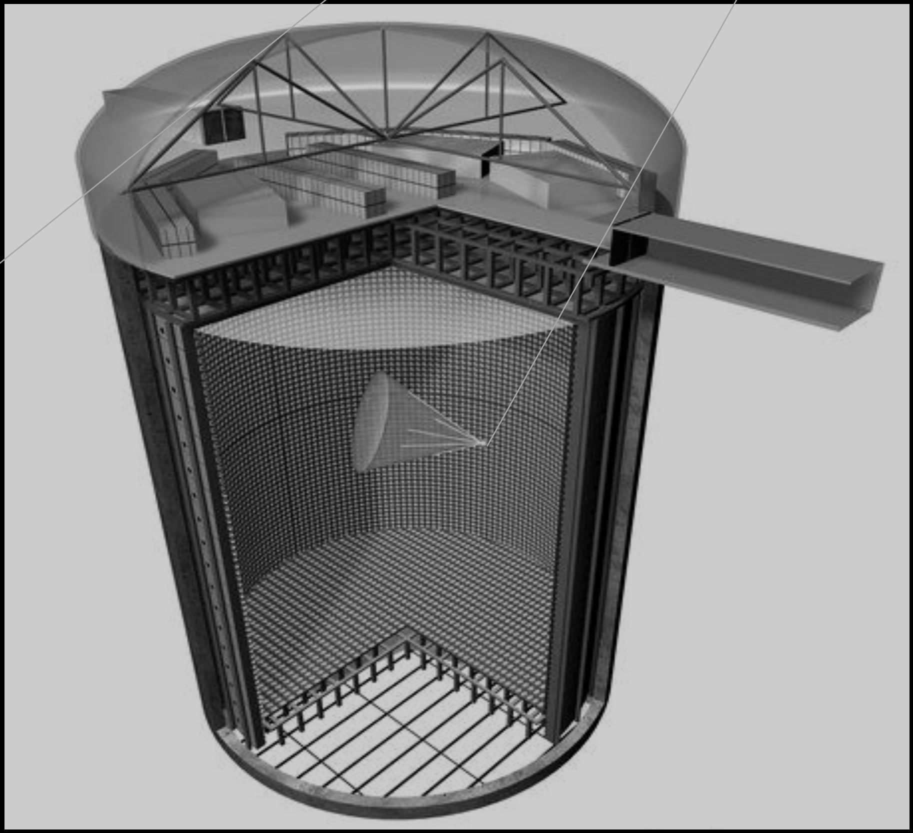
# スーパーカミオカンデ大気ニュートリノデータ解析におけるMLの利用

Roger Wendell
Kyoto University
ML@HEP 2022
2022.07.09



# Introduction

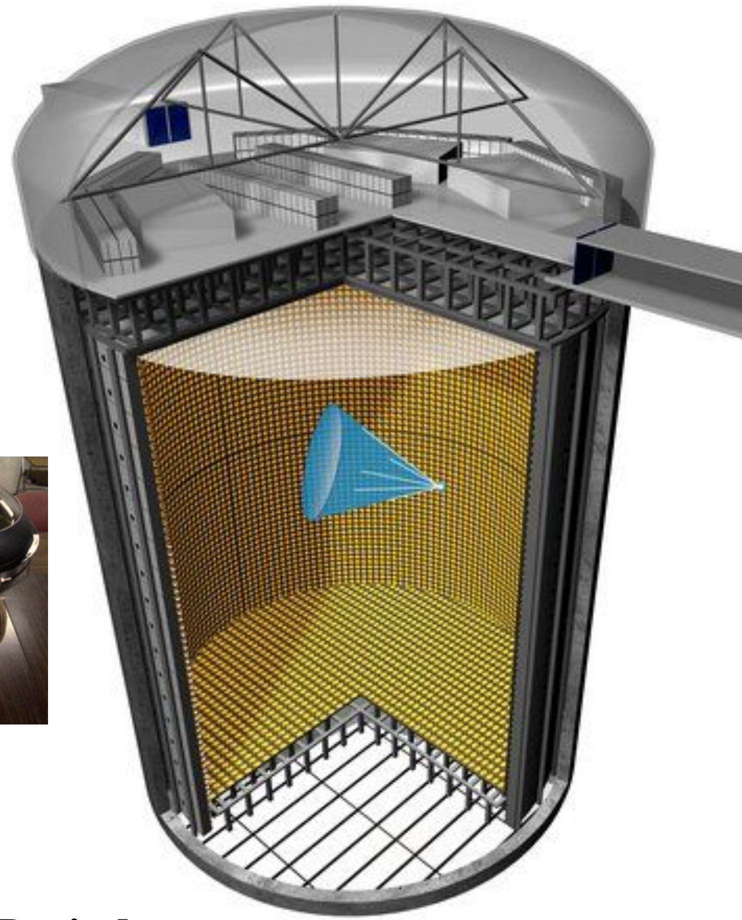
- Briefly explain recent applications of ML to analyses at Super-Kamiokande (SK) and perhaps give a few words about future prospects

- In general, SK has relied on zeroth-generation (?) techniques for separation of signals and backgrounds
  - Primarily TMVA-based neural networks, single-classifier problems

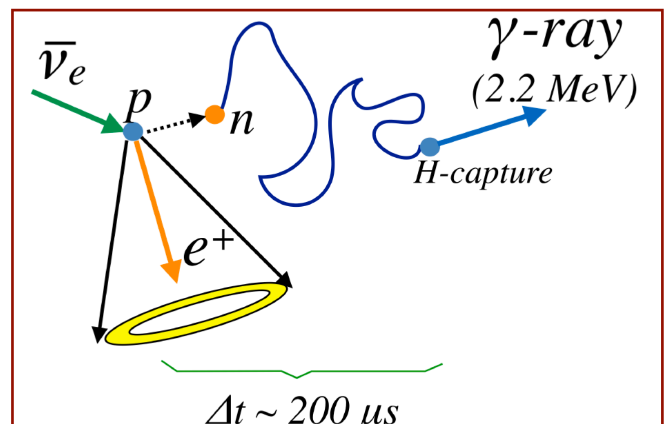
- Some exploration of other models, but no public results yet
  - In principle particle identification and kinematic reconstruction can also be done with ML
  - C.F. WatChMaL.org

- Topics
  - Tau Neutrino Appearance [~GeV]
  - Neutron Tagging [~**M**eV ]
  - Multi-Ring event classification
  - Boosted Dark Matter [~GeV]
  - Nucleon Decay Searches [~GeV]

# Super-Kamiokande:



- 22.5 kton fiducial volume
- Optically separated into
  - Inner Detector 11,146 20” PMTs
  - Outer Detector 1885 8” PMTs
- No net electric or magnetic fields
- Excellent PID between showering (e-like) and non-showering (μ-like)
  - →PMT hit pattern, timing, charges
  - < 1% MIS ID at 1 GeV
- Multipurpose physics




**Four Run Periods:**

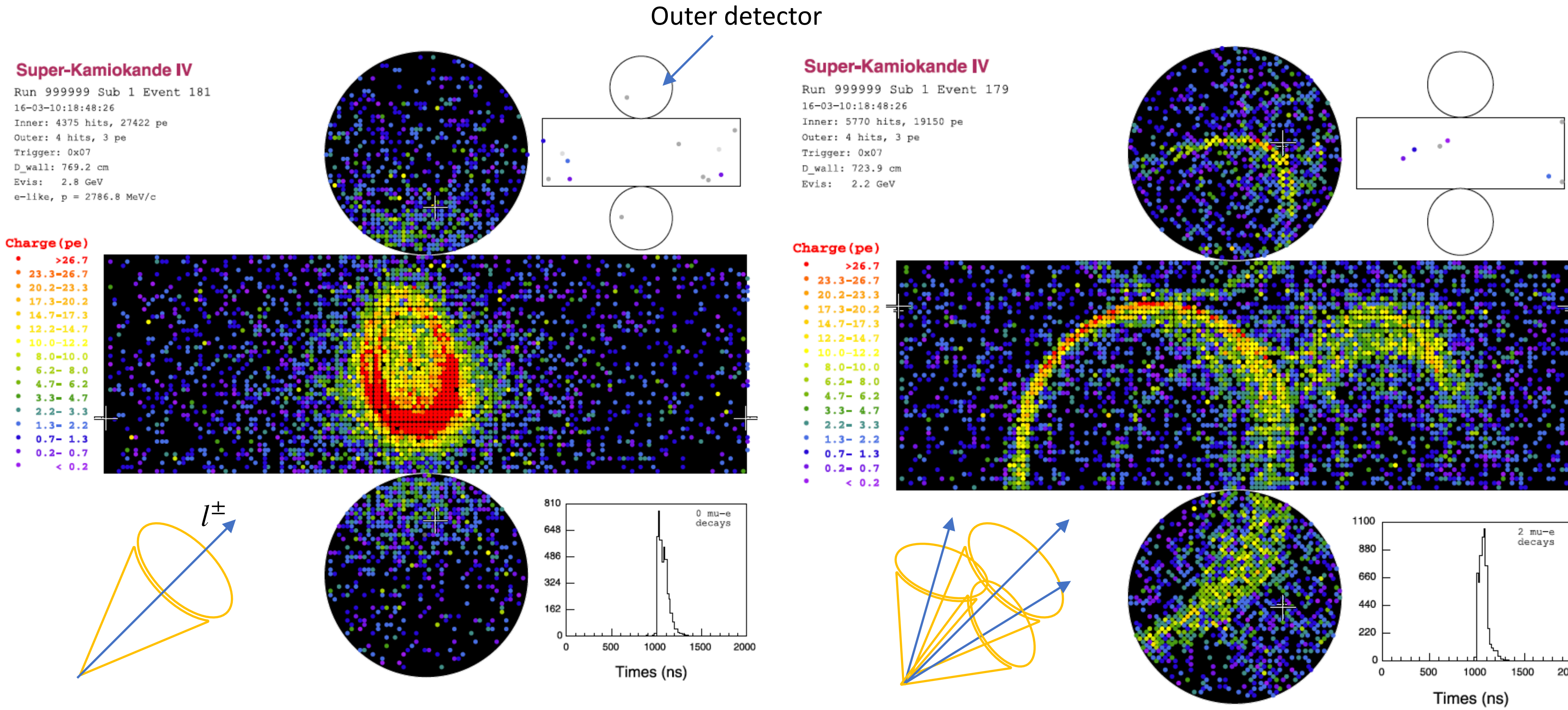
**SK-I (1996-2001)** **SK-II (2003-2005)**

**SK-III (2005-2008)** **SK-IV (2008-2018)**

**SK-V (2019-2020)** **SK-VI (2020-2022) [0.01% Gd]**

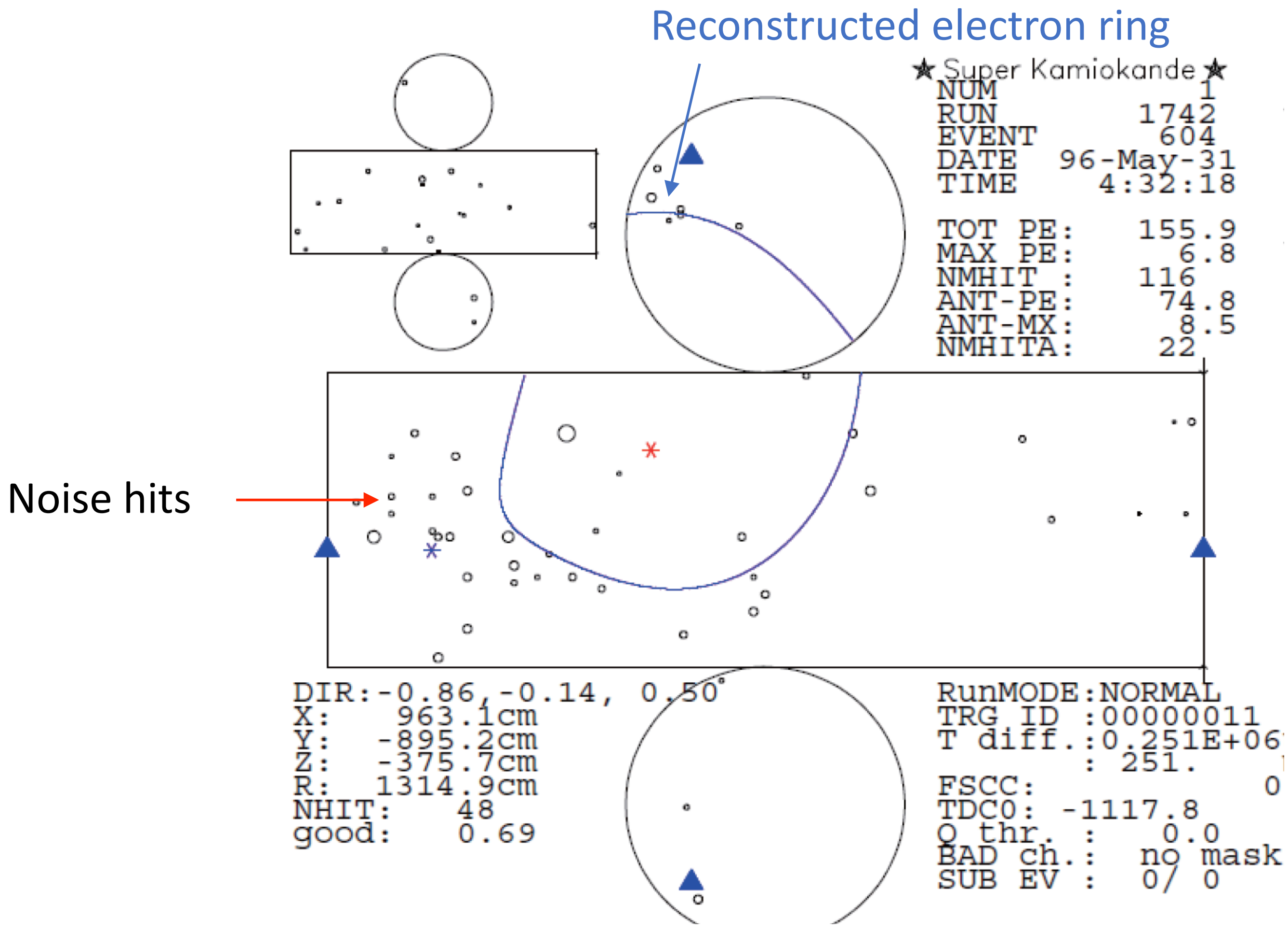
**Upgrade Complete Now operating as (SK-Gd)**

# The Fundamental Problem : E > 100 MeV



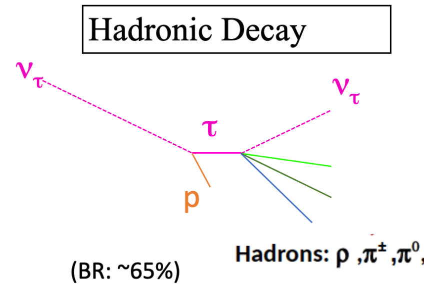

- No (3+1)-dimensional track information
  - All particle parameters estimated from (**2**+1) projection of (3+1) info.
- Overlapping particles confuse the reconstruction algorithm
- When energy deposition is small, competition between signal hits and noise

# The Fundamental Problem : E < 100 MeV

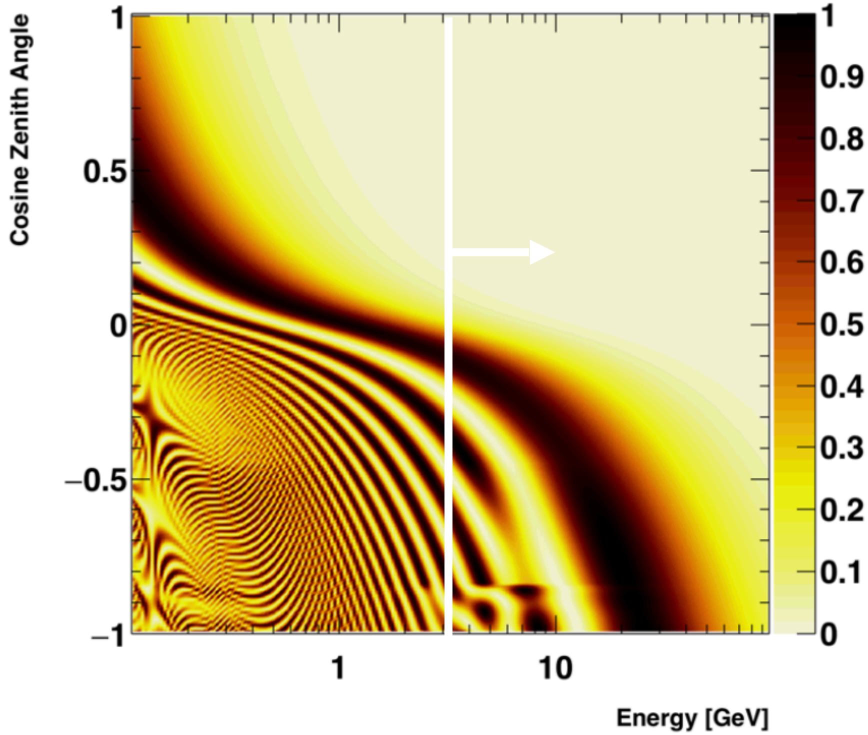



- No (3+1)-dimensional track information
  - All particle parameters estimated from (2+1) projection of (3+1) info.
- Overlapping particles confuse the reconstruction algorithm
- When energy deposition is small, competition between signal hits and noise

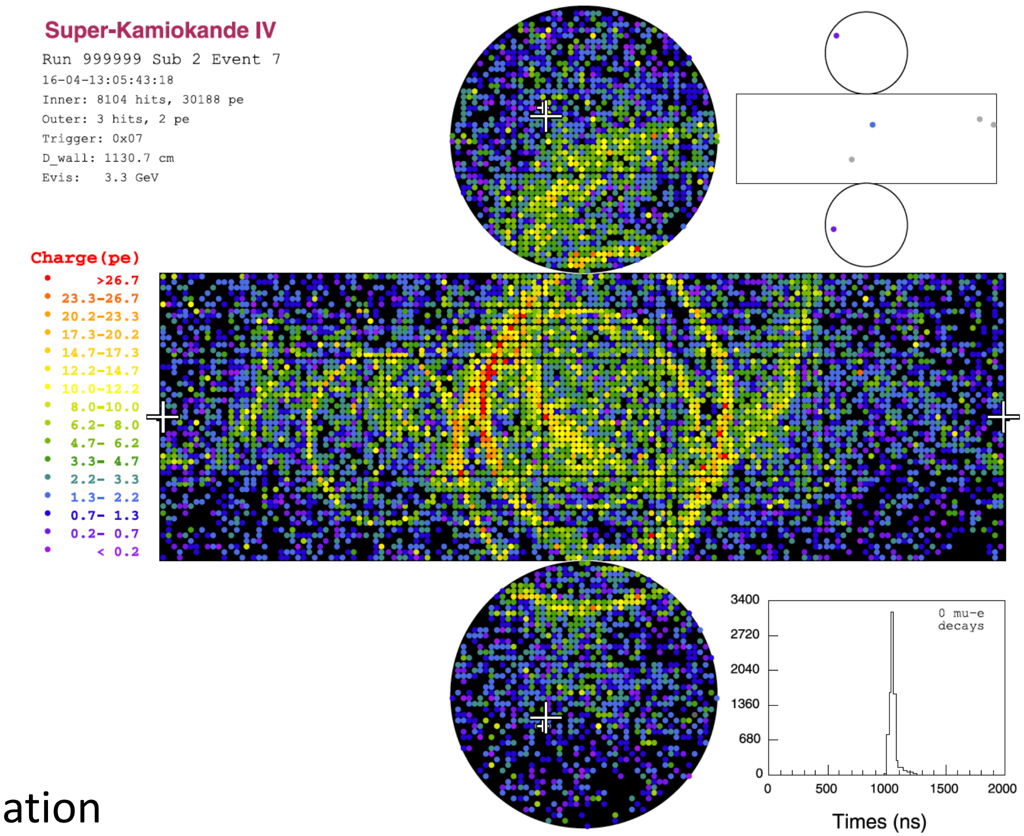
# Searching for $\nu_\tau$ : Prototypical ML




3 Flavor $P(\nu_\mu \to \nu_\tau)$

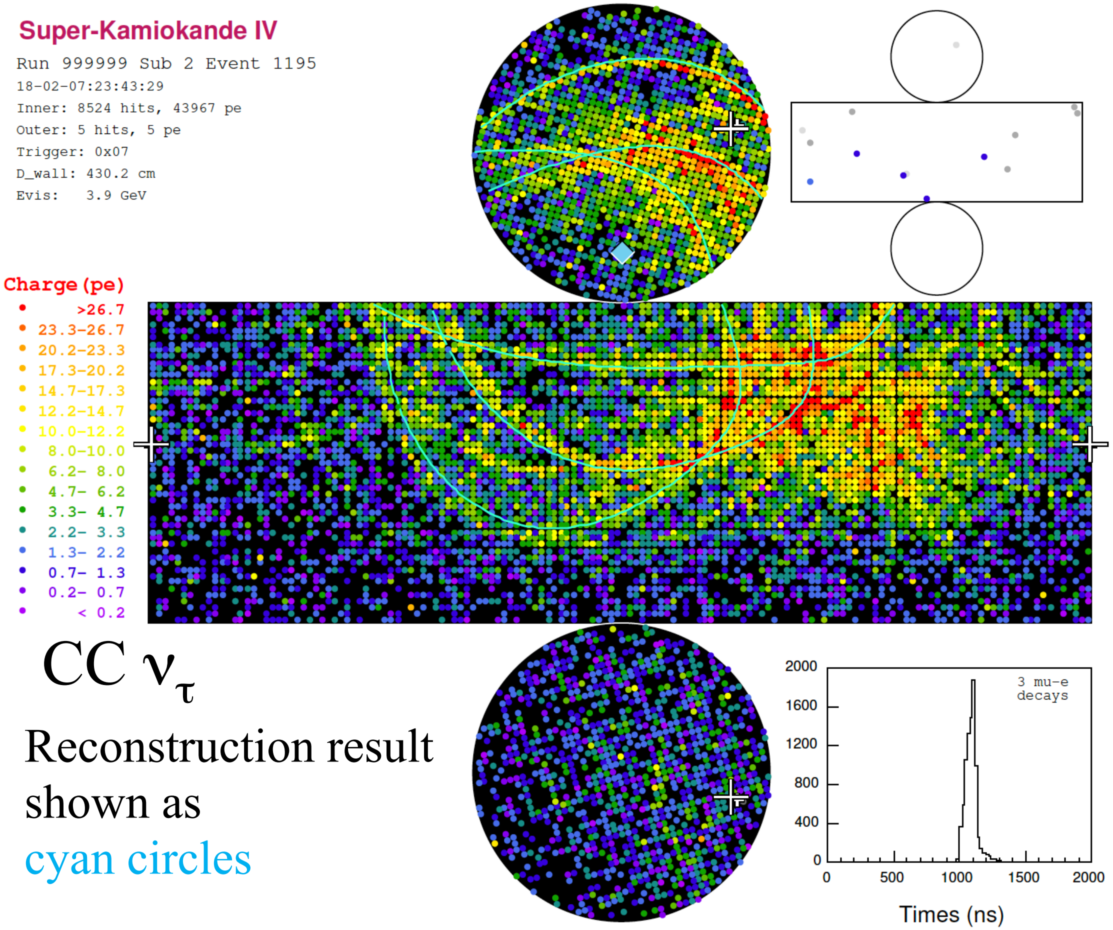



CC $\nu_\tau$




- Direct signal of atmospheric neutrino oscillation
- Large background for neutrino mass-hierarchy search
- → Complicated event topologies with many overlapping charged particles
- → Backgrounds $\nu_{\mu,e}$ DIS similar (but more forward) topology

# Search for Tau Neutrinos at SK: Selection



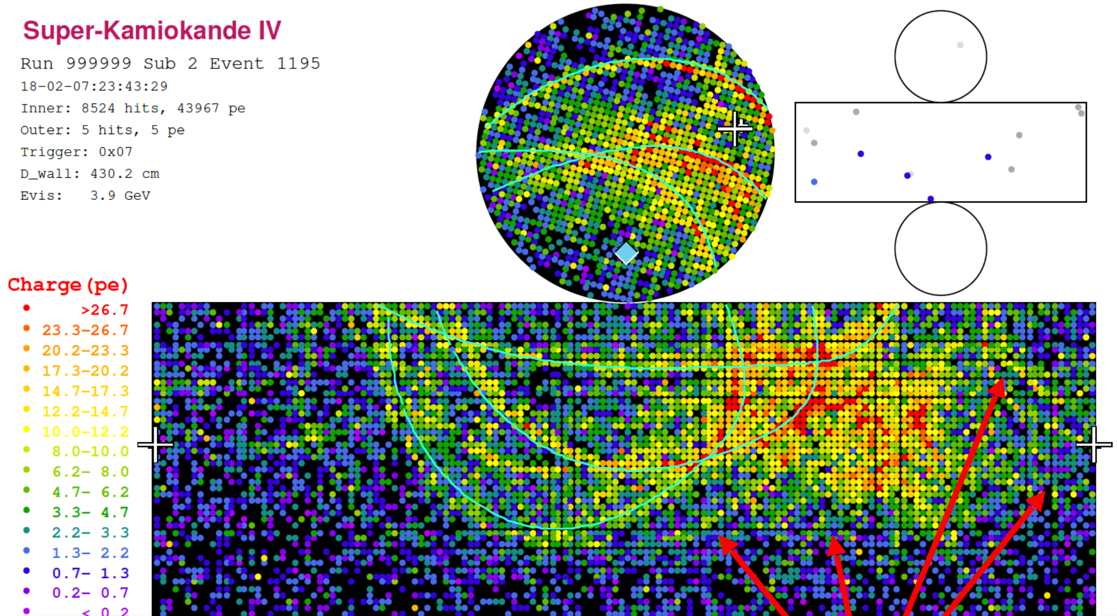

CC $\nu_\tau$

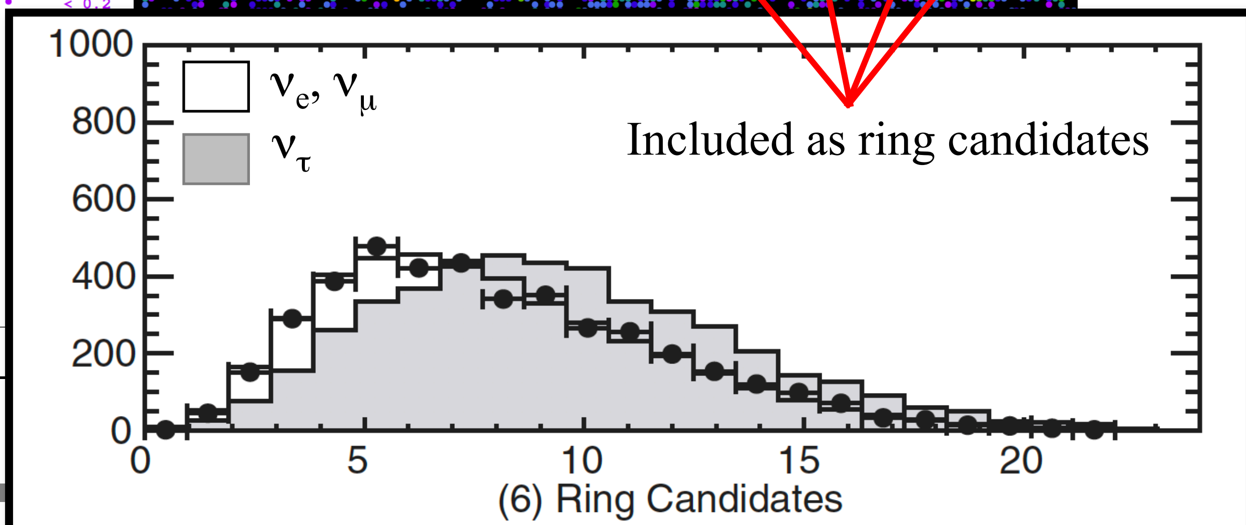
Reconstruction result shown as cyan circles

- Pre-selection
  - Fully-Contained
  - Fiducial Volume (22.5kton)
  - Visible Energy > 1330.0 MeV
- Seven Neural Network Variables
  - Log10 (Visible Energy)
  - Particle ID Likelihood
  - Number of Decay Electrons
  - Distance to farthest electron
  - Event sphericity
  - Number of Ring Candidates
  - Fraction of momentum in leading ring

- Difficult to fully reconstruct all particles, so count ring pieces
- Typically $\tau$ decay produces more pions above C threshold
  - Good problem for pattern recognition, CNN

# Search for Tau Neutrinos at SK: Selection



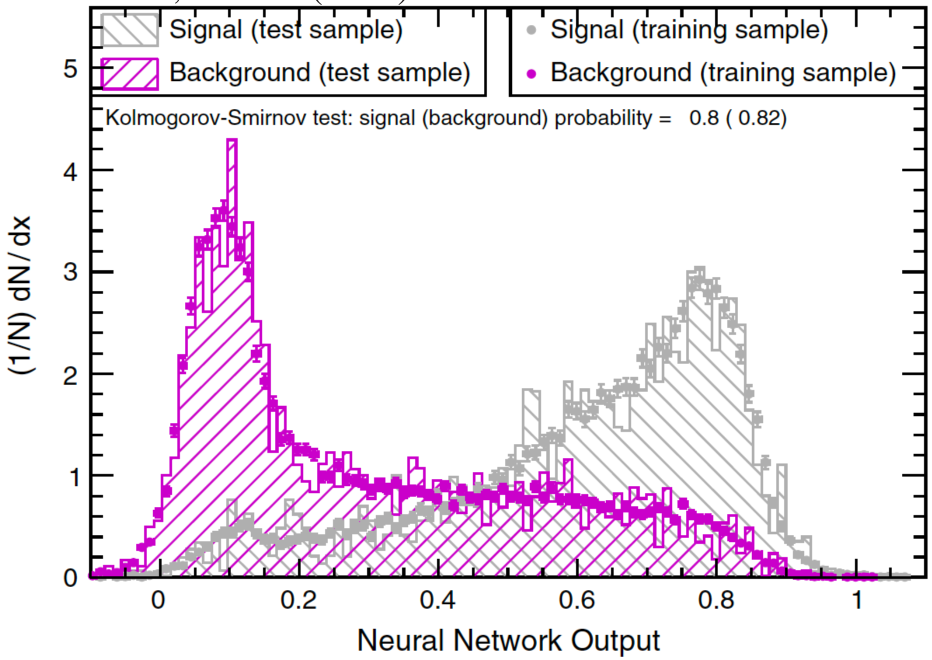




- Pre-selection
  - Fully-Contained
  - Fiducial Volume (22.5kton)
  - Visible Energy > 1330.0 MeV
- Seven Neural Network Variables
  - Log10 (Visible Energy)
  - Particle ID Likelihood
  - Number of Decay Electrons
  - Distance to farthest electron
  - Event sphericity
  - Number of Ring Candidates
  - Fraction of momentum in leading ring

- …unt ring pieces
- Typically τ decay produces more pions above C threshold
  - Good problem for pattern recognition, CNN

# Search for Tau Neutrinos at SK: Performance

PHYS. REV. D 98, 052006 (2018)

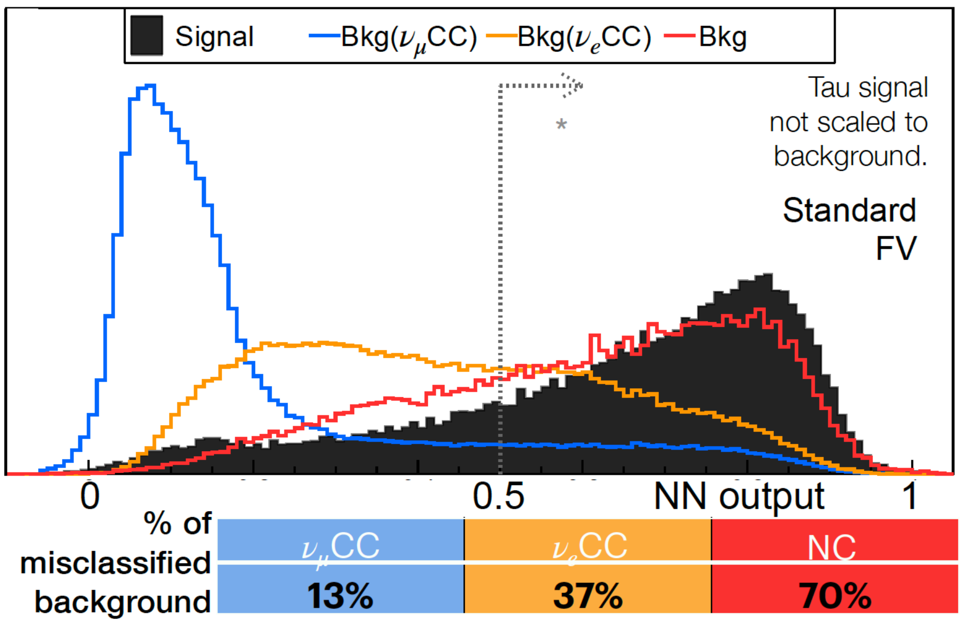


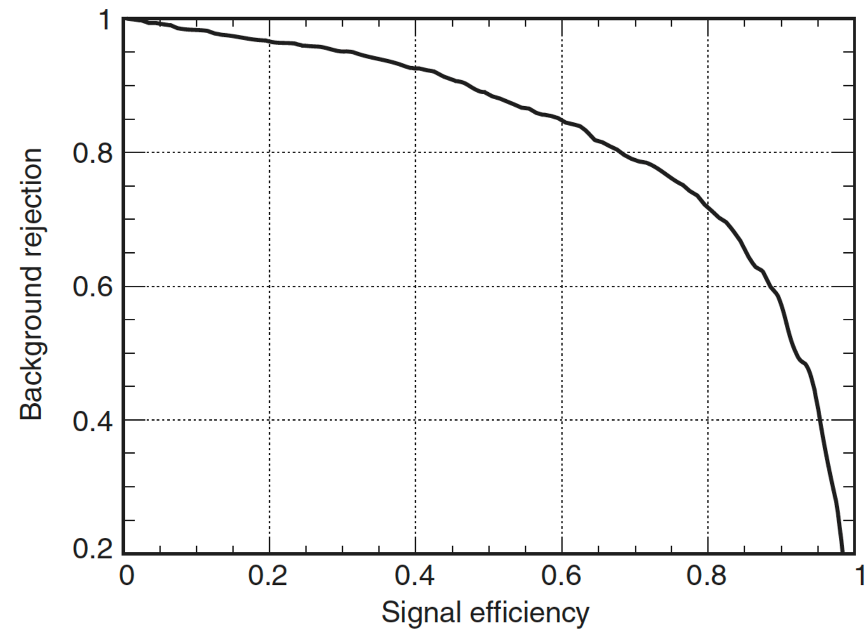

| NN > 0.5 | Background | Signal ($\nu_\tau$) |
|---|---|---|
| Efficiency | 28% | 76% |
| Purity | 95.3% | 4.7% |
| Rate [Mton· year]$^{-1}$ | 8467 | 422 |

**ROOT TMVA**

- MLP
- 7 Node Input Layer
- 10 Node Hidden Layer
- 1 Node Output
- Mean Square Estimator
- Back Propagation

# Search for Tau Neutrinos at SK: Performance



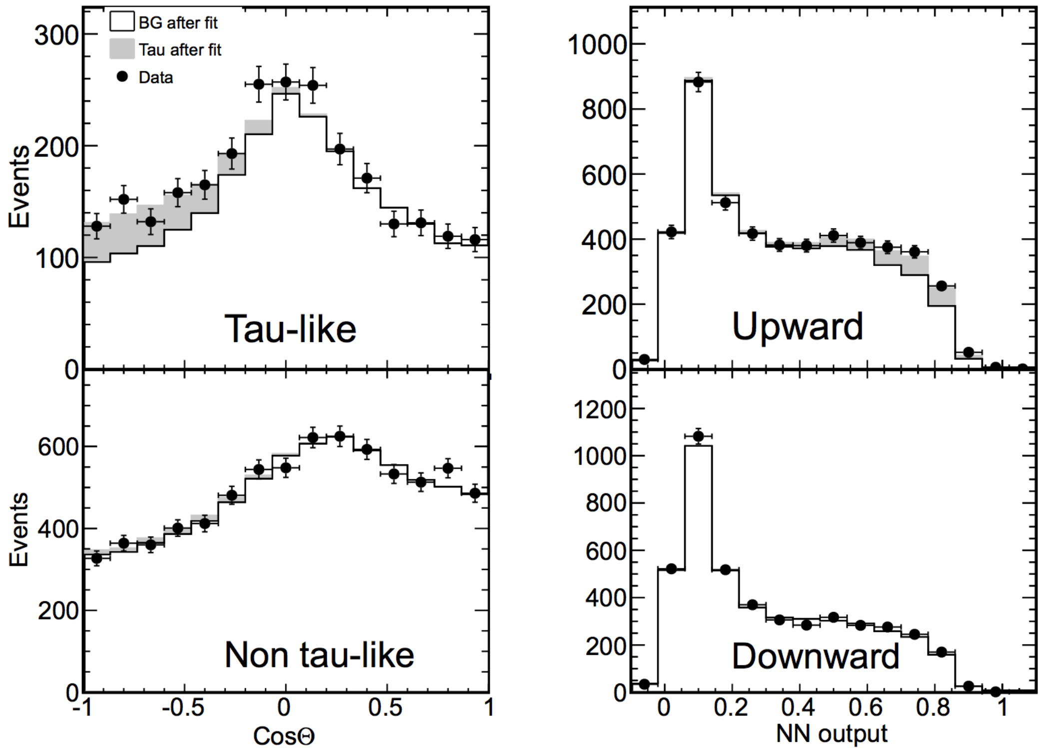

| NN > 0.5 | Background | Signal ($\nu_\tau$) |
|---|---|---|
| Efficiency | 28% | 76% |
| Purity | 95.3% | 4.7% |
| Rate $[\text{Mton}\cdot\text{year}]^{-1}$ | 8467 | 422 |

| Decay mode | Branching ratio (%) | Tau-like fraction (%) |
|---|---|---|
| $e^-\bar{\nu}_e\nu_\tau$ | 17.83 | $67.3 \pm 2.2$ |
| $\mu^-\bar{\nu}_\mu\nu_\tau$ | 17.41 | $42.6 \pm 2.6$ |
| $\pi^-\nu_\tau$ | 10.83 | $84.7 \pm 3.8$ |
| $\pi^-\pi^0\nu_\tau$ | 25.52 | $81.0 \pm 2.1$ |
| $3\pi\nu_\tau$ | 18.29 | $88.7 \pm 2.5$ |
| Others | 10.12 | $90.5 \pm 3.4$ |

- Neutral Current (NC), isotropic hits similar to signal
- With 1st-gen(+) ML: reconstruct $\rho_{770}$ , better $\mu, \pi, e, \gamma$ separation?

# Search for Tau Neutrinos at SK :

PHYS. REV. D **98**, 052006 (2018)



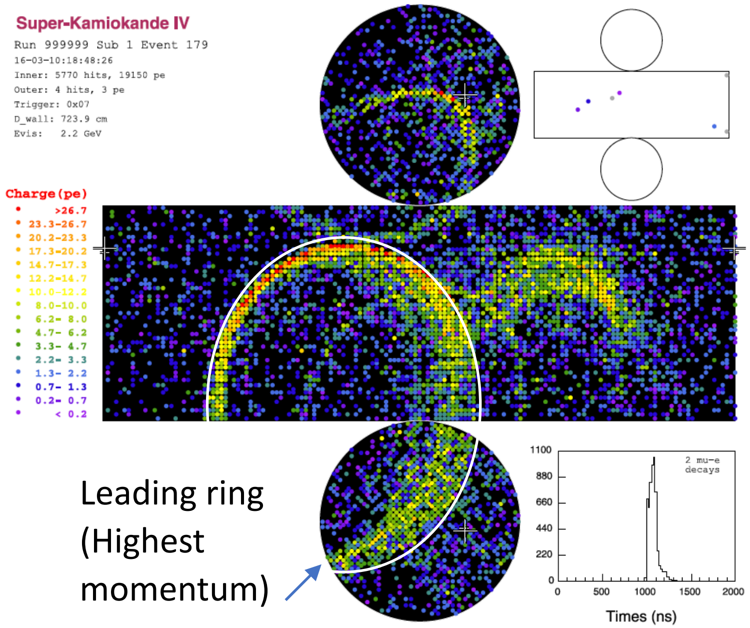

$$\text{Data} = \text{PDF}_{\text{BG}} + \alpha \times \text{PDF}_{\text{tau}} + \sum \epsilon_i \times \text{PDF}_i$$

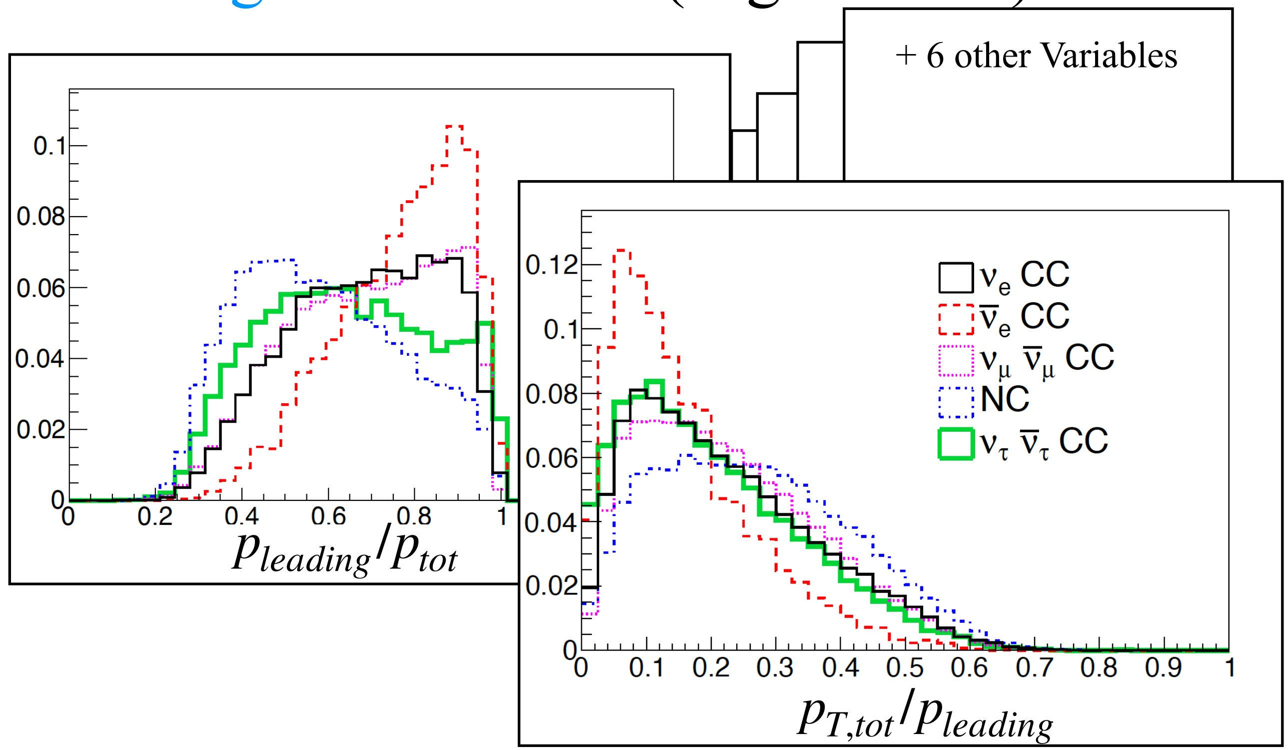
$\alpha =$ 1.47±0.32 (stat+syst)

4.6σ rejection of no τ appearance

- Fit 2-dimensional PDFs (cos θ, Neural Network ), while simultaneously varying systematic error templates
  - <u>No cut</u> to separate tau-like and non-tau-like
- Uses 328 kton-yr exposure (1996-2018 data)

# General Classification of Multi-Ring Events: BDT (Light GBM)




- Separate atmospheric $\nu$ sample into pieces with different sensitivity
  - $\nu_e : \bar{\nu}_e$ - Mass hierarchy, $\delta_{CP}$ ; $\nu_\mu$ - atm. Mixing ; NC - background
- Like tau problem, hard to identify leading lepton among many rings
- Adopt BDT
  - First multi-classifier at SK adopted in 2020
  - Improved sensitivity and reduced training time (MLP:20min →BDT:10s)

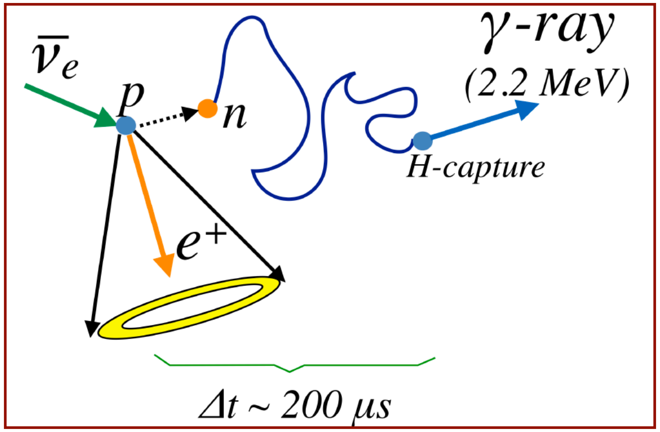
# General Classification of Multi-Ring Events: BDT (Light GBM)

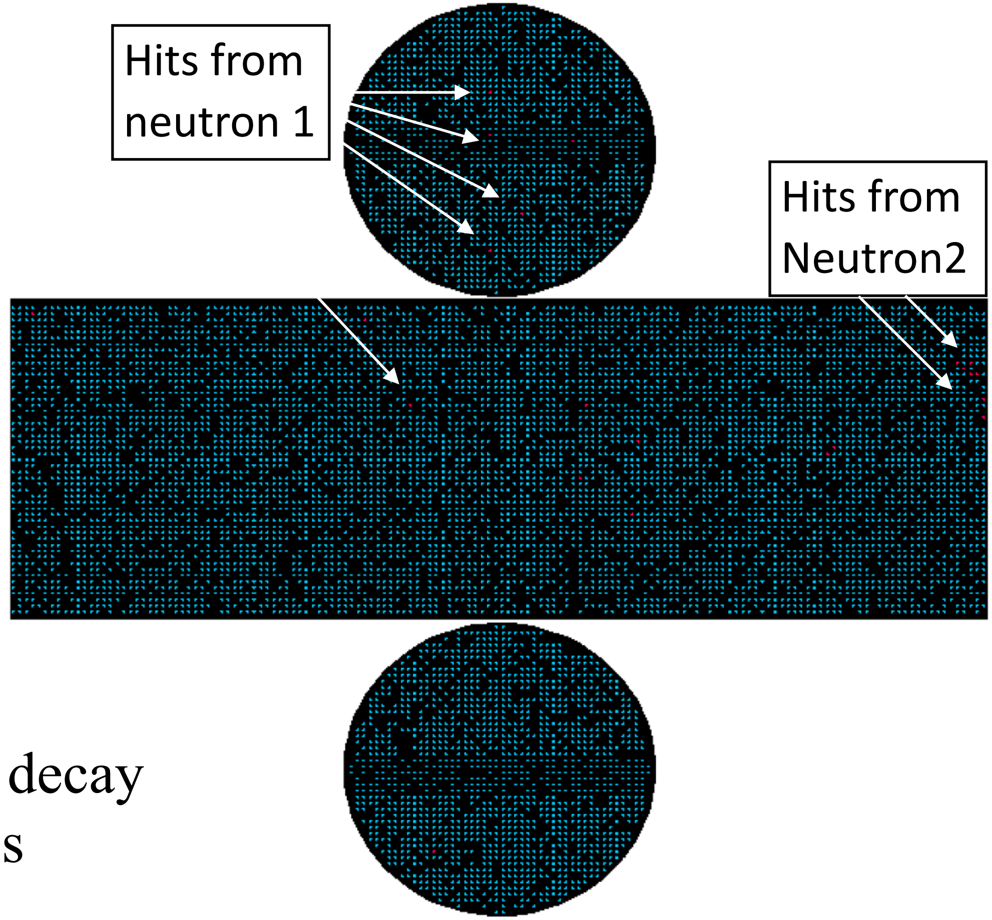
| % | BDT Signal Eff. | BDT Purity | LLR Signal Eff. | LLR Purity |
|---|---|---|---|---|
| $\nu_e - like$ | 45.8 | 46 | 34.4 | 46 |
| $\bar{\nu}_e - like$ | **62.4** | **32** | 60.4 | 24 |
| $\nu_\mu - like$ | 84.4 | 91 | 77.0 | 93 |
| NC+$\nu_\tau$ | 44.2 | **54** | 55.6 | 44 |
| **Train Time** | **10 s** | | | |
| $\Delta\left(\Delta\chi^2_{MO}\right)$ | **1.1** | | **0** | |

- Separate atmospheric $\nu$ sample into pieces with different sensitivity
  - $\nu_e$ : $\bar{\nu}_e$ - Mass hierarchy, $\delta_{CP}$ ; $\nu_\mu$ - atm. Mixing ; NC - background
- Like tau problem, hard to identify leading lepton among many rings
- Adopt BDT
  - First multi-classifier at SK adopted in 2020
  - Improved sensitivity and reduced training time (MLP:20min →BDT:10s)

# Neutron Tagging

Integrated Event Display $t_\nu + [18{,}540]\mu s$ with two neutrons (red hits) and dark noise (blue hits)

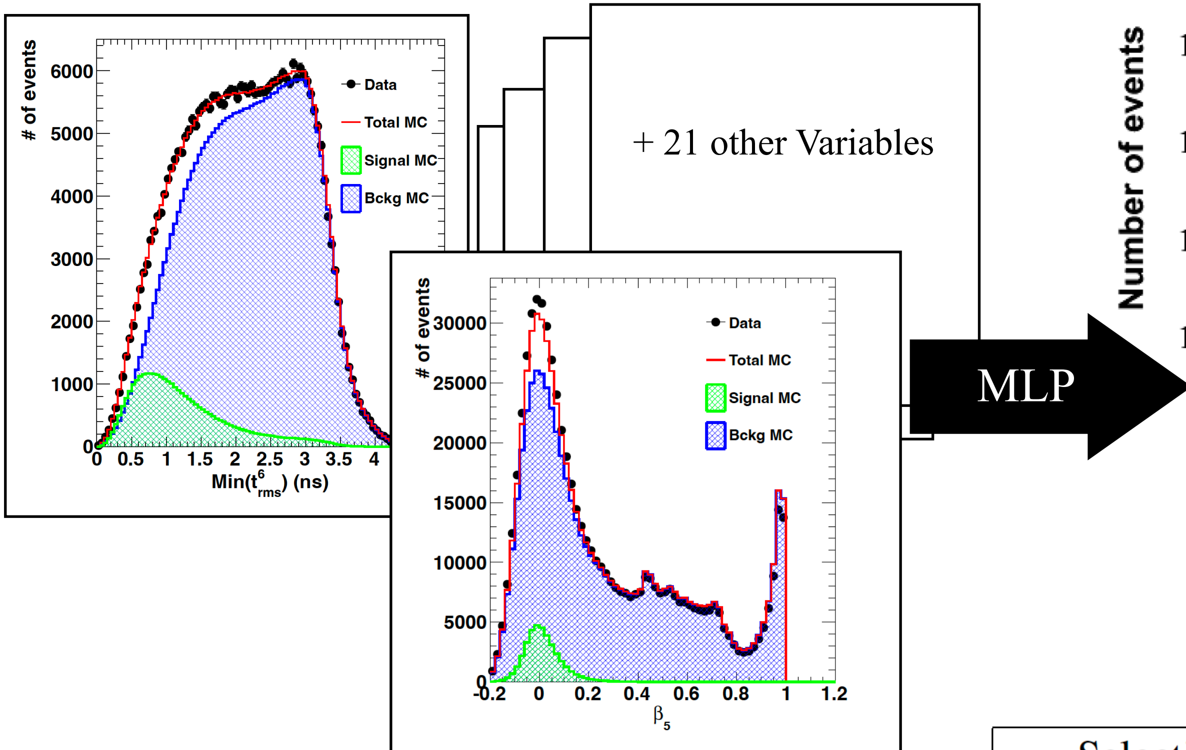


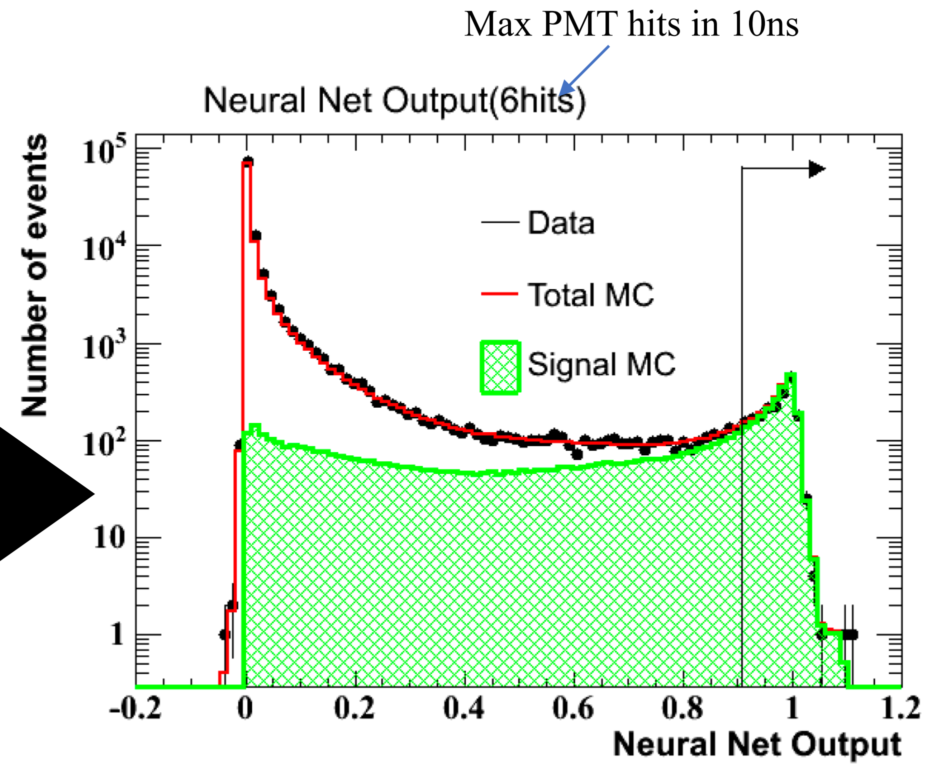




- Neutrons useful for
  - Neutrino-antineutrino separation
  - Reducing backgrounds to proton decay
  - Tagging supernova relic neutrinos

- Capture on hydrogen produces 2.2 MeV gamma ~ 10 PMT hits
- Similar number expected from dark hits, but not along a Cherenkov ring
  - BG model taken directly from data (random trigger)

# Neutron Tagging

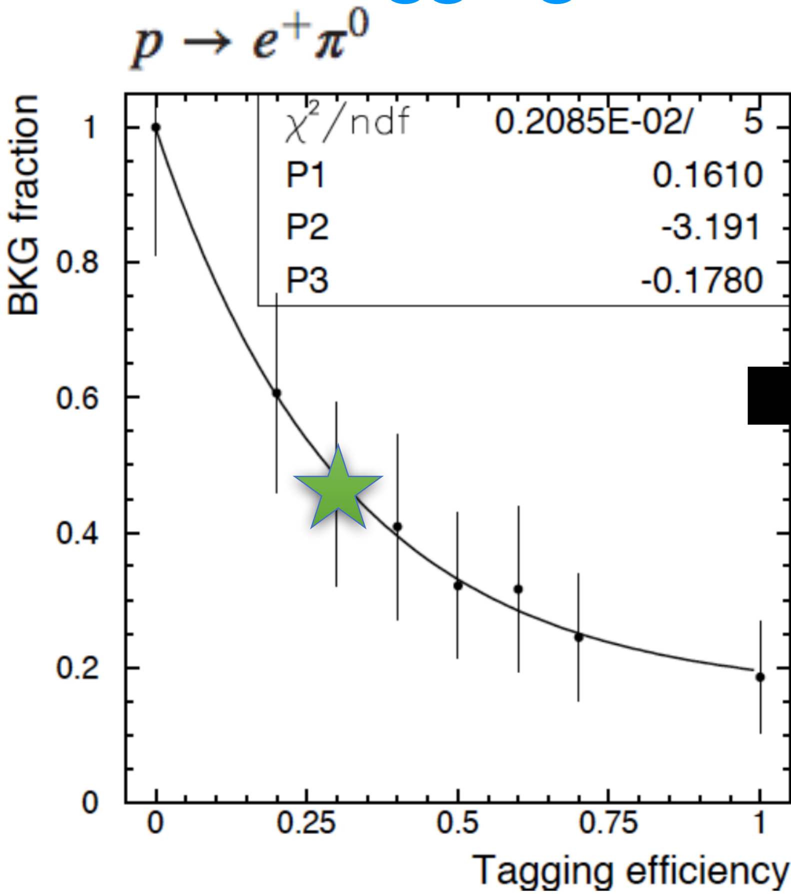


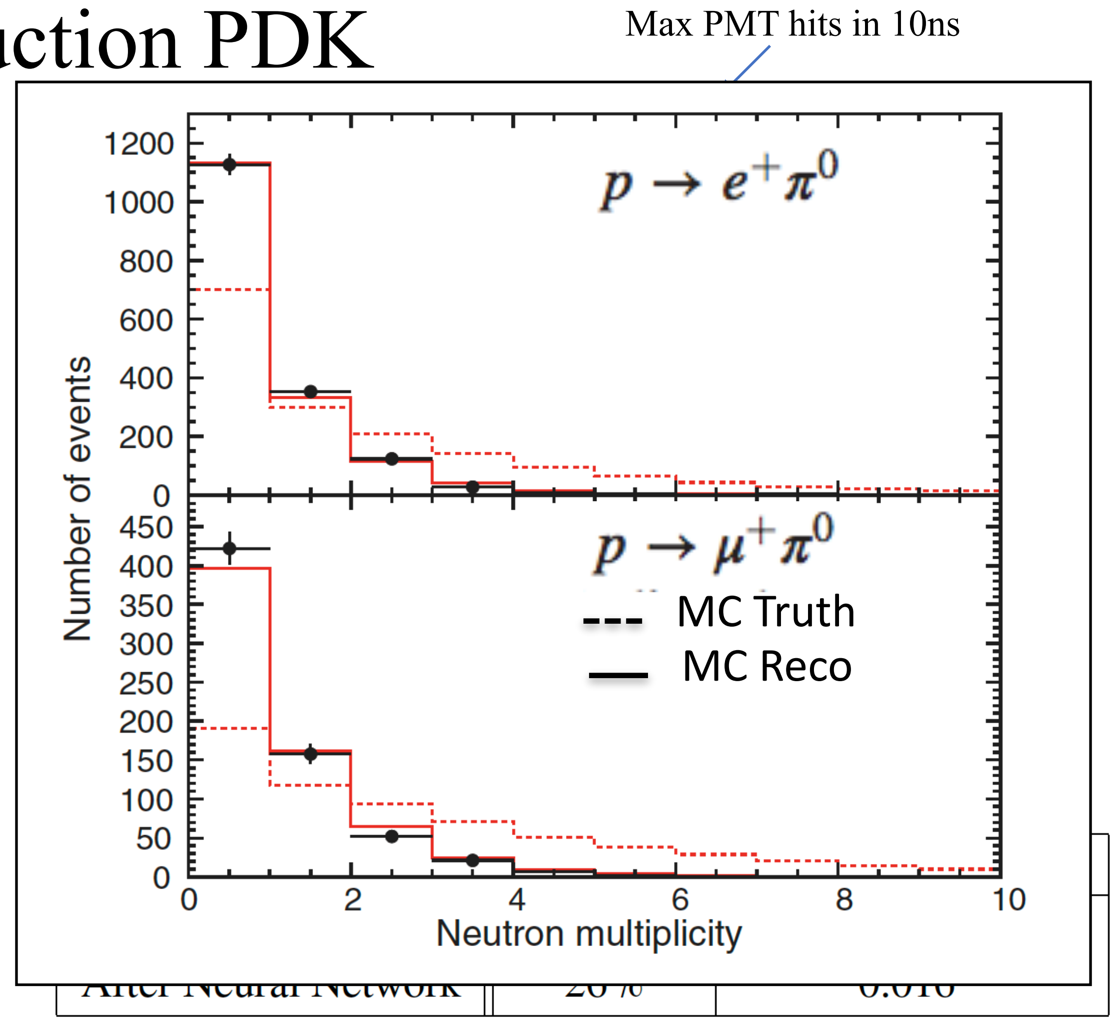




| Selection stage | Efficiency | Background / Event |
|---|---|---|
| Initial Selection | 49% | 22 |
| After Neural Network | 26% | 0.016 |

- Successfully observe neutrons, but with low efficiency (26%) in pure water
  - (N.B. in 0.01% Gd-loaded water ~ 40~50%)
- NN is sensitive to changes in dark rate, water transparency, and $n - \nu$ distance
  - Can these be improved with 1st-gen ML?

# Neutron Tagging: BG Reduction PDK



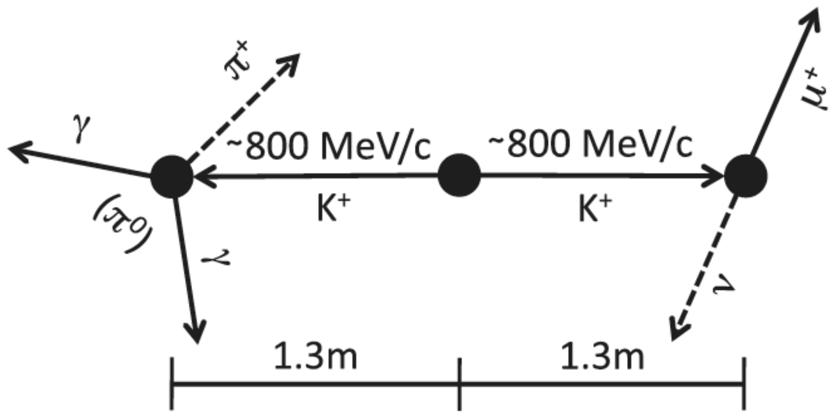




- Successfully observe neutrons, but with low efficiency (26%) in pure water
  - (N.B. in 0.01% Gd-loaded water ~ 40~50%)
- NN is sensitive to changes in dark rate, water transparency, and $n - \nu$ distance
  - Can these be improved with 1st-gen ML?

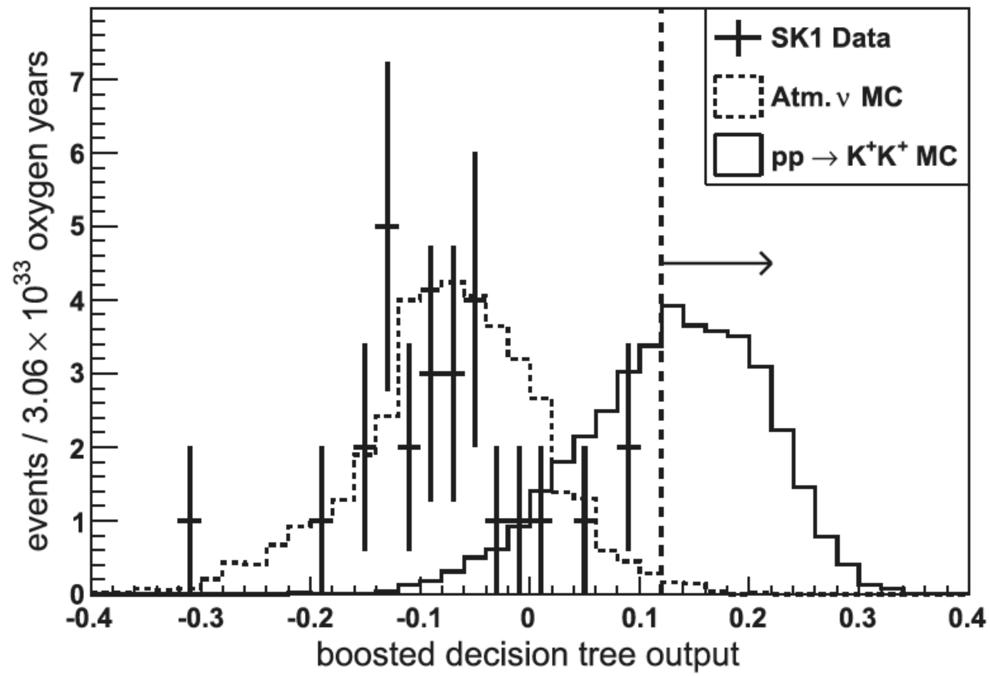
# Other Applications

Dinucleon Decay Search Search

$$pp \rightarrow K^{+}K^{+} \rightarrow \mu^{+}\nu_{\mu}\pi^{+}\pi^{0}$$




PRL **112,** 131803 (2014)




- R-Parity-violating mode
  - Predicts many particles and rings (K+ above C threshold)
- TMVA-based BDT
- 32 input variables , 500 trees
- Achieve 12.6% signal efficiency

$$\tau/BR_{pp\rightarrow K^{+}K^{+}} > 1.7 \times 10^{32}\text{years.}$$

# Summary and Conclusions

- So far SK has mostly adoped "0th-generation" ML for simple classification problems
- Expect some improvement in performance using more modern algorithms and techniques
  - Many problems lend them selves to image processing and pattern recognition
- Some effort towards Multiple-label classification problems implemented and kinematic estimation
- More advanced ML techniques likely to improve many aspects of the experiment
  - …underway.

# Supplements