Using TMVA for Shower Classification

JEF UPDATE : Oct 22

University of Regina | Faculty of Science

Key MVA Training Variables explained

nHits: Number of blocks in a shower with energy deposited

E_9/E_{25}: Energy deposited in 9 blocks (3x3) over energy deposited in 25 blocks (5x5) centred around the block with the highest energy (similarly E_1/E_9)

sumUSh: normalized second moments of the energy distribution within a shower about $\hat{\mathbf{u}}$ (similarly sumVSh)

asymSh: asymmetry ratio between sumUSh and sumVSh

speedSh : effective velocity -> the distance over the time of the difference in the interaction point and the shower

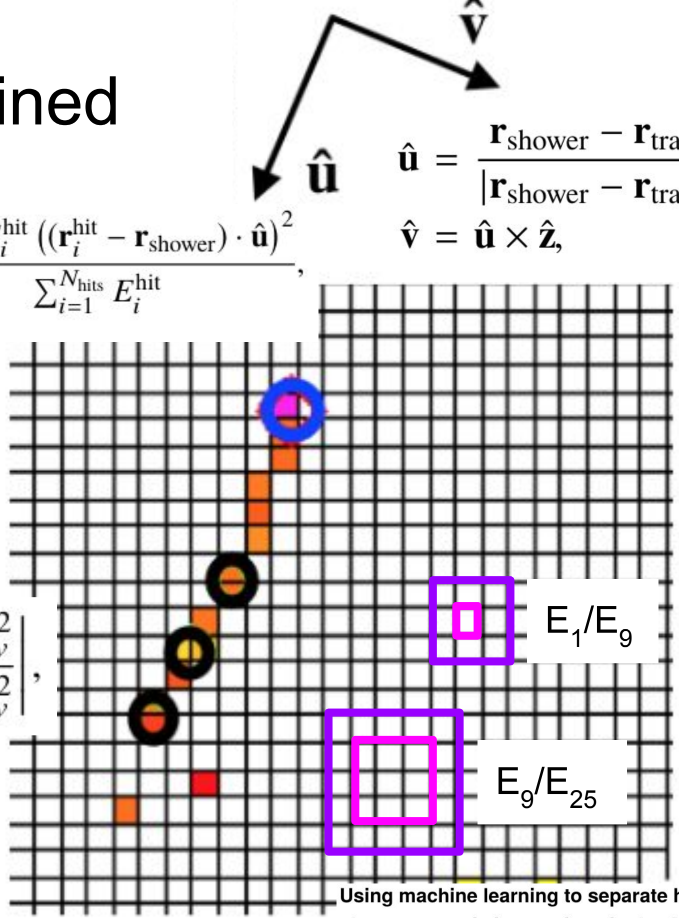
dtSh (Δt) : the difference in shower time and the impact cluster track time.

$$\hat{\mathbf{u}} = \frac{\mathbf{r}_{\text{shower}} - \mathbf{r}_{\text{track}}}{|\mathbf{r}_{\text{shower}} - \mathbf{r}_{\text{track}}|},$$

$$\hat{\mathbf{v}} = \hat{\mathbf{u}} \times \hat{\mathbf{z}},$$

$$\sigma_u^2 = \frac{\sum_{i=1}^{N_{\text{hits}}} E_i^{\text{hit}} \left((\mathbf{r}_i^{\text{hit}} - \mathbf{r}_{\text{shower}}) \cdot \hat{\mathbf{u}}\right)^2}{\sum_{i=1}^{N_{\text{hits}}} E_i^{\text{hit}}},$$

$$A_{uv} = \left|\frac{\sigma_u^2 - \sigma_v^2}{\sigma_u^2 + \sigma_v^2}\right|,$$



Using machine learning to separate hadronic and electromagnetic interactions in the GlueX forward calorimeter

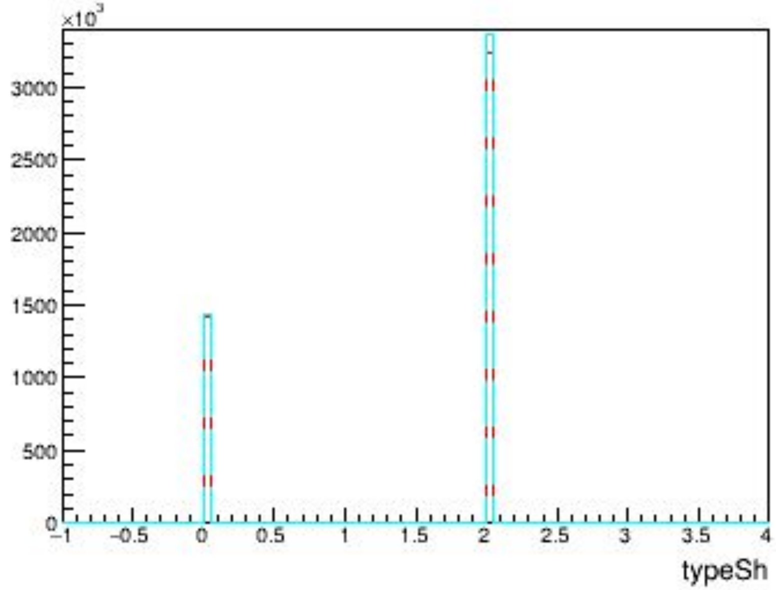
Rebecca Barsotti and Matthew R. Shepherd

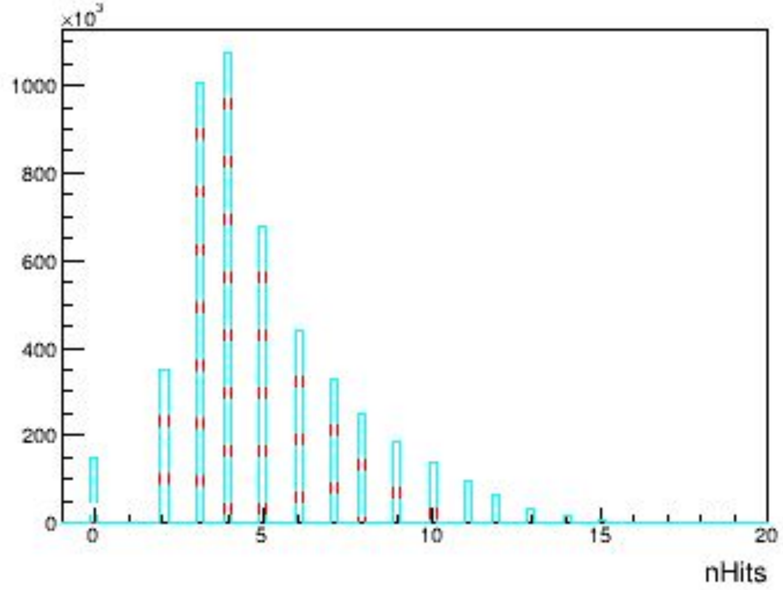
Department of Physics, Indiana University, Bloomington, IN 47405

Differents EBCAL Cuts Overlayed

EBCAL < 0.05
EBCAL < 0.1
EBCAL < 0.2

Type 0 - Signal - True Photons
Type 2 - Background - Splitoff Photons

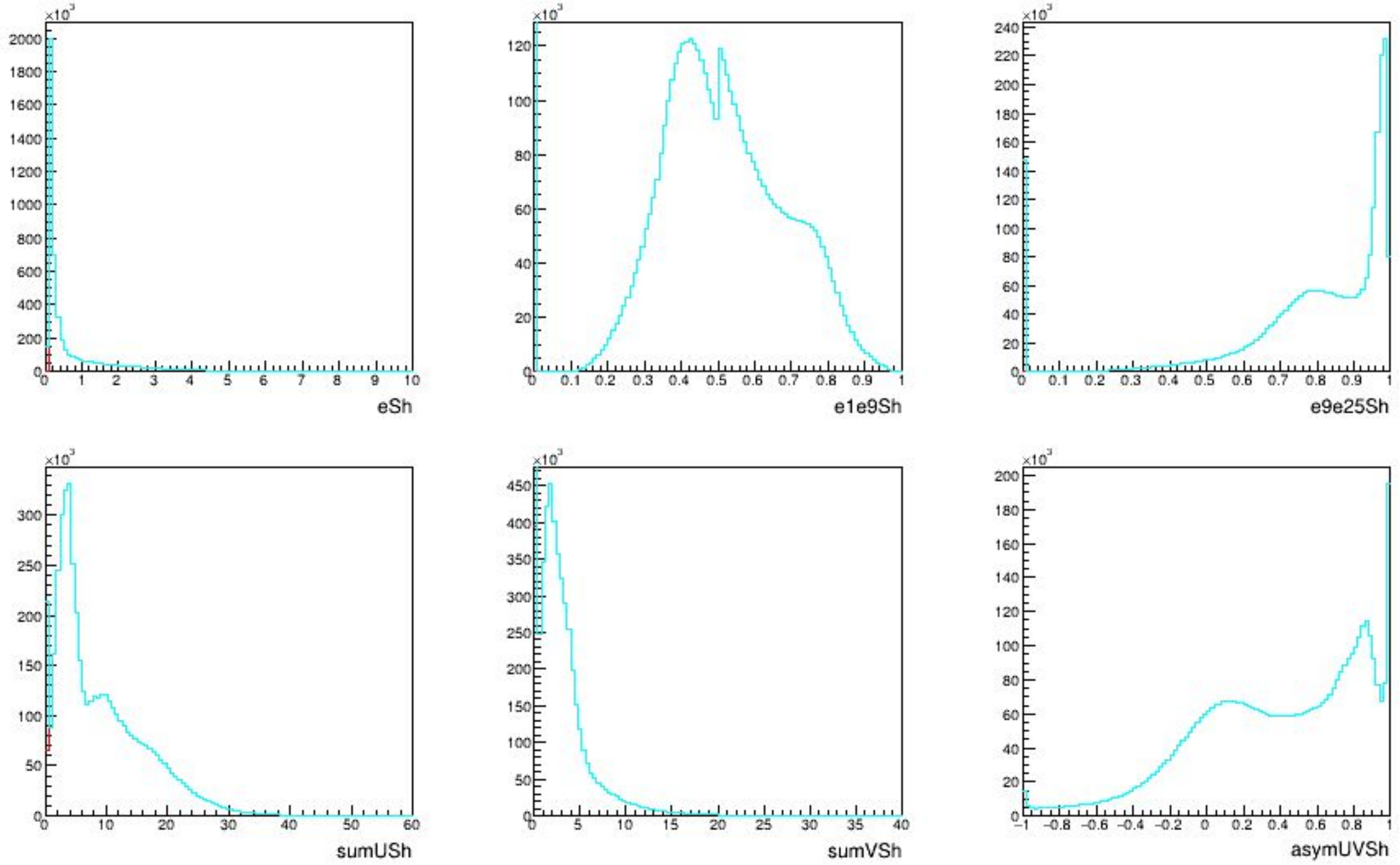




See no difference between different BCAL Energy cuts

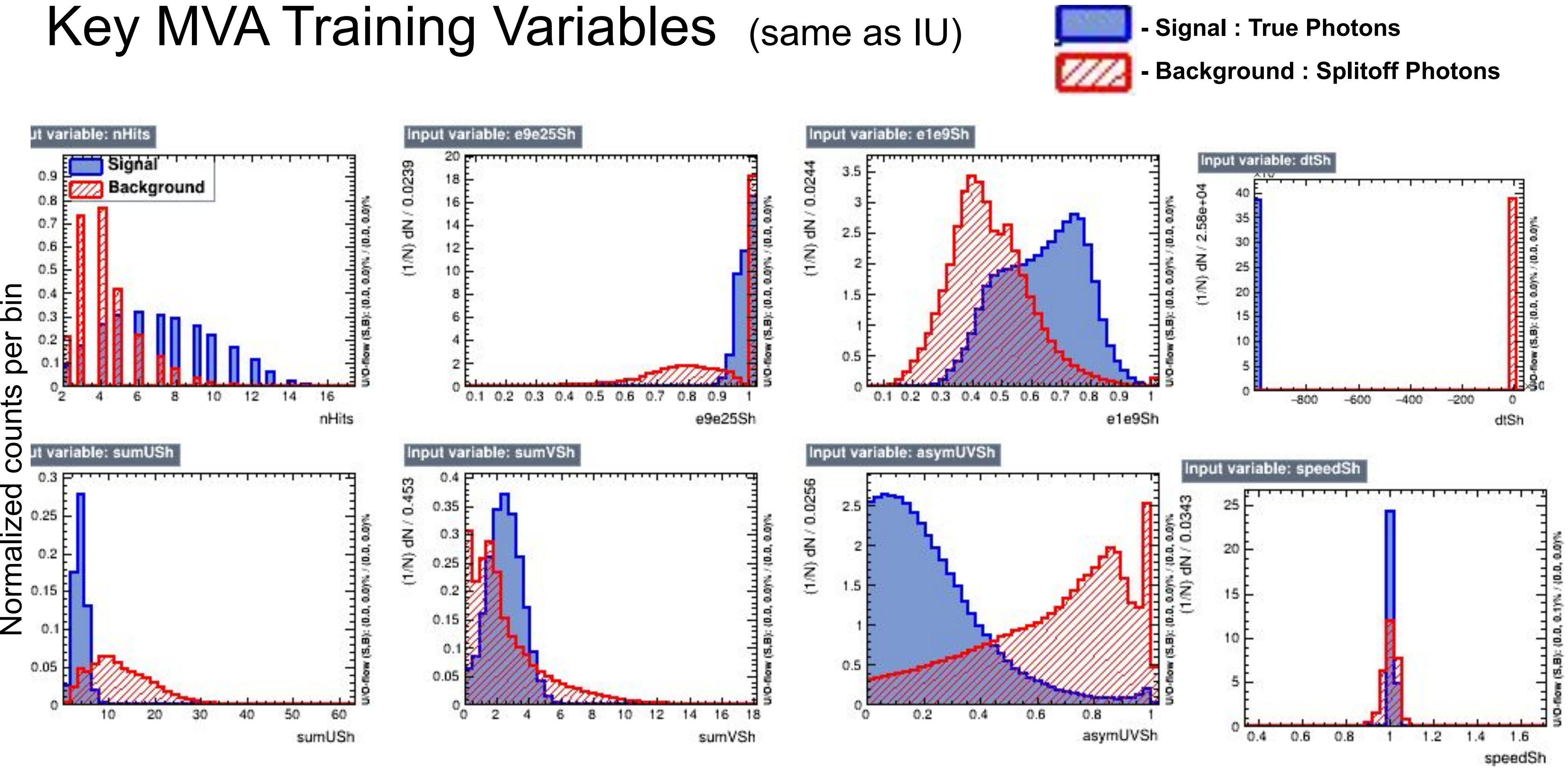
Differents EBCAL Cuts Overlayed

EBCAL < 0.05
EBCAL < 0.1
EBCAL < 0.2



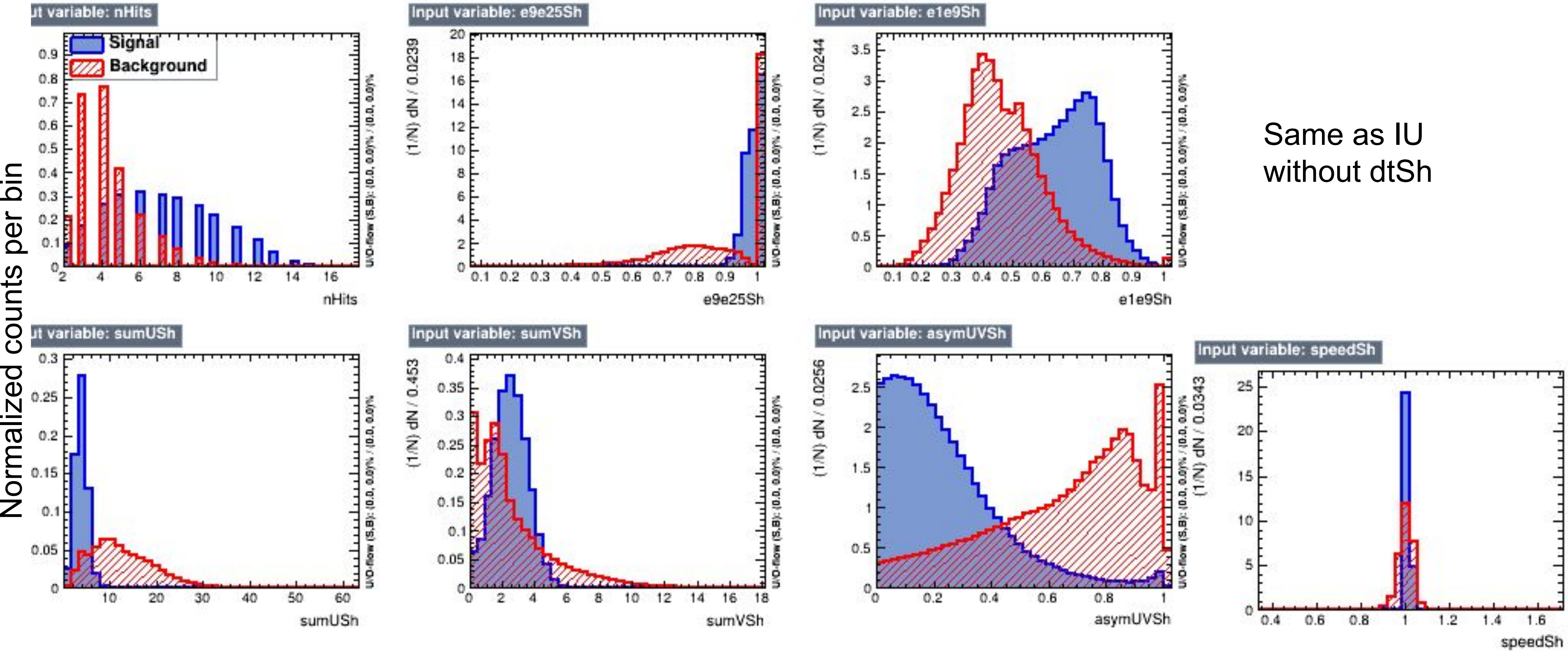
See no difference between different BCAL Energy cuts in key training variables

Key MVA Training Variables (same as IU)



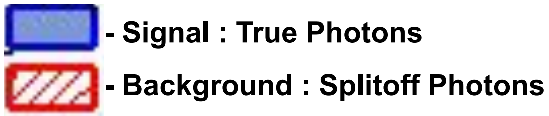
Current Training Set

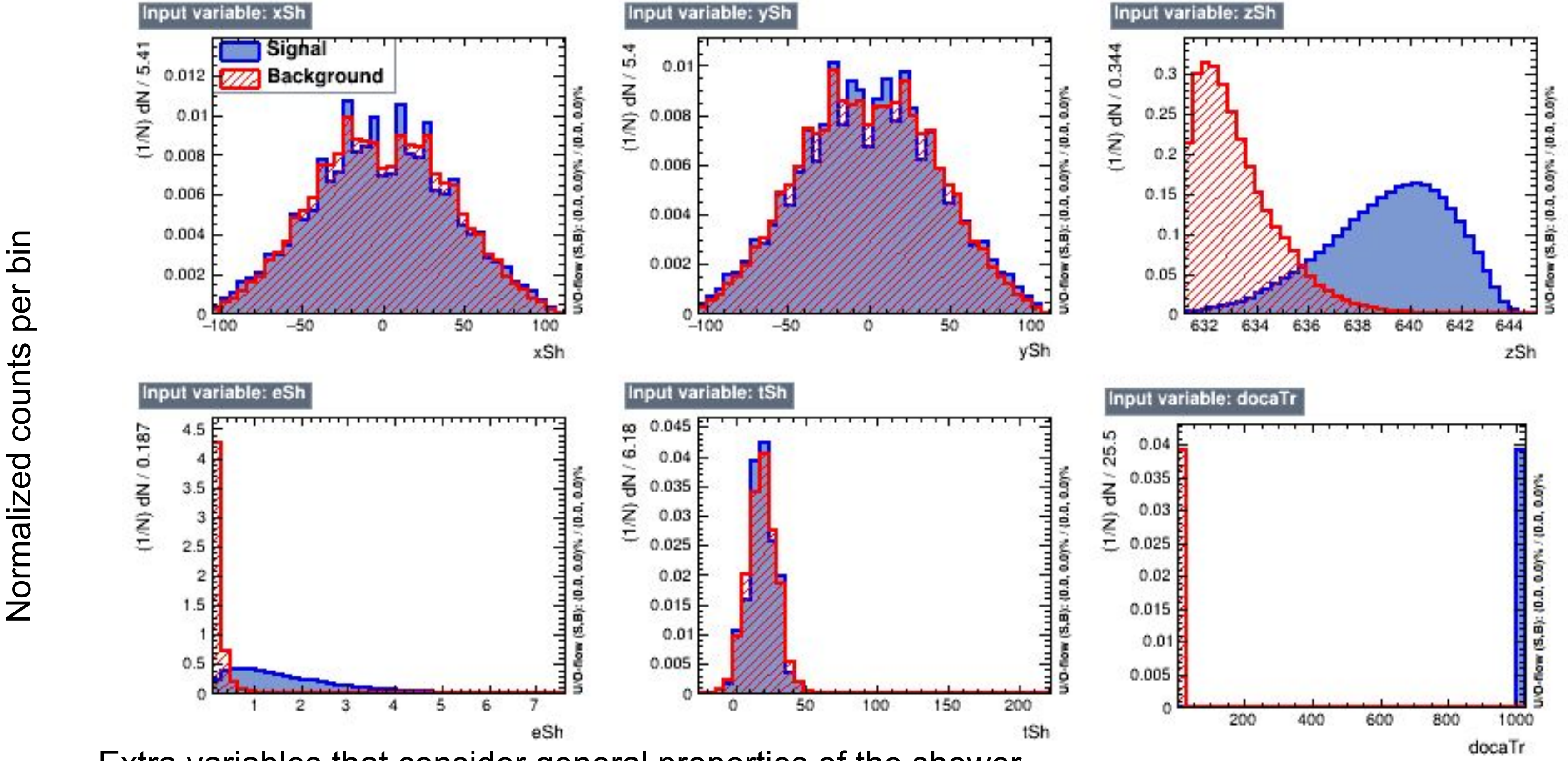
- **Signal : True Photons**
- **Background : Splitoff Photons**



Same as IU without dtSh

Additionally Considered Variables



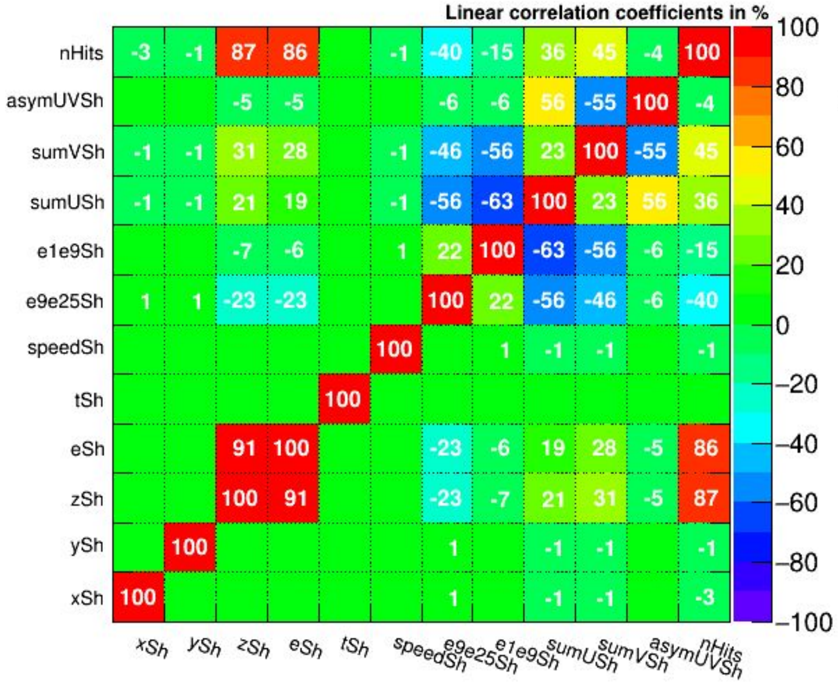


Extra variables that consider general properties of the shower

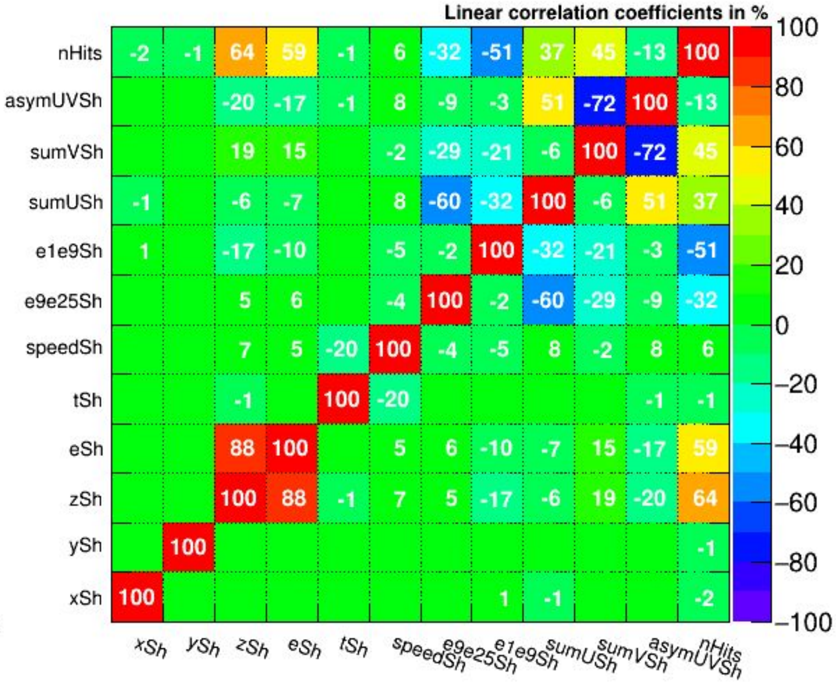
docaTr : distance of closest approach of the shower to the track

Correlation matrices for all variables

Correlation Matrix (signal)

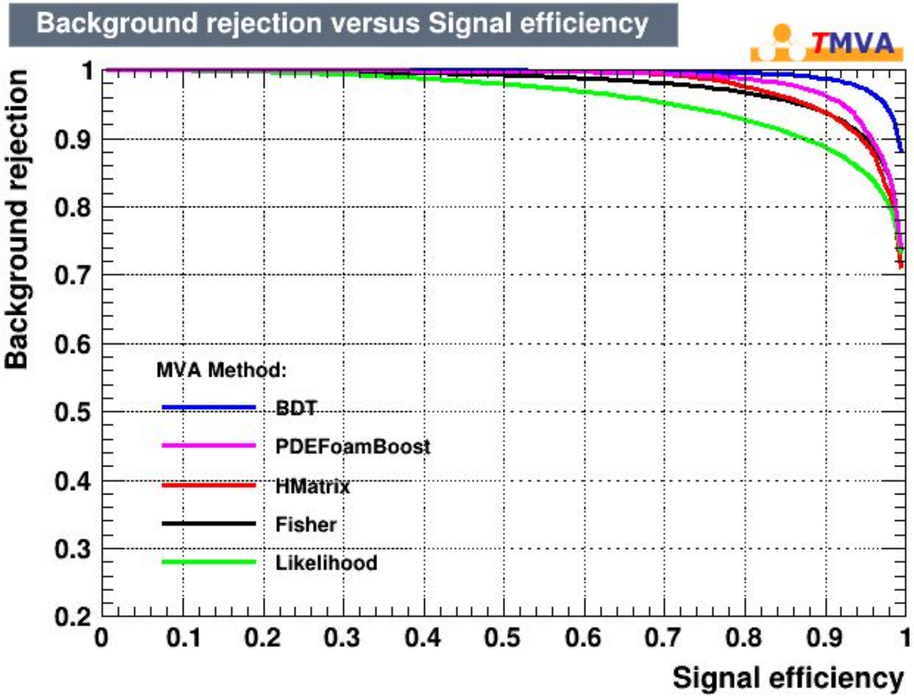


Correlation Matrix (background)



Model Performance Comparison to IU

(almost) same input features
Same MVA training parameters



UofR - no dtSh

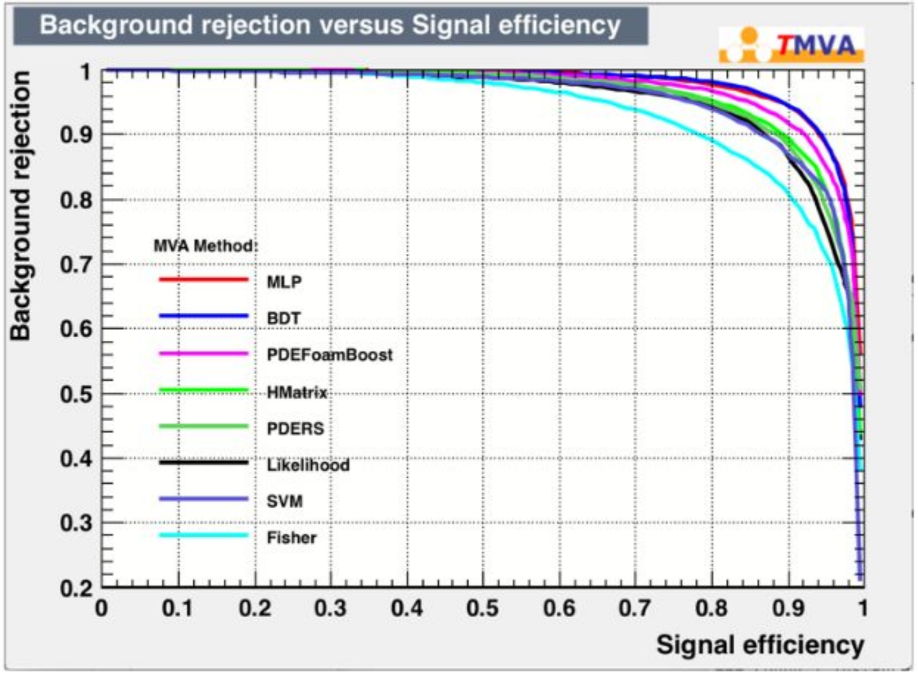
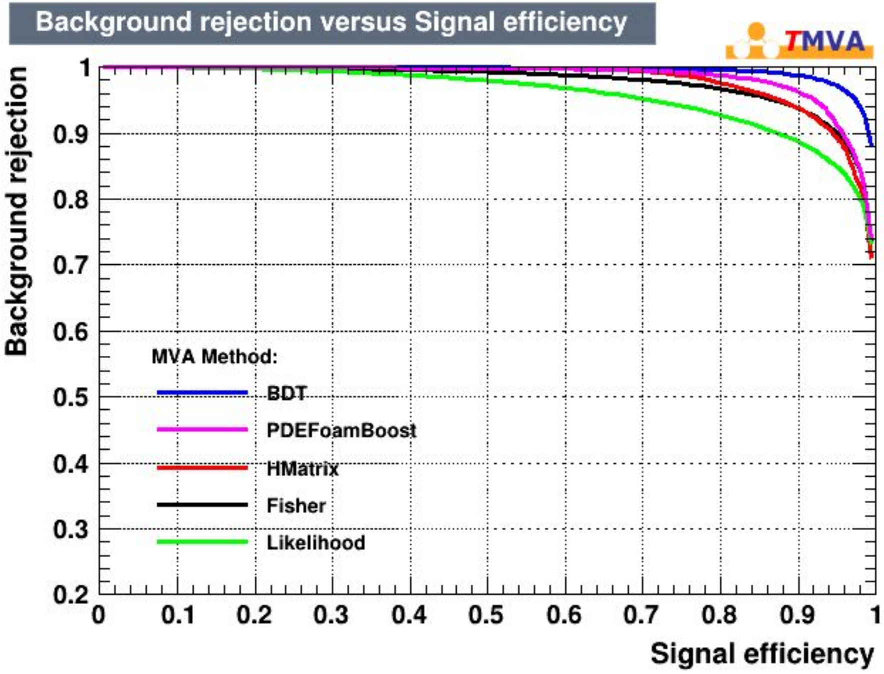


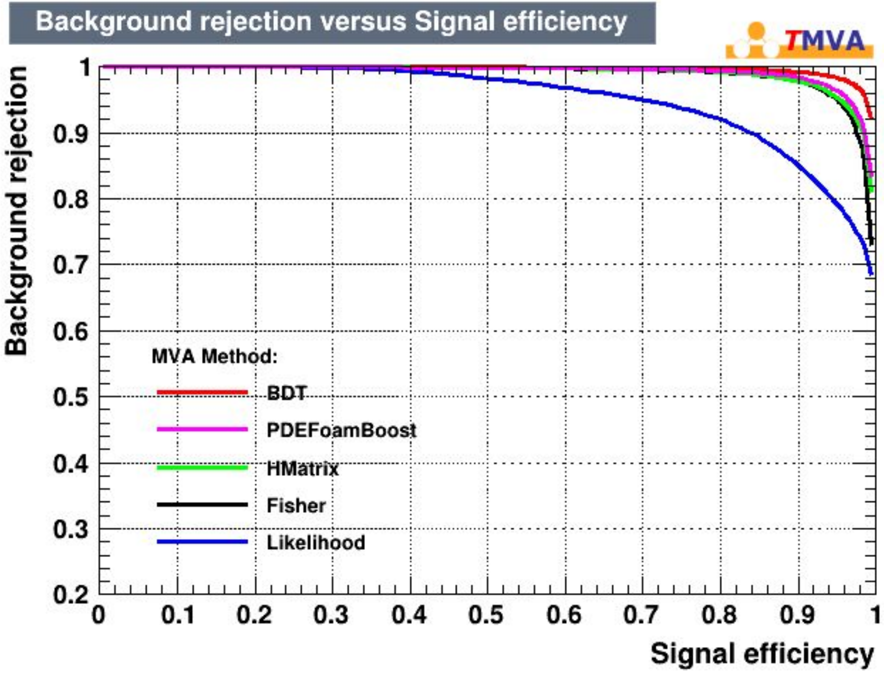
Fig 5: IU : https://arxiv.org/abs/2002.09530

Model Performance Comparison : Feature Selection

Different input features
Same MVA training parameters



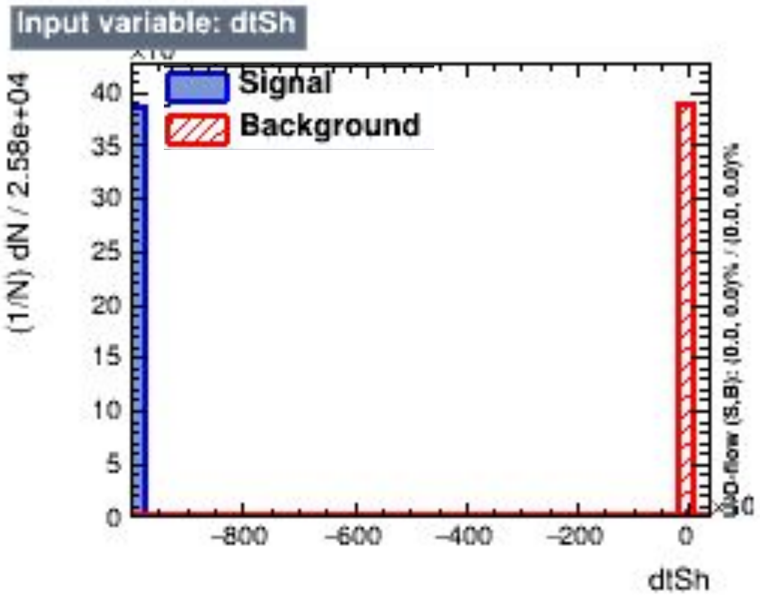
Considering only key training variables

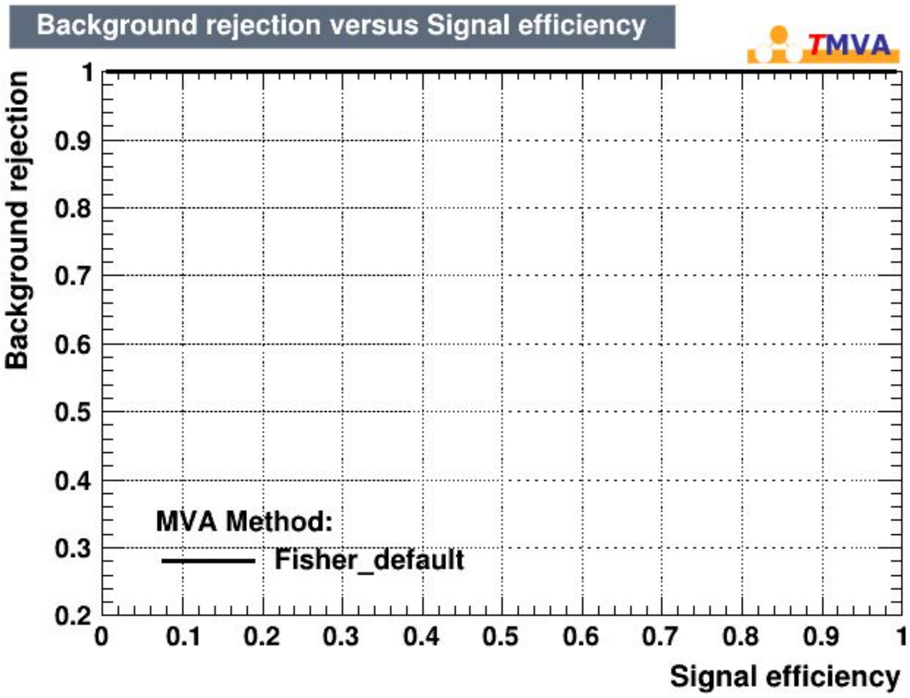


Considering all training variables

Using dtSh as a Feature

The problem with using dtSh as a training variable is that it directly correlated to shower type





All Input Variables