RICH particle identification using Machine Learning

Armen Gyurjinyan
INFN - Laboratori Nazionali di Frascati

Hall B collaboration meeting, November 7-10,2023

Outline

- Motivation
  - RICH Alignment
- Data preprocessing for machine learning
  - Data filtering
  - Physics cuts to select subsample
- Machine learning model
  - Model and data
  - Input/Output features
- Preliminary results
  - Comparison between CLAS12 Event Builder, RICH pass2, RICH NN

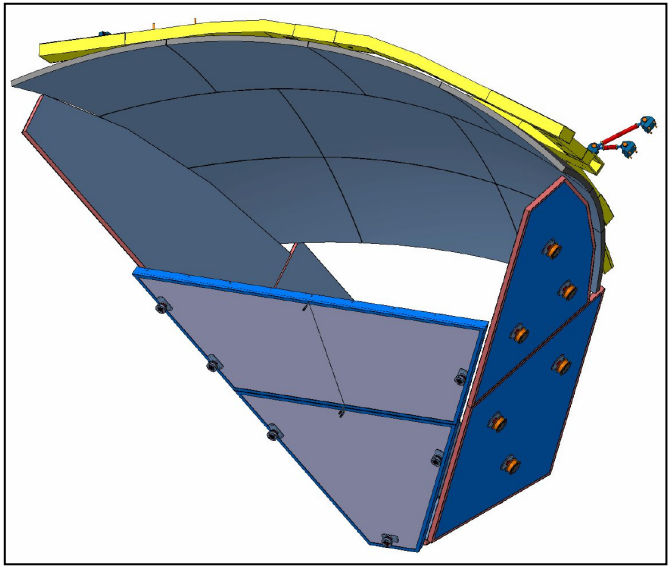
RICH alignment

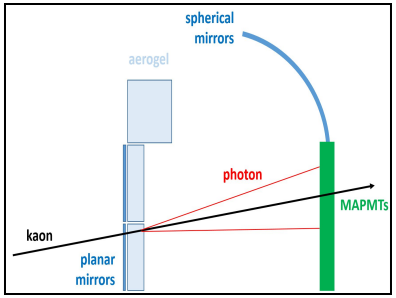
RICH composition
- 3 aerogel planes
- 7 planar mirrors
- 10 spherical mirrors
- MAPMT

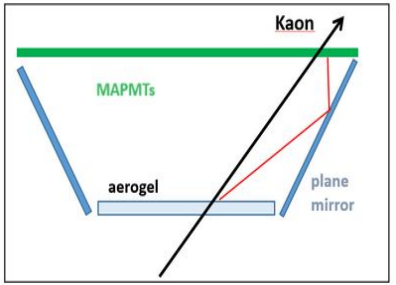
6 alignment parameter (x,y,z, Θ_x, Θ_y,Θ_z) per element

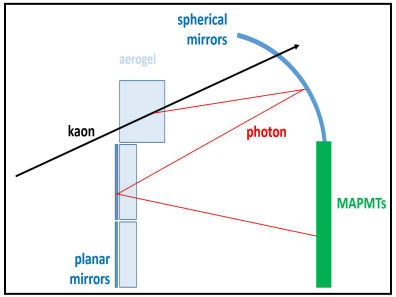
Total 126 parameters

However sensitive parameters are z, Θ_x, Θ_y









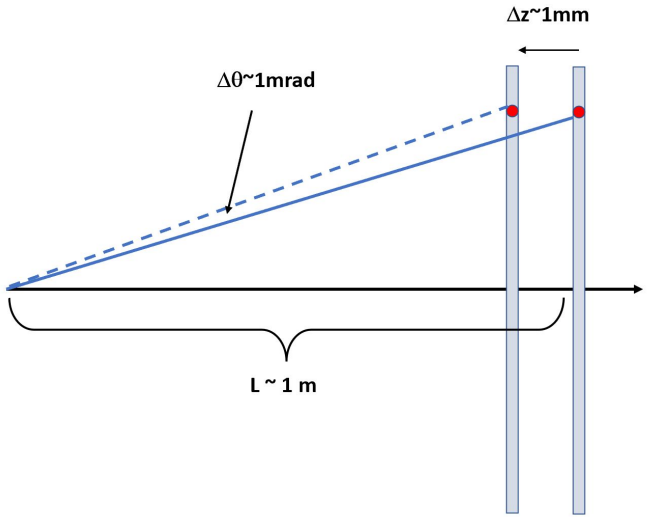
RICH alignment

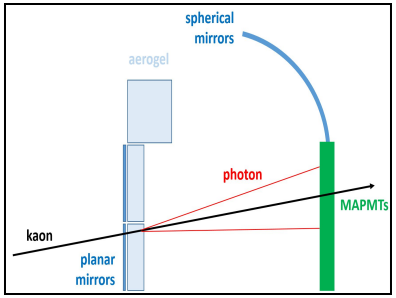
RICH composition
- 3 aerogel planes
- 7 planar mirrors
- 10 spherical mirrors
- MAPMT

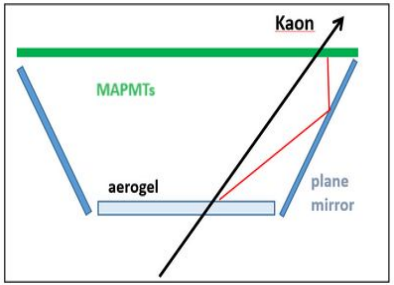
6 alignment parameter (x,y,z, Θ_x,Θ_y, Θ_z) per element

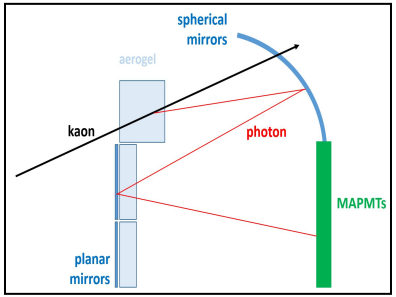
Total 126 parameters

However sensitive parameters are z, Θ_x, Θ_y

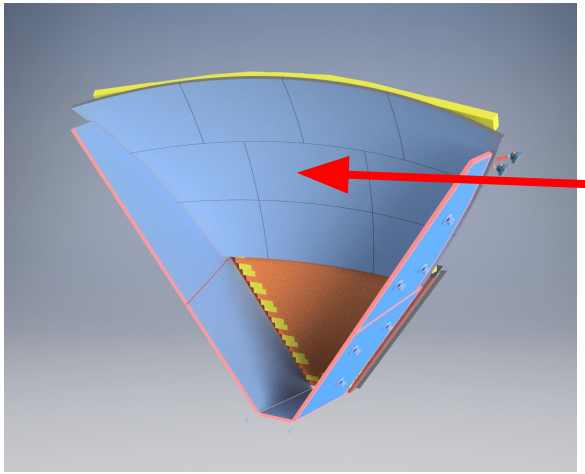






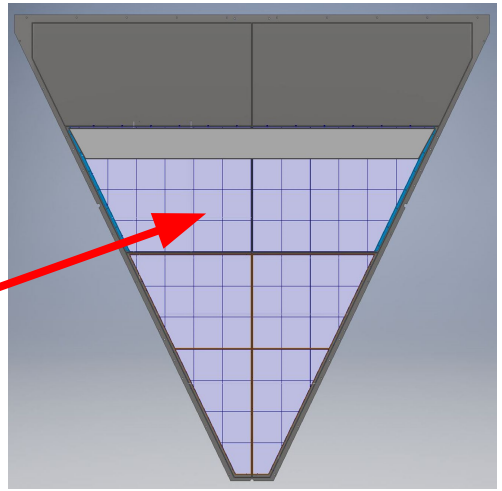


RICH alignment limitations

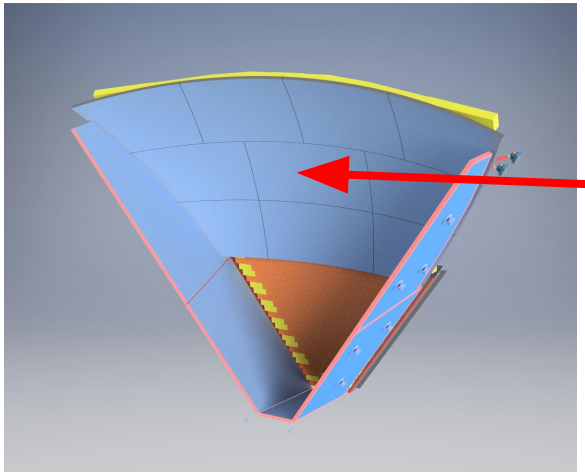


Spherical mirrors upper and middle rows are not aligned in pass2 cooking.

Aerogel Layer 2 not aligned in pass2 cooking

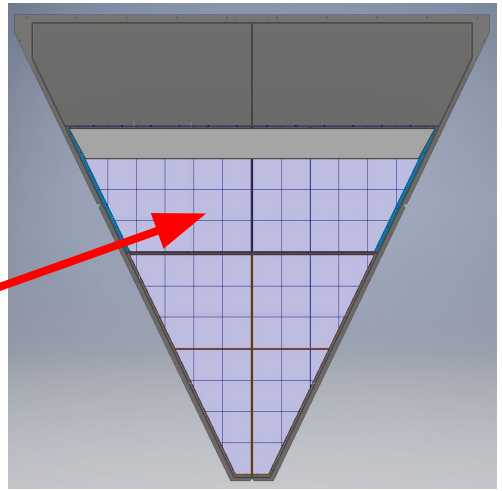


RICH alignment limitations



Spherical mirrors upper and middle rows are not aligned in pass2 cooking.

Aerogel Layer 2 not aligned in pass2 cooking



New Approach with Neural Networks to bypass alignment task!

Data filtering

ep -> eph+ (π-)

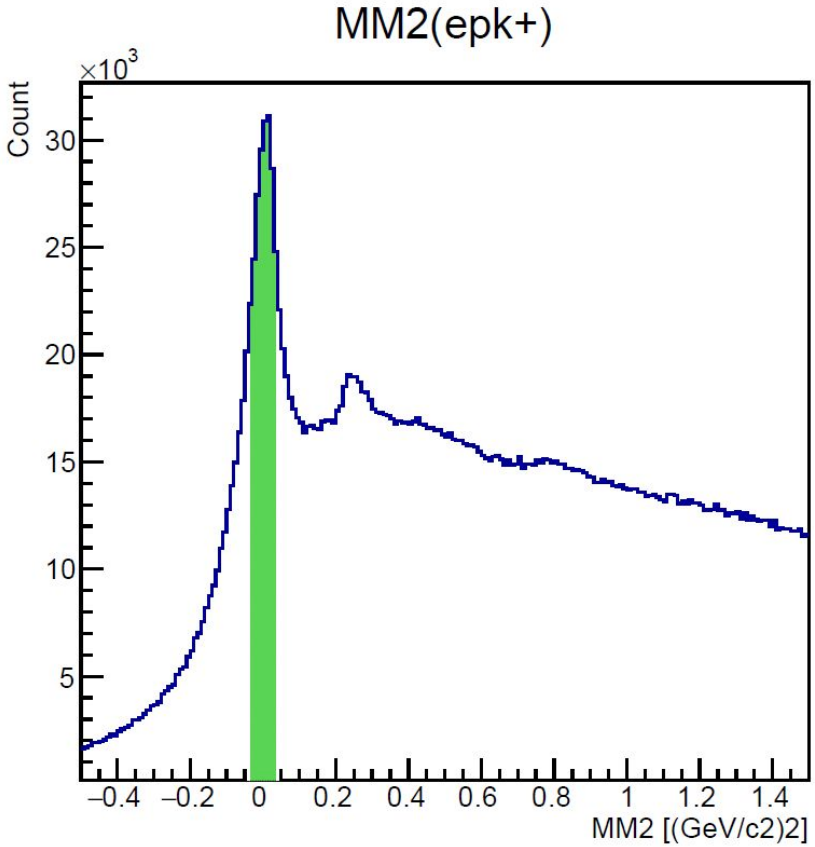
Track based filters

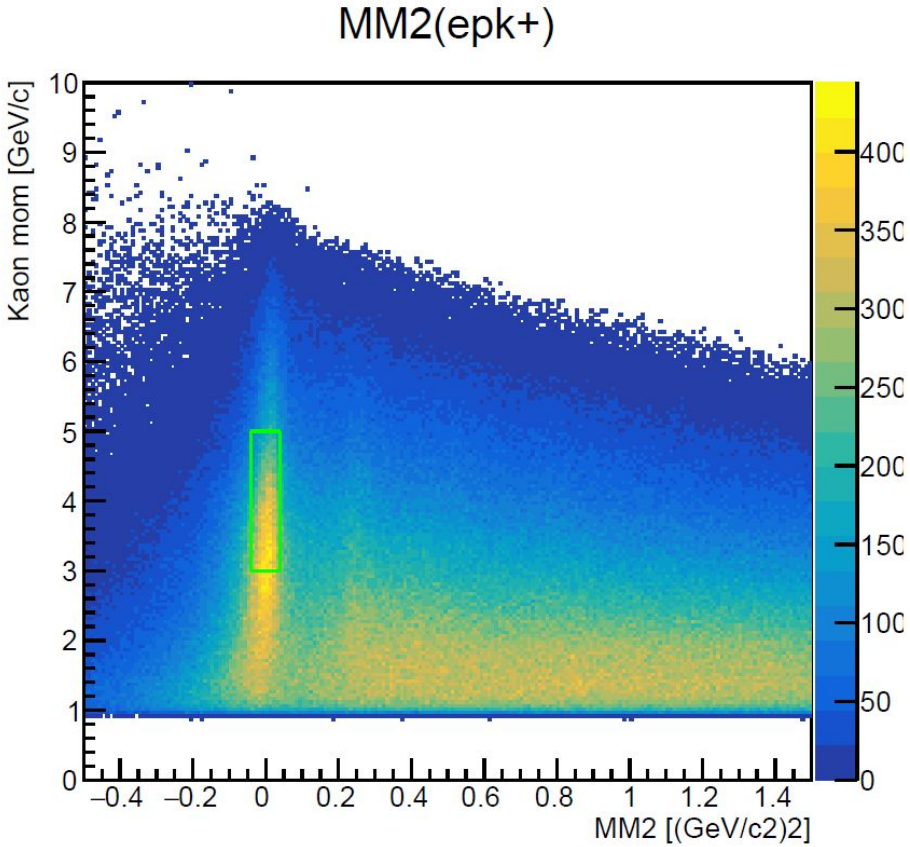
- 1.5 GeV < E(e) < 8 GeV
- **One** charged particle in the RICH
- At least one hit on MAPMT
- CLAS12 EB identifies as kaon or pion
- Missing **π-** cut for reactions **h+** kaon or pion

Hit based filters

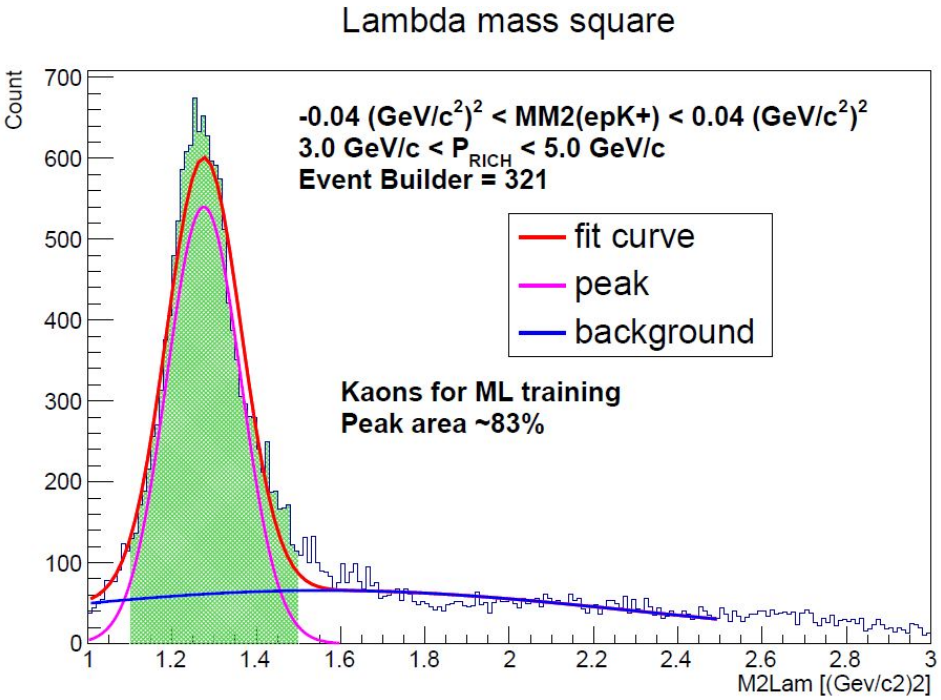
- Remove noisy anode hits
- Remove background hits based on timing

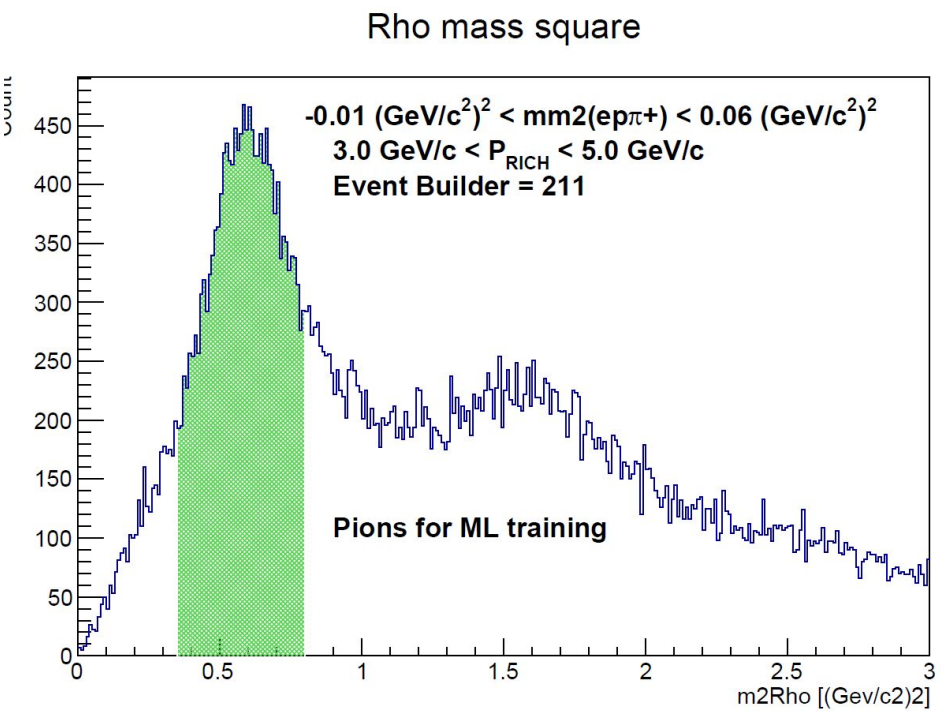
Kaon/Pion training data selection



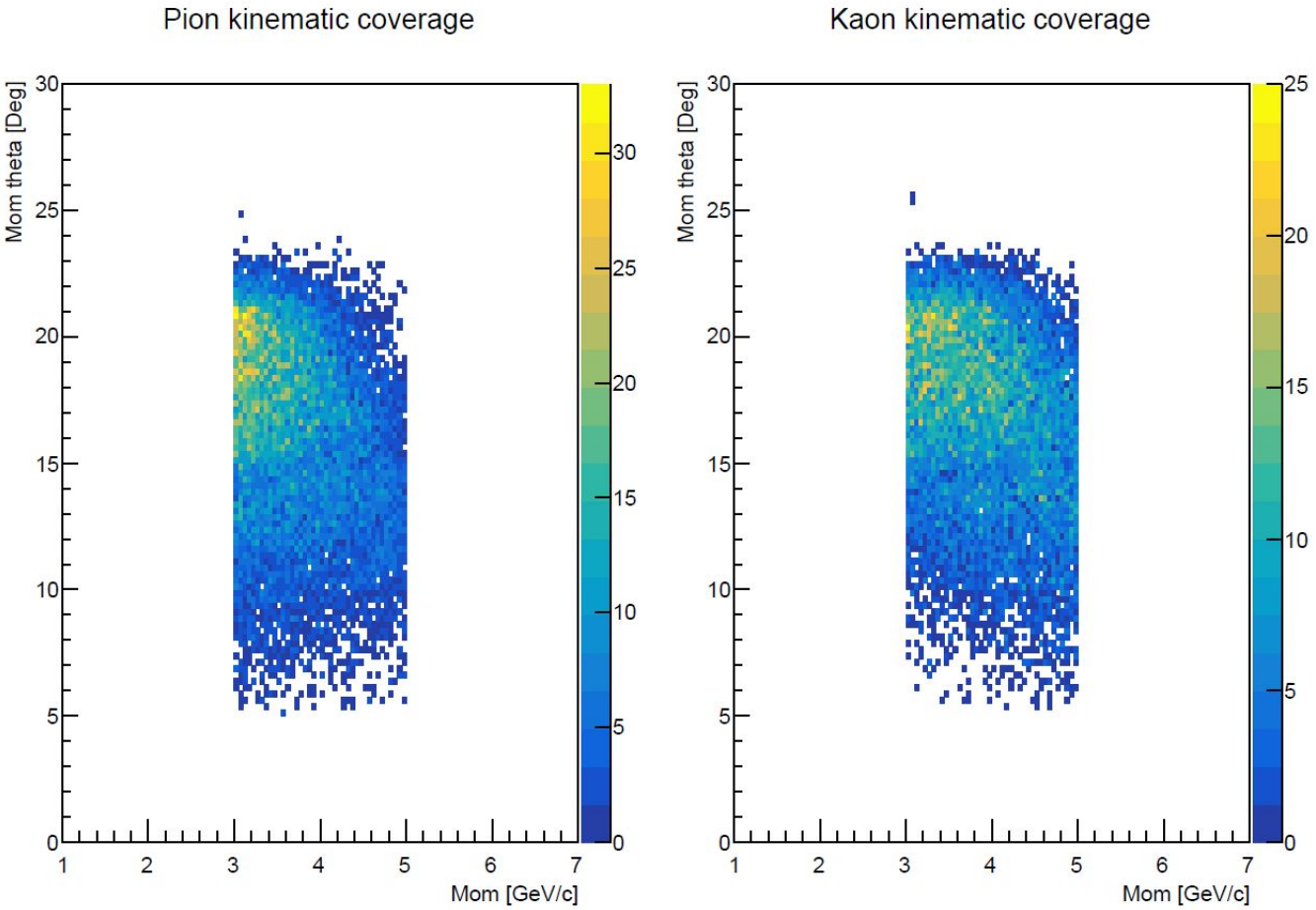


Kaon/Pion training data selection

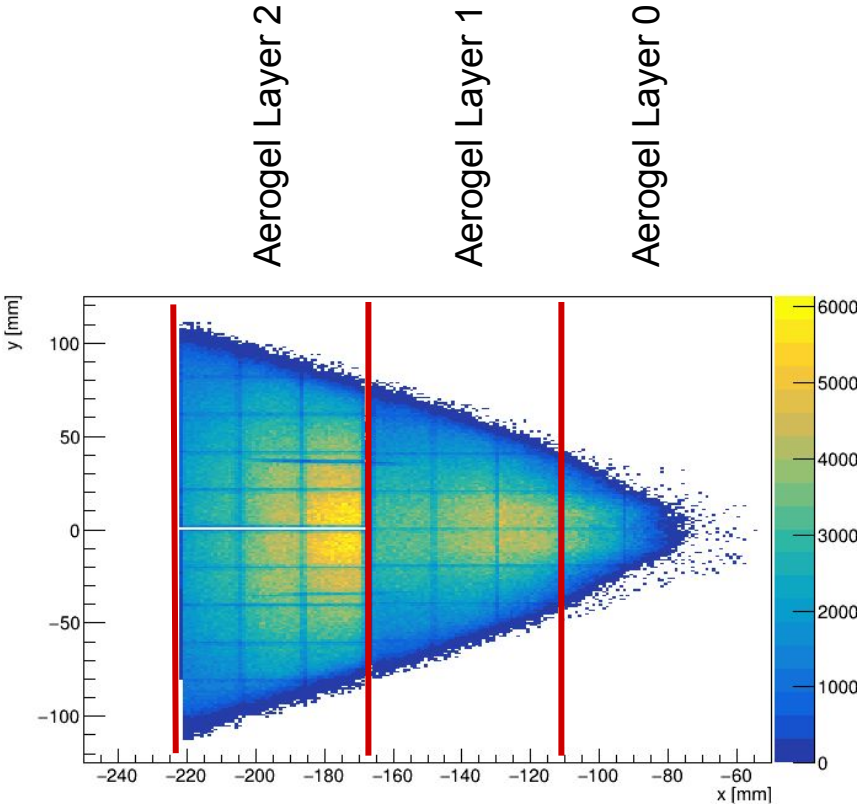


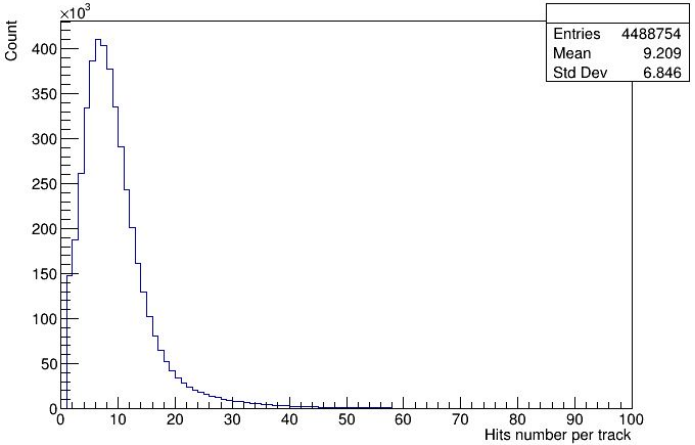


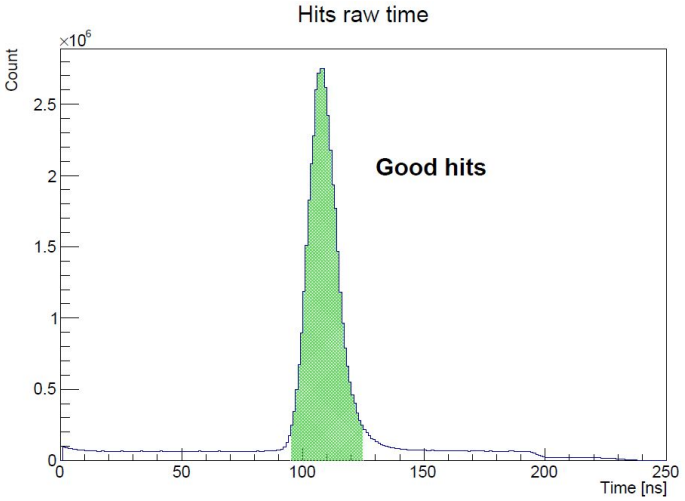
Kinematic coverage in training sample



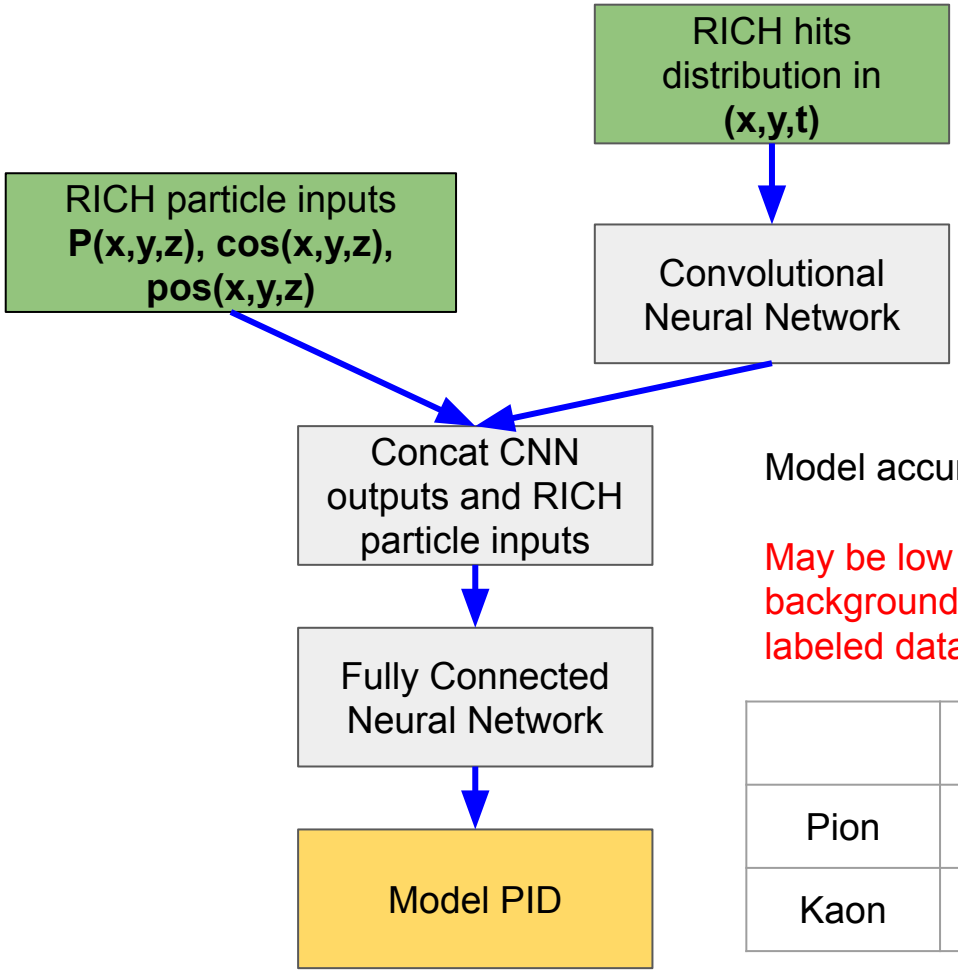
Hits distribution







Machine learning model



Training sample

Kaons : 14283

Pions : 16131

Average 10 hits per event

Model accuracy 66.7 %.

May be low because of background and wrong labeled data.

	Precision	Recall
Pion	67.8 %	66.7 %
Kaon	65.6 %	66.6 %

Precision - What is the probability that the model will predict label correctly.

Recall - What percentage of actual labels were predicted correctly.

Machine learning model

RICH hits distribution in **(x,y,t)**

RICH particle inputs **P(x,y,z), cos(x,y,z), pos(x,y,z)**

Convolutional Neural Network

Concat CNN outputs and RICH particle inputs

Fully Connected Neural Network

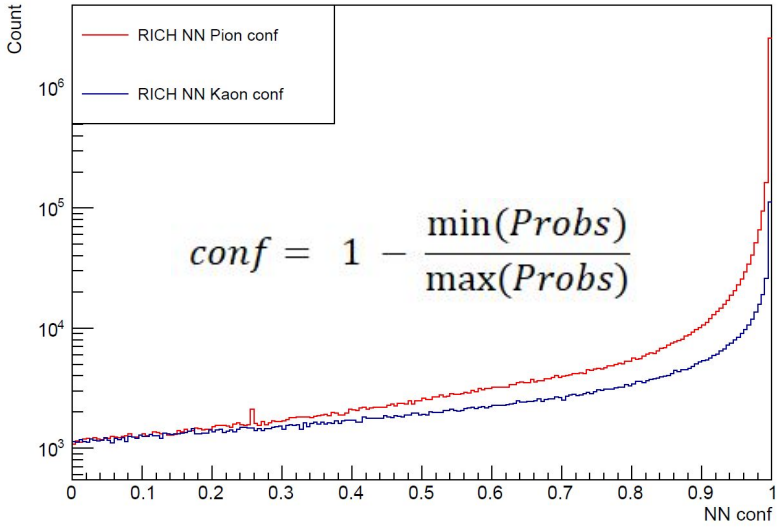
Model PID

Training sample

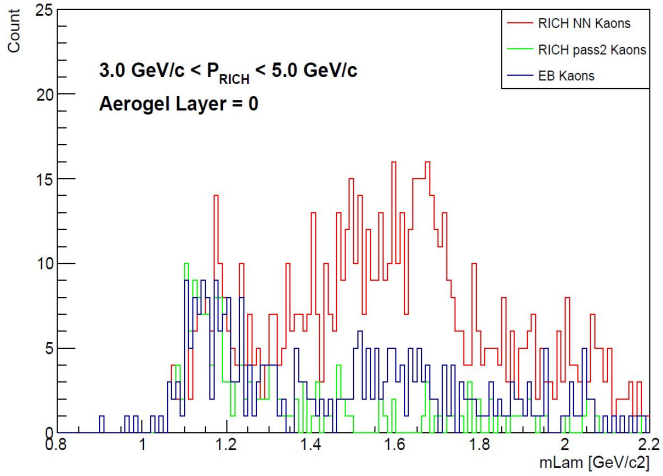
Kaons : 14283

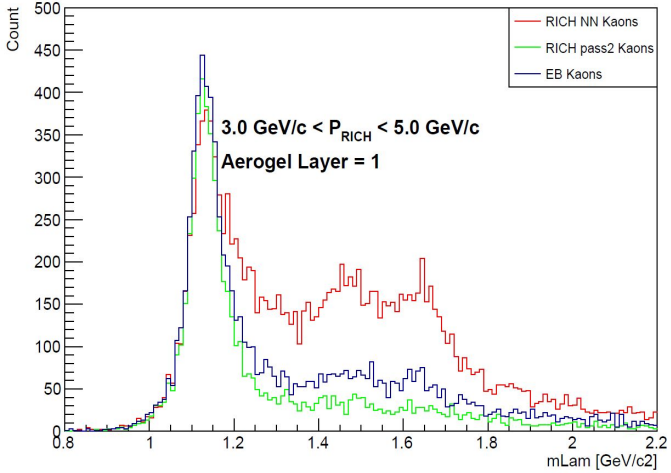
Pions : 16131

Average 10 hits per event

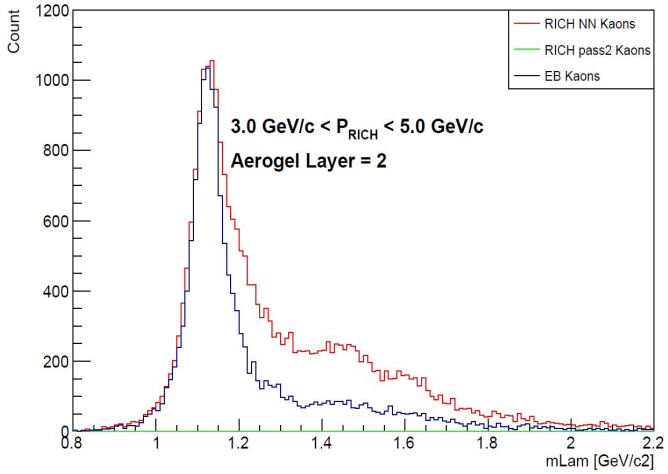


Preliminary results



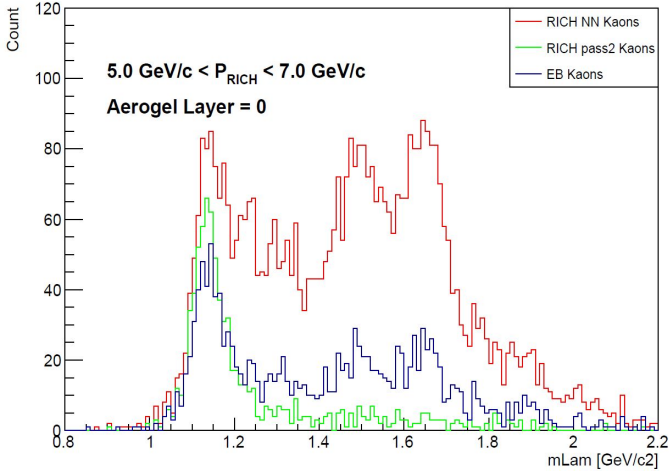


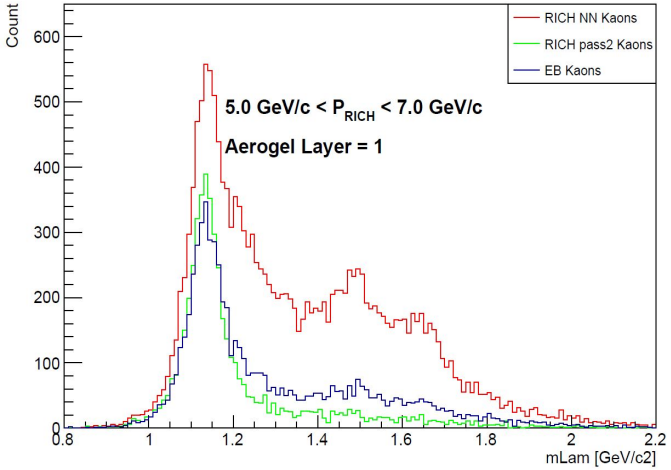
Training region



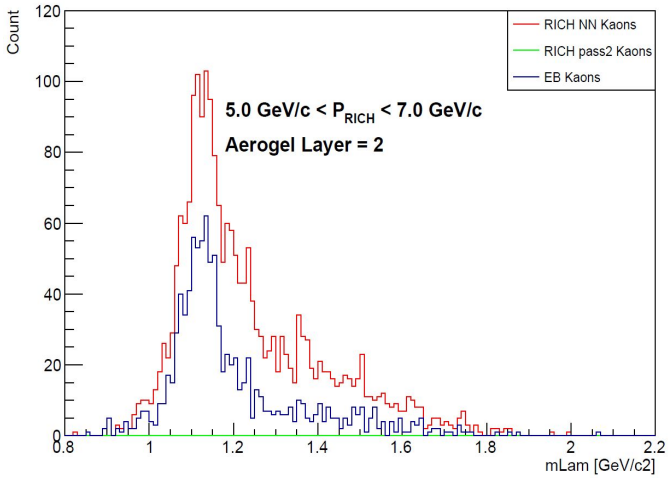
- RICH NN was able to reconstruct in all 3 aerogel layers, however background reduction still required
- Neural network is able to do predictions in aerogel layer 2, where RICH pass2 cooking is not aligned

Preliminary results





Outside training region based on momentum



- RICH NN was able to reconstruct in all 3 aerogel layers, however background reduction still required
- Neural network is able to do predictions in aerogel layer 2, where RICH pass2 cooking is not aligned
- Neural networks can be trained in one region and do prediction in other region as well.

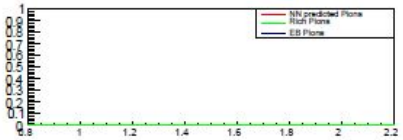
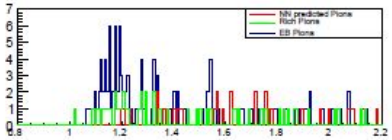
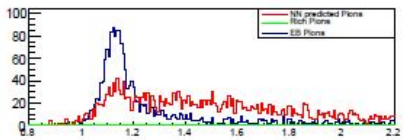
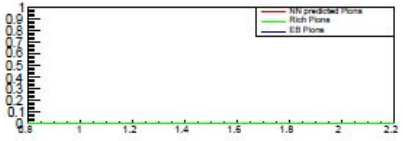
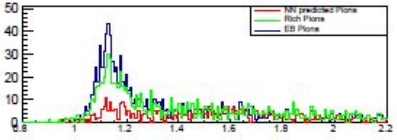
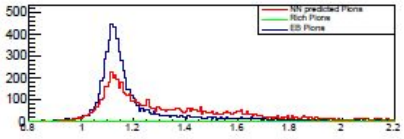
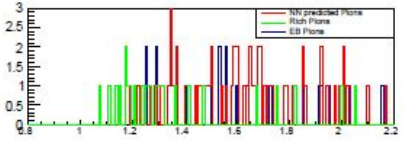
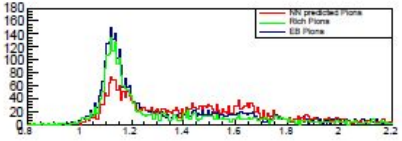
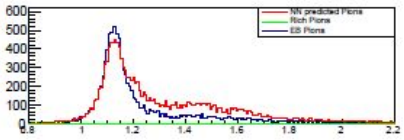
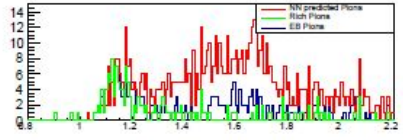
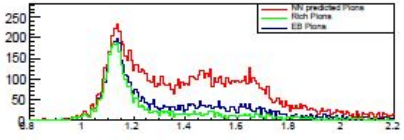
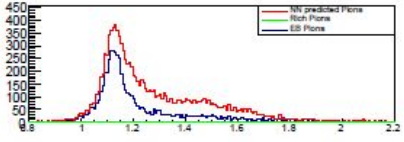
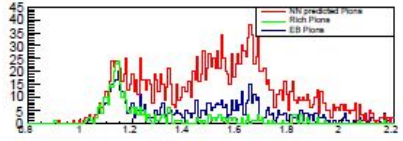
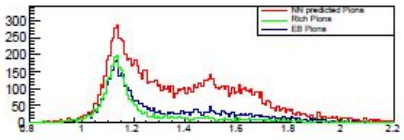
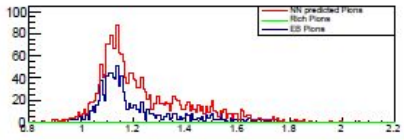
Conclusion and next steps

- **RICH NN was able to do predictions without alignment information in whole range. Further analysis is required to do background reduction, but important is that RICH NN was able to predict in the aerogel layer 2, where RICH is not aligned!**

- **Analyse the results to have better understanding on predictions.**

- **Improve the results based on data cleaning, more complex model selection or adding new input features.**

Thank you for you attention!
Questions?

	Aerogel Layer 0	Aerogel Layer 1	Aerogel Layer 2
P < 2 GeV/c			
2 GeV/c < P < 3 GeV/c			
3 GeV/c < P < 4 GeV/c			
4 GeV/c < P < 5 GeV/c			
5 GeV/c < P < 6 GeV/c			
6 GeV/c < P