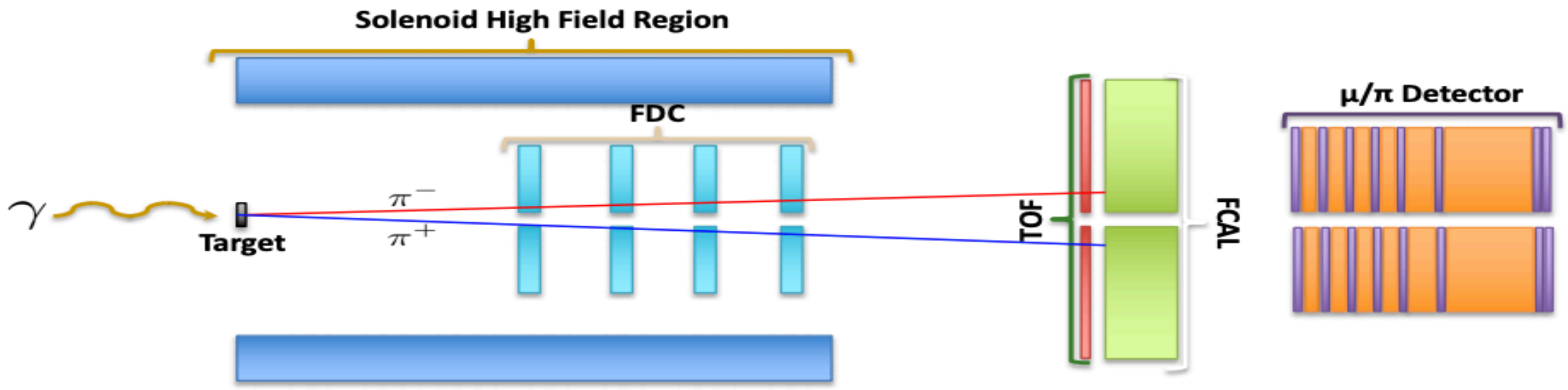
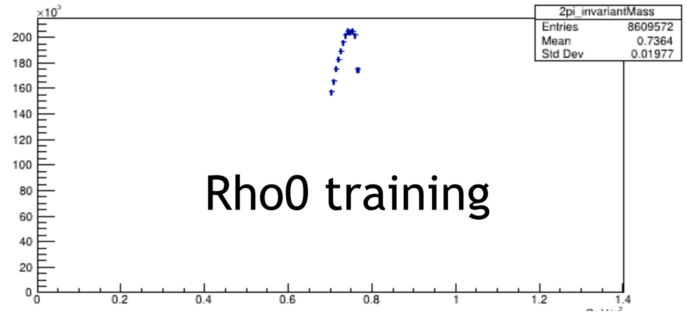
# $e/\pi/\mu$ Separation in CPP

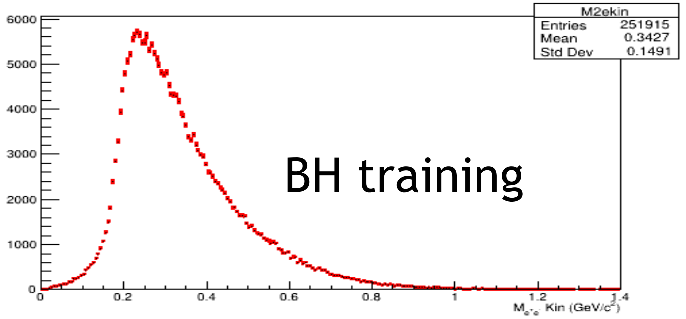
- In order to make a precision measurement of the charged pion polarizability, need to address backgrounds: Bethe-Heitler electrons and muons.
- Two independent studies for $e/\pi$ and $\mu/\pi$ identification
  - $e/\pi$ separation focuses on quantities from the forward calorimeter (FCAL).
  - $\mu/\pi$ separation combines information from FCAL and new detector system (MWPCs) into one number you can cut on. TMVA guided the mechanical design of the iron absorbers in the muon system.

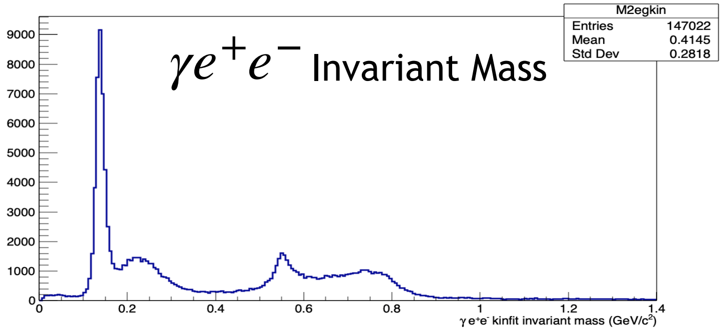


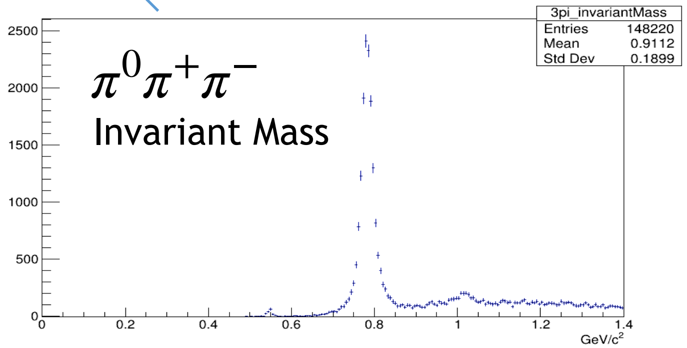
# MVA for Electron/Pion Separation Summary

- 2 Multi-layer perceptron neural nets—one for e-/π- separation, one for e+/π+.
- Train on rho0 pions (700 MeV < W < 770 MeV) and simulated Bethe-Heitler electron pairs
- Use $\pi^0 \rightarrow \gamma e^+ e^-$ and $\omega(782) \rightarrow \pi^0\pi^+\pi^-$ reactions as a way to test performance of the neural nets.



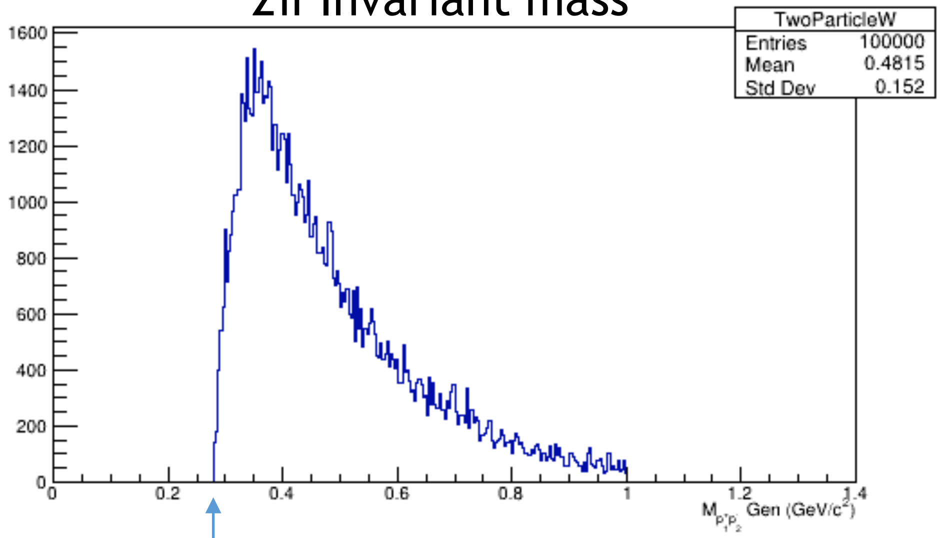
# Training Samples

We use $\rho^0$ pion data, and BH electron Monte Carlo for training our single track neural nets

# CHOOSING TRAINING SAMPLES FOR e/π NEURAL NET

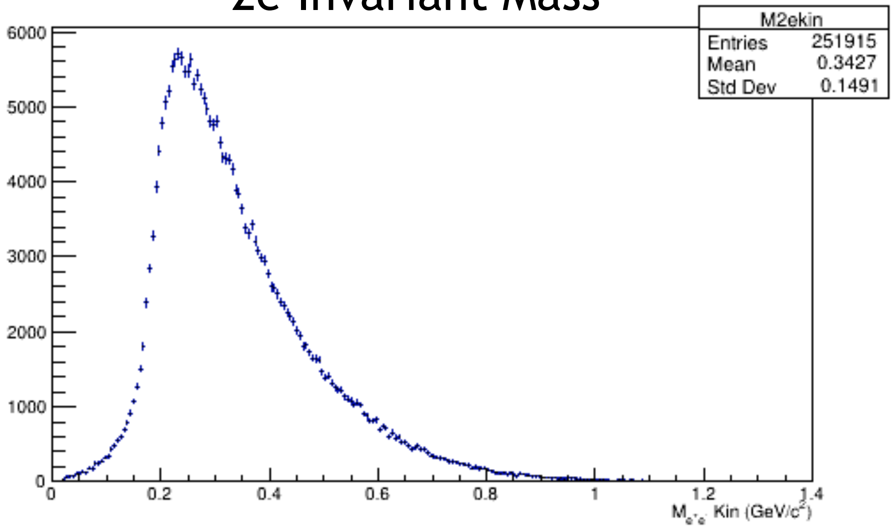
*For CPP Experiment, need to separate Primakoff Pions (signal) from BH electrons (background).*

Primakoff Pions
2π Invariant mass




2π threshold
π± mass = 139.5 MeV/c^2

Bethe-Heitler Electrons
2e Invariant Mass




Acceptance

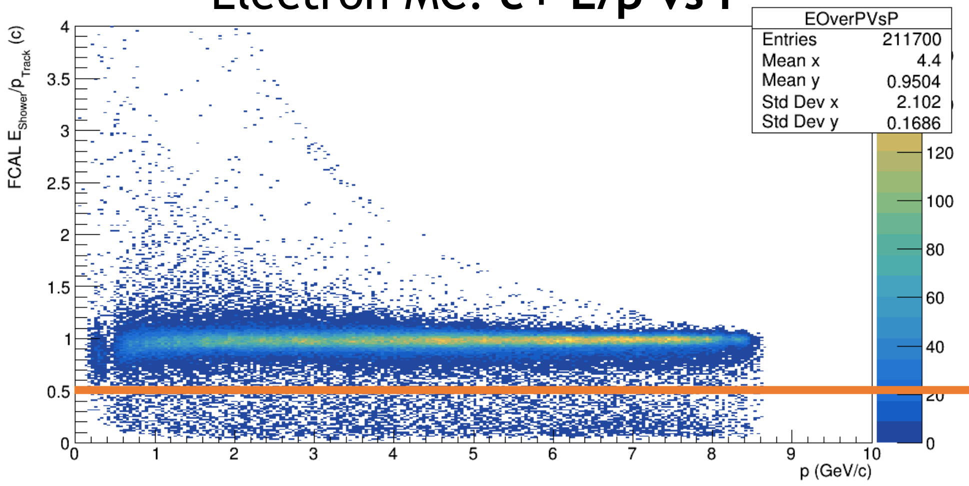
# Choosing Training Samples for e/$\pi$ Neural Net

- Data or Simulation?

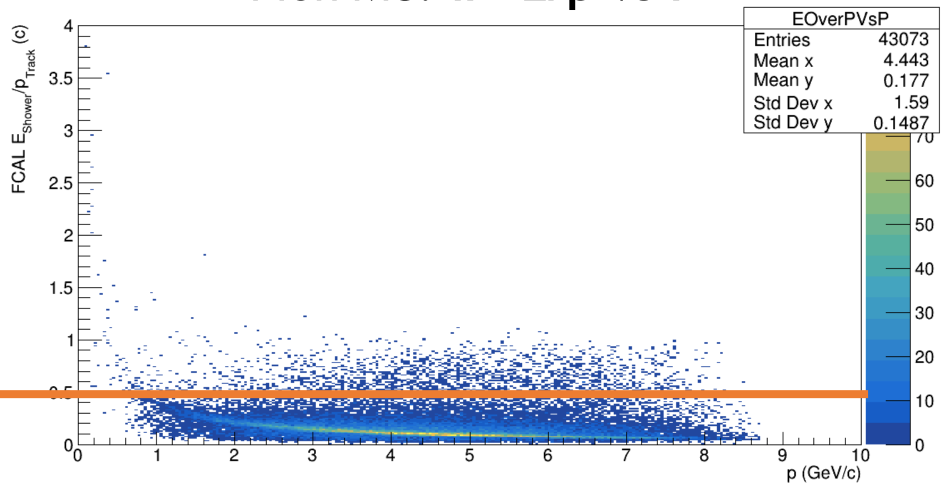
# Training Samples for e/π Neural Net

- Pion training: Data or *Simulation*?

Electron MC: **e+ E/p vs P**



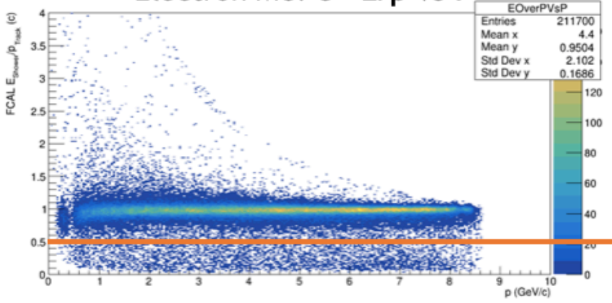

Pion MC: **π+ E/p vs P**



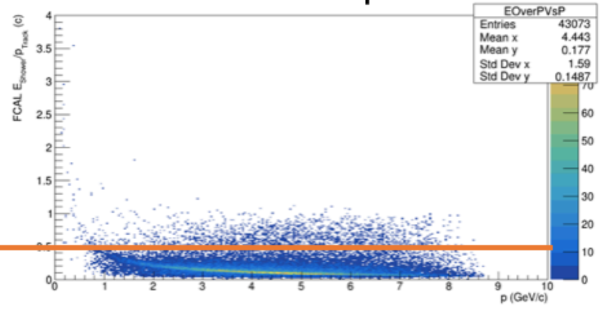
# TRAINING SAMPLES FOR e/π NEURAL NET

- Pion training: Data or *Simulation*?

Electron MC: **e+ E/p vs P**

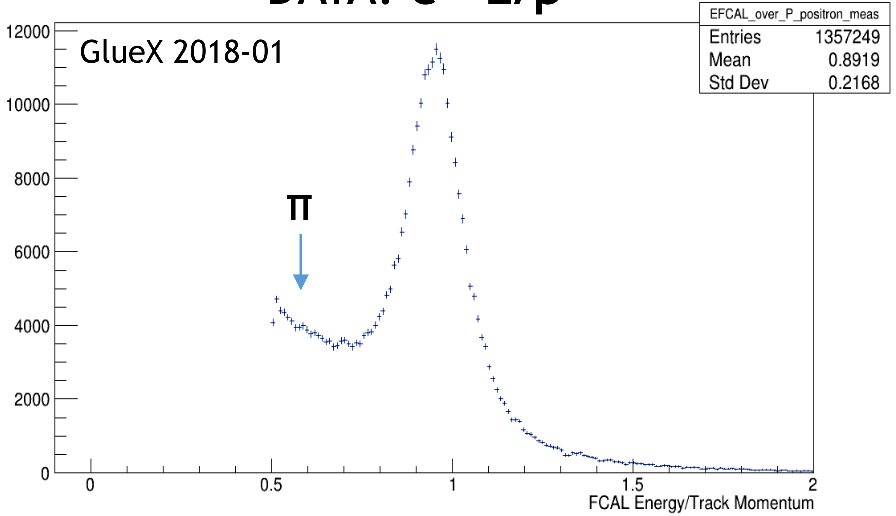



Pion MC: **π+ E/p vs P**

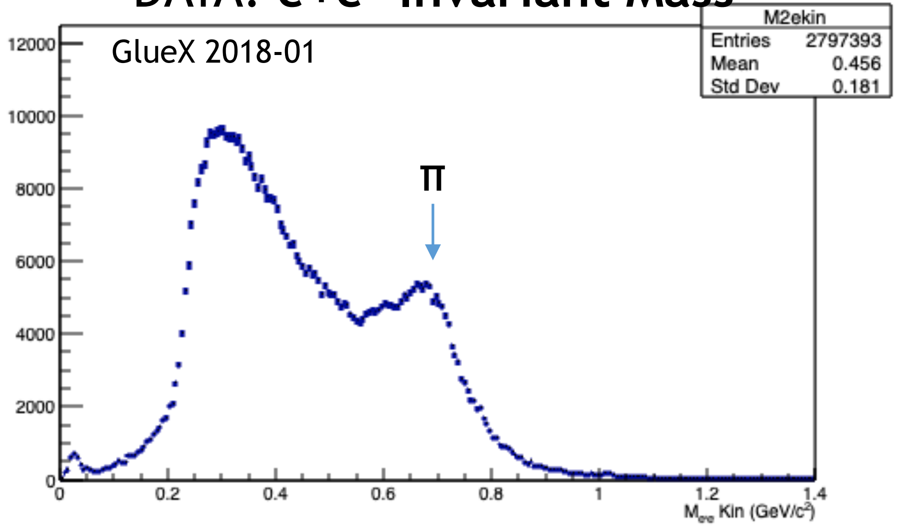



Require E/p > 0.5 for both e+e- tracks and look at data:

DATA: **e+ E/p**

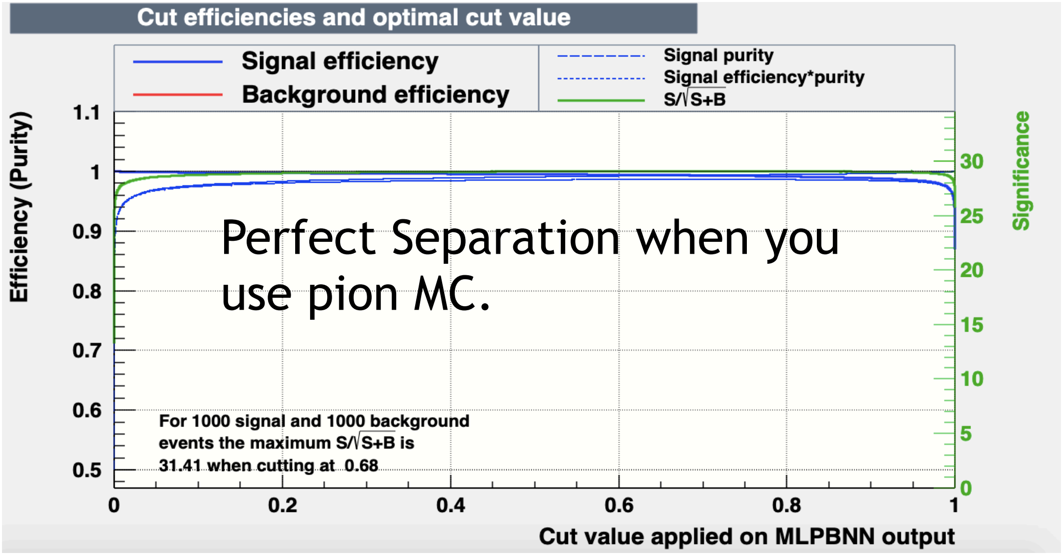


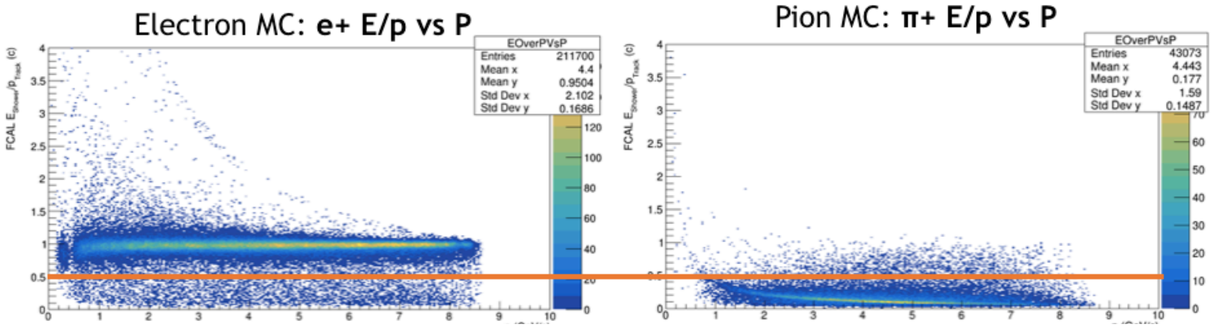

DATA: **e+e- Invariant Mass**



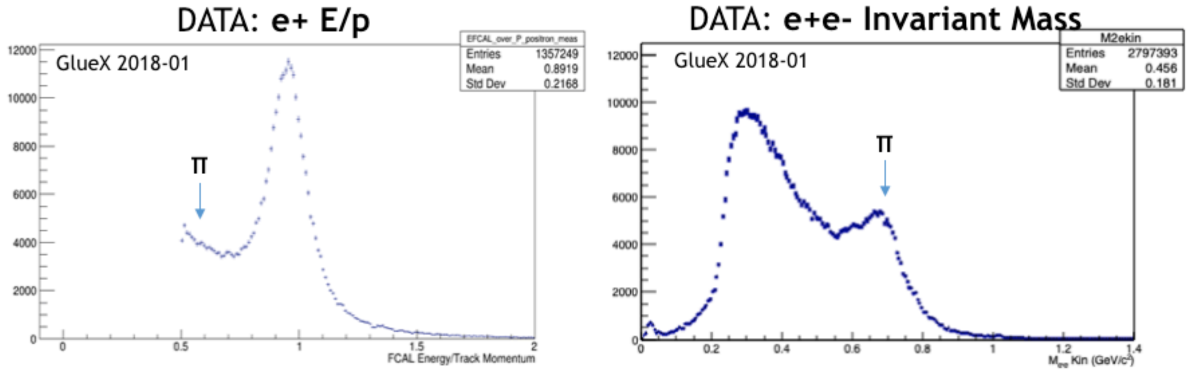
# Training Samples for e/π Neural Net

- Pion training: Data or *Simulation*?



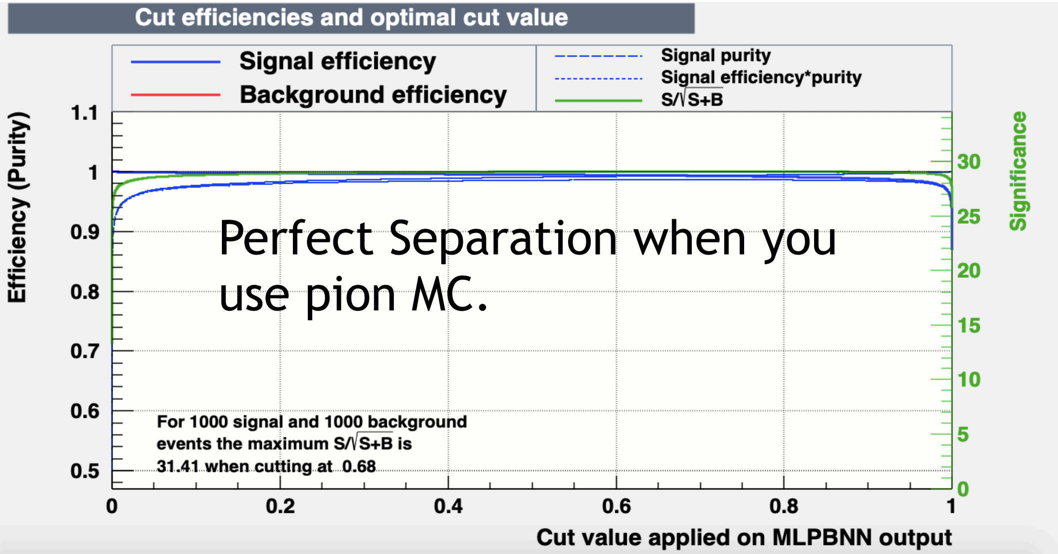



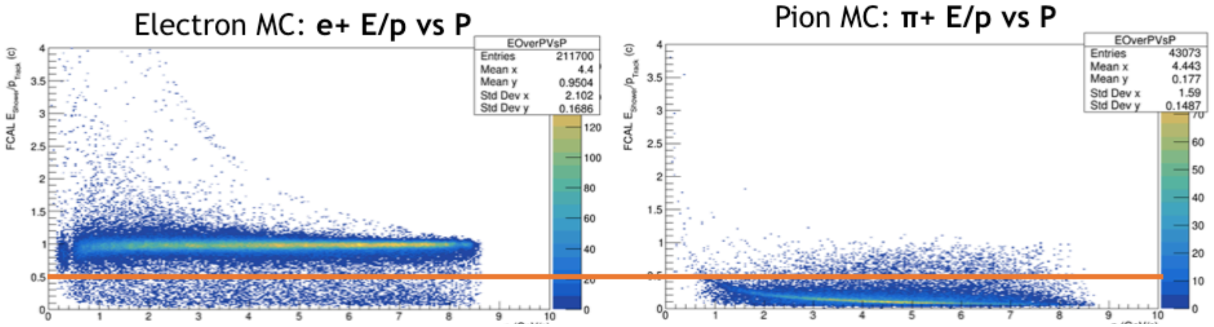

Require E/p > 0.5 for both e+e- tracks and look at data:



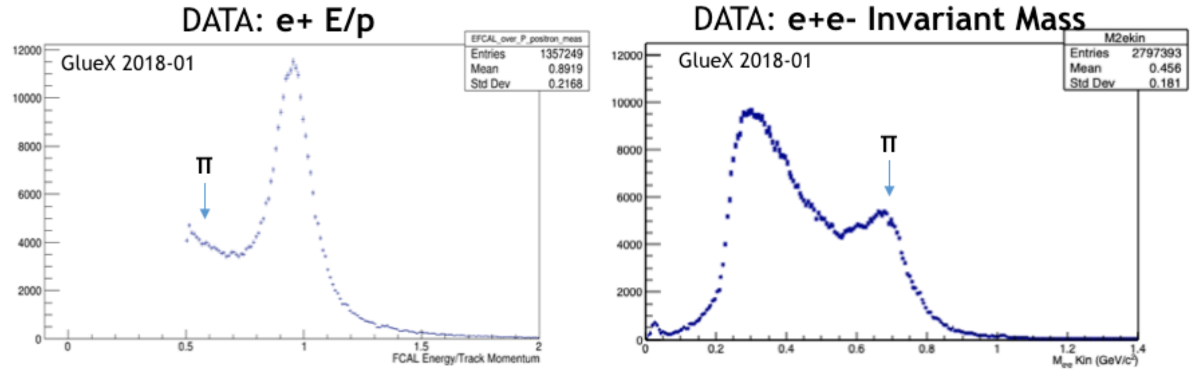
# TRAINING SAMPLES FOR e/π NEURAL NET

- Pion training: Data or *Simulation*?



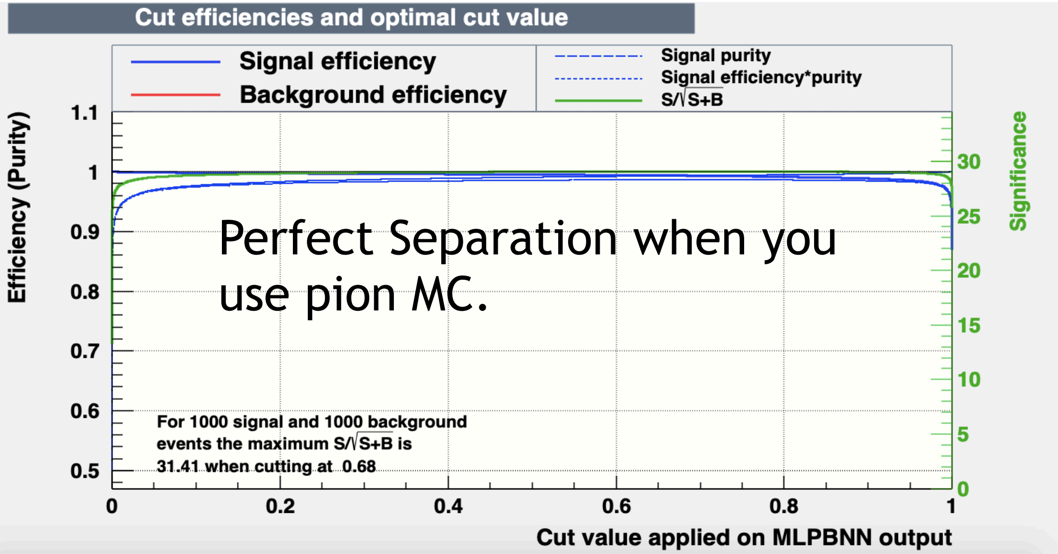

Require E/p > 0.5 for both e+e- tracks and look at data:

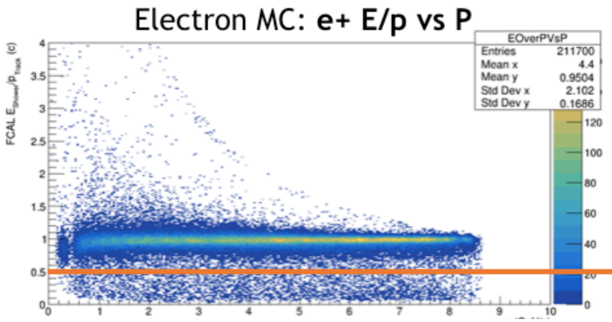


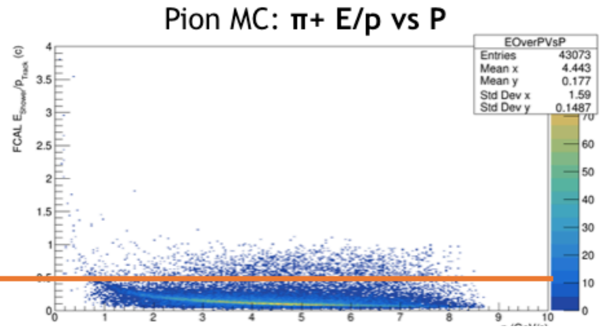

**Questions that must be answered:**
Is modeling of hadronic showers in calorimeter good enough?
Is sample size of pions too small to be representative?

# TRAINING SAMPLES FOR e/π NEURAL NET

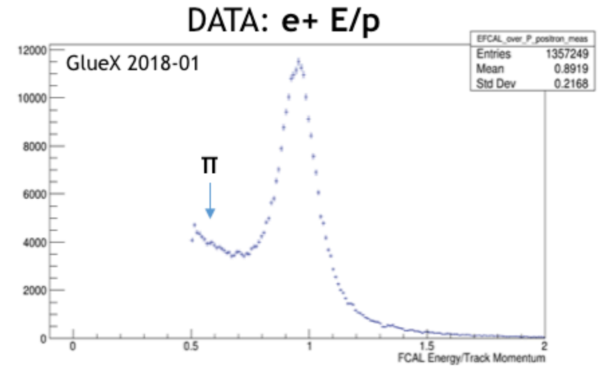
- Pion training: Data or *Simulation*?

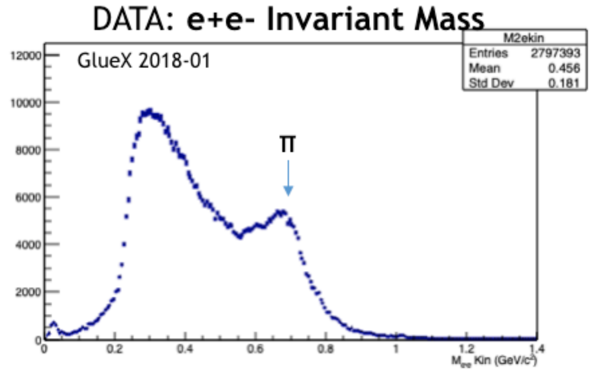



Require E/p > 0.5 for both e+e- tracks and look at data:



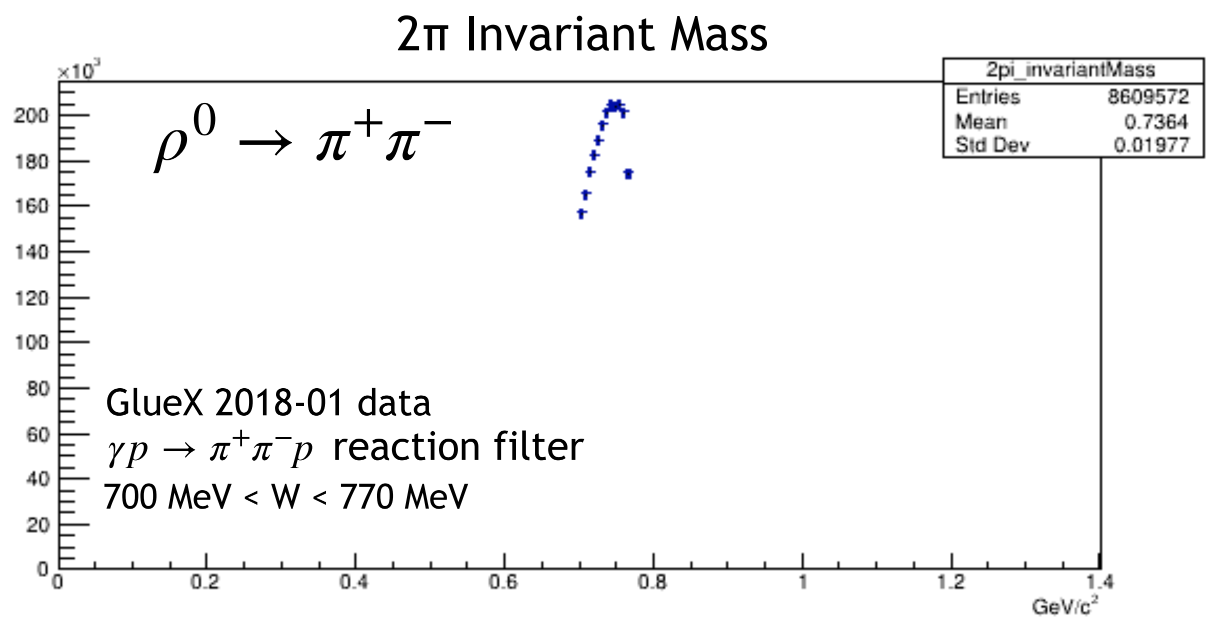

**Questions that must be answered:**
Is modeling of hadronic showers in calorimeter good enough?
Is sample size of pions too small to be representative?

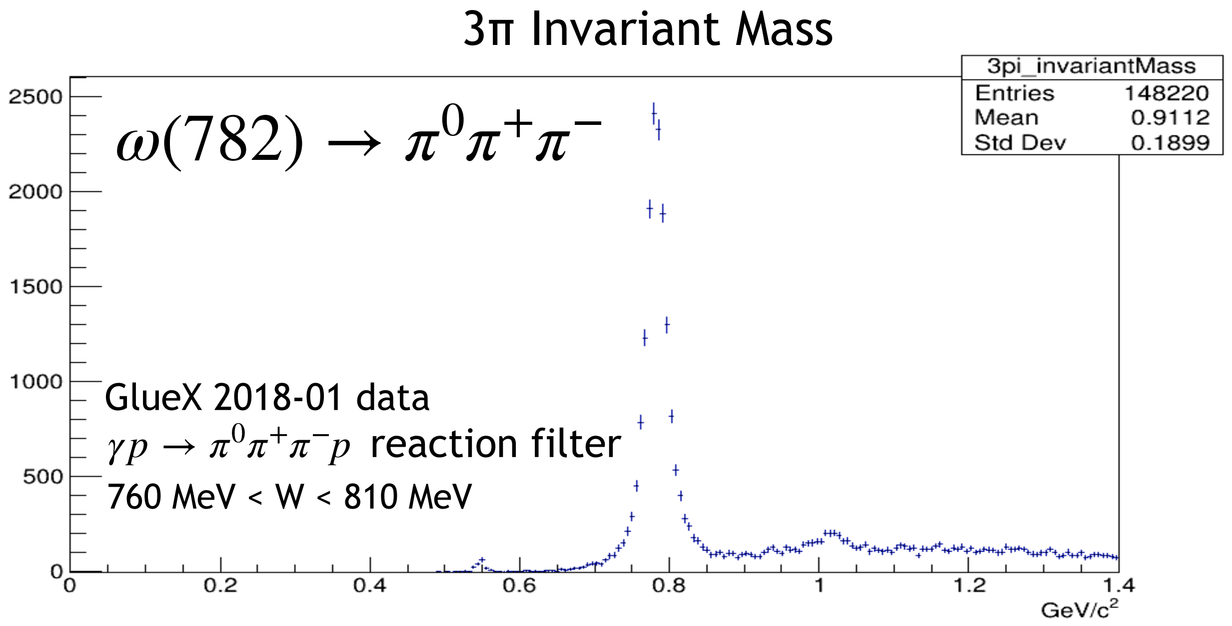
Can we just get a pure sample of $\pi^+\pi^-$ data for training?

# TRAINING SAMPLES FOR e/π NEURAL NET

- Pion training: *Data* or ~~Simulation~~?

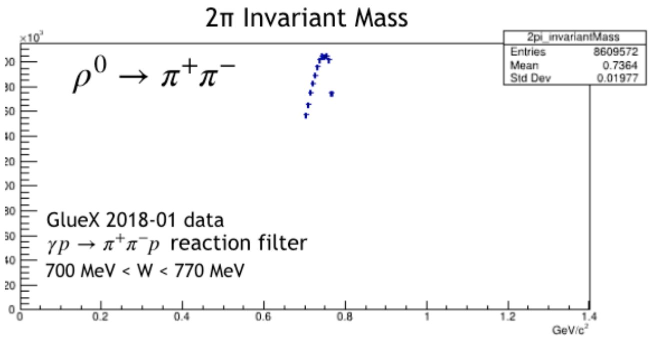
Can we just get a pure sample of $\pi^+\pi^-$ data for training? Yes.

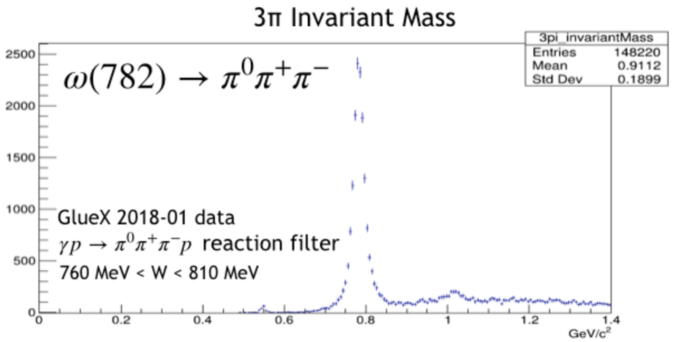


# TRAINING SAMPLES FOR e/π NEURAL NET

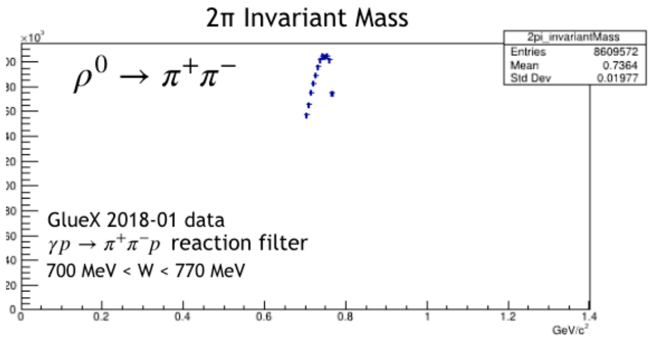
- Pion training: *Data* or ~~Simulation~~?
  - Which source to use?
    —> should be as close to Primakoff kinematics as possible.

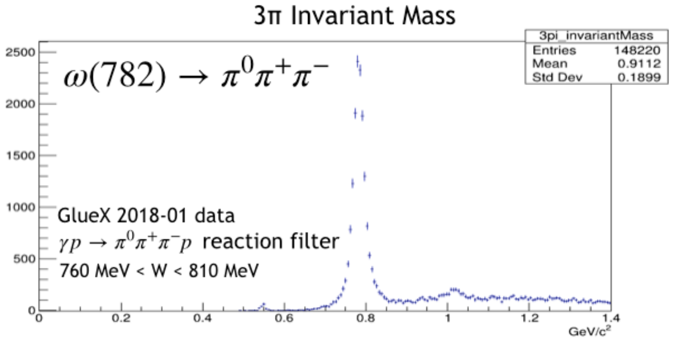


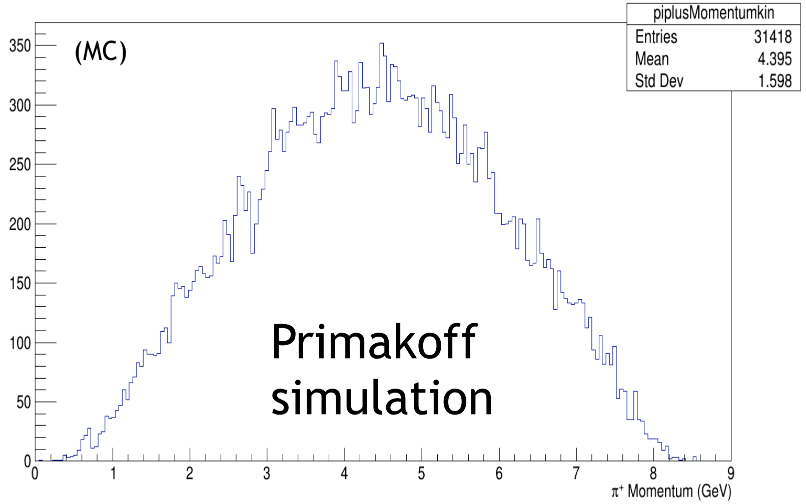
# Training Samples for e/π Neural Net

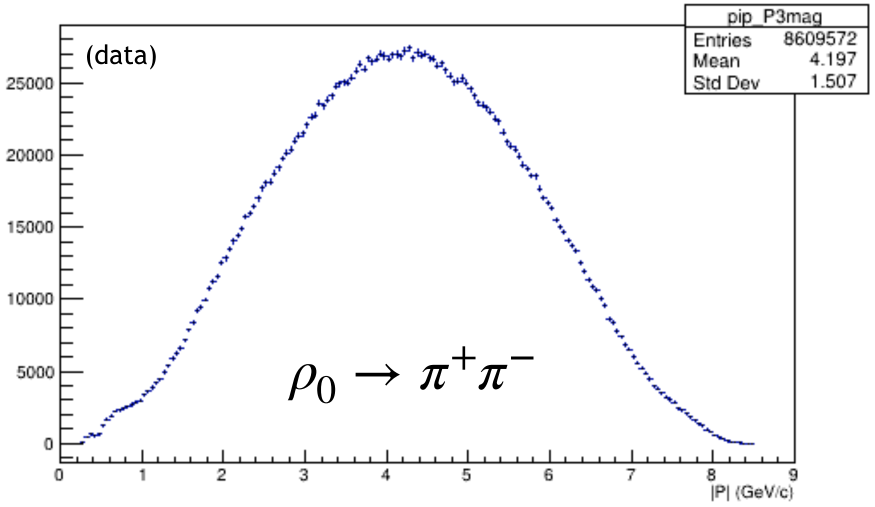
- Pion training: *Data* or ~~Simulation~~?
  - Which source to use?
    —> should be as close to Primakoff kinematics as possible.

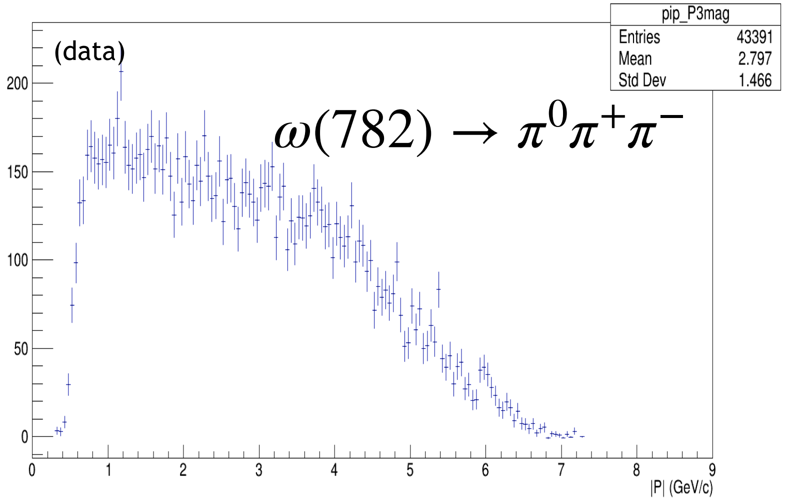



## π⁺ MOMENTUM

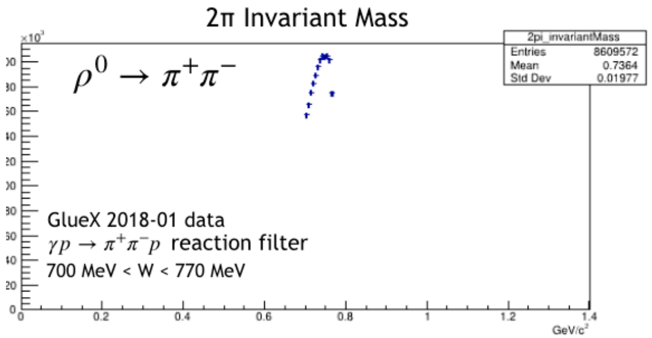


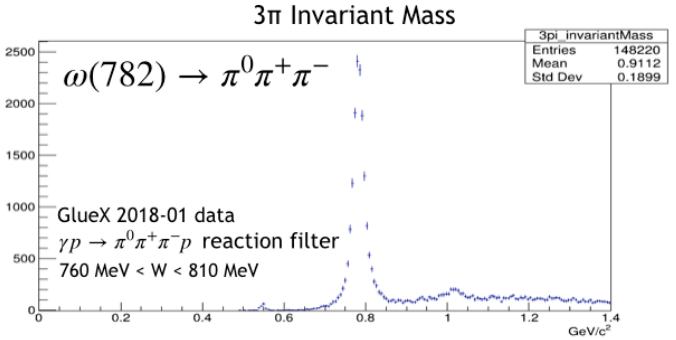



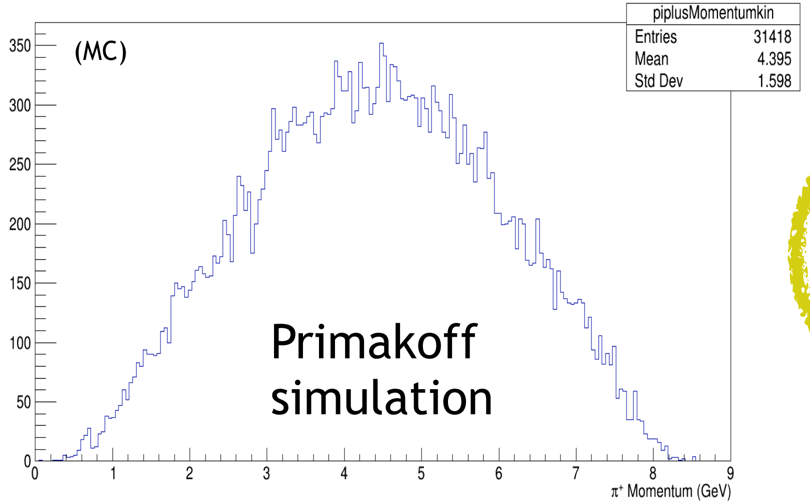
# Training Samples for e/π Neural Net

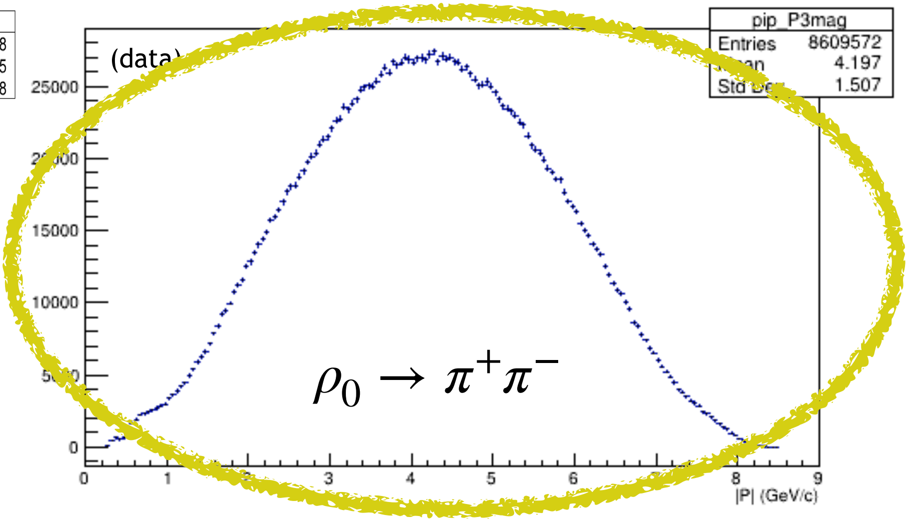
- Pion training: *Data* or ~~Simulation~~?
  - Which source to use?
    —> should be as close to Primakoff kinematics as possible.

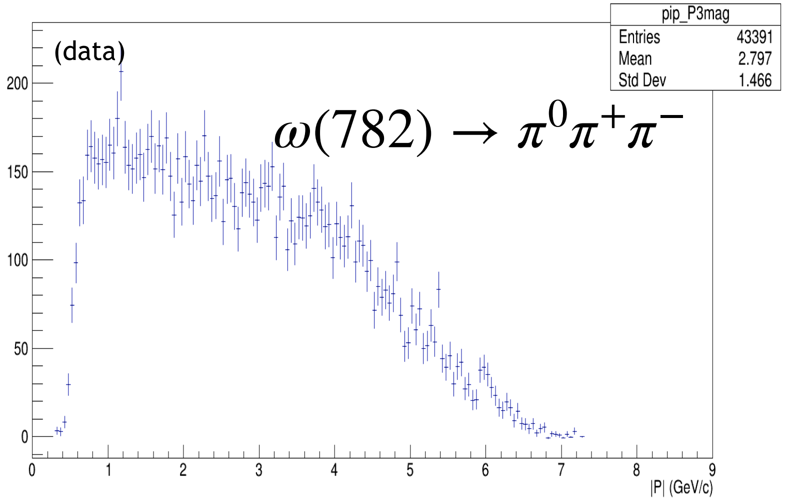



## π⁺ MOMENTUM



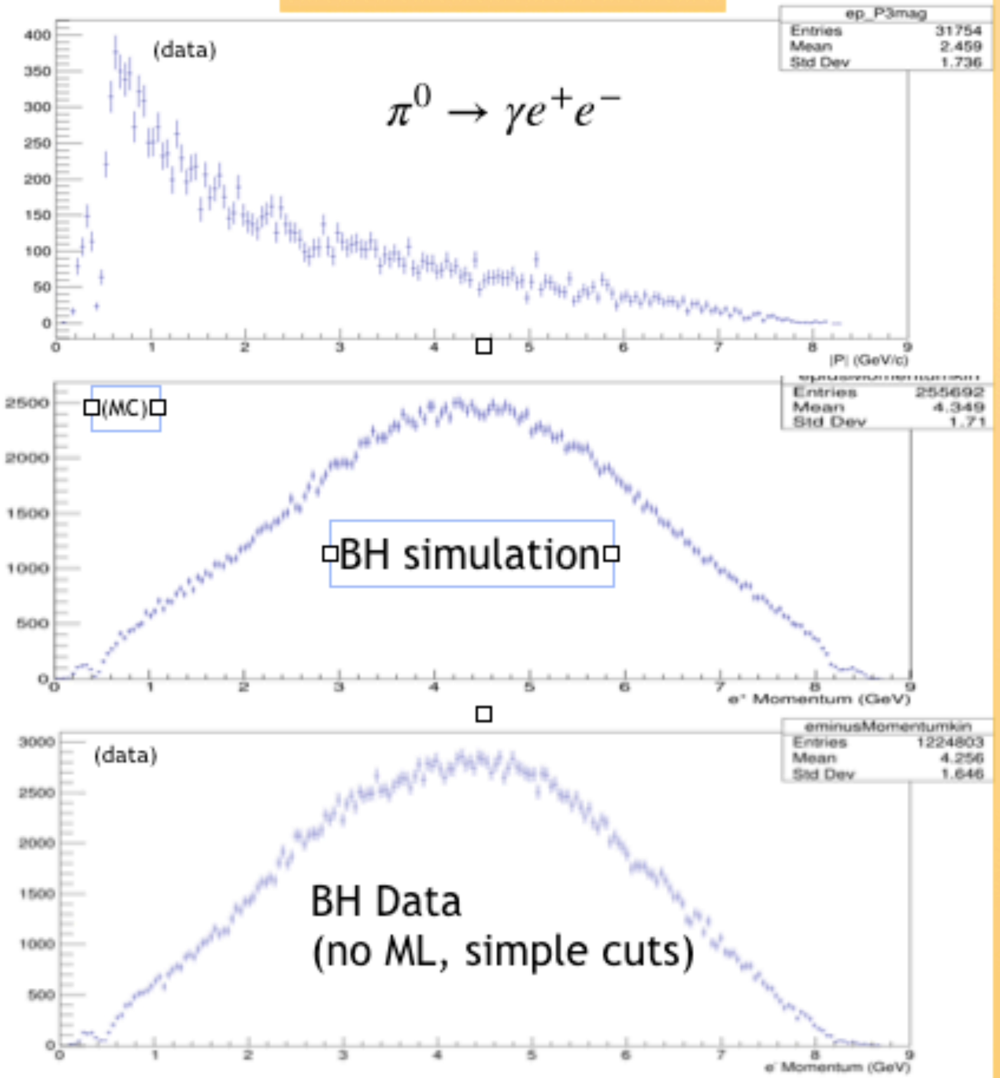



# Training Samples for e/π Neural Net

- Pion Training: **$\rho^0$ Data**
- Electron Training: Data or Simulation?
  - Bethe-Heitler electrons are the background in CPP, so ideally we should train on them. Are either BH in GlueX data or MC viable for training? Data from other source?
  - Pure BH data samples are hard to obtain for the same reason we can easily train on pion data—contamination from $\rho^0$ pions dominates.
  - QED is very well understood—no reason to suspect that the modeling of calorimeter showers is inadequate.

# POSITRON MOMENTUM



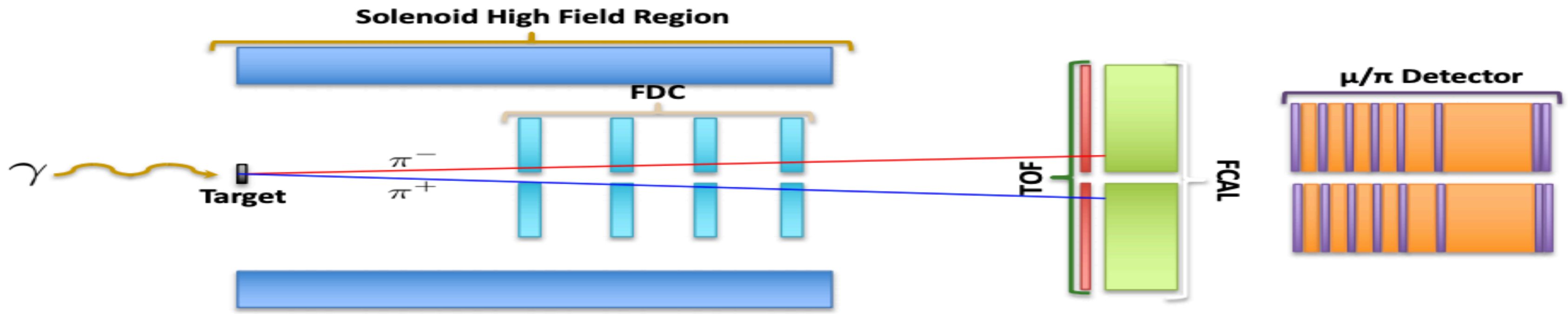
# Training Variables

We train on 3 variables with quantities constructed from the particle momentum and the forward calorimeter (FCAL):

E/p, E9/E25 Shower Ratio, FCAL DOCA

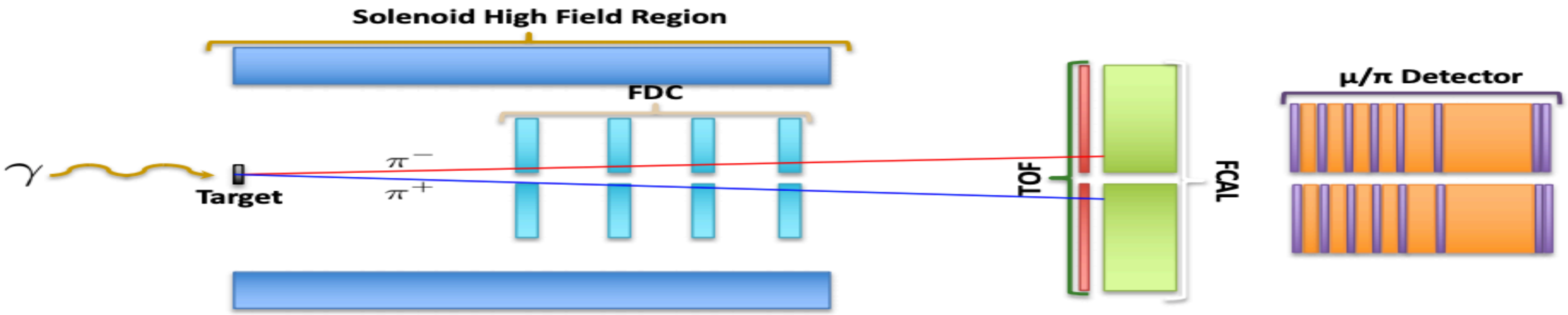
# Training Variable Selection

- The CPP measurement will be made with low t, forward going Primakoff pions. Therefore the e/π neural net can only use quantities from forward detectors—FDC, TOF, FCAL.



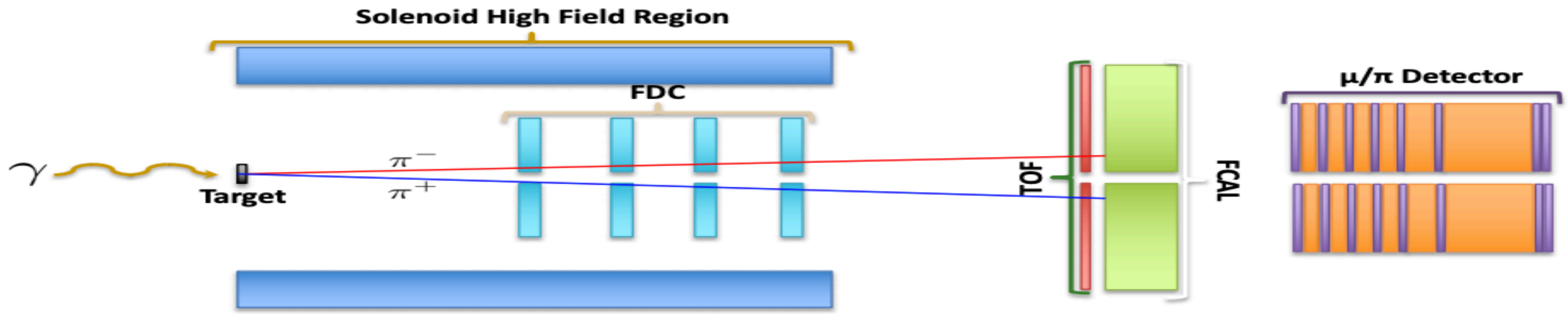
# Training Variable Selection

- The CPP measurement will be made with low t, forward going Primakoff pions. Therefore the e/π neural net can only use quantities from forward detectors—FDC, TOF, FCAL.
- Of the forward detectors, FCAL is (probably) the only useful one for e/π separation.



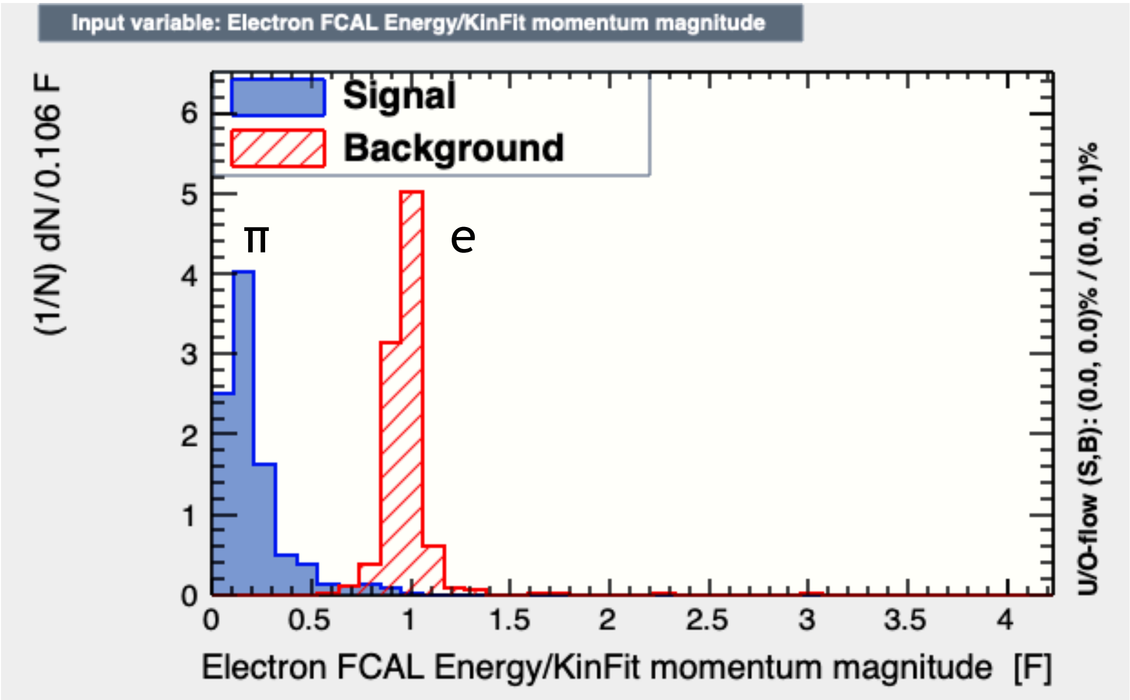
# TRAINING VARIABLE SELECTION

- The CPP measurement will be made with low t, forward going Primakoff pions. Therefore the e/π neural net can only use quantities from forward detectors—FDC, TOF, FCAL.
- Of the forward detectors, FCAL is (probably) the only useful one for e/π separation.
- If the NN finds useful information elsewhere, the concern would be that it's relying on kinematics of $\rho^0$ versus BH, and would struggle to identify pions from other sources, e.g. Primakoff pions.



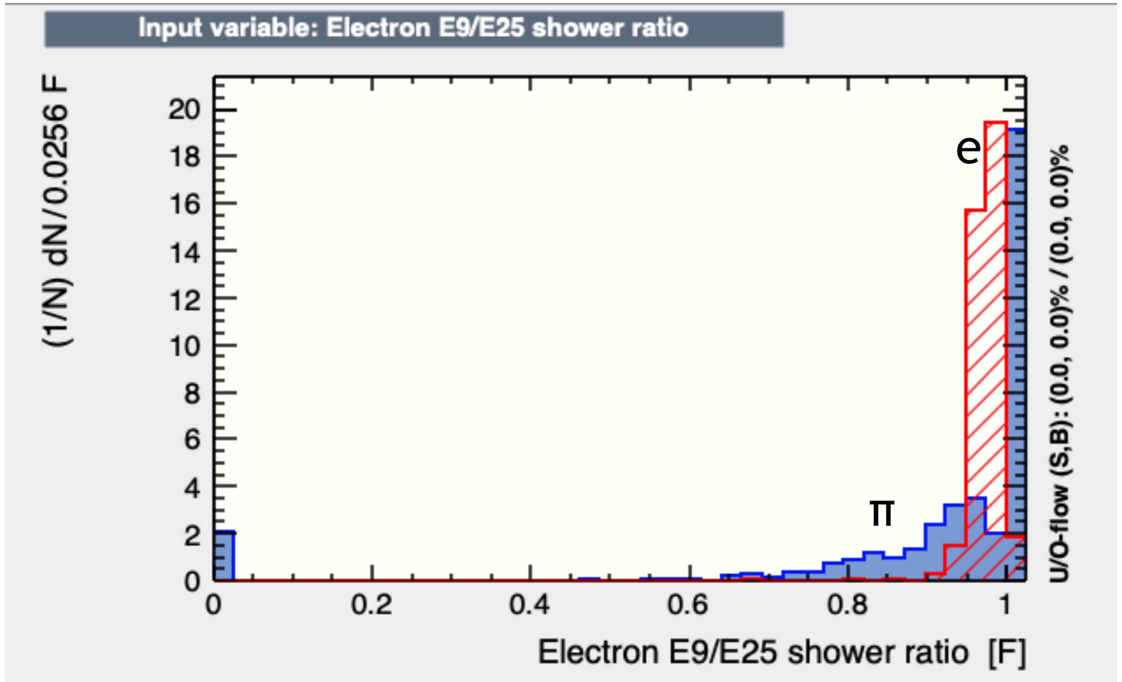
# VARIABLE 1: E OVER P

E is the shower energy deposited in the forward calorimeter that is associated with the charged track.
P is the magnitude of the 3-momentum of that charged track.



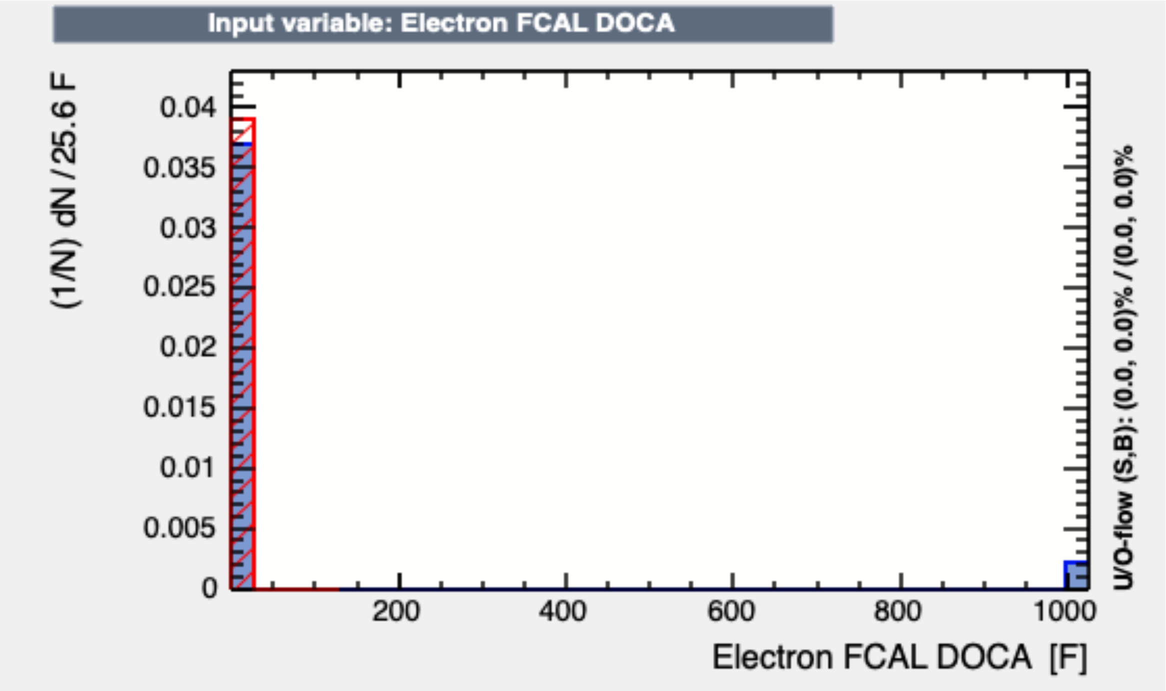
# Variable 2: E9/E25 Shower Ratio

E9 and E25 are the sum of the energies in a 3x3 and 5x5 square around the crystal in which the center of the shower lies.



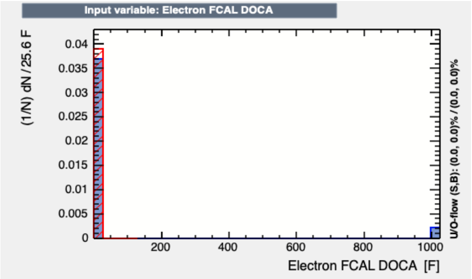
# VARIABLE 3: FCAL DOCA

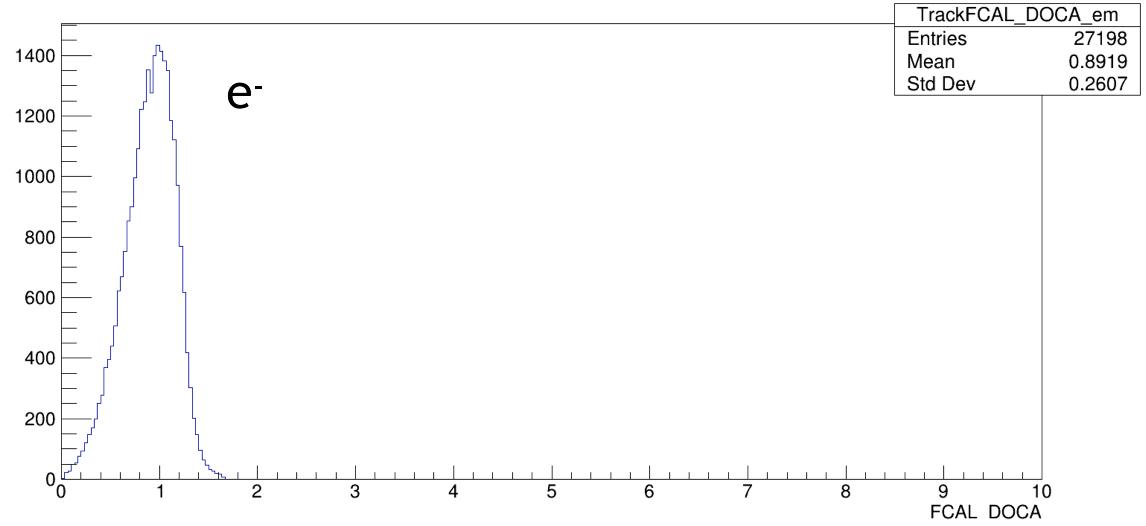
DOCA is the distance of closest approach. It is the distance in centimeters from the projection of the track through FCAL to the centroid of the shower. Tracks with no shower are given DOCA of 1000.

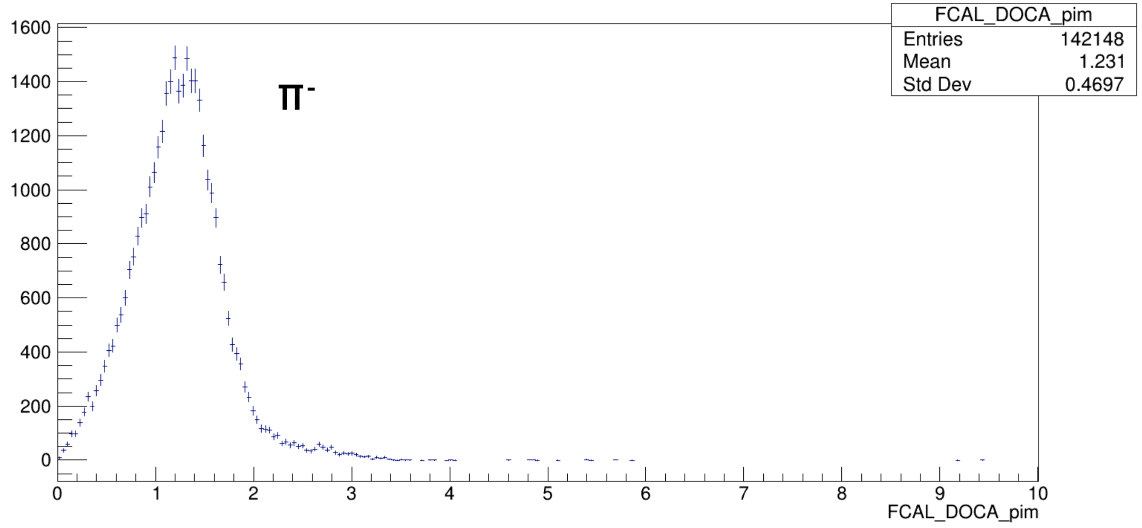


# VARIABLE 3: FCAL DOCA

DOCA is the distance of closest approach. It is the distance in centimeters from the projection of the track through fcal to the centroid of the shower.



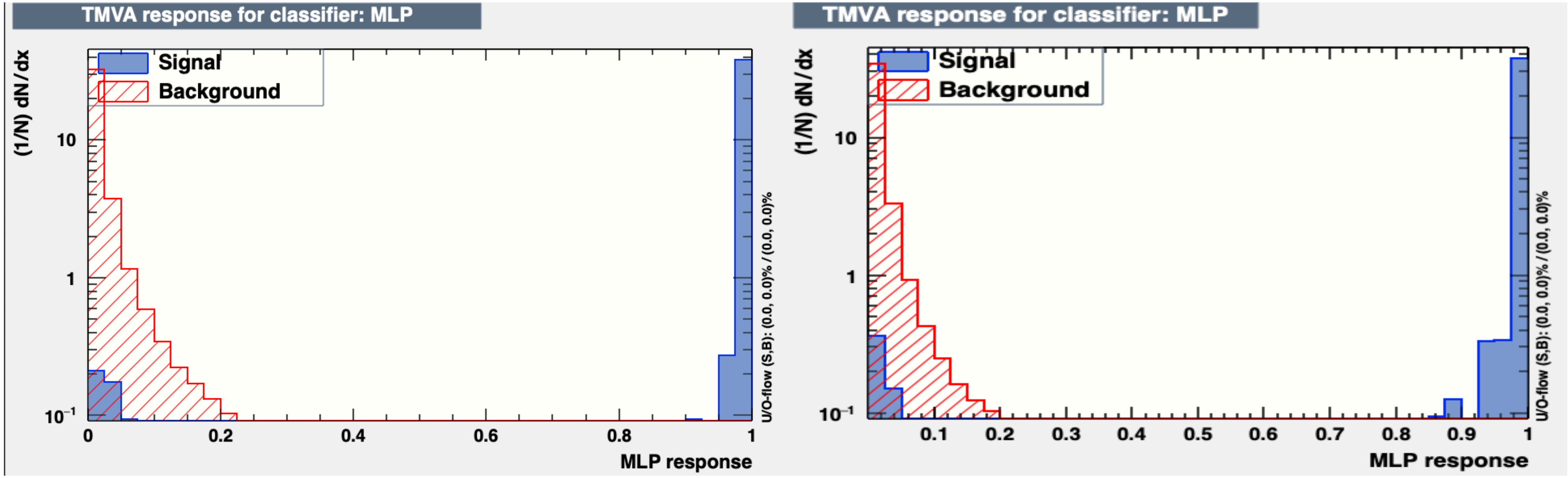



# Two Single-Track Neural Nets or One Two-Track Neural Net?

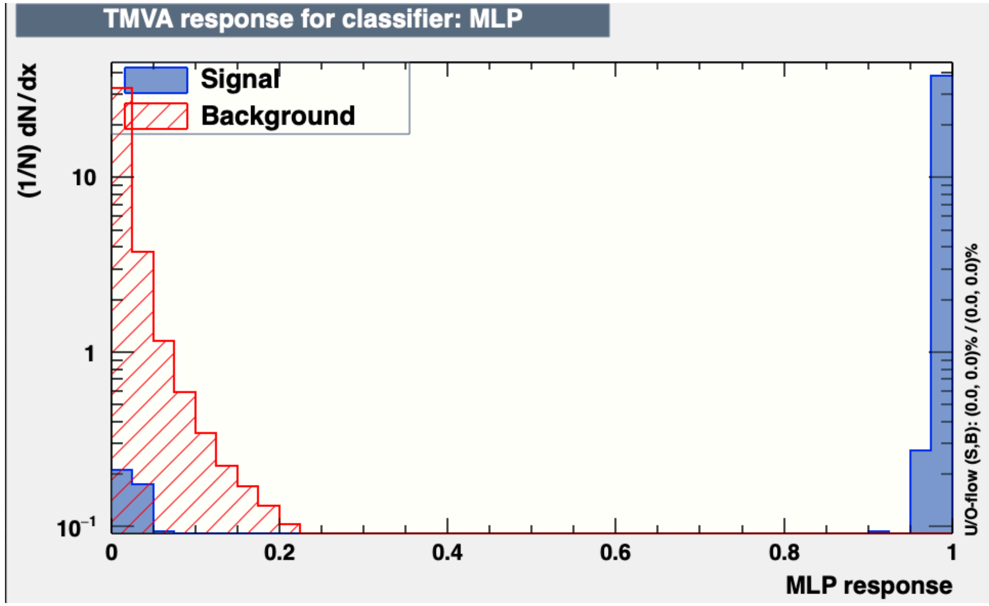
# Training Methods

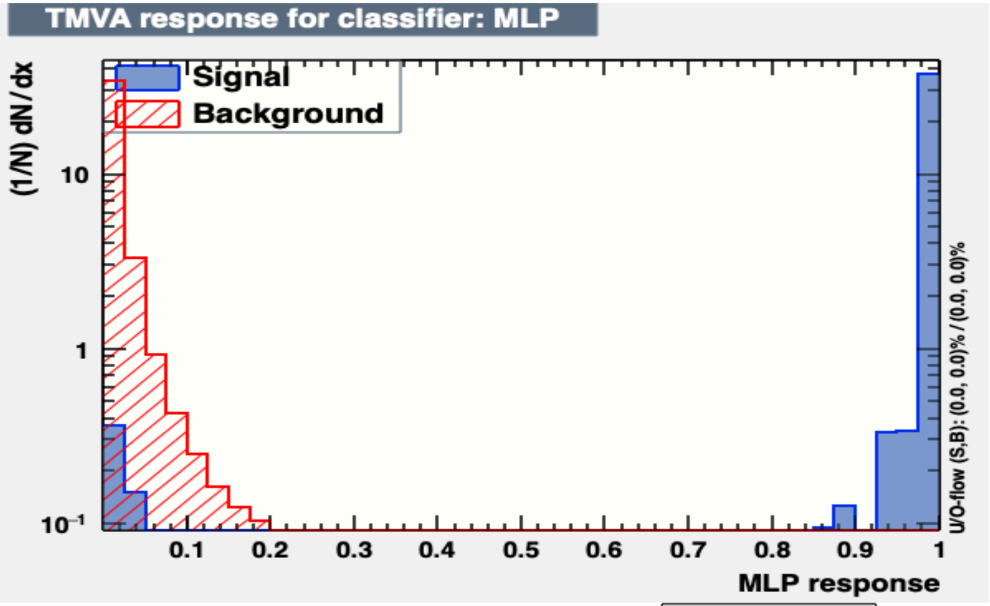
MLP, BDT, MLPBNN, LD

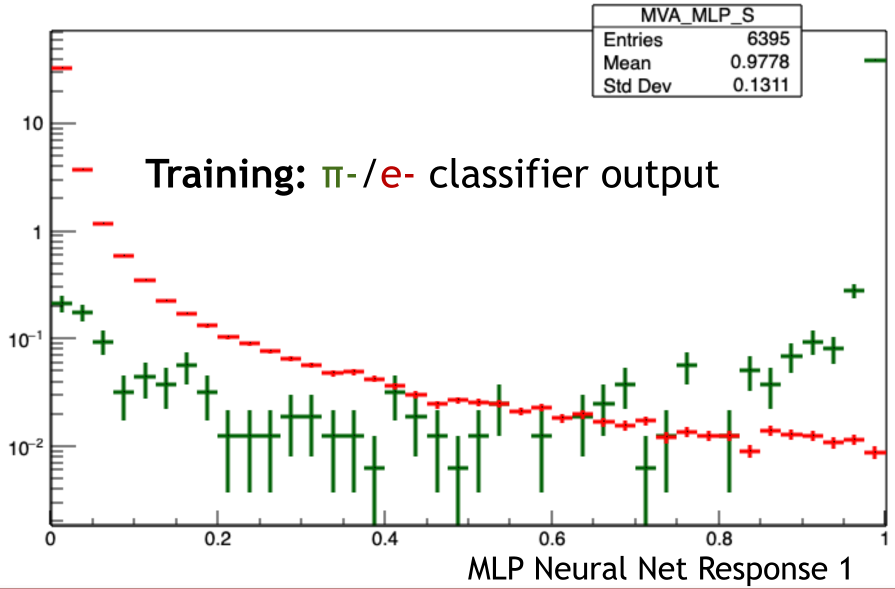
# Classifier Outputs for Training Samples

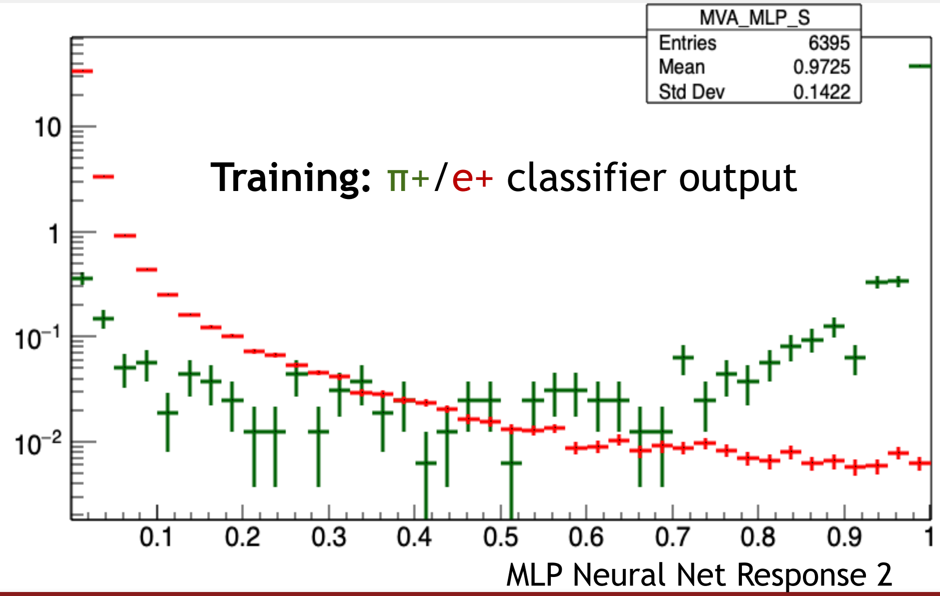


# Classifier Outputs for Training Samples

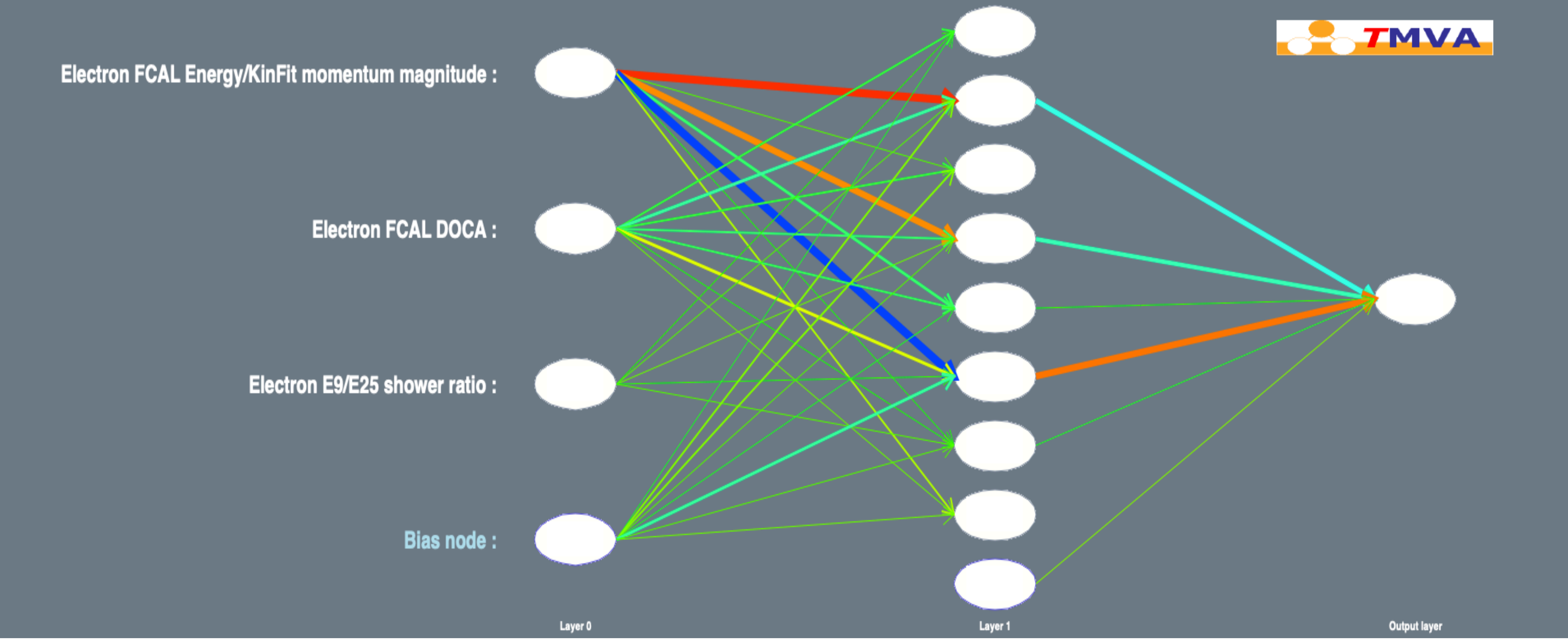


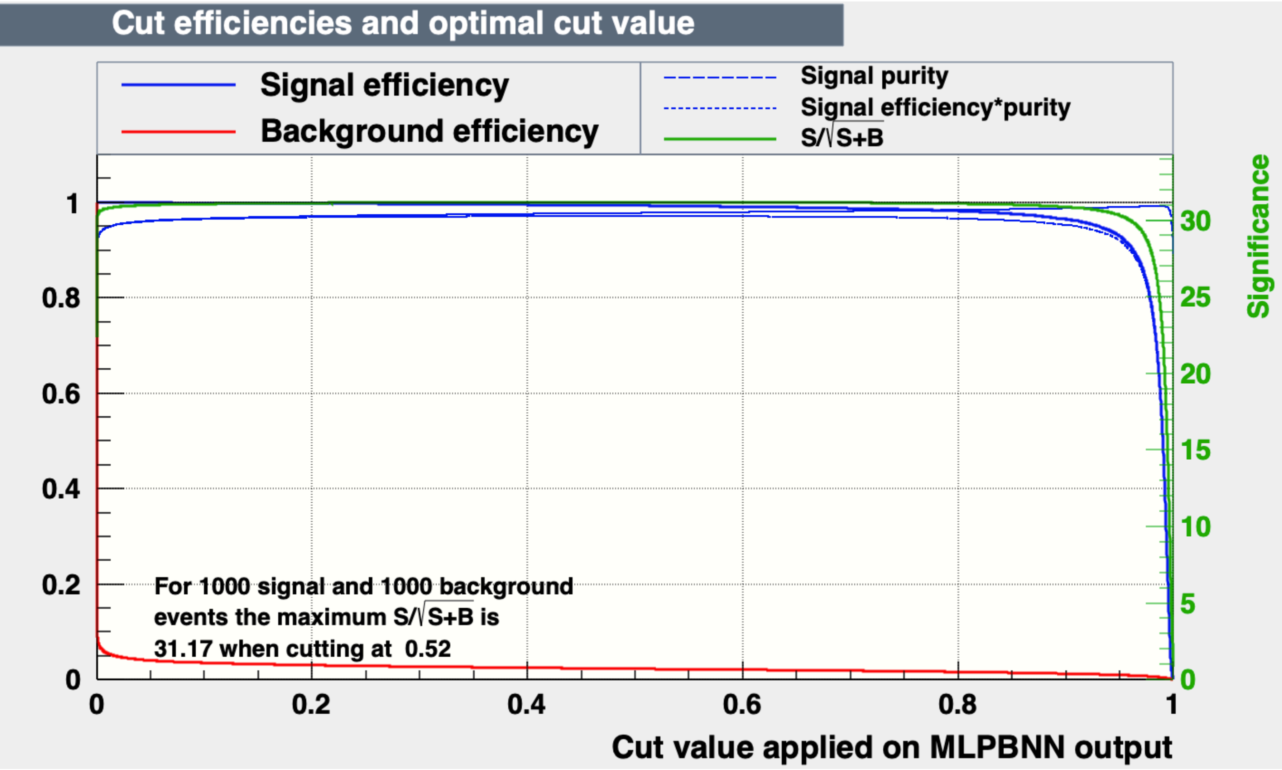



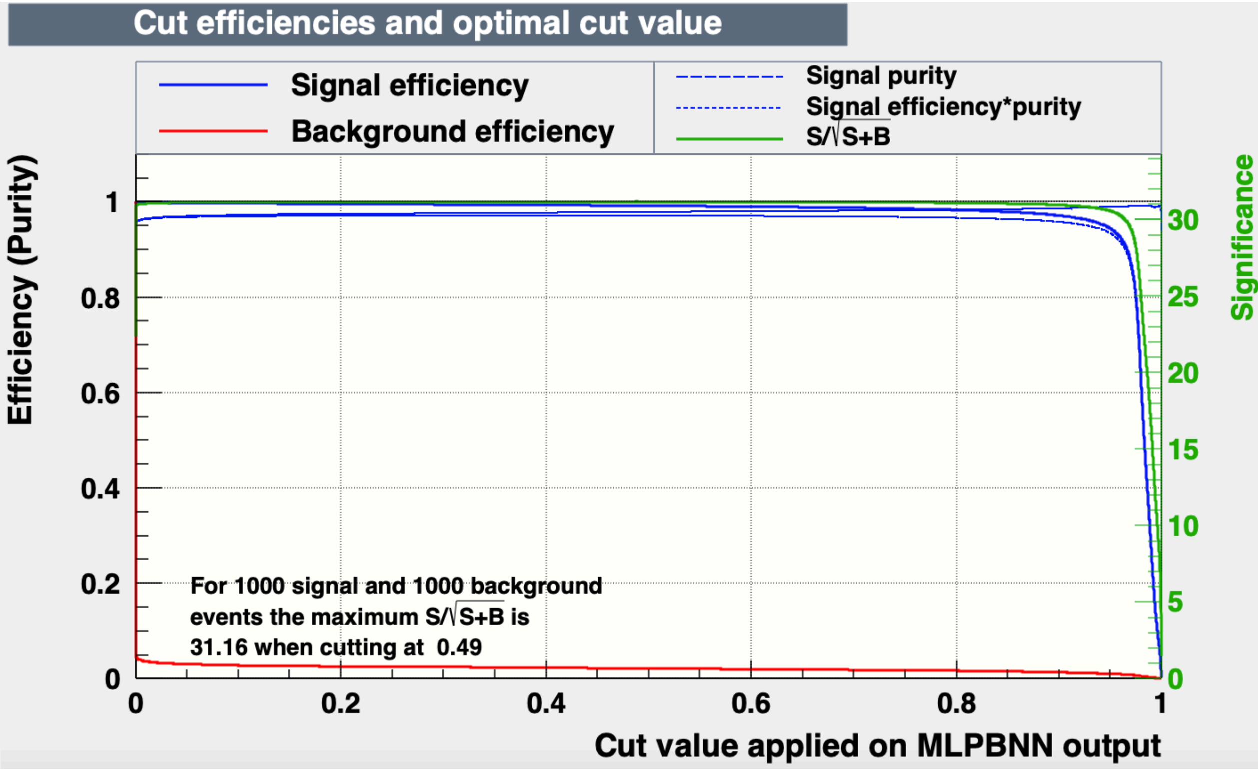




Bottom graphics plot the same data as above, but values less than 10e-1.

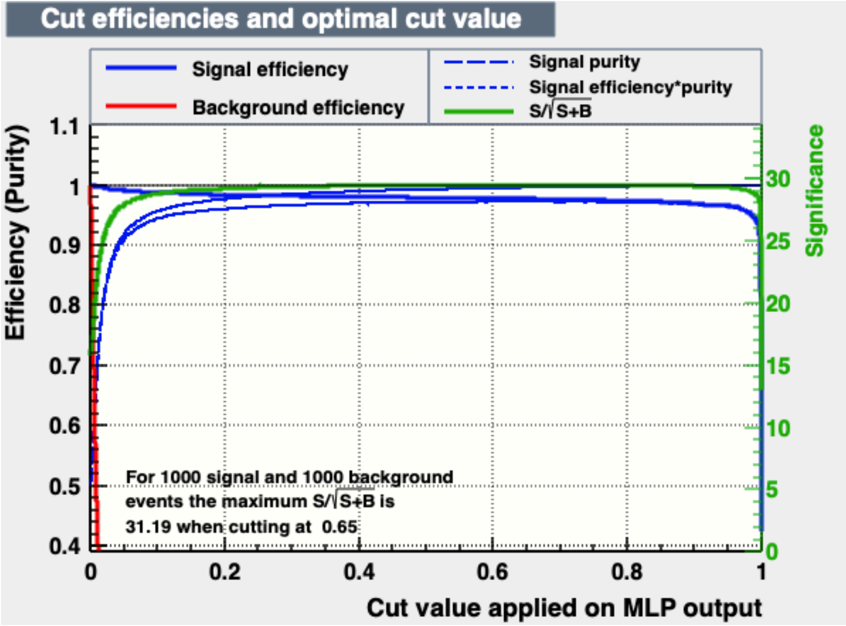


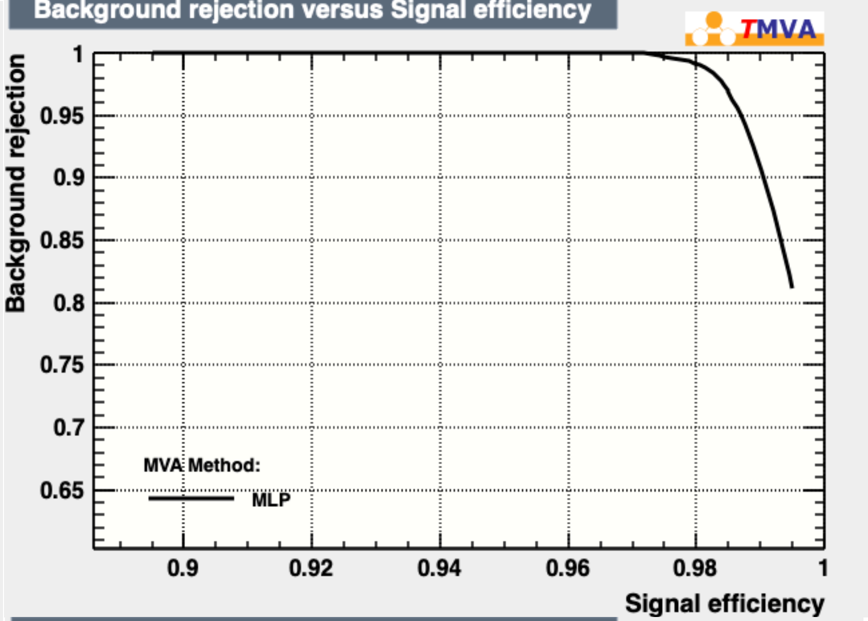
TMVA
Electron FCAL Energy/KinFit momentum magnitude :
Electron FCAL DOCA :
Electron E9/E25 shower ratio :
Bias node :
Layer 0
Layer 1
Output layer

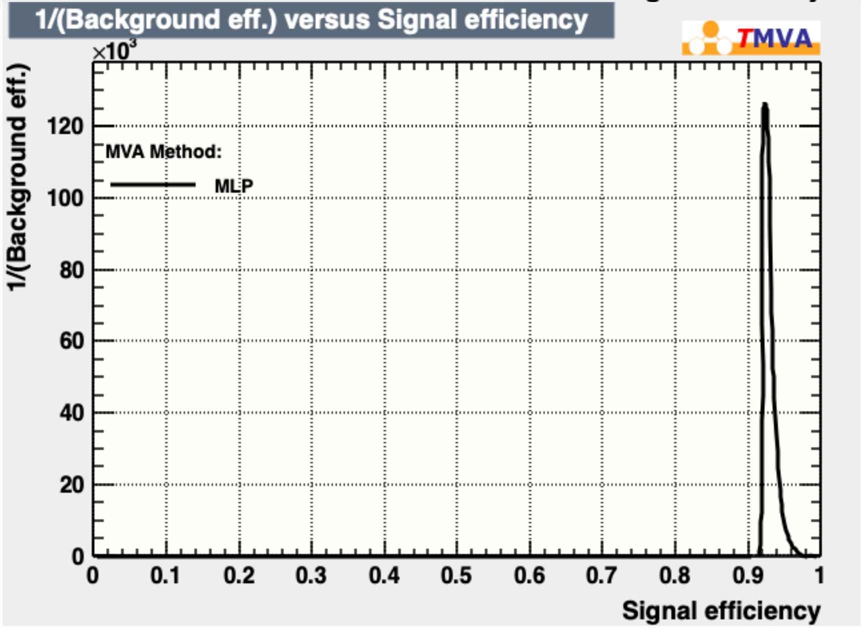

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.17 when cutting at 0.52
Cut value applied on MLPBNN output


Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Efficiency (Purity)
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.16 when cutting at 0.49
Cut value applied on MLPBNN output

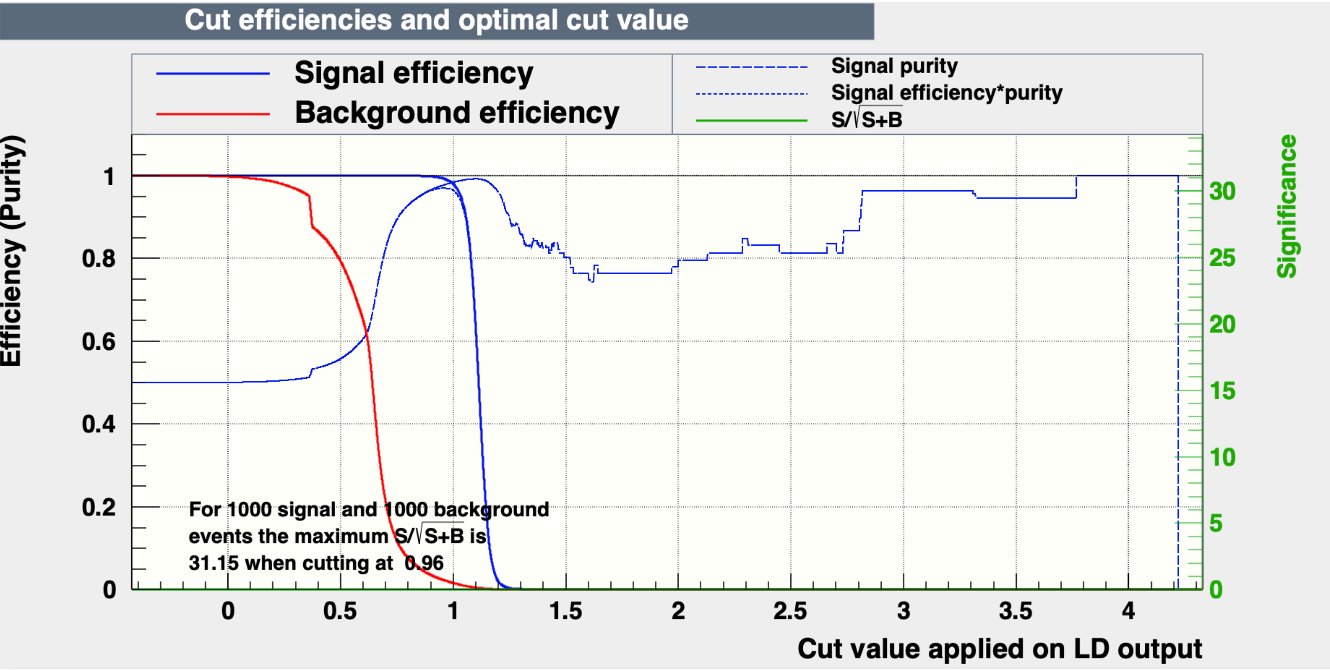
DIRC IN

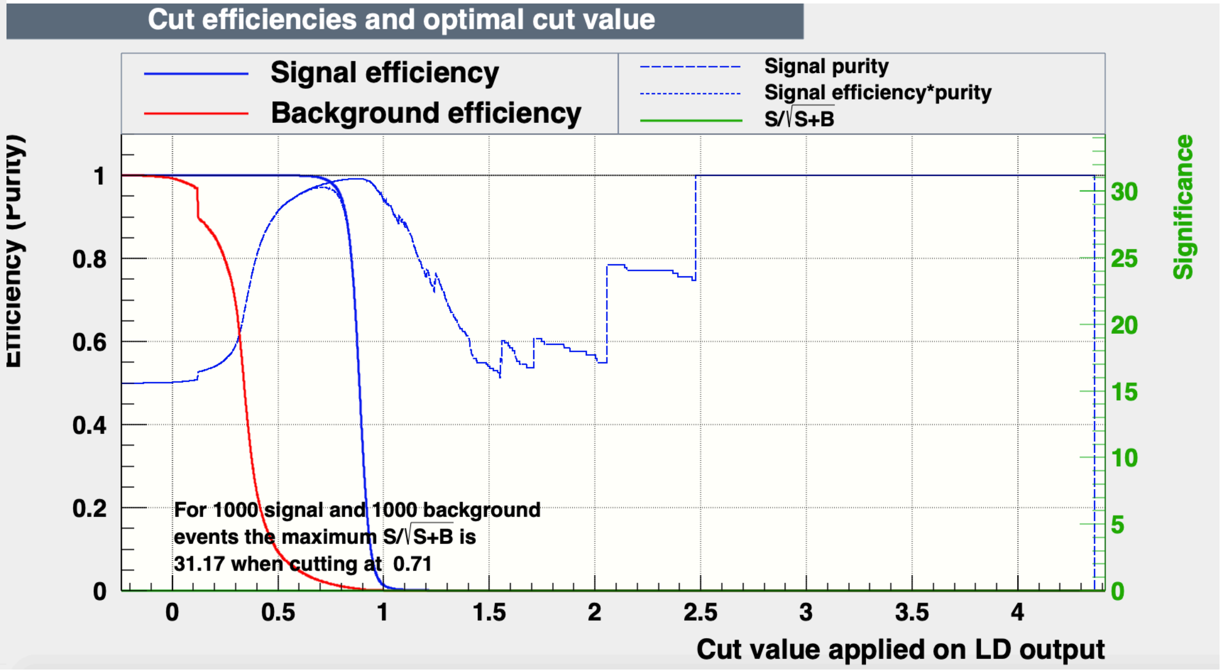

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Efficiency (Purity)
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.19 when cutting at 0.65
Cut value applied on MLP output


Background rejection versus Signal efficiency
TMVA
Background rejection
MVA Method:
MLP
Signal efficiency

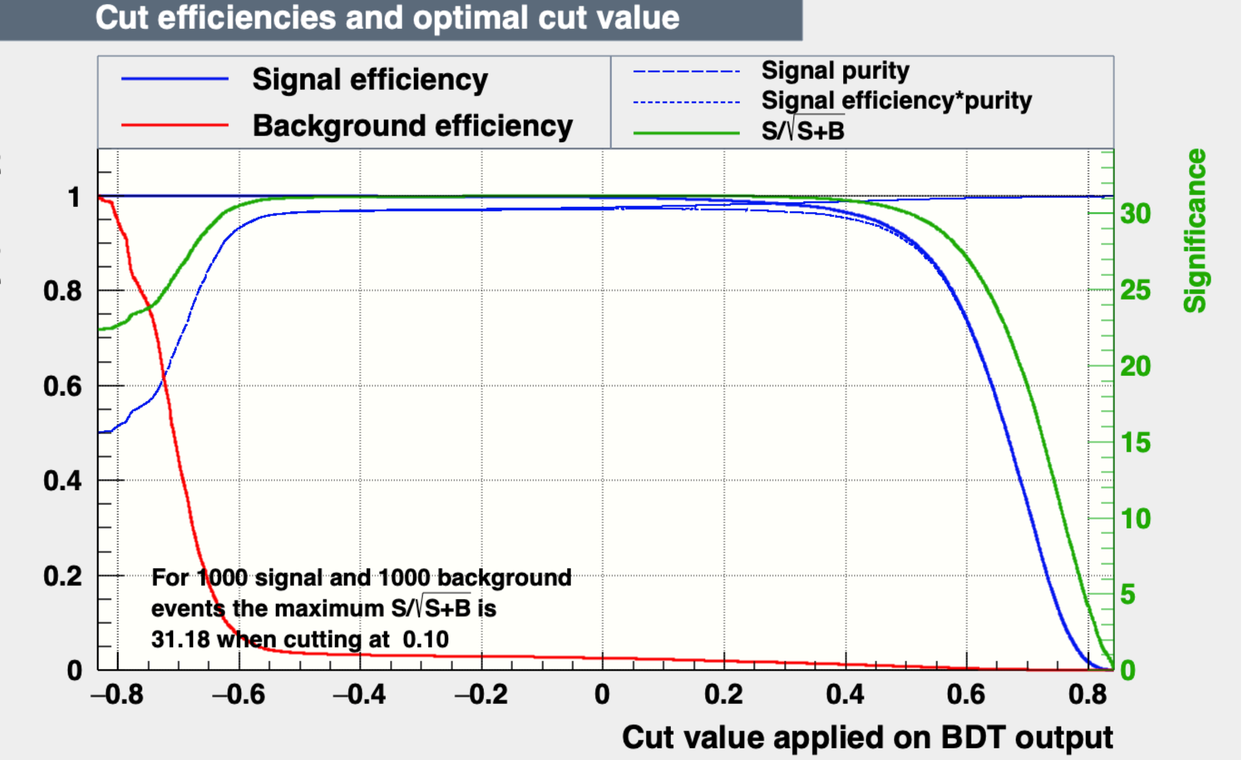

1/(Background eff.) versus Signal efficiency
TMVA
×10³
1/(Background eff.)
MVA Method:
MLP
Signal efficiency

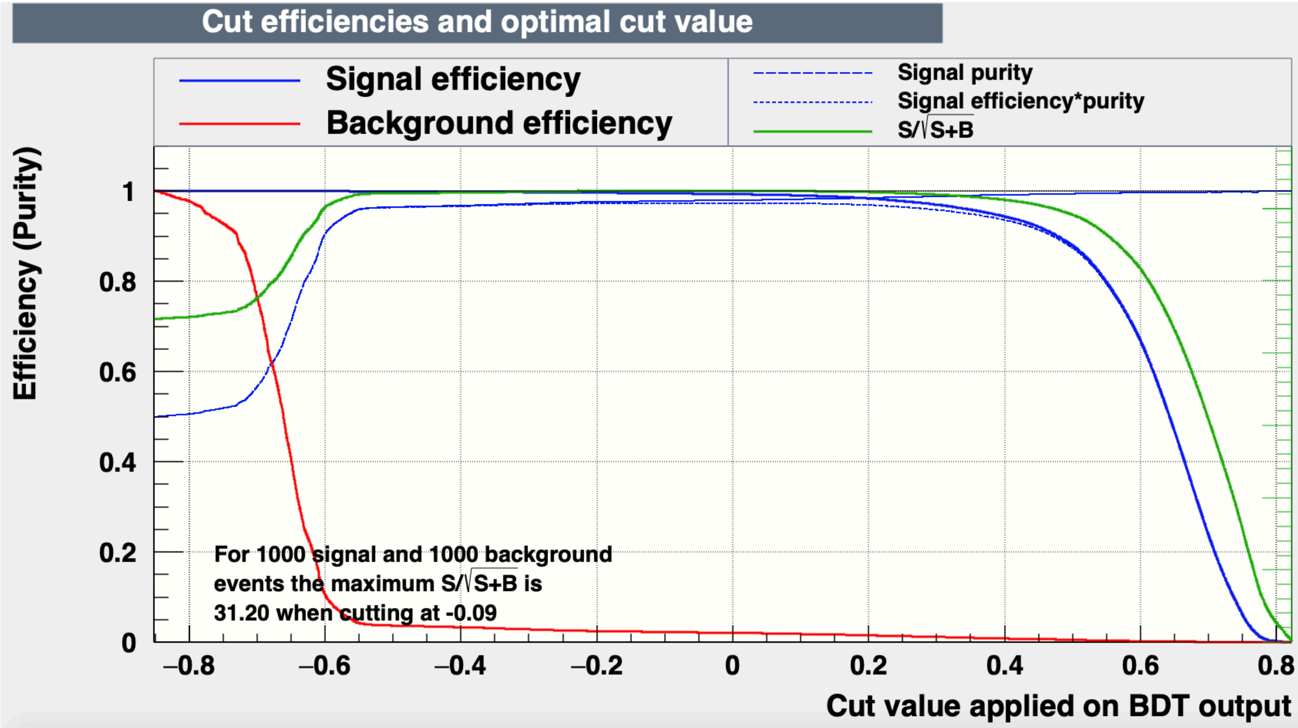
# Old ROC Curves/DIRC study

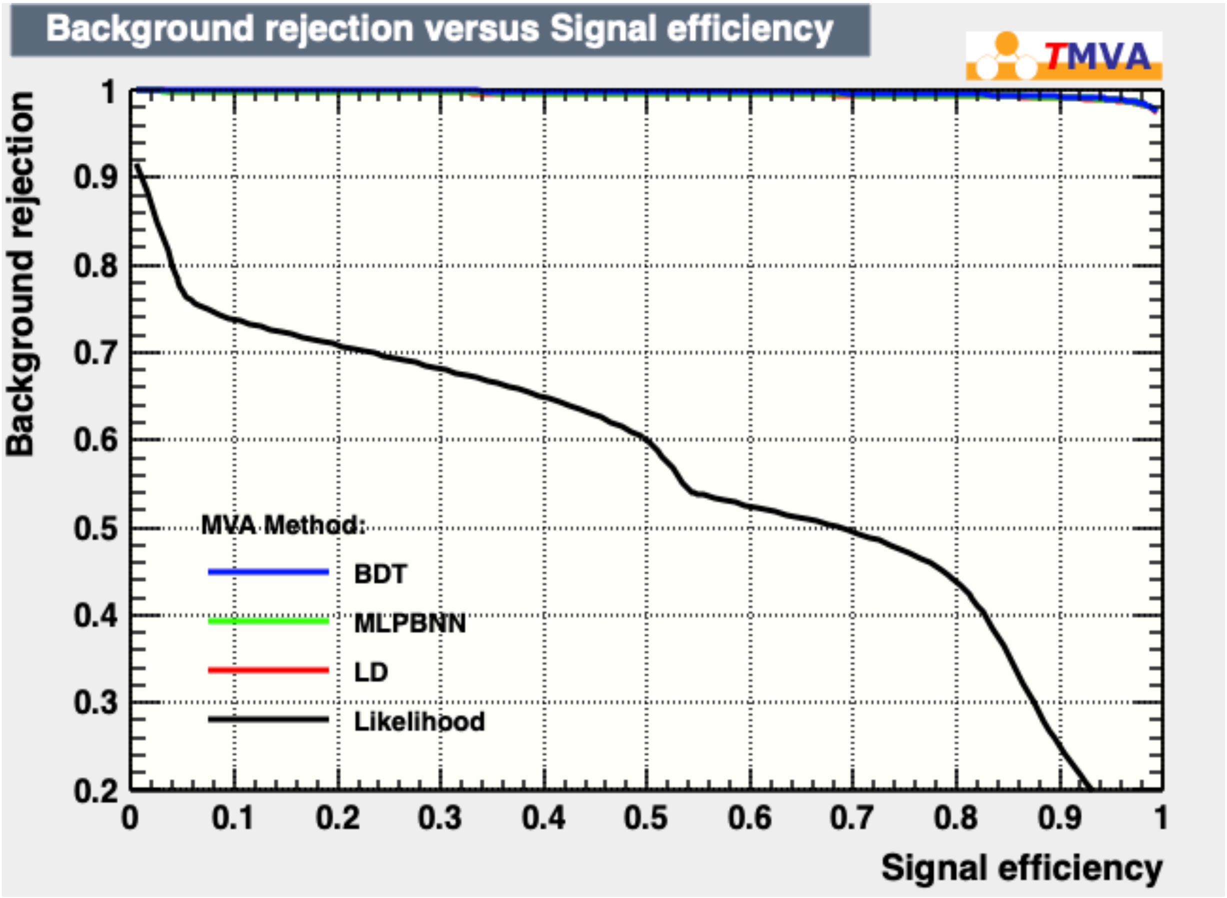

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Efficiency (Purity)
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.15 when cutting at 0.96
Cut value applied on LD output

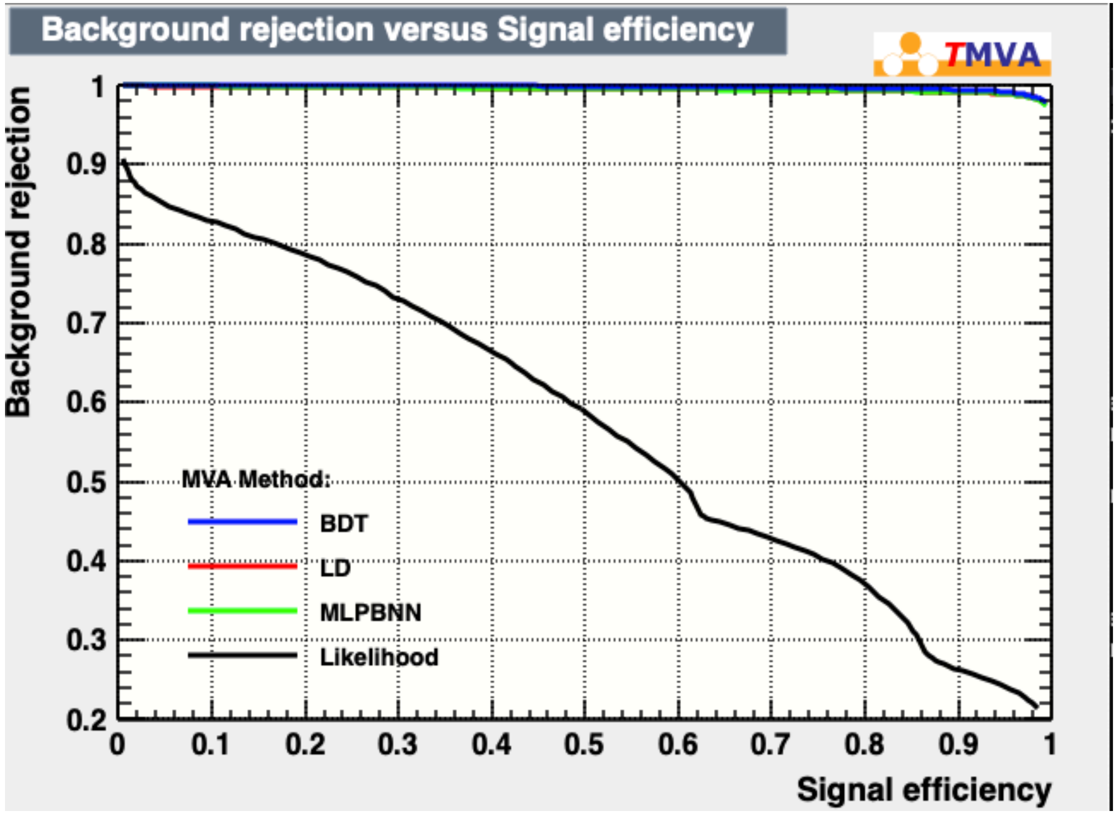

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.17 when cutting at 0.71
Cut value applied on LD output

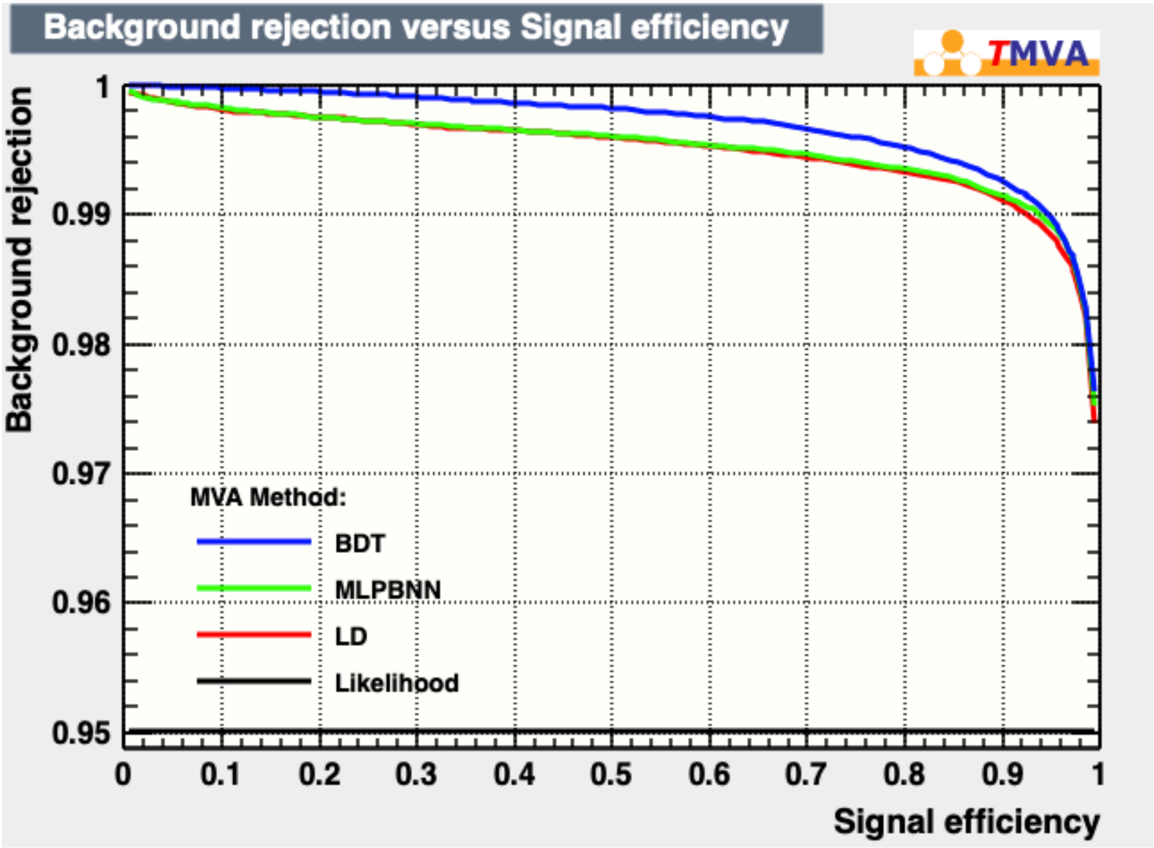
DIRC IN

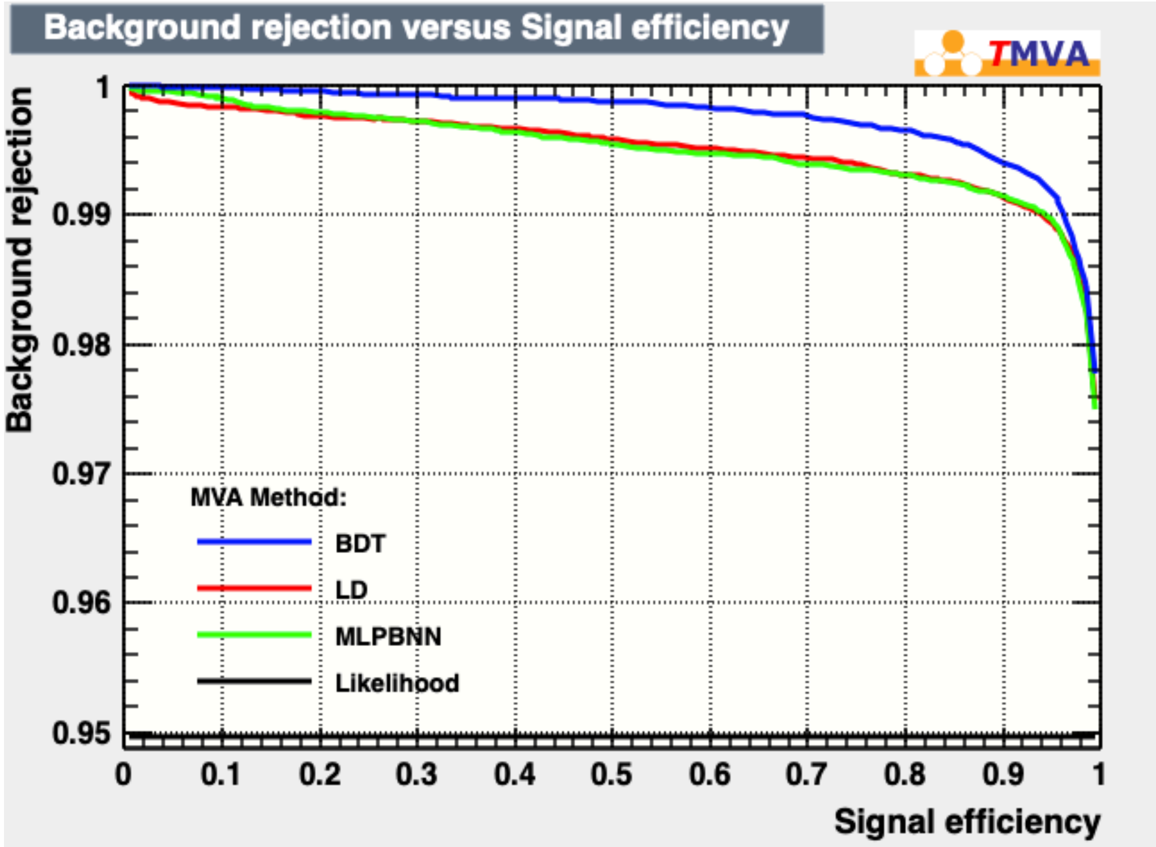

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Significance
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.18 when cutting at 0.10
Cut value applied on BDT output

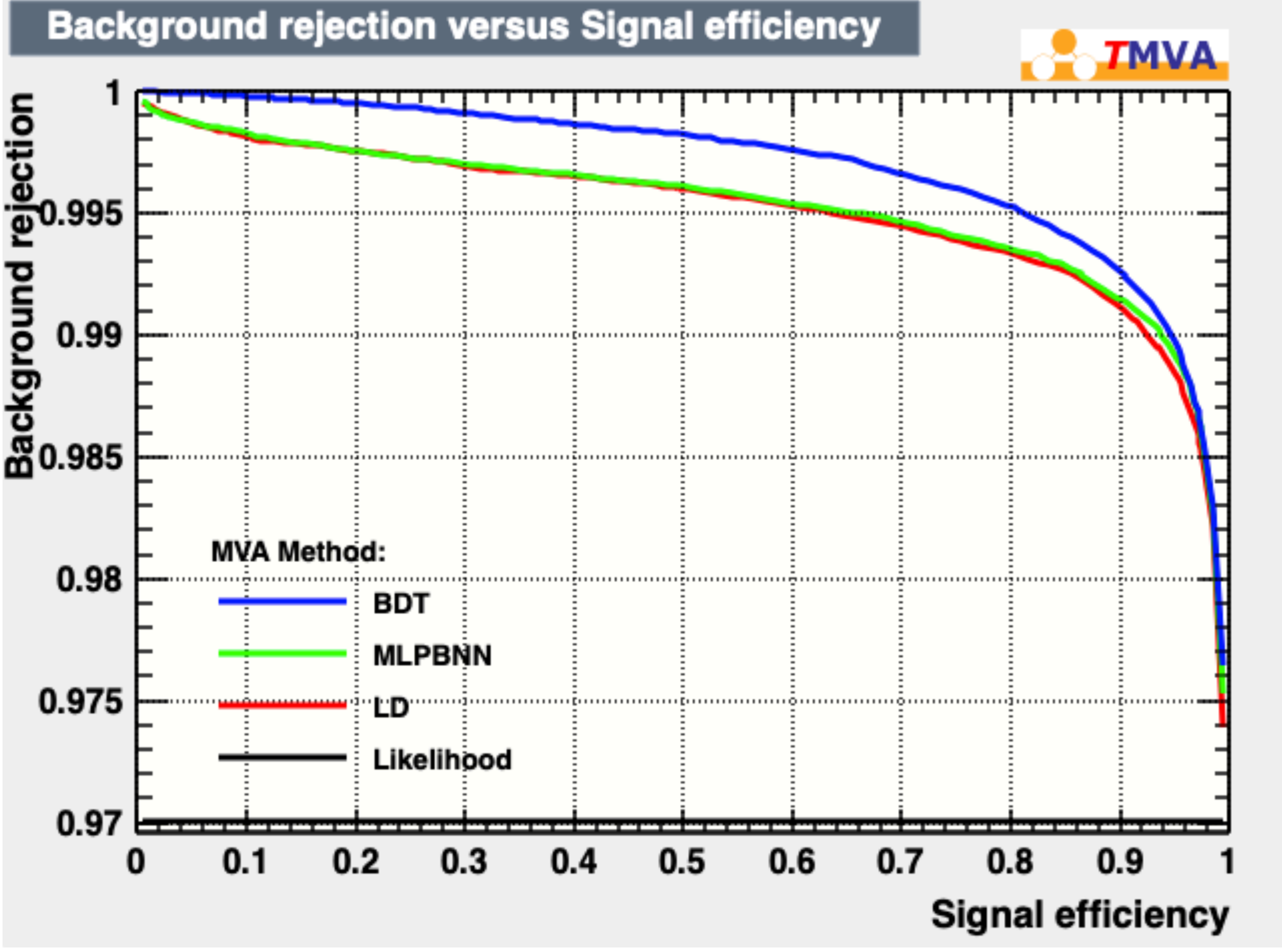

Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/√S+B
Efficiency (Purity)
For 1000 signal and 1000 background
events the maximum S/√S+B is
31.20 when cutting at -0.09
Cut value applied on BDT output


Background rejection versus Signal efficiency
TMVA
Background rejection
Signal efficiency
MVA Method:
BDT
MLPBNN
LD
Likelihood

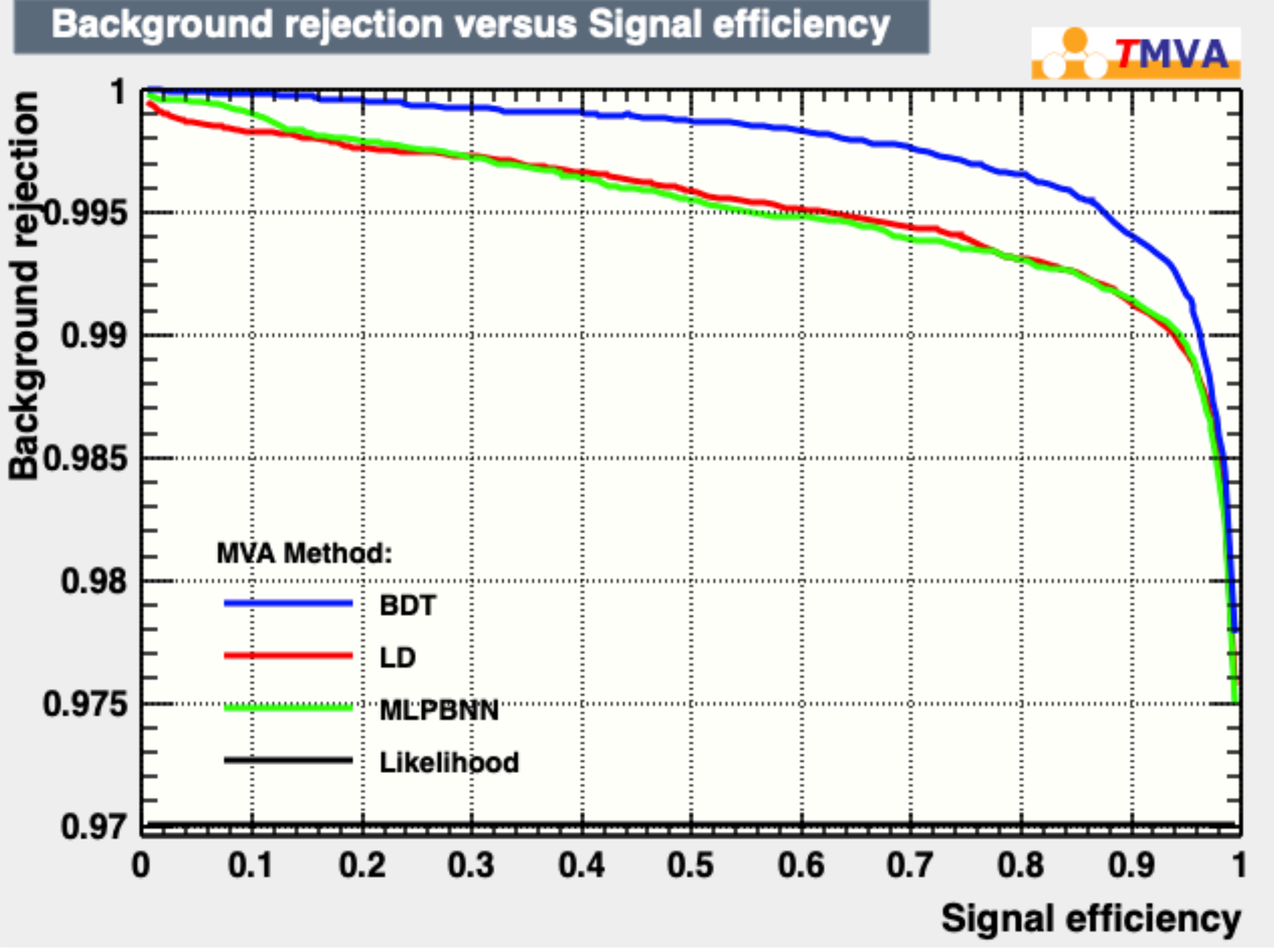

Background rejection versus Signal efficiency
TMVA
Background rejection
Signal efficiency
MVA Method:
BDT
LD
MLPBNN
Likelihood


Background rejection versus Signal efficiency
TMVA
Background rejection
Signal efficiency
1
0.99
0.98
0.97
0.96
0.95
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
MVA Method:
BDT
MLPBNN
LD
Likelihood

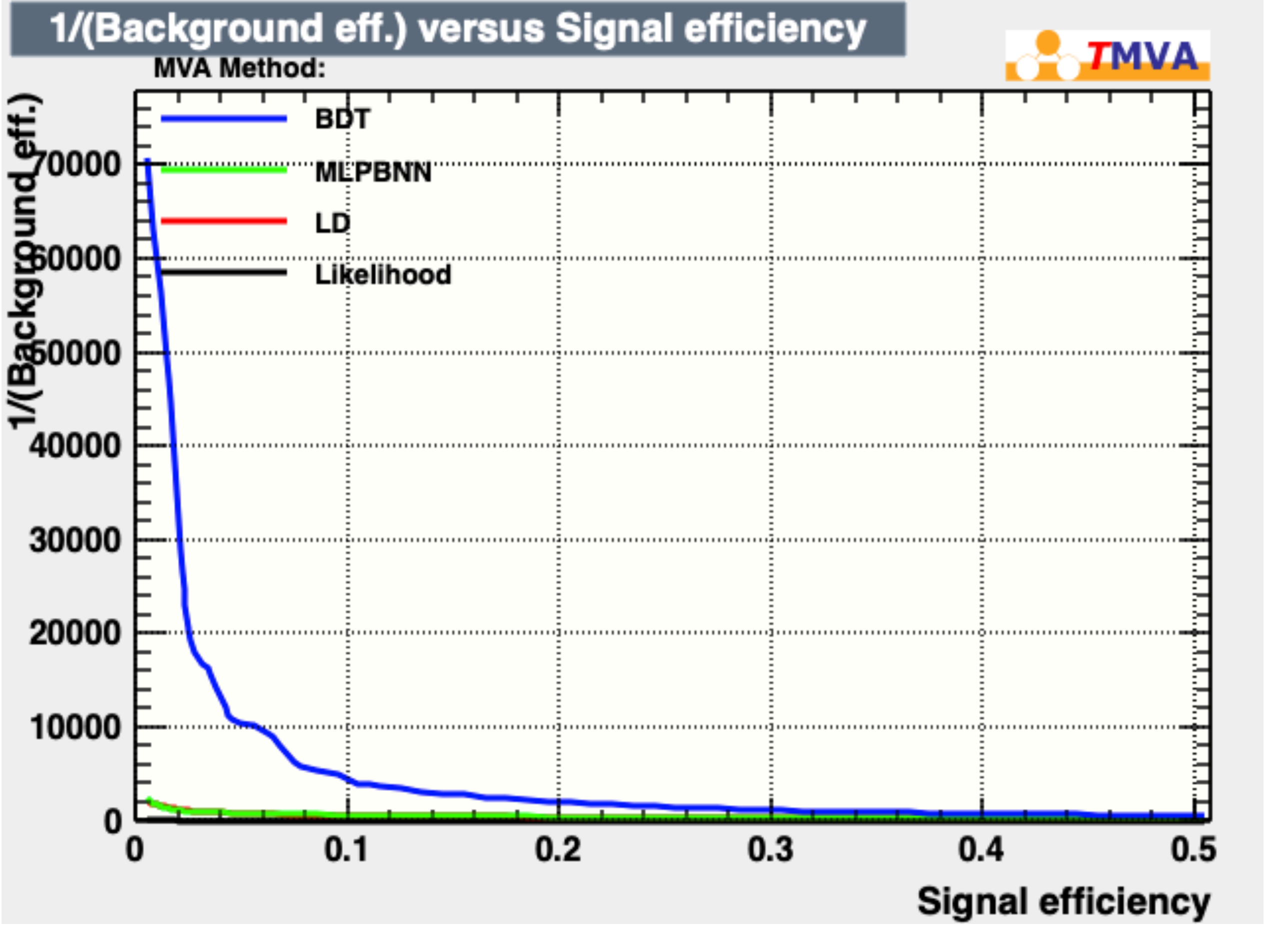

Background rejection versus Signal efficiency
TMVA
Background rejection
Signal efficiency
1
0.99
0.98
0.97
0.96
0.95
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
MVA Method:
BDT
LD
MLPBNN
Likelihood



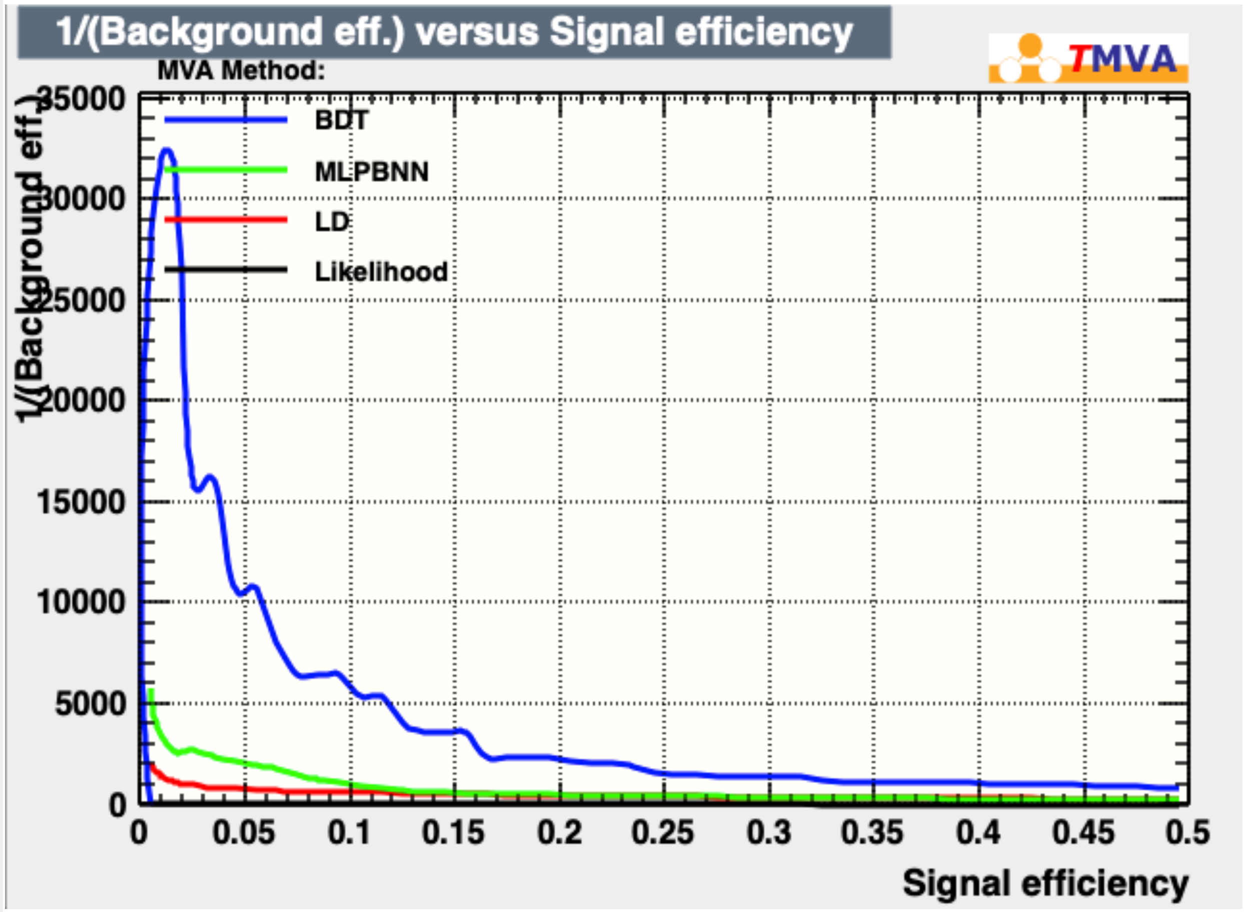

DIRC OUT



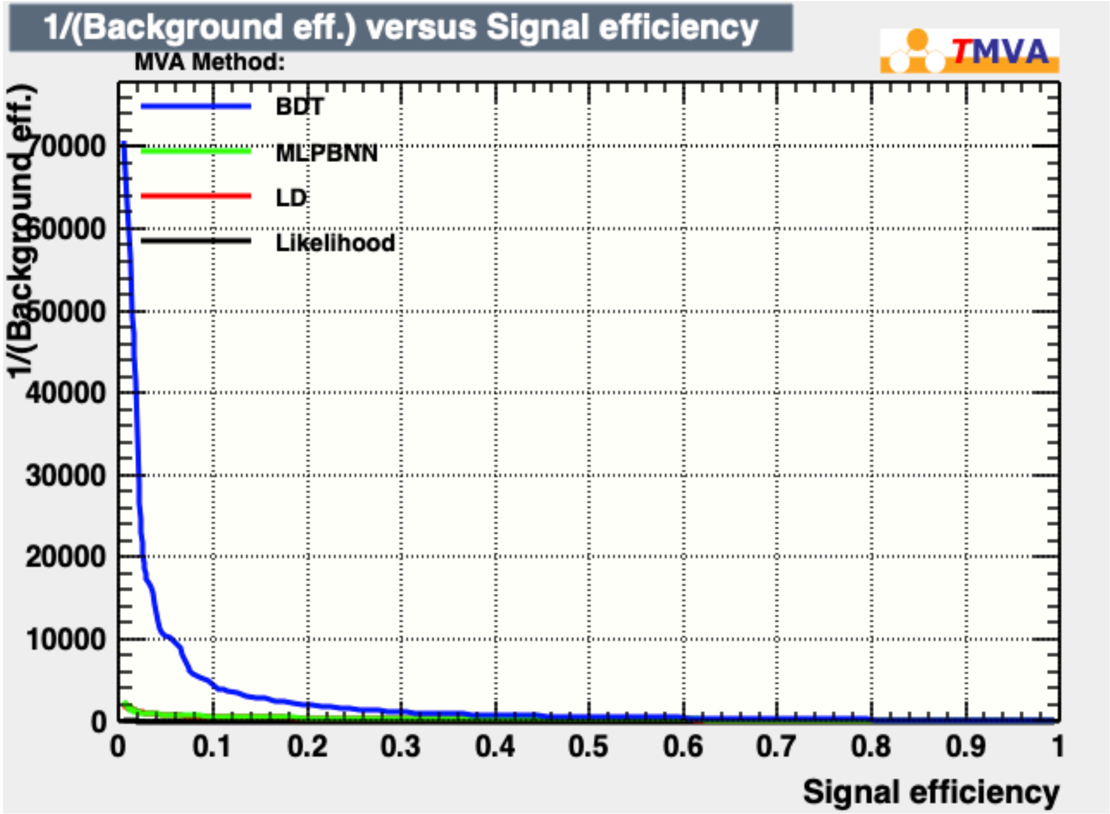

DIRC IN

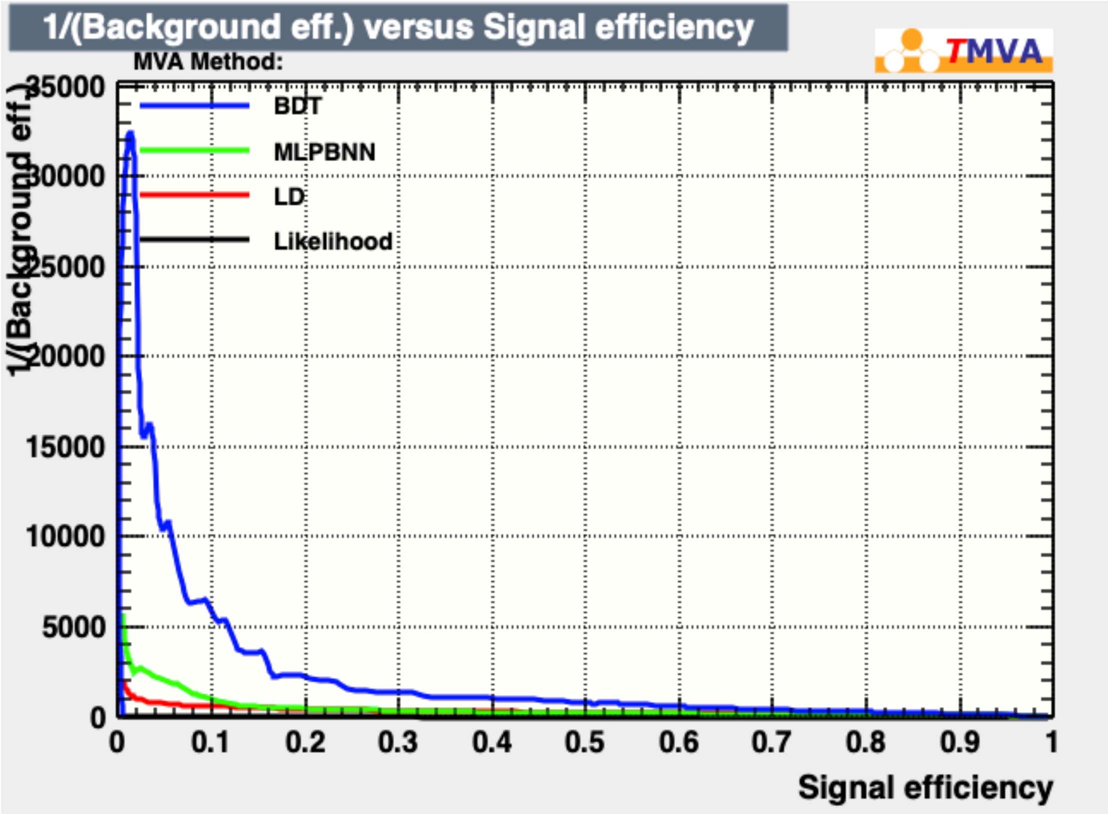



DIRC OUT




DIRC IN



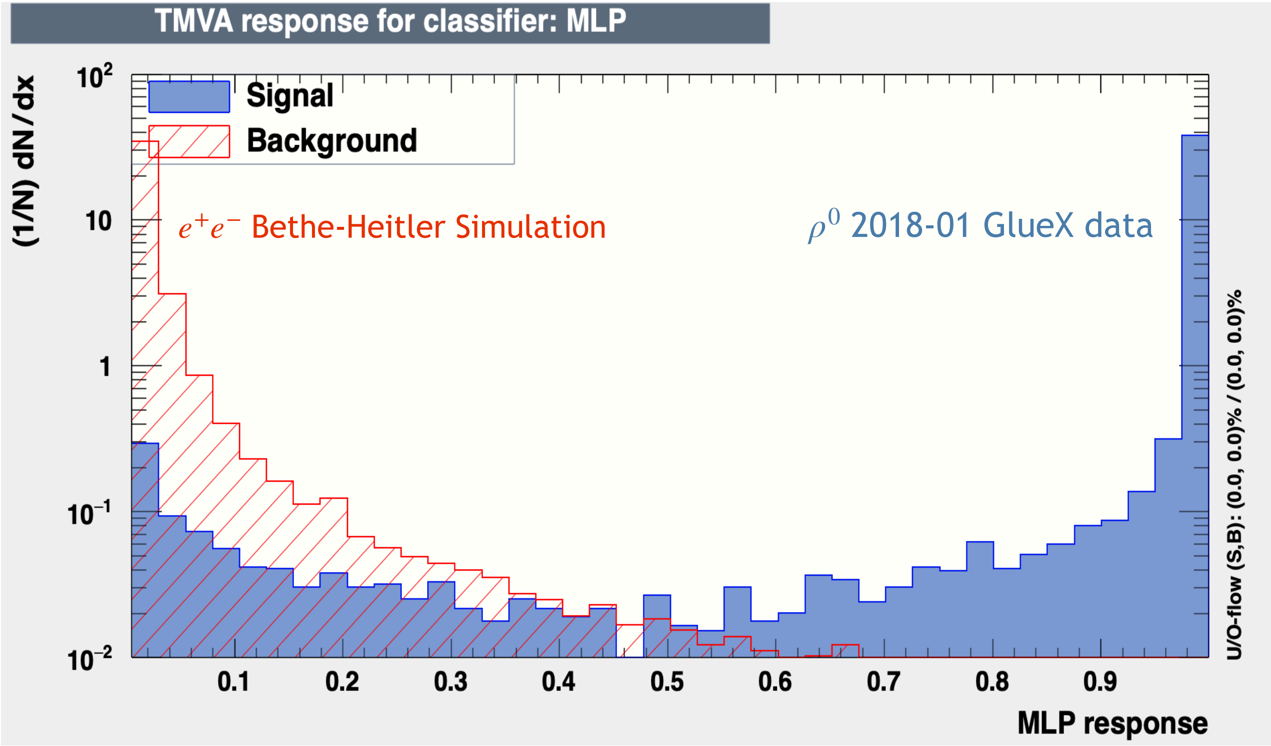




DIRC OUT

DIRC IN

# Extreme Neural Net Response Regions

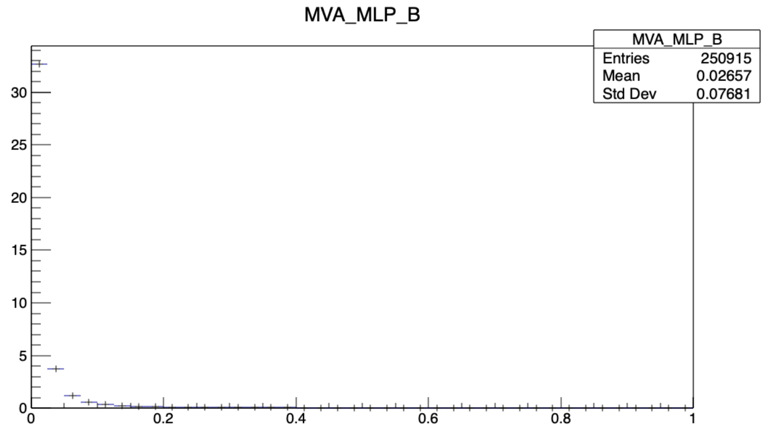
Pions identifying as electrons. How many misclassified electrons would appear in a given interval

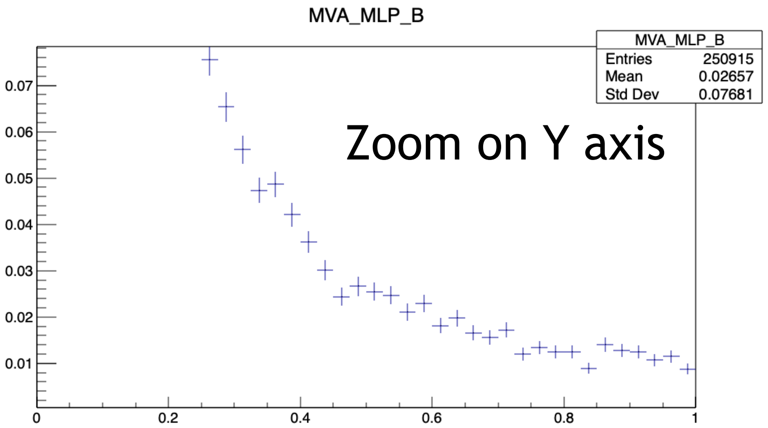


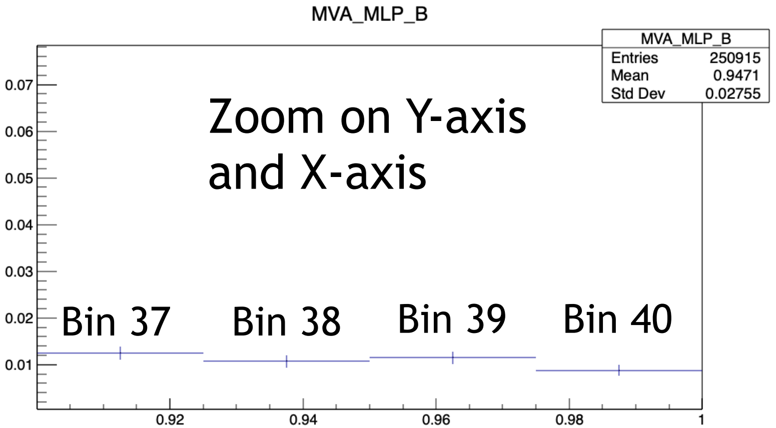

Suppose we wanted to only look at pions with a NN response above 0.9.
In the training sample, do we have perfect separation from electrons?

If we were to do the pi0 study again, where we classify the pi0->ge+e- peak, selecting for pions, how many misclassified electrons could we expect to see above 0.9?

# Simulated **electrons** (used in training) MLP Neural net response values.



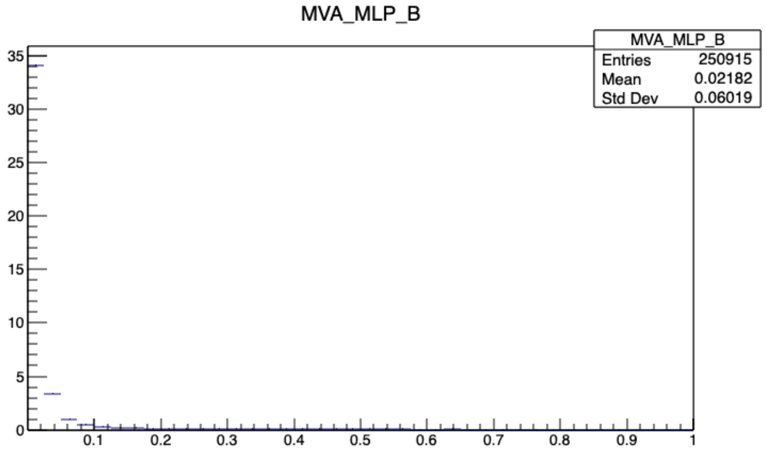

There are 40 bins total.

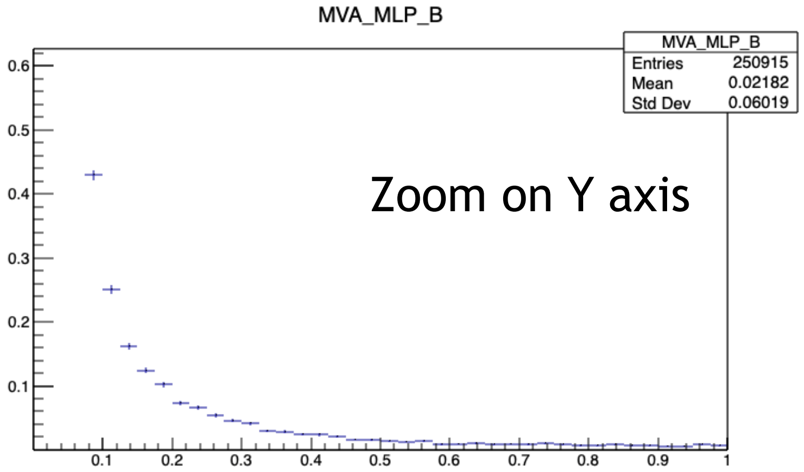
$$\int_1^{40} N_{BH}(x)dx = 39.999664$$

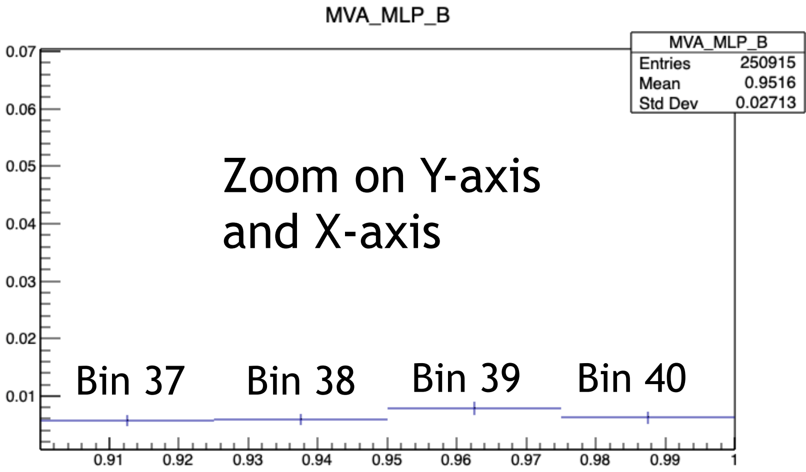
***Likelihood of e- track getting misclassified as π- = 0.1%***

$$\frac{\int_{37}^{40} N_{BH}(x)dx}{\int_1^{40} N_{BH}(x)dx} = \frac{0.030926548}{39.999664} = 0.0010840325$$

Simulated **positrons** (used in training) MLP Neural net response values.



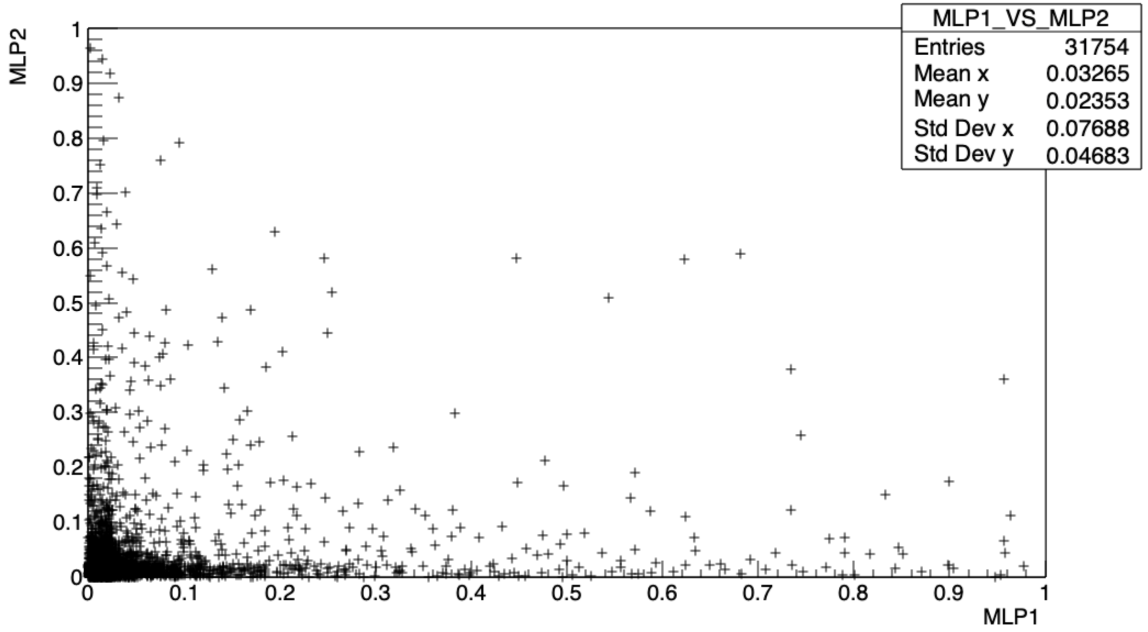

$$\frac{\int_{37}^{40} N_{BH}(x)dx}{\int_{1}^{40} N_{BH}(x)dx} = \frac{0.025666261}{40.000309} = 0.00064165652$$

***Likelihood of e- track getting misclassified as π- = 0.1%***

***Likelihood of e+ track getting misclassified as π+ = 0.06%***

$$0.1\% \times 0.06\,\% = 0.006\,\%$$

# Classifying pi0->ge+e-




| Selecting for Pions | | |
|---|---|---|
| NN cut | omega->pi0pi+pi- (%) | pi0->ge+e- (%) |
| NN > 0.9 | 89.27888272 | 0 |
| NN > 0.6 | 92.98702496 | 0 |
| NN > 0.55 | 93.35576502 | 0.006298419097 |
| NN > 0.4 | 94.32370768 | 0.01259683819 |

LIKELIHOOD OF MISCLASSIFICATION AT NN > 0.9 WAS

$$0.1\% \times 0.06\,\% = 0.006\,\%$$

ZERO PI0 events surviving does not contradict this prediction.

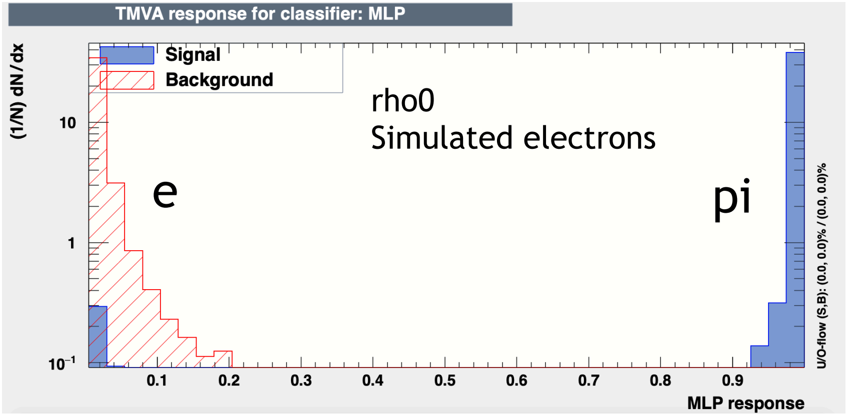
Can actually set NN response cut for pions much lower than I originally thought.

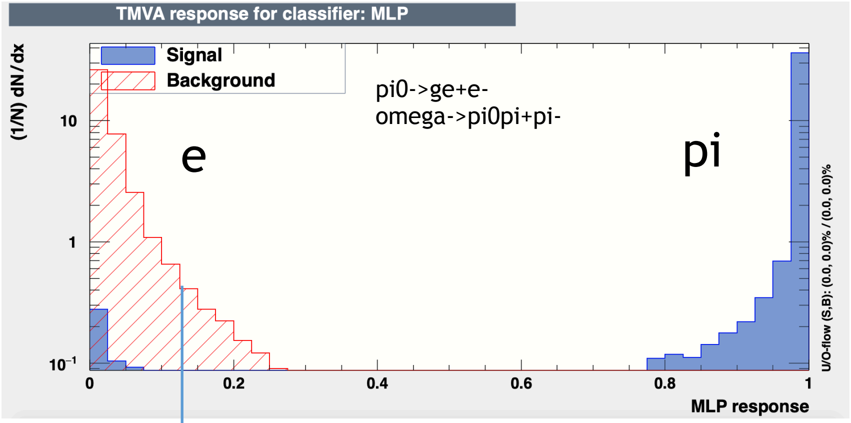
# 2. Low NN response PIONS

Pions have a pileup under electron peak in two separate training sample sources
*(BH MC/rho vs omega/pi0)*

a.) Are the number of events of the low NN response pions the same between the two training sample sources?

b.)
Contamination?
Physics?



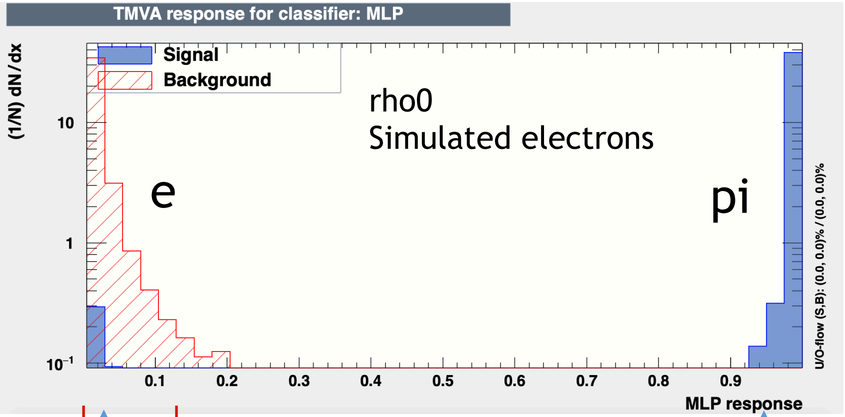

LET'S CALL EVERYTHING BELOW BIN 5 THE "LOW NN RESPONSE REGION" OR NNR = 0.12

*a.) Are the number of events of the low NN response pions the same between the two training sample sources?*

Integral procedure:
MVA response histograms have **40 bins.**

Background peak integral: bins 1 to 5
Signal peak integral: bins 31 to 40

I keep the bounds of integration the same for both rho0 and omega(782)

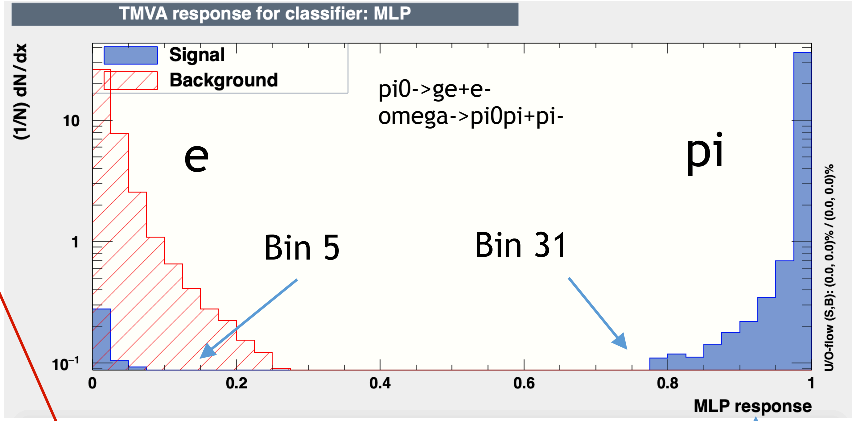



Background peak

Signal peak

$$\frac{\int_1^5 N_{\rho_0}(x)dx}{\int_{31}^{40} N_{\rho_0}(x)dx} = \frac{0.76404053}{39.046492} = 0.019567456$$

1000*.019 = 19 events +/- 4 events
(sqrt(19) ~ 4, statistics)




Background peak

Signal peak

$$\frac{\int_1^5 N_{\omega(782)}(x)dx}{\int_{31}^{40} N_{\omega(782)}(x)dx} = \frac{0.67999367}{38.519641} = 0.017653167$$

***roughly 2% in both training methods***

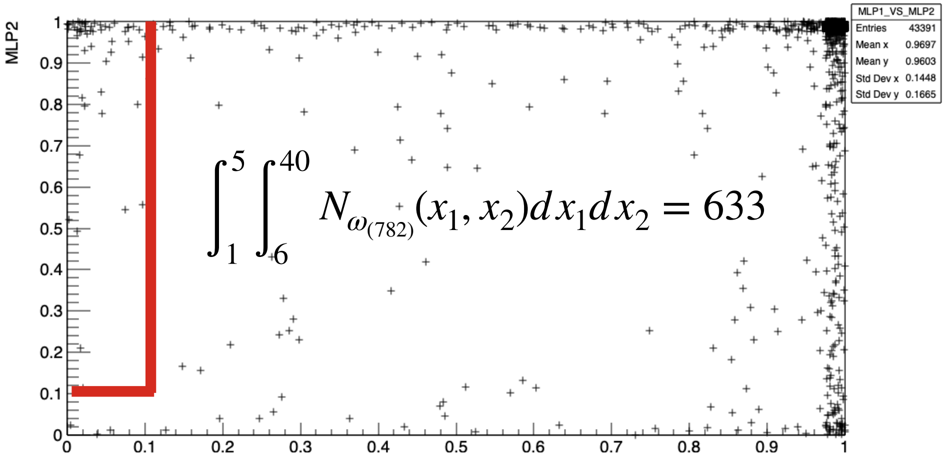
LET'S CALL EVERYTHING BELOW BIN 5 THE "LOW NN RESPONSE REGION" OR NNR = 0.12

# Conclusions

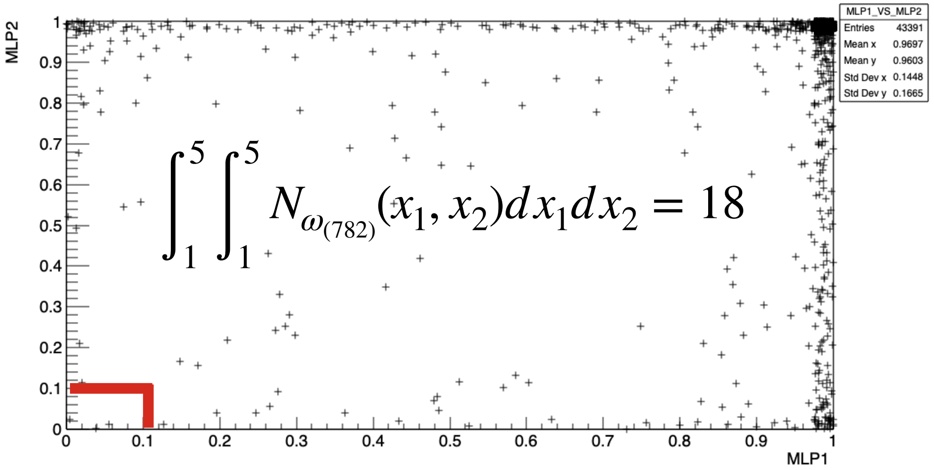
- There is no significant contamination in the rho0 training vs the omega(782) training.

# **Integration of 3 complementary regions:** *how many pion events have at least 1 track in the extreme e+e- NN response territory?*

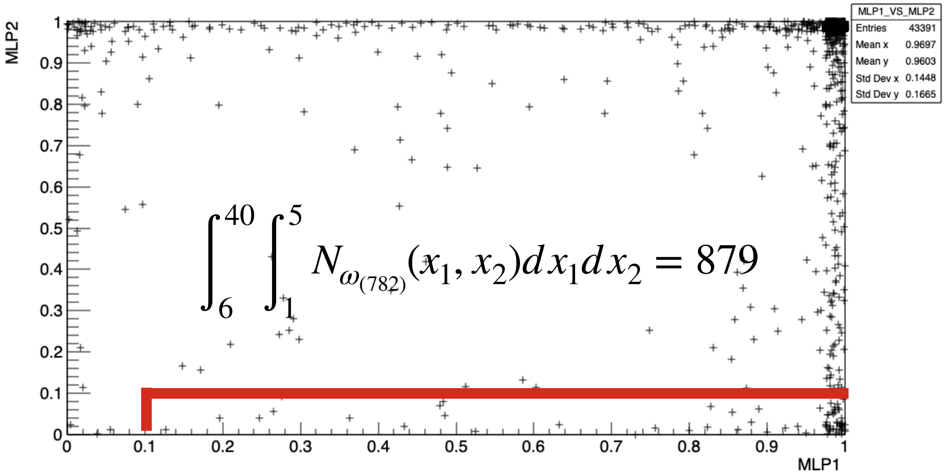
Trained on simulated BH pairs and $\rho^0 \rightarrow \pi^+\pi^-$ events
Classifing $\omega(782) \rightarrow \pi^0\pi^+\pi^-$ events




$$\int_1^5 \int_6^{40} N_{\omega(782)}(x_1, x_2)dx_1dx_2 = 633$$




$$\int_1^5 \int_1^5 N_{\omega(782)}(x_1, x_2)dx_1dx_2 = 18$$



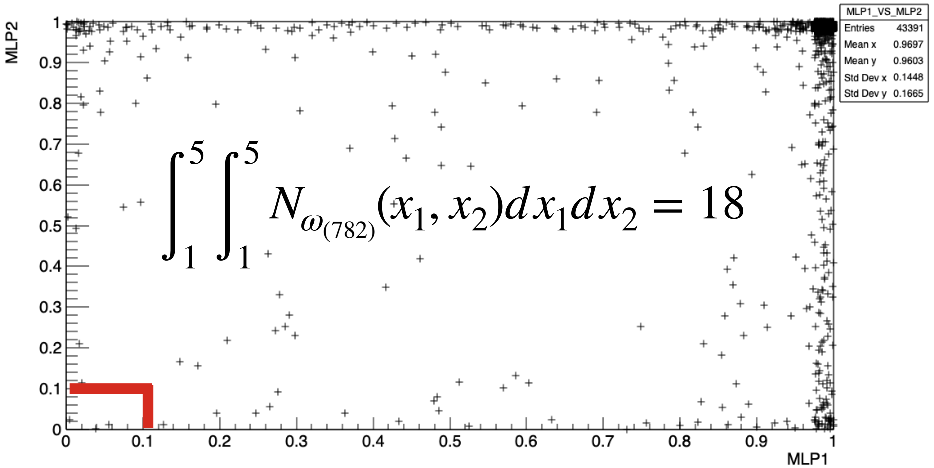

$$\int_6^{40} \int_1^5 N_{\omega(782)}(x_1, x_2)dx_1dx_2 = 879$$

$$\frac{18 + 633 + 879}{43391} = \frac{1530}{43391} = 0.035$$

3.5% of events need to be explained.

# 3.5% of events need to be explained.

1. Double low NN response



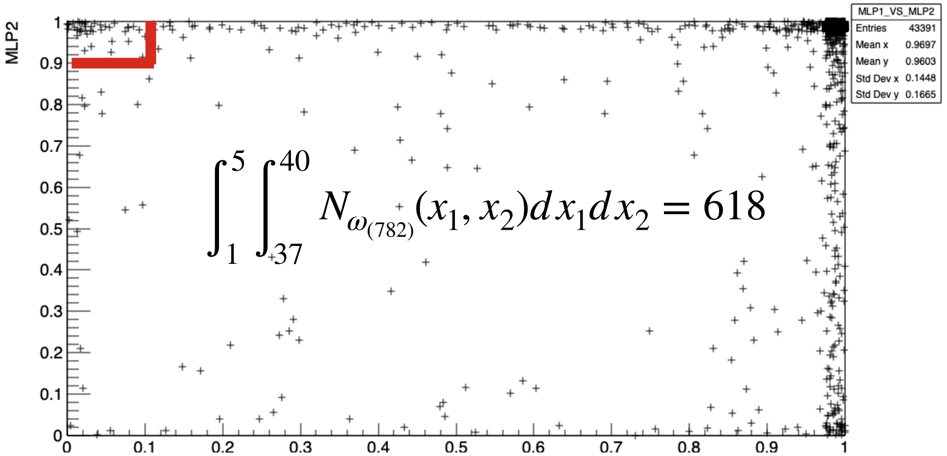

.04% is basically 2% of 2%. Could just be events where both events charge exchange reaction

$0.04\,\%$ of $\omega(782) \rightarrow \pi^0\pi^+\pi^-$ events have both tracks with low NN response values.

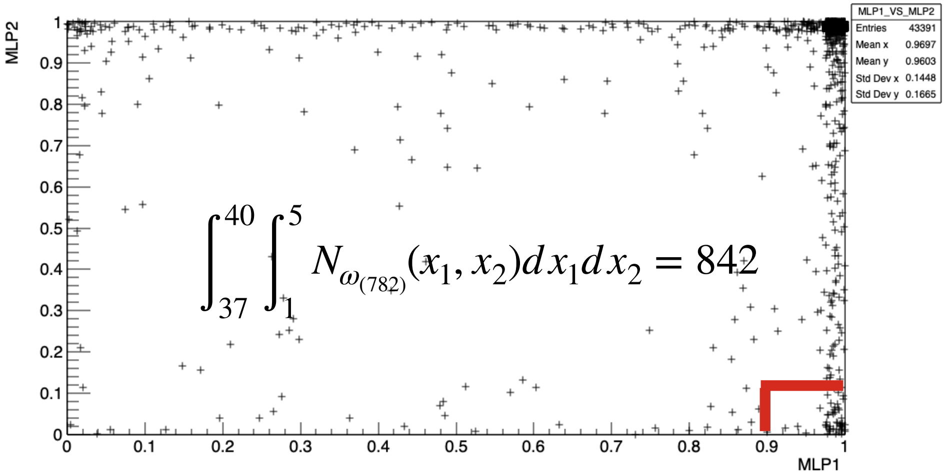
Let's assume (for no good reason other than playing a game) that this is contamination. Now 3.46% events need to be accounted for.

# 3.46% of events need to be explained

2. One high NN response, one low. (NN1 > 0.9, NN2 < 0.12)




$$\int_{1}^{5}\int_{37}^{40} N_{\omega(782)}(x_1, x_2)dx_1dx_2 = 618$$




$$\int_{37}^{40}\int_{1}^{5} N_{\omega(782)}(x_1, x_2)dx_1dx_2 = 842$$

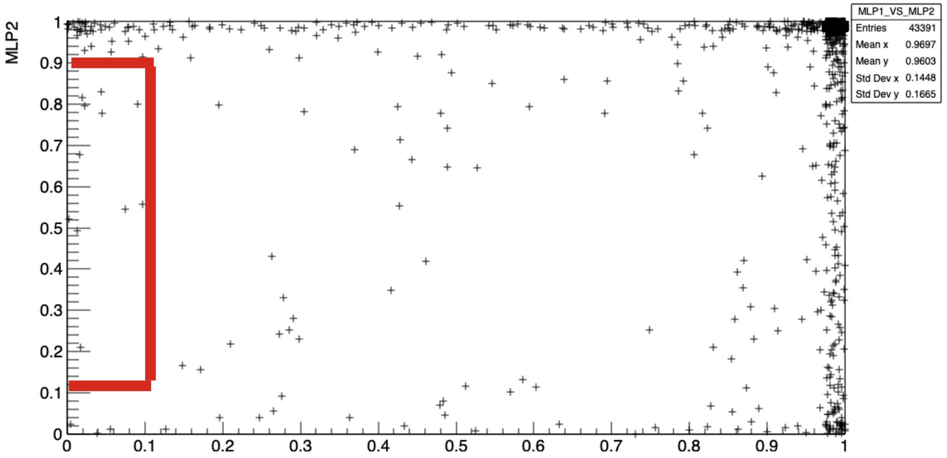
$$\frac{618 + 842}{43391} = 0.033647531$$

Let's assume (again, for no good reason)
3.3% of events have charge exchange reaction (π+ or π-)

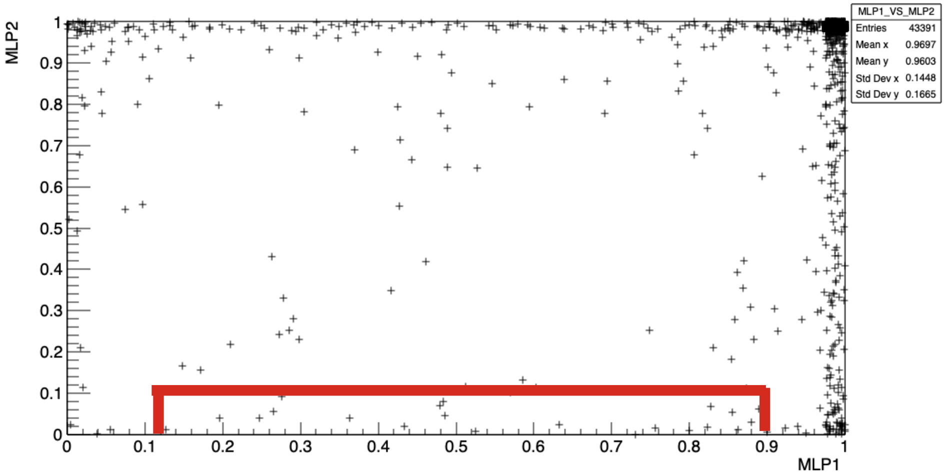
Now 0.16% of events need to be accounted for.

# This leaves 0.16% of events left populating “no man’s land”

2. “No man’s land”




Actually let’s pair 0.16% with the 0.04% that we called contamination into a 0.2% group




***Is it reasonable?***

3.3% total charge exchange reaction (π+ and π-)
0.2% contamination/unknown

# Benchmark Study 1

Use alternative pure sources of $e^+e^-$ (from $\pi^0$ Dalitz decay) and $\pi^+\pi^-$ (from $\omega(782)$ decay) to test Neural Net classifier output distributions from training.

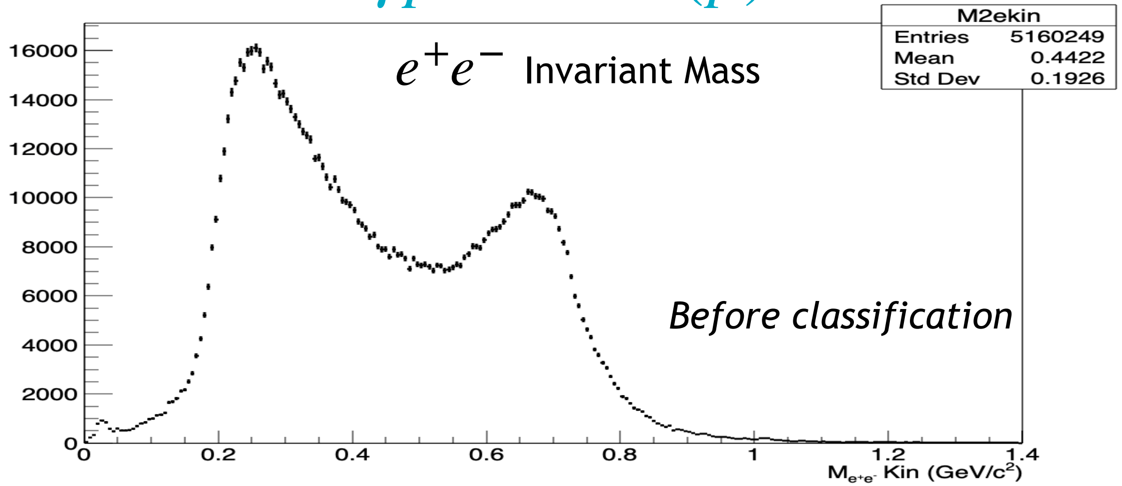
# Benchmark Studies

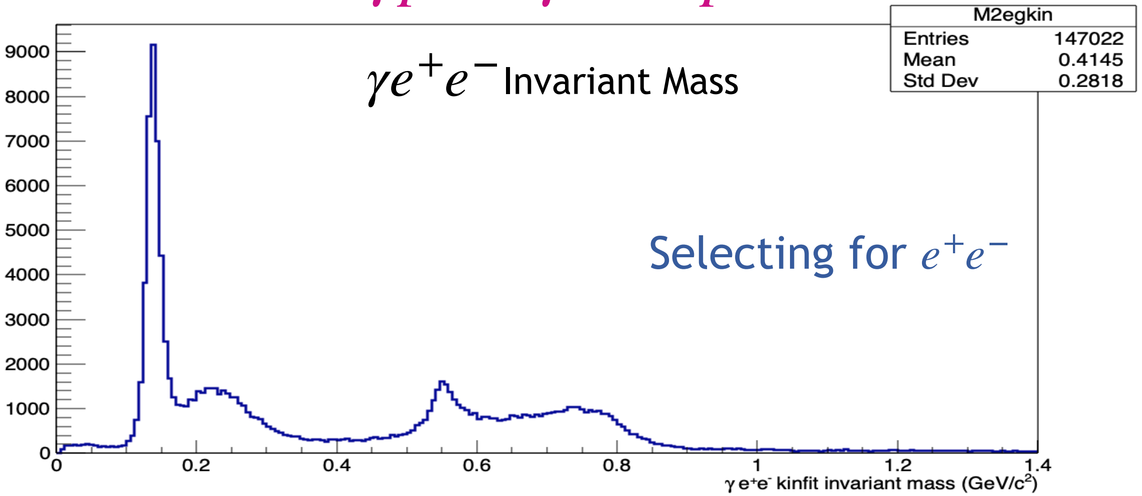
**LEFT:** 2018 GlueX data containing BH pairs and Rho0. Use NN to classify and separate.

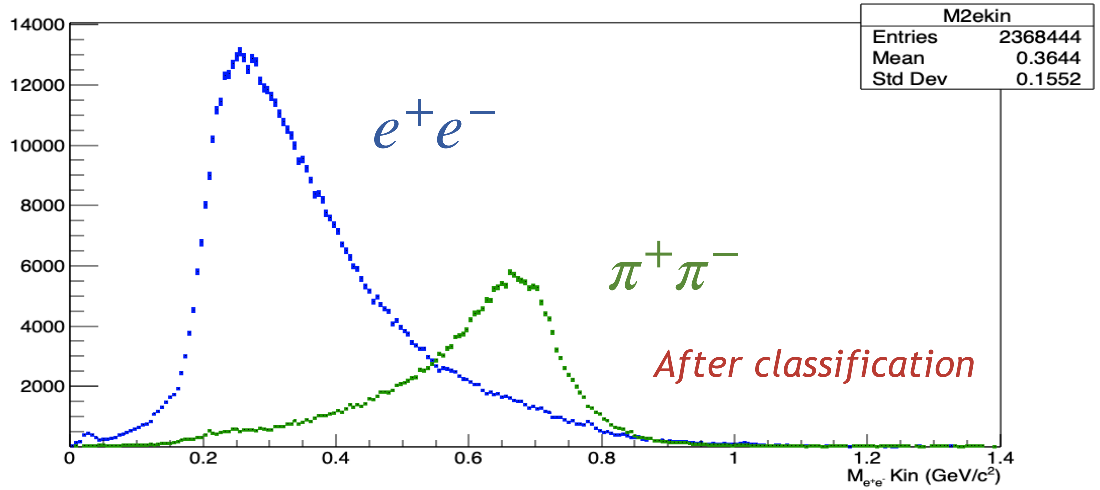
**RIGHT:** 2018 GlueX data containing pi0 Dalitz decay. Select for pions and see how many e+e- pairs from pi0 get through.

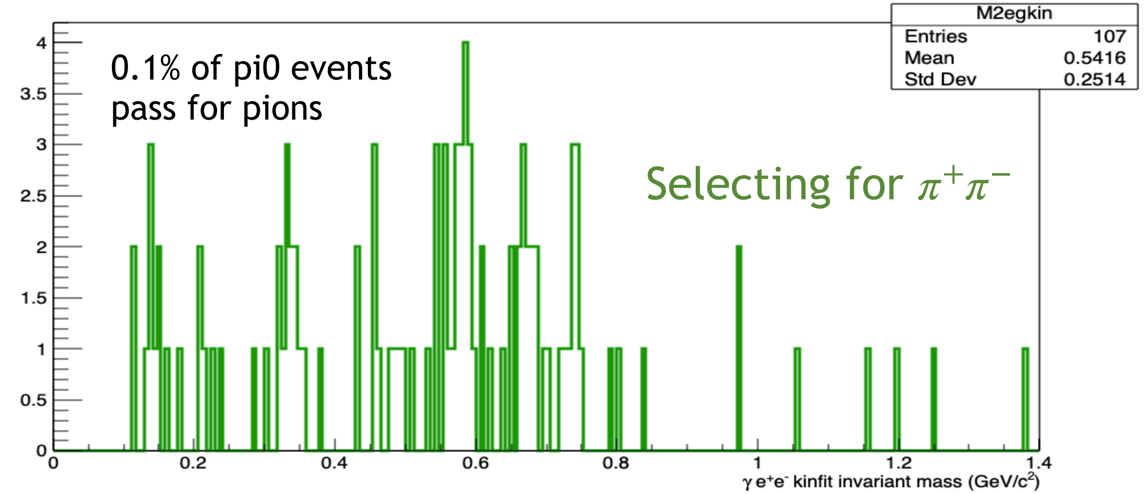
$\gamma p \rightarrow e^+e^-(p)$

$\gamma p \rightarrow \gamma e^+e^- p$



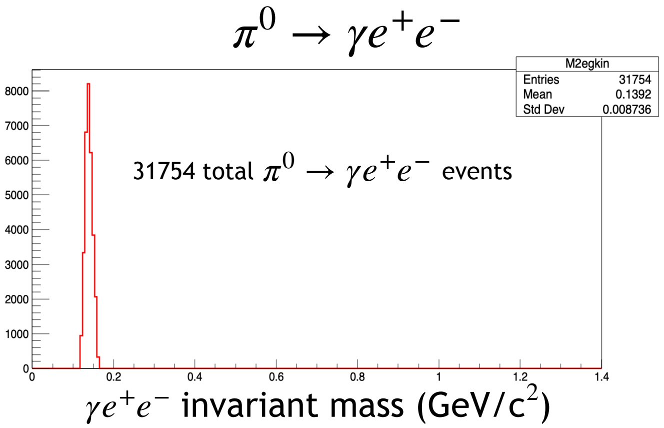



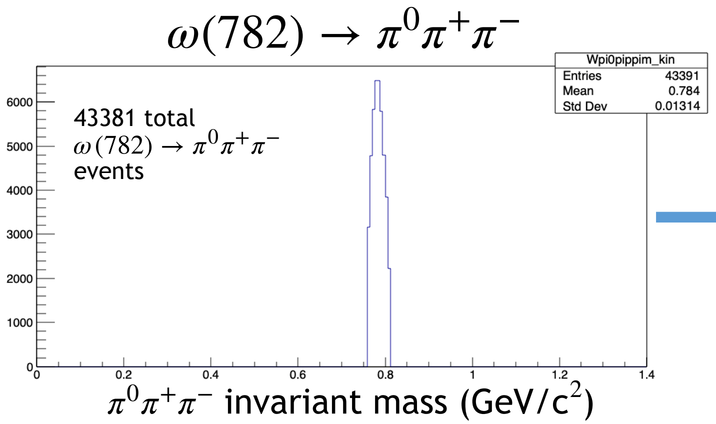



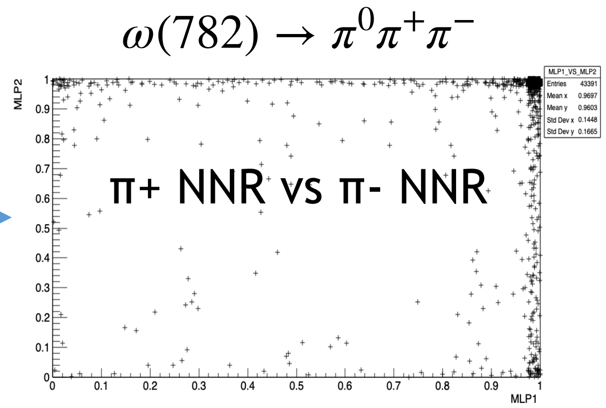



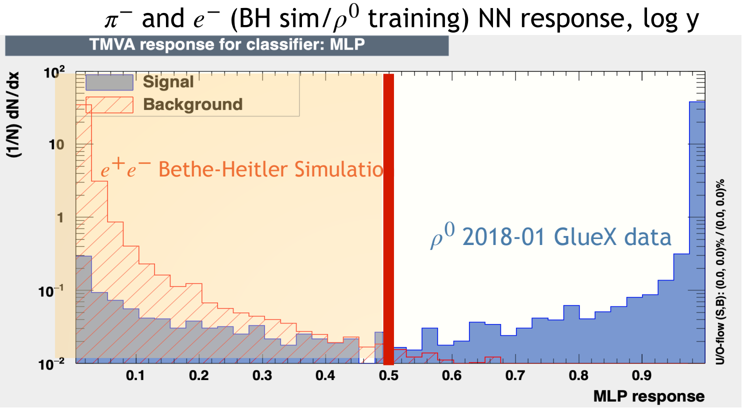

**Same neural net and cut on NN response used in both studies**

1: Get neural net response for e+e- from π0 and π+π- events from omega



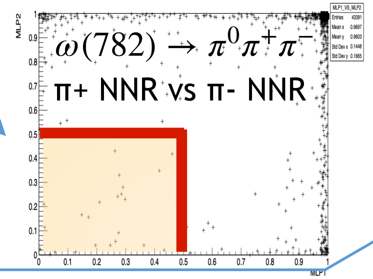



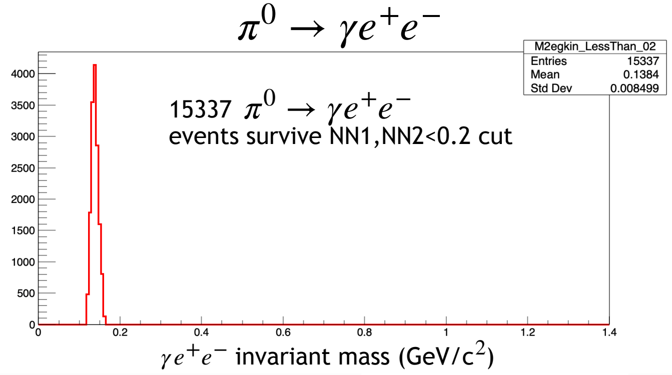



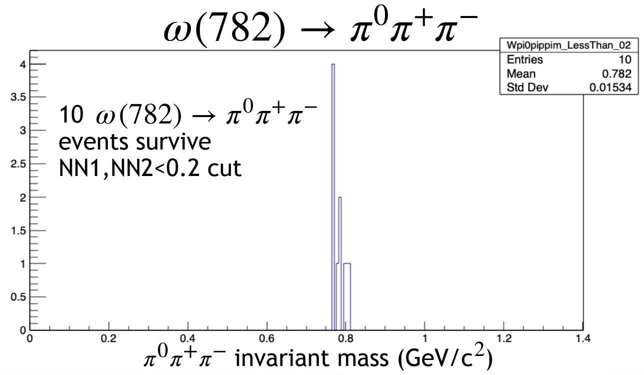

## 2. Pick a neural net response to be the integral boundary. Integrate all events to the left of that "cut point")

**since we're applying the same NNR cut to *both* tracks, it's really integrating within the square below:




3. Record how many events of each type get integrated for the cut point. Repeat steps 2 and 3 with a new cut point.



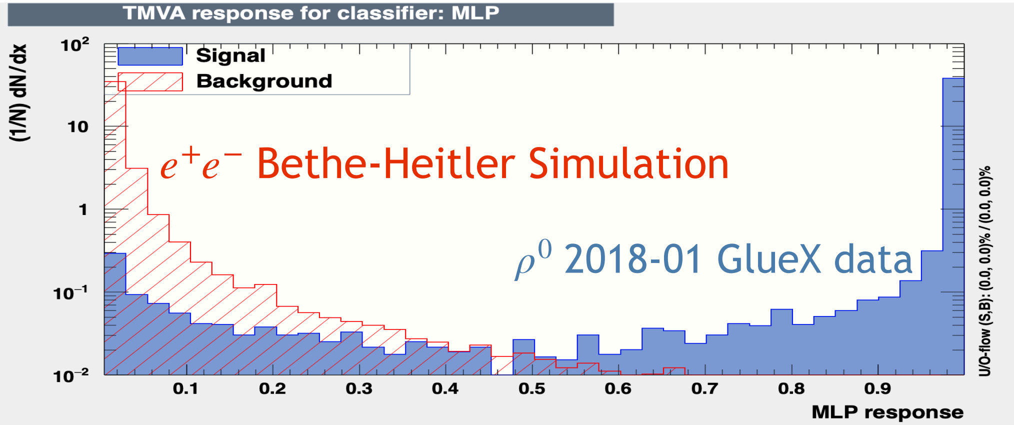




| NN< | omega->Pion Efficency | pi0->Electron Efficiency | Pions survived | e+e- survived |
|---|---|---|---|---|
| 0.01 | 0.01613237768 | 20.81627511 | 7 | 6610 |
| 0.02 | 0.02304625383 | 48.29942684 | 10 | 15337 |
| 0.03 | 0.02996012998 | 64.04232538 | 13 | 20336 |
| 0.04 | 0.03226475536 | 72.84121685 | 14 | 23130 |
| 0.05 | 0.03456938075 | 78.32713989 | 15 | 24872 |
| 0.06 | 0.03917863151 | 82.0904453 | 17 | 26067 |
| 0.07 | 0.03917863151 | 84.9278831 | 17 | 26968 |
| 0.08 | 0.03917863151 | 87.09768848 | 17 | 27657 |

# Two Integration Sweeps:

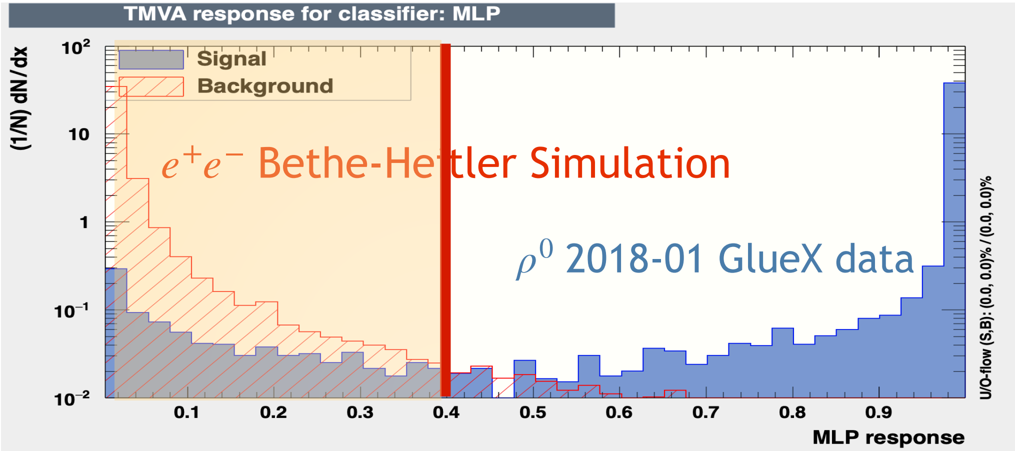
**On this slide I'm showing the NN response for a single track from the training of the NN. This is only for building intuition on what regions will accept/reject pions and electrons. The actual NN response plots of the events I'm integrating are on subsequent slides.

$\pi^-$ and $e^-$ (BH sim/$\rho^0$ training) NN response, log y



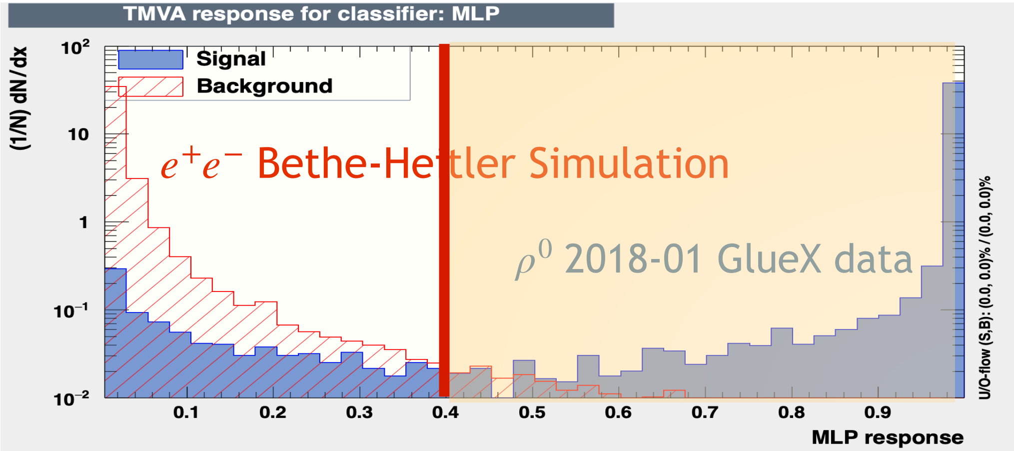

## 1. 'Less than' sweep: select e, reject π

$\pi^-$ and $e^-$ (BH sim/$\rho^0$ training) NN response, log y

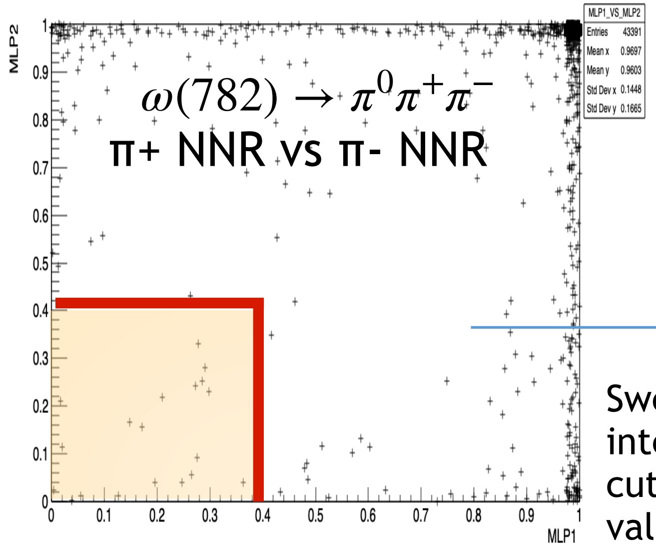



## 2. 'Greater than' sweep: select π, reject e

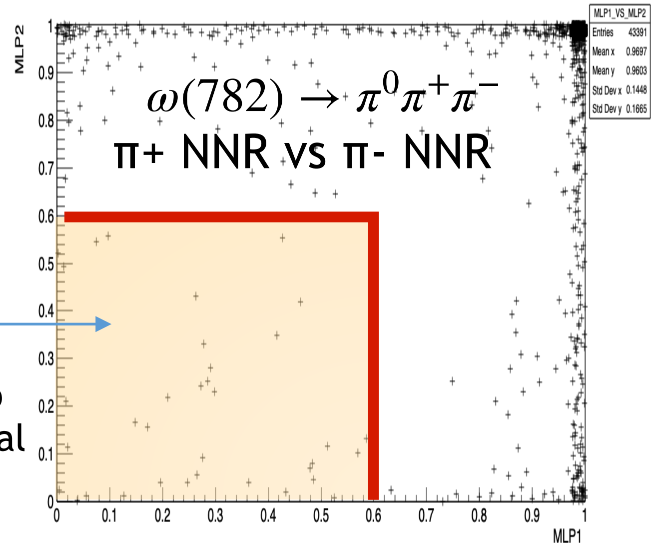
$\pi^-$ and $e^-$ (BH sim/$\rho^0$ training) NN response, log y

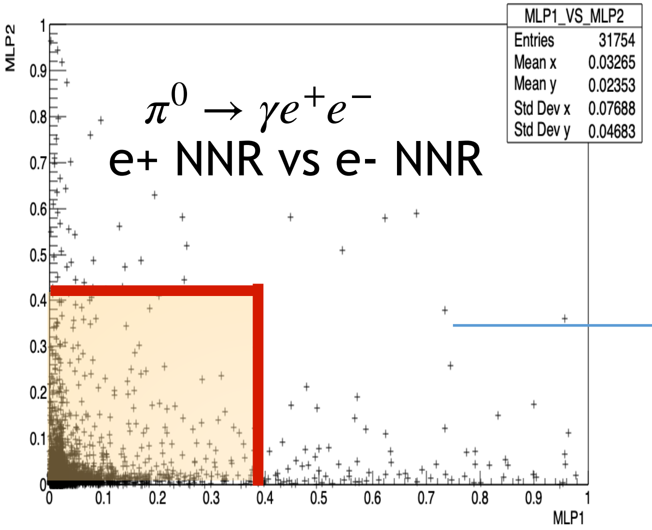


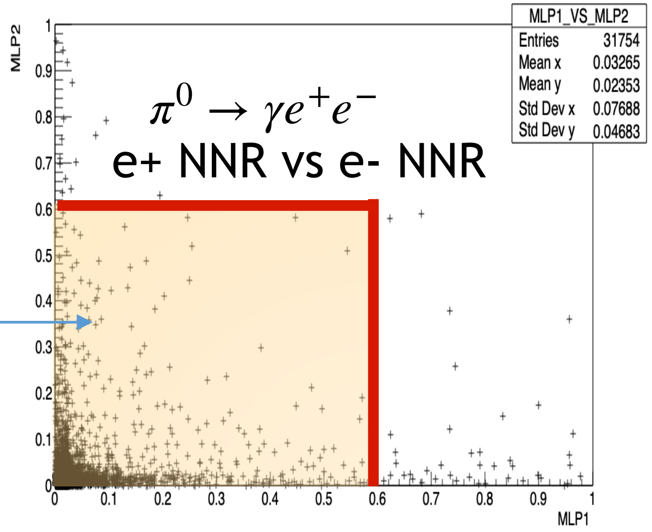
# 1. 'Less than' sweep: ***select e, reject π***



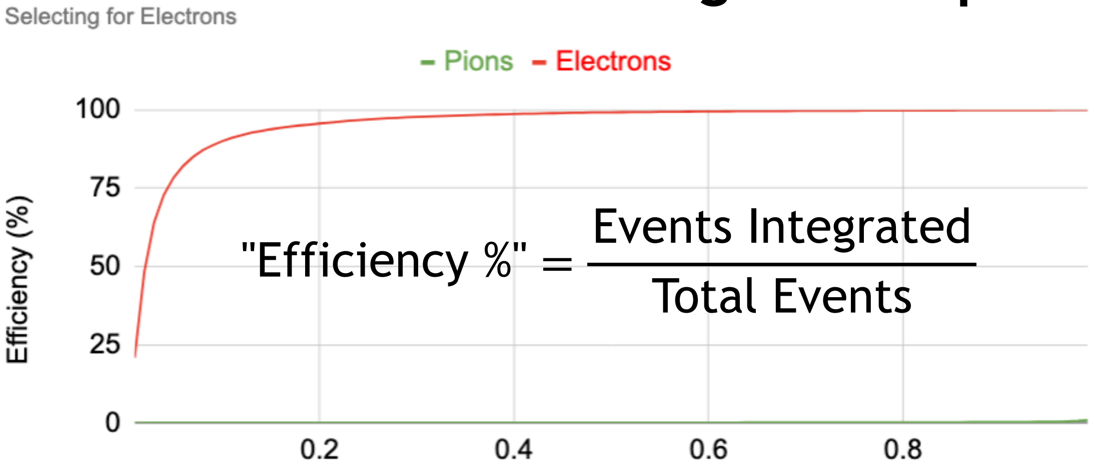

Sweep integral cutoff valus

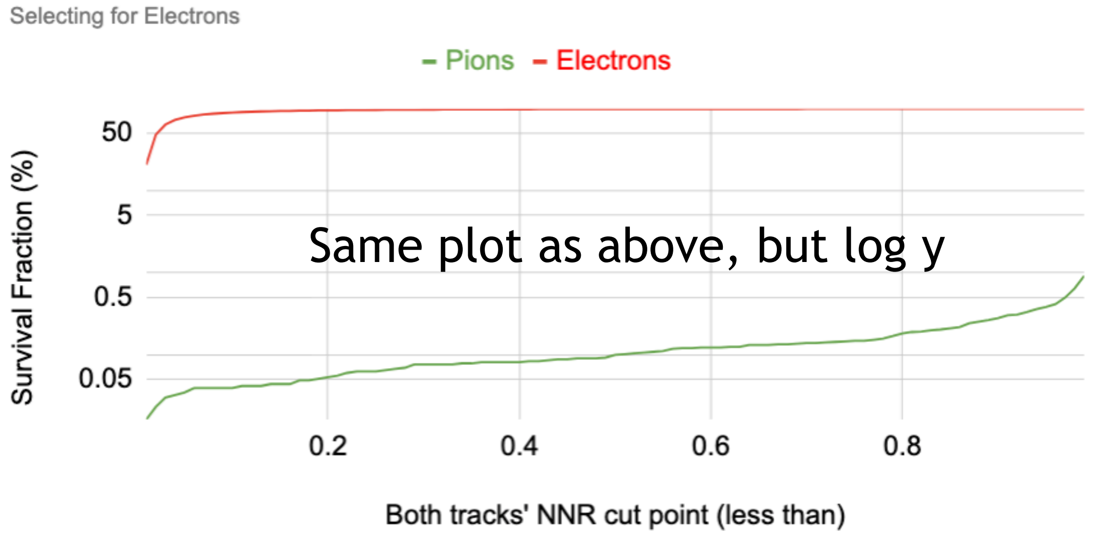


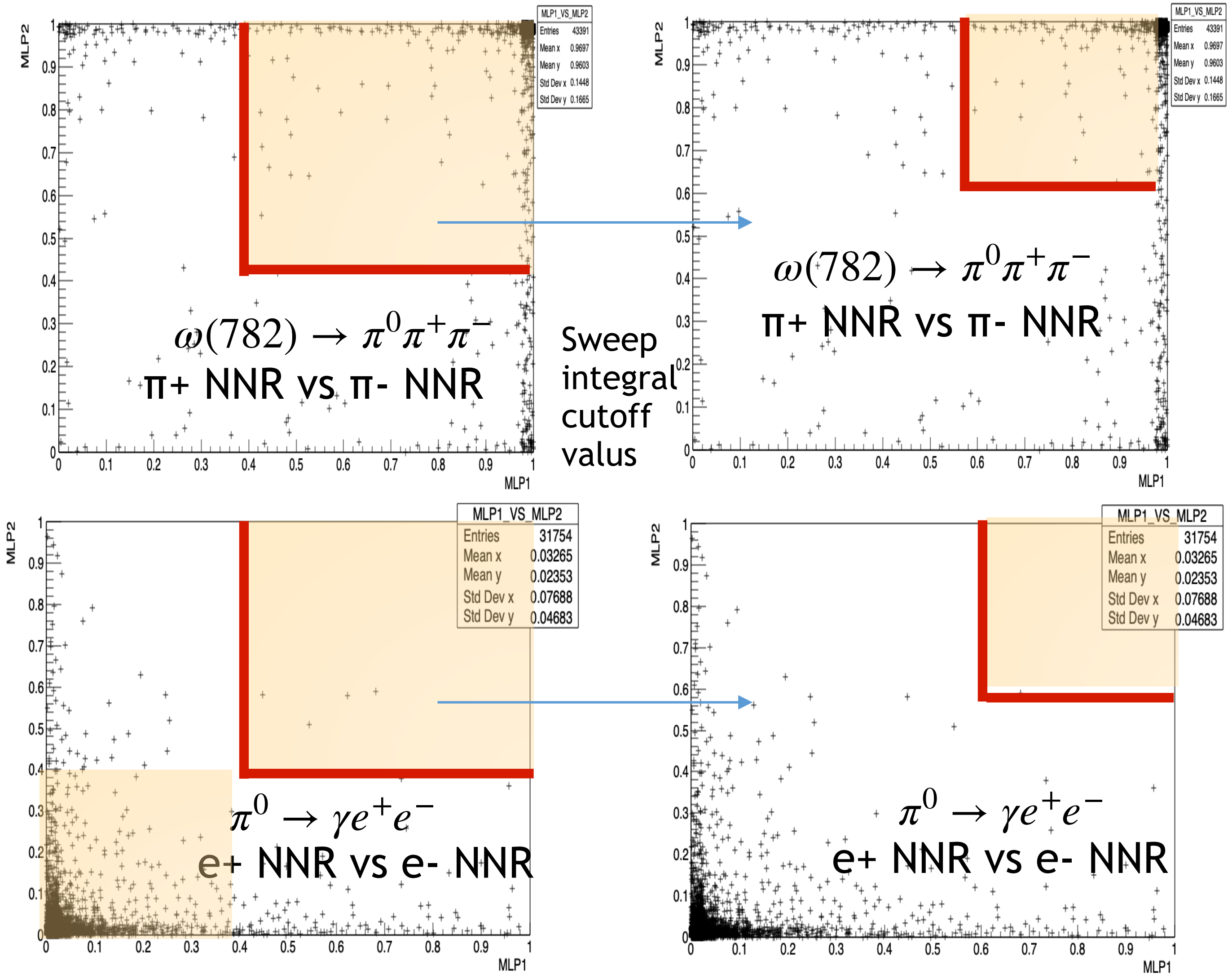




## Results from integral sweep

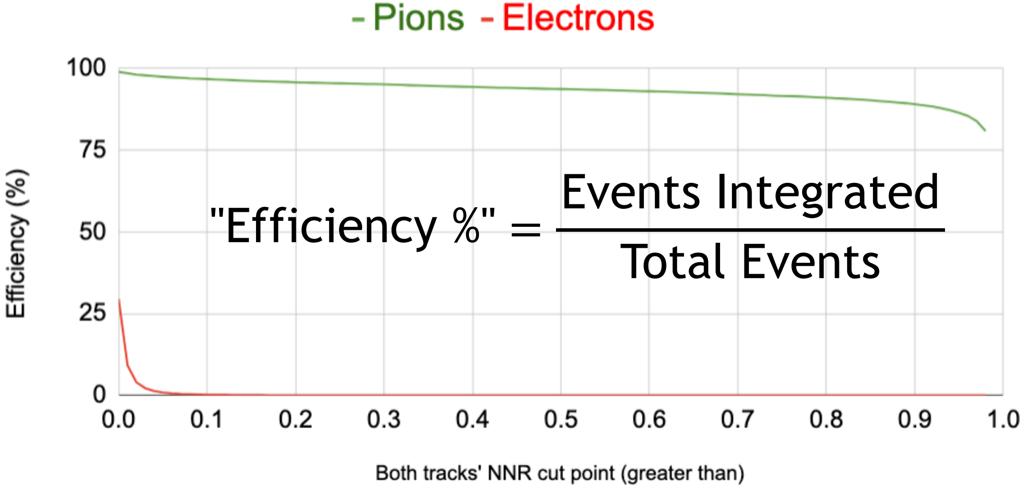


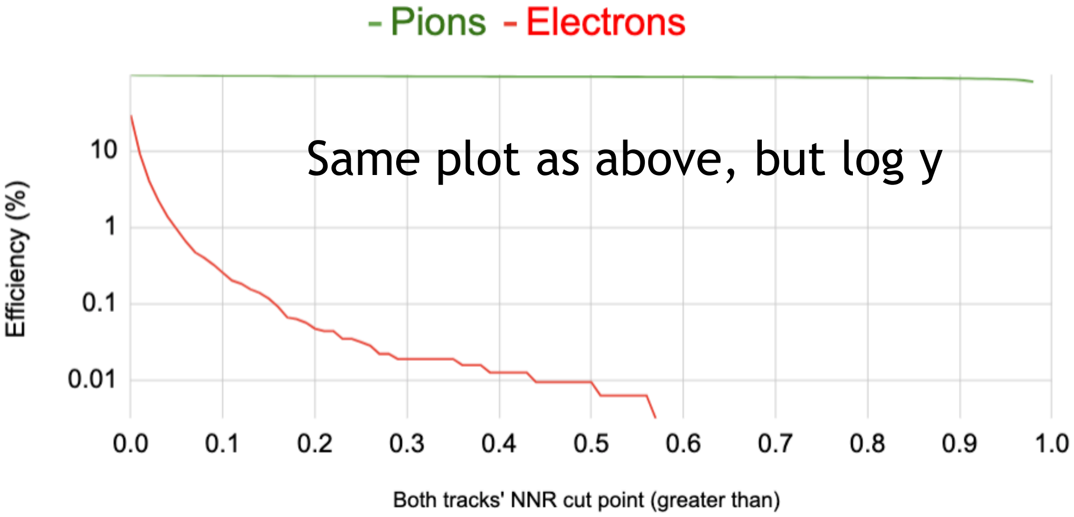



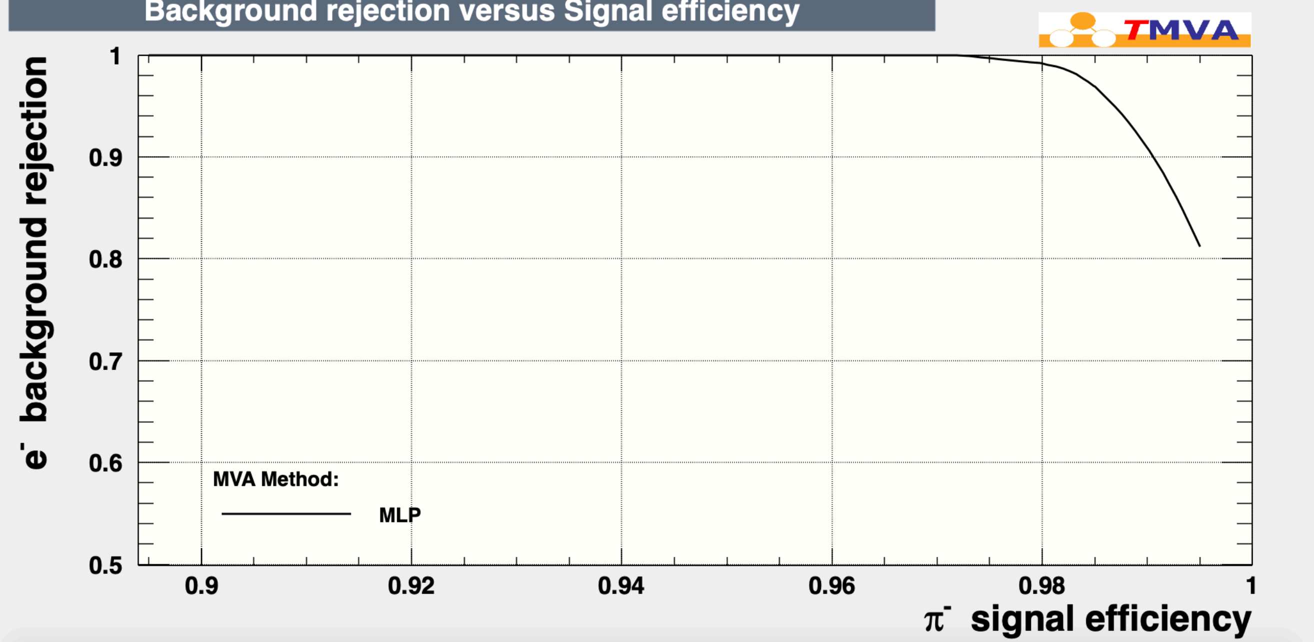
# 2. ‘Greater than’ sweep: ***select π, reject e***

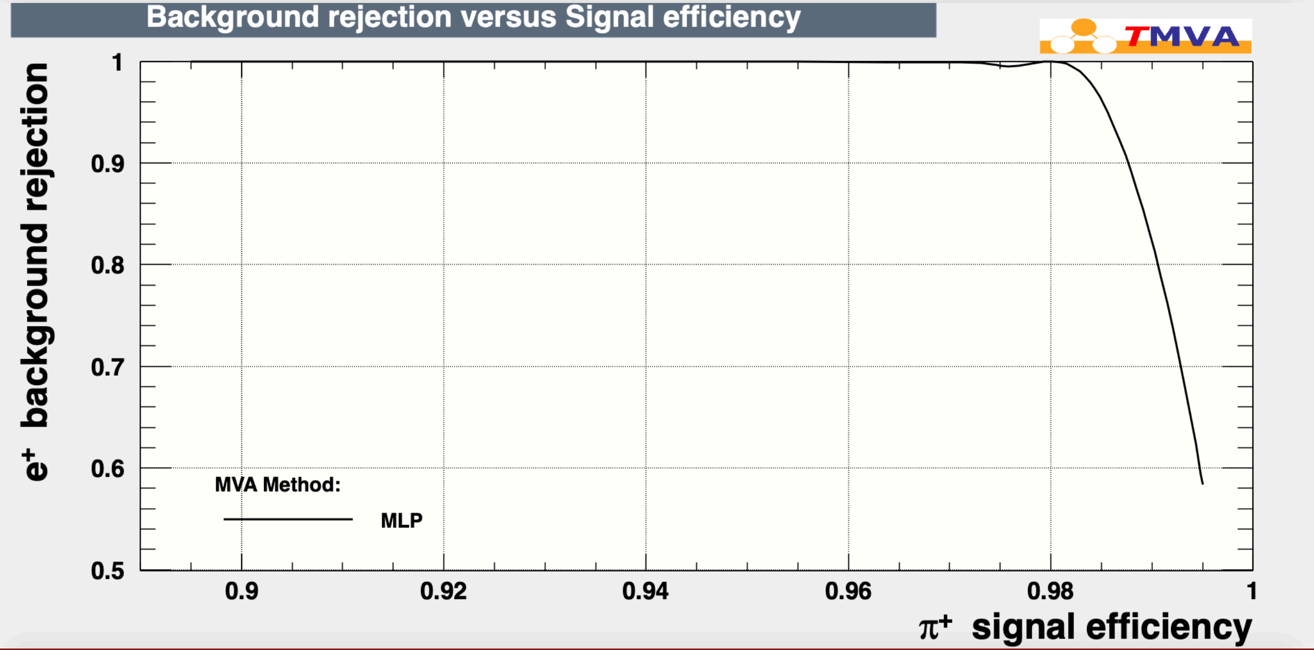


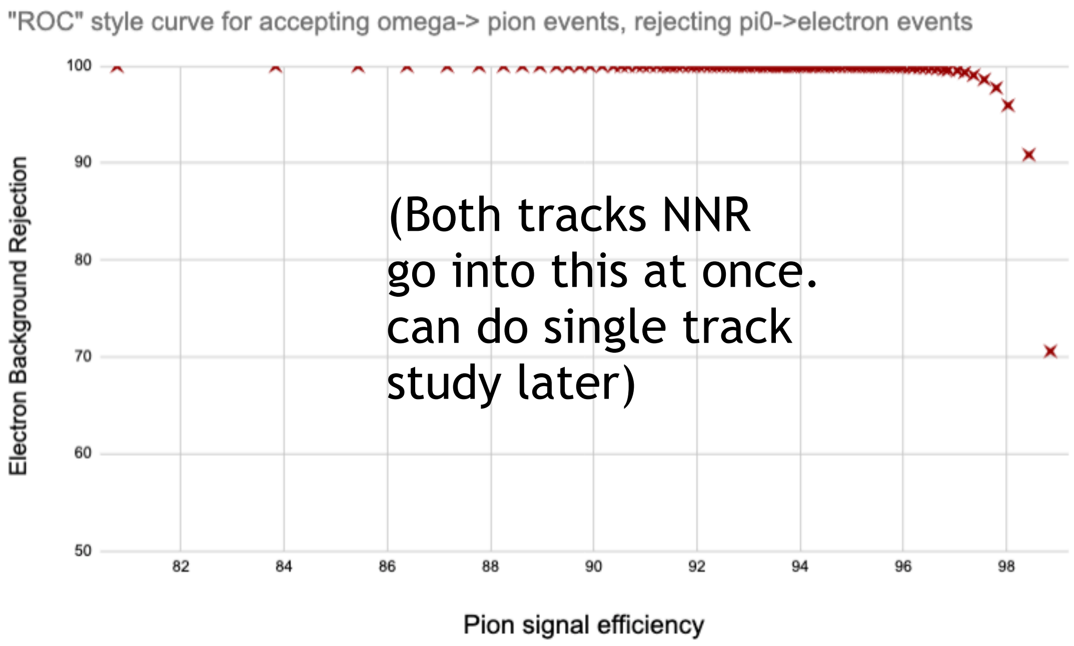

## Results from integral sweep





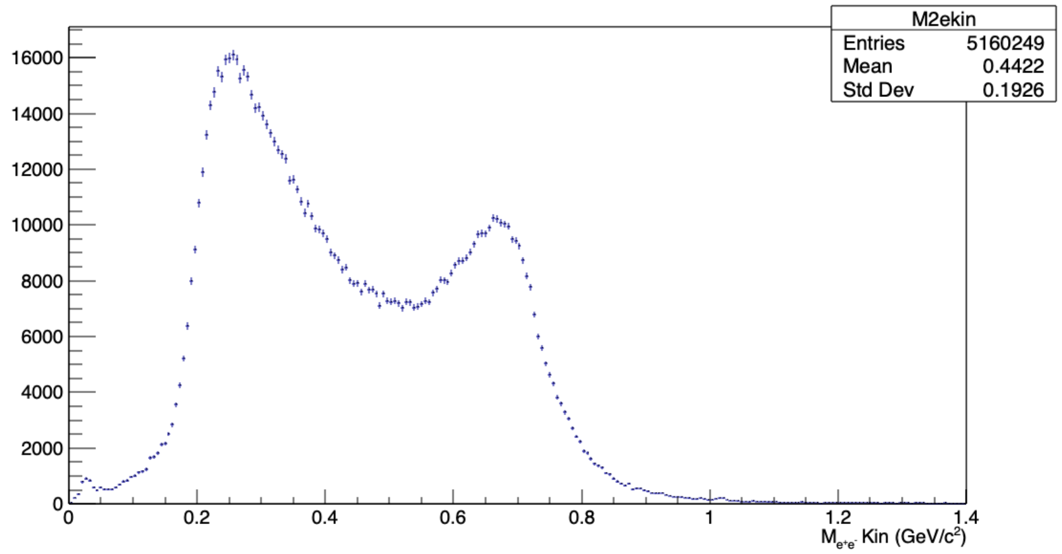



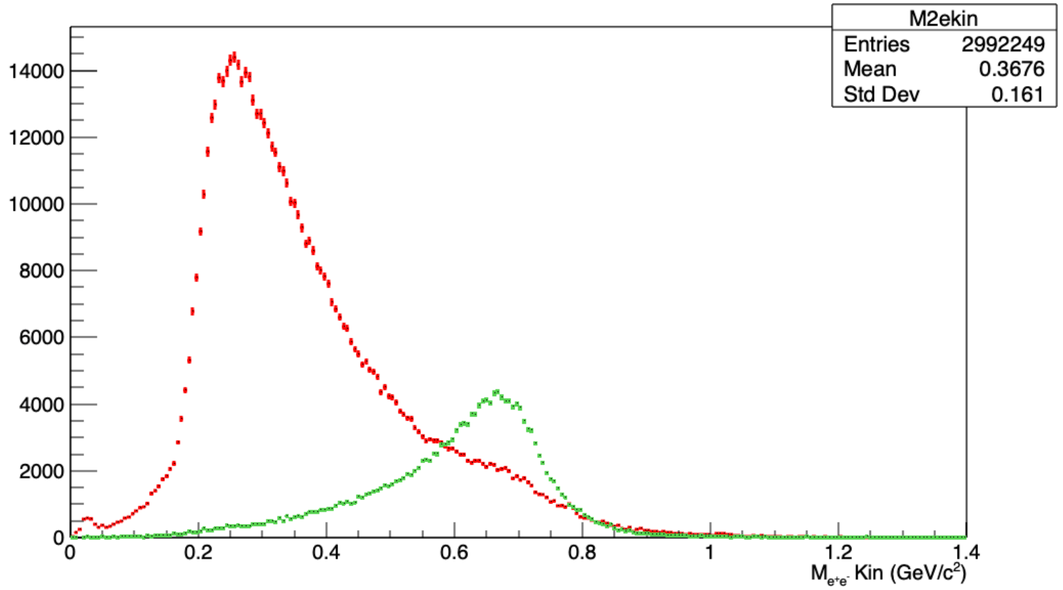

(Both tracks NNR go into this at once. can do single track study later)

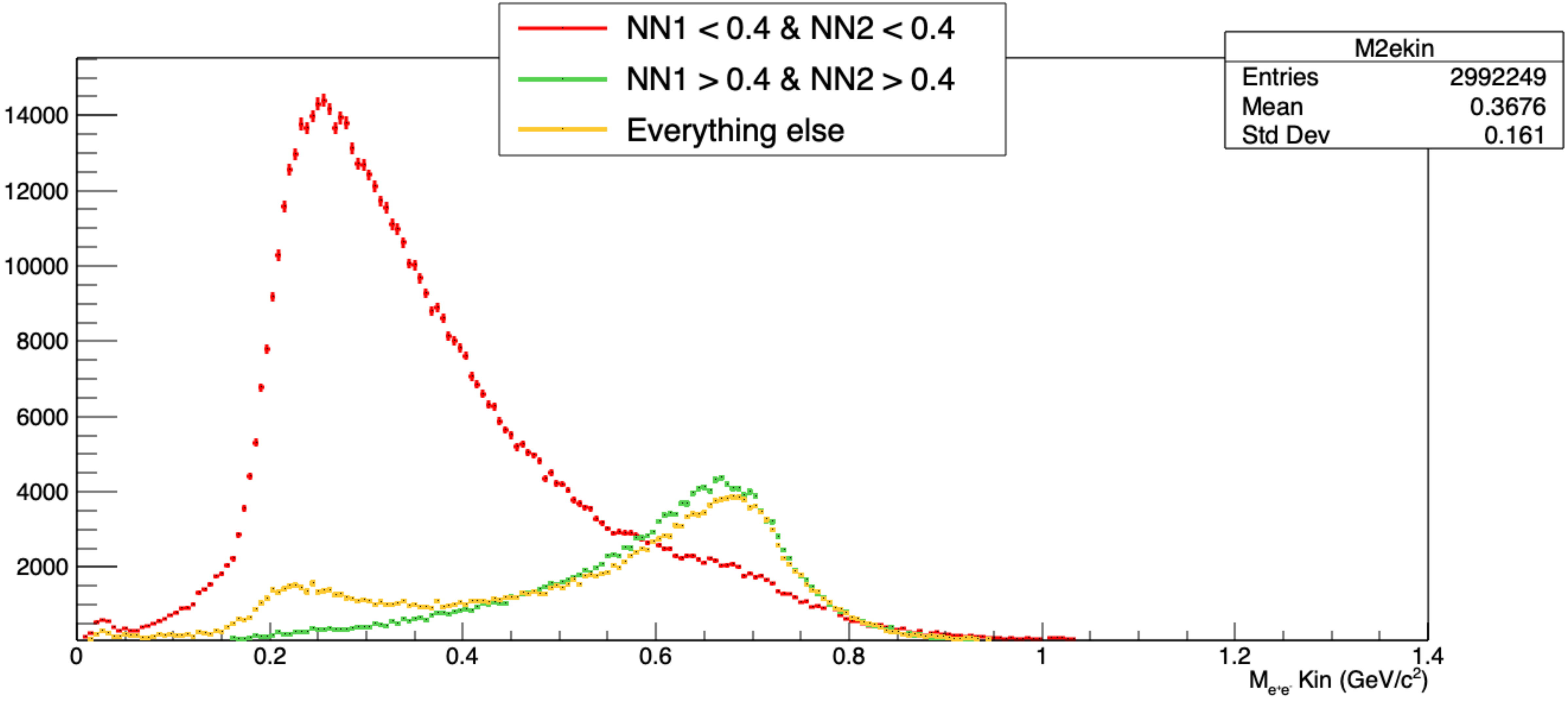
Differences: Training ROC curves can get 100% background rejection efficiency at 97% signal efficiency,
whereas the double track data classification requires you to go down to 93%

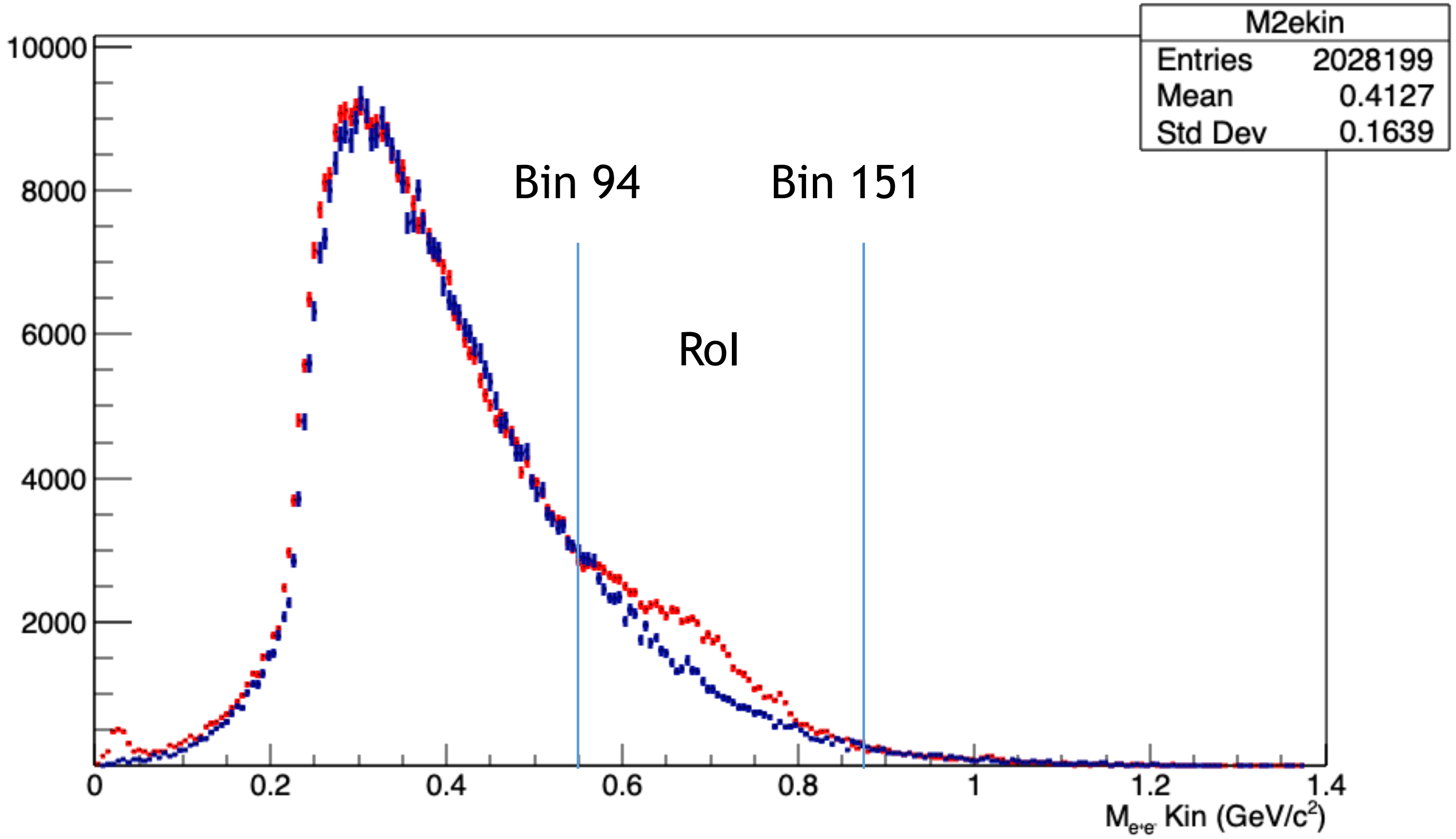
# Benchmark Study 2

Classify GlueX decay containing both $\rho^0$ pions and BH electrons.


M2ekin
Entries 5160249
Mean 0.4422
Std Dev 0.1926
16000
14000
12000
10000
8000
6000
4000
2000
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
$M_{e^+e^-}$ Kin (GeV/$c^2$)


M2ekin
Entries 2992249
Mean 0.3676
Std Dev 0.161
14000
12000
10000
8000
6000
4000
2000
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
$M_{e^+e^-}$ Kin (GeV/$c^2$)

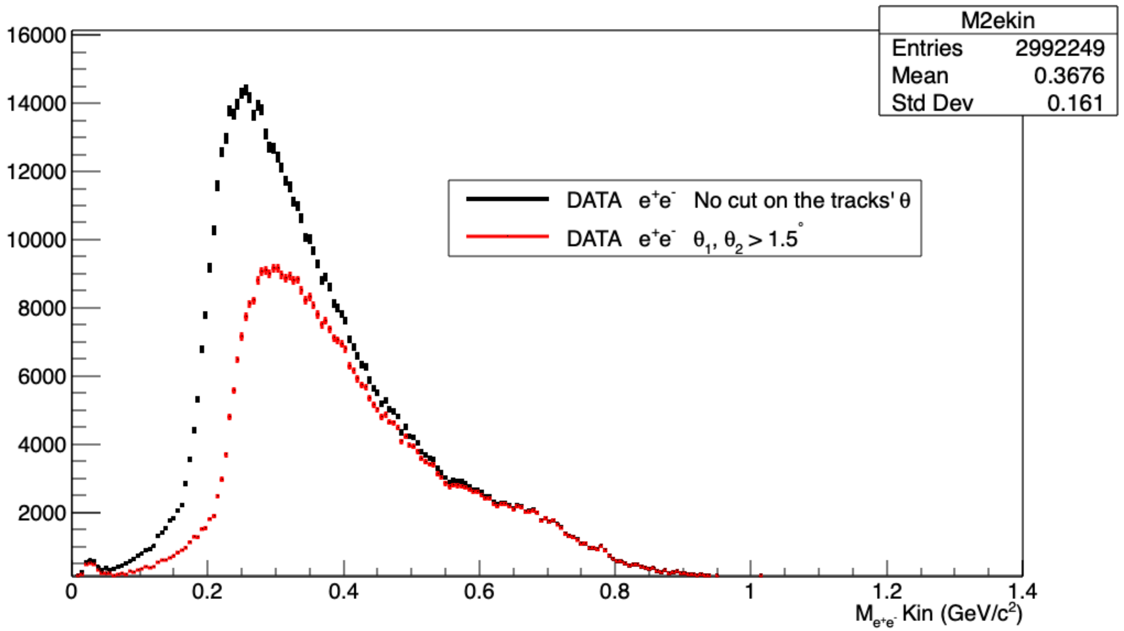

NN1 < 0.4 & NN2 < 0.4
NN1 > 0.4 & NN2 > 0.4
Everything else
M2ekin
Entries 2992249
Mean 0.3676
Std Dev 0.161
14000
12000
10000
8000
6000
4000
2000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
$M_{e^+e^-}$ Kin (GeV/c$^2$)




$$\frac{\int_{94}^{151} W_{data} dn}{\int_{94}^{151} W_{MC} dn} = \frac{86855.500}{68714.912} = 1.263997$$

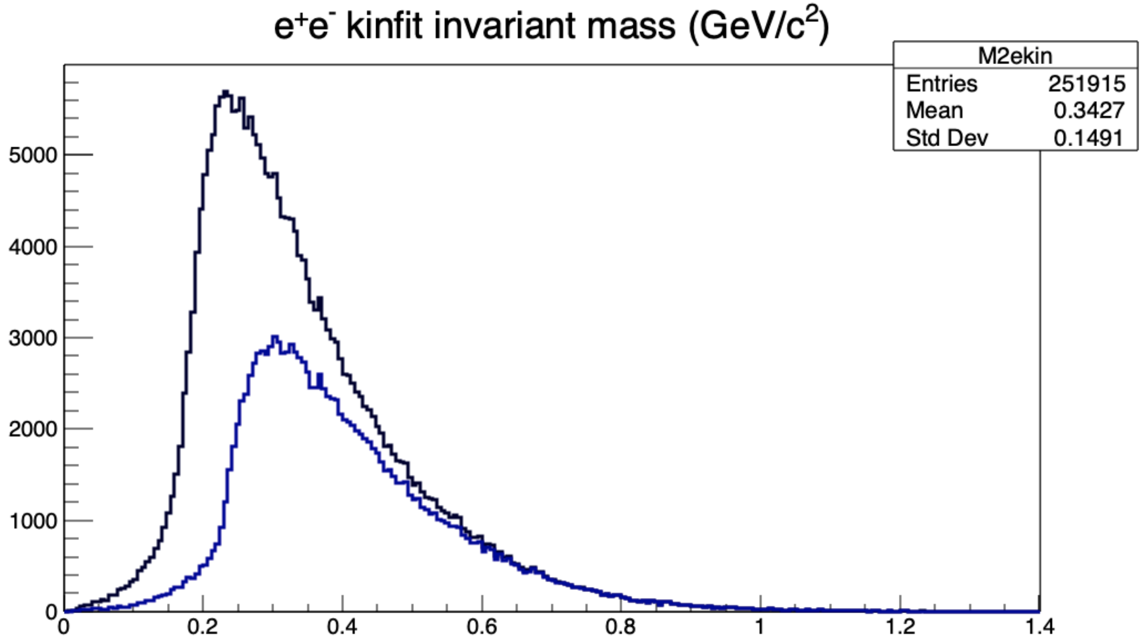
# Comparing theta cut ON with theta cut OFF in DATA and MC

## DATA

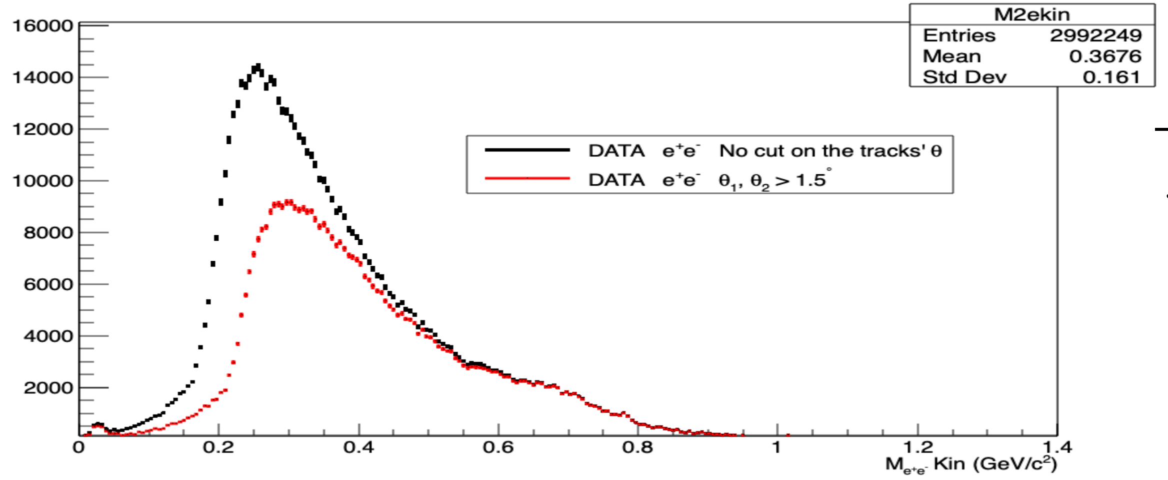



$$\frac{\int_{94}^{151} W_{\text{data}}\, dN}{\int_{94}^{151} W_{\text{data},\ \theta>1.5}\, dN} = 1.0215933$$

## MC

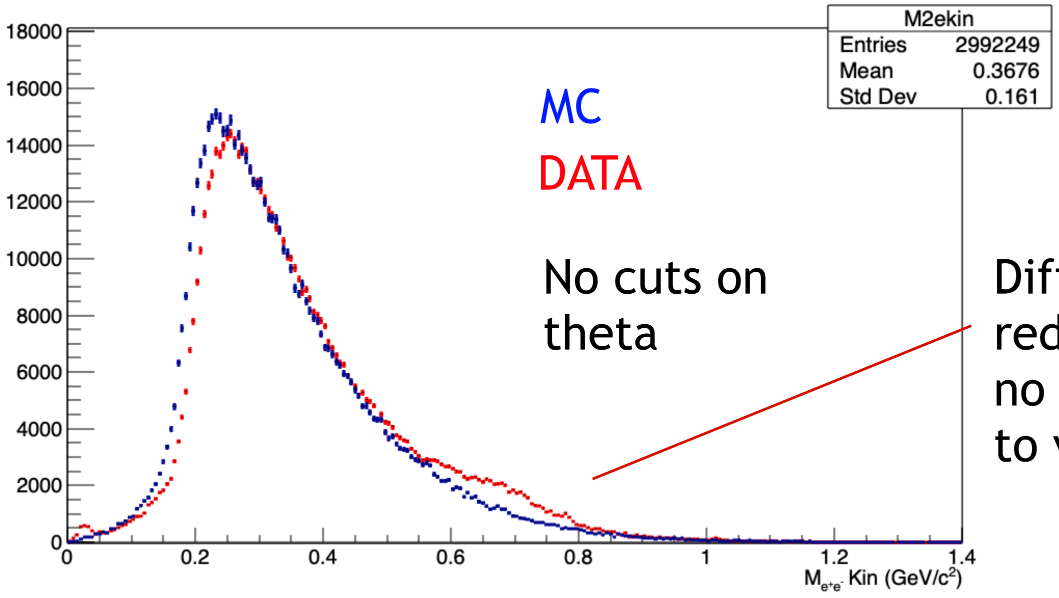



$$\frac{\int_{94}^{151} W_{\text{MC}}\, dN}{\int_{94}^{151} W_{\text{MC},\ \theta>1.5}\, dN} = 1.0523269$$

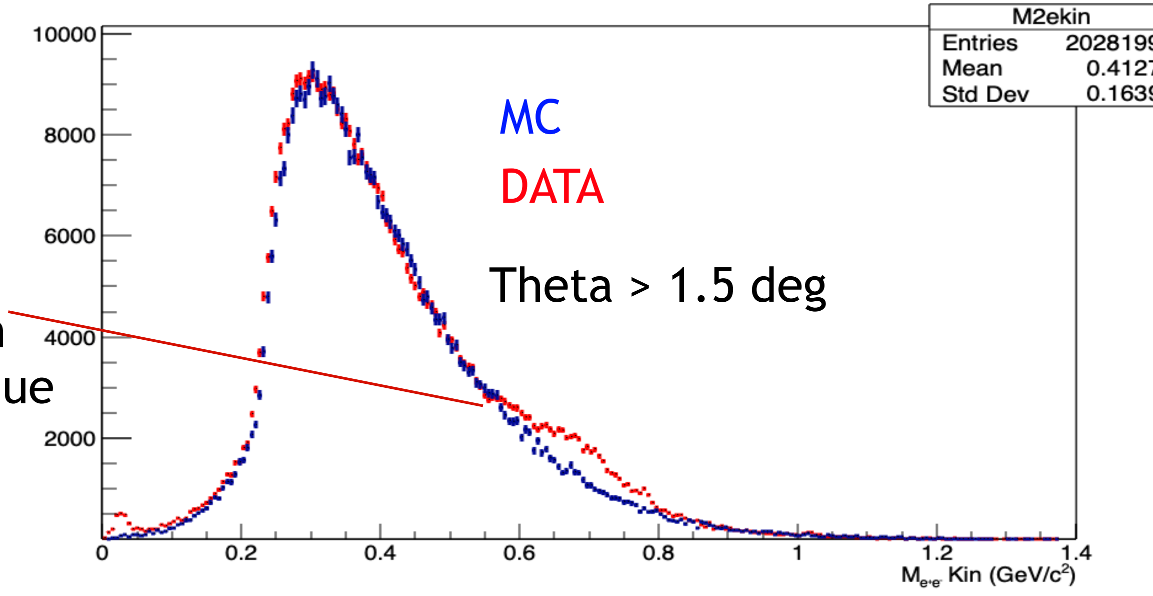


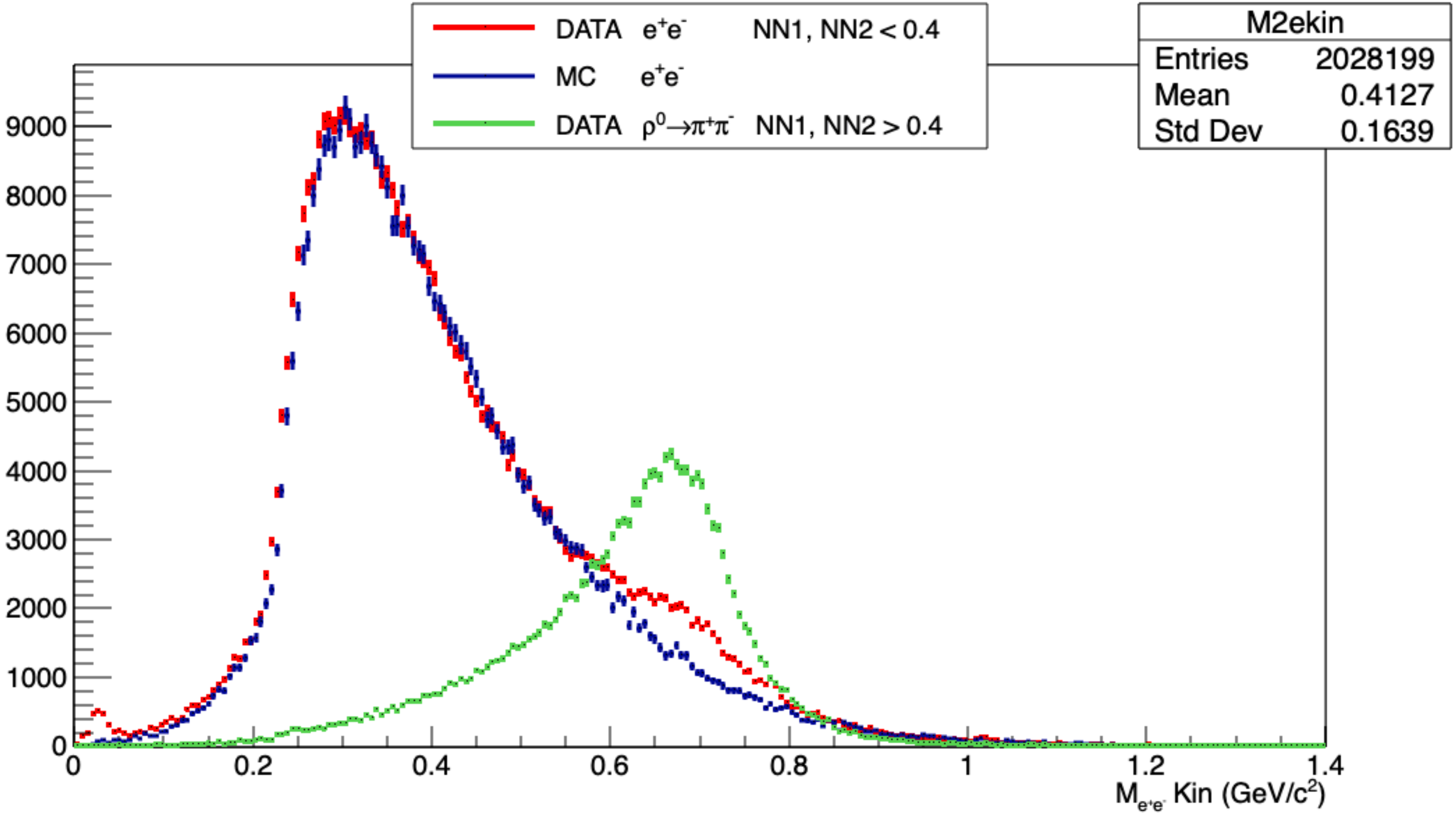

$$\frac{\int_{94}^{151} W_{\text{data}}\, dN}{\int_{94}^{151} W_{\text{data},\ \theta>1.5}\, dN} = 1.0215933$$

2% difference in RoI

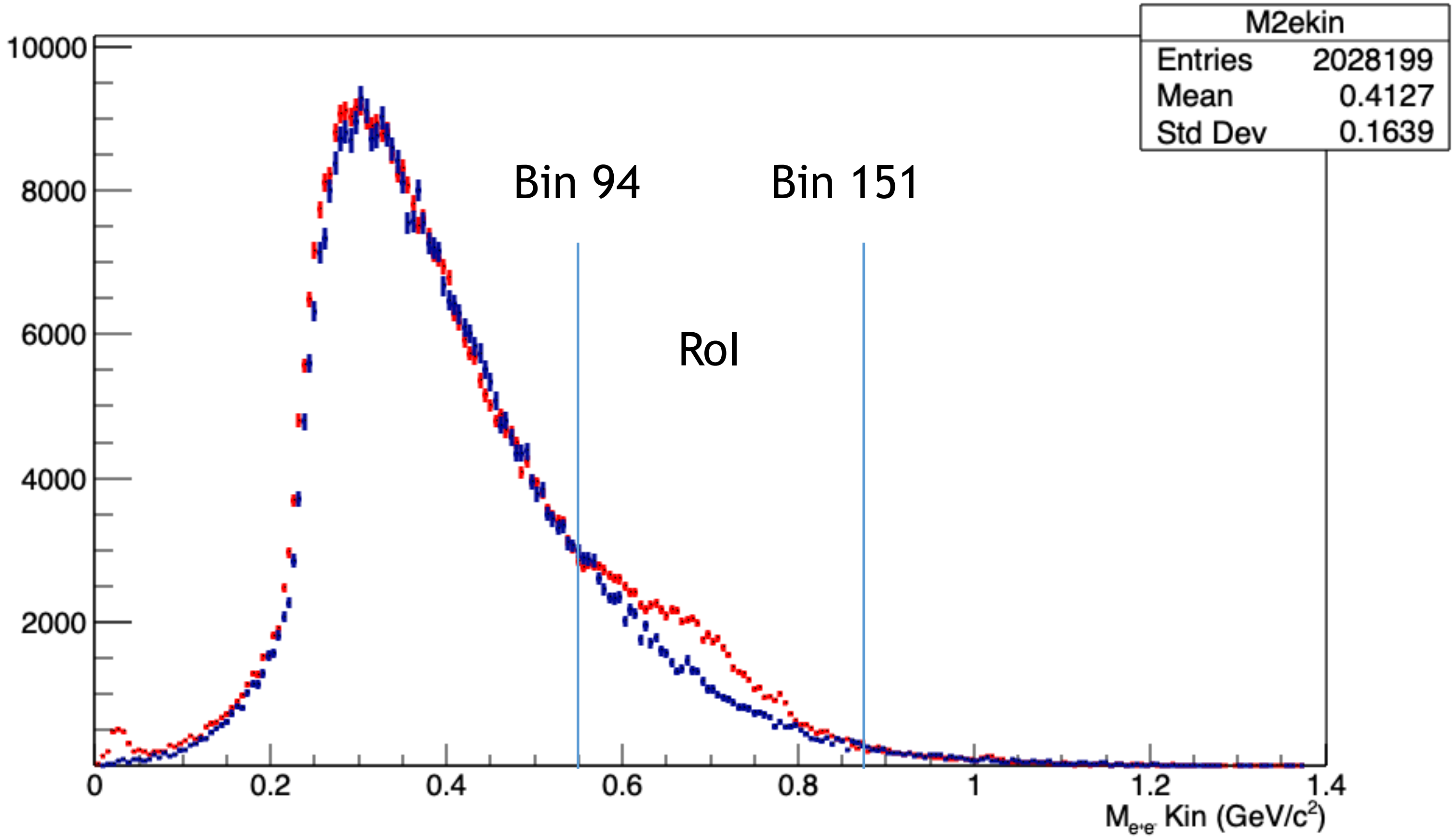



Difference in red curve between no cut/theta cut due to vertical scale

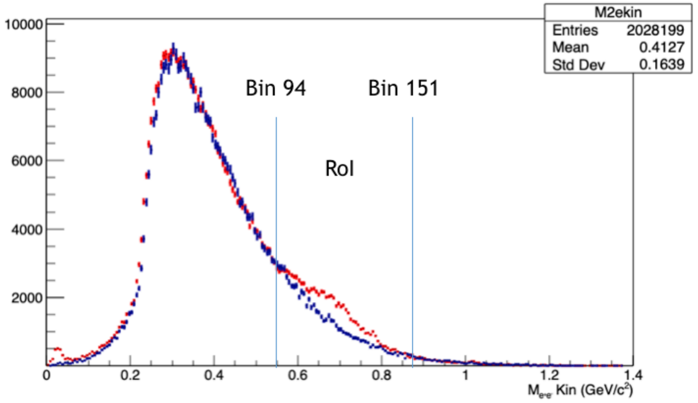




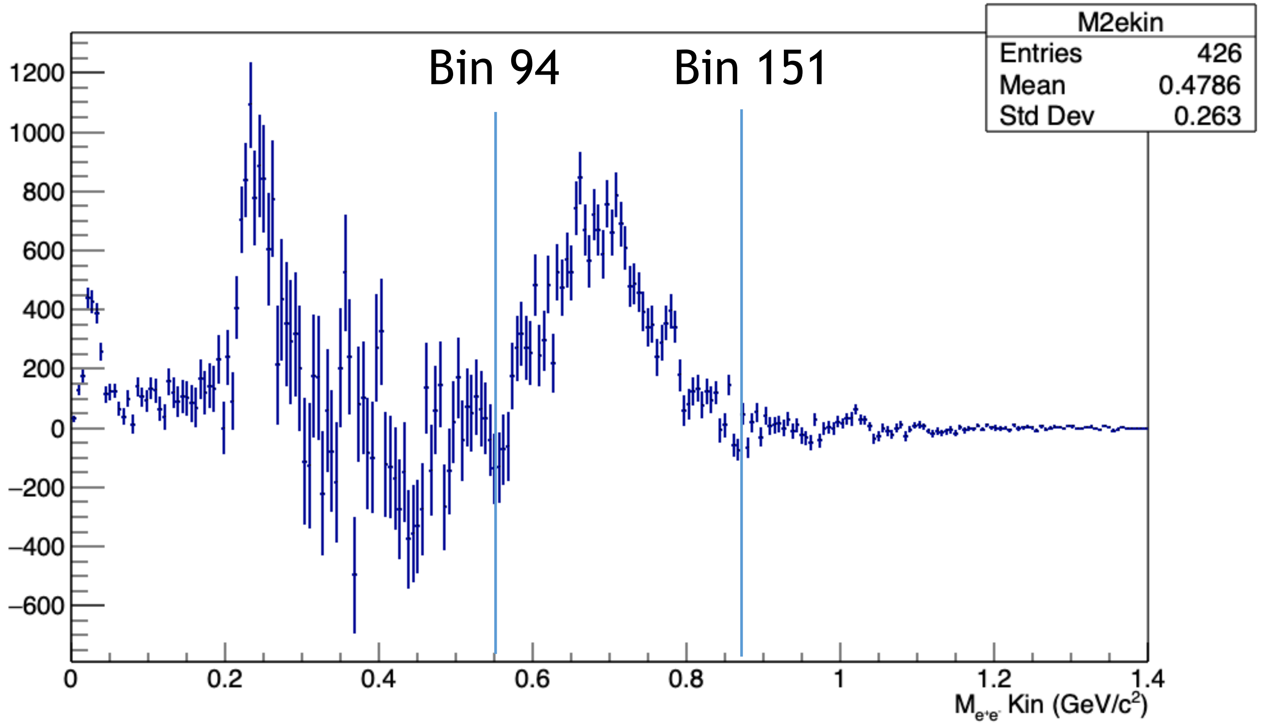
DATA e+e- NN1, NN2 < 0.4
MC e+e-
DATA ρ0→π+π- NN1, NN2 > 0.4
M2ekin
Entries 2028199
Mean 0.4127
Std Dev 0.1639
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Me+e- Kin (GeV/c2)



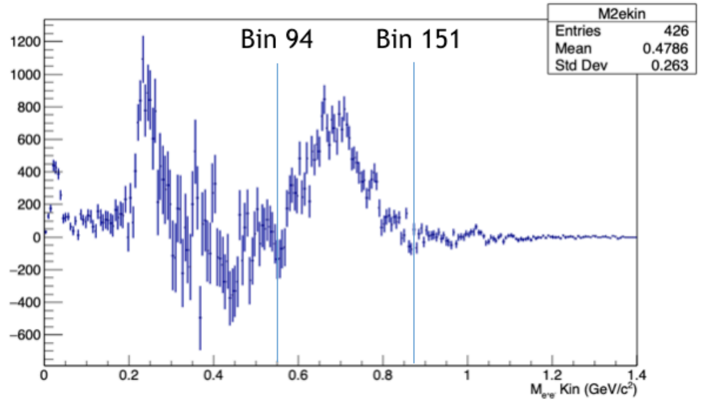

$$\frac{\int_{94}^{151} W_{data} dn}{\int_{94}^{151} W_{MC} dn} = \frac{86855.500}{68714.912} = 1.263997$$



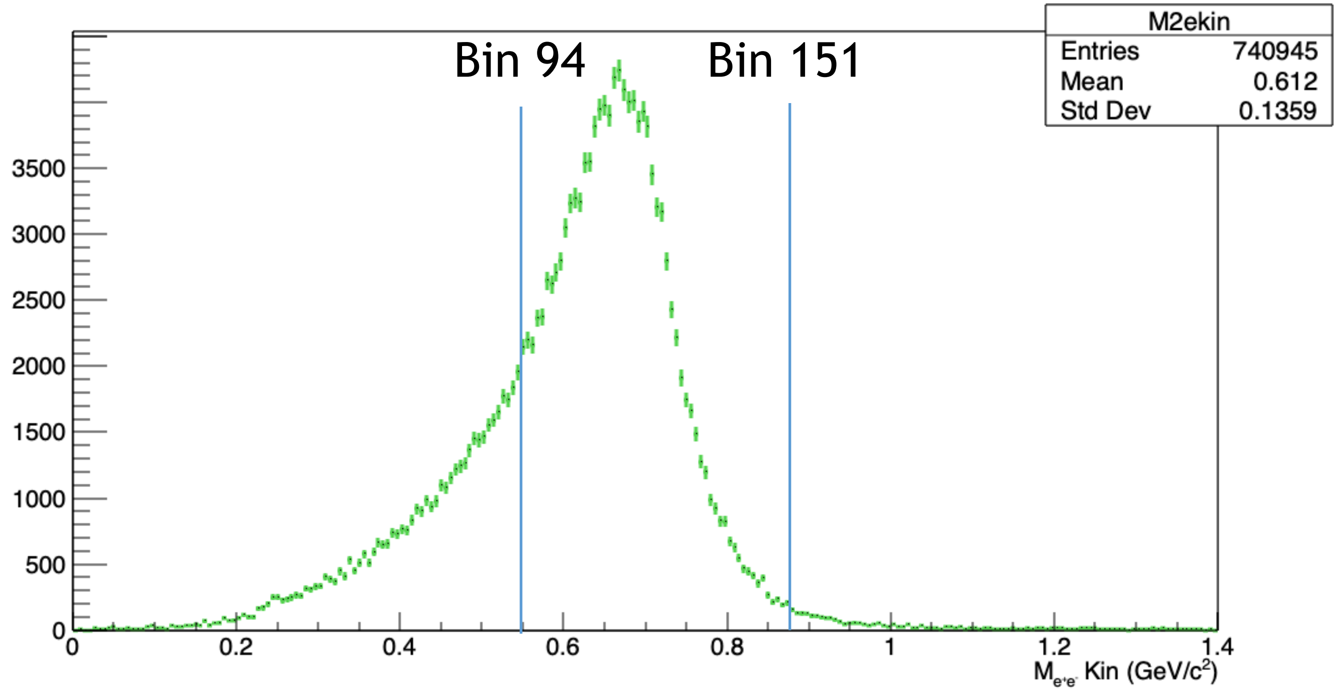

Subtract the MC distribution from the data and plot the difference



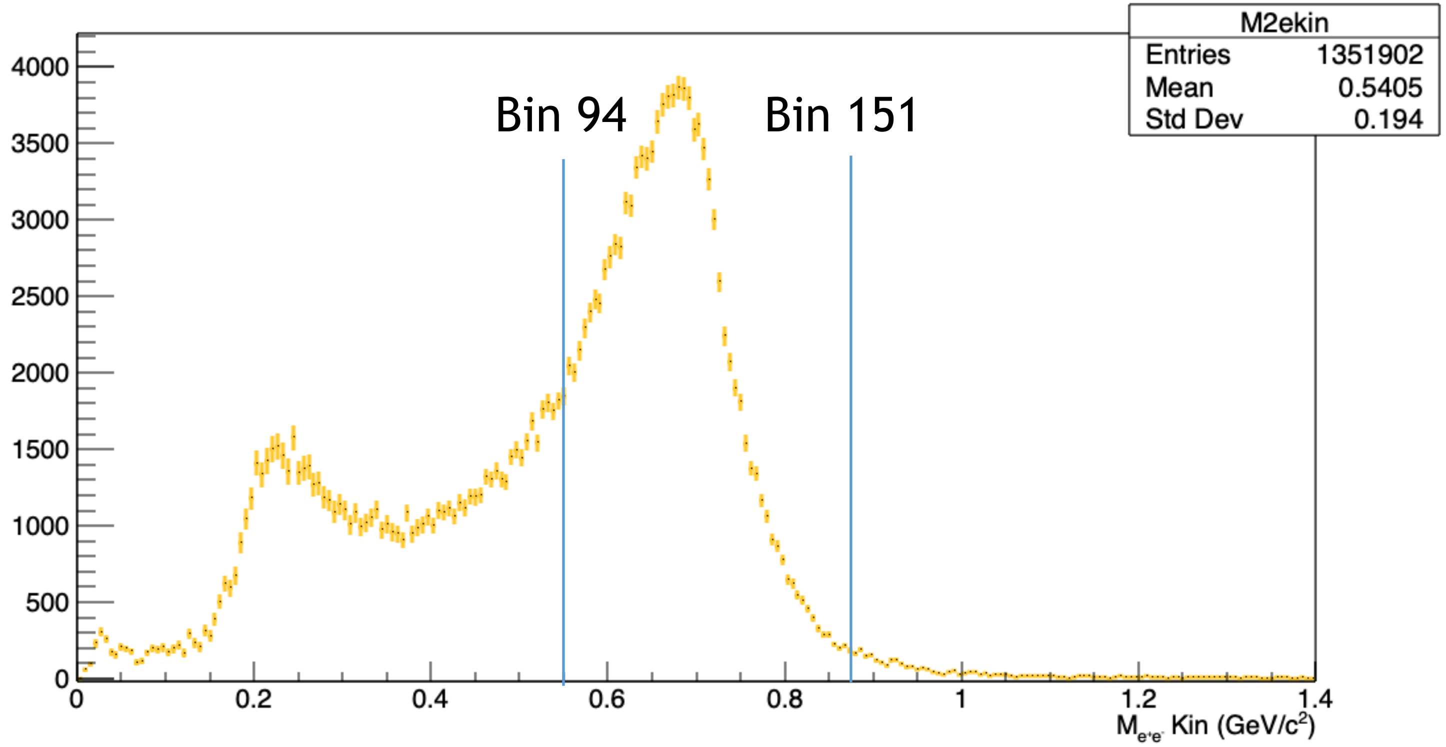

$$\int_{94}^{151} (W_{data} - W_{MC})dn = 18140.588$$



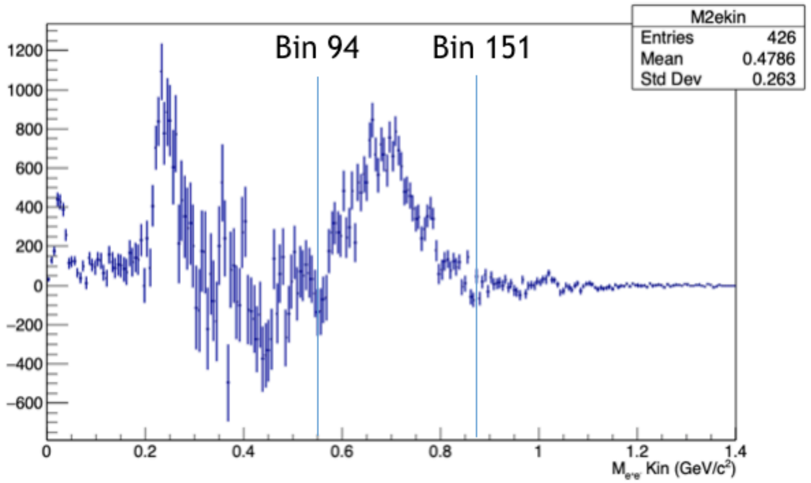

$$\int_{94}^{151} (W_{data} - W_{MC})dn = 18140.588$$

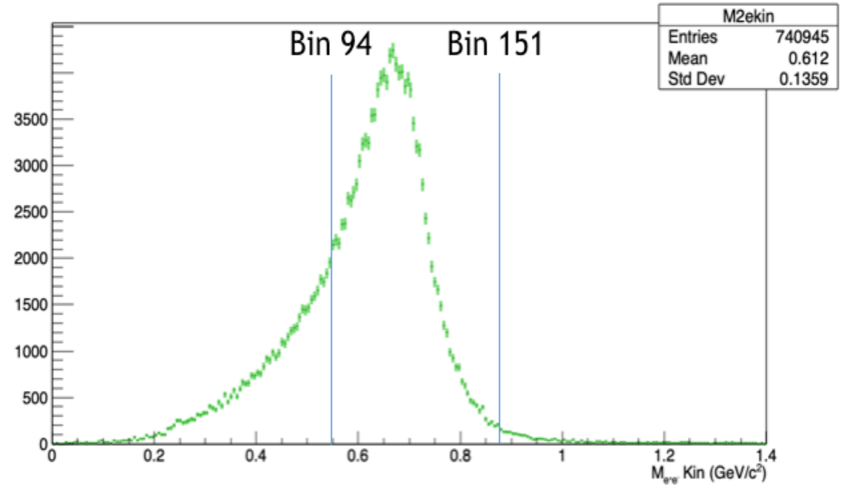


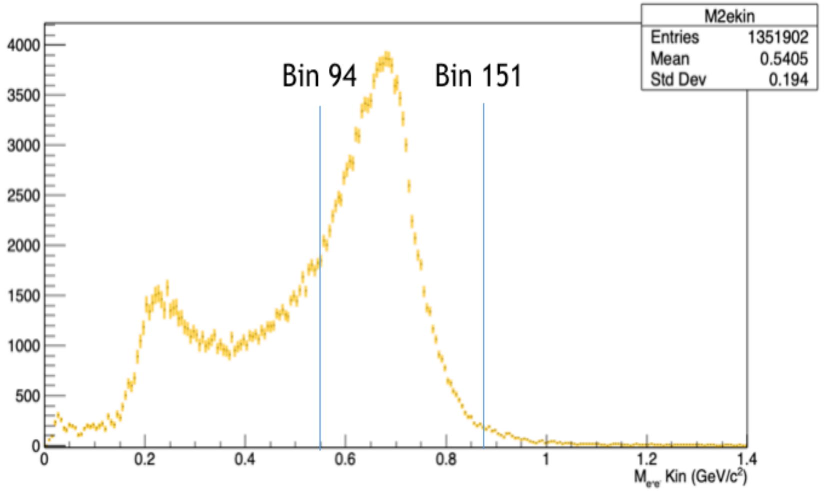

$$\int_{94}^{151} W_{\rho} dn = 125071.50$$




$$\int_{94}^{151} W_{\rho \text{rejected}} dn = 116950.00$$




$$\int_{94}^{151} (W_{data} - W_{MC})dn = 18140.588$$

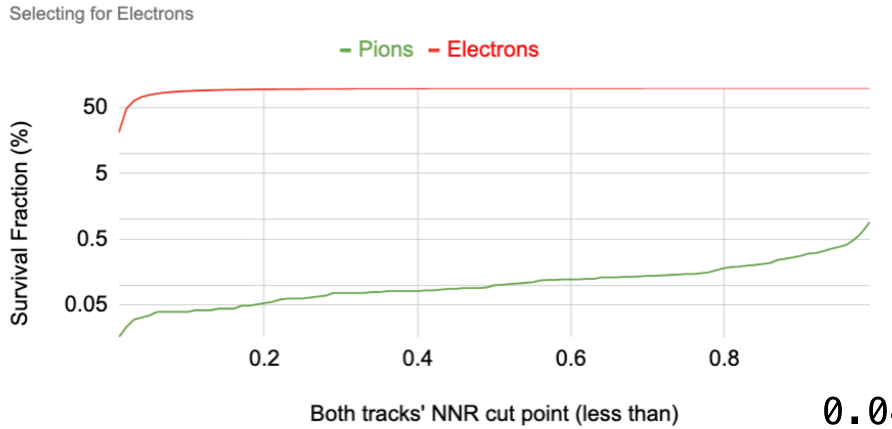
$$\int_{94}^{151} W_{\rho} dn = 125071.50$$

$$\int_{94}^{151} W_{\rho\text{rejected}} dn = 116950.00$$

Assume difference is purely pion contamination from Rho0. Does the assumption hold up to previous studies?

$$\frac{\int W_{diff}}{\int (W_{diff} + W_{\rho} + W_{rej})} = \frac{18140.588}{18140.588 + 125071.50 + 116950.00} = 0.069728023$$

6.9% survival fraction under this assumption.
Let's compare to omega/pi0 study:




0.046214761

| Cut point | Pion % | Electron % |
|---|---|---|
| 0.4 | 0.08066188841 | 98.68677962 |