Likelihood and Deep Learning Analysis of the electron neutrino event sample at Intermediate Water Cherenkov Detector (IWCD) of the Hyper-Kamiokande experiment

42nd International Conference on High energy Physics (ICHEP), 2024

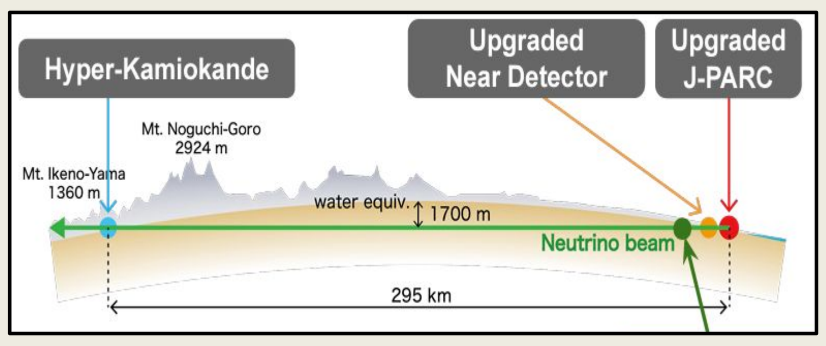
Tanima Mondal*, of the Hyper-Kamiokande Collaboration

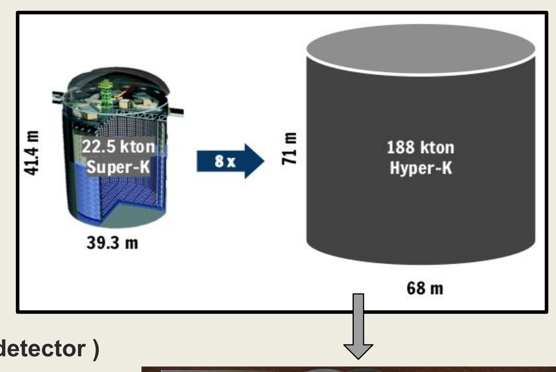
*Ph.D Research Scholar, Dept. of Physics, IIT Kharagpur, India

Introduction

Hyper-Kamiokande Experiment

- Hyper-Kamiokande (HyperK/ HK) [1] is the next generation underground water Cherenkov neutrino experiment, a successor of Super-Kamiokande (SK).
- Far detector ~8x SK Fiducial Volume → Higher statistics (20x times the rate of T2K) → Upgrade of neutrino beam (750 kW → 1300 kW).





(Intermediate Water Cherenkov detector)

HK - Water Cherenkov detector

Physics Goals –

1. Measuring matter asymmetry in the universe through measuring CP violation (CPV) in the neutrino sector —

$$P(\nu_\mu \to \nu_e) \neq P(\bar{\nu}_\mu \to \bar{\nu}_e)$$

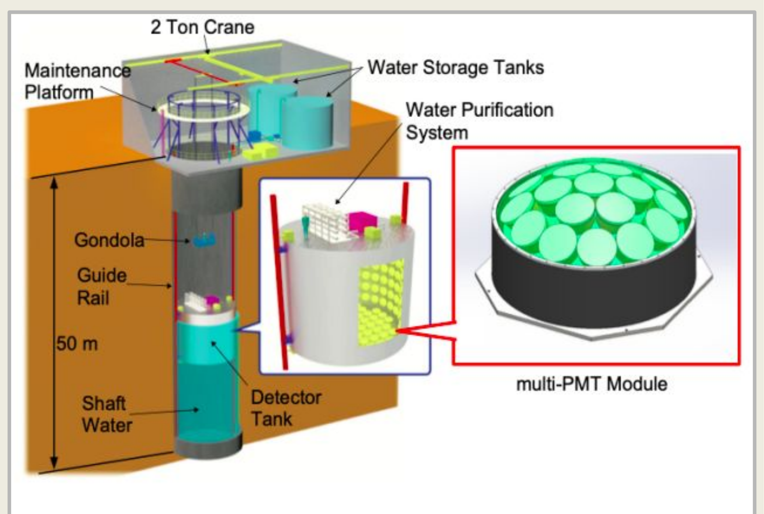
2. Study neutrino oscillation - through observing accelerator, atmospheric, solar, supernova and astrophysical neutrinos, nucleon decay searches.

Intermediate Water Cherenkov Detector

- An intermediate detector, located ~ 750 m from source.
- Measures of flux and cross-section of mostly un-oscillated beam.
- 600 ton water Cherenkov detector (WCD), moves vertically in ~50 m tall pit.
- Spans range of angles (1°- 4°)[1] off-axis from ν beam for different ν energy spectra.
- High-resolution multi-PMT photon detectors(~ 500 multi-PMT modules (mPMTs), an array of 19 smaller 8 cm diameter PMTs).

IWCD Detector Geometry

Inner Detector (ID)	Current geometry
ID Radius	400 cm
ID HalfLength	300 cm
Baseline	~750 m from source



Approved Hyper-K project includes IWCD

Goals —

➢ Study neutrino interaction rate peaked at different energies, higher precision.

➢ To determine ~ 1% of anti(neutrino) ν_e components in the beam, to have improved event selection.

Aim of this work —

→ Analysing Particle Identification (PID) performance for IWCD current geometry to study ν_e event samples using fiTQun → Improving sensitivity and precision of ν_e cross-section measurements.

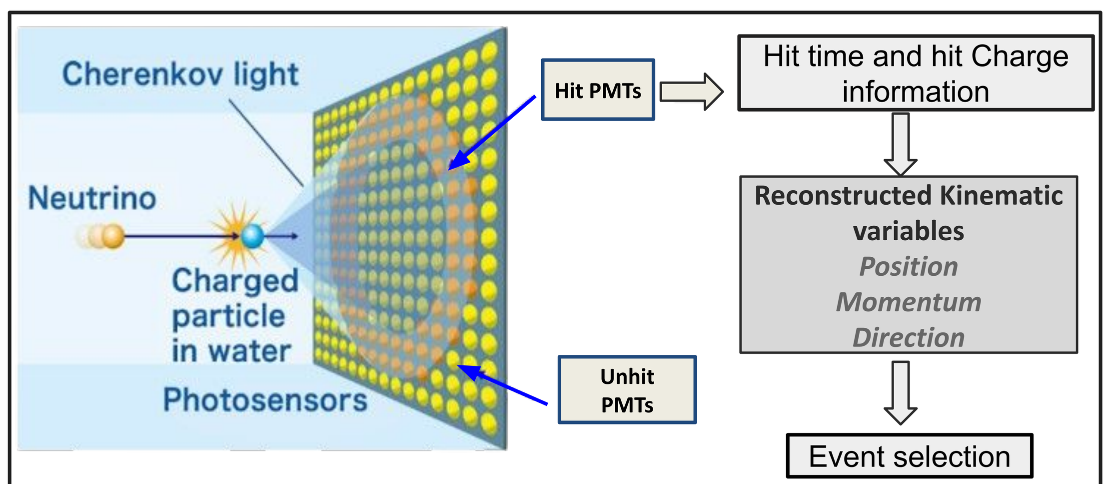
→ Building up a new Machine Learning (ML) analysis framework for IWCD event selection improving both purity and efficiency of the signal over fiTQun.

Constraints —

➢ Conventional neutrino beam contains only 1.5% of ν_e, challenging to measure its cross-section.

➢ Intrinsic flux of $\nu_e/\bar{\nu}_e$ produce single electron sample, use to constrain $\sigma_{\nu_e}/\sigma_{\nu_\mu}$, $\sigma_{\bar{\nu}_e}/\sigma_{\bar{\nu}_\mu}$ → CP violation search

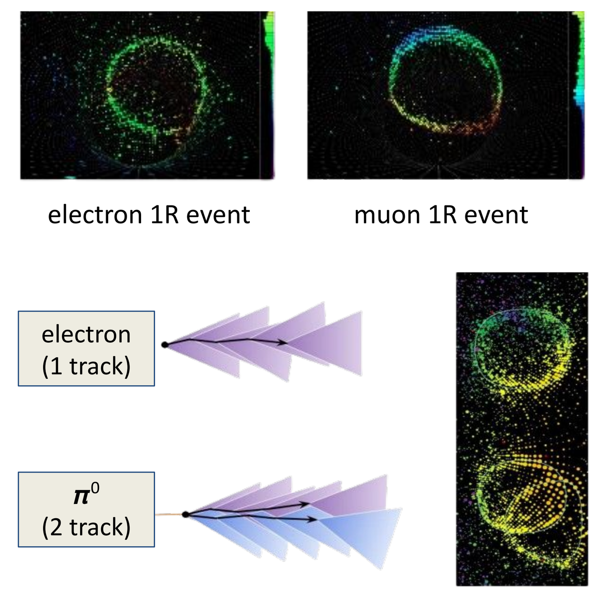
Water Cherenkov Detector and Particle Identification



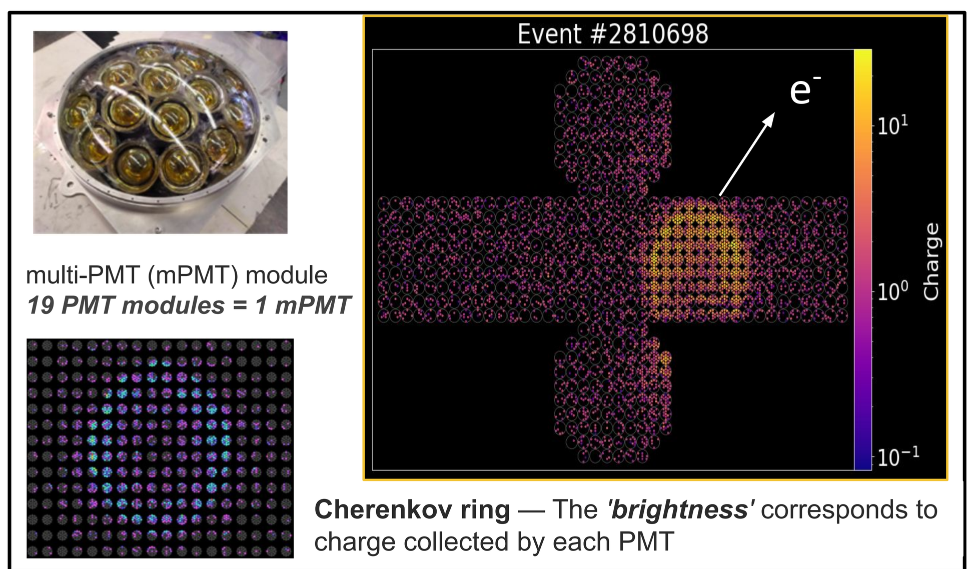
- Inside WCD, ν's are detected either through CC or NC interaction.
- Hyper K far detector aiming to measure ν_μ to ν_e (anti-neutrino).

Main Signal — $\nu_e CC0\pi$; **Backgrounds**— NCπ^0, ν_μCC events, entering γ

Particle Identification



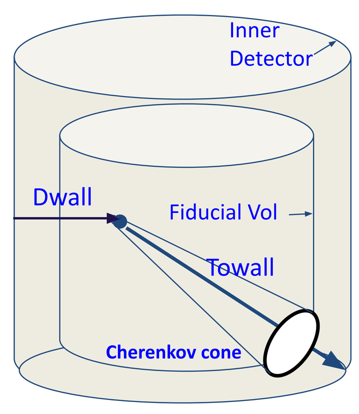
mPMT Data Vizualization



multi-PMT (mPMT) module
19 PMT modules = 1 mPMT

Cherenkov ring — The ***'brightness'*** corresponds to charge collected by each PMT

Reconstruction Parameters



Fiducial volume cuts for ν_e has strong impact on PID accuracy —
Dwall cut > 100 cm
Towall cut > 100 cm

Kinematic variable cuts for ν_e — ***Reconstructed momentum*** > 100 MeV/c

Other cuts – Muon containment, One-ring candidate and subevent cuts

IWCD Event Reconstruction with fiTQun

fiTQun for event reconstruction

- fiTQun → Likelihood-based reconstruction software for WCD.
- Above a particular Cherenkov Threshold, a particle can be identified as electron, muon.

➢ For a given event hypothesis Γ' (e^-, μ^-, π^0), Likelihood function L can be maximized varying its kinematic parameter θ' —

$$L(\Gamma', \theta') = \prod_j^{unhit} P_j(unhit|\Gamma', \theta') \prod_i^{hit} \{1 - P_i(unhit|\Gamma', \theta')\} \times f_q(q_i|\Gamma', \theta') f_t(t_i|\Gamma', \theta')$$

Event Selection —

- Samples enriched with ν_e interactions, mostly contained within detector's fiducial volume.
- Detector covers off-axis angle of 2.39°, with exposure of 10^{17} Proton on Target (POT) in Forward Horn Current (FHC) mode.

Purity, efficiency of signal & event rates after cut selection

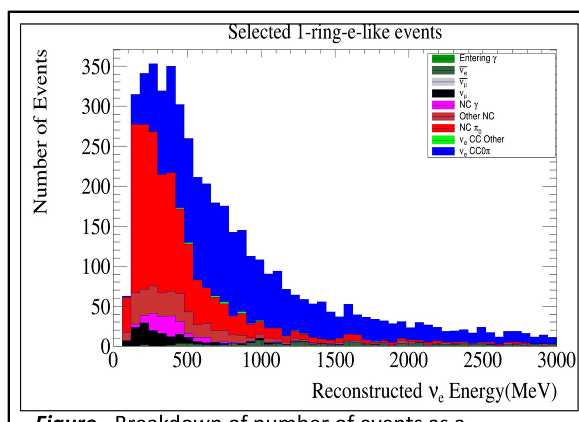


Figure - Breakdown of number of events as a function of reconstructed electron neutrino energy.

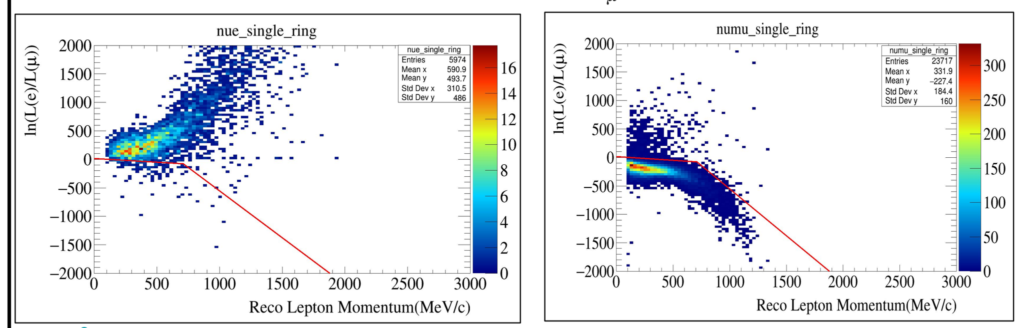
- Significant population of NC events are dominating at low reconstructed energy regime., specifically NC π^0 and NC γ.

Type of events	fiTQun cuts
ν_e CC0π	2535
ν_e CC other	24
$\bar{\nu}_e$	157
ν_μ CC	144
$\bar{\nu}_\mu$	1
Total NC	2100
NC π^0	1544
NC γ	121
Purity	**51.09%**
Efficiency	**69.47%**

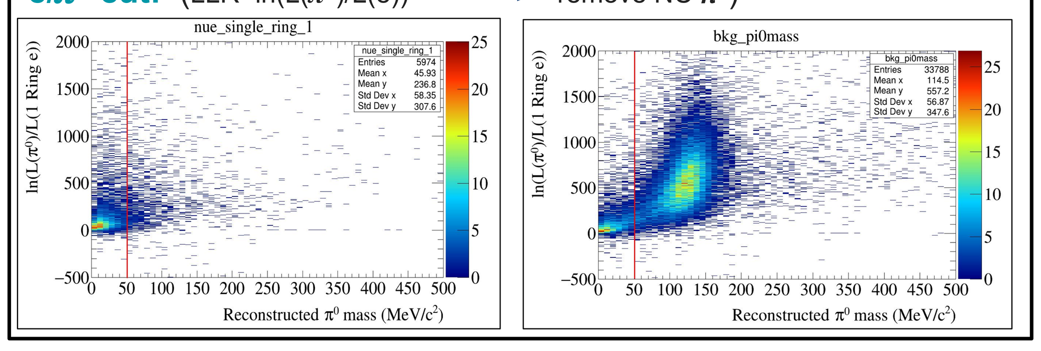
IWCD Cut selection —

Relative Likelihood-Method → To discriminate background muon and pions, and to achieve pure ν_e candidates.

e/μ cut: (LLR=ln(L(e)/L(μ)) → remove ν_μ CC)



e/π^0 cut: (LLR=ln(L(π^0)/L(e)) → remove NC π^0)

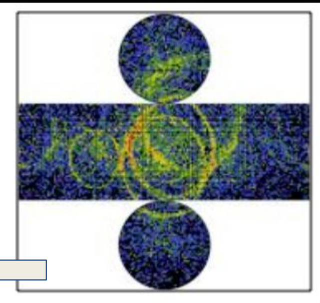


Limitations of fiTQun and importance of Machine Learning in event selection

Limitations in discriminative power of likelihood method:

γ rays from π^0 can be misidentified as a single e⁻like ring. ($\gamma \to e^+ + e^-$) Can not identify any double ring events.

More computational time → fiTQun (CPU): ~ 1 event per minute.



➢ **ML contribution in reconstruction**

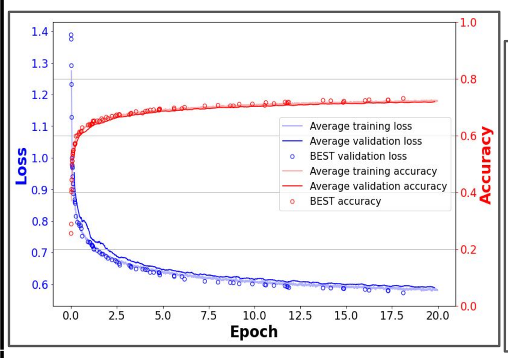
- Identification of particle types, and reconstruction of kinematic variables.
- Separating and reconstructing single and multi-ring events, Efficient computational time.

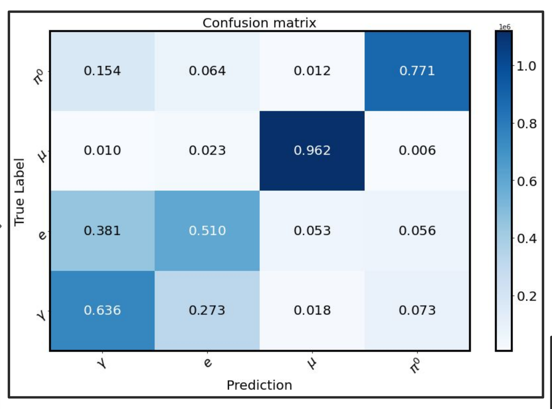
Machine Learning (ML) for Electron Neutrino Event Selection

ML training and application to classify IWCD particle-gun events

ML Simulation Dataset —

- 4 classes of particles ($e^-, \mu^-, \gamma, \pi^0$), 3 million of each events are generated.
- The kinetic energy is uniformly distributed upto 1 GeV above Cherenkov Threshold.
- Individual event data represents charge and time registered by PMTs in the IWCD detector.





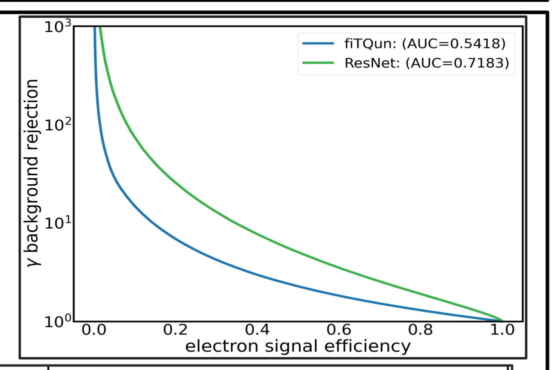
Model Performance

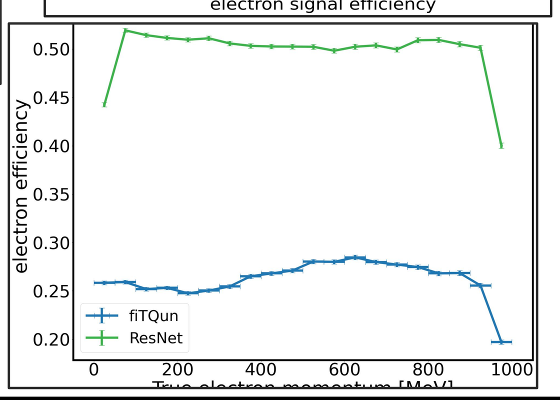
A) Confusion Matrix:

Perform classification by comparing model's predictions to the actual outcomes.

$P_e(e^-) = 0.510$ → 51%

$P_e(\gamma) = 0.381$ → 38.1%





B) ROC curve –

- Probability curve, measure performance at a threshold.
- fiTQun AUC: 0.5418, ResNet AUC: 0.7183
- ResNet models show significant gains over the traditional fiTQun model [3].

ML Training — *Convolutional Neural Network* -ResNet-18 model)

- Model is trained [2] for the equivalent of 20 epochs of 9 million events.

Event rates before and after ML cuts

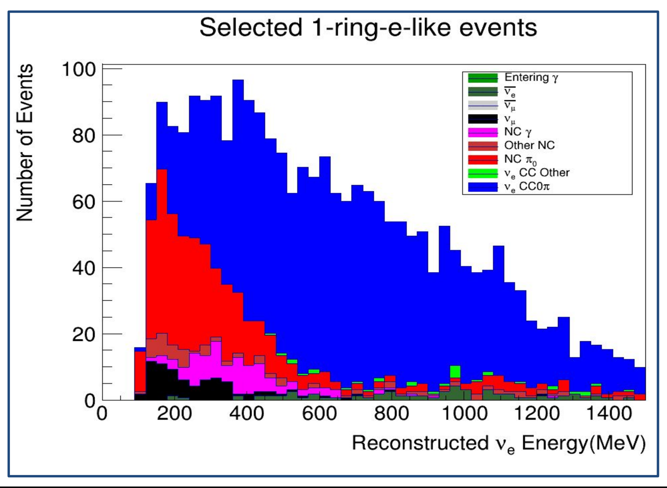
Cuts Applied	ν_e CC	ν_μ CC	NC	NC π^0	NC γ	ν_e CC other	$\bar{\nu}_e$	$\bar{\nu}_\mu$	Purity	Efficiency
Before ML cut	3653	14468	20650	16473	3534	137	345	40	0.09%	—
P(μ) cut	3559	368	19515	16336	133	120	345	3	9.66%	97.42%
P(π^0) cut	3222	81	1899	916	95	103	330	3	57.15%	90.56%
P(e) cut	2856	77	1342	627	93	79	288	0	61.53%	78.18%

Comparison - ML/fiTQun —

Ratio	ν_e CC	ν_μ CC	NC	NC π^0	NC γ	ν_e CC other	$\bar{\nu}_e$	$\bar{\nu}_\mu$
ML/fiTQun	1.13	0.54	0.64	0.41	0.77	3.29	1.83	0

- ***Purity*** percentage improved by ~ 20.43% with ML over fiTQun.
- ***Efficiency*** also improved ~ 12.53% with ML over fiTQun.

Breakdown of events after ML cuts

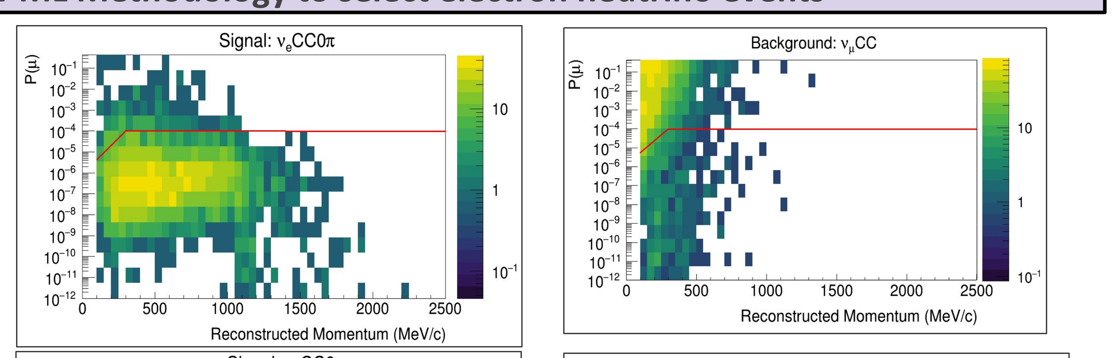


Application of new ML Methodology to select electron neutrino events

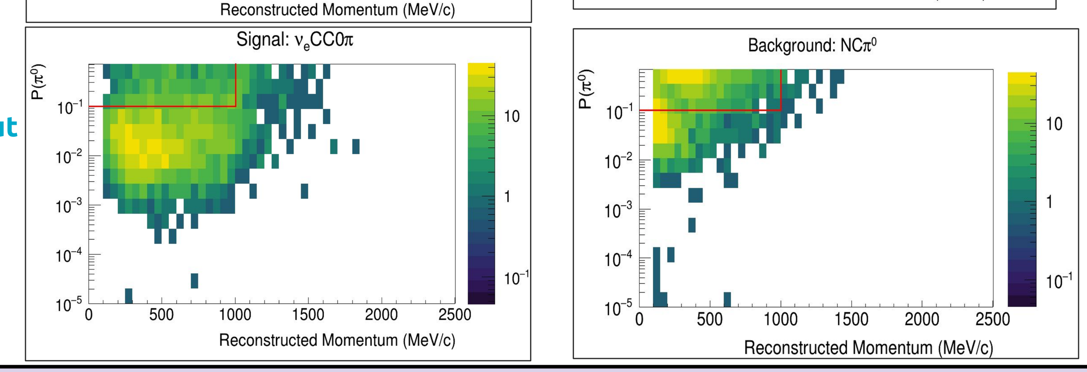
Methodology —

- Perform Model evaluation
  4 softmax variables as an output — P(e), P(μ), P(γ), P(π^0).
- Determined three suitable discriminator across P(μ), P(π^0) and P(e) softmax values.
- Compare fiTQun with ML.

P(μ) cut



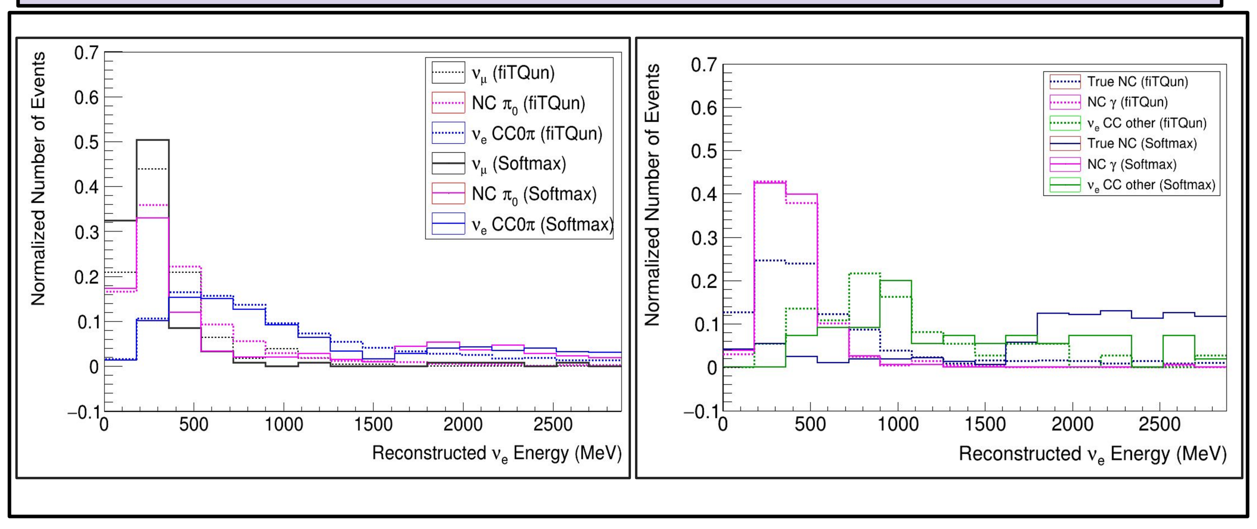
P(π^0) cut



Event rates after ML cuts

Type of events	ML cuts
ν_e CC0π	2856
ν_e CC other	79
$\bar{\nu}_e$	288
ν_μ CC	77
$\bar{\nu}_\mu$	0
Total NC	1342
NC π^0	627
NC γ	93
Purity	**61.53%**
Efficiency	**78.18%**

Visualization of Overall Comparison between fitQun/ML



Conclusions

- fiTQun cut variables reduces the background events significantly.
- However, newly developed analysis framework with ML techniques, which improved purity and signal efficiency over fiTQun event selection algorithm.
- ML softmax cuts significantly improved signal efficiency and purity over fiTQun.

Future Perspectives —

- Apply more complex and automated cuts with ML into our analysis to further improve the purity and signal efficiency.

References

1. Abe, K.; et al. Hyper-Kamiokande Design Report 2018. [arXiv:physics.ins-det/1805.04163], https://doi.org/10.48550/arXiv.1805.04163
2. WatChMaL. https://github.com/WatChMaL/WatChMaL. GitHub repository.
3. Prouse, N. et al. Machine Learning Techniques to Enhance Event Reconstruction in Water Cherenkov Detectors. NuFact 2022, https://doi.org/10.3390/psf2023008063

Poster Essence In Audio (Hear Me!)

mtanima14@gmail.com