

Variable Classes

Observables

Invariant masses
Production/Decay angles
CoM kinematics

Used to extract physics

Don't want to mess with these

Discriminatory

Exclusive Process
Missing Mass
Missing Energy
Missing Momentum
Decay Processes
+ Invariant masses

Used to remove backgrounds

They still need to be useable after
any ML (or other) cuts

Particle Identification

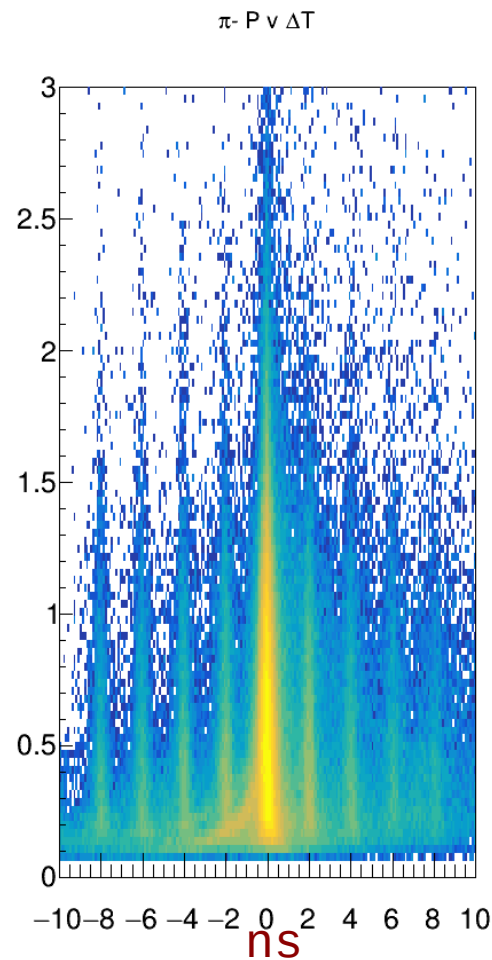
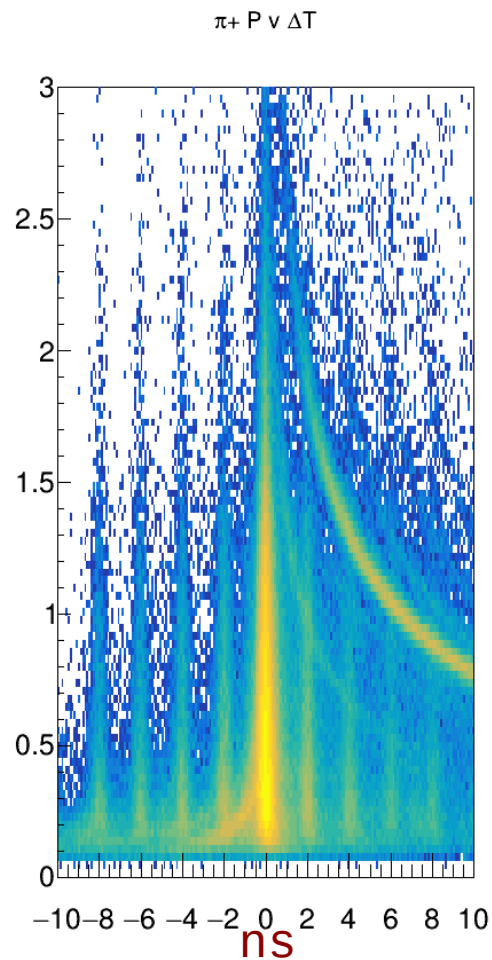
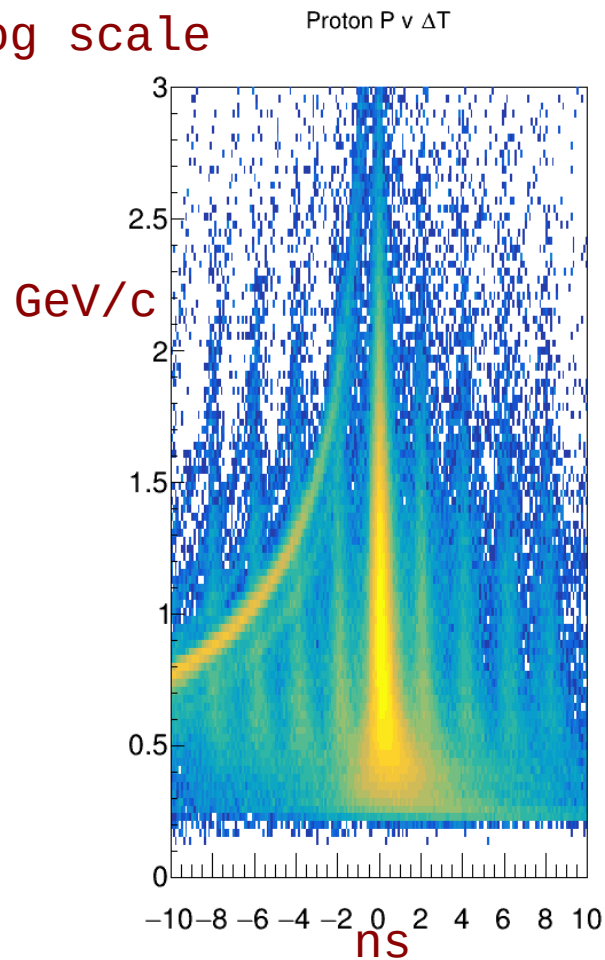
Momentum, Position
Time of Flight
Delta Energy
Cherenkov ...

Used to minimise backgrounds

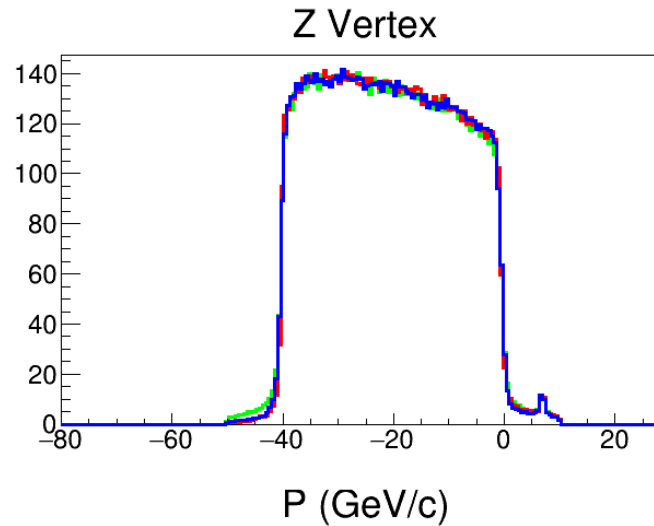
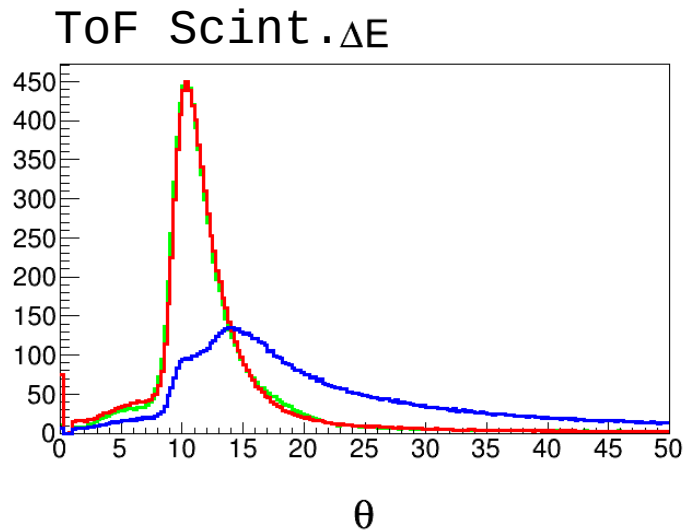
Best candidate for ML algorithms?¹

Example from CLAS ΔT PID

*Log scale



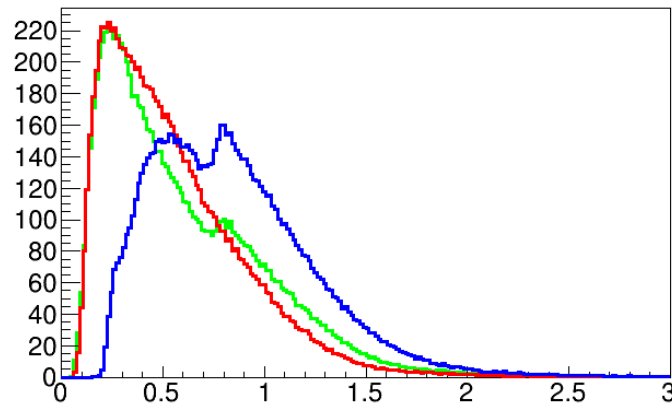
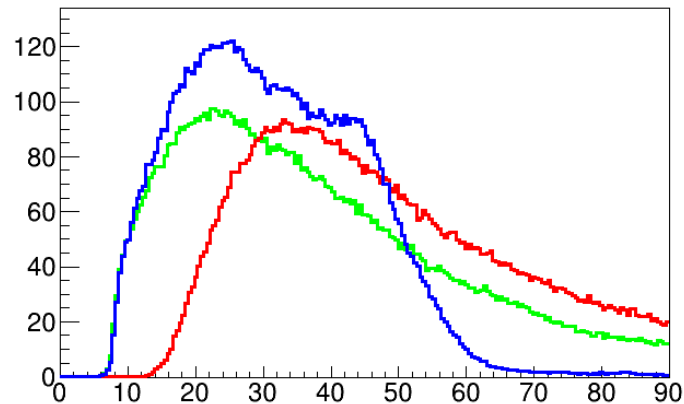
Other Input Variables



Proton

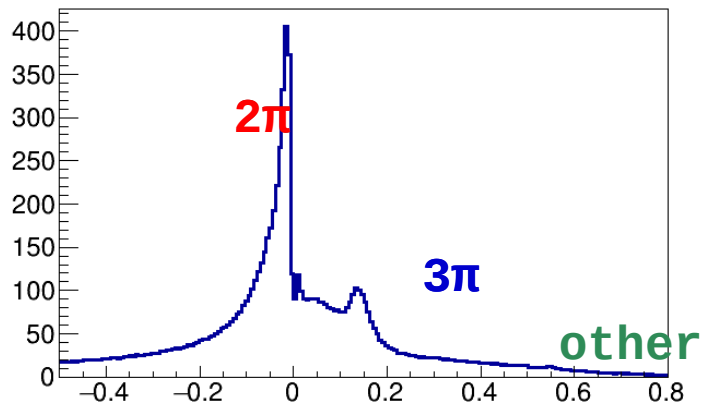
π^+

π^-

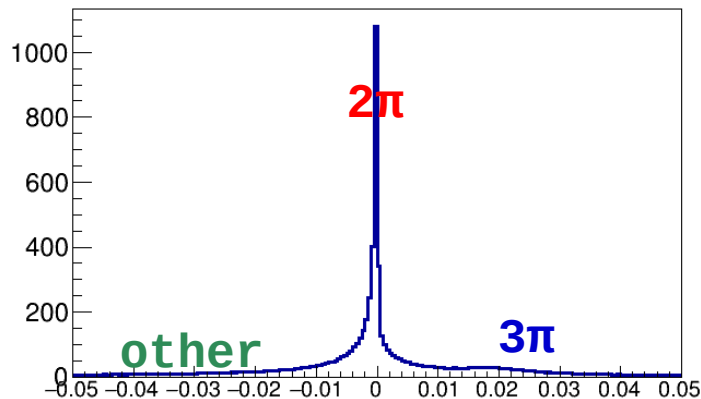


Exclusivity variables

Missing Mass ($p\pi+\pi^-$)

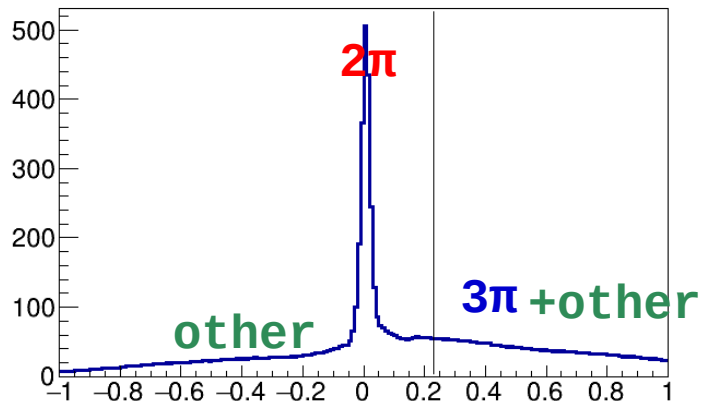


Missing Mass Squared ($p\pi+\pi^-$)

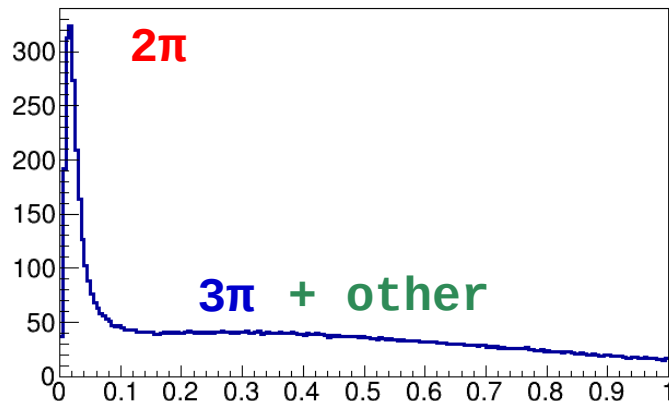


NOT FOR
TRAINING WITH

Missing Energy ($p\pi+\pi^-$)



Missing Momentum ($p\pi+\pi^-$)



GOAL :
Using PID variables
Retain all 3π
Remove backgrounds

Train with Simulation

Signal : **Simulated** 3π (Truth Matched)

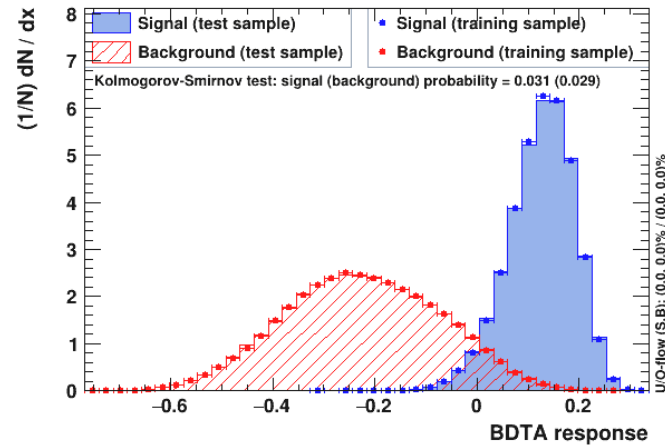
Background : All **Experimental Data** Events +(Simulated Non Truth)

Training variables + E_γ

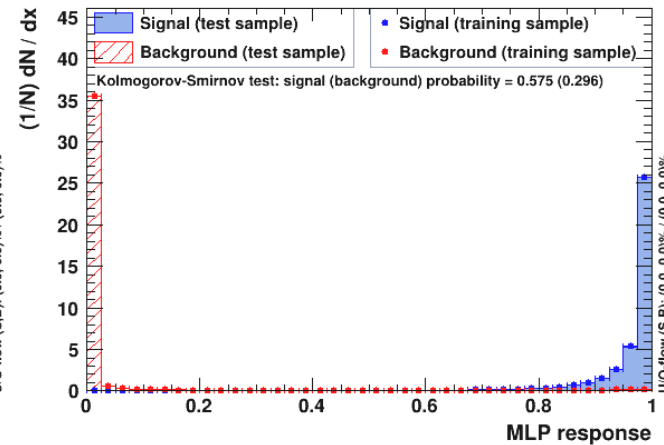
Particle(ΔTime , ΔEnergy , P , V_z , θ , φ)

200k Training and Test data

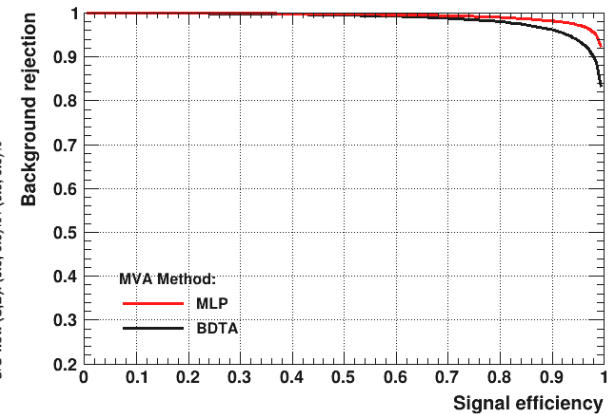
TMVA overtraining check for classifier: BDTA



TMVA overtraining check for classifier: MLP



Background rejection versus Signal efficiency

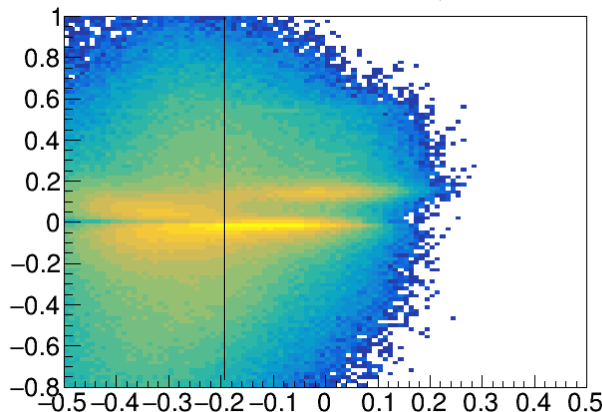
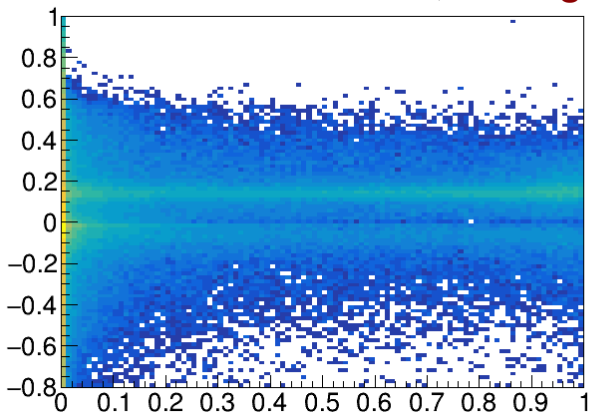


Apply Simulated training

MissMass V MLP (with E_γ)

Log z scale

MissMass V BDT (with E_γ)

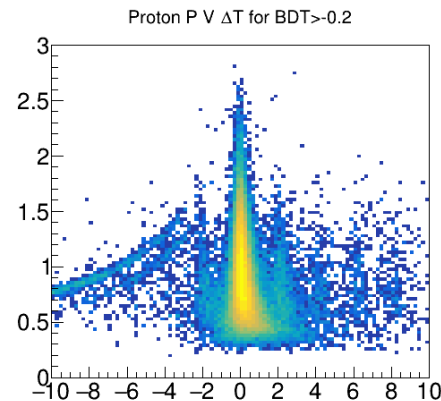
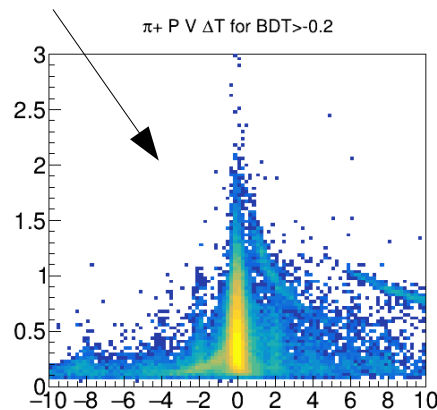


Real Experiment
distributions shifted to
background values
compared to Simulation

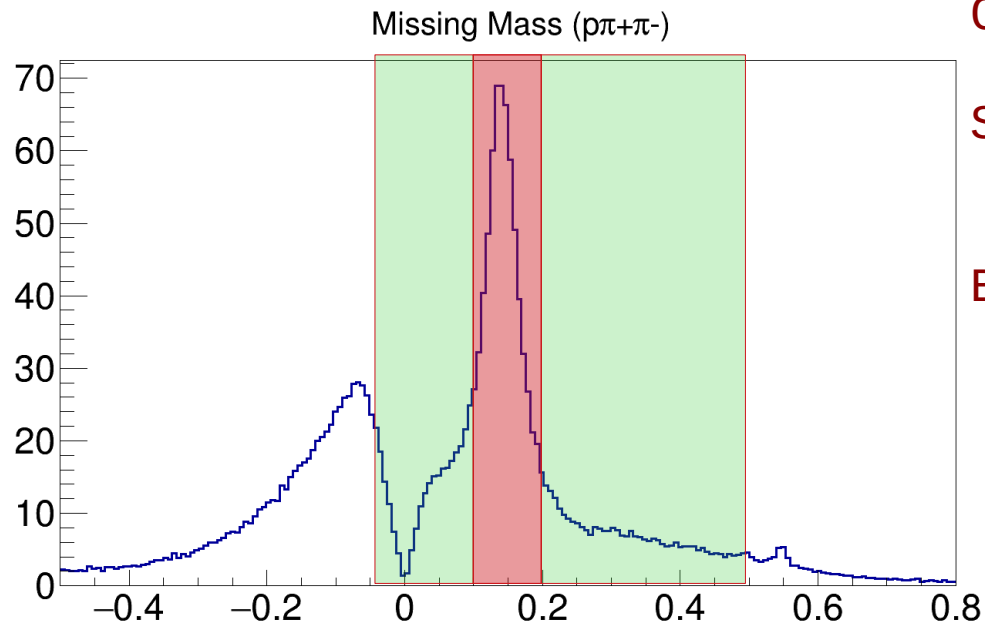
Simulation must be very
Realistic!

Small separation between 2π and 3π
- OK PID variables similar

Cuts still clean up
data, but must use
low BDT value



Mixed events training



Cut on region with high signal density

Signal = (Missing mass peak $\sim \pi^0$)

Background = sidebands inc 2π

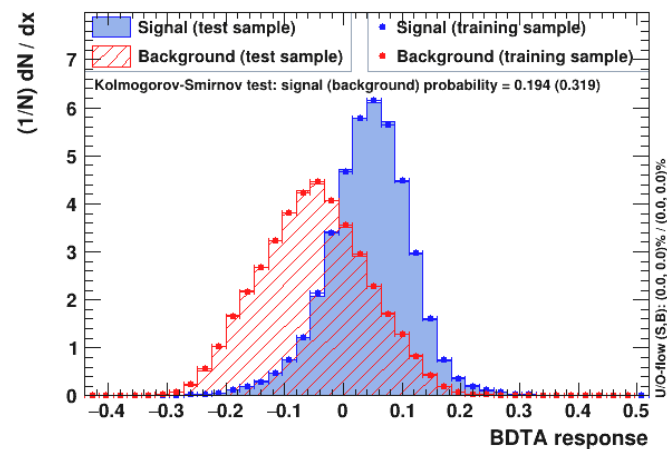
Train with Mixed Events (III)

Signal : $0.1 < \text{MissMass} < 0.2$

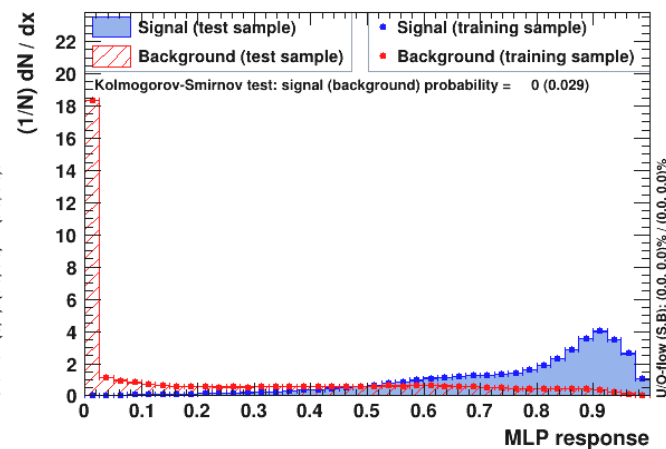
Background : $-0.04 < \text{MissMass} < 0.1$ and $0.2 < \text{MissMass} < 0.5$

Variables + E_y

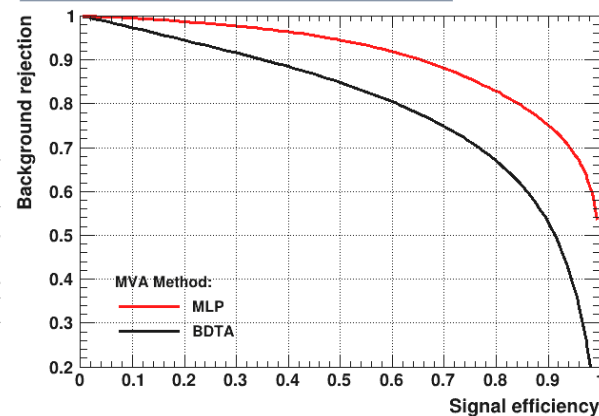
TMVA overtraining check for classifier: BDTA



TMVA overtraining check for classifier: MLP



Background rejection versus Signal efficiency



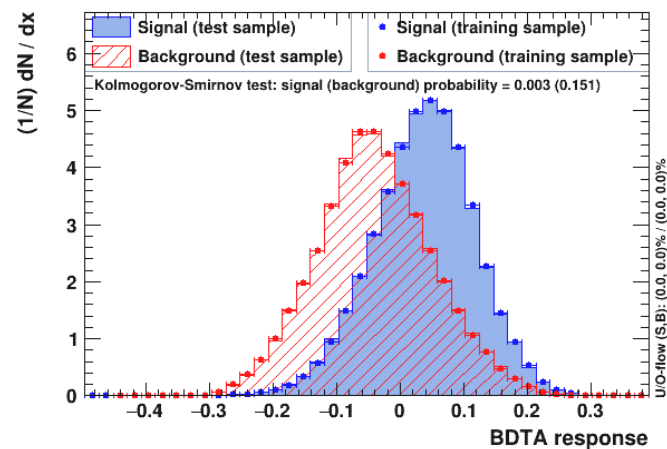
Train with Mixed Events (IV)

Signal : $0.1 < \text{MissMass} < 0.2$

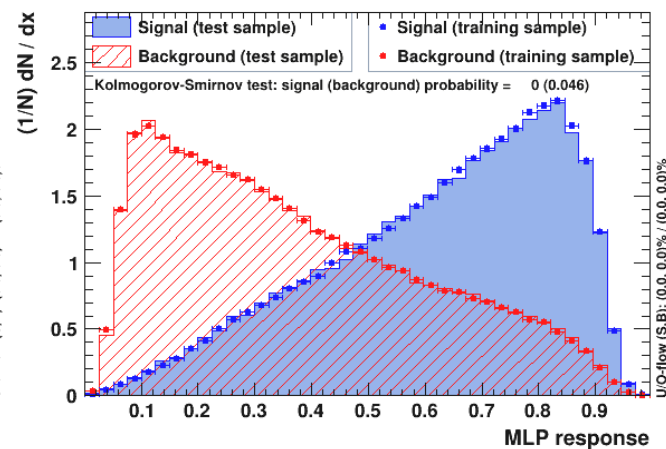
Background : $-0.04 < \text{MissMass} < 0.1$ and $0.2 < \text{MissMass} < 0.5$

Variables - E_y

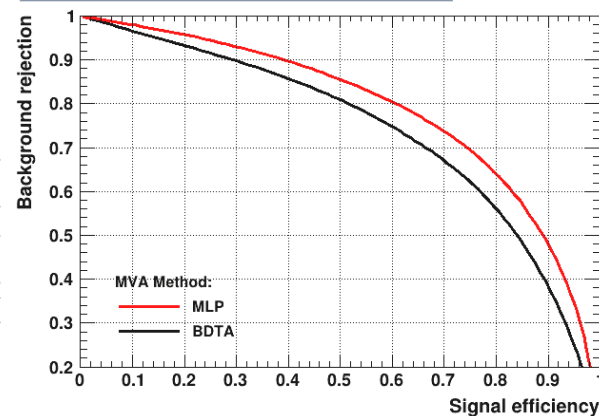
TMVA overtraining check for classifier: BDTA



TMVA overtraining check for classifier: MLP

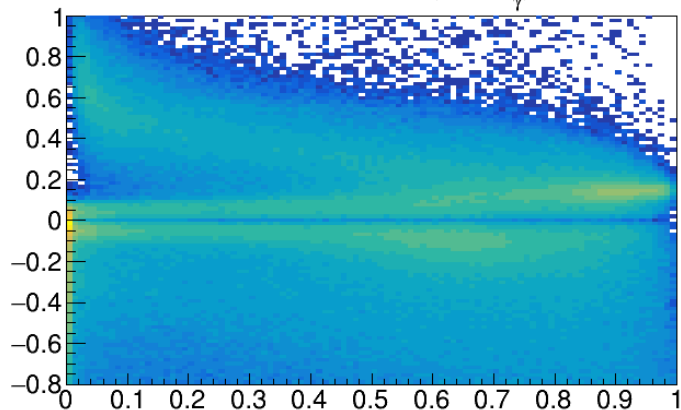


Background rejection versus Signal efficiency

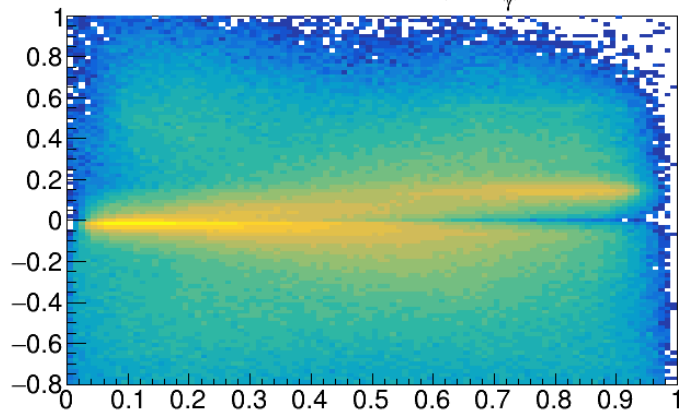


Miss Mass V Responses

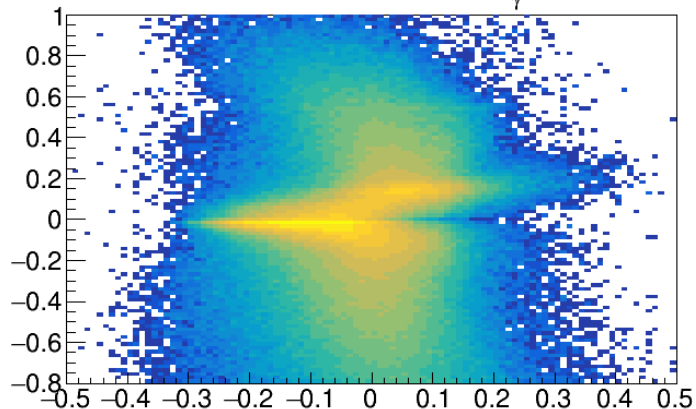
MissMass V MLP (with E_γ)



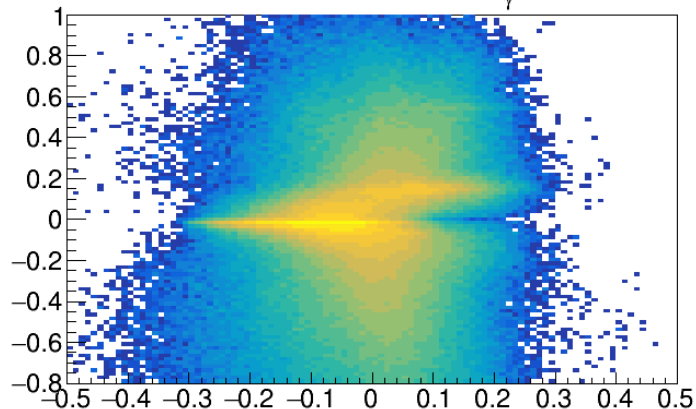
MissMass V MLP (no E_γ)



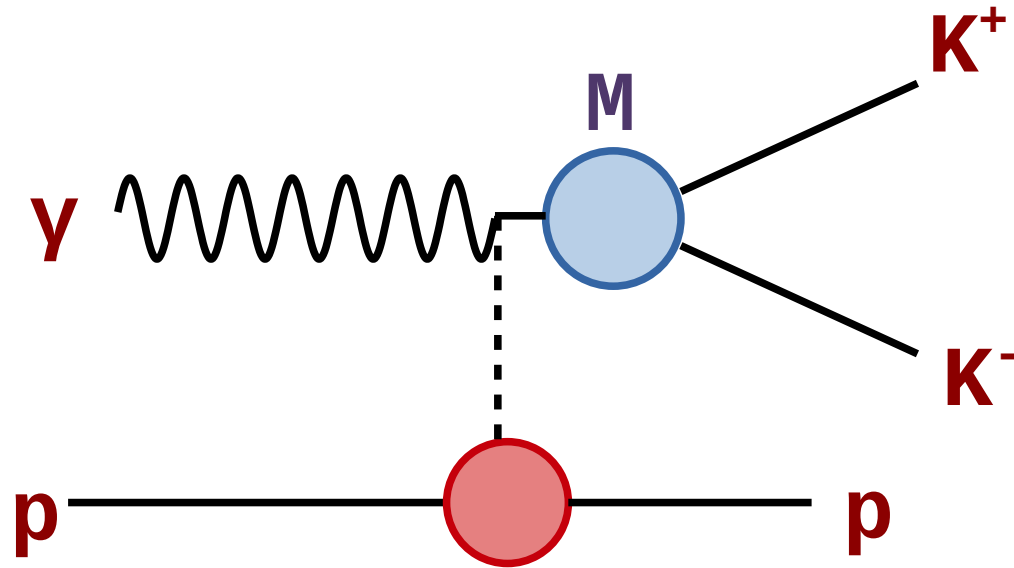
MissMass V BDT (with E_γ)



MissMass V BDT (no E_γ)



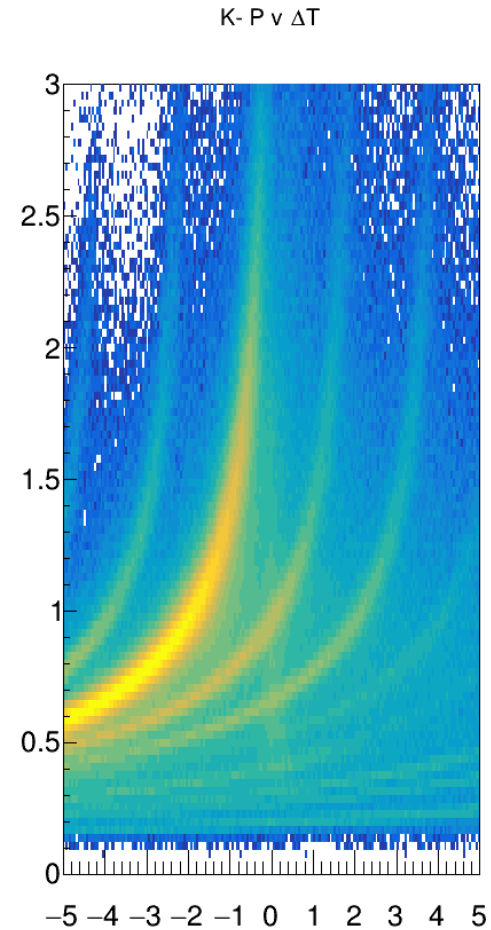
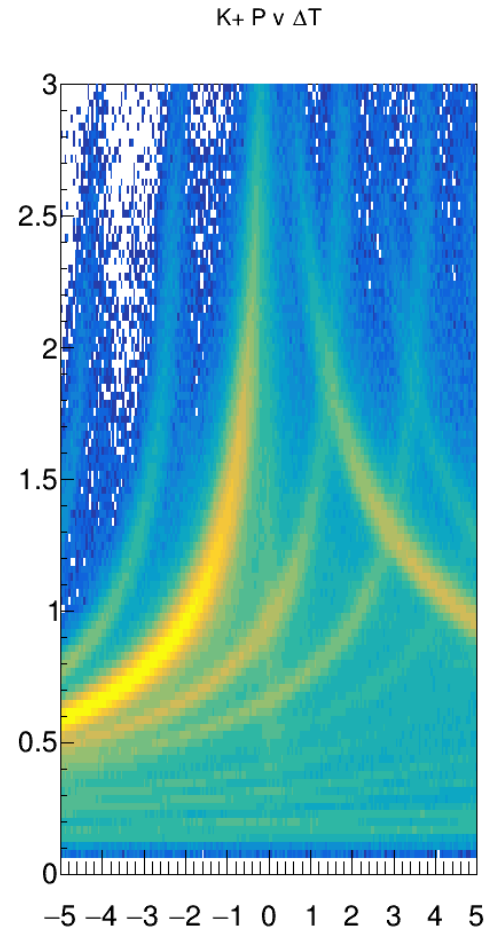
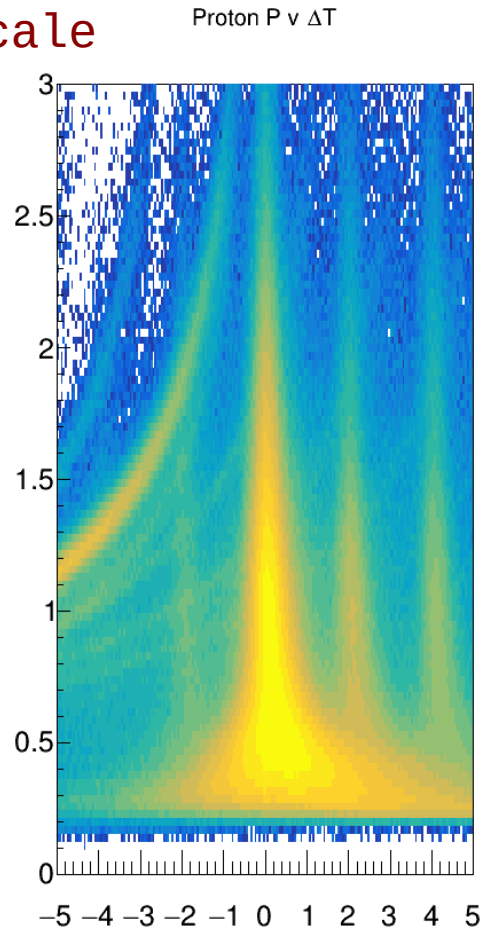
Example Reaction



Large background from
Pions IDed as Kaons

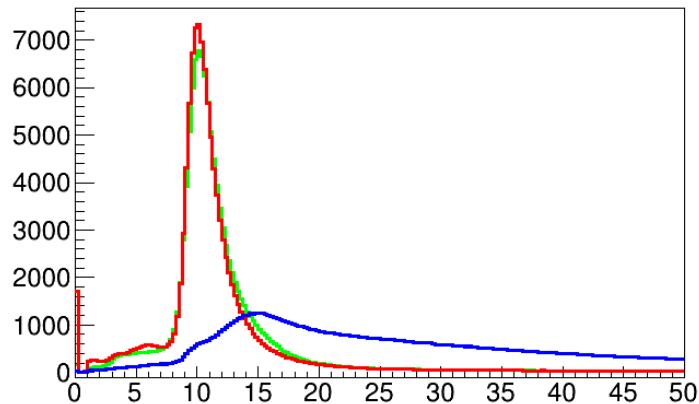
Δ Time PID

*Log scale



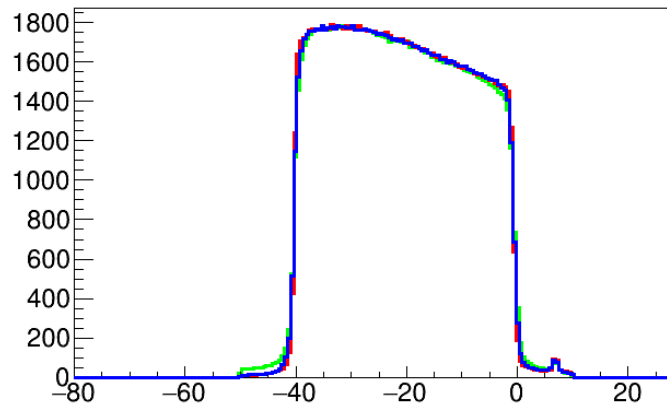
Other Input Variables

ToF Scint. ΔE



θ

Z Vertex

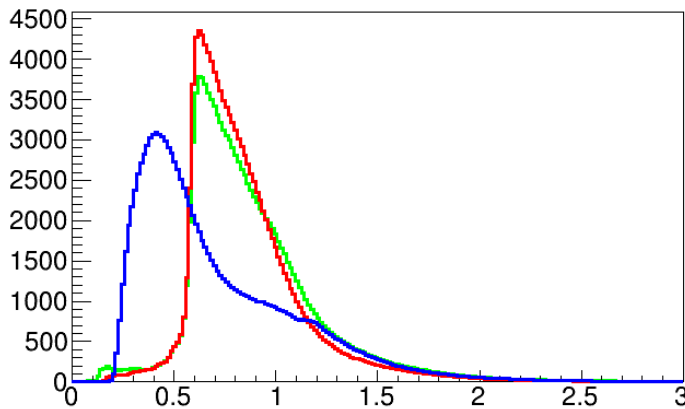
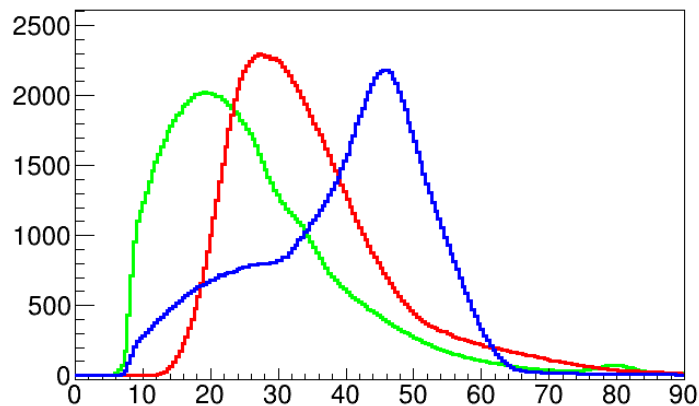


Proton

K^+

K^-

Momentum



Train BDT with sWeights

SPlot – technique for disentangling different event species using a discriminatory variable (generalised side-band subtraction)

M. Pivk, F.R. Le Diberder, Nucl.Inst.Meth.A 555, 356-369, 2005

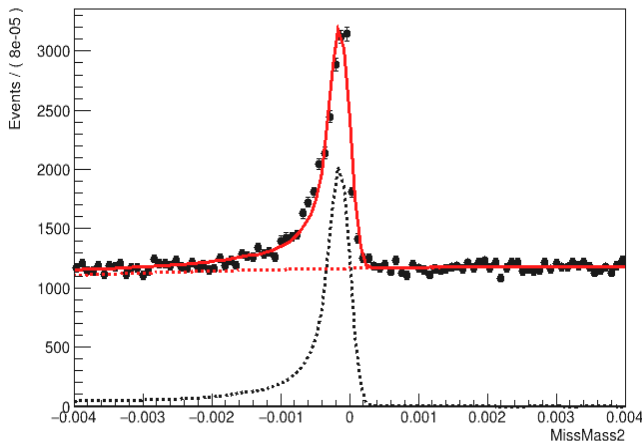
Used RooStats implementation

Exclusive

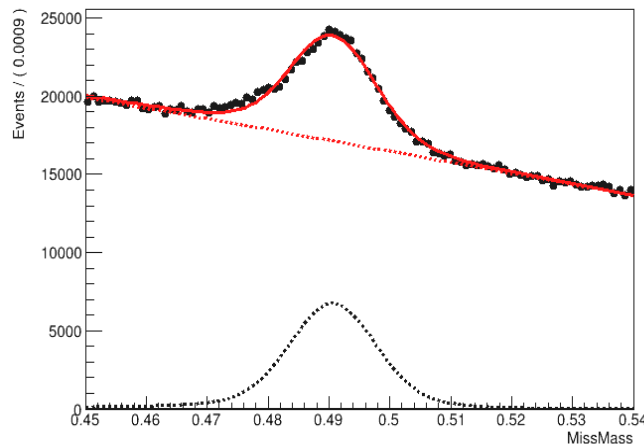
Missing K-

Missing K+

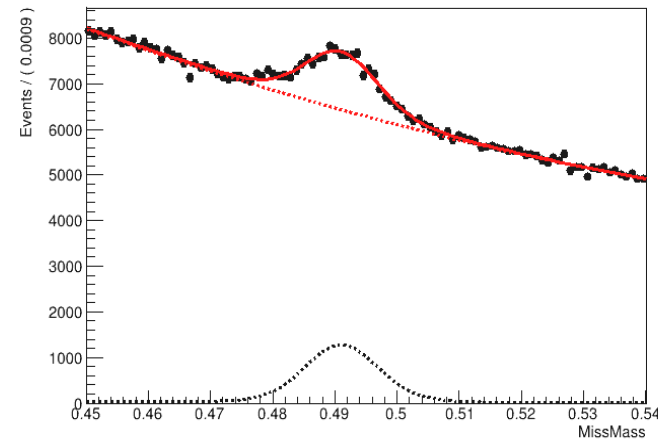
Fit components for MissMass2



Fit components for MissMass



Fit components for MissMass



In TMVA BDT accept negative event weights
MLP does not

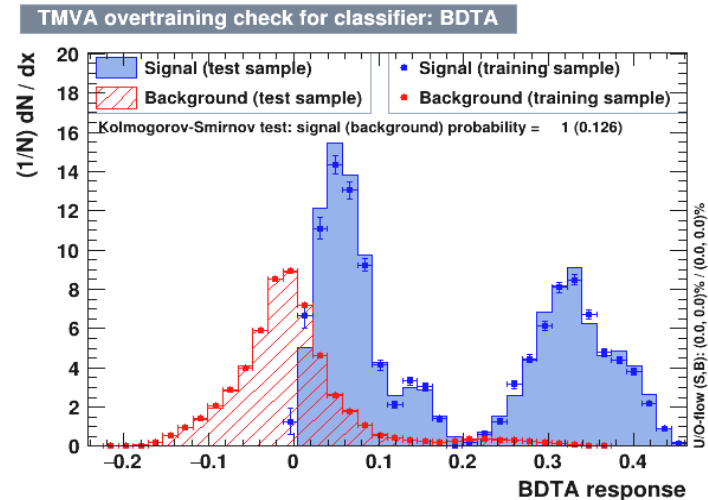
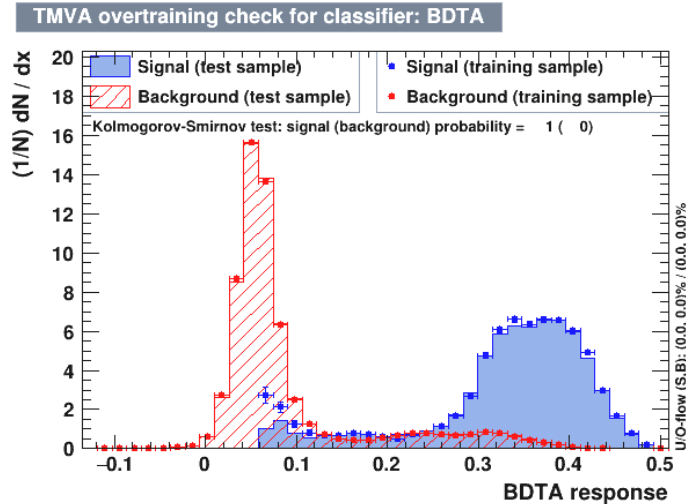
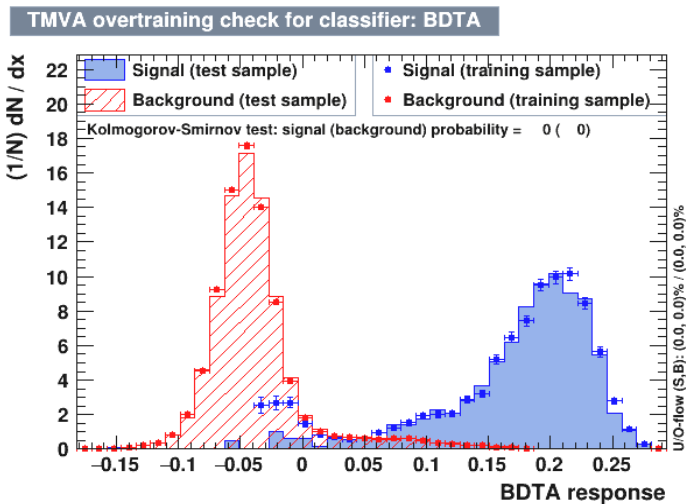
$${}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}$$

Train BDT with sWeights

80k test and 20k train
Exclusive

300k test and train
Missing K-

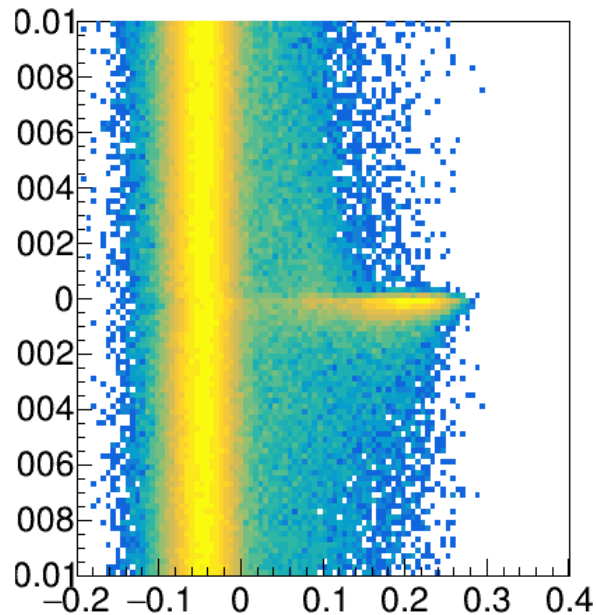
300k test and train
Missing K+



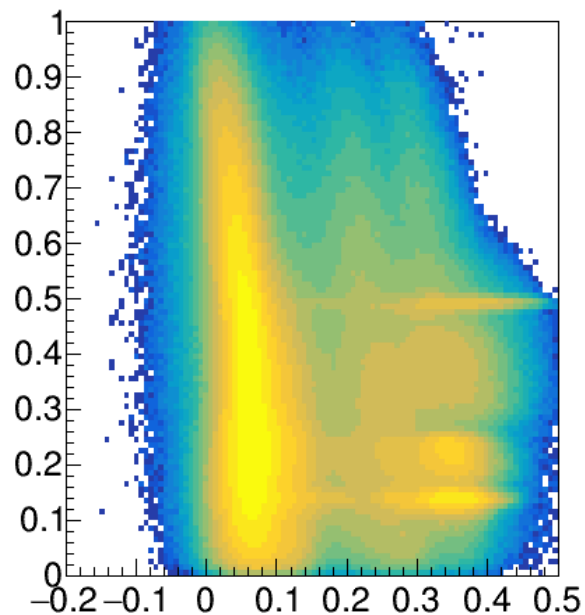
2 types of events
in Missing K+ signal
weights
Seperated by training

Exclusivity variables V BDT

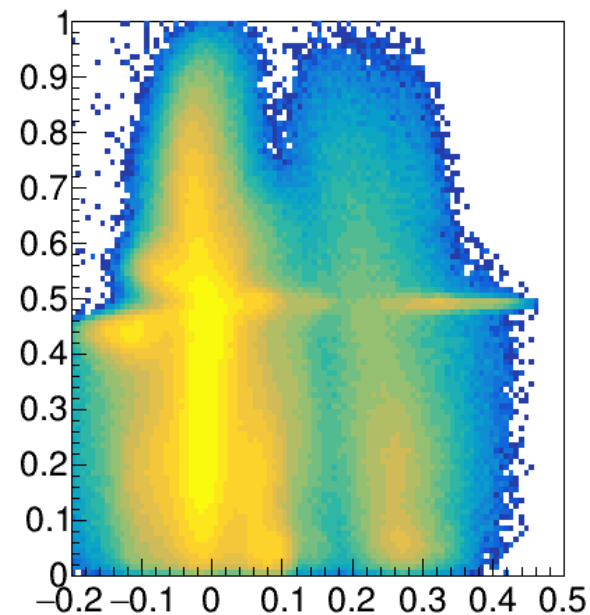
Exclusive MissMass2 V BDT



MissK - MissMass V BDT

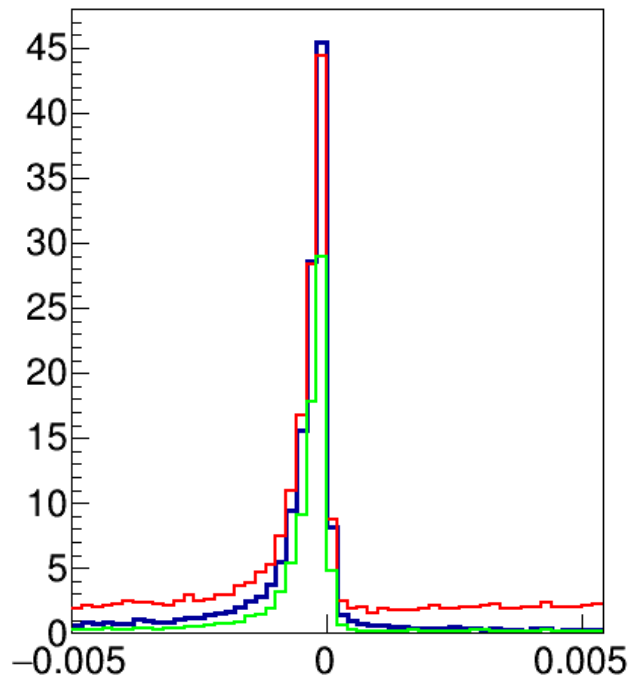


MissK+ MissMass V BDT

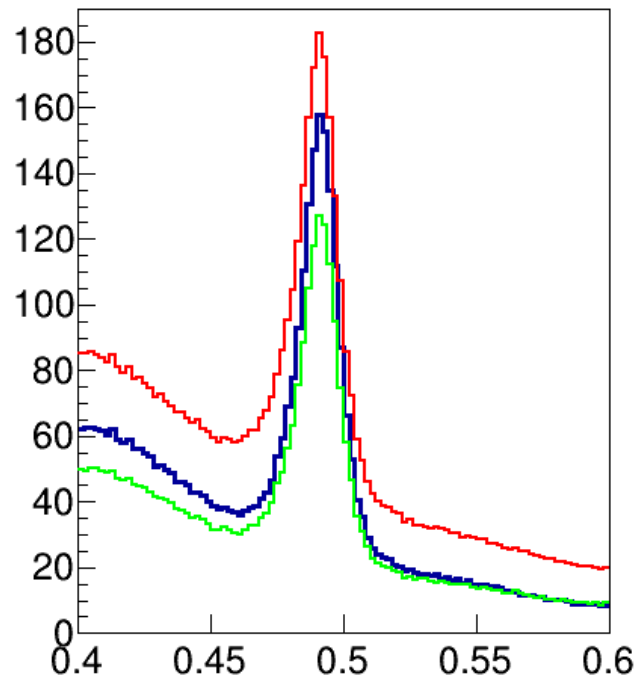


Event exclusivity with BDT cut

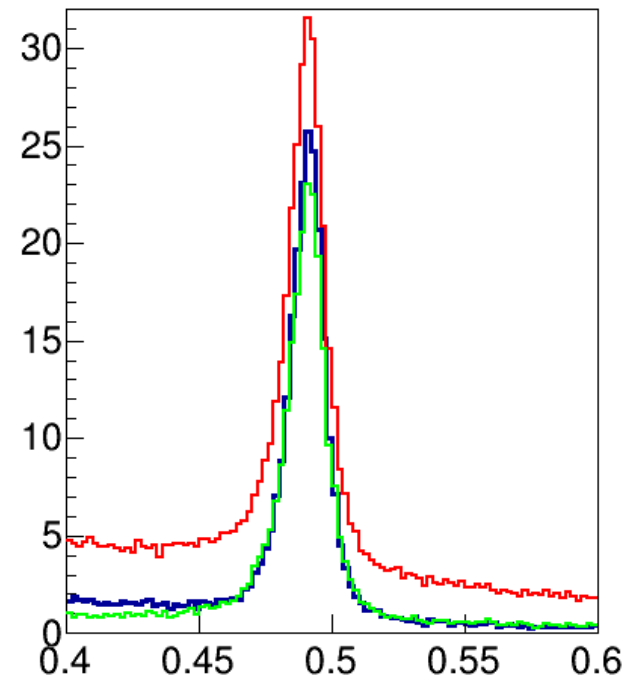
Exclusive MissMass Squared BDT>0.1



MissMass(pK-) BDT>0.28



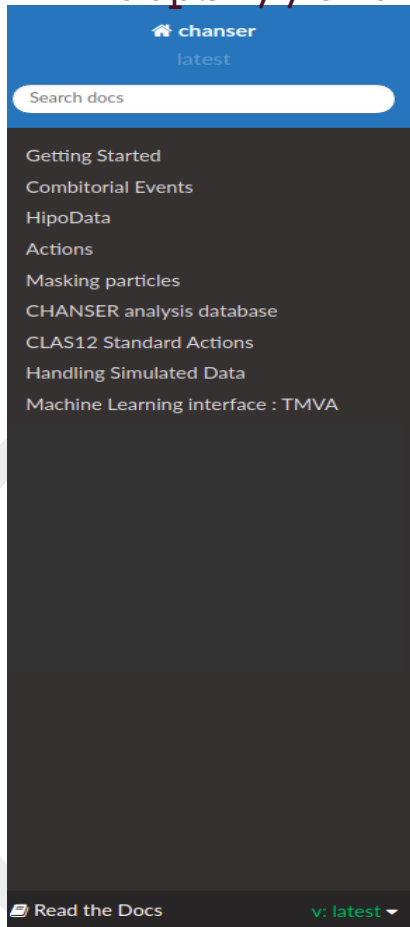
MissMass(pK+) BDT>0.28



Loose Cuts Based(1ns)
Tight Cuts Based(0.5ns)
BDT Cut

CHANSER CLAS12 HIPO ANALYSER

<https://chanser.readthedocs.io/en/latest/>



» Welcome to chanser's documentation!

[Edit on GitHub](#)

Welcome to chanser's documentation!

- [Getting Started](#)
 - [Installation](#)
 - [Creating your own final state class](#)
 - [Configuring an analysis](#)
 - [Running the analyse](#)
- [Combinatorial Events](#)
 - [Combinatorials of the first kind](#)
 - [Combinatorials of the second kind](#)
 - [Combinatorials of the third kind](#)
 - [Combining Combinatorials of different kinds](#)
 - [Setting combinatorial behaviour for configured FinalState objects](#)
- [HipoData](#)
- [Actions](#)
 - [StartTime Actions](#)
 - [Particle Data Actions](#)
 - [Particle Cut Actions](#)
 - [Particle Correction Actions](#)
 - [Quick Usage](#)
- [Masking particles](#)
 - [Masking Calorimeter split-off clusters](#)
 - [Masking Radiated photons](#)
- [CHANSER analysis database](#)
 - [Including Parameter Lookup in Code](#)
 - [Defining Parameters in .db files](#)
- [CLAS12 Standard Actions](#)
 - [Standard Run Group A Actions](#)

1. Be able to
Automate everything

2. Automate
everything

3. including ML

<https://github.com/dglazier/chanser/wiki/TMVA>

<https://indico.jlab.org/event/472/contributions/9070/attachments/7345/10144/AnalysisToolsCppROOT.pdf>

