Status of Application of ML techniques in IWCD and WCTE

WCTE Collaboration Meeting Updates

Date: 22 July 2022

Speakers-
Tanima Mondal
Sunanda
Arnab Sarker

Motivation

- The Machine learning (ML) technique is an effective tool to deduce detailed information from a complex image.

- ML techniques has wide application in:
    - ➢ Particle type identification
    - ➢ Reconstruction of single/multi-ring events
    - ➢ Reconstruction of particle kinematic variables

- ML ResNet model shows significant performance over fiTQun.

- The performance of ResNet Model is analysed using IWCD geometry for particle gun events.

- ML techniques will be applied to the IWCD event selection and WCTE particle gun dataset.

Intermediate Water Cherenkov Detector(IWCD)

- IWCD is a sub-kiloton scale water Cherenkov detector.
- Nominally designed to measure neutrino interactions before oscillation effect is significant.
- Vertical Moving detector, 1-4 degrees off-axis spanning.

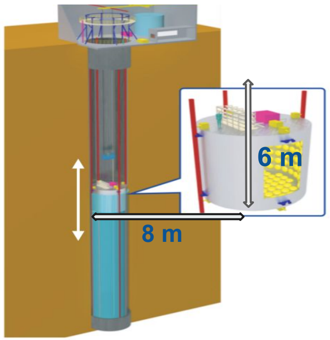


Figure: IWCD Detector (**Short tank geometry)**

Rebaseline of IWCD geometry:

	Current Geometry
ID Radius	400 cm
ID HalfLength	300 cm
Baseline	~ 750 m from source

Physics Goals for Rebaseline:

- **Study neutrino interaction rate peaked at different energies, higher precision.**
- **To identify ~1% of anti(neutrino) ν_e components in the beam.**
- **Mitigate neutrino beam pile up events.**

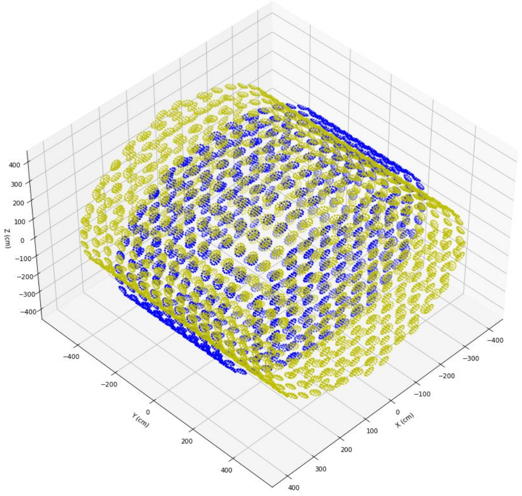


Figure: ML Techniques used to generate Long and short tank geometry

IWCD constraint:

- Conventional neutrino beam contains only 1.5% ν_e, challenging to measure ν_e cross-section.
- The electron (anti)neutrino intrinsic fluxes produce single-ring electron (1Re) samples, use to constrain cross-section ratio of $\sigma_{\nu_e}/\sigma_{\nu_\mu}$, $\sigma_{\bar{\nu}_e}/\sigma_{\bar{\nu_\mu}}$

Current Progress with IWCD:

Analysing fiTQun PID performance for IWCD current geometry to study ν_e event samples.

Sample selection criteria:

- Samples to be enriched with ν_e interactions.

Event selection:

- CC0$\pi\nu_e$ signal events, NCπ^0 and entering γ as background events.

Reconstruction Algorithm:

- fitQun: Evaluates best-fit likelihood

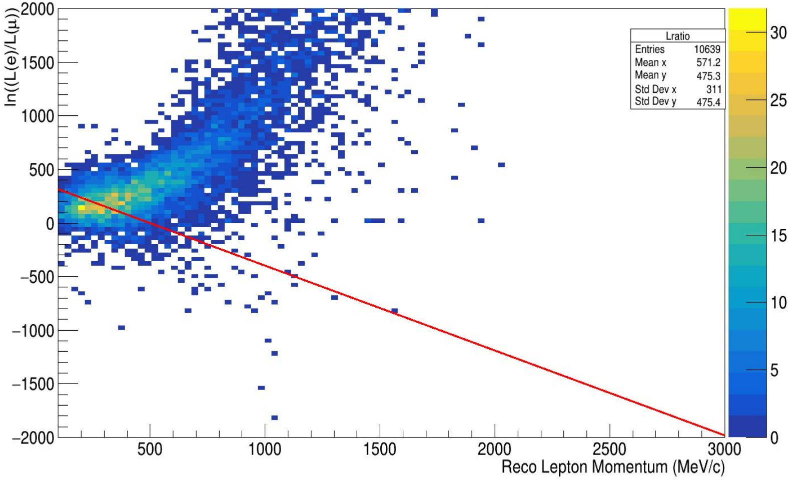
Cut tuning :

- Separate signal from Background

e-μ cut: LR = $\ln\left(L(e)/L(\mu)\right)$

π^0-e cut: LR = $\ln\left(L(\pi^0)/L(e)\right)$

Selection overview for e-mu cut:



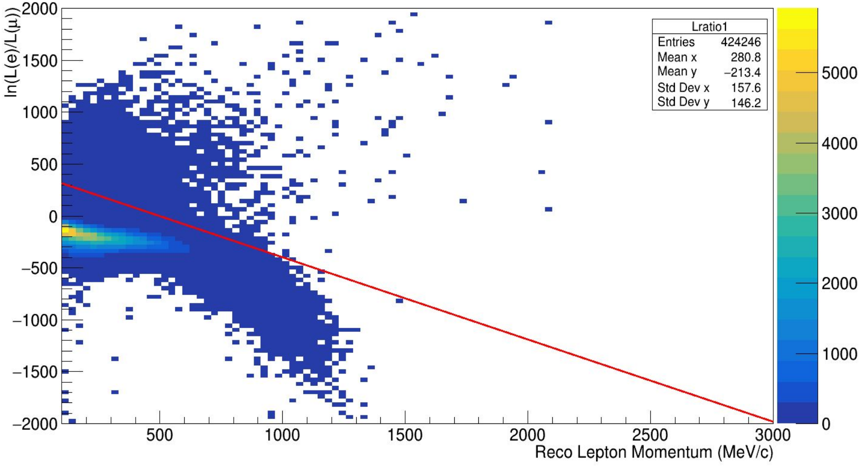


Figure: Distribution of events in reconstructed electron-muon likelihood ratio vs reconstructed lepton Momentum

➢ **True Event:** CC0$\pi\nu_e$

CC : NEUT Code < 30

Neutrino ID : ν_e = 12 (PDG)

Towall Cut > 50 cm

Dwall Cut > 75 cm

IsOneRingCandidate event

➢ **Background Event:** ν_μ CC

CC : NEUT Code < 30

Neutrino ID : ν_μ = 14 (PDG)

TruthMuonContainmentCut:

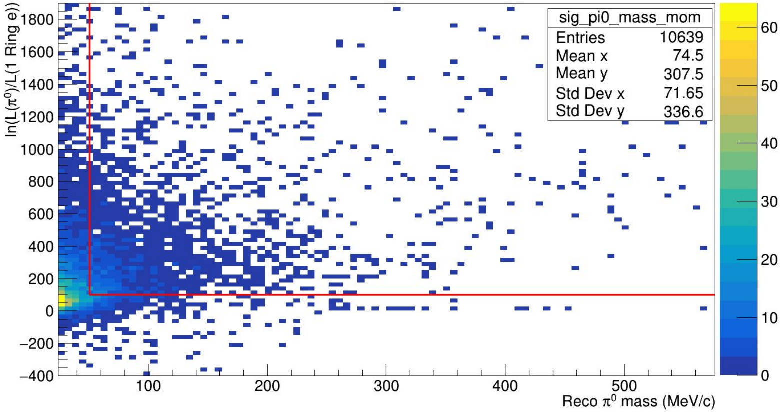
a) Towall > 50 cm

b) lepton mom p < 2 * wall

Cut tuning with lr = (316.6 - fq1rmom[0][1]*0.7657)

Selection overview for π^0-e cut:

WCTE will measure properties of π±, and π^0s from charge exchange.



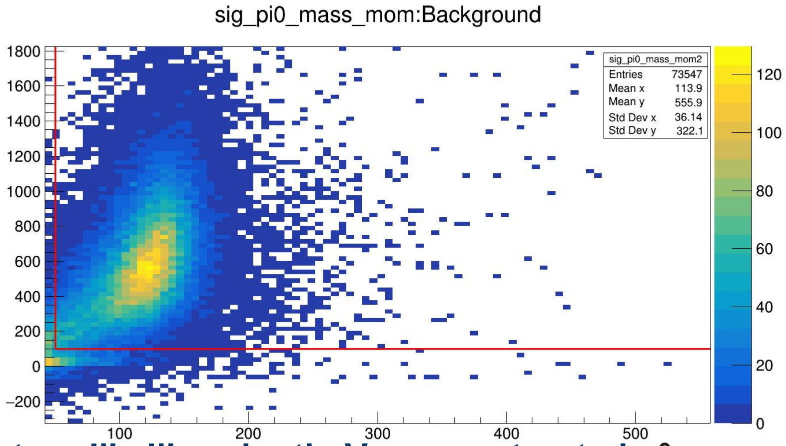


Figure: Distribution of events in reconstructed π^o-1 ring electron likelihood ratio Vs reconstructed π^o mass

- ➢ **True Event:** CCo$\pi\nu_e$

CC : NEUT Code < 30
Neutrino ID : ν_e = 12 (PDG)
Towall Cut > 50 cm
Dwall Cut > 75 cm
IsOneRingCandidate event

- ➢ **Background Event:** NCπ^o

1. **CC :** NEUT Code > 30
2. **Neutrino ID :** π^0 = 111 (PDG)
3. **Towall** > 50 cm

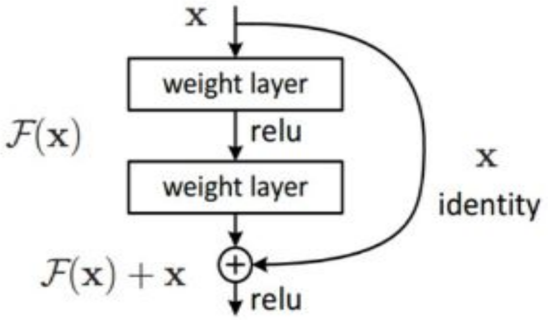
π^0 mass >50 Mev/c and LR >100

Cut line will be tuned for Current detector geometry

Machine learning Techniques

ResNet Model:

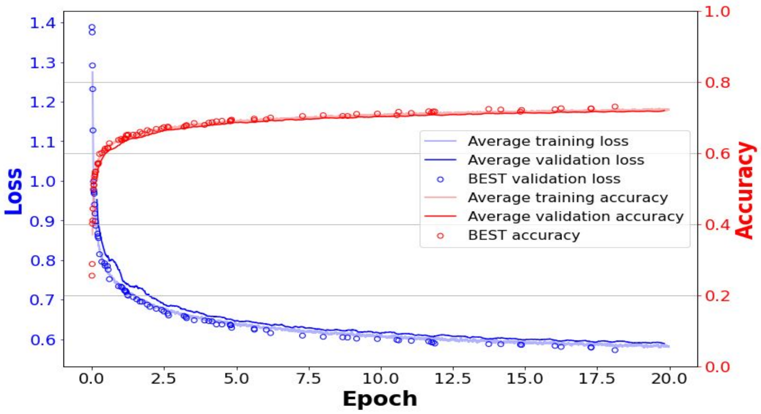
- ResNet is a Convolutional Neural Network (CNN) architecture
- Solve Complex problems, stack some additional layers.
- Improve accuracy
- Boost the performance of Neural Network



Application:

Identifying four kinds of particle gun events (e^-, μ^-, γ, π^0) simulated using WCSim software.

- **Loss** signifies how well the model is trained in corresponds to the actual data.



Results

- With the short tank IWCD geometry, ResNet model is trained to distinguish between two classes of particle: e^- and γ

Validating the events

- For each type of particles,
  - ➢ ~ 9,000,000 events were produced for e^-
  - ➢ ~ 9,000,000 events were produced for γ

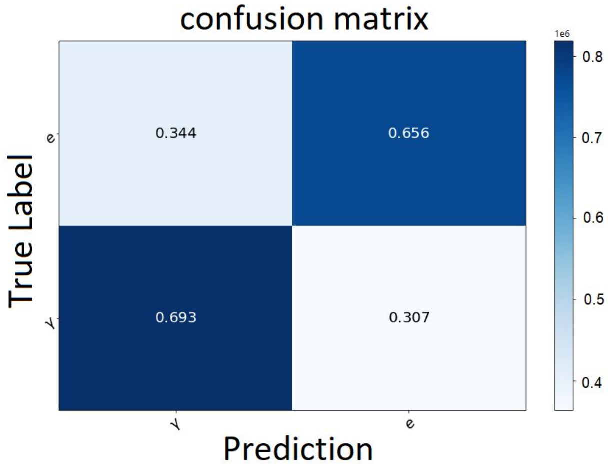


Figure: **Confusion Matrix**

$P_e(e)$=0.656

$P_e(\gamma)$= 0.344

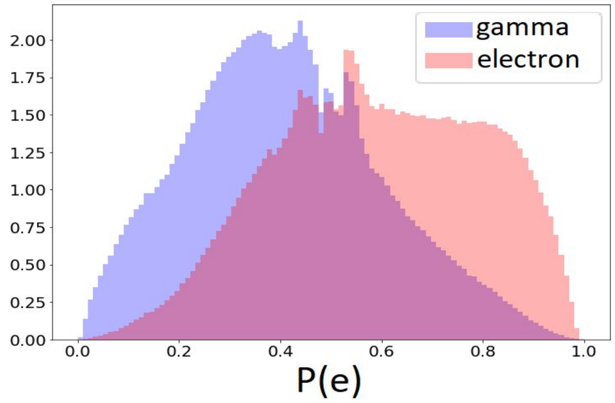
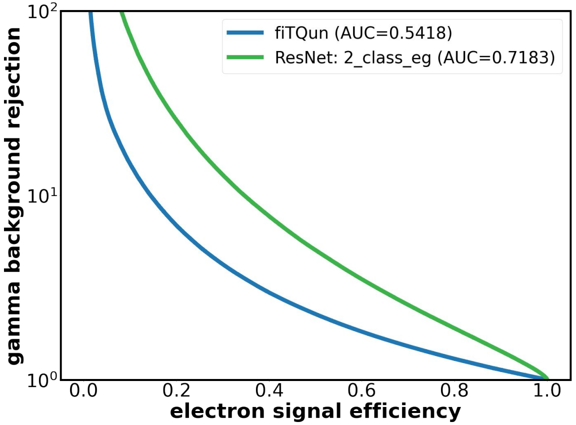


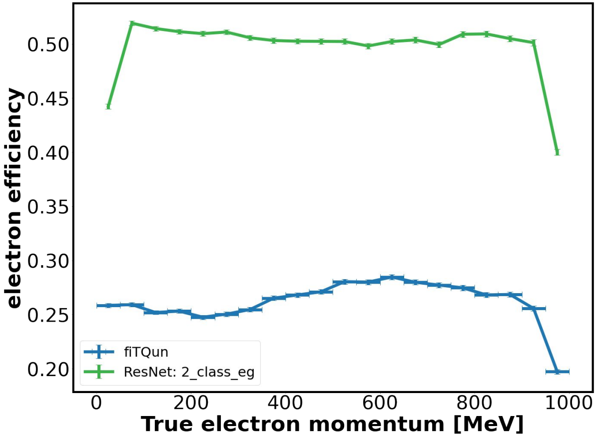
Figure: **Predicting electron as a true particle and electron being classified as γ**

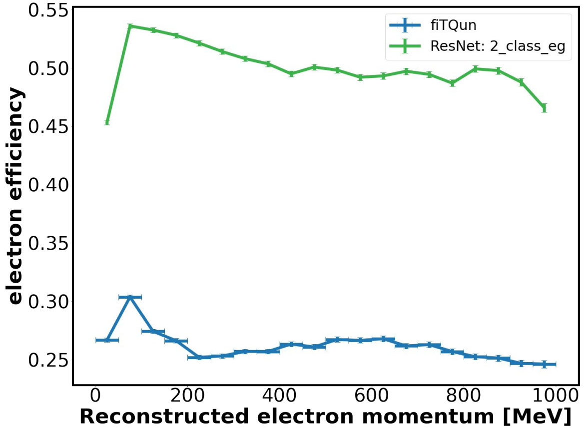
- **ROC curve: (2 Class)**

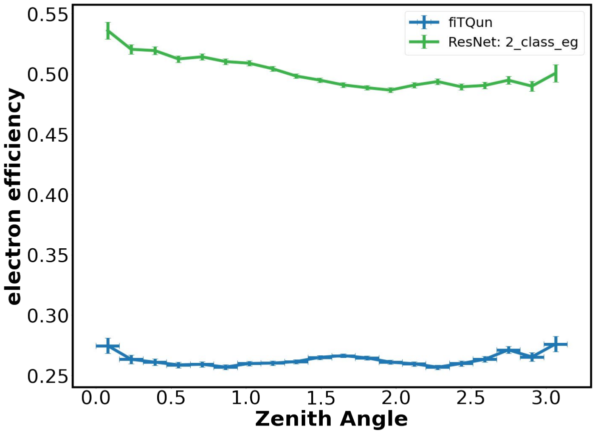
Comparing ResNet and fiTQun performance in e^-/γ identification:

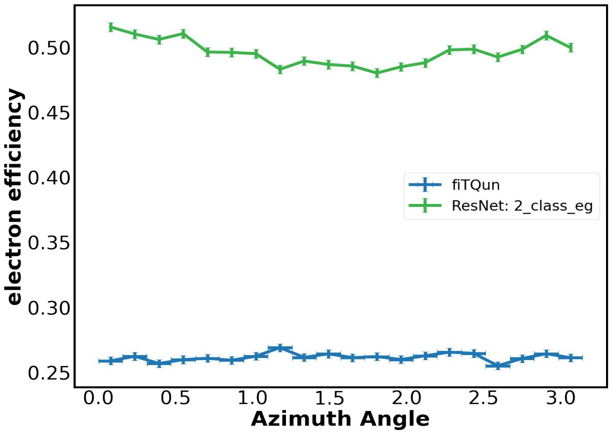
(fiTQun AUC: 0.5418, ResNet AUC: 0.7183)



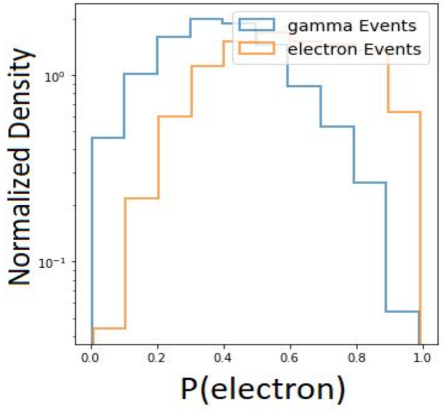








Results



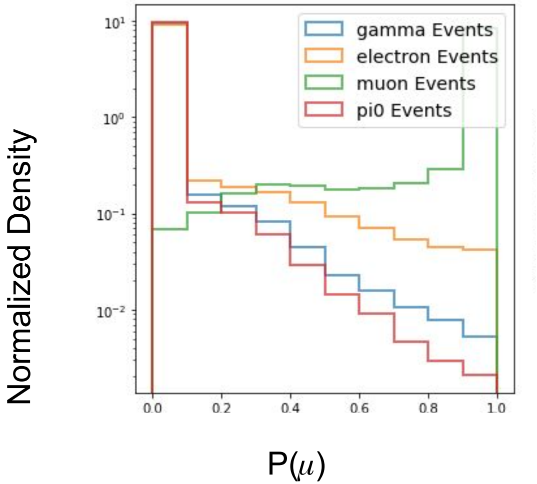
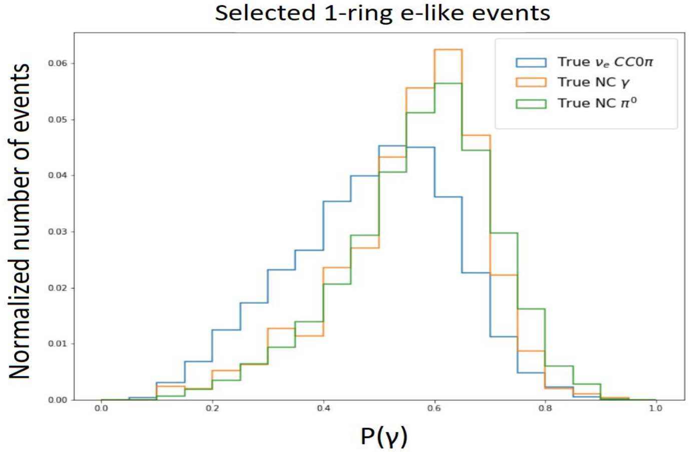


Figure: Distribution of particles for 2-class and 4-class analysis

- Using ML techniques, NC components are already well separated from signal for IWCD preliminary geometry



Ref. A.Oshlianskyi,'Electron neutrino analysis for IWCD long tank geometry for Hyper-Kamiokande experiment'

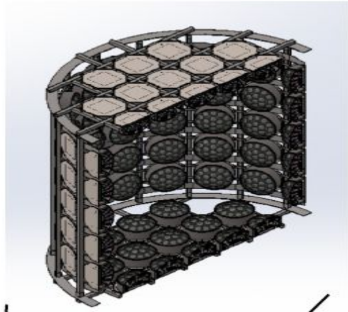
- Signal efficiency and background rejection for fiTQun and Softmax cuts

Fraction	True ν_e $CC0\pi$	True ν_e CC $other$	True NC	True NC γ	True NC π^0	True $\nu_{e,ws}$	True $\nu_{\mu,ws}$	True ν_μ
Softmax/fiTQun	1.018306	0.991455	0.812033	1.054341	1.100492	1.034594	0.0	0.595286

- **ML techniques will be applied to IWCD short tank data for signal-background separation**
- **Potential for tagged γ beam at WCTE to verify e/γ discrimination performace.**

Water Cherenkov Test Experiment (WCTE)

- A 50 ton scale proposed water Cherenkov detector which will operate at CERN T9 beam.
- Part of Hyper-K development program for the IWCD
- Study detector calibration & response with known particle fluxes of 0.2 GeV/c - 1 GeV/c



Current Geometry

Cylinder Height ~ 3.40m
Cylinder Radius ~ 1.90m
Number of mPMT's ~ 102

Physics Goals

- Test and evaluate IWCD performance with new technologies.
- Reduce Detector Systematics Error in Water Cherenkov Detectors.
- Develop better calibration techniques.
- Test Event reconstruction performance with mPMT's modules.

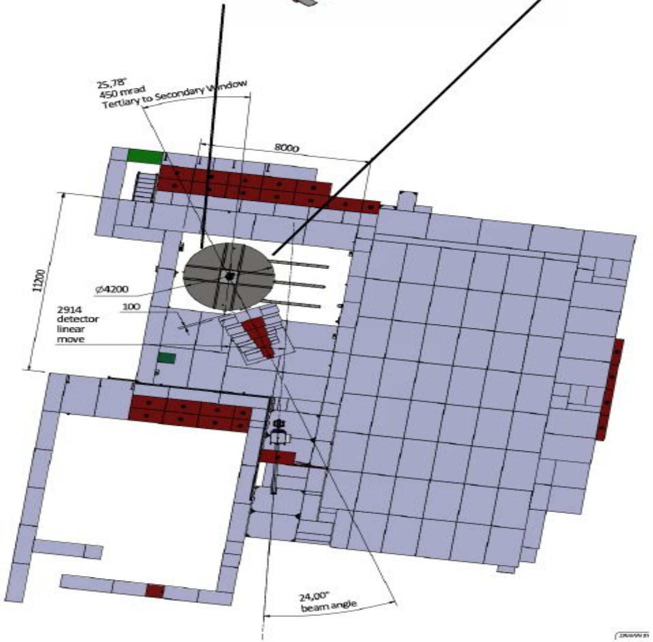


Fig. T9 CERN area

Detector Dimensions & Beam Configurations

- Reduction of detector dimension (to fit the wall and ceiling)
- Reconstruction performance remains unhampered

Config	Columns	Rows	Height (mm)	Diameter (mm)	ID height (mm)	ID diameter (mm)
Original	18	5	4320	4022	3539	3621
Reduced diam 1	18	5	4200	3800	3539	3439
Reduced diam 2 (16c-5r)	16	5	4200	3800	3539	3427
Reduced height and diam (16c-4r)	16	4	3400	3800	2739	3427

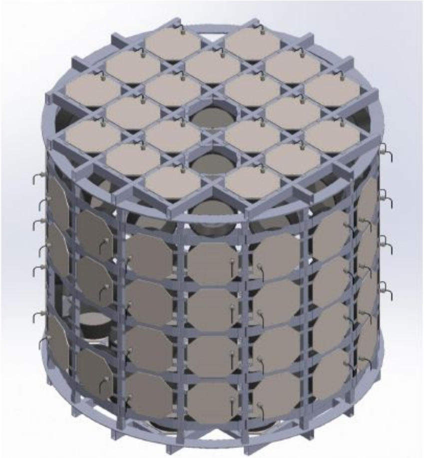


Fig. WCTE detector geometry

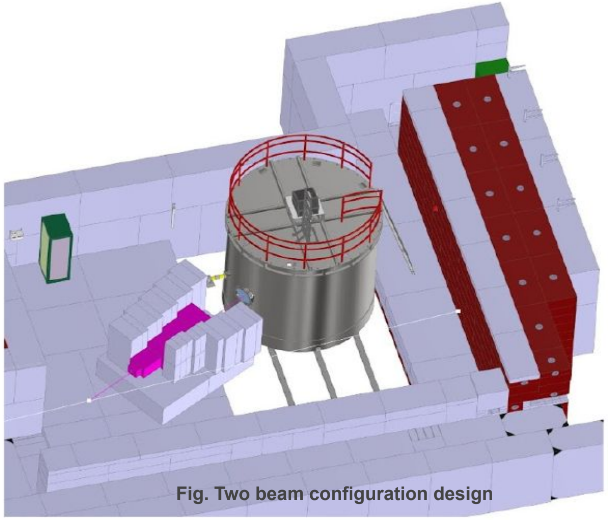


Fig. Two beam configuration design

WCTE will run with two beam configuration

- **Tertiary Beam (0.2 – 1.2 GeV/c)**
  - Access low momentum pion and proton fluxes.
- **Secondary Beam (~0.4 GeV/c to ~1.5 GeV/c)**
  - Detector is set in the beam line.
  - Access e^-, μ^- and proton fluxes.

Simulation of particles for WCTE

- WCTE/WCSim (https://github.com/laurenanthony2/WCSim)
- Using the Current geometry(4r,16c)

```
#Use mPMTs settings (uncomment/delete the above)
#/WCSim/WCgeom nuPRISM_mPMT
#/WCSim/WCgeom nuPRISMBeamTest_mPMT ## this is 18c5r from the original design
#/WCSim/WCgeom nuPRISMBeamTest_18c_mPMT ## this is 18c5r from CAD
#/WCSim/WCgeom nuPRISMBeamTest_16c_mPMT ## this is 16c5r from CAD
/WCSim/WCgeom nuPRISMBeamTest_16cShort_mPMT ## this is 16c4r from CAD
#/WCSim/WCgeom nuPRISMShort_mPMT
```

Using the WCTE geometry

- A visualization of the run is shown in the figure.
- Run simulation for 1 million $\mathbf{e^-}$ and $\boldsymbol{\mu^-}$ events.
  - Energy uniformly spread over 0 - 1000 MeV
  - Point of origin uniformly spread over the detector.
  - Uniformly spread over φ and cosθ.
- Conversion of *wcsim.root* files to *.npz*
- And *.npz* file to *.h5* file for direct use in Machine learning.

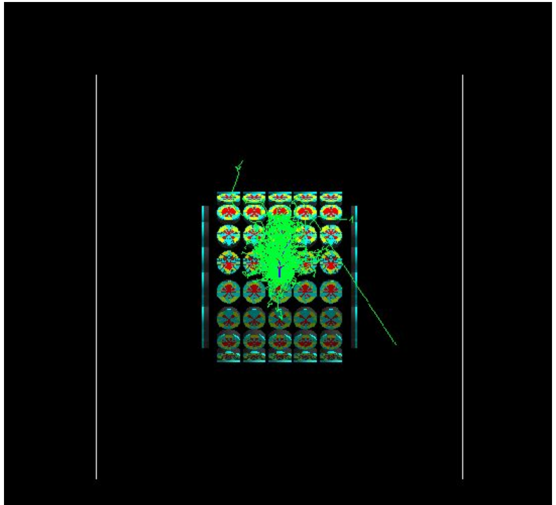


Fig - Visualization of event run

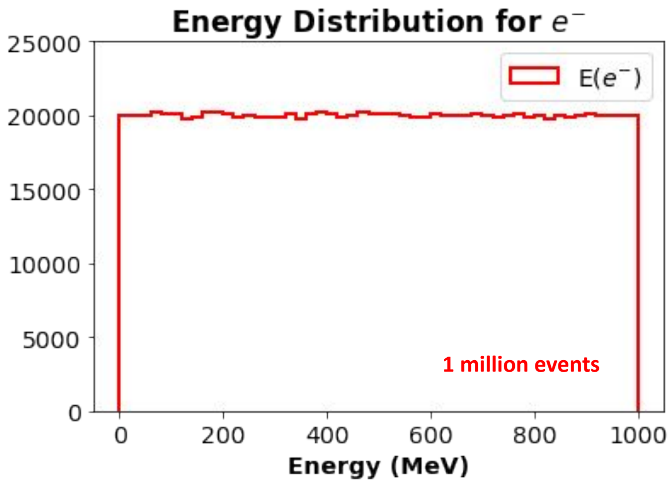
Current Progress with WCTE

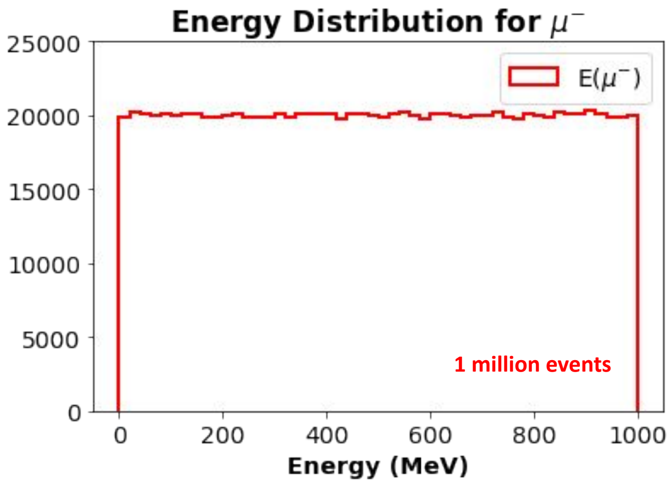
- Data generation of 1million **e**$^-$ and **μ**$^-$ events is completed.
- Simulated events- used for Machine Learning(ML) training.
- Development of the ML pipeline.

Data Exploration: Validating the events generated from the *.h5* datafile

➢ Energy Distribution Plots

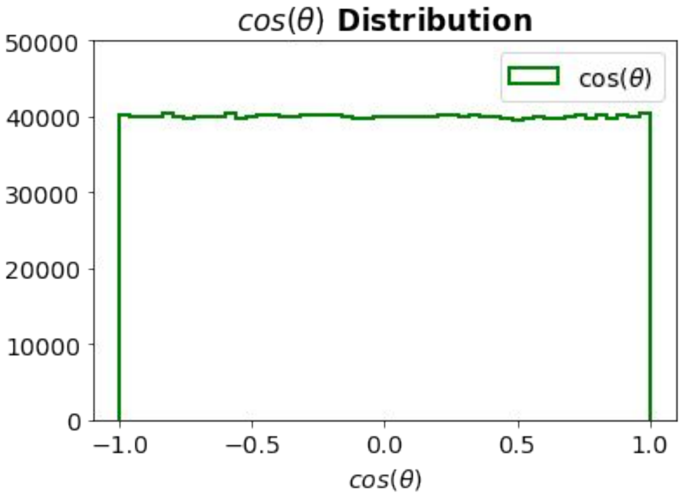
- Generated **e**$^-$ and **μ**$^-$ events (**Each 1 million**)
- Uniformly distributed Energy between **0 - 1 GeV**

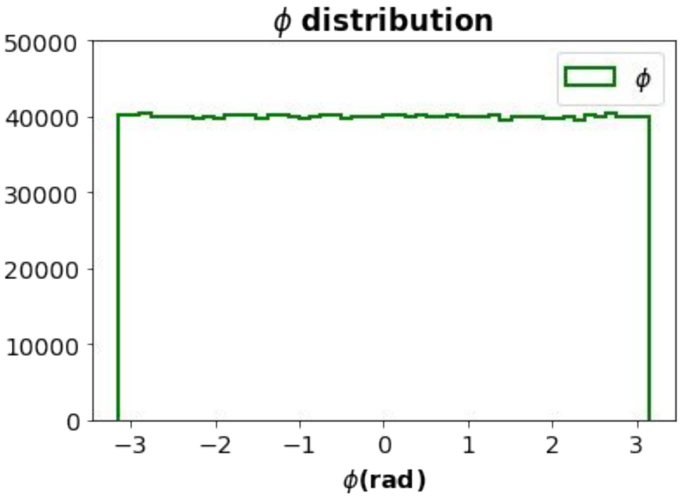




➢ Angle(θ,φ) Distribution Plots

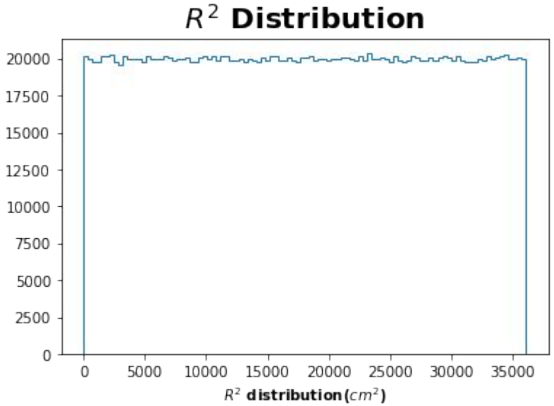
- Events are isotropically distributed over θ
- Uniformly distributed over φ

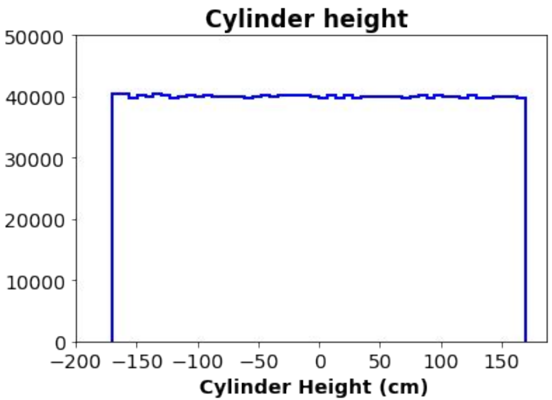


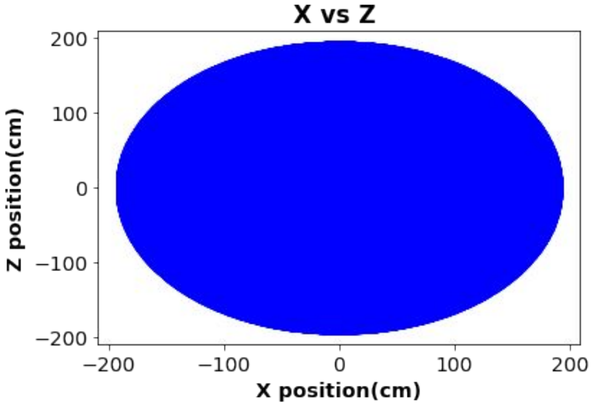


➢ Position(x, Height, z) Distribution Plots

- Point of origin is uniformly distributed over the cylinder height.
- **R^2** is uniform from the simulation
- Uniformly distributed over XZ-plane







ML Pipeline Preparation

- Pipeline will help in better implementation of the ML model.
- Initial steps
  - Preparing the Data.
  - Mapping PMT's in 3D detector to a 2D Image.
  - Start the ML model building.

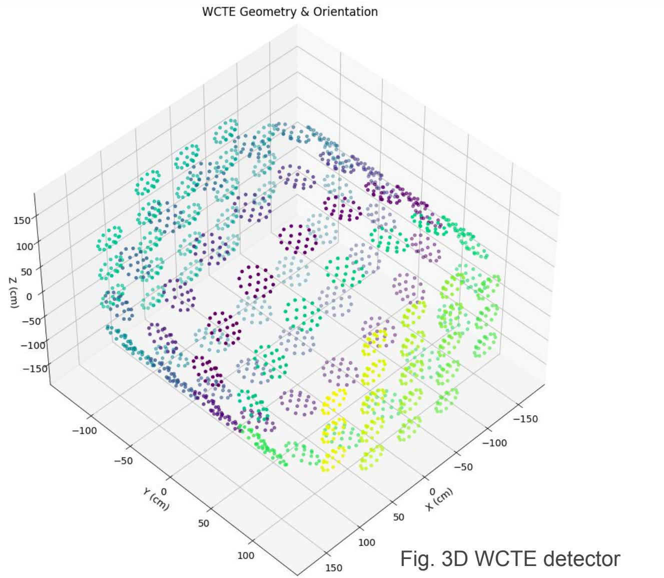


Fig. 3D WCTE detector

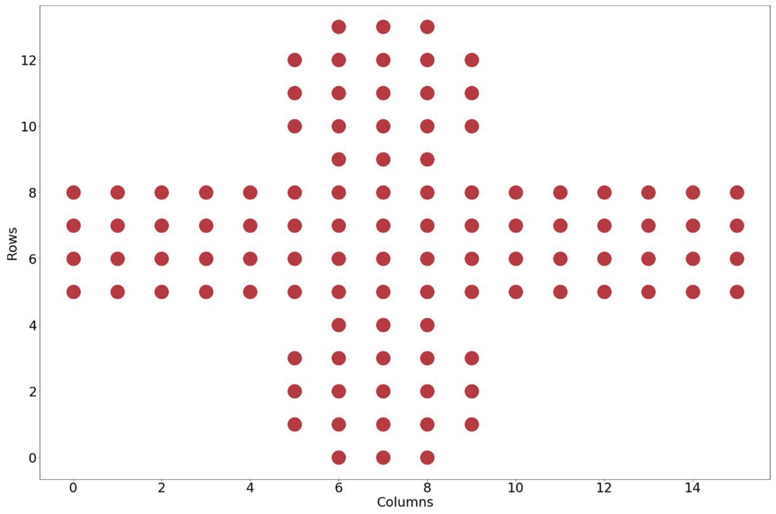


Fig. 2D mapping of detector

Future Plan of Work

IWCD

- ❑ Analyse fiTQun particle identification techniques over current IWCD production.
- ❑ Produce new training sample based on IWCD new detector geometry, to train ML ResNet Model.
- ❑ Eventually applying ML PID techniques to IWCD event selection.

WCTE

- ❑ Development of ML data pipeline.
- ❑ Initiate ML training with 1 million e^- and μ^- events data.
- ❑ Finally, apply Machine Learning Algorithms
  - ➔ For Event Reconstruction
  - ➔ Particle Identification Analysis.

Acknowledgement

Akira Konaka

Debanjan Bose

Lauren Anthony

Moon Moon Devi

Nick Prouse

Patrick De Perio

Reetanjali Moharana

Ryosuke Akutsu

Xiaoyue Li

Thank You for your attention!