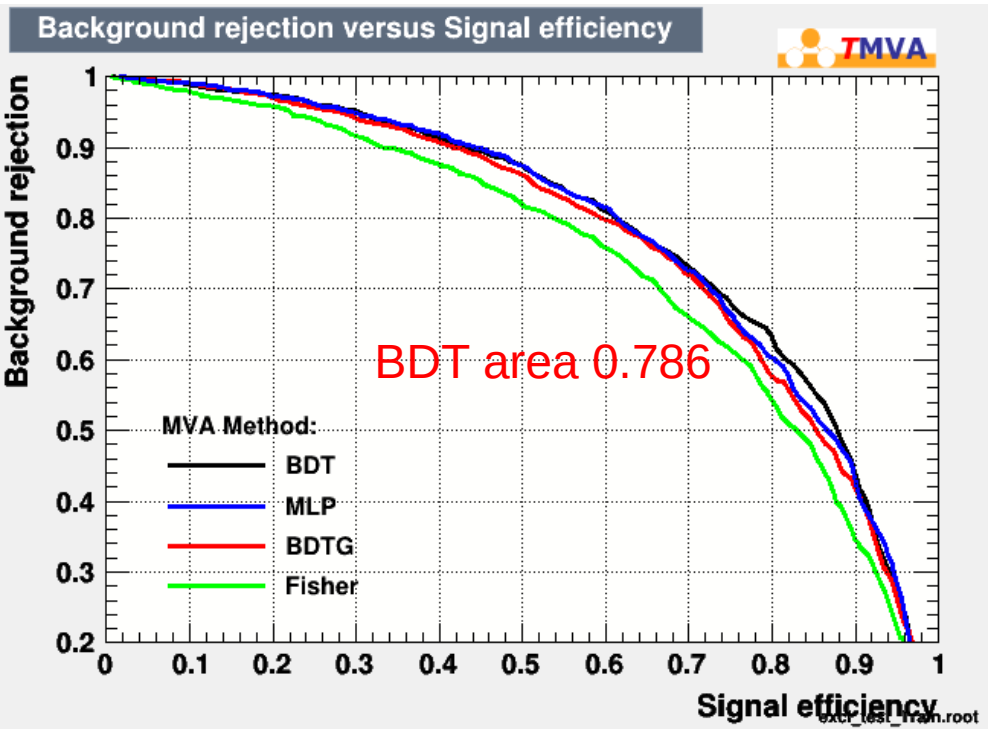
N_{B+B-} = 109

N_{B0B0} = 54

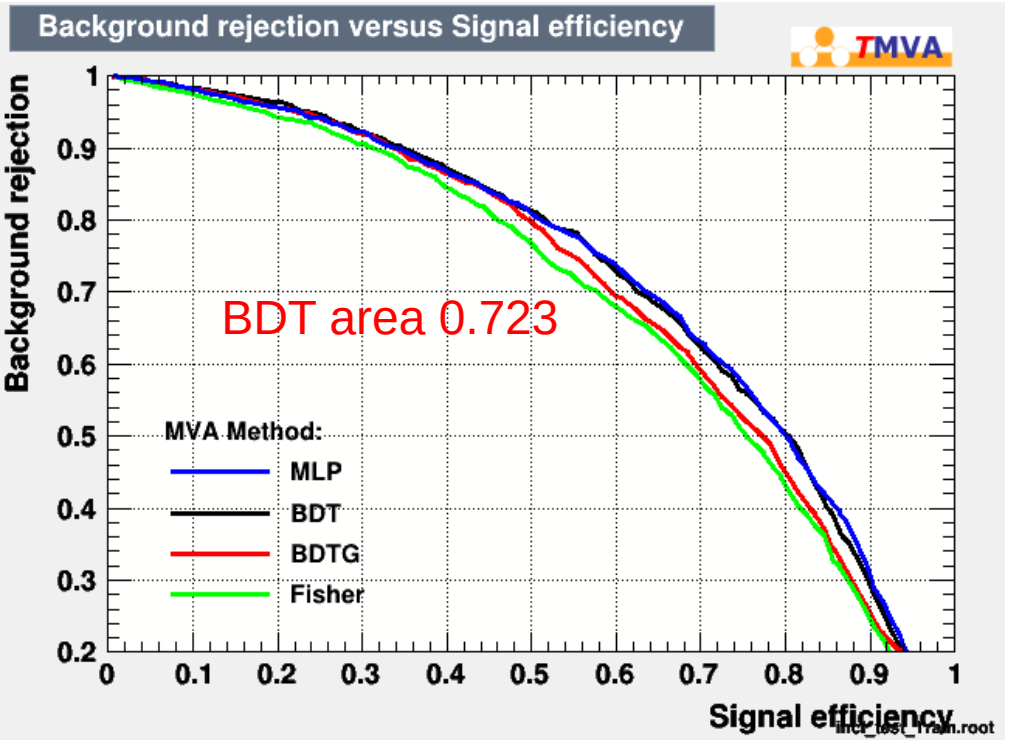
Back up

ROC (test train samples)

Approach 1

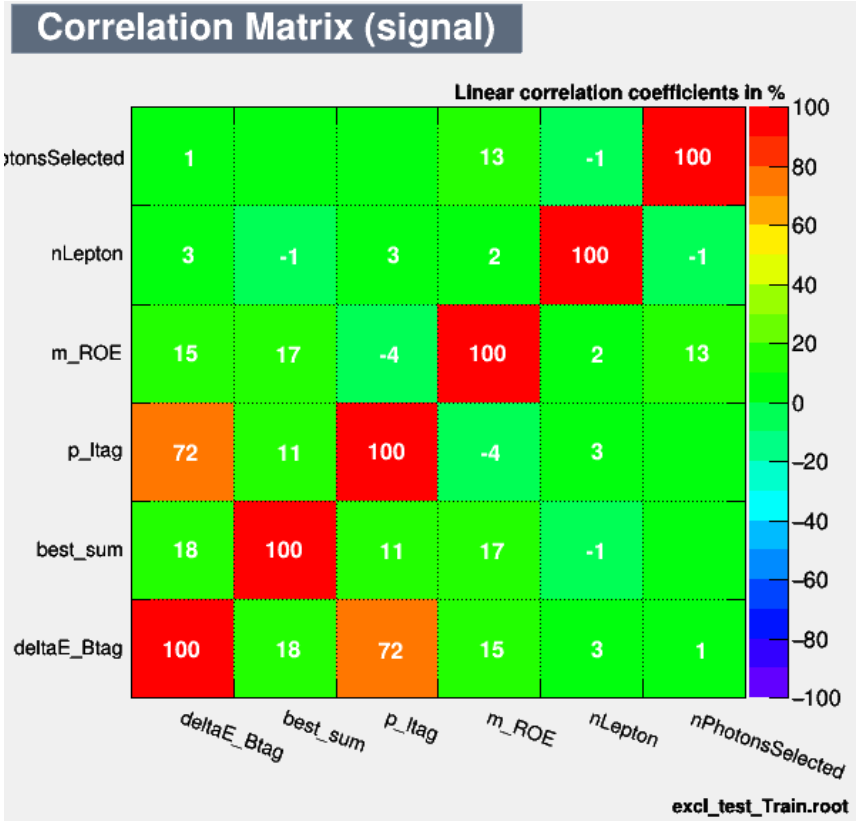


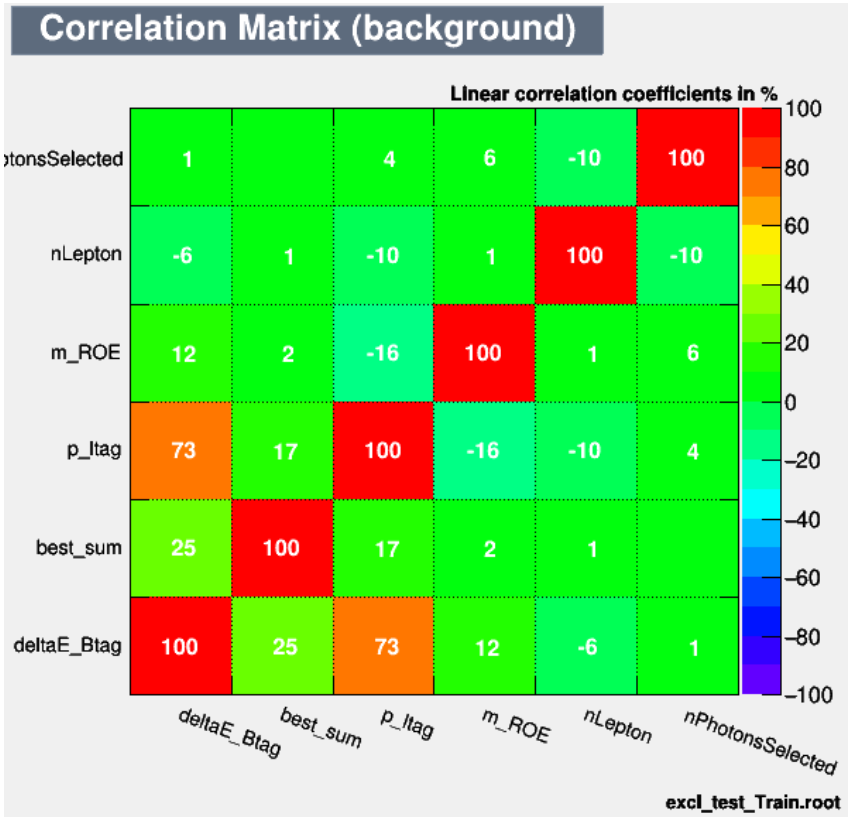
Approach 2



ROC (test train samples)

Approach 1





ROC (test train samples)

Approach 2