PROCEEDINGS OF SCIENCE

Enhancing Atmospheric Background Reduction using Convolutional Neural Networks in DSNB searches at Super-Kamiokande Gd

Soniya Samani[a,*] **for the Super-Kamiokande Collaboration**

[a] *Department of Physics, University of Oxford,*
Keble Road, Oxford, OX1 3RH, United Kingdom

E-mail: soniya.samani@physics.ox.ac.uk

The detection of the Diffuse Supernova Neutrino Background (DSNB) flux will provide invaluable insights into constraining cosmological models, core-collapse dynamics and neutrino properties. The Super-Kamiokande-Gd (SK-Gd) experiment currently exhibits the best sensitivity for discovery due to enhanced neutron tagging capability with 0.011% $Gd_2(SO_4)_3 \cdot 8H_2O$, as per this analysis. While the Inverse Beta Decay (IBD) interaction is identifiable in SK-Gd, the low-energy signal is dominated by atmospheric neutrino backgrounds. This study explores a novel approach to background reduction by leveraging topological features of SK events with the discriminative power of Convolutional Neural Networks (CNNs). Well-established techniques for data pre-processing, event selection and feature extraction are used to train CNNs on IBD and atmospheric Neutral Current (NC) Monte-Carlo (MC) events. The preliminary performance of two CNN models highlights the potential of using machine-learning techniques to improve the DSNB signal efficiency.

XVIII International Conference on Topics in Astroparticle and Underground Physics (TAUP2023)
28.08_01.09.2023
University of Vienna

[*] Speaker

PoS(TAUP2023)223

1. The Diffuse Supernova Neutrino Background

Direct observation of a Core-Collapse Supernova (CCSN) guarantees a substantial burst of detectable neutrinos, with the flux surpassing optical emissions by $O(10^3)$. Given that the CCSN burst rate is 2-3 per century within <10 kpc, efforts also focus on measuring the integrated neutrino flux from more distant SN in the observable universe [1]. This forms the Diffuse Supernova Neutrino Background (DSNB). The isotropic and time-independent signal of the DSNB flux consists of neutrinos of all flavours with expected energies up to 40 MeV [1]. The primary detection channel is via the Inverse Beta Decay (IBD) of electron anti-neutrinos ($\bar{\nu}_e + p \rightarrow e^+ + n$), which has the largest cross-section below 30 MeV [2]. This process produces a characteristic coincidence event signature from the (prompt) positron and (delayed) neutron capture final-state particles. Since the signal region is saturated by cosmic muon spallation and atmospheric neutrino backgrounds, a robust detector with neutron identification is required.

2. The Super-Kamiokande Detector

Super-Kamiokande (SK) is a water Cherenkov detector situated in the Kamioka mine in Japan, shielded by a 1000 m rock overburden [3]. It consists of a welded stainless-steel tank, 39.3 m in diameter and 41.4 m in height, filled with 50 ktons of ultra-pure water. Support structures optically separate the tank into two regions containing the Inner Detector (ID) and the Outer Detector (OD). The ID is instrumented with 11,129 20-inch photomultiplier tubes (PMTs), providing a 40% photocathode coverage for effective event detection. An outward-facing array of 1,885 8-inch PMTs operates as the OD, serving as an active veto against backgrounds. To mitigate backgrounds from radioactivity near the detector wall, a 22.5 kton fiducial volume is defined for most analyses. With the addition of $Gd_2(SO_4)_3 \cdot 8H_2O$ in SK at 0.011% concentration in 2020, neutron captures on Gd emit an 8 MeV γ-cascade with an average time constant of $116.5 \pm 0.2\,\mu s$ [4]. This SK-VI phase achieves a 50% tagging efficiency on Gd that enables enhanced sensitivity to the DSNB flux.

3. Atmospheric Neutrino Background

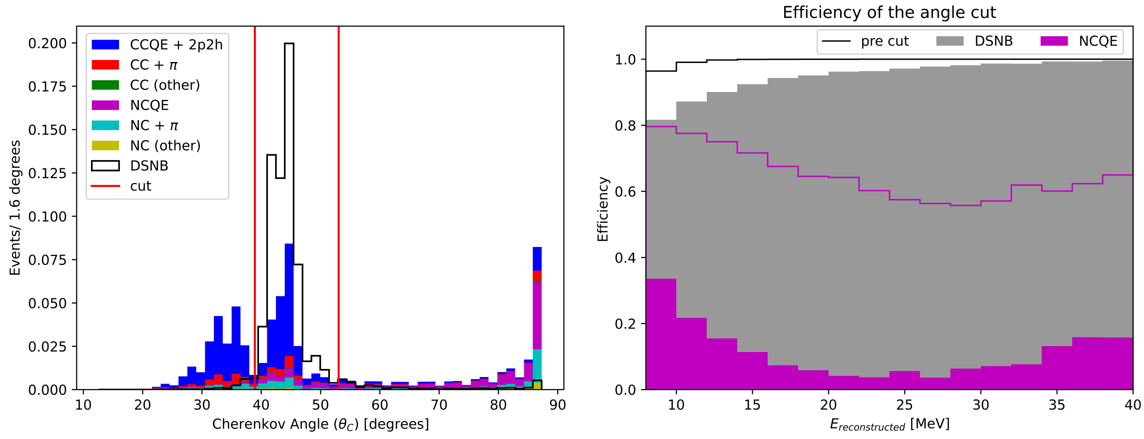
Neutral Current Quasi-Elastic (NCQE) interactions of atmospheric neutrinos with oxygen nuclei is a dominant background below 20 MeV, where the predicted DSNB flux is high [6]. For $E_\nu \geq 200$ MeV, the nucleon knock-out processes,

$$\nu_l(\bar{\nu}_l) + ^{16}O \rightarrow \nu_l(\bar{\nu}_l) + n + ^{15}O^* \qquad \nu_l(\bar{\nu}_l) + ^{16}O \rightarrow \nu_l(\bar{\nu}_l) + p + ^{15}N^*$$

induces emission of primary γ-rays during the de-excitation of residual nuclei, with energies between 1-10 MeV [5]. Outgoing neutrons undergo hadronic interactions with other nuclei in the detector triggering the production of secondary γ-rays and additional neutrons. The combination of prompt γ-rays and a subsequent neutron capture produces a delayed-coincidence event signature that closely mimics an IBD signal. Therefore, NCQE interactions of atmospheric neutrinos significantly contribute as a challenging residual background in the analysis.

Implementing cuts on the Cherenkov angle (θ_C) is currently the most powerful discriminator for removing this background. Highly-relativistic IBD positrons reconstruct with $\theta_C \sim 42°$ whereas NCQE events emit several energetic γ-rays that cannot be individually resolved in SK [6]. Instances of such events with multiple Cherenkov cones can mis-reconstruct with a large θ_C peaked close to

PoS(TAUP2023)223



(a) **(b)**

Figure 1: (a) Normalised θ_C distribution for the DSNB signal, based on the Horiuchi+09 flux model (black) and the atmospheric neutrino MC sample separated by each Charged Current (CC) and Neutral Current (NC) interaction channel. This includes contributions from CCQE, two-particle, two-hole (2p2h) states (blue), CC and NC interactions with pion (π) production (red and cyan), all other NC and CC interactions (green and yellow) and NCQE interactions (magenta). The cut points are shown as red lines. (b) DSNB and NCQE MC efficiency before and after the θ_C cut.

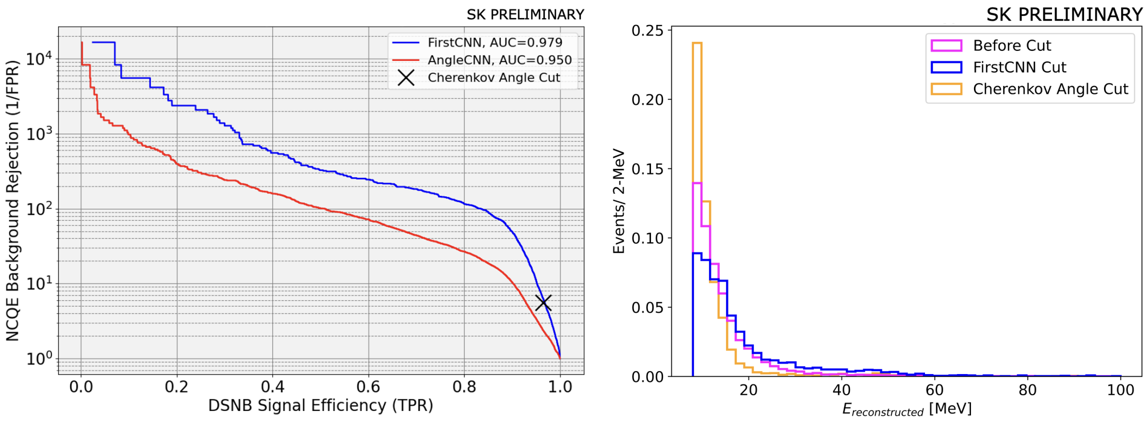
90°, as shown in Figure 1a. Currently, the cuts-based analysis removes the high-angle component by selecting an optimised θ_C region between $[38°, 53°]$ at 95% signal efficiency. However, in Figure 1b, it is evident that the approach retains $\sim$ 40% of low-energy background events in the final sample. The limitations of the traditional analysis cuts motivates investigation into alternative methods.

4. Background Reduction with Machine-Learning

Convolutional Neural Networks (CNNs) for the characterisation of IBD and NCQE MC events, based on PMT hit patterns, have been developed for atmospheric background rejection. The ResNet18 architecture is selected for its advantageous use of residual learning and shortcut connections [7]. The network's input consists of a 147 x 150 pixel map containing the barrel and endcaps to form an SK event display-style image. Each pixel represents a single PMT with a two-channel input for the timing and charge within a 300 ns time slice after the event trigger time. This defines the prompt window used to train the CNN for e^+/γ classification. Model training, validation, and testing were completed using an IBD signal and NCQE background MC sample containing 400,000 events. SK's flagship neutron tagging algorithm is applied to implement a single tagged neutron selection criterion on the NCQE sample.

Two separate CNN models are developed to assess the capabilities of machine-learning techniques as either alternatives or supplements to conventional analysis cuts. The models are described below:

- **FirstCNN:** Evaluates a CNN's discriminatory power by training on all events, with no positron cuts applied, to directly compare against the performance of the θ_C cut.
- **AngleCNN:** Investigates a CNN's potential to further reduce background after applying the θ_C cut to the training sample.



(a) **(b)**

Figure 2: (a) NCQE background rejection as a function of signal efficiency of FirstCNN and AngleCNN ResNet18-based models. The 95% θ_C cut efficiency is shown for the R.O.I. on the curve. (b) Normalised energy distribution of NCQE MC with no cuts (magenta), θ_C cut (orange) and FirstCNN cut (blue) applied.

Figure 2a displays the Receiver Operating Characteristic curves for both CNN models tested on a dataset with no θ_C cut applied. The FirstCNN model marginally outperforms the θ_C cut at the 95% signal efficiency baseline. Closer inspection of the energy spectra indicates that FirstCNN adapts an energy-dependent cut inferred from the PMT hit distribution (Figure 2b). Further studies, not shown here, confirmed that FirstCNN demonstrates superior background reduction compared to AngleCNN when tested across another dataset with the θ_C cut applied, highlighting the effectiveness of diverse dataset training. These results support integrating a CNN method into the full analysis to leverage its effectiveness at NCQE background rejection. Ongoing efforts aim to refine the FirstCNN model and include it in the full energy bin-by-bin analysis to set constraints on DSNB flux upper limits.

References

[1] John F. Beacom, *The Diffuse Supernova Neutrino Background*, *Annual Review of Nuclear and Particle Science*, **60**(1), 439–462, 2010.

[2] G. L. Fogli, E. Lisi, A. Mirizzi, D. Montanino, *Probing supernova shock waves and neutrino flavor transitions in next-generation water-Cerenkov detectors*, *JCAP*, **04**, 002, 2005.

[3] Y. Fukuda et al., *The Super-Kamiokande detector*, *Nucl. Instrum. Meth. A*, **501**, 2003.

[4] K. Abe et al., *First gadolinium loading to Super-Kamiokande*, *Nucl. Instrum. Meth. A*, **1027**, 166248, 2022.

[5] K. Abe et al., *Measurement of the neutrino-oxygen neutral-current interaction cross section by observing nuclear deexcitation γ rays*, *Phys. Rev. D*, **90** (7), 072012, 2014.

[6] K. Abe et al., *Diffuse supernova neutrino background search at Super-Kamiokande*, *Phys. Rev. D*, **104** (12), 122002, 2021.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, *Deep Residual Learning for Image Recognition*, 2015, arXiv: 1512.03385