Machine Learning-Based Approaches for Initial Neutron Tagging in Hyper-Kamiokande

Sergio Suárez

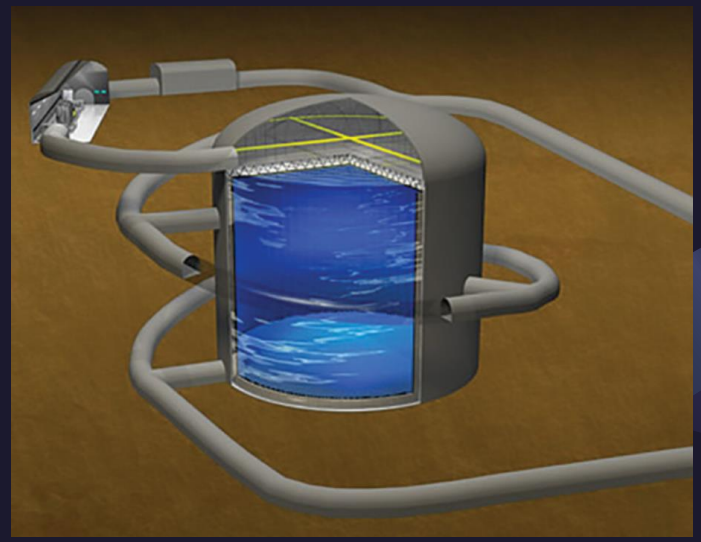
Jornadas del ICTEA 2025
ICTEA Days 2025

Introduction

Hyper-Kamiokande (HK) is the next-generation water Cherenkov neutrino detector currently under construction in Japan.

The nominal design of Hyper-Kamiokande consists of a massive **underground cylindrical tank** containing **258 kilotons of ultra-pure water**, with a **fiducial volume of 187 kilotons**—more than eight times larger than Super-Kamiokande.

HK will be located about **650 meters underground** in the **Kamioka area,** using the depth to shield the detector from cosmic ray interference. It will also be paired with the **J-PARC** accelerator facility, which will provide a controlled beam of neutrinos for long-baseline experiments.



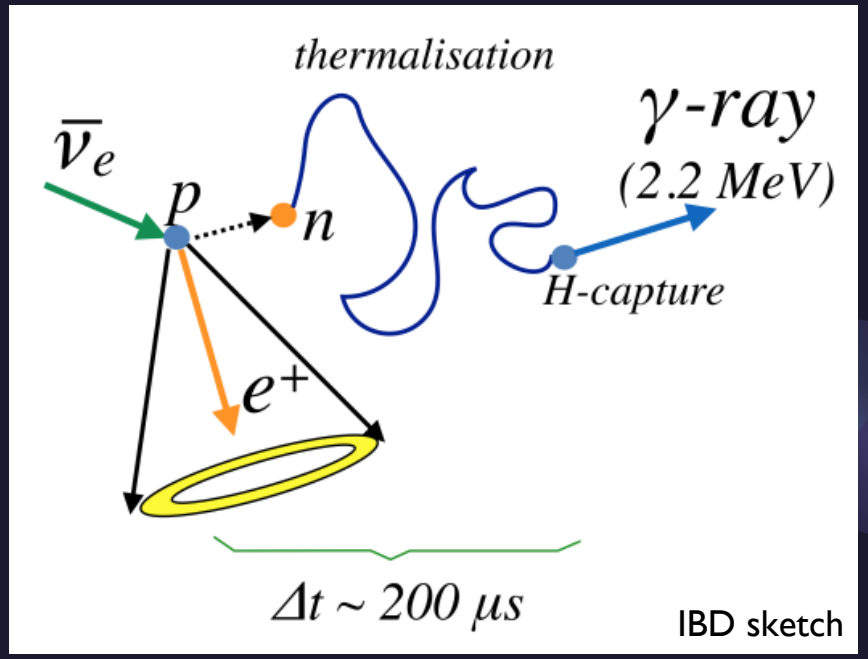
- **Neutrino Oscillations:** Precisely measure how neutrinos change flavors mass ordering and mixing parameters.
- **CP Violation in the Lepton Sector:** Test whether neutrinos violate charge-parity symmetry
- **Proton Decay:** Search for extremely rare decay processes
- **Supernova Neutrinos**
- **Solar and Atmospheric Neutrinos**

Introduction

Neutrons produced in the detector thermalize and are captured by hydrogen

Given the amount of H in HK, it is guaranteed that all thermalised neutrons will be captured by hydrogen

Once hydrogen captures a neutron, it is turned into deuterium, emitting a 2.2 MeV photon from its de-excitation with a lifetime of 200 μs.

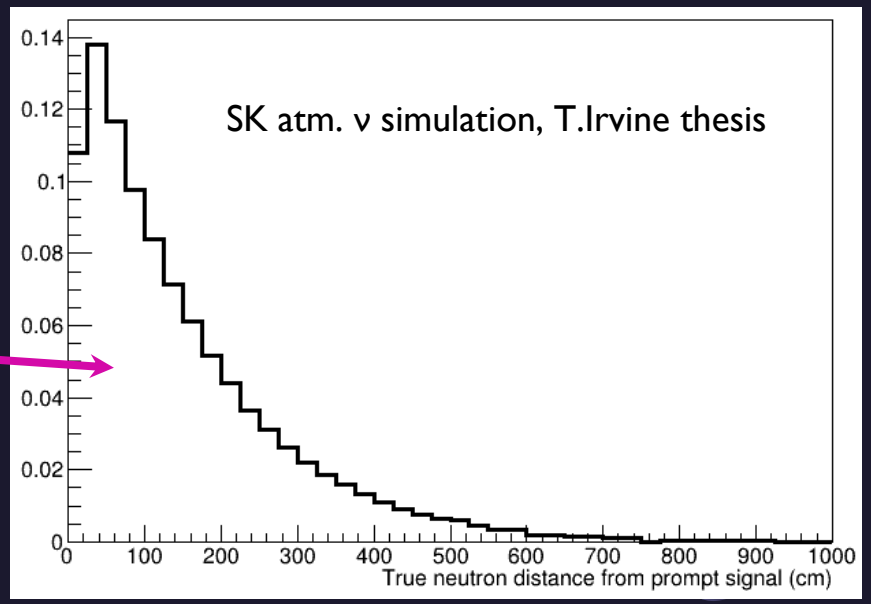


IBD sketch

Being this such a small signal (below the lowest energy trigger), only neutrons produced by in a previous triggered interaction are looked for

Simulation

- Simulated 2.2 MeV γ uniformly distributed in the HK ID using WCSim (hybridPMT version)
- Nominal HK configuration, i.e. 20k B&L PMTs with noise levels of 4.2 kHz
- As the reconstructed prompt signal is used to compute the TOF in the search for neutron's delayed signal, it is important to account for the distance travelled by the neutron during the thermalization
  - this depends on the energy of the generated neutrons
  - conservatively we are assuming neutrons generated from atmospheric neutrinos
  - this is included after the WCSim simulation in the TOF calculation



Selection of neutron candidates with Neural Networks

UP TO NOW

- Data sorted in time windows (10 ns)
- Threshold in the number of hits in 10 ns window to be considered as candidate
  - Results around ~58% True Positive Rate (TPR) and ~3% False Positive Rate (FPR)

OUR APPROACH

- Data sorted in 30 ns windows.
- Classification of whether the data of each PMT hit comes from signal or from background noise.
- We are planning to solve both detection of the signal, and a further step, identificate inside the signal which PMT measurements correspond to signal.

Selection of neutron candidates with Neural Networks

OUR APPROACH

- As inputs for the NN we use not only the timing information but also the charge of the hit PMTs
    - A small issue! Sometimes, several hits happen in the same ns. We choose to add the intensities to give the network all the information.

- Flag each PMT hit whether it comes from noise or signal

```
t-TOF (ns) | Q (pe) | Label (1 signal, 0 bkg)
126.599  1.10166 0
129.3098 2.00506 1
146.43456 1.38051 1
146.73961 0.813318 0
147.06379 0.47011 1
148.03917 1.06994 1
```

Neutron Tagging with Neural Networks

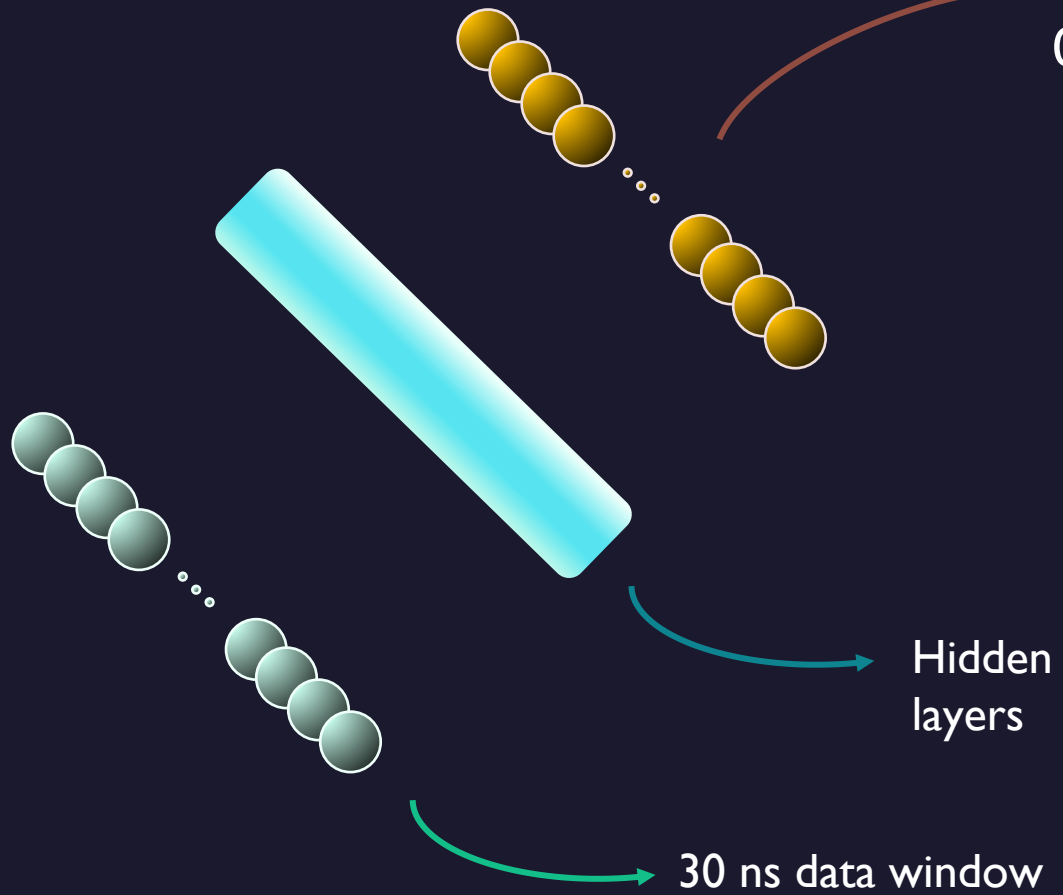
Improving traditional approach for detection only

Setting a rule to define if there is a signal in the data window, as a consequence of each ns identification data

Using what we learned from the problem to set a network that identifies the nature of each ns data in a window of 30ns

A conservative rule: at least 3 identifications are required to consider a detection

Neutron Tagging with Neural Networks



30 outputs, indentification of each ns
0 or 1 corresponding to if the hit is from bkg or sgn

If 3 or more sgn identifications

Detection of the sample as signal

Results over simulated data

SOME NUMERIC DATA RESULTS OVER SEVERAL ROC CUTS

IDENTIFICATION	10% CUT	30% CUT
TPR	0.8207	0.6175
TNR	0.966	0.9907
FPR	0.034	0.0093

DETECTION	60% CUT	90% CUT
TPR	0.7537	0.6294
TNR	0.9726	0.9893
FPR	0.0274	0.0107

Comparison with previous results ~3% of FPR, that achieved ~58% of TPR

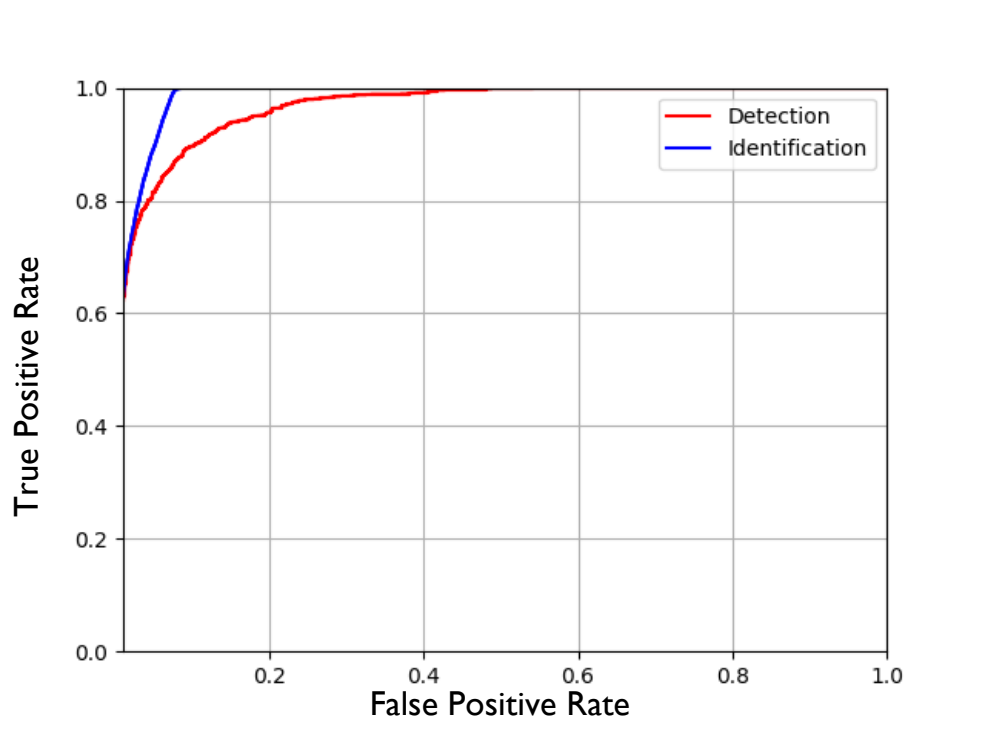
At ~1% of FPR, we are achieving up to ~62% of TPR

Identification of each ns hit has awesome results, ~3% of FPR with ~82% of TPR

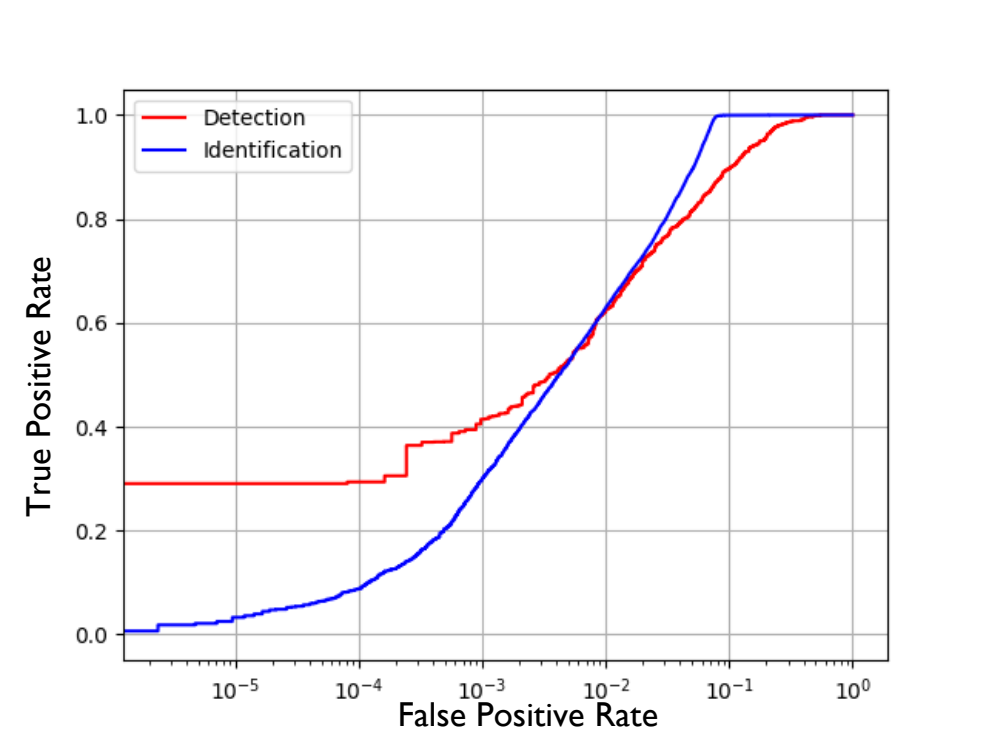
A more conservative cut in Identification of each ns hit, allows ~1% of FPR with ~62% of TPR

Results over simulated data

ROC CURVE



LOGARITHMIC



Summary

- Detection of neutro sigal has increased, at ~3% of FPR, from ~58% of TPR to ~75%
- Identification of which PMTs hits are from a positive signal are classified, at ~3% of FPR, with ~82% of TPR

Goal

- Try out more HK configurations (mPMTs, afterpulse, increase dark rate...)
- Feed the second stage of the reconstruction (FLOWER, LEAF) with the actual hit information
- Currently waiting for the start of the operations in HK for the possible implementation.
- … Also other related projects are being developed:
  - Tau identification (with Sara Rodriguez)
  - Low Energy Trigger (with Lluis Marti, Pablo Fernandez, Nataly Ospina)