Low Momentum Muon and Pion Separation for Higgsino Reconstruction

Hale Sert

DESY
Hamburg University

ILD Software and Analysis Meeting

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Collaborative Research Centre 676
Particles, Strings, and the Early Universe
The Structure of Matter and Space-Time

Light Higgsino Scenario

Motivated by naturalness which requires μ at the electroweak scale

Scenario contains

- 3 light higgsinos: $\tilde{\chi}_1^{\pm}$ & $\tilde{\chi}_1^0$ & $\tilde{\chi}_2^0$
- Almost mass degenerate: $\Delta M(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0)$ & $\Delta M(\tilde{\chi}_2^0, \tilde{\chi}_1^0)$ $\sim$ a (sub) GeV
- All other supersymmetric particles are heavy up to a few TeV

Benchmark point:

dm770

Mass Spectrum	
Particle	Mass (GeV)
h	127
$\tilde{\chi}_1^0$	166.59
$\tilde{\chi}_1^{\pm}$	167.36
$\tilde{\chi}_2^0$	167.63
H's	$\sim 10^3$
$\tilde{\chi}$'s	$\sim 2 - 3 \times 10^3$

$\Delta M(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = 0.77$ GeV

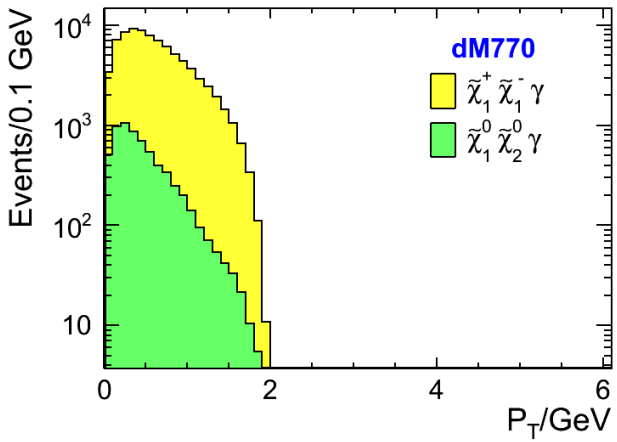
Production Processes & Decay Modes

Production Processes

- $e^+e^- \to \tilde{\chi}_1^+\tilde{\chi}_1^-$
- $e^+e^- \to \tilde{\chi}_1^0\tilde{\chi}_2^0$

Decay Modes

- $\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0 W^{\pm *}$
- $\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 Z^{0*}$
- $\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 \gamma$



P_T dist. of final state particles at generator level

- Muons & Pions are very soft!

Separation of Signal Processes

Exclusive decay modes:

- $\tilde{\chi}_1^+\tilde{\chi}_1^+ \to 2\tilde{\chi}_1^0\ W^{+*}\ W^{-*}$
  - semileptonic final state (35%)
- $\tilde{\chi}_1^0\tilde{\chi}_2^0 \to 2\tilde{\chi}_1^0\ Z^{0*}/\gamma$
  - photonic final state (74%)

In semileptonic decays

- $\mathrm{BR}(\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ \pi) \approx 60\ \%$
- $\mathrm{BR}(\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ \mu^-\nu_\mu) \approx 13\ \%$

- Muon & Pion separation plays an important role in this analysis

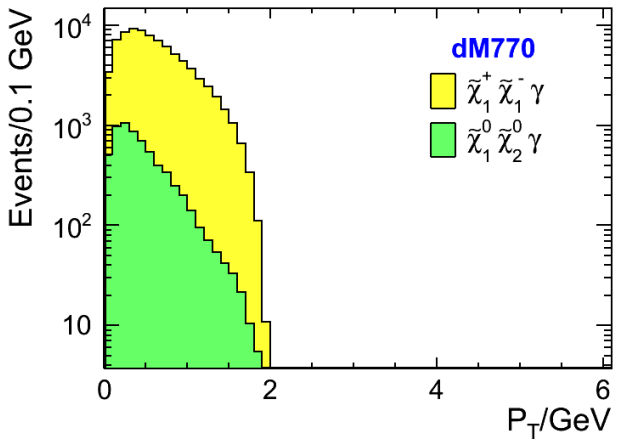
Production Processes & Decay Modes

Production Processes

- $e^+e^- \to \tilde{\chi}_1^+\tilde{\chi}_1^-$
- $e^+e^- \to \tilde{\chi}_1^0\tilde{\chi}_2^0$

Decay Modes

- $\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0 W^{\pm *}$
- $\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 Z^{0*}$
- $\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 \gamma$



P_T dist. of final state particles at generator level

- Muons & Pions are very soft!

Separation of Signal Processes

Exclusive decay modes:

- $\tilde{\chi}_1^+\tilde{\chi}_1^+ \to 2\tilde{\chi}_1^0\ W^{+*}\ W^{-*}$
  - semileptonic final state (35%)
- $\tilde{\chi}_1^0\tilde{\chi}_2^0 \to 2\tilde{\chi}_1^0\ Z^{0*}/\gamma$
  - photonic final state (74%)

In semileptonic decays

- $\mathrm{BR}(\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ \pi) \approx 60\ \%$
- $\mathrm{BR}(\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0\ \mu^-\nu_\mu) \approx 13\ \%$

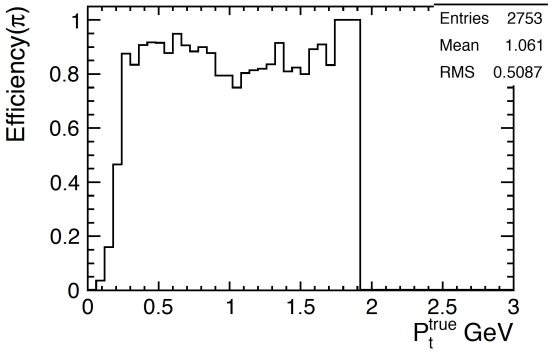
- Muon & Pion separation plays an important role in this analysis

Muon and Pion Reconstruction in Full Simulation

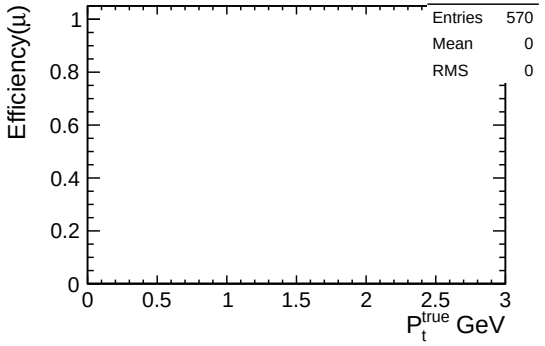
- Since muons are very soft, they don't leave a proper signature in the calorimeters. They get stuck in the ECAL or in the HCAL.

Reconstruction Efficiency of

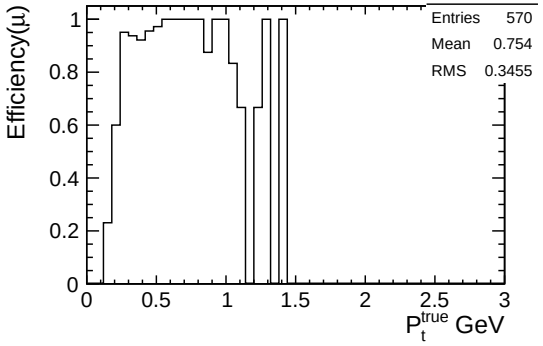
Pions



Muons



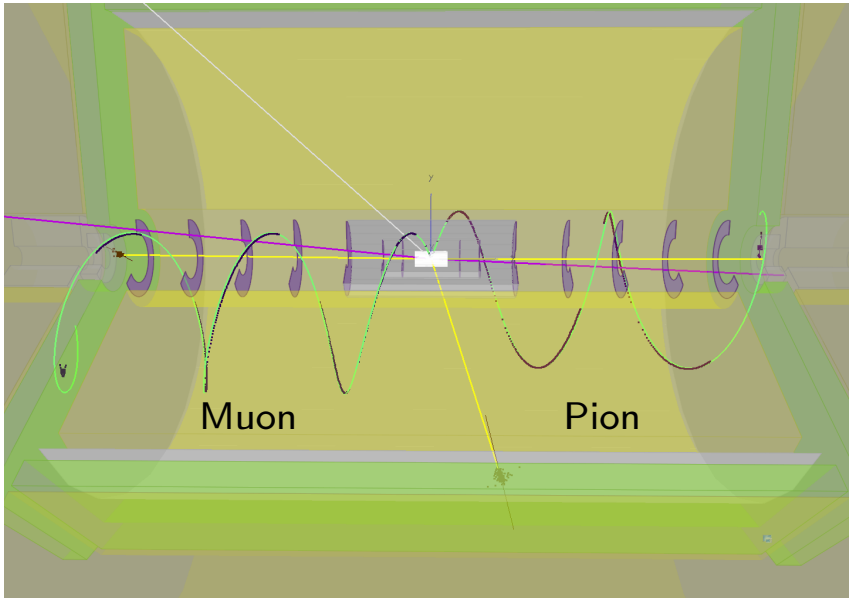
Muons as Pions

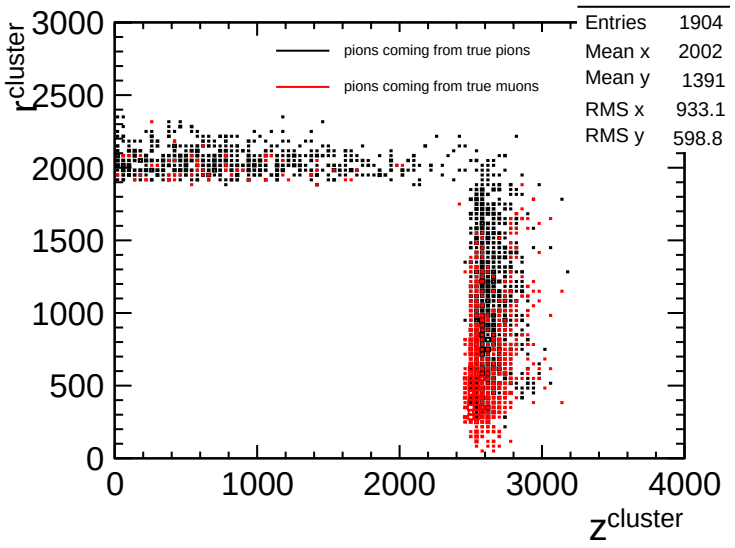


- In most of the cases they are reconstructed as pions
- Pandora PFO reconstruction does not work for separation of muon & pion at low momentum region
- Need to separate them

Muon and Pion Separation in Calorimetry

Checked the behaviours of muon and pion using event display

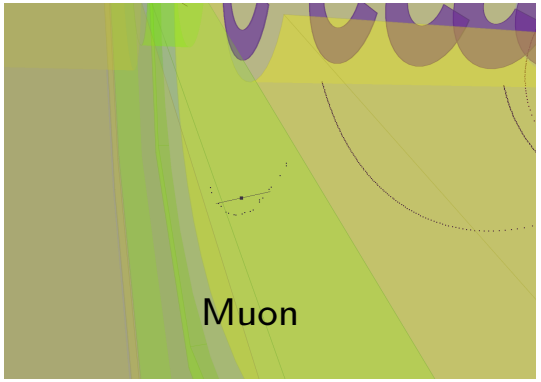


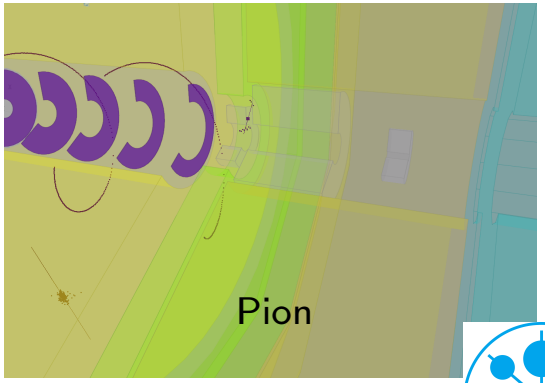


- The particles curl and travel along the magnetic field lines, and hit the endcap calorimeters

Cluster properties are studied

- Observed that the cluster shape of muon and pion is different





Differences in Cluster Shape

To examine the cluster shape, generate pure μ & π samples using particle gun

- From 0.2 GeV to 2 GeV momentum in the interval of 0.1 GeV
- Shot gun directly to the calorimetry to get rid of decaying of particles & curling of low energetic particles during their way on TPC

- Determine variables showing the differences of the cluster shapes:

** Depth of the cluster

$$Z_{cluster} = D_{cluster} \times \cos\alpha$$

** Cluster energy/track(true) momentum

** Radius distribution of hits considering cluster position as a reference point

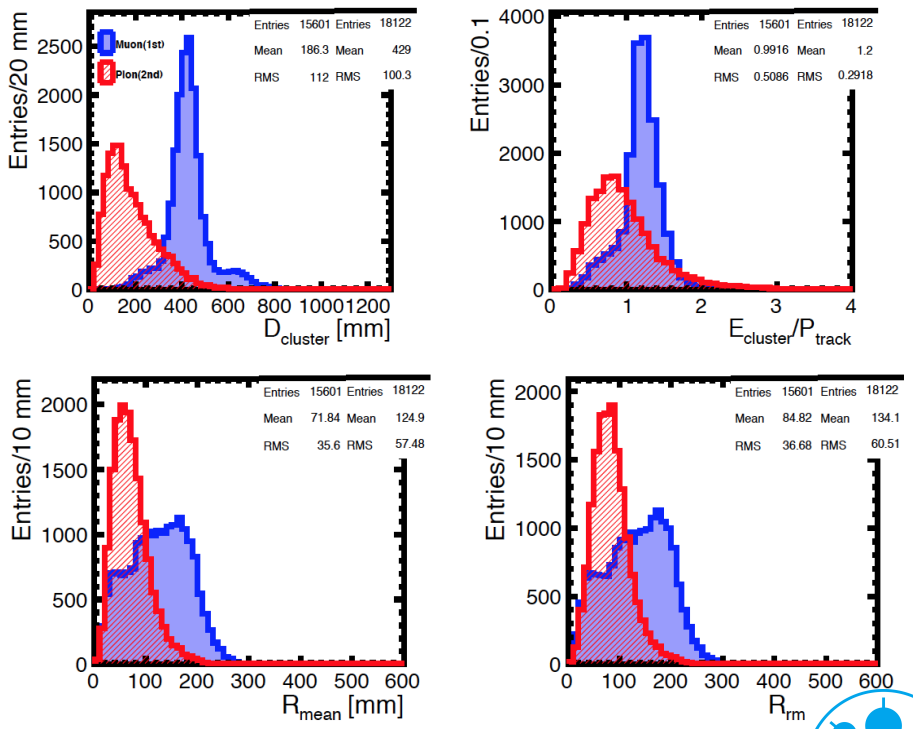
$$R_{hit} = \sqrt{((x_{hit} - x_{clus})^2 + (y_{hit} - y_{clus})^2)}$$

- To have a variable changes per cluster one could consider the below variables

** Mean Value of R_{hits}

** RMS Value of R_{hits}

1 GeV P

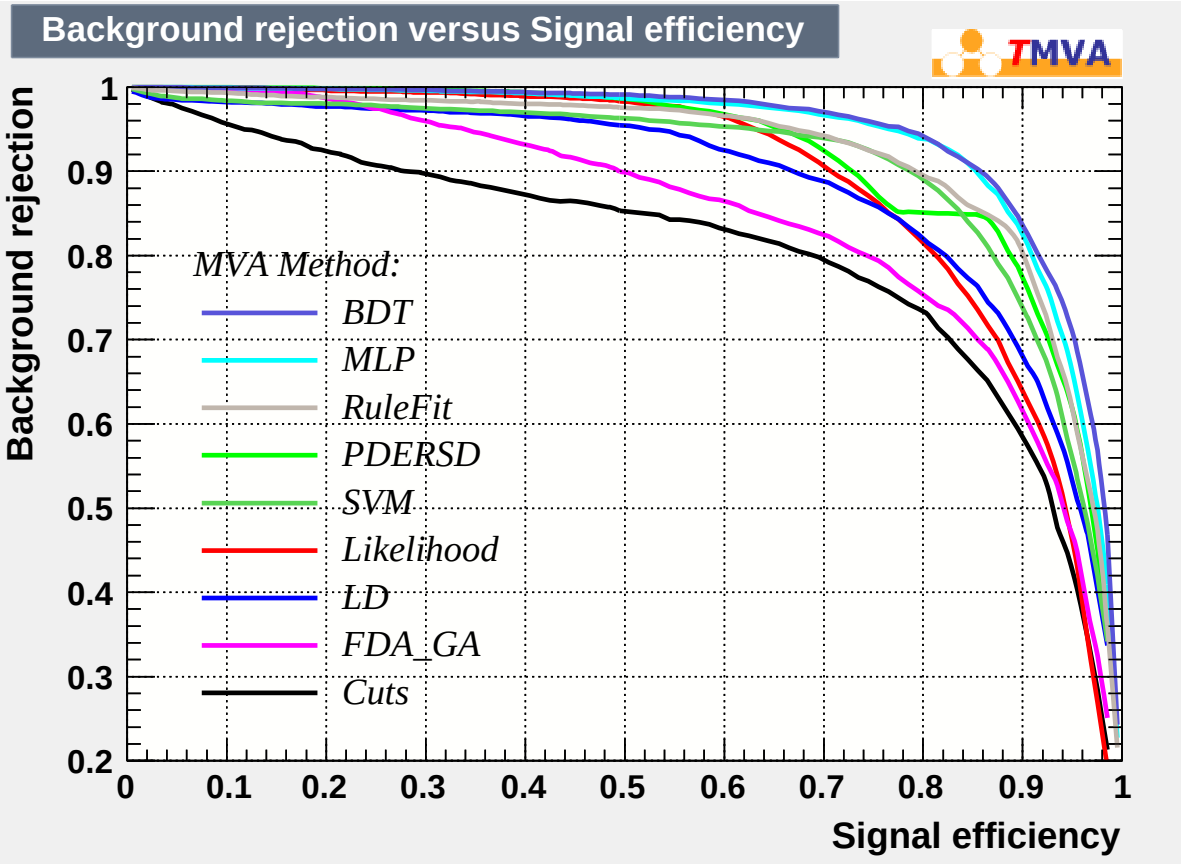


Multivariate Analysis – Testing All Methods in TMVA

Checked all the methods to choose the best one

ROC curve Comparison of All TMVA Methods (ROC-Receiver Operating Characteristic)

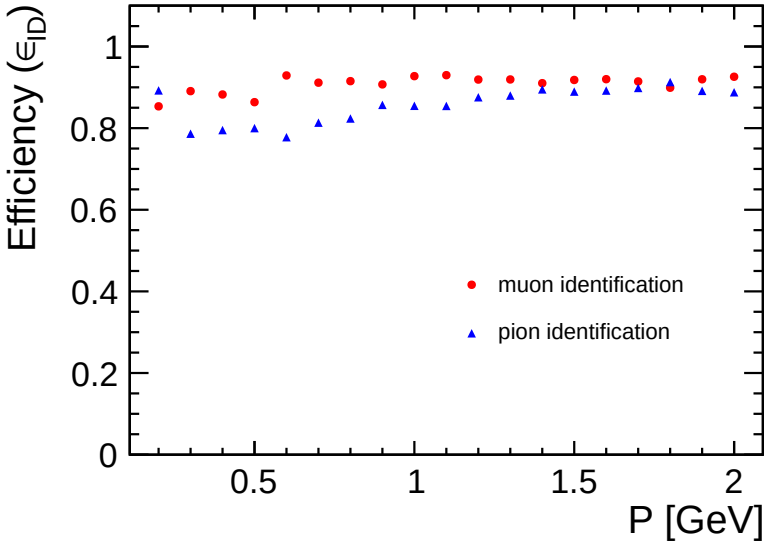
P = 0.2 GeV



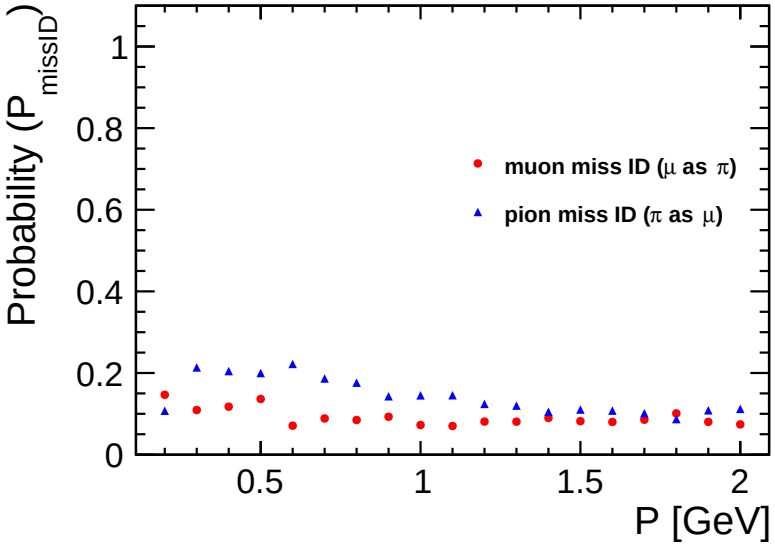
➤ BDT is the best method!

Identification Efficiencies with BDTG Method

Identification Efficiency



Miss Identification Probability



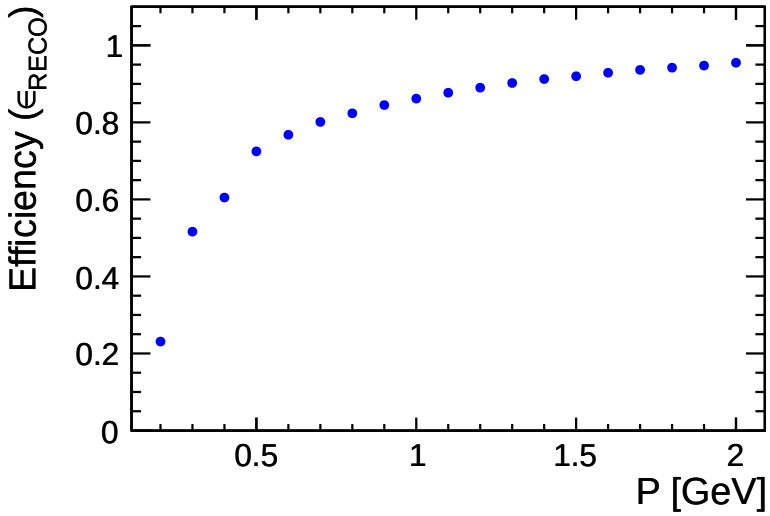
- $\epsilon^{\pi}_{ID} = \frac{N_{MC_{\pi(rc\pi)}}}{N_{MC_{\pi(rc\pi)}} + N_{MC_{\pi(rc\mu)}}}$
- $\epsilon^{\mu}_{ID} = \frac{N_{MC_{\mu(rc\mu)}}}{N_{MC_{\mu(rc\mu)}} + N_{MC_{\mu(rc\pi)}}}$
- $\epsilon^{\pi}_{missID} = \frac{N_{MC_{\pi(rc\mu)}}}{N_{MC_{\pi(rc\pi)}} + N_{MC_{\pi(rc\mu)}}}$
- $\epsilon^{\mu}_{missID} = \frac{N_{MC_{\mu(rc\pi)}}}{N_{MC_{\mu(rc\mu)}} + N_{MC_{\mu(rc\pi)}}}$

- Efficiencies do not depend on momentum so much
- Muons can be identified with 90 % efficiency
- Pions can be identified with 85 % efficiency

Reconstruction Efficiency of Particle Gun Samples

Having PFO efficiency

- $\epsilon = \frac{Nevt_{reconstructed}}{Nevt_{generated}}$



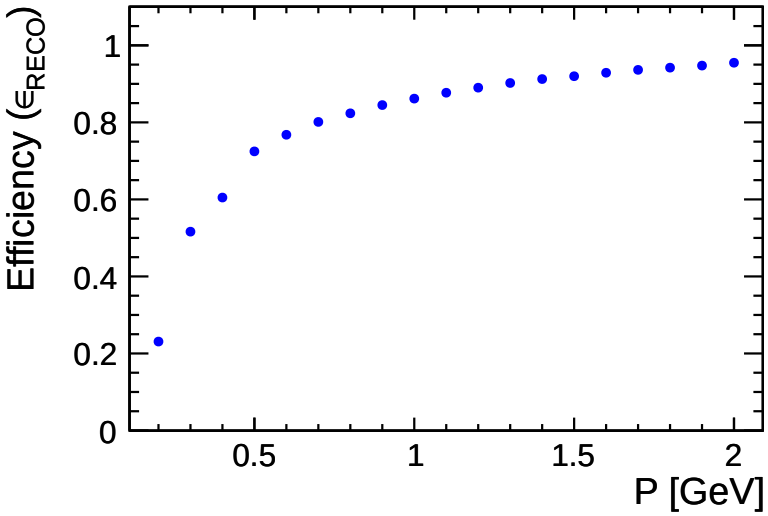
There are some events which has no reconstructed particles

- Checked if events have any hits or not => Most of them have hits!
- But they don't create cluster due to some cuts on Pandora (they are neutral no track info)
  - $N_{hits} > 5$
  - $E_{hadrons} > 0.25$
- Decided to use calorimeter hits and checked if that helps

Reconstruction Efficiency of Particle Gun Samples

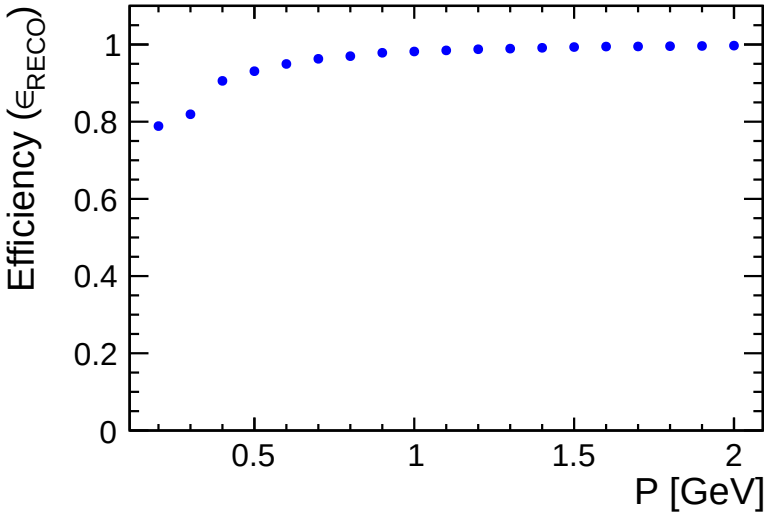
Having PFO efficiency

- $\epsilon = \frac{Nevt_{reconstructed}}{Nevt_{generated}}$



Having hit efficiency

- $\epsilon = \frac{Nevt_{hits}}{Nevt_{generated}}$



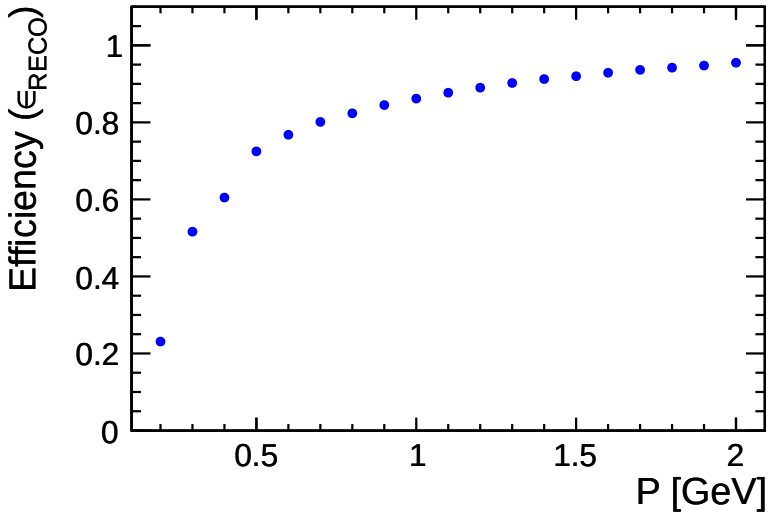
There are some events which has no reconstructed particles

- Checked if events have any hits or not => Most of them have hits!
- But they don't create cluster due to some cuts on Pandora (they are neutral no track info)
  - $N_{hits} > 5$
  - $E_{hadrons} > 0.25$
- Decided to use calorimeter hits and checked if that helps

Reconstruction Efficiency of Particle Gun Samples

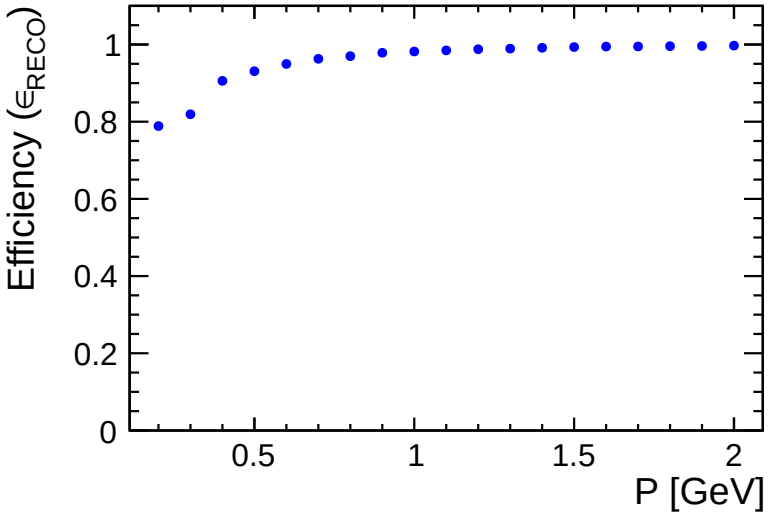
Having PFO efficiency

- $\epsilon = \frac{Nevt_{reconstructed}}{Nevt_{generated}}$



Having hit efficiency

- $\epsilon = \frac{Nevt_{hits}}{Nevt_{generated}}$



There are some events which has no reconstructed particles

- Checked if events have any hits or not => Most of them have hits!
- But they don't create cluster due to some cuts on Pandora (they are neutral no track info)
  - $N_{hits} > 5$
  - $E_{hadrons} > 0.25$
- Decided to use calorimeter hits and checked if that helps

Selection of the proper hits

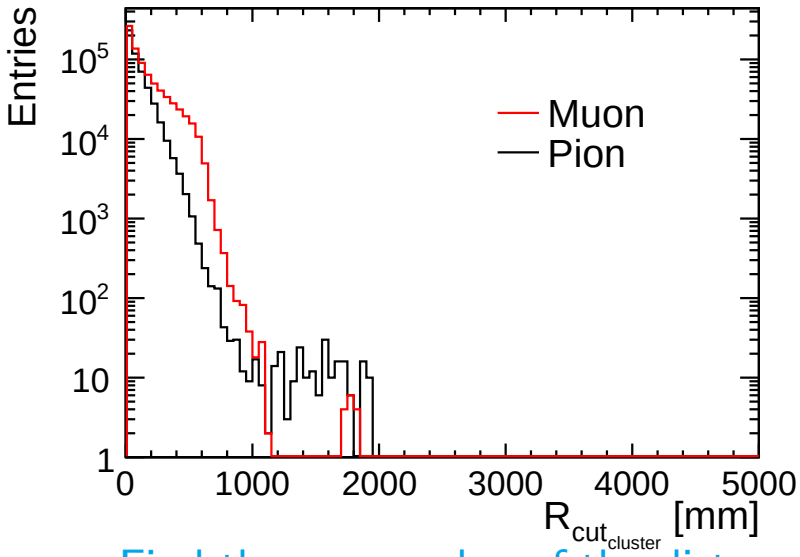
The noise hits needs to be suppressed. Therefore, we have required

- $R_{cut} < R_{cut_{cluster}}$ to determine a region around the MC particle
- $N_{hits} > 2$, at least 3 hits within the selected region

- To choose the region around the place where the particle are created
  - $R_{cut} = \sqrt{(mcvtx - htposx)^2 + (mcvty - htposy)^2}$

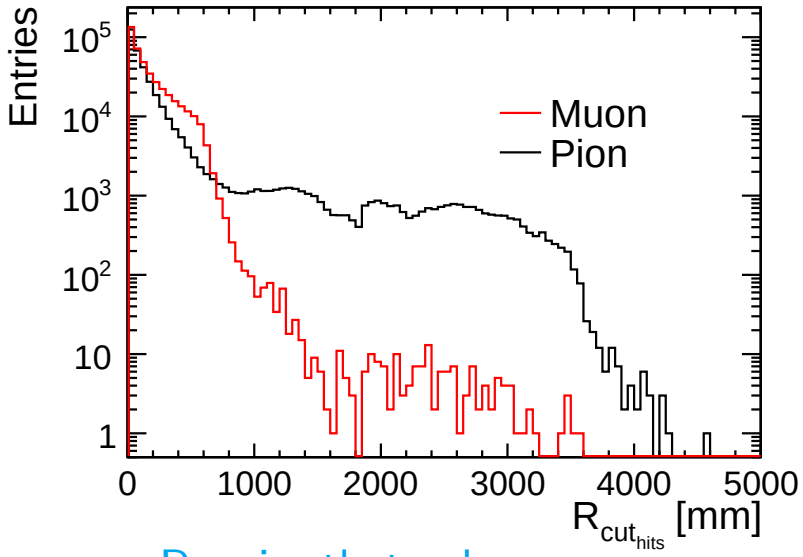
where $mcvtx, mcvty$ are the x and y position of the production vertex of MC particles with maximum momentum

Hits assigned to a cluster



Find the max. value of the dist.

All hits



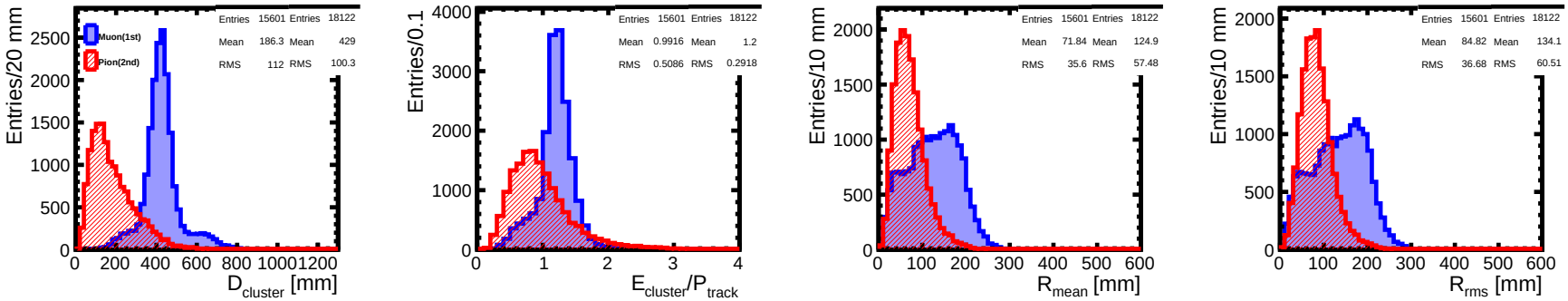
Require that value as a max

Used parameters for the separation are calculated using 3 different cases:

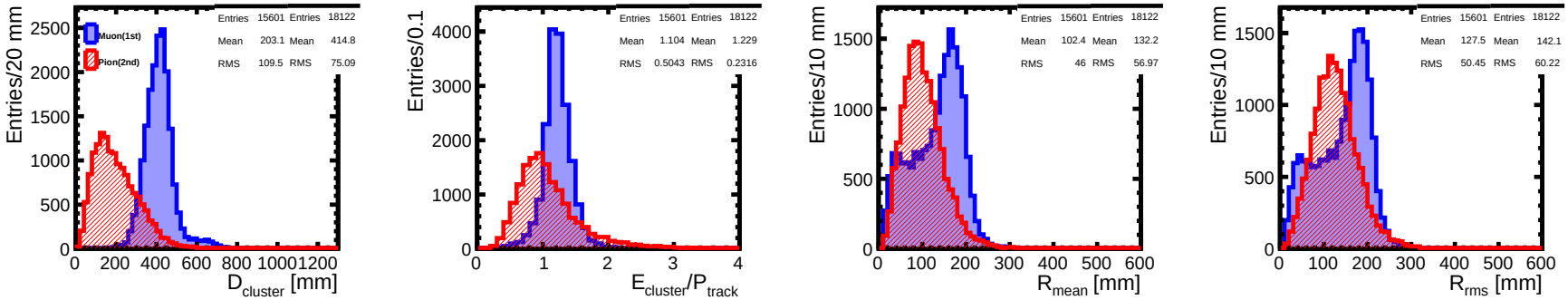
- Hits assigned to cluster = Cluster parameters via Pandora PFA
  [Pandora cluster hits]
- Hits of the events which have cluster
  [Hits – events having cluster]
- Hits of all the events, even if there is no Pandora cluster
  [Hits of all the events]

TMVA input variables @ 1 GeV Momentum

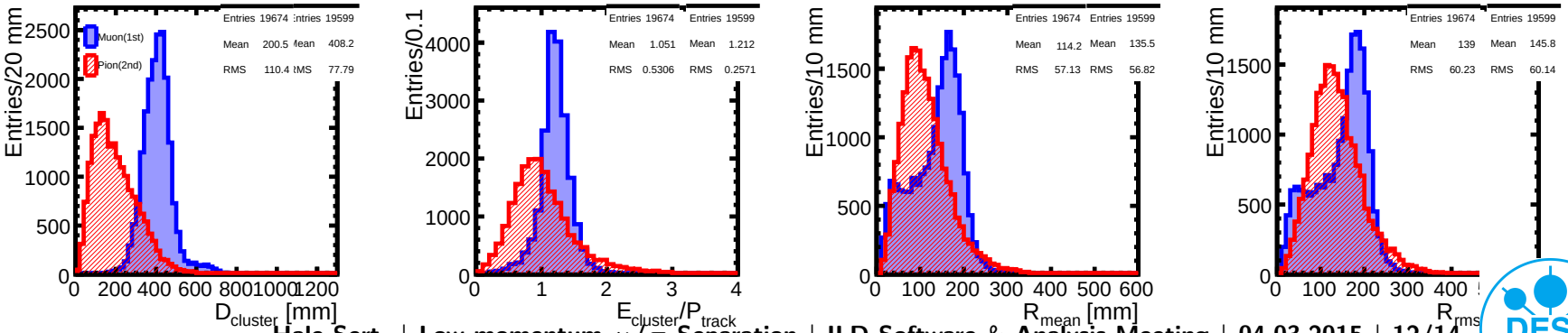
Pandora Cluster hits



Hits – events having cluster [15601 + 18122 = 33723 / 40000 ≈ 0.84]

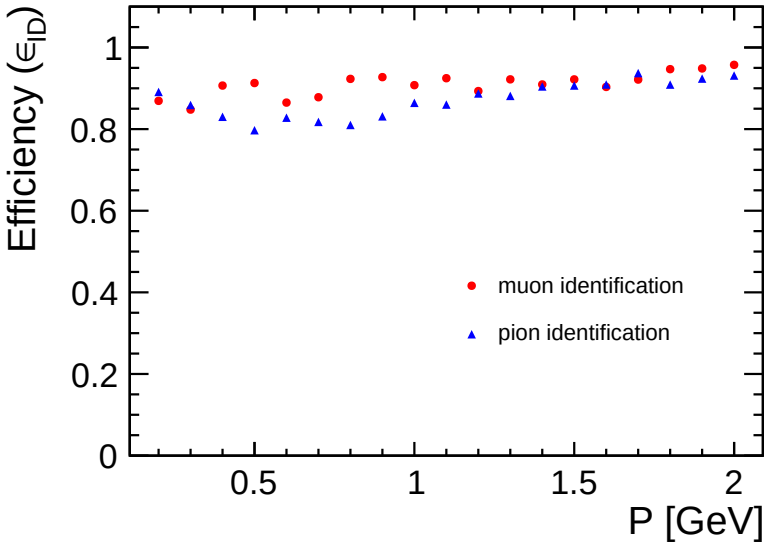


Hits of all the events [19674 + 19599 = 39961 / 40000 ≈ 0.98]

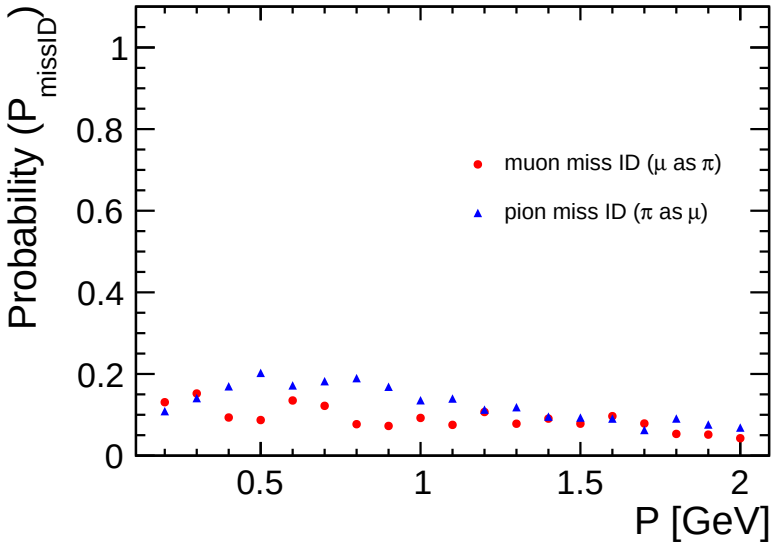


Identification Efficiencies using Calorimeter Hits with BDTG

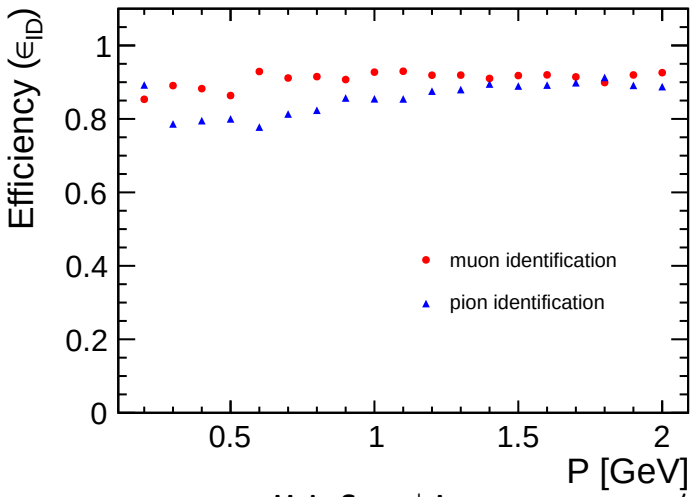
Identification Efficiency

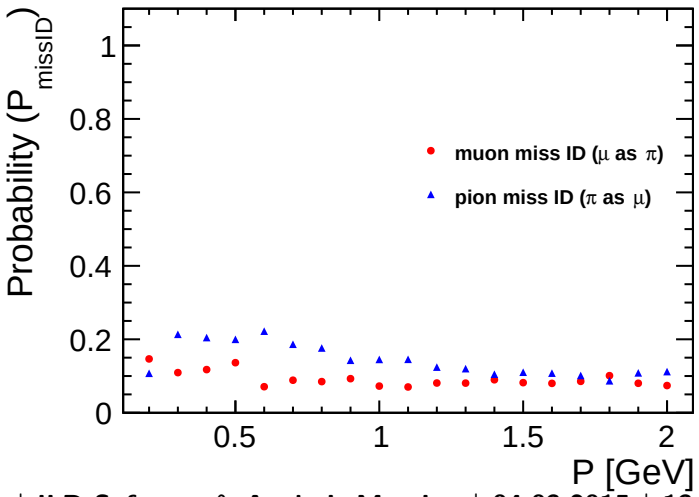


Miss Identification Probability



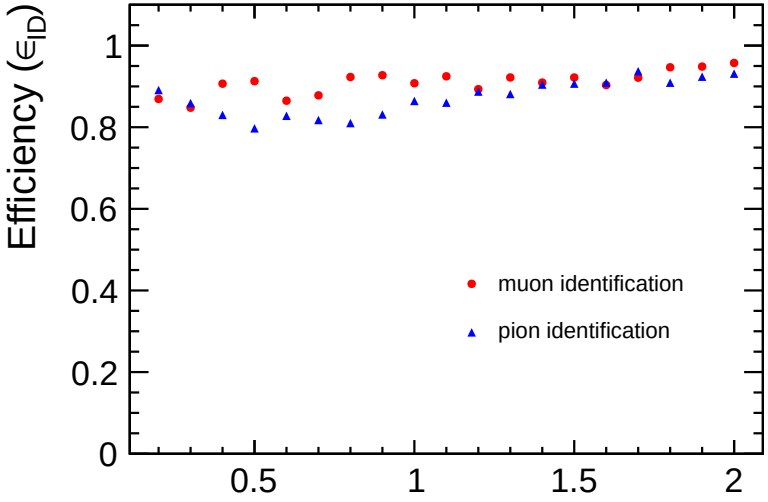
Using Cluster Parameter



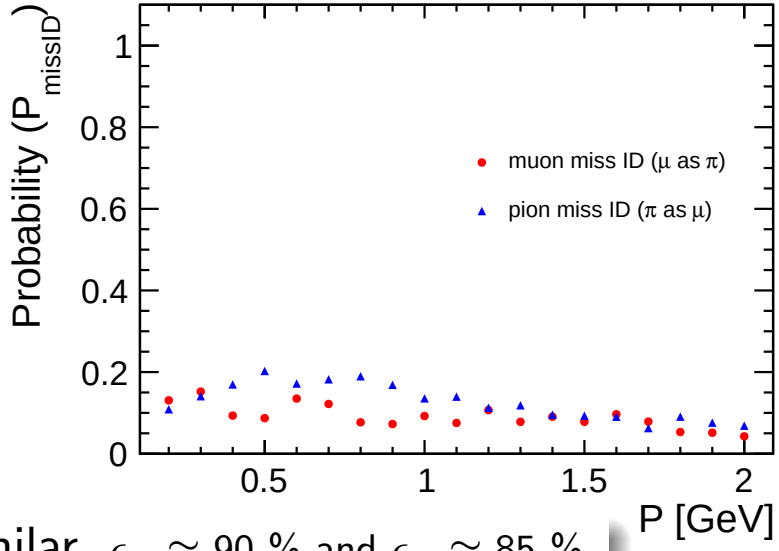


Identification Efficiencies using Calorimeter Hits with BDTG

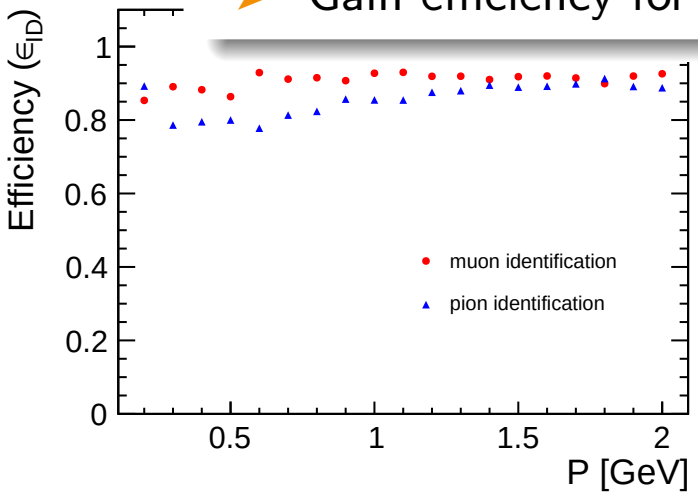
Identification Efficiency

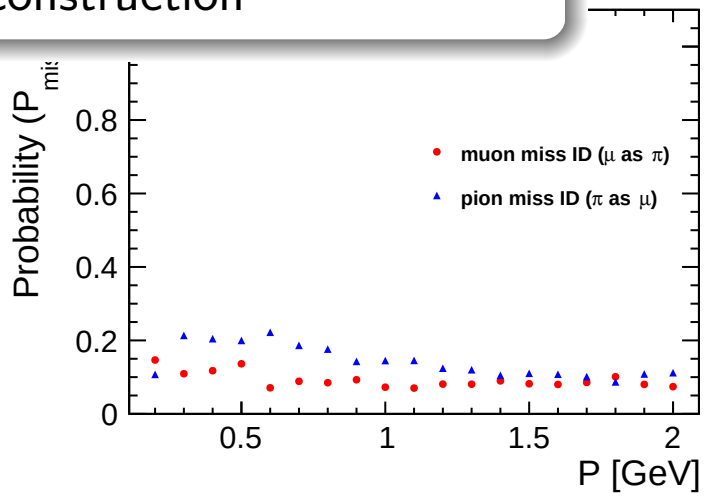


Miss Identification Probability



- ID efficiencies looks similar, $\epsilon_\mu \approx 90$ % and $\epsilon_\pi \approx 85$ %
- Gain efficiency for the reconstruction

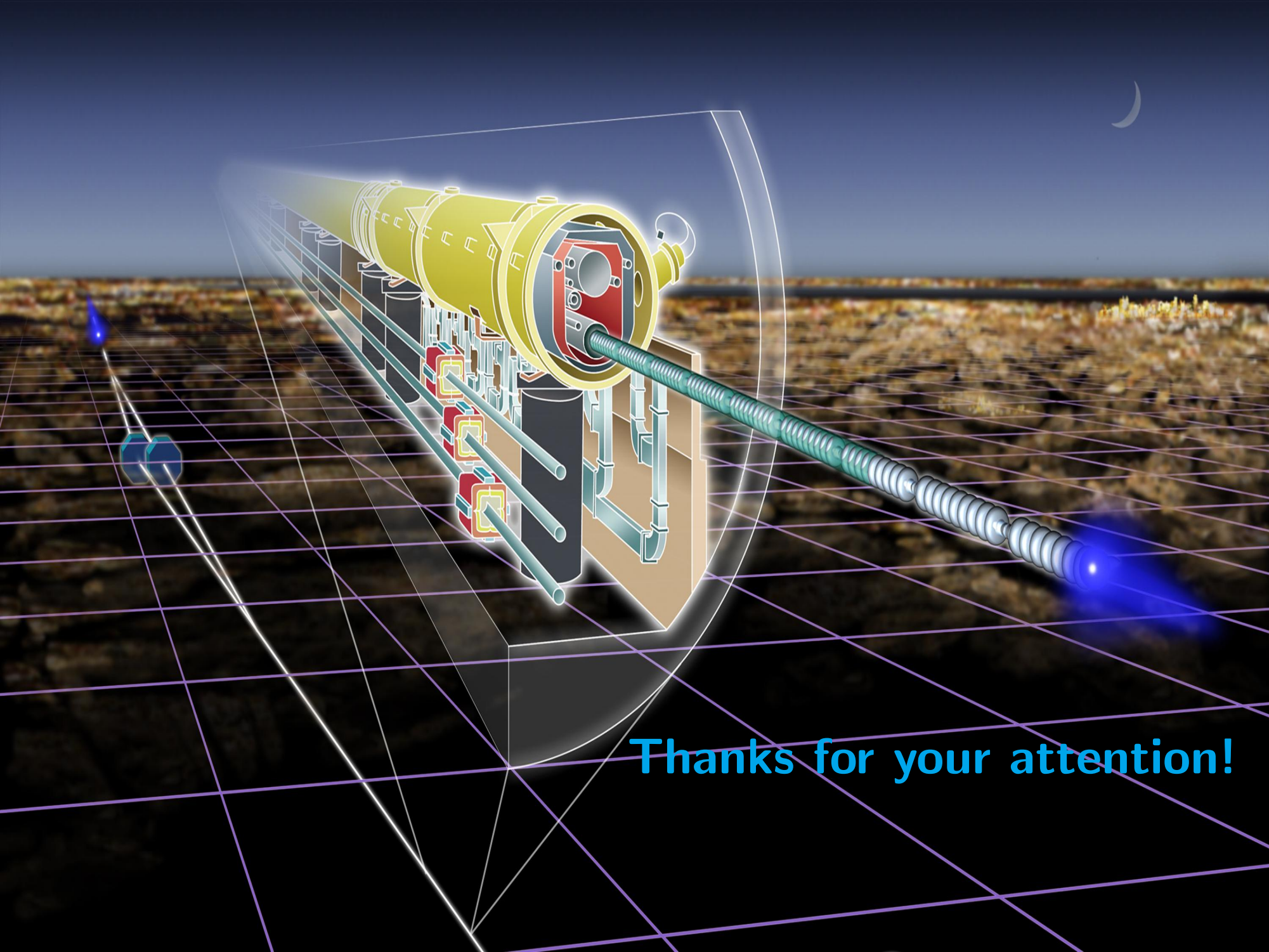




Summary

Developed a method to separate μ & π for $P < 2$ GeV

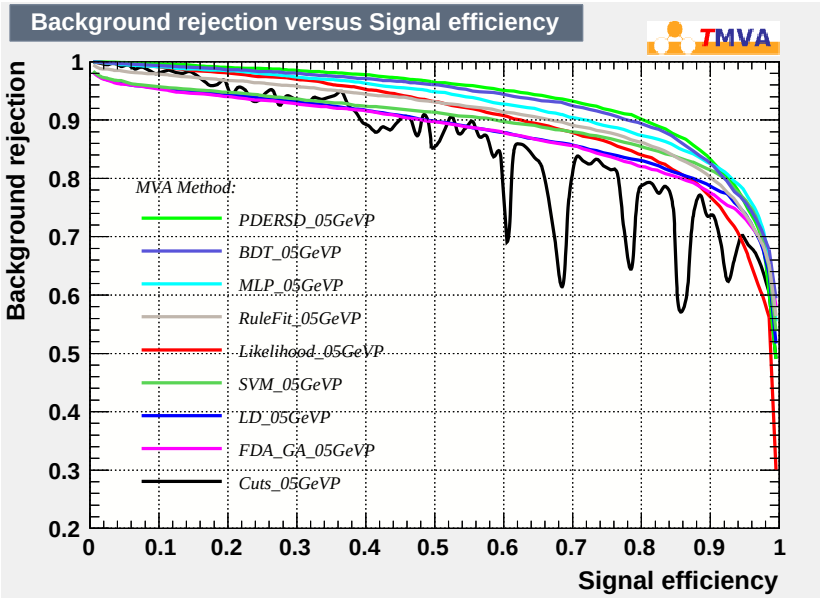
- It is based on 4 cluster shape variables
- Boosted decision tree (BDTG) method is used
- It works well with either cluster parameters or hits
- It gives $\epsilon_\mu \approx 90$ % and $\epsilon_\pi \approx 85$ % in both cases


Thanks for your attention!

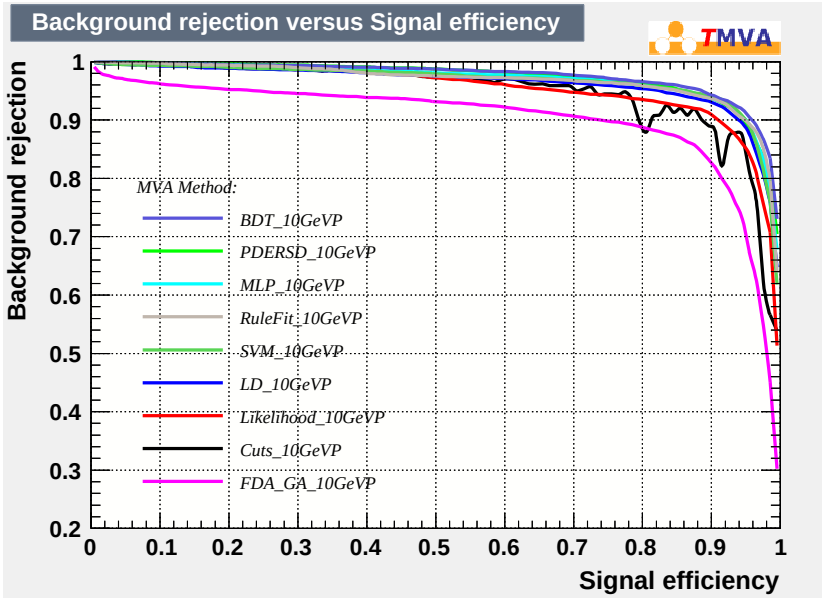
Multivariate Analysis – Testing All Methods

Checked it for other momenta..

P = 0.5 GeV



P = 1 GeV



PDERSD – Probability density estimator range search (multi-dimensional likelihood)

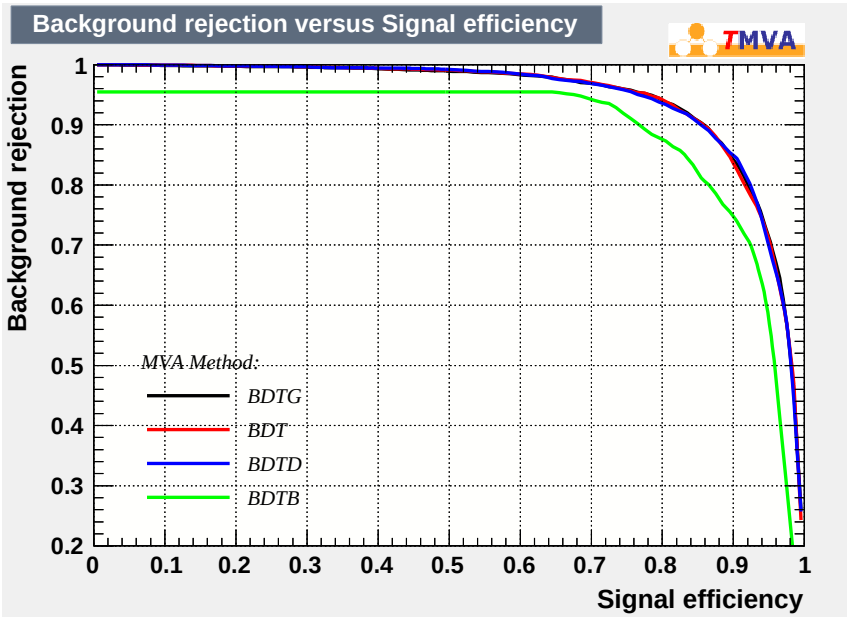
- BDT gives the best result!

Checking BDT(Boosted Decision Tree) Methods

4 available BDT methods

- Adaptive Boost (BDT)
- Gradient Boost (BDTG)
- Decorrelation of variables + Adaptive Boost (BDTD)
- Bagging (BDTB)

P = 0.2 GeV

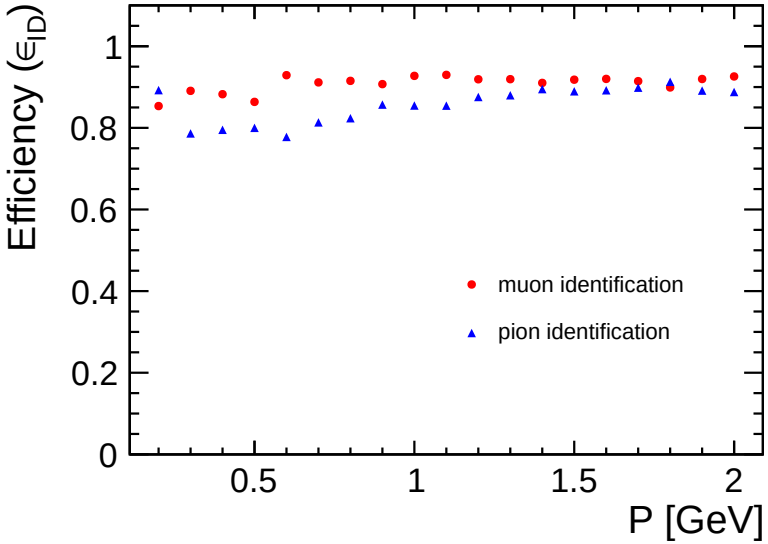


- BDTB is clearly out
- All other 3 methods have similar behaviour
- The usual BDT gives an error during training
  - Complains about $err > 0.5$
- Test the remaining methods to choose one of them

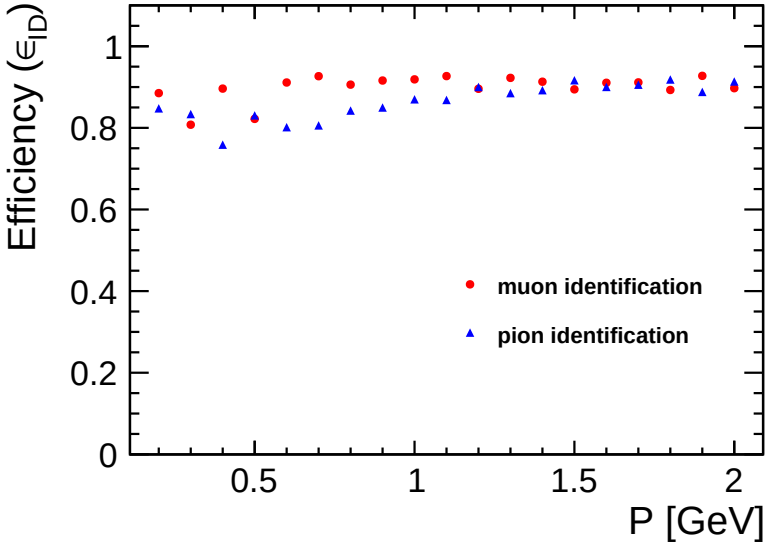
Comparison of BDT Methods

Particle Identification Efficiencies for each momentum values

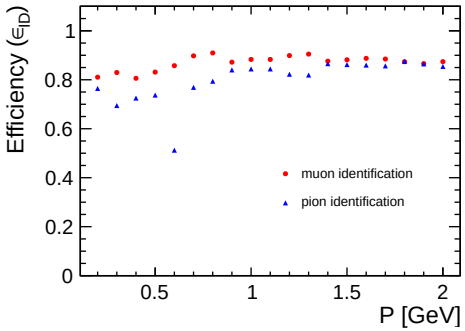
BDTG method



BDTD method



Likelihood method



- Both are better than Likelihood method
- BDTG looks less dependent on momentum
- Choose BDTG for the analysis!