Advanced new tool for background rejection in KamLAND geo-neutrino analysis using machine learning methods

#68

RCNS, Tohoku University
Taichi Sakai
sakai@awa.tohoku.ac.jp

[Geo-neutrino]

Geo-neutrino : Electron-antineutrinos from natural radioactive decays in the Earth

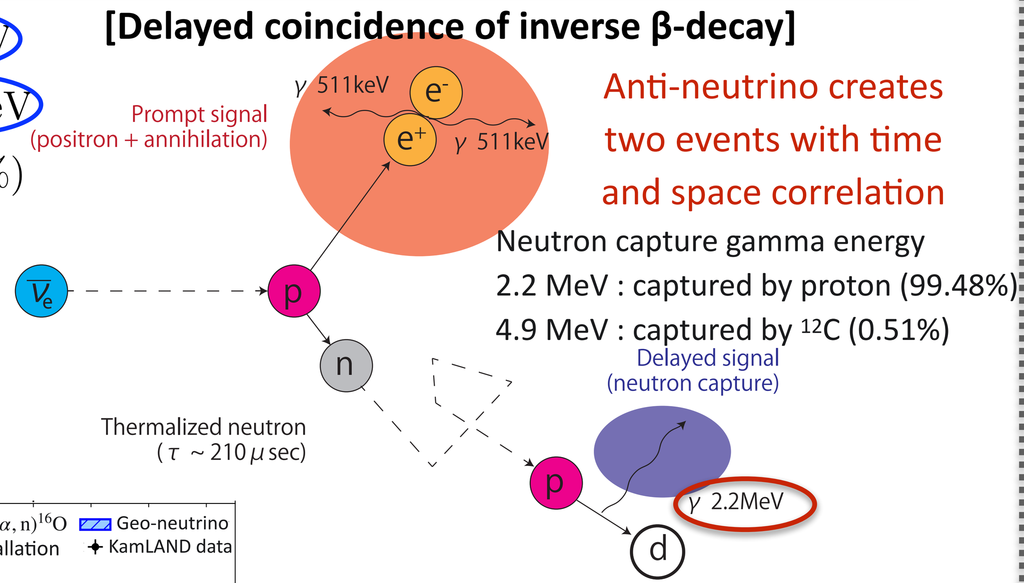
$^{238}U \rightarrow {}^{206}Pb + 8\alpha + 6e^- + 6\bar{\nu}_e + 51.7\,\mathrm{MeV}$

$^{232}Th \rightarrow {}^{208}Pb + 6\alpha + 4e^- + 4\bar{\nu}_e + 42.7\,\mathrm{MeV}$

$^{40}K \rightarrow {}^{40}Ca + e^- + \bar{\nu}_e + 1.311\,\mathrm{MeV}\ (89.28\%)$

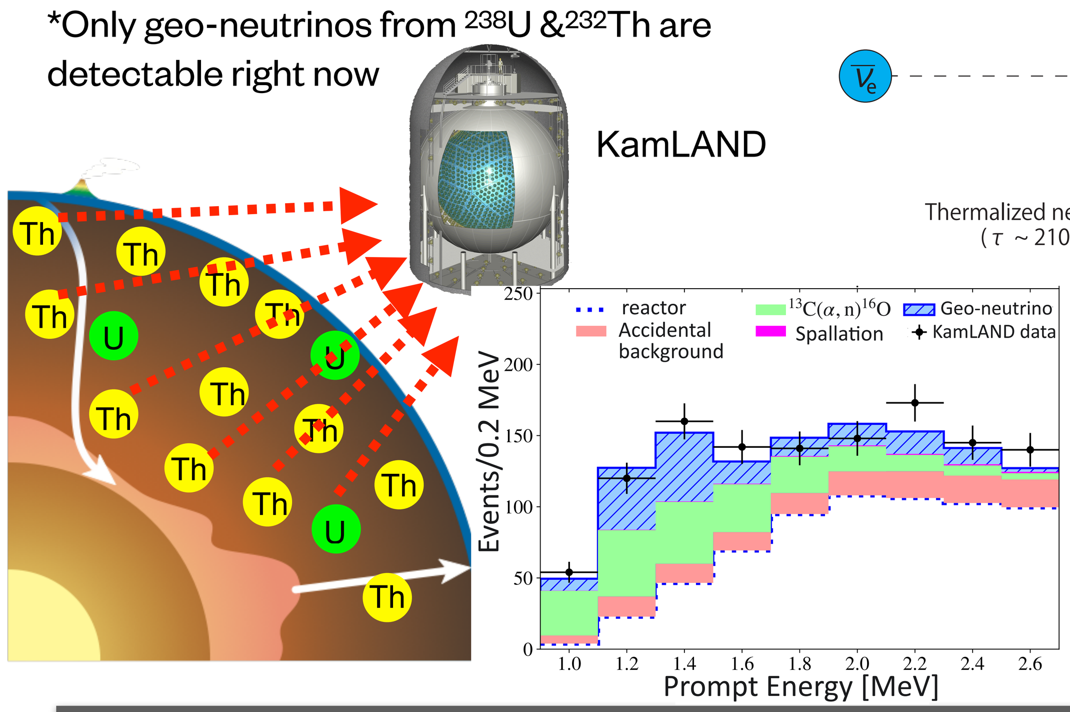
*Only geo-neutrinos from ^{238}U & ^{232}Th are detectable right now

[Delayed coincidence of inverse β-decay]



Anti-neutrino creates two events with time and space correlation

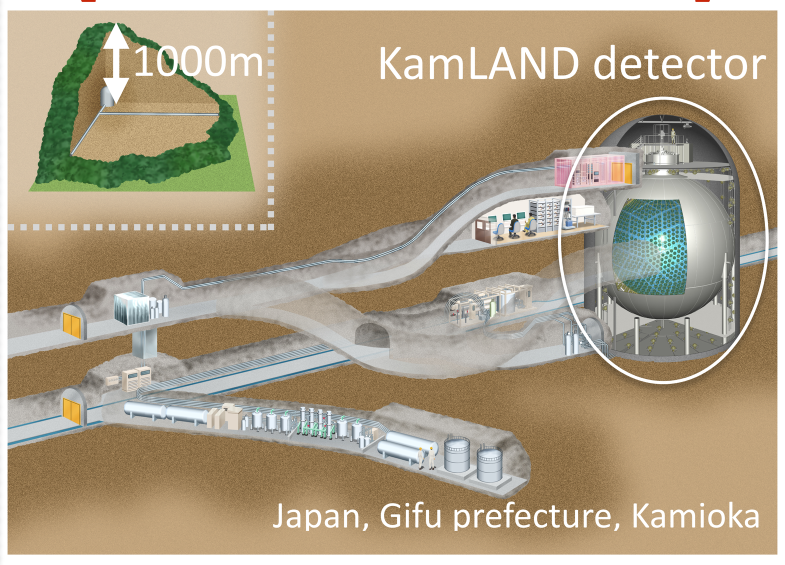
Neutron capture gamma energy
2.2 MeV : captured by proton (99.48%)
4.9 MeV : captured by ^{12}C (0.51%)



Number of geo-nu ∝ amount of U Th radiogenic heat

- **Geo-neutrino is powerful tool to directly measure the amount of radiogenic heat sources**

[KamLAND detector]



Depth: 2700 m.w.e
Muon flux = 1.50×10^{-7} (cm^{-2}s^{-1})

Low background environment
^{238}U: ~ 5.0×10^{-18} g/g
^{232}Th: ~ 1.3×10^{-17} g/g

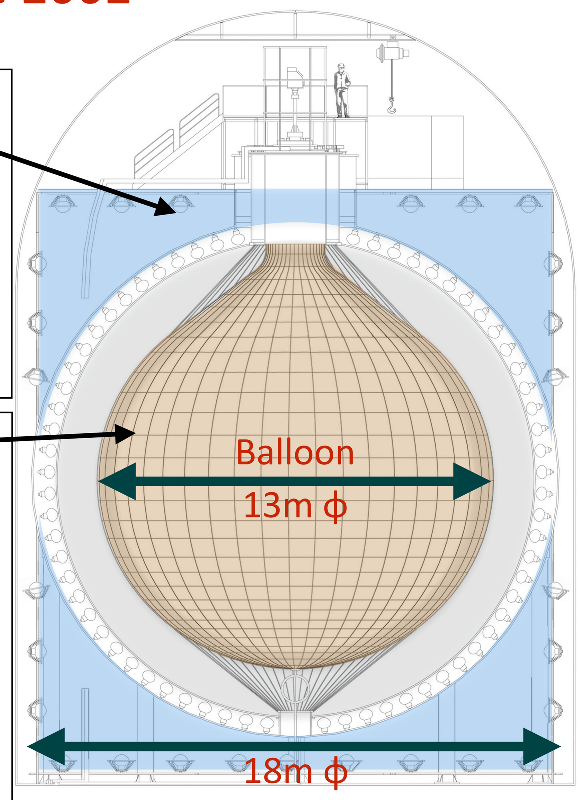
In operation since 2002

Outer detector
- 3.2 kton Pure water
- 20 inch PMTs (#140)

Shielding and active to muon

Inner detector
- 17inch PMT(#1325)
- 20inch PMT(#554)
- Buffer oil
- 1 kton LS in the balloon

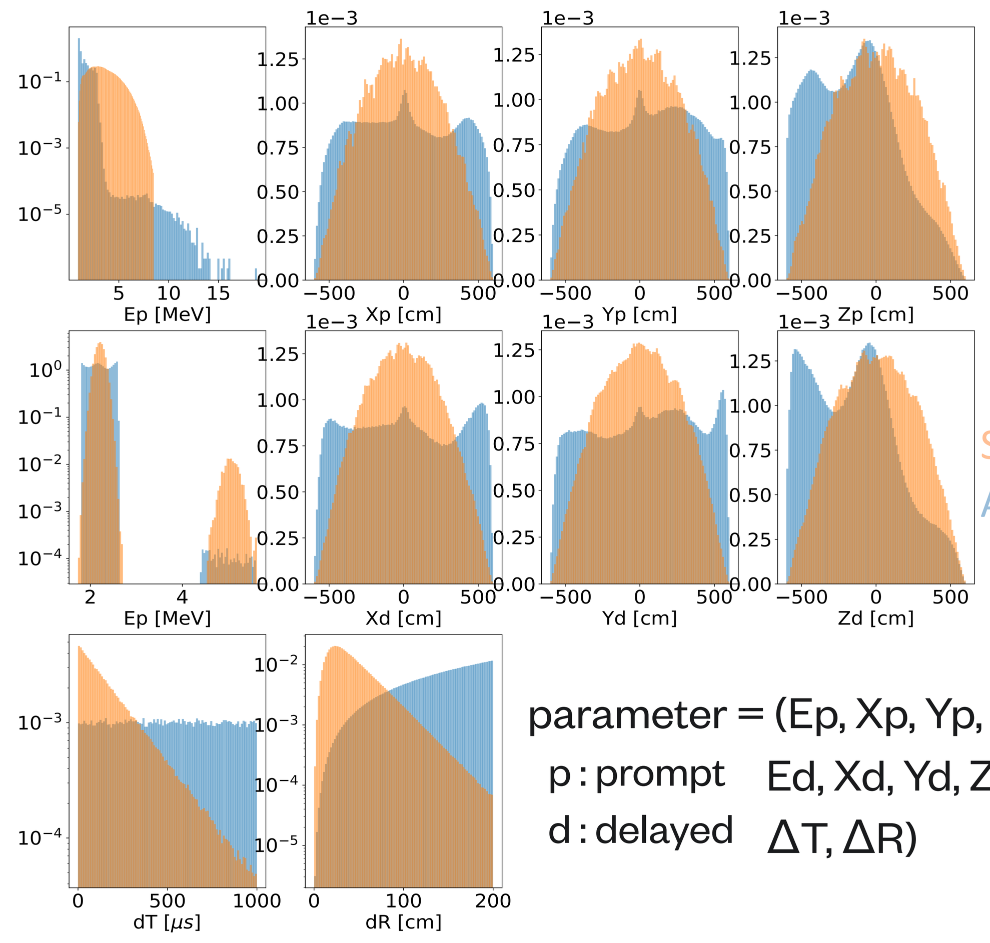
Neutrino detection



[BDT-likelihood selection]

BDT : Boosted Decision Tree

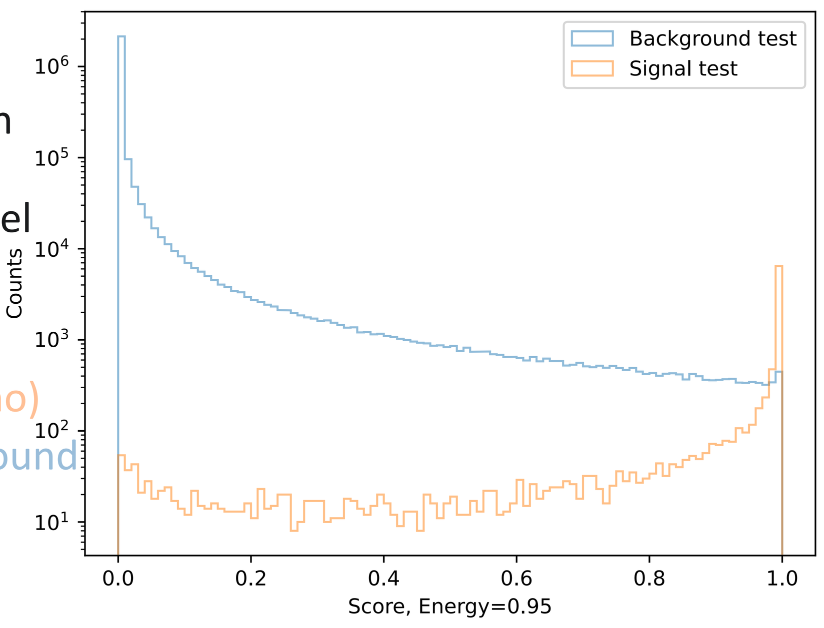
Input parameters



Each data is given a score by the trained BDT model

Signal(anti-neutrino)
Accidental background

Output histogram



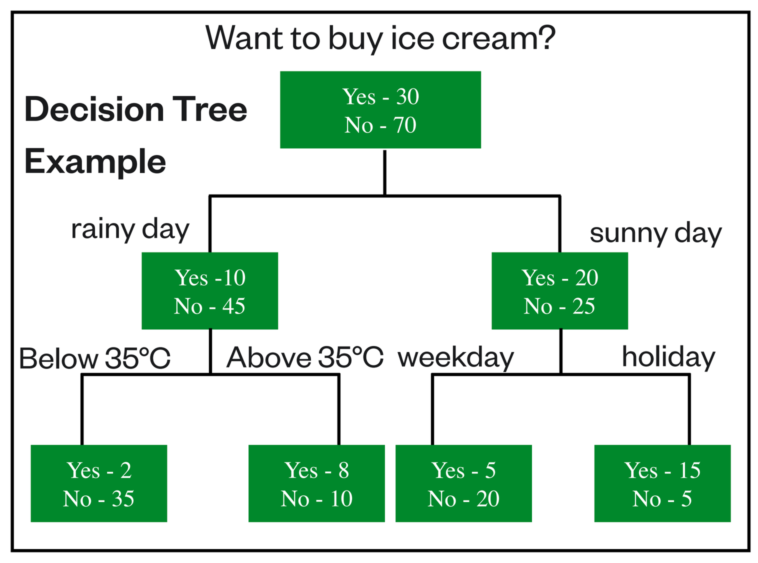
parameter = (Ep, Xp, Yp, Zp, Ed, Xd, Yd, Zd, ΔT, ΔR)
p : prompt
d : delayed

Old likelihood selection
Ep(76 bin), Ed(3 bin), Rp(5 bin), Rd(5 bin), ΔT(1 bin), ΔR(16 bin) → 1200 bin

Improvement points
- Binned analysis → un-binned analysis
- Consider correlations between parameters
- Input parameters taking into account detector asymmetry (R → X,Y,Z)

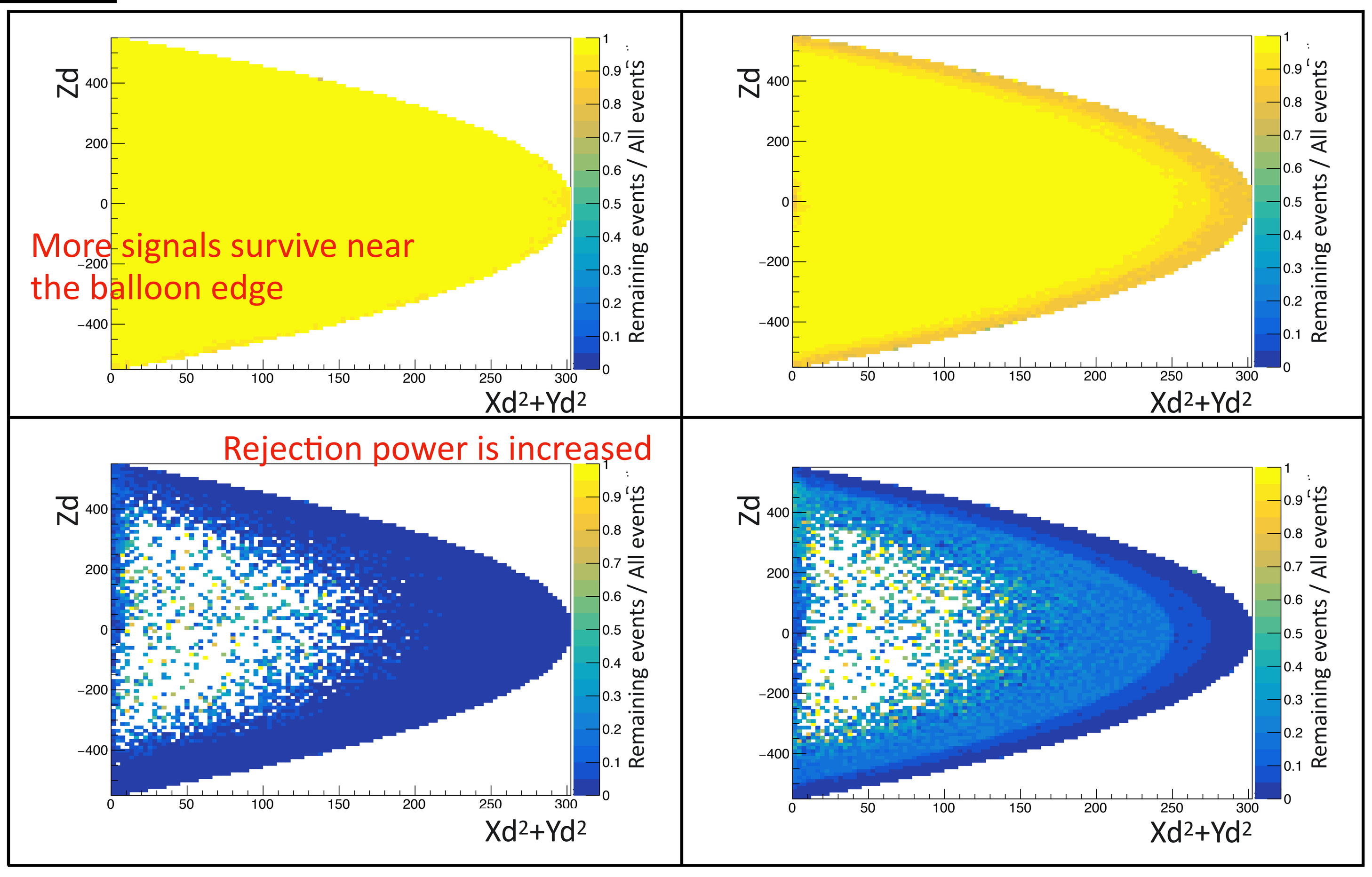
*Training data is generated by simulation(signal) and data-driven(background)
*Hyper-parameters are tuned using optuna

Decision tree :
A decision tree is a model that has a tree-like structure, where each "leaf" represents a question. Data is split based on the conditions of these questions.
(Ex. "leaf" : dT < 200μs , Ed > 2 MeV …)
GBDT (Gradient Boosting Decision-Tree) :
A method to reduce the error (loss) by adding weak learned decision tree. There are several powerful models.
XGBoost (Extreme Gradient Boosting):
Robust and prevents over-learning.

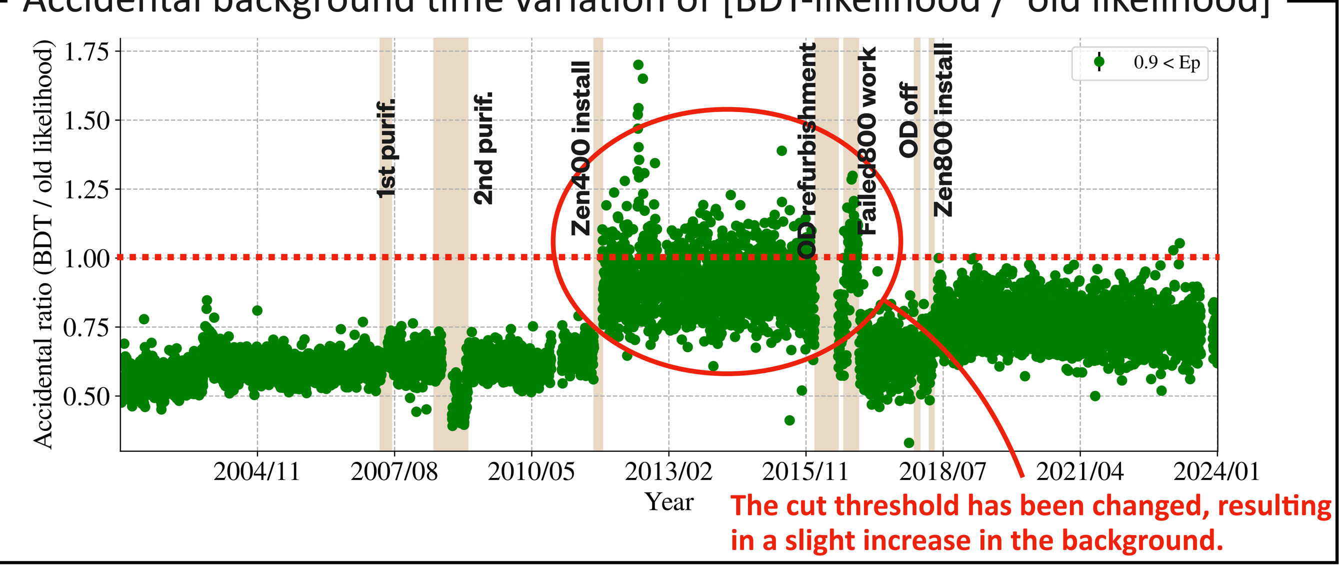


Result

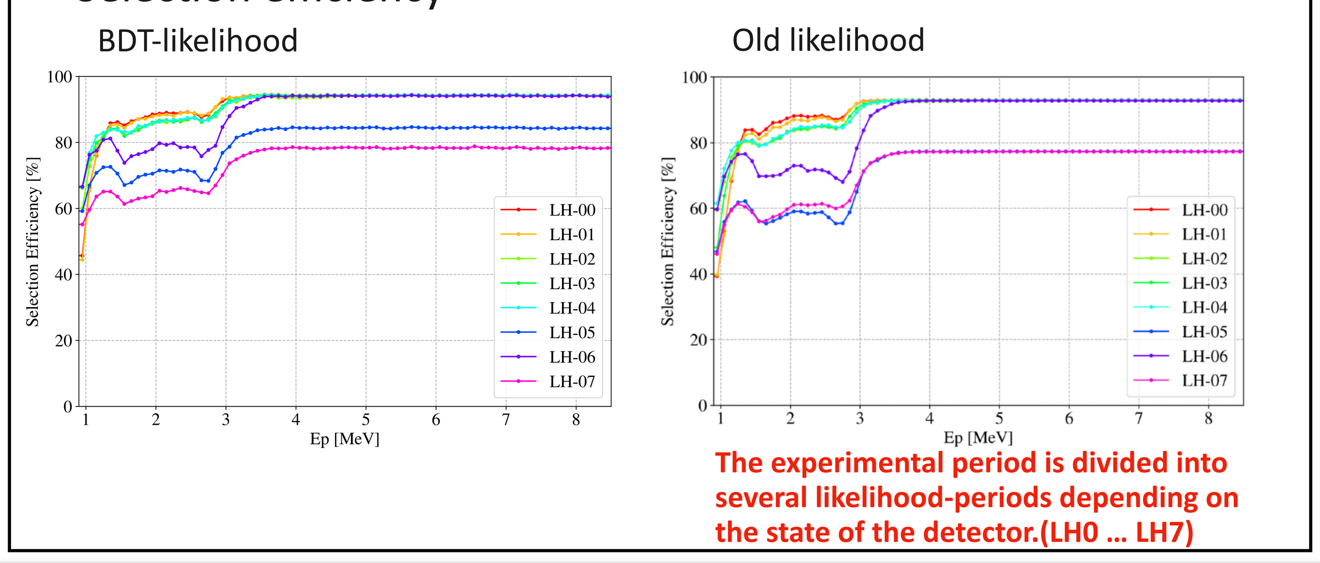
Remaining ratio



Result



Result

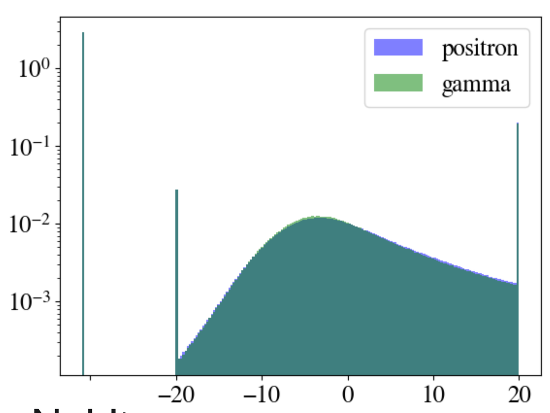


Background : 40% decrease
Signal : 8% increase
Significant improvement in S/N ratio

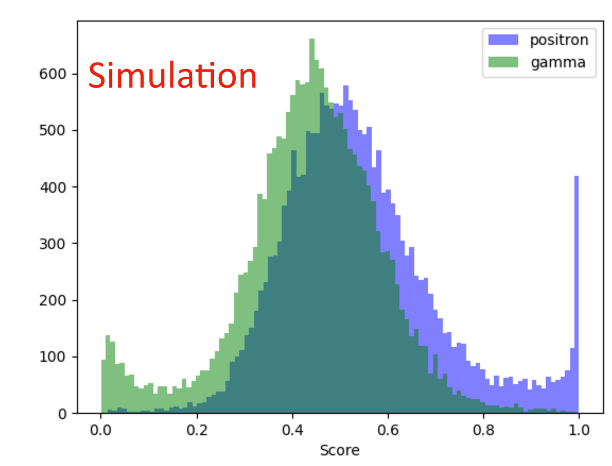
[Particle identification using graph neural network]

- GNN(graph neural network) model
- Training data is created by geant4

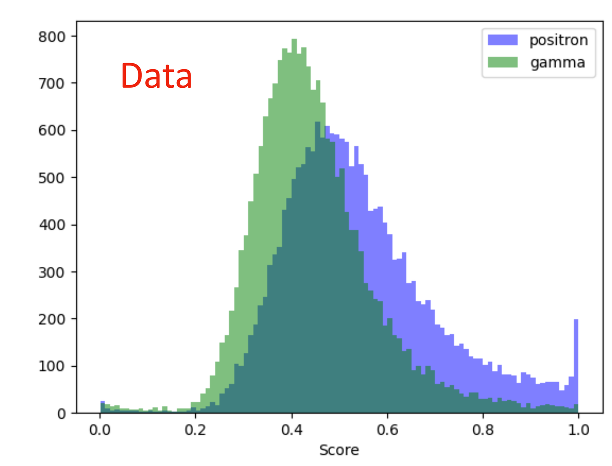
positron(annihilation, ortho-positronium), gamma
Positrons make ortho-positronium at about 50% and lifetime of ortho-positronium is about 3ns in Kam-LS.
Geant4 does not create ortho-positronium, so I made the events using electron and 2γ



Nohit
Time spectrum [ns]

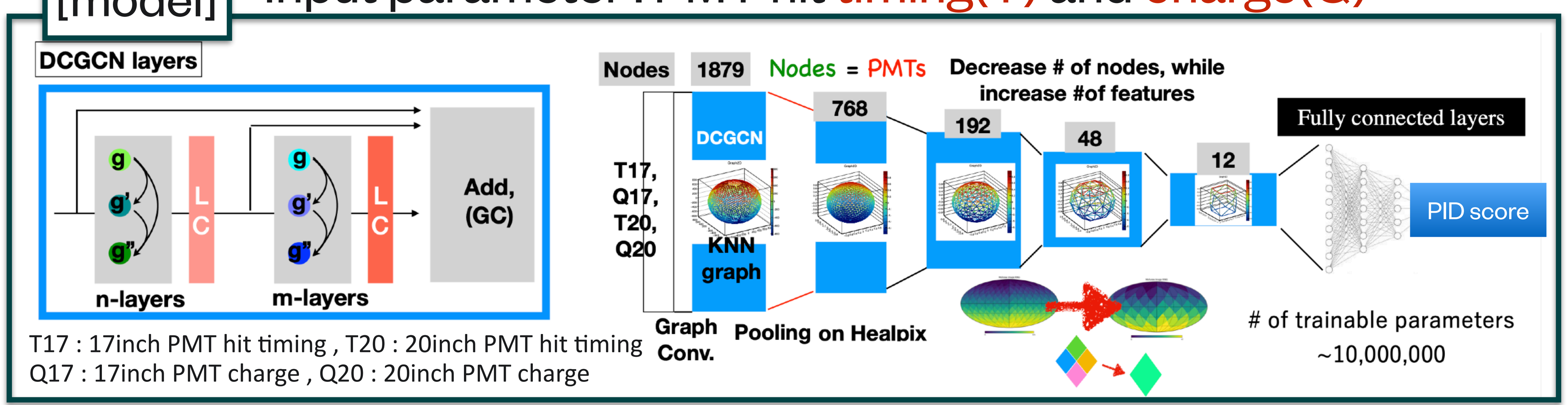


score (simulation data 0.9 - 8.5 MeV)



score (real data)
Positron : ^{11}C positron (1.2 -2.0 MeV)
Gamma : spallation neutron gamma (1.8 - 2.6 MeV)

[model] Input parameter : PMT hit timing(T) and charge(Q)



T17 : 17inch PMT hit timing , T20 : 20inch PMT hit timing
Q17 : 17inch PMT charge , Q20 : 20inch PMT charge

[summary]

* Significant improvements have been achieved by using new tools with Decision Tree
* The use of neural networks showed the feasibility of PID.
* Increased light intensity would improve PID accuracy
* Machine learning may also enable analyses that were previously challenging, such as neutrino directional detection.
* The new analysis using machine learning will be applied to the final data analysis of KamLAND.

Neutrino 2024

This work was supported by Graduate Program on Physics for the Universe (GP-PU), Tohoku University.