PROCEEDINGS OF SCIENCE

New Approaches of first selection for Neutron Tagging in Hyper-Kamiokande

Sergio Luis Suárez Gómez[a,*] **and Pablo Fernández Menéndez**[b,c]

a Department of Mathematics, University of Oviedo, C/ Leopoldo Calvo Sotelo 18, Oviedo, Spain

b University of Liverpool, Department of Physics, Liverpool, United Kingdom

c Donostia International Physics Center DIPC, San Sebastián / Donostia, E-20018, Spain

E-mail: suarezsergio@uniovi.es, pablo.fernandez@dipc.org

Hyper-Kamiokande (HK) is a next-generation neutrino experiment with a large-scale water-Cherenkov far detector approved in Japan. Its physics program addresses some of the most challenging questions in fundamental physics like the precise measurement of the neutrino oscillation parameters (solar, atmospheric, accelerator), the investigation of astrophysical neutrino sources (supernovae and Diffuse Supernova Bursts (DSNB)), and the search for proton and exotic nucleon decays.
Since over a decade, HK's predecessor, Super-Kamiokande, has proven the importance of neutron-tagging in a large variety of measurements, improving the limits of DSNB and proton-decay searches, and enhancing the sensitivity to the atmospheric oscillation parameters.
Neutrons produced in the interaction of an HK event thermalize and are eventually captured by hydrogen, emitting a 2.2 MeV photon. This signal is too weak for HK's trigger threshold; therefore, the delayed neutron signal is searched by scanning all the hit PMTs after the prompt signal.
The developed method in this study feeds this information into a neural network, providing as output which of the hit PMTs are more likely to have received the neutron capture signal. This not only improves the candidate selection efficiency and purity, but also provides valuable information about the hit PMTs, identifying the most relevant ones for the subsequent fitting process.
This new technique improves the primary selection of neutron signals from 58% to 75% compared to the usual procedure based on a number-of-hits threshold.

The 22nd International Workshop on Neutrinos from Accelerators (NuFact2021)
6–11 Sep 2021
Cagliari, Italy

[*]Speaker

PoS(NuFact2021)241

1. Introduction

Hyper-Kamiokande (HK) is the next-generation of water-Cherenkov neutrino and proton decay detector to be built in Japan. The nominal design of HK consists of a tank containing 258 kton of ultra-pure water (187 kton fiducial volume) and optically divided into outer detector (OD) mainly devoted as cosmic-ray veto, and inner detector (ID), instrumented with 20,000 50-cm PMTs facing inwards [1].

Over the years, neutron tagging capabilities of these type of detectors have been proven to be very useful for improving many of their physics analysis results.
In this study, we show how Artificial Neural Networks (NNs) can be used at the hit PMT level to detect very low energy signals. The selection of NNs as technique to address the problem relies on their capability to discern signal from background in complex classification problems [2].

2. Neutron Tagging in Hyper-Kamiokande

Hydrogen is present in the water molecules of the HyperK detector and has a sizeable cross-section of 0.329 ± 0.004 barn, for the capture of thermal neutrons (Figure 1). The amount of hydrogen in HK guarantees that all thermalised neutrons produced inside the tank 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.

Over the last years, the usage and development of neutron tagging on hydrogen has been playing an increasingly important role on the physics studies of Super-Kamiokande [3].

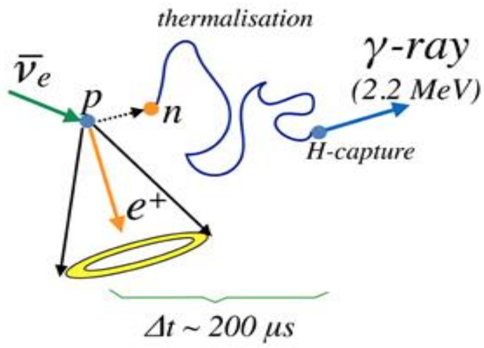

Figure 1. Hydrogen thermalisation of neutrons

3. Neural Network for Neutron Candidate Selection

To develope this technique in the context of the HK experiment, we simulate the signal of neutron captures on hydrogen using the WCSim software.
This signal is too low to activate the detector's trigger, therefore a scan of the hits, after the prompt signal, is conducted searching for a neutron signal candidate.

The inputs for the NN are the charges of the hit PMTs from a given time-window, and the output will be the selection of each hit as a signal or background.

3.1 Simulations

The simulations were performed to obtain 2.2 MeV γ uniformly distributed in the ID volume using official HK version of WCSim; for the time being, the nominal HK configuration is used, i.e. 20k B&L PMTs with noise levels of 4.2 kHz.

The time of hit PMTs is corrected by the time-of-flight (TOF) from the reconstructed vertex of the signal. In the simulation, we account for the distance neutrons travel during the thermalization process following the distribution below of neutrons generated from atmospheric neutrinos (Figure 2, [4]).

Further, we also account for the delayed time of the neutron capture with respect to the prompt signal, an exponential with a lifetime of 200 μs.

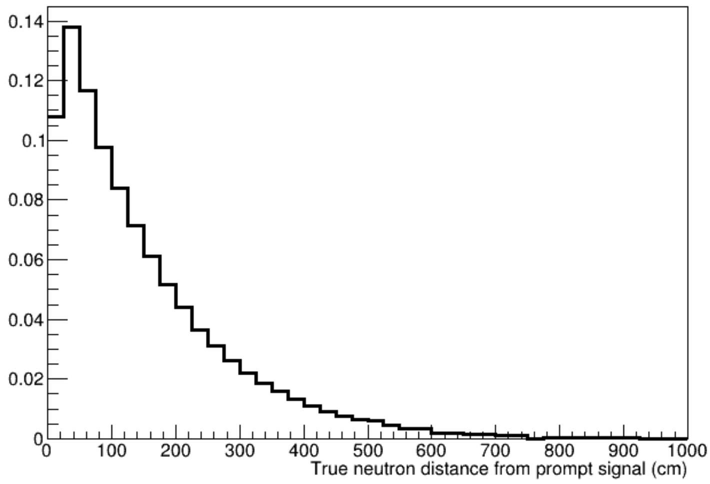

Figure 2. Neutron travel distance from the initial atmospheric neutrino interaction vertex.

3.2 Neutron Candidate Selection

Up to now, the approaches for the selection of neutron candidates is performed sorting the data in segments of 10 ns, with a threshold in the number of hits (usually 7) in this time-window to be considered as candidate. The results obtained round 58% True Positive Rate (TPR) at a 3% False Positive Rate (FPR).

The presented approach suggests data to be sorted in 30 ns windows. These time-windows are chosen to be large enough to contain a complete neutron capture signal, and small enough to avoid multiple neutron captures.

The aim of the work is to solve both, detection of the neutron capture signal and, as further step, the identification of the individual hit PMTs receiving that signal inside the time window. The detection is performed with NNs, using as inputs not only the timing information but also the charge of the hit PMTs.

The data is then turned into an array of 30 elements (one element per ns), with values zero if there is no hit or the the charge for the hit PMT at that time. Considering the issue that sometimes several hits happen in the same ns, it was chosen to add the charges to give the network all the information.

We checked that adding the intensities significantly improves the performance of the method. The label data is prepared as flagging each PMT hit whether it comes from noise or signal.

A rule to define if there is a signal in the data window was set; at least 3 identifications are required within a given time-window to consider a detection (neutron candidate).

4. Results and Discussion

Since the outputs of the neural network are continuous between 0 and 1, the boundary that defines the output as positive or not may imply a difference in the efficiency of the results. Using different cuts for those values allows to select safer classifications in detriment of the number of detections. The statistics are shown for some cuts in Table 1.

Identification	**10% cut**	**30% cut**	**Detection**	**60% cut**	**90% cut**
TPR	0.8207	0.6175	TPR	0.7537	0.6294
TNR	0.966	0.9907	TNR	0.9726	0.9893
FPR	0.034	0.0093	FPR	0.0274	0.0107

Table 1. Most relevant cuts found for both the identification and detection classification.

In comparison with previous results ~3% of FPR, that achieved ~58% of TPR, we select neutron candidates ~62% of TPR at a ~1% FPR.
Identification of each ns hit has even better results, ~3% of FPR with ~82% of TPR, opening the door to study other definitions of the requirement for detection.

The complete analysis from the different cuts can be extracted from the behavior of the ROC curves shown in Figure 3, and a detailed logarithmic ROC curve in Figure 4.

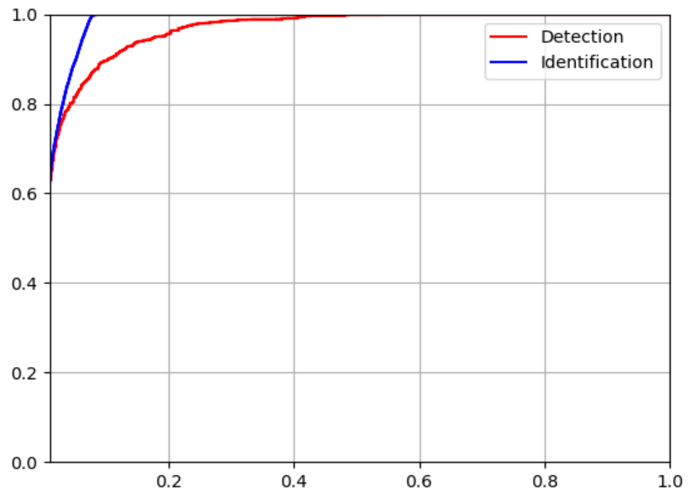

Figure 3. ROC Curve of results from detection and identification classification.

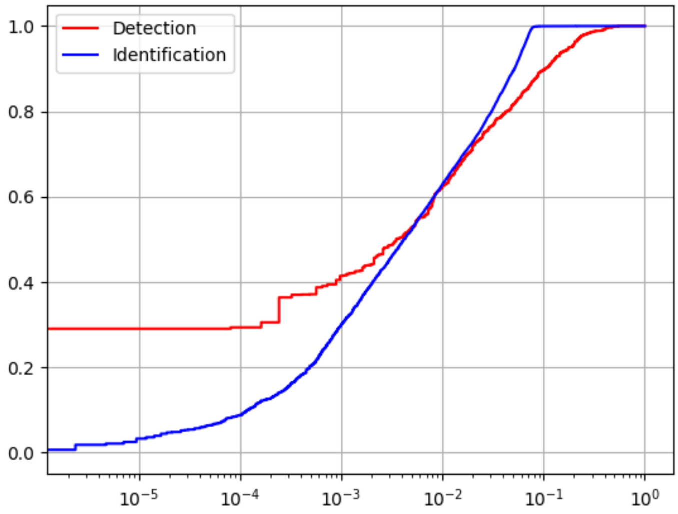

Figure 4. Logarithmic ROC Curve of results from detection and identification classification.

Some variations that can be done for this approach include the selection of the adequate time window, and its consequent modification to the input number of neurons for the NN.

Also, we are considering that 3 PMT identifications are required to consider a detection of the signal. This means that in segments (data samples) where the signal got cut between segments, and only 1 or 2 PMT hits are detected, they are not considered as signal (nor for the simulated data or the neural network output). This should be improved in future implementations of the approach.

5. Conclusion

This work successfully achieves the goal of how to give a simultaneous answer to both detection of a signal and its correspondent PMT hits identification problem, when events with energies below the trigger of HK are considered, by means of using artificial intelligence techniques.

This new technique improves the primary detection of the delayed neutron signal from 58% to 75% at the same false positive rate of 3%.

References

[1] K. Abe *et al.*, *Hyper-Kamiokande design report*, arXiv Prepr. arXiv1805.04163, 2018.
[2] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. **1**. MIT press Cambridge, 2016.
[3] K. Abe *et al.*, *Diffuse supernova neutrino background search at Super-Kamiokande, Phys. Rev. D*, vol. **104**, no. 12, p. 122002, 2021.
[4] T. J. Irvine, *Ph.D. thesis*, Univ. Tokyo, 2014.

PoS(NuFact2021)241