ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Application of Machine Learning techniques to improve event reconstruction in Super-Kamiokande

Neutrino Physics and Machine Learning 2024
27 Jun. 2024

Nicola Fulvio Calabria
Politecnico and INFN Bari

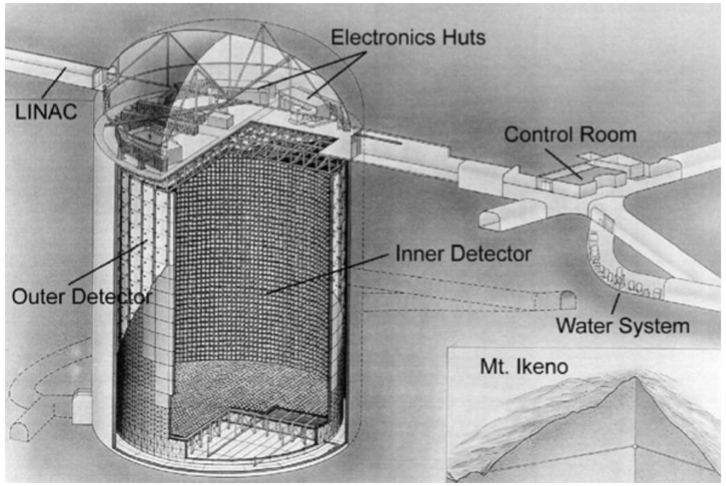
Outline

- Super-Kamiokande overview
- Reconstruction in Super-Kamiokande
- Proton decay and scientific motivation for reconstruction with Machine Learning algorithms
- Model training
- Preliminary results
- Conclusions and plans

Super-Kamiokande (SK), Kamioka mine, Japan

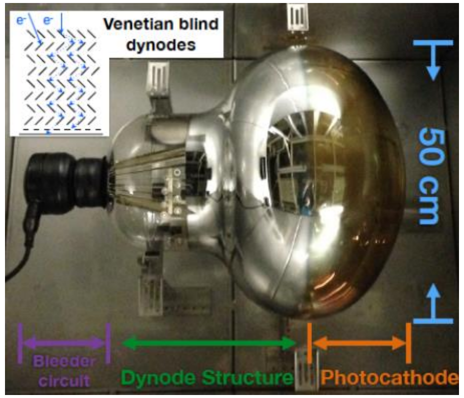
39 m x 40 m cylindric tank filled with 50 kton of ultrapure water, of which 22.5 kton inside Fiducial Volume, divided into two optically insulated sections:

- **Inner Detector (ID):** 11k 50 cm Photomultiplier Tubes (PMTs) (40% coverage) facing inwards.
- **Outer Detector (OD):** 2k 20cm PMTs facing outwards



Some research topics in SK:

- **Proton decay**
- Neutrino oscillations (2015 Nobel Prize)
- Neutrino astrophysics

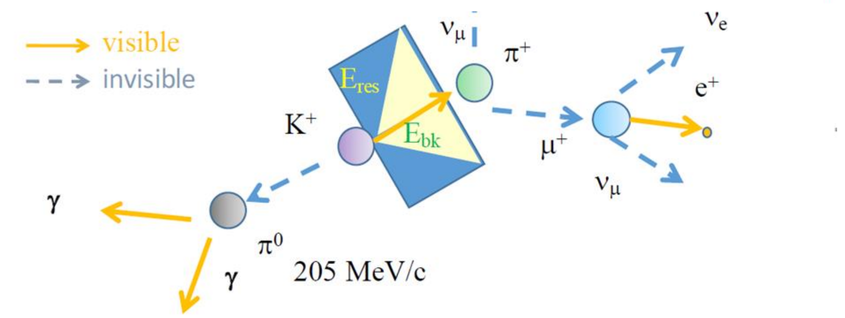


Proton decay p -> ν K^+ as a case study in SK

p -> ν K^+
Partial lifetime limit: 5.9 x 10^{33} yrs

Reference Study with APFit:
PHYSICAL REVIEW D90,072005 (2014)
«Search for proton decay via p→νK using 260 kiloton · year data of Super-Kamiokande»

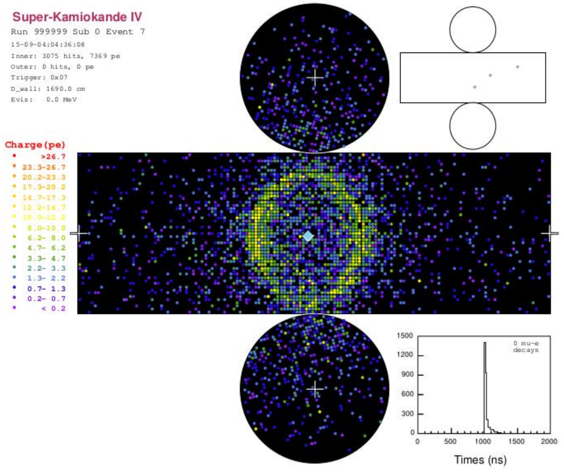
$K^+ \rightarrow \pi^+\pi^0$: Hadronic decay channel in water
$BR_{K^+\rightarrow\pi^+\pi^0}$ ~ 20%



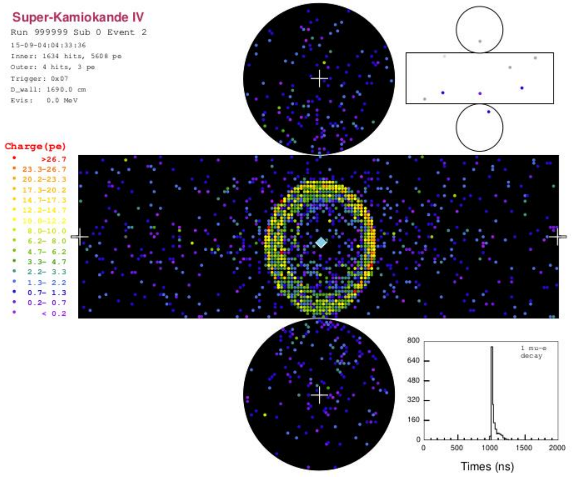
Upon trigger, for each hit PMT, charge produced and time of the hit are collected (event)

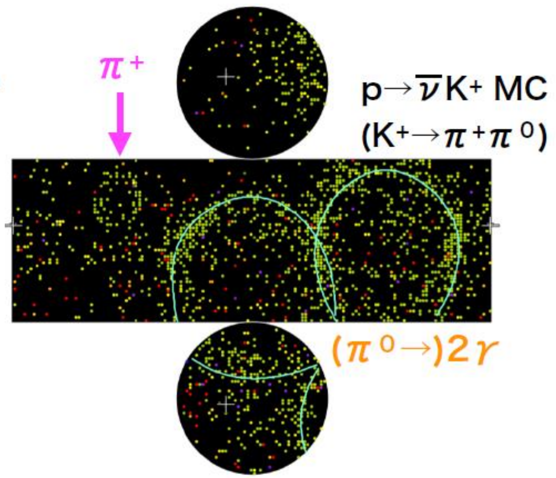
Atmospheric neutrino interaction events are background for this analysis

Showering (e-like)



Non showering (muon-like)

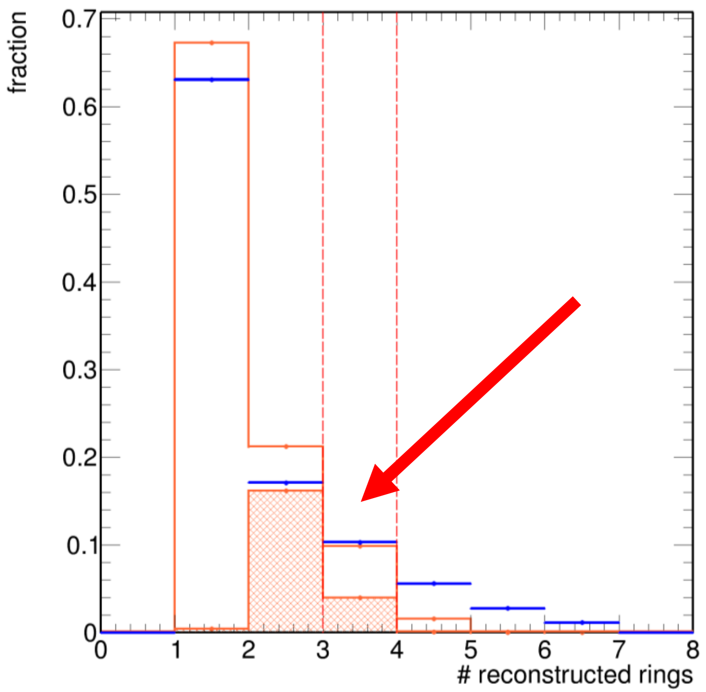




Reconstruction in SK

	APfit	fiTQun
Type of fit	Sequential (vertex, ring counting, PID, michel-e tagging)	Single log-likelihood function minimization $$L(\mathbf{x}) = \prod_j^{\text{unhit}} P_j(\text{unhit}\vert\mathbf{x}) \prod_i^{\text{hit}} [1 - P_i(\text{unhit}\vert\mathbf{x})]\, f_q(q_i\vert\mathbf{x}) f_t(t_i\vert\mathbf{x})$$
Used by	Super-Kamiokande	T2K, MiniBooNE, Super-Kamiokande, Hyper-Kamiokande
Max # rings	5	6
PID	$e^\pm$, $\mu^\pm$	$e^\pm$, $\mu^\pm$, $\pi^\pm$
CPU time per SK event	< 1 min/event	~ 10 min/event

fiTQun makes the reconstruction of charged kaon kinematics possible (charged pion PID)



SK ATM-ν MC (background)
SK PDK MC
True $K^+ \to \pi^+\pi^0$, $\pi^0 \to \gamma\gamma$ (hatching) (signal)

Exposure (kton*yr)	#BG	BG sys. Err. (%)	Eff. (%)	Eff. Sys. Err. (%)
200	0.03 ± 0.02	50.9	2.9 ± 0.02	26.1

Low-background analysis in this proton decay channel with fiTQun is possible.

We aim to increase signal selection efficiency by improving ring detection.

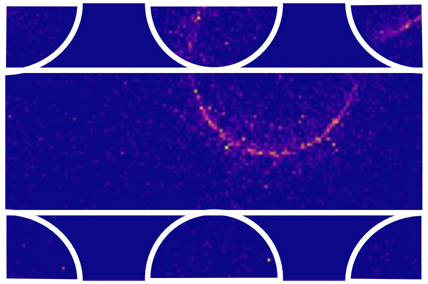
Machine Learning algorithms are interesting candidates for this purpose.

Results and plot from N.F. Calabria PhD Thesis, ‘Search for proton decay in Super-Kamiokande and perspectives in the Hyper-Kamiokande experiments’, 2023, Università degli Studi di Napoli.

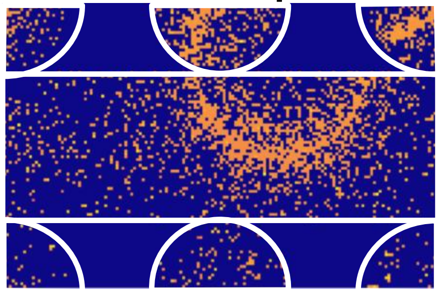
Reconstruction of electron events in Super-Kamiokande with Machine Learning

- Preliminary study using a custom ResNet-18 based Neural Network in PyTorch.

Hit Charge Bitmap



Hit Time Bitmap



Normalize each layer (subtract hits mean, divide by hits std. dev.)

Conv 3D, output 64
Kernel (2x7x7)
Stride 2
Padding (0x3x3)

From ResNet-18

3×3 max pool, stride 2

$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$

$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$

$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$

$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$

average pool

Hit charge mean and std. dev
Hit time mean and std. dev

Sequential:
- Linear (4, 4)
- Tanh

Sequential:
- Linear (516, 512)
- Tanh
- Linear (512, 512)
- Tanh
- Linear (512,7)
- Cosmetic output layer

Reconstruction:
- Momentum
- Direction (x, y, z)
- Vertex (x, y, z)

Model Training

Dataset:
4 M (Train/Validation: 80%/20%) + 1M (Test) electron events generated with SKDETSIM
Momentum: 0 – 1000 MeV/c isotropic
Vertex: uniform in ID volume with distance from wall 100 cm

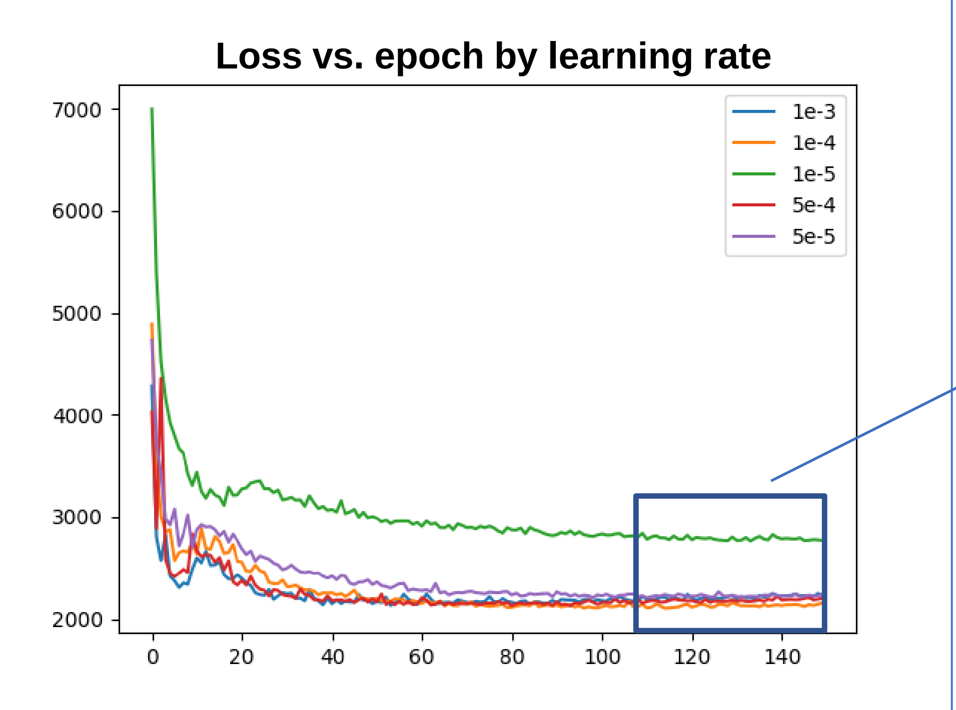
Hardware:
Tesla A100 40 GB

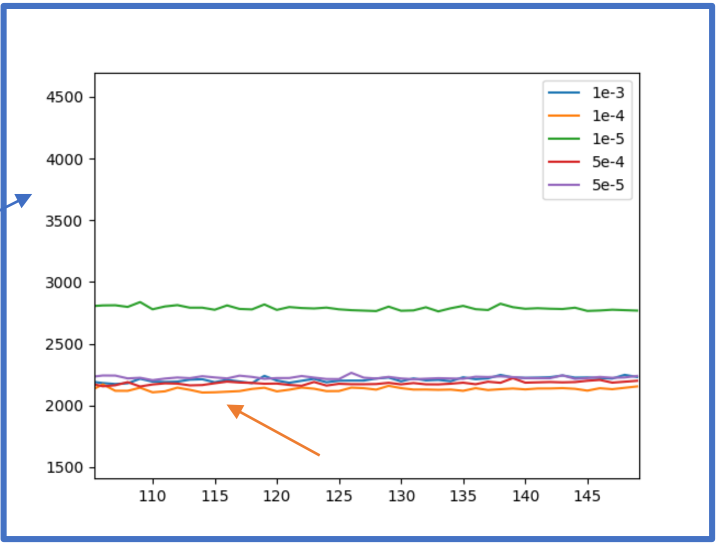
Optimizer:
ADAM

Loss:
MSELoss

Learning rate:
Coarse grid search, 150 epochs per trial

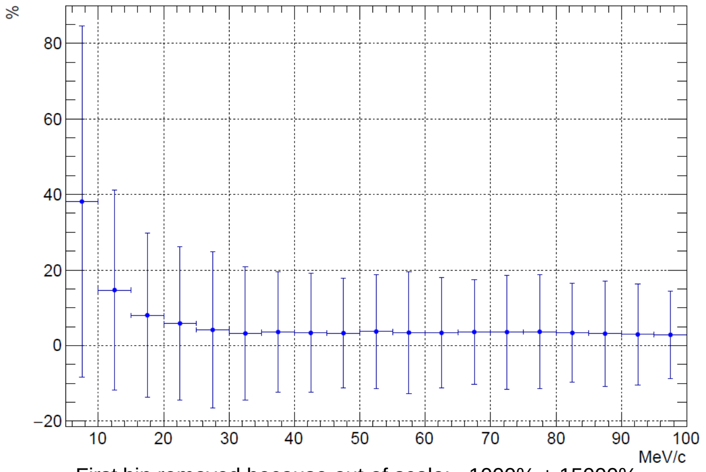
Best candidate chosen:
Learning rate 1×10^{-4} after 115 epochs
(~30 minutes per epoch)





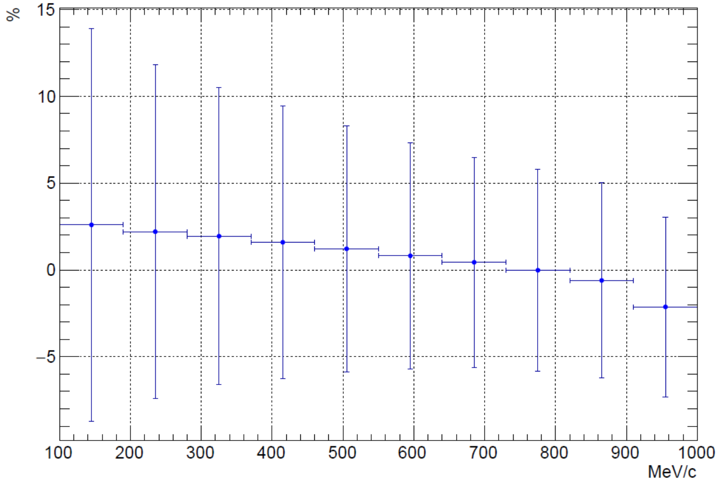
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
INFN

Momentum Bias vs. True Momentum



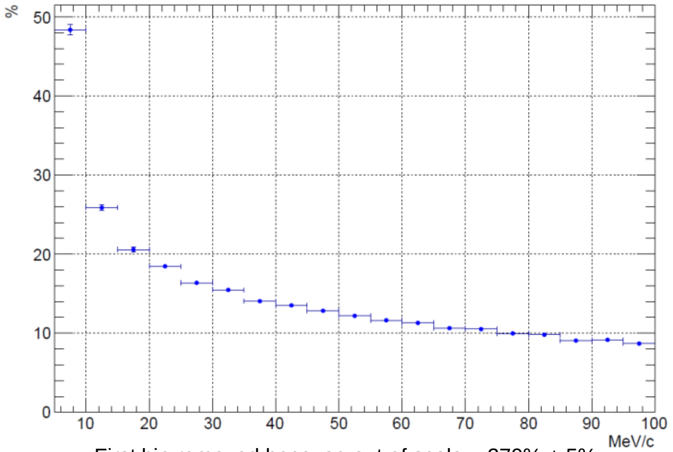
First bin removed because out of scale: ~1000% ± 15000%

Momentum Bias vs. True Momentum



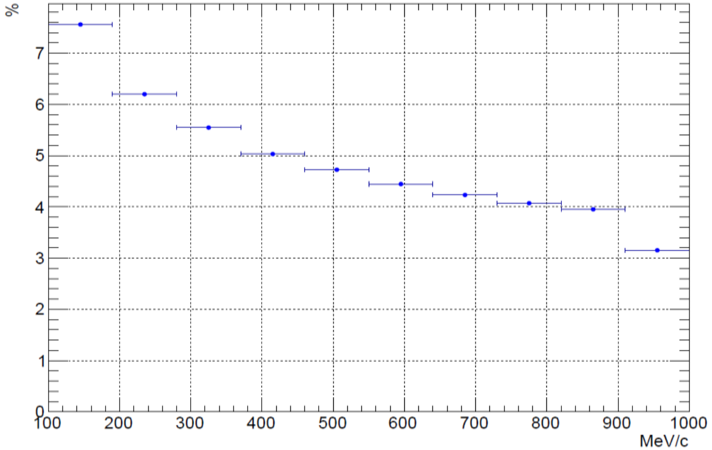
PRELIMINARY

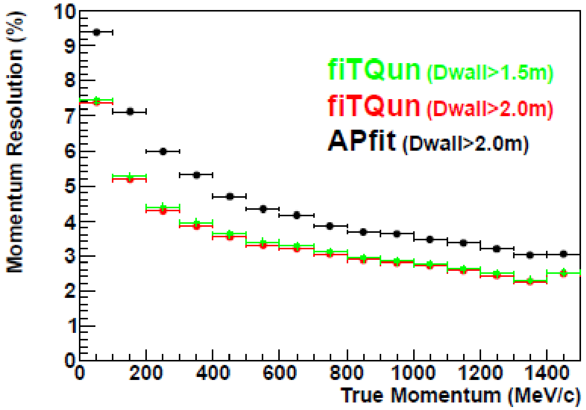
Momentum Resolution vs. True Momentum



First bin removed because out of scale: ~270% ± 5%

Momentum Resolution vs. True Momentum



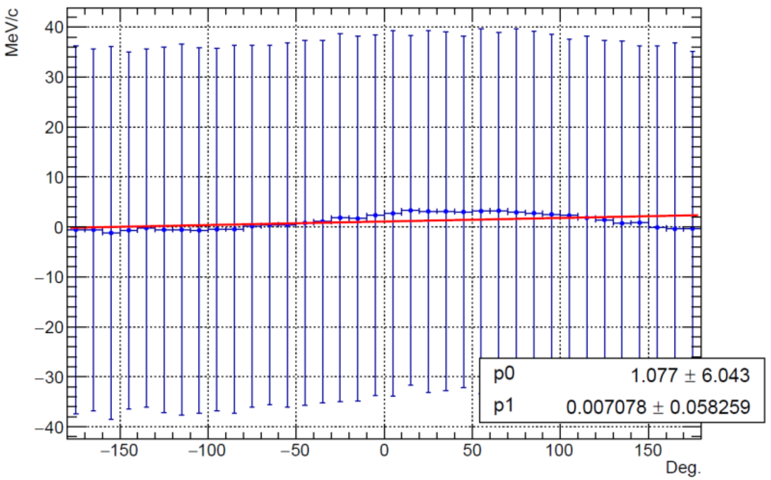


This plot and following fiTQun references are from Y. Suda PhD Thesis «Search for proton decay using an improved event reconstruction algorithm in Super-Kamiokande», 2017.

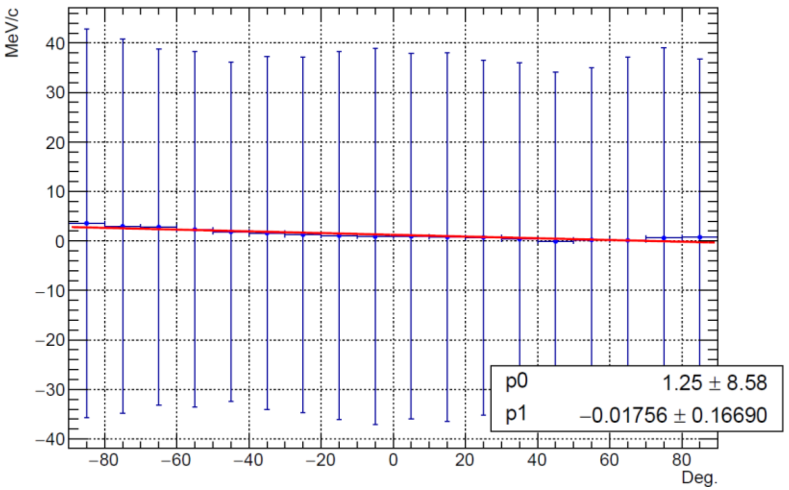
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
INFN

PRELIMINARY

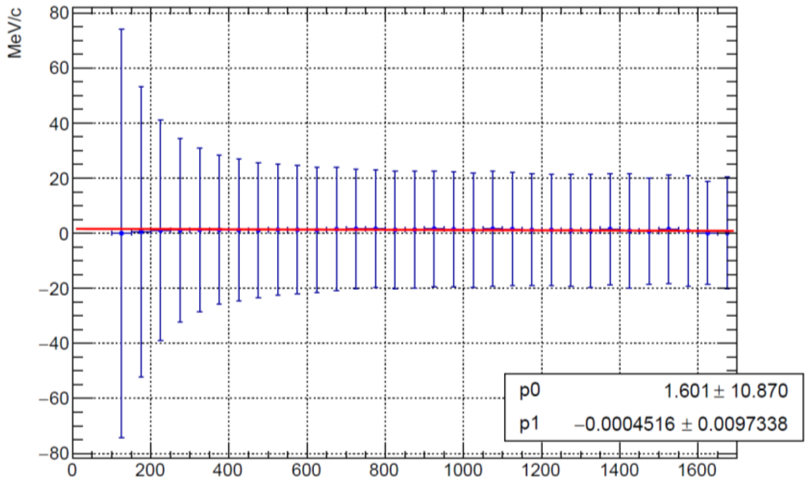
Momentum Bias vs. True Azimuth



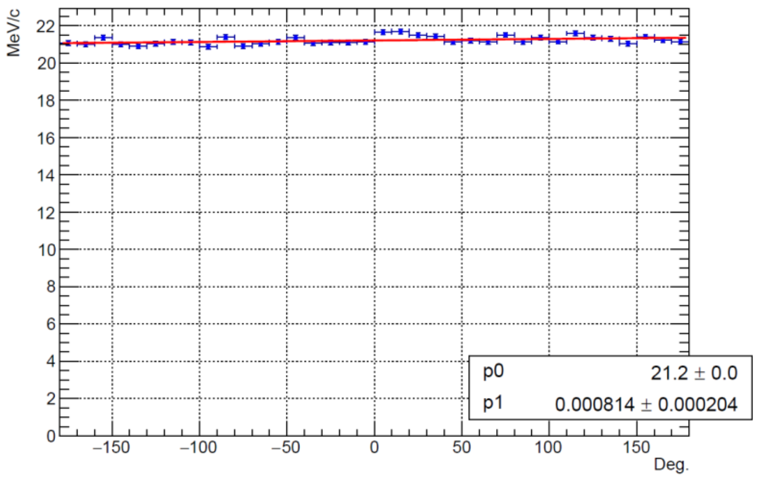
Momentum Bias vs. True Elevation



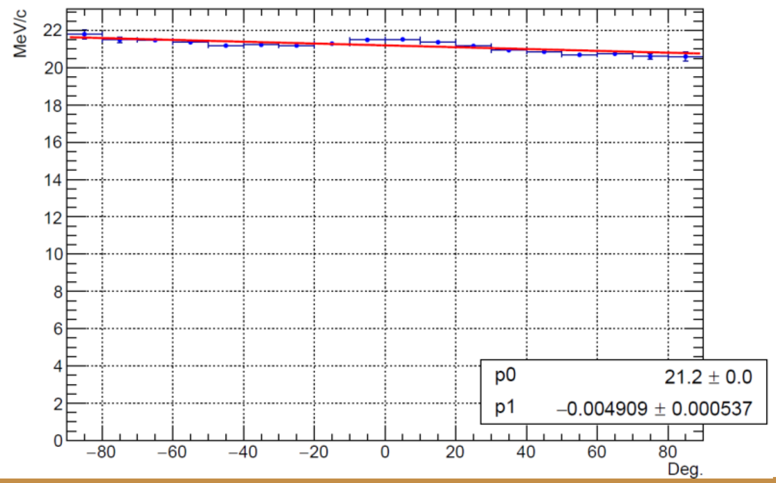
Momentum Bias vs. True Distance from ID Wall



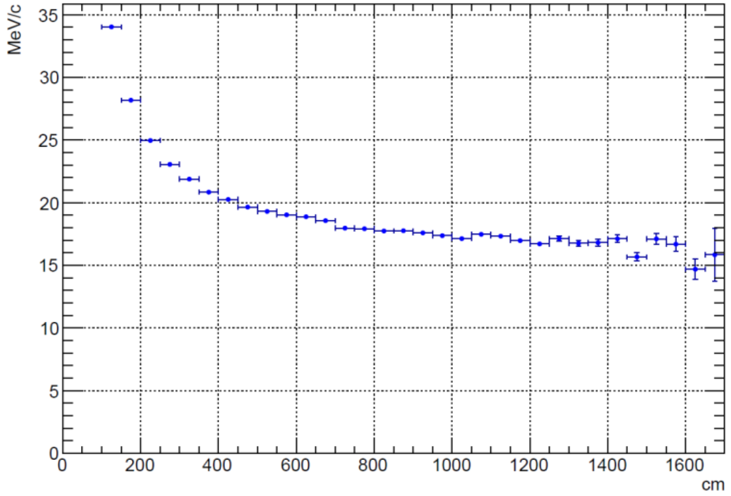
Momentum Resolution vs. True Azimuth



Momentum Resolution vs. True Elevation

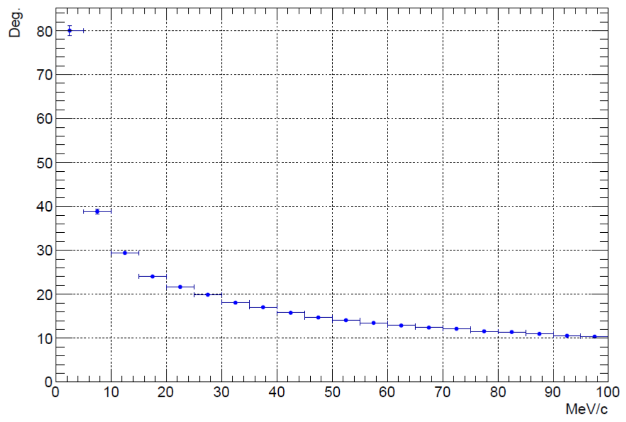


Momentum Resolution vs. True Distance from ID Wall

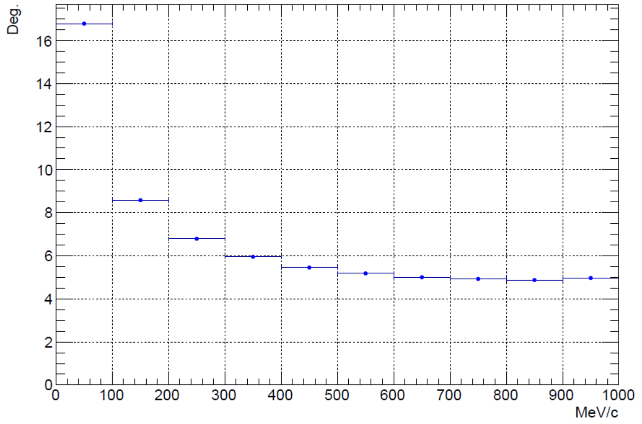


PRELIMINARY

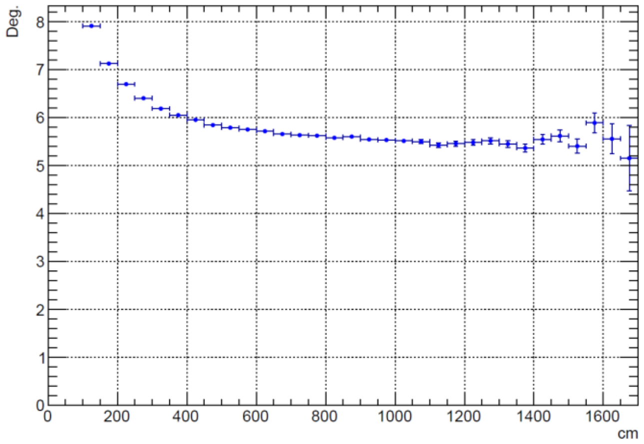
Angular Resolution vs. True Momentum



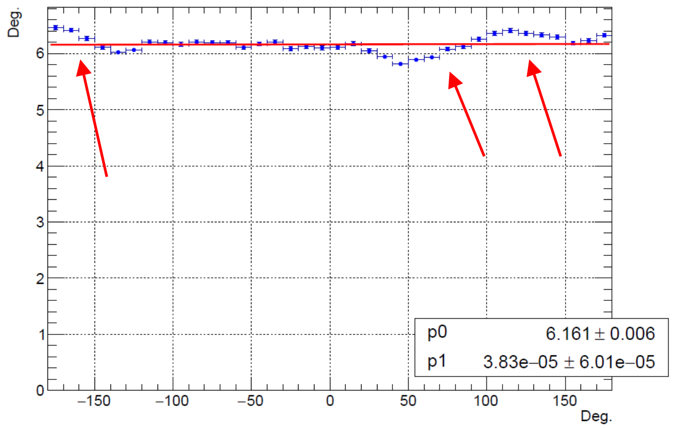
Angular Resolution vs. True Momentum



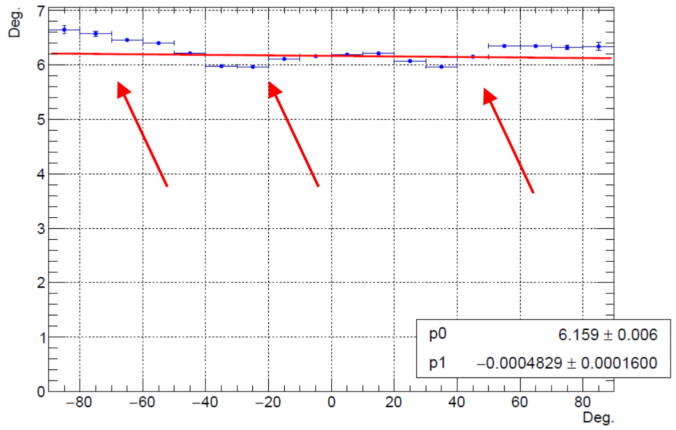
Angular Resolution vs. True Distance from ID Wall



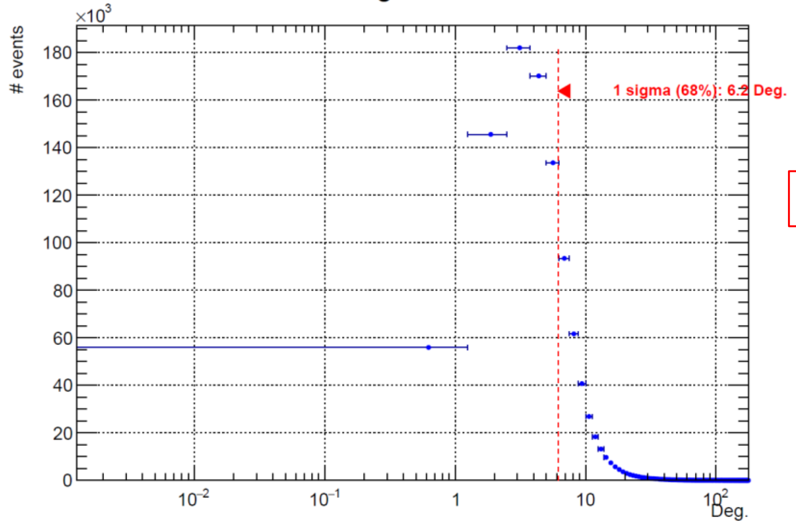
Angular Resolution vs. True Azimuth



Angular Resolution vs. True Elevation



Overall Angular Resolution



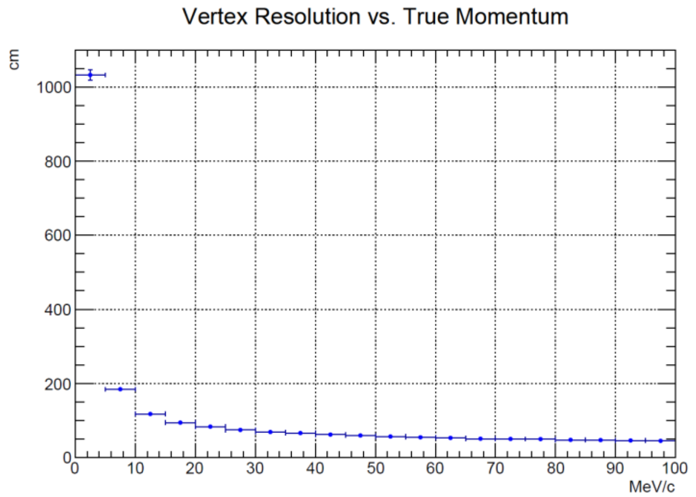
fiTQun: ~3°

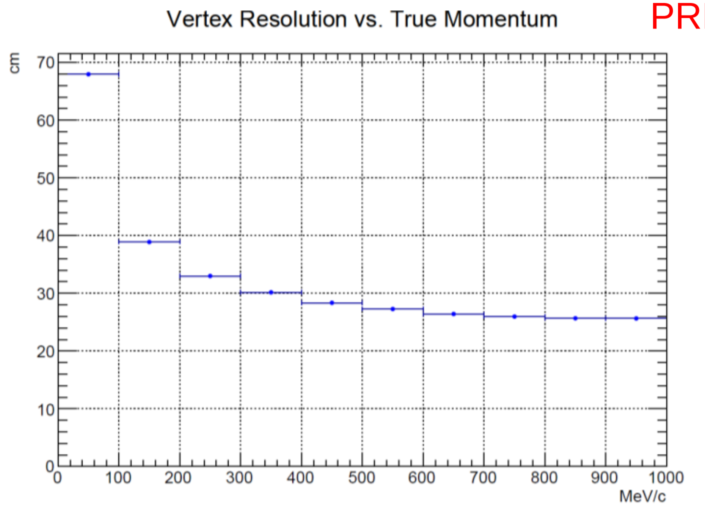
Finanziato dall'Unione europea NextGenerationEU

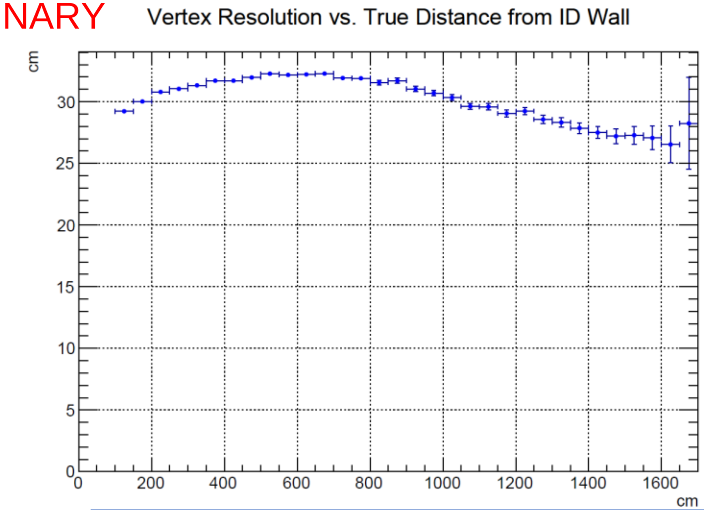
Ministero dell'Università e della Ricerca

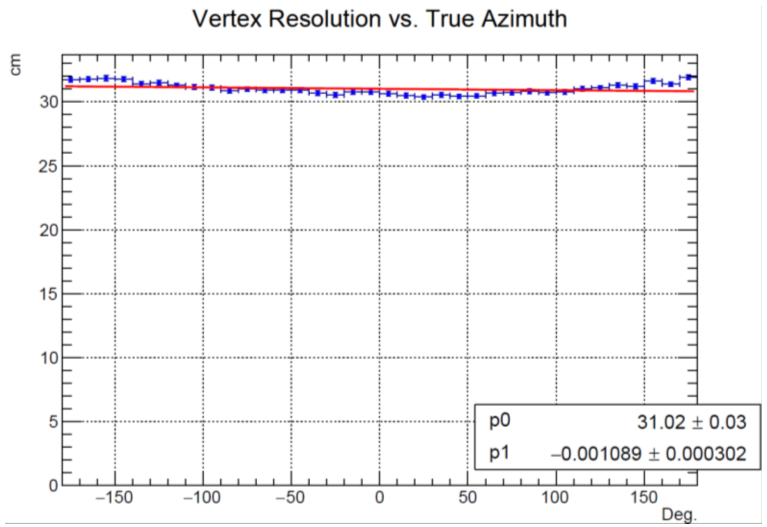
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

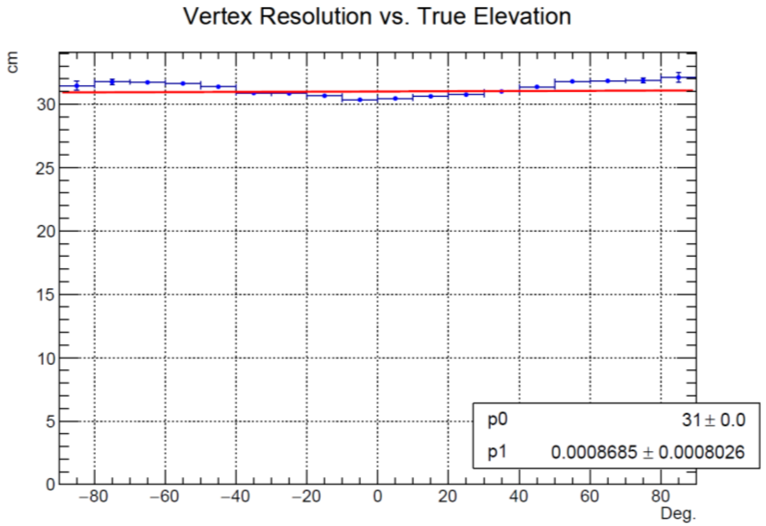
PRELIMINARY

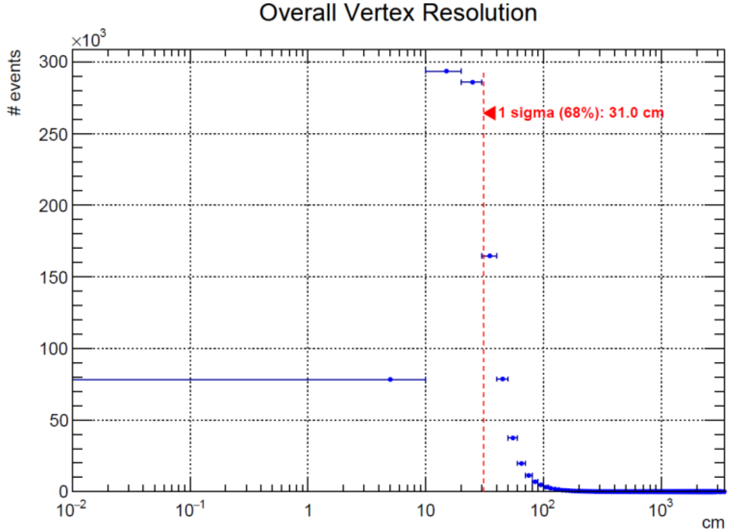












fiTQun: 19.5 cm

Conclusions and plans

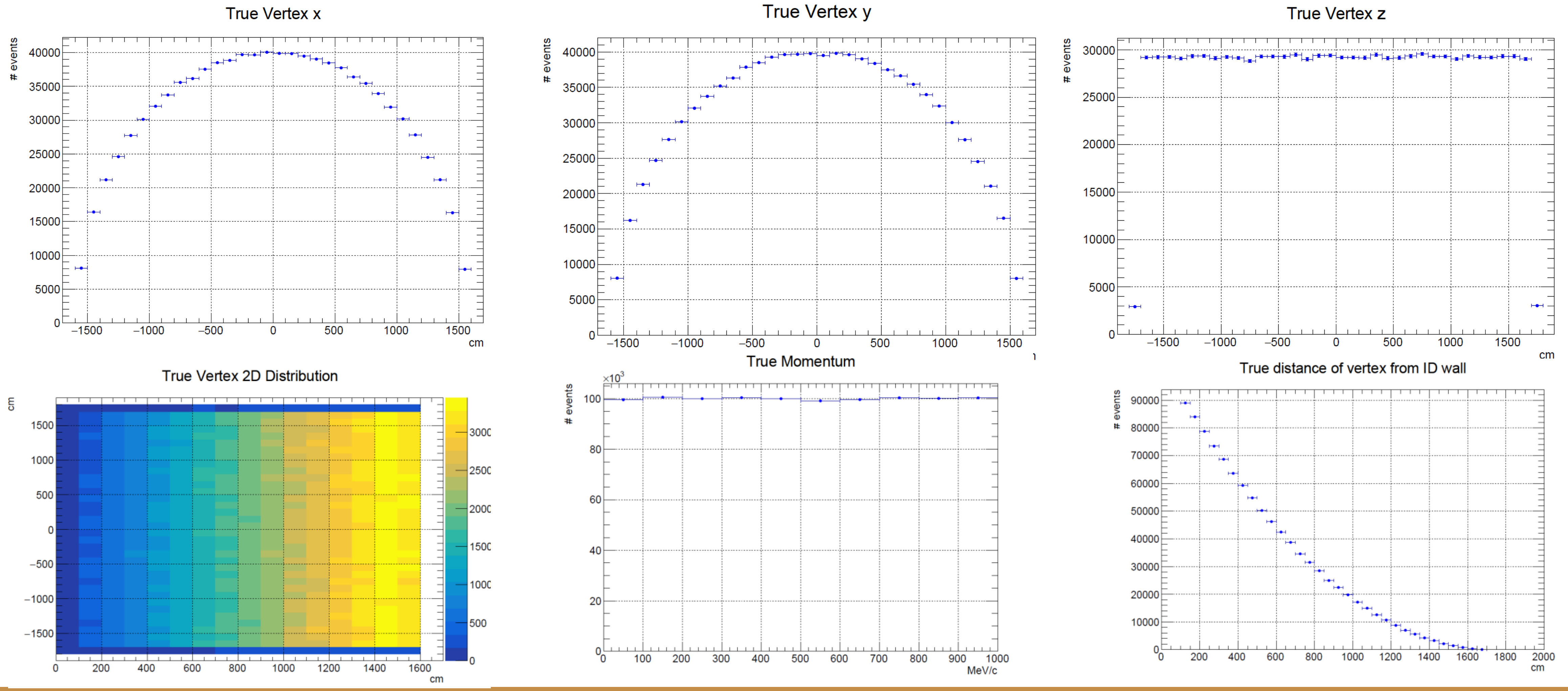
- I trained and tested a preliminary Machine Learning model for reconstruction in Super-Kamiokande with an electron-only dataset
- There are some features that need to be understood well
- There is much room for improvement and optimization
- Plan to extend this study to muons and charged pions
- It would be interesting to extend this study to an alternative architecture based on Graph Neural Networks

“This work is (partially) supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU”

THANK YOU!

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
INFN

Test dataset MC truth check



Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | INFN

Test dataset MC truth check