

# W Meeting EndCap Cuts study

Amani Kraishan

5/31/2017

# Toolkit for Multivariate Data Analysis (TMVA)

All multivariate techniques in TMVA belong to the family of “supervised learning” algorithms.

To make use of **training events**, for which the desired output is known, **to determine** the **mapping function** that either describes a decision boundary (classification) or an approximation of the underlying functional behavior defining the target value (regression).

# Input Info

- The training and testing is performed with the use of user-supplied data sets in form of ROOT trees.
  - The **Signal tree** has data from **the MC**.
  - The **background tree** has data from **Run 13 data**.
- Both trees have the same leaves.

The cuts I studied:

- **2x2/4x4**, **2x2/near\_cone**, **RSMD**, and **Signed  $p_T$  Balance**

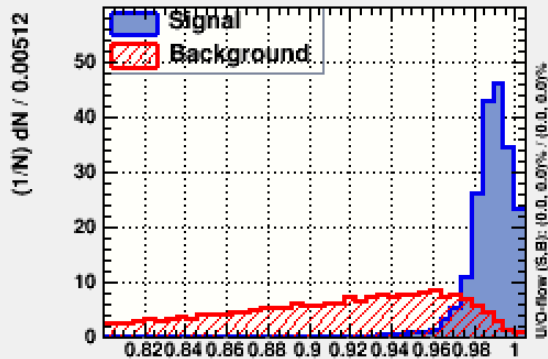
# TMVA methods

## Rectangular cut optimization

- The simplest and most common classifier for selecting signal events from a **mixed sample of signal and background** events is the application of an ensemble of rectangular cuts on discriminating variables.
- The optimization of cuts performed by TMVA **maximizes the background rejection** at given signal efficiency

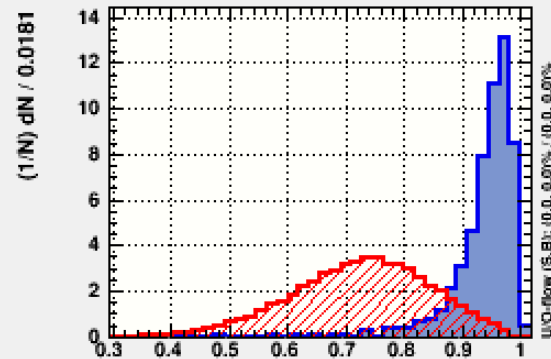
# Input Variable Distributions

Input variable: Et24F



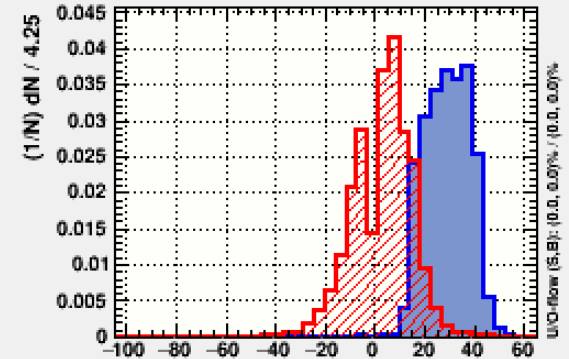
2x2/4x4

Input variable: cone



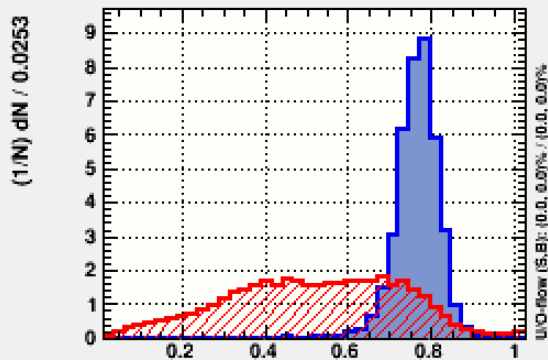
2x2/near\_cone

Input variable: ptb



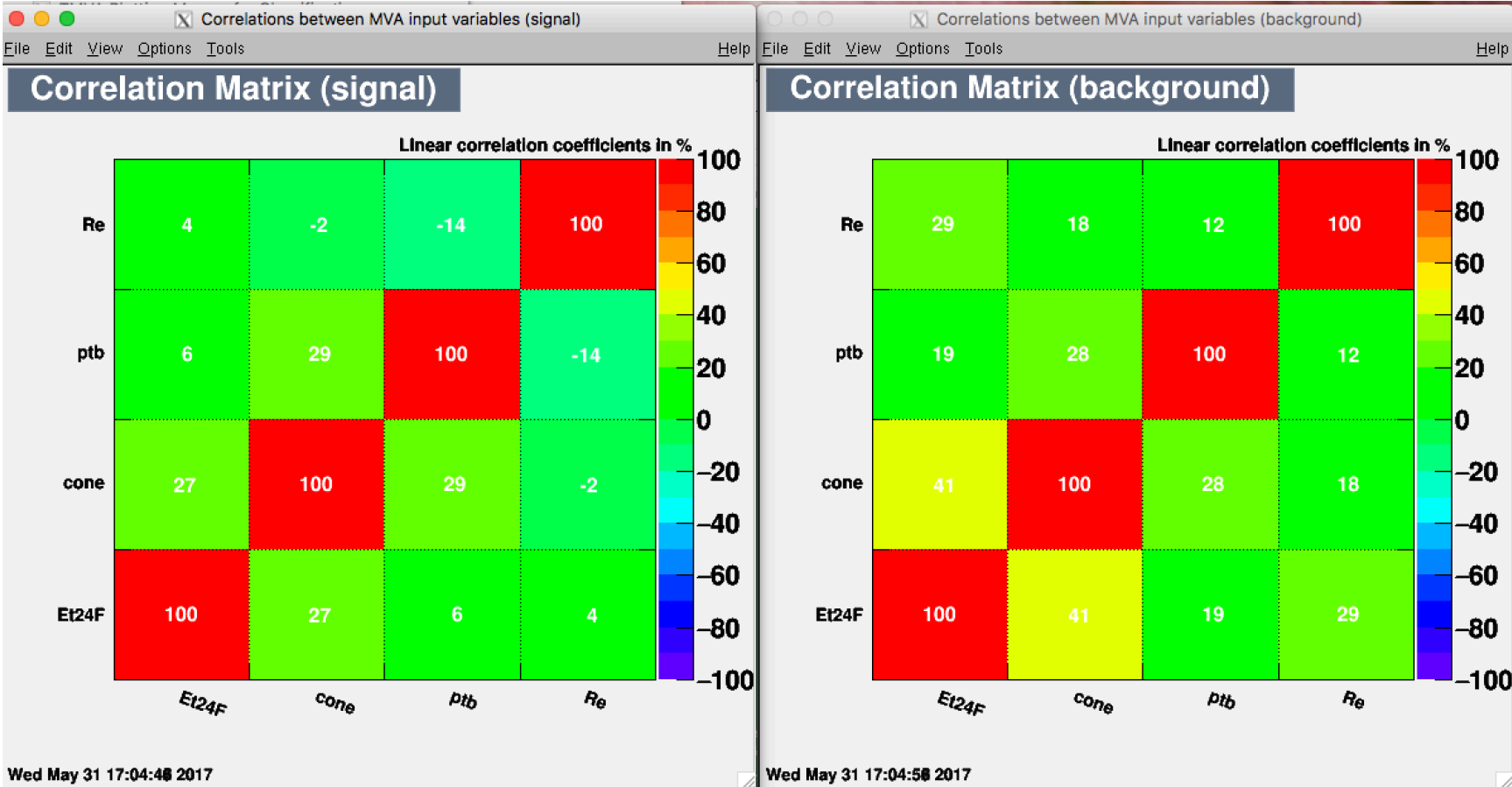
pT\_balance

Input variable: Re



RSMD

# Correlation between input variables



# Cuts value

```
--- Cuts : -----  
--- Cuts : Cut values for requested signal efficiency: 0.7  
--- Cuts : Corresponding background efficiency : 0.00316667  
--- Cuts : Transformation applied to input variables : None  
--- Cuts : -----  
--- Cuts : Cut[ 0]: 0.978956 < Et24F <= 1e+30  
--- Cuts : Cut[ 1]: 0.790692 < cone <= 1e+30  
--- Cuts : Cut[ 2]: 21.3128 < ptb <= 1e+30  
--- Cuts : Cut[ 3]: 0.574082 < Re <= 1e+30  
--- Cuts : -----  
--- Cuts : -----  
--- Cuts : Cut values for requested signal efficiency: 0.8  
--- Cuts : Corresponding background efficiency : 0.006  
--- Cuts : Transformation applied to input variables : None  
--- Cuts : -----  
--- Cuts : Cut[ 0]: 0.980532 < Et24F <= 1e+30  
--- Cuts : Cut[ 1]: 0.720738 < cone <= 1e+30  
--- Cuts : Cut[ 2]: 15.8237 < ptb <= 1e+30  
--- Cuts : Cut[ 3]: 0.624296 < Re <= 1e+30  
--- Cuts : -----  
--- Cuts : -----  
--- Cuts : Cut values for requested signal efficiency: 0.9  
--- Cuts : Corresponding background efficiency : 0.0124333  
--- Cuts : Transformation applied to input variables : None  
--- Cuts : -----  
--- Cuts : Cut[ 0]: 0.970945 < Et24F <= 1e+30  
--- Cuts : Cut[ 1]: 0.809673 < cone <= 1e+30  
--- Cuts : Cut[ 2]: 14.795 < ptb <= 1e+30  
--- Cuts : Cut[ 3]: 0.637274 < Re <= 1e+30  
--- Cuts : -----
```

# Input Variable Ranking

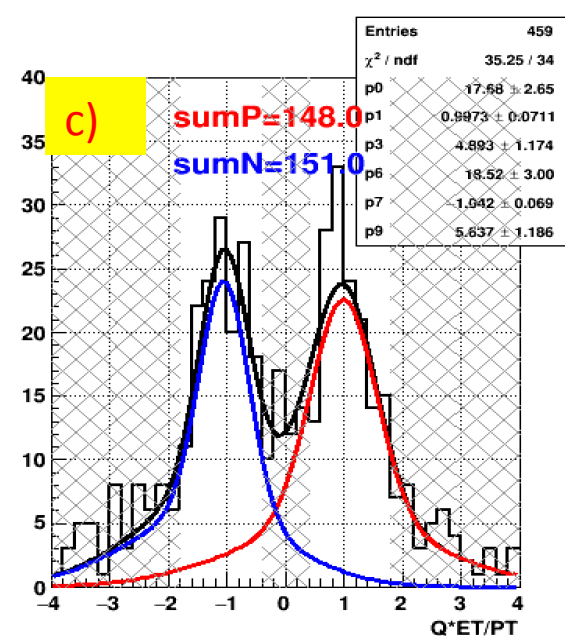
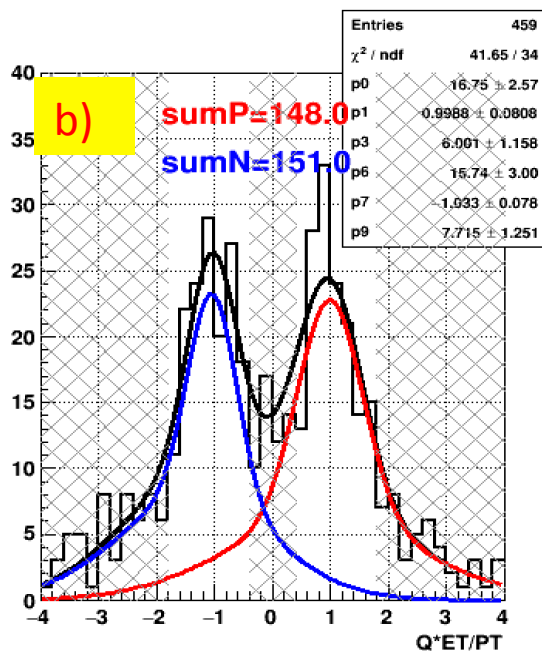
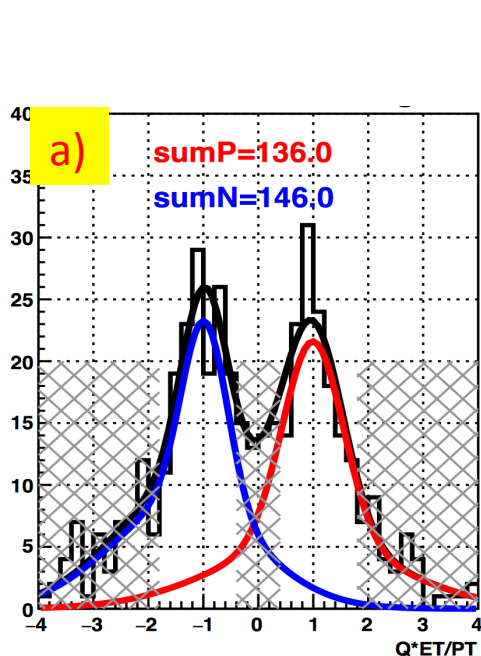
```
--- TFHandler_Factory : Create scatter and profile plots in target-file directory:
--- TFHandler_Factory : TMVA.root:/InputVariables_Gauss_Deco/CorrelationPlots
--- TFHandler_Factory :
--- TFHandler_Factory : Ranking input variables (method unspecific)...
--- IdTransformation : Ranking result (top variable is best ranked)
--- IdTransformation : -----
--- IdTransformation : Rank : Variable : Separation
--- IdTransformation : -----
--- IdTransformation : 1 : ptb : 7.429e-01
--- IdTransformation : 2 : Et24F : 7.397e-01
--- IdTransformation : 3 : cone : 7.396e-01
--- IdTransformation : 4 : Re : 5.755e-01
--- IdTransformation : -----
--- Factory :
```

## How discriminating is a variable?

```
--- Fisher : Ranking result (top variable is best ranked)
--- Fisher : -----
--- Fisher : Rank : Variable : Discr. power
--- Fisher : -----
--- Fisher : 1 : ptb : 2.055e-01
--- Fisher : 2 : cone : 1.987e-01
--- Fisher : 3 : Et24F : 1.515e-01
--- Fisher : 4 : Re : 1.050e-01
--- Fisher : -----
--- Factory :
```



# Charge Separation



Charge mis\_id (fit options)

W+

W-

a)W cut(L)

0.086

0.127

b)TMVA cut(L)

0.077

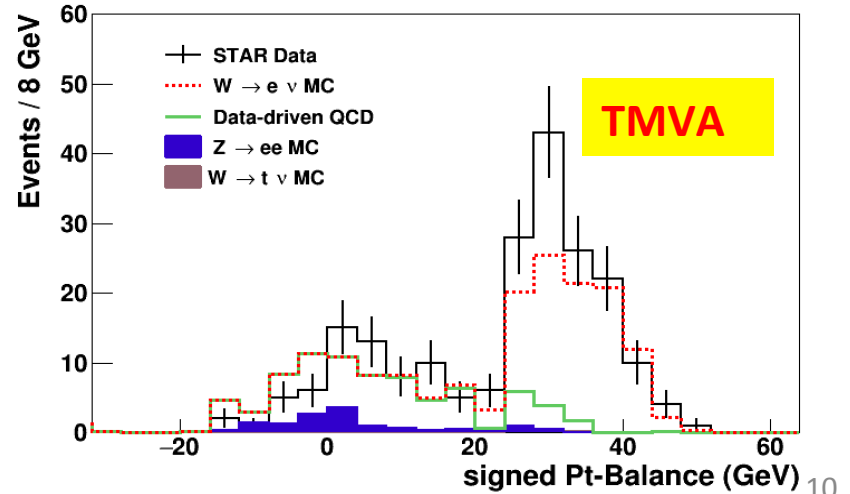
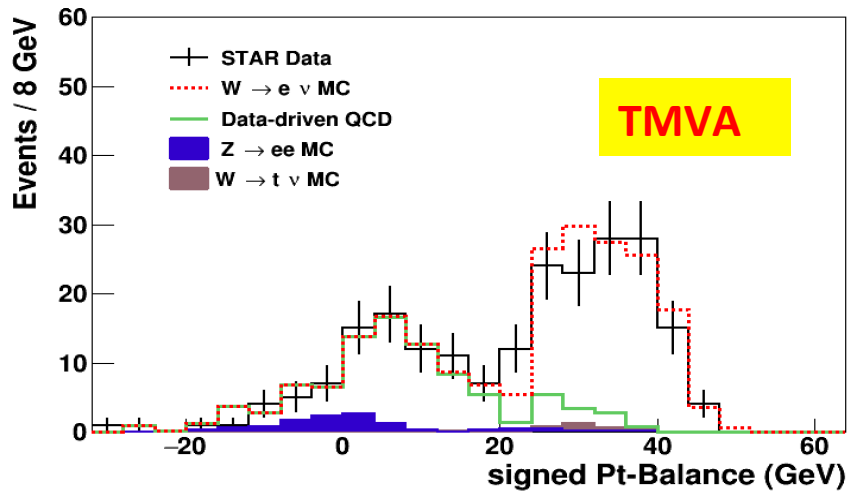
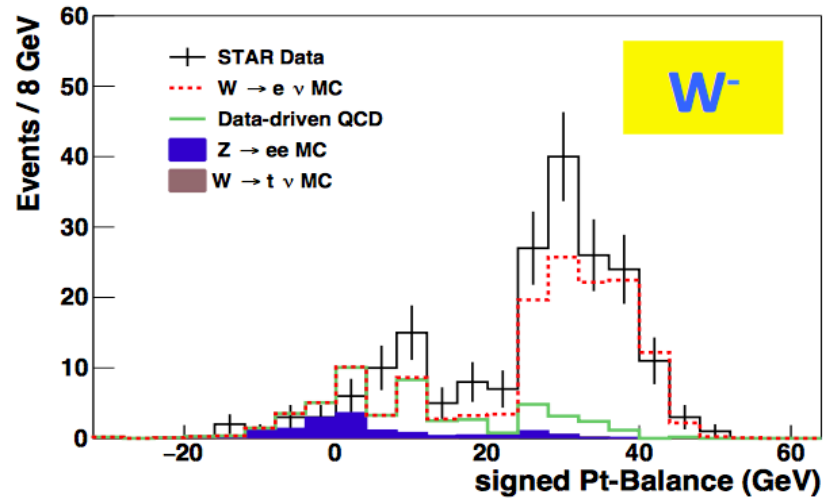
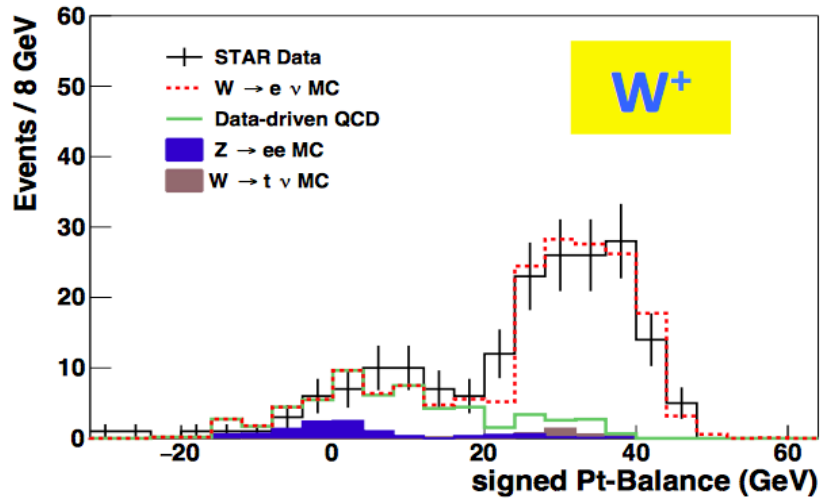
0.137

c)TMVA cut(P)

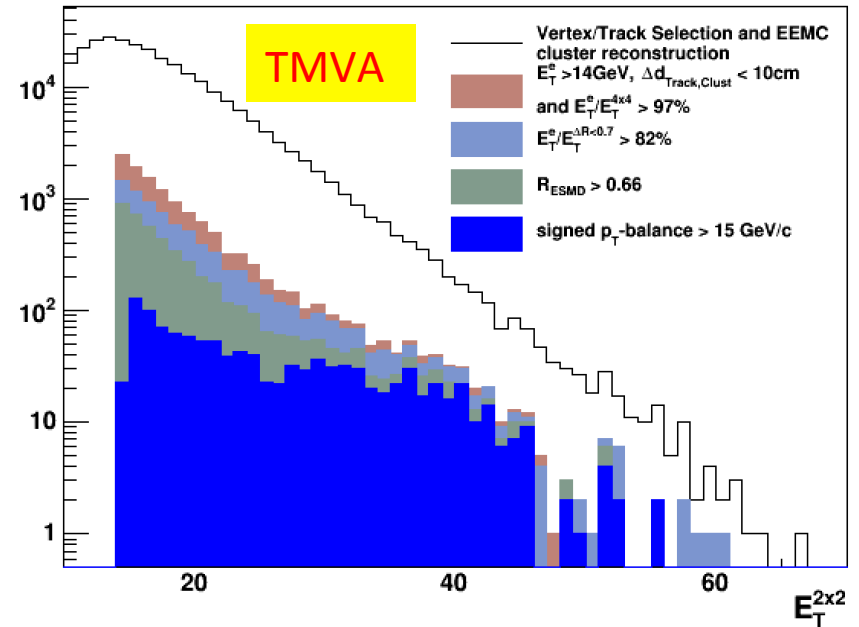
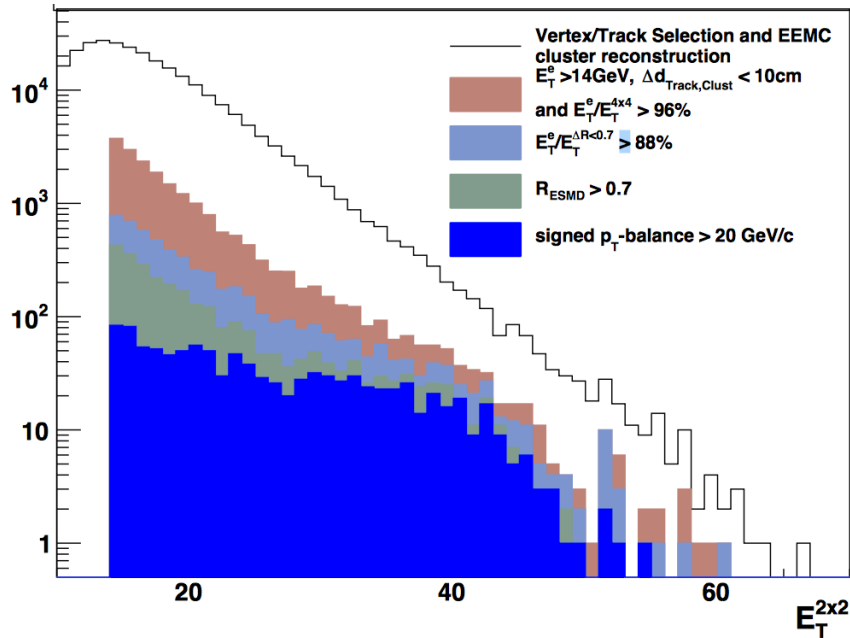
0.059

0.125

# Background



# W Selection Histogram



|        | 2x2/4x4 | 2x2/cone | RSMD | Signe_P <sub>T</sub> |
|--------|---------|----------|------|----------------------|
| W cuts | 96%     | 88%      | 0.7  | 20                   |
| TMVA   | 97%     | 82%      | 0.66 | 15                   |