

# The Search for CP-Symmetry Violation at DUNE

## Final Project Report

2133055

*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*

13/03/2025

The precise measurement of the leptonic CP-violating phase at the Deep Underground Neutrino Experiment (DUNE) requires accurate classification of neutrino interactions at the far detector as either charged-current muon neutrino, charged-current electron neutrino, or neutral current. The use of liquid argon time projection chambers in the far detector at DUNE, and the enhanced spatial and calorimetric resolutions which these can offer, allows for the use of advanced machine learning techniques, such as convolutional neural networks, to be utilised when distinguishing signal and background interaction channels. Such techniques and their capabilities are explored in this paper. A comparative study of four machine learning approaches for track-shower classification identified convolutional neural networks as the most effective, achieving an F1 score of 0.976. Following this, an algorithm for event classification in the far detector is presented, combining three convolutional neural networks, each trained for a specific classification task, in a decision tree based structure. The proposed algorithm achieved an accuracy of  $0.884 \pm 0.006$  and  $0.928 \pm 0.005$  in the classification of charged-current muon neutrino and charged-current electron neutrino interactions respectively.

## I. Introduction

Charge-Parity (CP) violation refers to the phenomenon where processes such as particle interactions are not invariant under a CP-transformation. CP violation was first observed in quarks in a 1964 experiment studying the  $2\pi$  decay of the  $K_0^2$  meson [1]. Observation

of CP violation in the leptonic sector, or measuring the CP phase parameter ( $\delta_{CP}$ ) to be nonzero (or  $\pi$ ), could have profound consequences. T2K have measured values of  $\delta_{CP}$  differing from 0 and  $\pi$  to  $3\sigma$  [2]. Future experiments such as DUNE and Hyper-Kamiokande will seek to obtain a measurement of  $\delta_{CP}$  at  $5\sigma$  [3, 4].

DUNE is a cutting-edge experiment and international collaboration involving scientists and engineers from across the globe, aiming to answer some of the most fundamental questions in the universe. DUNE has the potential to uncover groundbreaking new physics across astrophysics, cosmology, and particle physics through its three main physics goals [5]. The first goal of DUNE, is to carry out precise measurements of neutrino oscillations, with the intention of establishing the neutrino mass hierarchy, the octant in which the mixing angle  $\theta_{23}$  lies in, as well as measuring  $\delta_{CP}$  to better than three standard deviations over more than 75% of the range of possible values of  $\delta_{CP}$  [5]. Secondly, the detectors will allow for the observation of supernova neutrinos and solar neutrinos, which could provide insight into the early stages of core-collapse [6, 7]. The final main goal will be to search for proton decay in several decay modes. If observed, this would agree with the predictions of the grand unification of forces [5, 8, 9, 10].

Observation of CP violation in the leptonic sector could explain the matter-antimatter asymmetry observed in the universe. Sakharov first proposed in 1966 [11] that CP violation is necessary to account for the baryon asymmetry observed in the universe [12]. The extent of CP violation in quarks is insufficient to account for this asymmetry [2], however, the large mixing angles in the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix suggest the potential for significant CP violation in the leptonic sector, which could offer a plausible explanation for baryon asymmetry [13]. Observing CP violation at DUNE would support the leptogenesis mechanism, which identifies CP violation in the leptonic sector as a necessary condition for explaining the baryon asymmetry of the universe [14, 15].

To measure  $\delta_{CP}$  at DUNE, it will be necessary classify the event displays at the far detector as either charged-current electron neutrino ( $CC\nu_e$ ) or charged-current muon neutrino ( $CC\nu_\mu$ ) for the  $\nu_\mu$  beam, and similarly  $CC\bar{\nu}_e$ ,  $CC\bar{\nu}_\mu$ , for the  $\bar{\nu}_\mu$  beam. It will also be necessary to be able to distinguish background neutral current interactions that may occur. In this study, various machine learning (ML) techniques are deployed to address this challenge.

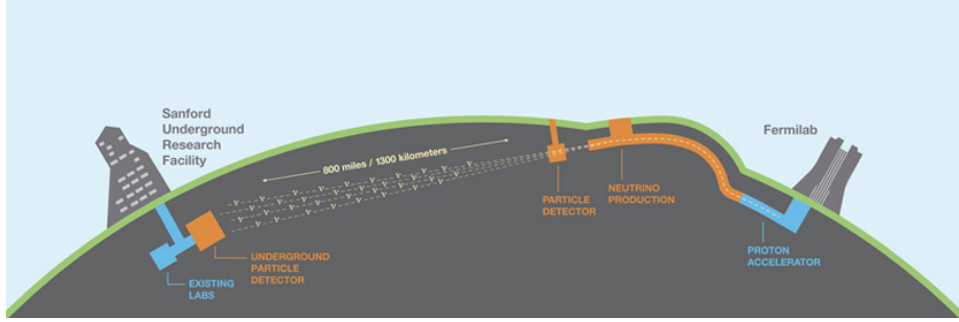


Figure 1: Illustration of the configuration of the far and near detectors, and the Long-Baseline Neutrino Facility (LBNF) at Fermilab [16]

## II. Theory

### A. DUNE Experimental Setup

DUNE will consist of a high intensity neutrino source, a far detector, and a near detector, the arrangement of which is depicted in Fig. 1.

#### A.1 The Neutrino Source

The high intensity neutrino source will be delivered by the Long-Baseline Neutrino Facility (LBNF) project hosted at Fermilab, and will deliver the world's most intense neutrino beam [16]. Many researchers contend that investigating neutrino oscillations over long baselines presents a promising approach to uncovering leptonic CP violation [17, 18, 19]. The DUNE beamline is capable of delivering 10 microsecond pulses of  $\nu_\mu$  or  $\bar{\nu}_\mu$  every 0.89 seconds [20], and has been specifically designed to allow for maximal sensitivity to CP violation [21].

The neutrino beam will be generated as follows: high-energy protons are directed towards a graphite target, where they collide and produce particles such as pions and kaons, which are then focused using magnetic horns and will decay into neutrinos [22].

#### A.2 The Near Detector

The near detector will be located 574m downstream of the neutrino beam source and consist of 3 main components: a Liquid Argon Time Projection Chamber (LArTPC), a multipurpose detector, and an on-axis beam monitor [16]. The main purpose of the near detector is to reduce systematic errors by precisely measuring the initial neutrino beam flux and flavour composition. By comparing the measured neutrino energy spectra at the beam source before any oscillation occurs with those at the far site, the various energy-

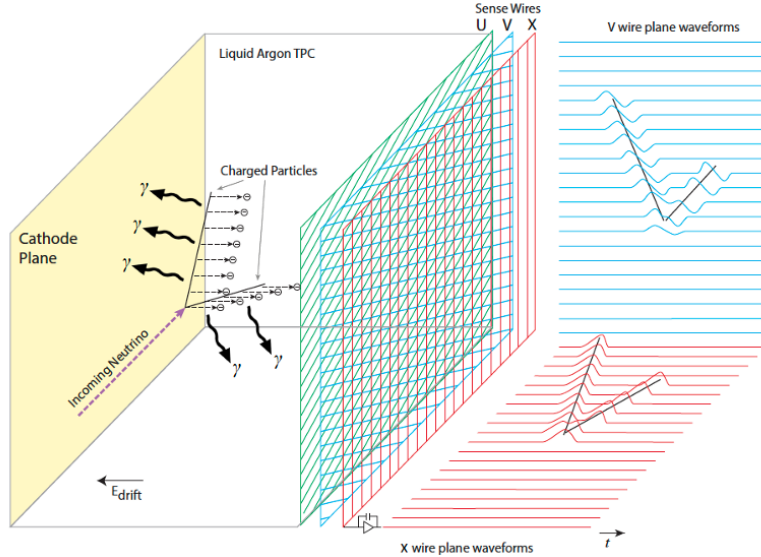


Figure 2: Diagram illustrating the key components of a LArTPC detector and an example interaction taking place in the detector. Interactions produce trails of ionisation electrons and scintillation photons. The electrons, under the influence of a horizontal electric field drift to the anode wire planes, resulting in electrical pulses on each wire, which are then reconstructed to obtain ‘images’ of the interaction. [16]

dependent effects that modulate the beam spectrum can be disentangled. This approach allows for the reduction of systematic uncertainties to the level necessary for discovering CP violation. [23]

### A.3 The Far Detector

The far detector is located 1300km downstream of the neutrino beam source,  $\sim 1.5$ km underground at the Sanford Underground Research Facility. The far detector was specifically positioned 1300km from the beam source to provide optimal sensitivity to CP violation and measurement of  $\delta_{CP}$ , while still being at a sufficient distance to investigate the neutrino mass hierarchy. At shorter separations, CP sensitivity is reduced by ambiguities from unknown mass ordering. At longer baselines CP sensitivity is harmed by matter effects that increase with distance [16, 24, 25].

The DUNE far detector will consist of four LArTPC detectors, each containing 17kt of liquid argon [26]. One of the main purposes of the far detector is to detect neutrino oscillations by comparing the neutrino energy spectra at the far and near detectors.

## B. LArTPC Detectors

The LArTPC was first proposed by C. Rubbia in 1977 as a concept for making precision neutrino physics measurements [27]. LArTPCs offer high-resolution imaging of neutrino interactions, enabling precise measurements of neutrino properties. They are capable of

tracking particle trajectories on sub-millimeter scales over volumes of the order of  $10^4\text{m}^3$  [20]. A diagram illustrating the key components of a LArTPC detector are shown in Fig. 2. The use of LArTPC detectors at DUNE, and the spatial and calorimetric resolution they provide, opens the possibility for state-of-the-art machine learning techniques such as convolutional neural networks to be used to accurately distinguish signal  $\text{CC}\nu_e$  and  $\text{CC}\nu_\mu$  interactions from background NC interactions.

Liquid argon is used because it is dense, transparent to scintillation light, and provides excellent ionisation properties. It serves as both the target for incoming particles and the medium for tracking their trajectories. When neutrinos from the source reach the liquid argon at the far detector, they interact with the argon nuclei, producing charged particles such as electrons, protons and muons. These interactions produce trails of ionisation electrons and scintillation photons. The scintillation light is detected by photodetectors, which helps in timing the events. The ionisation electrons, under the influence of an electric field passing across the chamber, drift to the anode wire planes. The anode wire plane consists of three layers of wire oriented at  $35.9^\circ$ ,  $-35.9^\circ$ , and  $0^\circ$  to the vertical, as depicted in Fig. 2. When an electron hits these wires they induce an electric pulse, which is collected and recorded. [28]

### B.1 Simulated Event Samples

The DUNE simulation uses the GENIE event generator [5, 29] to model neutrino-nucleus interactions and produce samples of  $\text{CC}\nu_e$ ,  $\text{CC}\nu_\mu$  and NC interactions. GENIE outputs the kinematics of the incoming neutrino, the target argon nucleus, and the final-state particles produced.

Particle propagation in the liquid argon is simulated using GEANT4 [30]. The simulation generates three-dimensional energy depositions within the detector, which are then mapped as electrical pulses onto three two-dimensional wire readout planes corresponding to the  $u$ ,  $v$ , and  $w$  views, at  $35.9^\circ$ ,  $-35.9^\circ$ , and  $0^\circ$ , to the vertical, respectively.

The raw simulated wire signals are then processed to reconstruct the original neutrino interaction, using software such as Pandora [31], and outputs physical quantities such as hit locations and energies. This study uses cheated reconstruction files in its analysis, where the reconstruction process is informed by the truth information from the simulation, and is the data upon which this analysis is performed.

### C. Neutrino Oscillations

Neutrino oscillations are a phenomenon in which a neutrino has a nonzero probability of changing flavour as it propagates. This was first observed in the 1960s, and confirmed by the Sudbury Neutrino Observatory in 2001 [32, 33]. In the current theory of neutrinos there are three flavour eigenstates ( $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$ ), and three mass eigenstates ( $\nu_1$ ,  $\nu_2$ , and  $\nu_3$ ). The flavour eigenstates of a neutrino are the states associated with the charged leptons observed in weak interactions; they are the states neutrinos are produced and interact as. The mass eigenstates are the states with definite mass and time evolution.

These flavour and mass eigenstates are related via the PMNS matrix as follows [34]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (1)$$

Non-zero off-diagonal elements, coupled with non-zero differences in the masses of  $\nu_1$ ,  $\nu_2$ , and  $\nu_3$ , lead to the aforementioned phenomenon of neutrino oscillations.

If neutrinos are Dirac fermions, meaning they obey the Dirac equation, then the PMNS matrix can be parametrised with one complex phase,  $\delta_{CP}$ , and three mixing angles,  $\theta_{12}$ ,  $\theta_{13}$ , and  $\theta_{23}$ . Note that if neutrinos are Majorana fermions there are additional non-trivial phases, but these do not affect oscillations at an observable level [13].

In the  $\nu_\mu \rightarrow \nu_e$  and  $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$  channels. The probability of a flavour change from  $\nu_\mu$  to  $\nu_e$  ( $P_{\mu e}$ ) is given by the following expression [13]:

$$P_{\mu e} \approx A_{atm}^2 + A_{sol}^2 + \cos \theta_{13} A_{atm} A_{sol} B(\delta_{CP}) \quad (2)$$

where  $A_{atm} = \sin \theta_{23} \sin 2\theta_{13} \sin(\frac{\Delta m_{31}^2 L}{4E})$ ,  $A_{sol} = \cos \theta_{23} \sin 2\theta_{12} \sin(\frac{\Delta m_{21}^2 L}{4E})$  and  $B(\delta_{CP}) = \cos(\frac{\Delta m_{31}^2 L}{4E} + \delta_{CP})$  and similarly for the antineutrino mode:

$$P_{\bar{\mu} \bar{e}} \approx A_{atm}^2 + A_{sol}^2 + 2 \cos \theta_{13} A_{atm} A_{sol} B(-\delta_{CP}) \quad (3)$$

If  $\delta_{CP} \neq 0$  or  $\pi$ , and assuming non-zero  $U_{\alpha i}$  ( $\alpha = e, \mu, \tau$ ;  $i = 1, 2, 3$ ), then the PMNS matrix has complex components, indicating CP violation in the leptonic sector [16]. This manifests as different oscillation probabilities in neutrino and antineutrino channels, and  $P_{\mu e} \neq P_{\bar{\mu} \bar{e}}$  should be observed. Measurement of  $P_{\mu e}$  and  $P_{\bar{\mu} \bar{e}}$  requires the

accurate classification of particles in the far detector, making the accurate identification of events a crucial component for the precise measurement of  $\delta_{CP}$ .

## D. Particles in LArTPC

Neutrinos cannot be directly observed, instead they are detected by their weak interactions. Therefore, in order to identify an interaction as  $CC\nu_\mu$ ,  $CC\nu_e$ , or NC, it will be necessary to be able to identify muons and electrons in the LArTPC imaging detectors. This is because a  $CC\nu_\mu$  can be identified by a tier one muon in the event, a  $CC\nu_e$  can be identified by a tier one electron in the event, and a NC can be identified by an absence of either. A tier one particle refers to particles which are direct products of the neutrino interaction.

Below, some theory is outlined that will help in making these distinctions at the energies relevant to the DUNE oscillation analysis.

Note that tau interactions at the beam energies studied at DUNE are rare and therefore will not be considered in this study.

### D.1 Tracks

Muons are easily distinguishable from electrons and photons in a LArTPC detector because they exhibit a track-like topology, unlike the shower-like topology seen in photons and electrons as a result of electromagnetic showers (see D.3), so one key component of event classification will be to identify a particle as a track or a shower.

Particles such as muons, pions and protons exhibit a track like topology in the LArTPC detector. Tracks can be characterised by their focused hits along a principal axis which is a result of these particles interacting mainly via ionisation, and their high mass relative to electrons.

Particles which have track-like topologies in the LArTPC lose energy predominantly through ionisation as they travel through the liquid argon. Ionisation involves knocking electrons out of the argon atoms, producing a trail of localised charge along its path. The result is a series of hits which is reconstructed as a linear track.

Muons, pions and protons have a higher mass than electrons meaning they are less susceptible to random scattering, and maintain their directionality better leading to the straight paths observed in the LArTPC. On the other hand, showers involve widespread, branching interactions (see D.3), leading to hits in the LArTPC that are more sparsely

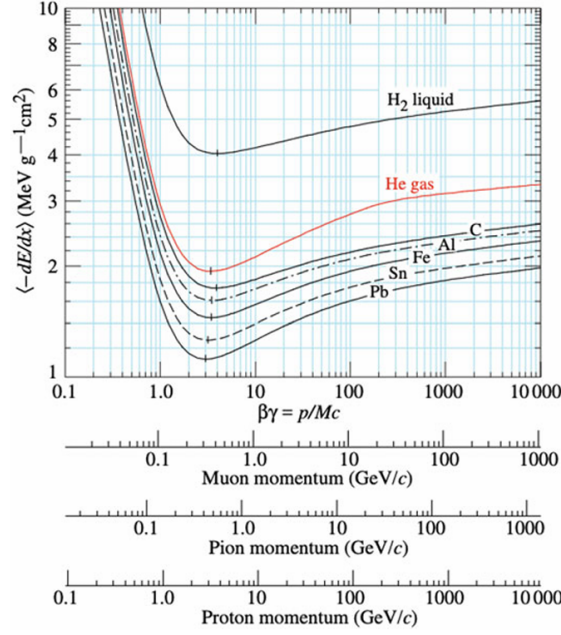


Figure 3: Bethe-Bloch curve as a function of momentum for muons, pions and protons in various media [37].

distributed away from it's principal axis.

Another characteristic of tracks is the energy deposition profiles that they display towards the end of their paths. The mean energy deposited per unit length is given by the Bethe-Bloch formula:

$$\left\langle -\frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (4)$$

where  $x$  is the distance travelled through the medium,  $\beta = v/c$ ,  $\gamma = (1 - \beta^2)^{-1/2}$ ,  $W_{\max}$  is the maximum possible energy transfer in a collision,  $\delta(\beta\gamma)$  accounts for impact the density of the medium has on the energy loss,  $I$  is the average ionisation potential of the atoms,  $Z$  is the medium's atomic number,  $z$  is the incident particle's charge,  $A$  is the atomic mass of the medium, and  $K$  is a constant [36].

As muons, pions and protons deposit energy, they move towards the low-energy rise in the Bethe-Bloch curve shown in Fig. 3. This manifests itself in a large energy deposition towards the tail end of their trajectories. This phenomenon is referred to as a Bragg peak [38] and is another key differentiator of tracks from showers. An illustration of a Bragg peak can be seen in Fig. 4.



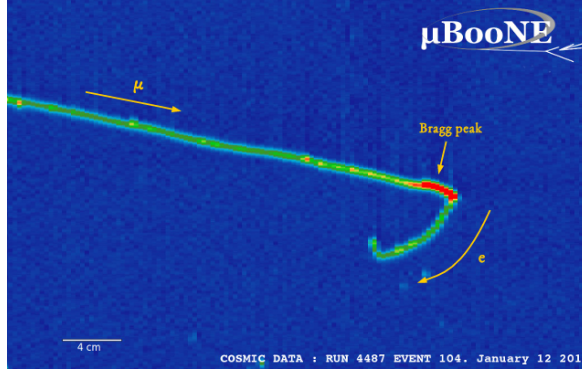


Figure 4: Event display for the process  $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$ . The neutrinos are not detectable so the only observed product is the Michel electron at the end of the Bragg peak. Here, colour is scaled off energy, with red indicating increased energy deposition [39]

## D.2 Muons

DUNE must still distinguish between muons and other particles that also have track-like topologies. One way this can be achieved is by looking at energy deposition rates. Muons in the DUNE energy range will predominantly be produced in the minimally ionising range and have energy deposition rates of approximately 2.1 MeV/cm in argon. Particles, such as protons, which are heavier than muons are likely to be created towards the low-energy rising region of Fig. 3. As a result, muons typically display longer (frequently leaving the detector), lower ionising tracks than those of heavier particles such as protons, and is one of the key ways in which we can distinguish muons from other tracks [40].

Despite muons being minimally ionising, in some cases they can decay within the LArTPC. The dominant muon decay mode is  $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$  [42]. The child electron that is produced is called a Michel electron. The observation of a Michel electron at the end of a long track is a signature of a muon, and can be seen in the event display in Fig. 4.

Additionally, muons may produce delta rays as they travel by scattering with electrons at rest. These secondary electrons, or delta rays, create short tracks that emanate from the primary muon track, providing further evidence for the presence of a muon.

Identifying these characteristics helps us accurately distinguish muons from other particles in a LArTPC detector.

## D.3 Electromagnetic Showers

An electromagnetic (EM) shower occurs when a high-energy electron, positron or photon passes through the liquid argon in the LArTPC detector. At high energies, the photoelec-

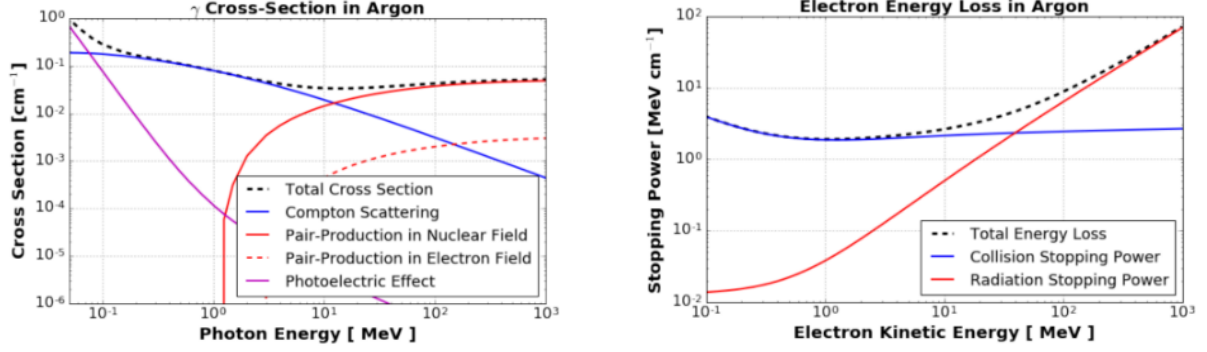


Figure 5: (Left) Cross section of photon interactions in argon for (red) pair-production in nuclear field, (red dotted) pair-production in electron field, (blue) Compton scattering, and (pink) photoelectric effect. (Right) Energy loss of electrons in Argon, for both (red) radiation stopping power and (blue) collision stopping power. [40]

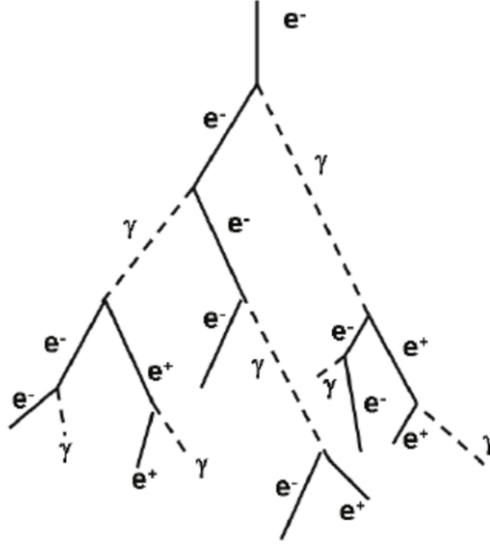


Figure 6: Cascading EM shower. The initial electron emits a photon via bremsstrahlung, which subsequently pair produces into an electron positron pair. This process is repeated forming a branching EM shower. [41]

tronic effect and Compton scattering are insignificant, and photons interact mainly via pair production, as illustrated in Fig. 5. Similarly, above a critical energy, electron energy loss is dominated by bremsstrahlung: the emission of photons due to the deceleration or deflection of a charged particle, such as an electron, by another charged particle, typically a nucleus [43]. These processes (pair production and bremsstrahlung) continue, resulting in a cascade of particles with decreasing energy until the photons no longer have sufficient energy to pair produce, and the electron energy drops below the critical threshold. An illustration of this process is seen in Fig. 6.

In our case, it will be necessary to distinguish between an electron induced EM shower, and a photon induced EM shower. A photon induced shower first travels a short distance

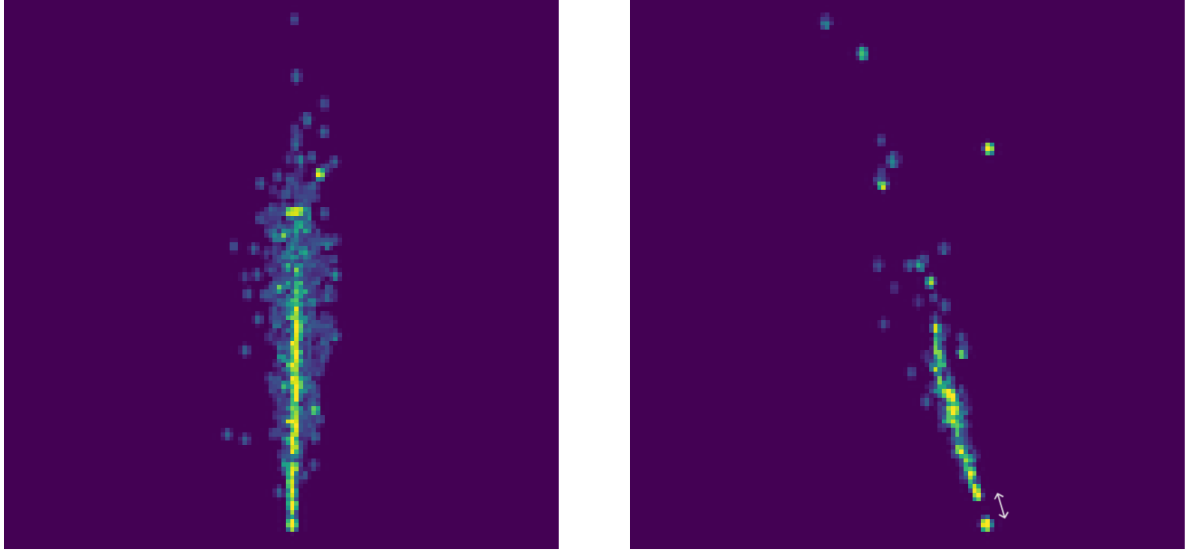


Figure 7: (Left) Display for an electron induced EM shower from simulated data files. (Right) Display for a photon induced EM shower, with the conversion gap highlighted.

before pair production occurs, resulting in a conversion gap between the interaction vertex and the start of the observed shower, as seen in the display in Fig. 7. One can also differentiate between the two types of showers by examining the energy deposition at the start of the shower. Since the photons convert via pair production, the energy deposition per unit length at the start of a photon induced shower ( $dE/dx \approx 4.6$  MeV/cm) is roughly double that of an electron induced shower ( $dE/dx \approx 2.3$  MeV/cm) [44].

### III. Methodology

#### A. Performance Characterisations

##### A.1 Bhattacharyya Coefficient

The Bhattacharyya coefficient (BC) is a measure of similarity between two probability distributions and is commonly used in statistics and machine learning [45]. It is defined as:

$$BC = \int \sqrt{F(x) \cdot Q(x)} dx \quad (5)$$

where  $F(x)$  and  $Q(x)$  are two probability density functions (PDFs), and  $BC$  is the Bhattacharyya coefficient. A  $BC$  close to zero indicates no overlap while a  $BC$  close to one indicates maximal overlap.

### A..2 Receiver Operator Characteristic Curves

A receiver operating characteristic (ROC) curve is a graphical representation used to evaluate the performance of a binary classification model [46]. When used as a performance classification in this paper, the ROC curve will be presented as a plot of purity against efficiency. Where efficiency is defined as:

$$\epsilon = \frac{TP}{TP + FN} \quad (6)$$

and purity is defined as:

$$\zeta = \frac{TP}{TP + FP} \quad (7)$$

where  $\epsilon$  is the efficiency,  $\zeta$  is the purity,  $TP$  is the number of true positives,  $FN$  is the number of false negatives, and  $FP$  is the number of false positives.

### A..3 F1 Score

The F1 score is a commonly used performance metric for assessing the performance of a ML algorithm in classification tasks [47]. It can be interpreted as the harmonic mean of the efficiency and impurity and is defined as follows:

$$F1 = \frac{TP}{TP + 0.5(FP + FN)} \quad (8)$$

where  $TP$ ,  $FP$ , and  $FN$  are the same as defined in 6 and 7. An F1 score has a maximum value of 1, representing perfect efficiency and purity, and a minimum of 0, indicating the worst performance.

## B. ML Methods

Below we outline the four supervised machine learning approaches that will be evaluated for the task of track-shower classification. The performance of these approaches will inform the selection of an approach for classifying the entire event.

The naive Bayes classifier (B..1), boosted decision tree (B..2), and artificial neural network (B..3) will make classifications based on variables for each particle, as defined in Section IV.A.. The CNN, however, will be trained on 128x128 single-channel images of individual particles in the  $(w, x)$  view.

### B..1 Naive Bayes Classifier

The naive Bayes classifier (NBC) is a widely used supervised machine learning technique, effective in distinguishing signal from background [48]. This method is based off Bayes theorem, expressed as:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)} \quad (9)$$

where, in the context of this study,  $y \in \{0, 1\}$ , indicates a particle being a track ( $y = 1$ : signal) or a shower ( $y = 0$ : background). The feature vector  $X = (x_1, x_2, \dots, x_n)$ , comprises of variables that encapsulate relevant physical characteristics of the particle.  $P(A)$  represents the probability of outcome  $A$ , and  $P(A|B)$  denotes the conditional probability of  $A$  given  $B$ .

The likelihood is then defined as:

$$\mathcal{L}_y = \frac{P(y|X)}{P(y|X) + P(y^c|X)} = \frac{P(X|y)P(y)}{P(X|y)P(y) + P(X|y^c)P(y^c)} \quad (10)$$

where  $y^c$  refers to the complement of  $y$ .

The simplifying assumption of independence among feature variables is made, often referred to as the ‘naive’ assumption, so that:

$$P(X = (x_1, x_2, \dots, x_n)|y) \propto \prod_{i=1}^n P(x_i|y) \quad (11)$$

where  $P(x_i|y)$  is derived from the individual PDFs for each variable.

Provided the input variables  $x_1, x_2, \dots, x_n$  exhibit weak correlations and the model is free of inaccuracies, such as unrepresentative PDFs, the expression in (10) achieves optimal track-shower separation for the chosen input variables [49].

### B..2 Boosted Decision Trees

A decision tree, illustrated in Fig. 8, is a classification algorithm structured as a sequence of branching nodes in a tree-like formation. At each node, the input variable that provides the optimal separation between signal and background is used to perform a binary split of the data at that node. Each branch ultimately terminates at a leaf node, which is classified as either signal or background based on the majority of events reaching that node. [49]

The term ‘boosting’ refers to an ensemble learning approach aimed at enhancing the

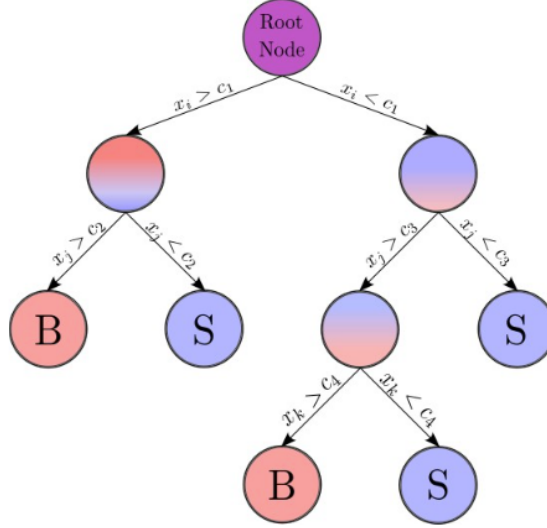


Figure 8: Illustration of a simple decision tree for a binary classification task classifying data as either signal (S) or background (B). At each node, a cut ( $c_{1,2,3,4}$ ) is applied to one of the input variables ( $x_{i,j,k}$ ) which splits the data at that node into two branches. The data passes down the tree subjected to cuts at each node until it reaches a leaf (terminating) node, where the classification is based on the majority of data at that node.

performance of weak classifiers by iteratively combining them into a more robust model. It operates by sequentially training multiple decision trees, with each tree focussed on correcting the errors made by the previous ones. Boosting improves the classification capability of a decision tree. In this study, the AdaBoost boosting algorithm is used [50].

Boosted decision trees (BDTs) allow for a more sophisticated separation of the  $n$ -dimensional phase space of variables compared with the NBC. In the NBC approach, a single cut splits the phase space into two regions, whereas BDTs are able to spit the phase space into numerous hypercubes, enabling a more detailed discrimination between signal and background [49]. It is also possible to learn different levels of importance for each variable and weight them accordingly rather than the uniform weighting in the NBC which provides further enhanced discrimination.

### B.3 Artificial Neural Networks

An artificial neural network (ANN) is a structure of interconnected nodes, referred to as neurons, designed to simulate the human brain. The architecture of an ANN consists of an input layer, hidden layers, and an output layer, as illustrated in Fig. 9. The input layer receives the input variables, with each neuron corresponding to a specific variable. The data is then propagated through one or more hidden layers, and ultimately to the output layer, which provides a classification prediction of either signal or background.

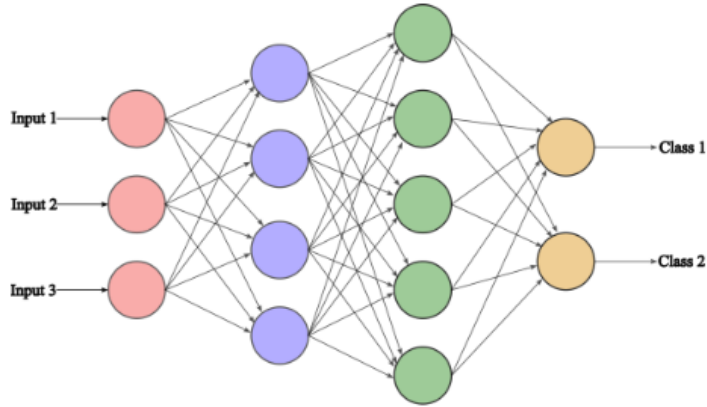


Figure 9: Illustrated example of a typical ANN structure. In this example the ANN receives three input variables to the input layer (red) which then passes through two hidden layers, the first containing four nodes (blue), and the second containing five nodes (green), before passing to the output layer (yellow) where it is classified as Class 1 or Class 2.

At each neuron within the hidden and output layers, a weighted sum of inputs from the preceding layer is computed, followed by the application of an activation function. During the training process, the weighting of connections between neurons are iteratively updated to optimise the network’s predictive performance. As with the BDT, an ANN allows for a more sophisticated splitting of the phase space of input parameters, in theory allowing it to exceed the performance of a NBC approach. [51]

In this analysis, the scikit-learn MLPClassifier was used, with the Adam optimisation algorithm [52], the logistic sigmoid activation function, and a network structure consisting of two hidden layers with ten neurons in the first layer and five neurons in the second layer.

#### B.4 Convolutional Neural Networks

A convolutional neural network (CNN) is a class of neural networks designed specifically for image recognition tasks and is particularly effective at capturing hierarchical and spatial features within images [53]. A CNN is composed of several layers, three of the key ones being: convolutional layers, pooling layers, and fully connected layers. The convolutional layer consists of a number of filters that are applied to the image via convolutions, enabling the detection of image features at various levels of complexity. Pooling layers reduce the spatial dimensions of the input image, reducing the computational complexity and number of parameters in the model. Fully connected layers, consist of neurons connected to all nodes in the preceding layer. These layers aggregate the extracted features and are used by the network to make final predictions.

In this study, shallow CNN architectures were used, consisting of only two convolutional layers in each model.

## IV. Results

### A. Variable Construction and Performance

A CC interaction contains a charged lepton with the same flavour as the incoming neutrino. In the LArTPC, a final state muon typically produces a long, straight trajectory, whereas an electron generates an EM shower. Consequently, accurately distinguishing between tracks and showers is a key component of event classification at DUNE.

In the study of track-shower separation, five variables were constructed and calculated from the simulated dataset to capture the calorimetric and topological properties that distinguish tracks and showers. These variables will form the basis for a naive Bayes classification method, whilst also being used in the training of a BDT and an ANN for the identification of particles as either tracks or showers.

The motivation for the selection of these variables is presented in the sections below, alongside a description of their implementation. Each section also includes a normalised histogram, illustrating the track-shower discrimination that was achieved with each of these variables. These histograms will serve as PDFs in the naive Bayes classifier approach to calculate  $P(x_i|y)$  in Eq. 10.

These histograms were generated by calculating each of the variables for 23,000 showers and 32,000 tracks, and placed into 20 bins.

#### A.1 RMS/Length

The RMS/length variable was defined to be the root mean square (RMS) of the perpendicular distances of each hit from the principal axis, normalised by the total trajectory length of the particle. The principal axis was determined by performing linear regression on the first ten hits to estimate the particle's direction of travel. The length of the trajectory was defined as the distance between the first hit and the hit where 86.9% of the particle's total energy had been deposited. This cutoff value of 86.9% was optimised using the Bhattacharyya coefficient to achieve a maximal track-shower separation, resulting in a BC score of 0.24.

This variable effectively captures the concentrated nature of track hits along a princi-



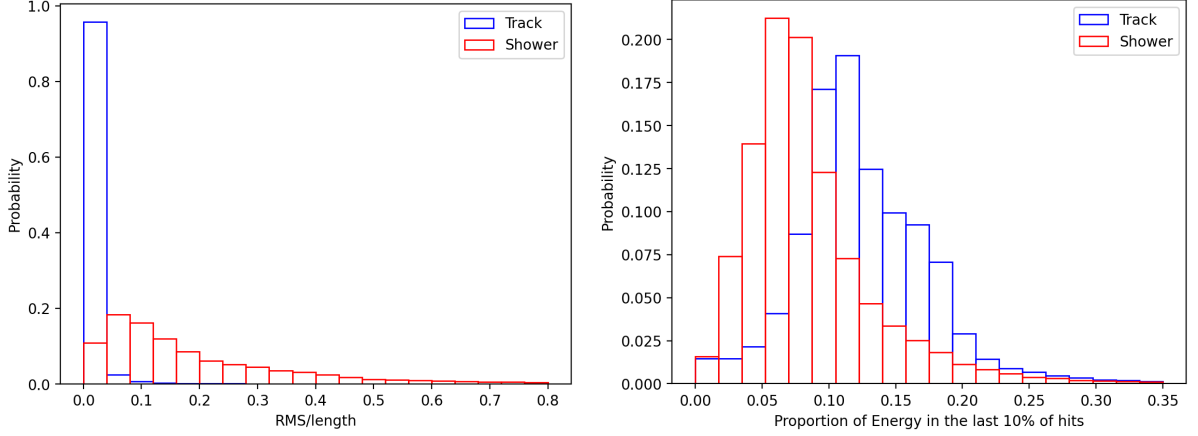


Figure 10: (Left) Normalised histogram showing RMS/length values for tracks (blue) and showers (red). (Right) Normalised histogram showing the proportion of total energy deposited in the last 10% of hits for tracks (blue) and showers (red).

pal axis compared to the more dispersed distribution observed in showers. This separation is illustrated in Fig. 10, in which RMS/length values for tracks are predominantly concentrated near zero, while values for showers are more broadly distributed.

## A.2 Residual Energy

As previously discussed, tracks and showers display distinct energy deposition profiles within the LArTPC. Specifically, tracks exhibit a rapid increase in  $dE/dx$  as they come to a stop, forming a pronounced Bragg peak. The residual energy variable was designed to be sensitive to this phenomenon and defined as the fraction of the particle's total energy deposited in the last 10% of it's hits. Similar to RMS/length, the 90% cutoff was selected to optimise track-shower PDF separation using the BC metric, resulting in a score of 0.73.

As expected, the results show that on average tracks deposit a higher fraction of their total energy in the last 10% of hits, as can be seen in Fig. 10.

One limitation of this variable arises from minimally ionising particles, such as muons, which often exit the detector without stopping. These particles, existing in the high-energy, flat end of the Bethe-Bloch curve (Fig. 3), exhibit consistent energy deposition along their path in the detector, lacking an observed Bragg peak. This contributes to the reduced BC score compared with RMS/length.

## A.3 Angular Distribution

The angular distribution variable captures the spatial spread of hits relative to the principal axis of the particle's trajectory. It is defined as the mean of the absolute angular

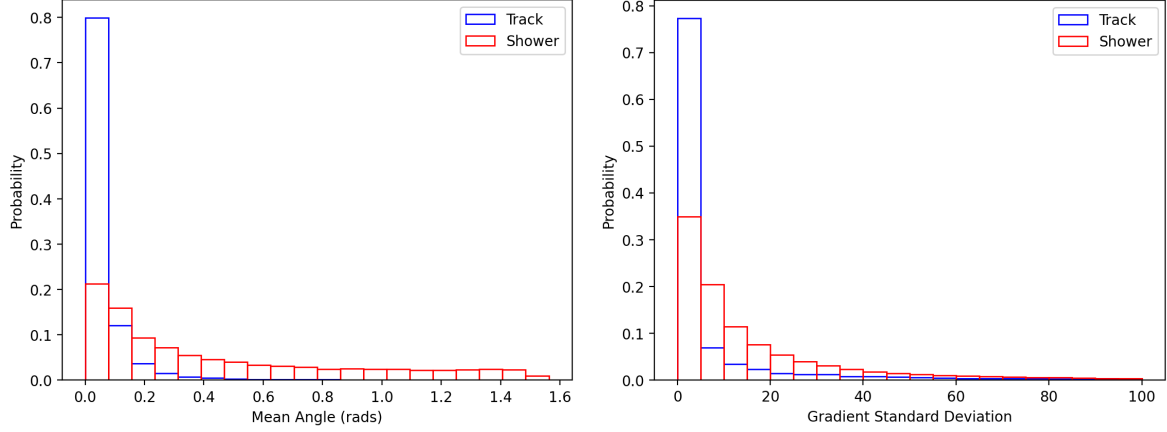


Figure 11: (Left) Normalised histogram showing the values of mean angle from the principal axis for tracks (blue) and showers (red). (Right) Normalised histogram showing the standard deviation of the gradients between every other hit for tracks (blue) and showers (red).

deviation of each hit from the principal axis, with the absolute values ensuring symmetry around the principal axis does not reduce the calculated mean.

Fig. 11 demonstrates the separation achieved by this variable, with track values concentrated near zero and shower values occupying higher bins spread over a wider range, yielding a BC score of 0.43.

#### A..4 Gradient Method

The gradient variable quantifies the variability of the particle’s trajectory by calculating the standard deviation of gradients between every other hit. Tracks exhibit greater variability in the gradients between hits due to their branching nature, compared to tracks which should produce a narrower distribution of gradients. The gradient variable aims to gain exposure to this distinction.

As shown in Fig. 11, the gradients for tracks are typically narrowly distributed, resulting in a lower standard deviation compared with showers, achieving a BC score of 0.55.

#### A..5 Number of Endpoints

As previously discussed, showers can be characterised by widespread branching interactions, typically exhibiting multiple endpoints, whereas tracks tend to follow a continuous paths with one, or sometimes zero, endpoints. To capture this property, the endpoints variable was constructed to count the proportion of hits classified as endpoints.

A hit was labeled an endpoint if no other hits were located within a semicircle of radius

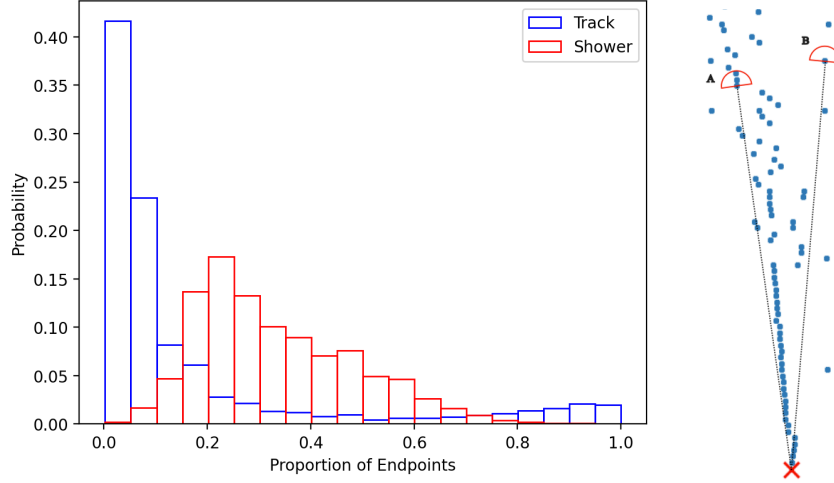


Figure 12: (Left) Normalised histogram showing the proportion of hits which are labelled endpoints for tracks (blue) and showers (red). (Right) Illustrated example of how a hit is determined to be an endpoint for an example shower.

$r$ , drawn normal to the particles first hit. The right hand figure in Fig. 12 illustrates this for a typical shower. For example, Point A in Fig. 12 does not qualify as an endpoint due to the presence of neighbouring hits within the localised semicircle, whereas Point B does satisfy the endpoint condition. The radius  $r$  was set as a global constant  $r = 1.1 \times \langle d \rangle$ , where  $\langle d \rangle$  was the average distance between consecutive hits for a sample of 30,000 tracks.

This variable provided the greatest track-shower PDF separation, achieving a BC score of 0.19. This separation is illustrated in the left hand figure in Fig. 12.

As expected, a higher proportion of showers' hits are labelled as endpoints. Notably however, the bins for tracks at the higher end of the scale, close to 1 are non-empty. This arises because some high energy tracks produced in the LArTPC can travel significant distances before interacting with the argon. These wider hit separations result in nearly all hits being labelled as endpoints. Muons, in particular, are prone to this behaviour as they typically generate long, sparse tracks.

## A.6 Variables Summary

As discussed in section B.1, Eq. 10 enables optimal track-shower separation provided the input variables are sufficiently decorrelated. The pairwise correlations among the input variables are presented in Table 1. Overall, the variables demonstrate weak correlations, meaning the naive assumption made in the NBC method is valid.

The residual energy variable exhibits a negative correlation with all other variables. This result aligns with its expected behaviour being the only of the five variables for which a higher value is characteristic of a track. This observation supports the validity of the

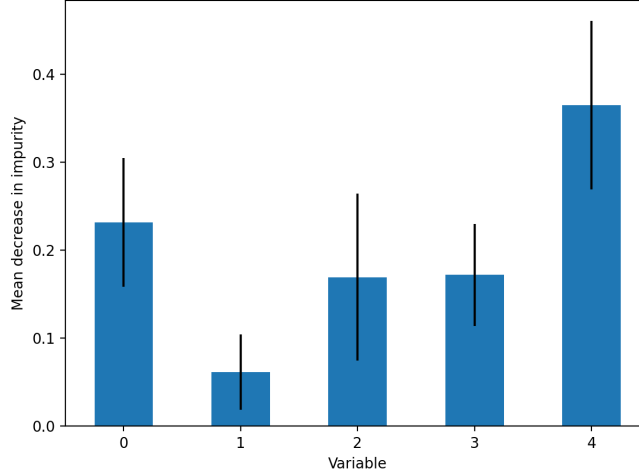


Figure 13: Variable importance as determined by a BDT. The importance values are calculated as the mean impurity decrease accumulated across all trees, with uncertainties represented by the standard deviation. The BDT was trained using 70,000 events for the classification of tracks and showers. Variables 0, 1, 2, 3, and 4 correspond to the RMS/length, residual energy, angular distribution, standard deviation of gradients, and number of endpoints variables respectively.

variable definitions and their implementations.

	$d_{\text{RMS}}$	$dE/dx$	$\theta_{\text{dist}}$	$\sigma_{\text{gradients}}$	$n_{\text{endpoints}}$
$d_{\text{RMS}}$	1.000	-0.051	0.152	-0.001	0.044
$dE/dx$	-0.051	1.000	-0.219	-0.002	-0.186
$\theta_{\text{dist}}$	0.152	-0.219	1.000	0.003	0.225
$\sigma_{\text{gradients}}$	-0.001	-0.002	0.003	1.000	-0.060
$n_{\text{endpoints}}$	0.044	-0.186	0.225	-0.060	1.000

Table 1: Correlation matrix showing the correlation between each of the input variables. Where  $d_{\text{RMS}}$  corresponds to the RMS/length,  $dE/dx$  corresponds to the residual energy,  $\theta_{\text{dist}}$  corresponds to the angular distribution,  $\sigma_{\text{gradients}}$  corresponds to the standard deviation of gradients and  $n_{\text{endpoints}}$  corresponds to the number of endpoints.

The ranking of variable importance was analysed when training a BDT with these variables as inputs. The resulting variable importances are presented in Fig. 13, where importance is defined as the mean decrease in impurity. As anticipated, the number of endpoints and RMS/length variables exhibited the largest mean decreases in impurity, consistent with their low BC scores and their strong discriminatory power in track-shower separation. Conversely the residual energy variable demonstrated the smallest mean decrease in impurity, which is attributable to its high BC score, and the practical limitations related to the detector topology discussed in Section A.2.

## B. Track vs Shower

In this section the performance of the track-shower classification task of different supervised ML approaches is compared, and subsequent justification is provided for proceeding

with a CNN based approach for the event classification task.

Fig. 14 presents ROC curves comparing the combined efficiency and purity at a selection of cuts for the NBC, BDT, ANN, and CNN, and Table. 2 summarises the F1 scores achieved by each approach. The CNN demonstrated the highest performance in the track-shower classification task, achieving the greatest joint efficiency and purity on the ROC curve, as well as the highest F1 score.

<b>ML Approach</b>	<b>F1 Score</b>
NBC	0.929
BDT	0.957
ANN	0.954
CNN	0.977

Table 2: F1 scores of each ML approach in the track shower classification task. F1 scores were calculated on a test data set containing 30,000 total images with 15,000 tracks and 15,000 showers. Errors in the results presented here were calculated to be negligible using binomial statistics and hence omitted.

The BDT and ANN shared similar performance, with F1 scores of 0.957 and 0.954, respectively. Both outperformed the NBC approach, illustrating the improved performance that can be achieved by taking a more sophisticated separation of the 5-dimensional input parameter space and the ability to filter inputs through complex networks of weights, rather than the single slicing and uniform weighting used in the NBC.

The CNN delivered by far the strongest performance, obtaining an F1 score of 0.977, beating the other methods that relied on the input variables detailed in Section IV.A.. This underscores both the promise and challenges associated with CNNs. The track-shower CNN consisted of approximately 2 million trainable parameters, allowing it to learn deeper, more complex features than the five manually engineered features that were fed into the NBC, BDT, and ANN.

This does however raise the question: what features is the CNN learning to achieve this superior performance? The CNN was trained and tested on single-channel 128x128 images in the  $(w, x)$  view, with each hit’s colour scaled by hit energy. Based on this input, one would expect the CNN is limited to learning topological and calorimetric features, both of which however were captured by the variables fed into the NBC, BDT and ANN. The precise ‘physics’ underlying the improved distinction remains uncertain and is a common drawback to using a CNN. To mitigate this it is important to ensure the training datasets are sufficiently large and representative to avoid over fitting the data rather than learning the physical patterns underpinning the particles true classification.

Potential next steps in this study could involve a comprehensive investigation into the

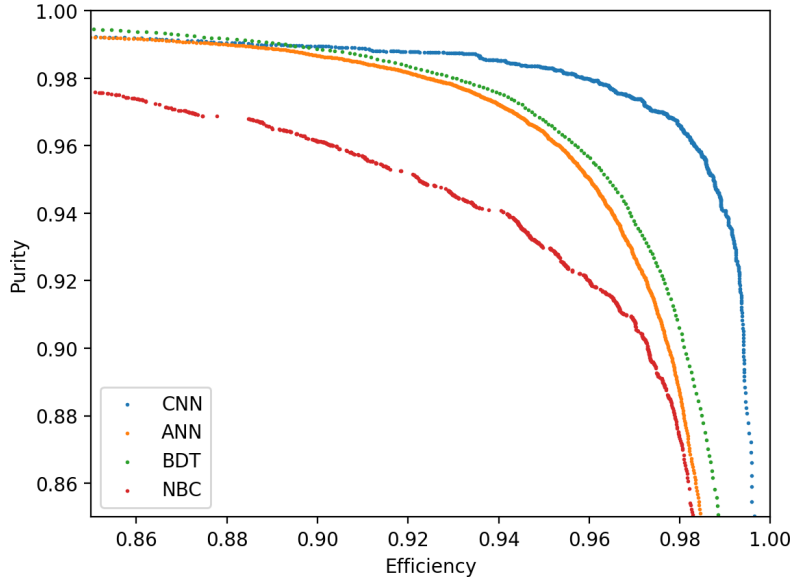


Figure 14: ROC curves comparing the efficiencies and purities achieved by the NBC (red), BDT (green), ANN (orange), and CNN (blue) for the track-shower classification task. ROC curve was formed by calculating the efficiency and purity for a selection of cut values, on a test data set of 15,000 tracks and 15,000 showers.

robustness and generalisability of the CNN model developed here and in subsequent parts of this paper. One of the primary concerns is the extent to which the CNN’s performance on simulated data will translate to real data once DUNE becomes operational. If the simulated data used to train these CNNs are unrepresentative of the complexities and noise inherent in real world data, the observed classification performance may drop significantly. A detailed robustness study could be conducted, where the simulated data is intentionally distorted to mimic potential real world deviations. This could include variations in noise level, energy deposition, or event topologies. Such an investigation could help quantify the CNN’s resilience and identify any vulnerabilities that could compromise its effectiveness.

Track-shower classification is a key tool for distinguishing muons and electrons, making it essential for the classification of events as  $CC\nu_e$ ,  $CC\nu_\mu$  or NC. The CNN’s superior performance in the track-shower classification task provides a strong foundation for event classification. Furthermore, its ability to learn complex and abstract features position it well for tackling more complex classification tasks.

### C. Event Classification

The CNN’s demonstrated superior performance in the track-shower classification task, meant that a CNN-based approach was employed to classify events as either  $CC\nu_e$ ,  $CC\nu_\mu$  or NC. The adopted methodology is illustrated in Fig. 15. This approach leveraged the

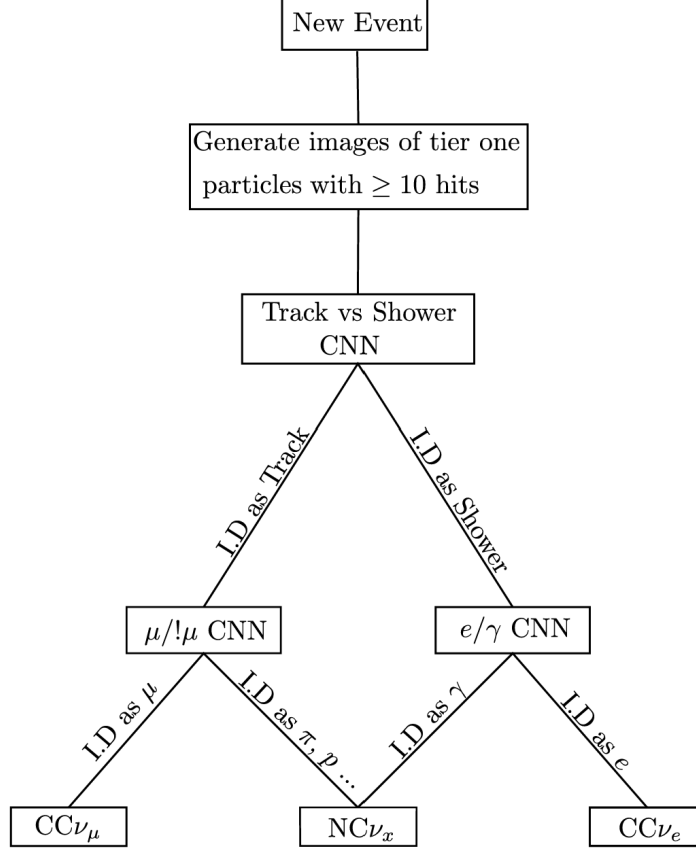


Figure 15: Classification approach for identifying events as  $CC\nu_e$ ,  $CC\nu_\mu$  or  $NC\nu_x$ . All tier-one particles with more than 10 total hits in an event are processed through the decision tree. Particles are assigned a predicted probability of being either a muon or not a muon (for those following the track branch of the tree), or an electron or photon (for those following the shower branch). The particle within the event with the highest probability of being a leading lepton (muon or electron) is subjected to a cut, specific to its classification as a muon or electron. The cut thresholds were informed by ROC curves plotted for each of the component CNNs and were chosen to maximize efficiency and purity in the respective muon and electron identification tasks.

high level of accuracy achieved in track-shower classification in a multi-CNN decision tree and required the training of two additional CNNs: one trained to classify tracks as muons or non-muons and another trained to classify showers as electrons or photons.

For event classification, only tier-one particles were considered, as the classification process relies exclusively on identifying this specific subset. Focusing the CNN training on tier-one particles enabled the model to learn the unique patterns characteristic of these interactions. Additionally, particles with fewer than 10 hits were excluded from the analysis, as such particles were typically unrepresentative and posed challenges for the CNNs during training.

The muon CNN, the electron CNN, and the track-shower CNN discussed in the previous section, were all trained on images generated in the same way. Several preprocessing steps were taken in the image generation process to highlight key physical features which

will allow the CNN to deliver strong performance on a smaller network, with lower statistics datasets.

Firstly, energy scaling was applied at a global fixed scale. This facilitated the identification of Bragg peaks, which are commonly observed in tracks, for the track-shower CNN. In the electron identification CNN, this approach enhanced sensitivity to the differing energy deposition rates observed at the start of electron-induced showers compared to photon-induced showers. In the muon CNN, the fixed energy scaling enabled the identification of muons, which typically exhibit an energy deposition rate of 2.1 MeV/cm, and the distinction of muons from particles with heavier rest masses such as protons, which have higher energy deposition rates.

Secondly, the interaction vertex was marked in the images. This was primarily introduced to highlight the conversion gap that is observed in photon-induced showers, aiding in the distinction between electrons and photons in the electron CNN.

To account for particle trajectories of varying lengths, a variable length scaling was adopted in the image preprocessing. While a globally fixed length scaling was trialled to emphasise the elongated nature of muon tracks, and potentially detect any Michel electrons in the event of decay, it resulted in suboptimal performance for track-shower and electron-photon separation due to the potentially shorter path lengths of these particles within the LArTPC meaning that key features could not be resolved. The inability to find a fixed length scaling that could simultaneously capture long muon tracks and resolve features of shorter trajectories highlighted one of the limitations of this approach. A fixed length scale could potentially become viable with an increased image resolution of  $224 \times 224$  pixels or higher, which would mitigate the issue of unresolvable smaller trajectories. This is a potential area of improvement that could be explored in a further study with more computational resources.

### C..1 Muon CNN

The muon CNN, had approximately 1 million trainable parameters, and was trained on single-channel 128x128 images of tier-one tracks in the  $(w, x)$  view, using a dataset of 37,000 images, of which 40% were muons. The model's performance was evaluated on 17,000 images of the same form, with the results summarised in Table 3. The CNN correctly identified  $92.9\% \pm 0.0020$  of true muons as muons.

A particle was classified as a muon if the CNN assigned it a classification probability



of greater than 0.54. This threshold was found to be the point on the ROC curve that maximized combined efficiency and purity, achieving an efficiency of  $0.919 \pm 0.0019$  and a purity of  $0.907 \pm 0.0021$ .

	<b>True Muon</b>	<b>True Not Muon</b>
<b>ID as Muon</b>	$0.929 \pm 0.002$	$0.151 \pm 0.012$
<b>ID as Not Muon</b>	$0.071 \pm 0.009$	$0.849 \pm 0.004$

Table 3: Column normalised confusion matrix summarising the proportion of correct classifications of the muon CNN for tier-one tracks. The performance was evaluated on 17,000 single channel 128x128 images of tier-one tracks in the  $(w, x)$  view.

## C..2 Electron CNN

The electron CNN consisted of approximately 2 million trainable parameters, and was trained on single-channel 128x128 images of tier-one showers in the  $(w, x)$  view, using a dataset of 34,000 images, evenly split between electrons and photons. The model’s performance was evaluated on 17,000 images, with results summarised in Table 4. The CNN achieved an accuracy of  $0.940 \pm 0.0019$  for true electrons and  $0.900 \pm 0.0035$  for true photons.

A particle was classified as an electron if the CNN predicted it to be an electron with a probability greater than 0.57; otherwise, it was classified as a photon. As in the previous section, this threshold was selected to maximise efficiency and purity, achieving an efficiency of  $0.927 \pm 0.0024$  and a purity of  $0.873 \pm 0.0028$ .

	<b>True Electron</b>	<b>True Photon</b>
<b>ID as Electron</b>	$0.940 \pm 0.002$	$0.100 \pm 0.011$
<b>ID as Photon</b>	$0.060 \pm 0.007$	$0.900 \pm 0.004$

Table 4: Column normalised confusion matrix summarising the proportion of correct classifications for the electron CNN for tier-one showers. The performance was evaluated on 17,000 single channel 128x128 images of tier-one showers in the  $(w, x)$  view.

## C..3 Entire Event

The full event classification approach, as outlined in Fig. 15, was implemented, and its results are presented in Table 5. Overall this approach achieved an accuracy of  $0.801 \pm 0.004$ .

Although this method achieves an accuracy of  $0.928 \pm 0.005$  and  $0.884 \pm 0.006$  for  $CC\nu_e$  and  $CC\nu_\mu$  classifications respectively, its performance is notably less effective at

	<b>True NC</b>	<b>True <math>\nu_e</math></b>	<b>True <math>\nu_\mu</math></b>
<b>ID as NC</b>	$0.559 \pm 0.009$	$0.051 \pm 0.004$	$0.107 \pm 0.006$
<b>ID as <math>\nu_e</math></b>	$0.221 \pm 0.008$	$0.928 \pm 0.005$	$0.009 \pm 0.002$
<b>ID as <math>\nu_\mu</math></b>	$0.220 \pm 0.008$	$0.021 \pm 0.003$	$0.884 \pm 0.006$

Table 5: Column normalised confusion matrix summarising the performance of the entire event classification algorithm. The performance was evaluated on 10,000 events, with an even splitting between each. Uncertainties presented here were calculated using binomial statistics.

background event rejection due to several contributing factors. Firstly, the nature of the classification method introduces a bias toward identifying events as  $CC\nu_e$  or  $CC\nu_\mu$ , as the particle with the highest probability of being a muon or electron is prioritised. Secondly, since NC classification does not rely on identifying a specific tier-one particle, it cannot leverage the strong performance of the track-shower CNN, which contributes significantly to the accuracy of  $CC\nu_e$  and  $CC\nu_\mu$  identification.

To mitigate this, an alternative approach could involve training an additional CNN to identify the leading lepton candidate and subsequently passing that particle through the decision tree shown in Fig. 15. This would reduce the impact of the bias that arises from selecting the particle with the highest probability of being a muon or electron.

Due to computational limitations, the method employed in this study uses shallower networks with fewer trainable parameters (the largest network contained only 2 million trainable parameters) and a smaller number of training images compared with similar studies. Other common image classification network architectures such as ResNet50 and AlexNet have approximately 25 and 60 million parameters, respectively [54]. Despite these limitations, the performance achieved is comparable to results reported in other studies. A recent study by A. Chappell and L. H. Whitehead [55] reported an F1 score of 0.85 and 0.84 in  $CC\nu_e$  and  $CC\nu_\mu$  identification respectively, using a ResNet18 model architecture with the Kaiming initialisation scheme. The model was trained on 100k 254x254 images of entire events with the  $(w, x)$ ,  $(u, x)$ ,  $(v, x)$  views mapped into the RGB channels. Other comparable studies have been trained on over three million 500x500 triple channel images [56]. Despite having fewer, and lower quality images, the event classification algorithm outlined in this paper reported comparable F1 scores of  $0.874 \pm 0.001$  and  $0.836 \pm 0.001$  in the  $CC\nu_e$  and  $CC\nu_\mu$  channels respectively. The overall F1 score, however was lower than that of the referenced study due to the weaker performance in NC identification.

The study cited above [55] also reported the successful application of transfer learning for neutrino event classification, outperforming the Kaiming initialised network. Whilst

each component CNN within the multi-CNN framework presented here delivered strong performance, a transfer learning approach could be taken to fine tune these CNNs and enhance performance in their relevant classification tasks, leading to improved event classification efficiency.

Nonetheless, the method presented here does show significant promise. With more computational resources, it would be feasible to incorporate a larger number of training images, aiding the CNN in generalisation; employ higher-resolution three-channel images, enhancing the network’s ability to learn patterns; and utilise deeper architectures with more trainable parameters, enabling the extraction of more complex image features. Additionally, future investigations could explore the implementation of a CNN designed to identify leading lepton candidates, and only passing these candidates through the decision tree outlined in Fig. 15, which could substantially improve NC identification.

An entire event classification was attempted based on the predicted classification by a CNN trained on images of the whole event. However, an accuracy of only  $0.674 \pm 0.006$  was achieved, due to computational limitations related to the size of the network, the number of training images, and their resolution. The alternative multi-CNN approach presented here demonstrated stronger performance under these constraints, and presents a promising computationally inexpensive solution.

## V. Conclusion

This study presented and evaluated a ML based approach for neutrino event classification at DUNE. The primary objective was to enhance the precision of identifying  $CC\nu_\mu$ ,  $CC\nu_e$ , and NC interactions, which is essential for accurately measuring  $\delta_{CP}$ .

A comparative analysis of four ML techniques identified the CNN as the most effective model for track-shower classification, achieving an F1 score of 0.976. This superior performance highlights the CNN’s ability to learn complex spatial and calorimetric features inherent to the interactions within the DUNE LArTPC detectors. Additionally, a decision tree based event classification algorithm was presented, combining three CNNs, each tailored to a specific classification task. This multi-CNN approach achieved an accuracy of  $0.928 \pm 0.005$  and  $0.884 \pm 0.006$  in identifying  $CC\nu_e$  and  $CC\nu_\mu$  interactions respectively.

Despite this success, the study highlighted potential areas for improvement, particularly in background rejection, where the algorithm’s performance was comparatively weaker due to inherent biases in the implementation of the decision tree approach. An

improvement in this area would be necessary for the precise measurement of CP violation at DUNE. Future work could involve training an additional CNN to specifically identify leading lepton candidates to reduce bias and improve NC identification.

Another area recommended for further study was an examination into the robustness of the performance achieved by the methods in this paper and the extent to which this would translate when applied to real world data. The risks of over fitting the training data rather than learning patterns representing physical characteristics of neutrino interactions highlights the importance of this next step. In the event that the simulated data upon which this analysis was performed turns out to be unrepresentative of the true data observed at DUNE, this investigation could reveal significant discrepancies in performance. By proactively identifying and addressing these potential limitations, it will be possible to develop a more resilient and reliable classification framework suitable for real-world application at DUNE.

In summary, the multi-CNN classification framework presented here offers a computationally efficient approach capable of achieving a high level of accuracy in the  $CC\nu_e$  and  $CC\nu_\mu$  channels, whilst offering a promising avenue of exploration for accurate event display classification at DUNE and the precise measurement of  $\delta_{CP}$ .

## References

- [1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138, (1964)
- [2] K. Abe *et al.*, Nature **580**, (2020), p.339–344
- [3] B. Abi *et al.*, Eur. Phys. J. C **80** 978, (2020)
- [4] K. Abe *et al.*, Hyper-Kamiokande Design Report, (2018)
- [5] B. Abi *et al.*, DUNE Far Detector Technical Design Report, (2020)
- [6] Jena, Deepika. “The DUNE Experiment.” PoS, vol. HQL2023, Jul. 2024
- [7] A. Ankowski *et al.*, arXiv:1608.07853 [hep-ex], (2016)
- [8] P. Kangacker, Physics Reports **72**, (1981), p.185-385
- [9] P. Nath, P. F. Perez, Physics reports **441**, (2007), p.191-317
- [10] J. C. Pati, International Journal of Modern Physics A **18** 22, (2003), p.4135-4155
- [11] A. D. SAKHAROV, Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki: Pis’ma v Redaktsiyu, **5**, (1967), p.32-35
- [12] L. Canetti *et al.*, New J. Phys. **14**, (2012)
- [13] C.Giganti, S.Lavignac, M.Zito, *Neutrino oscillations: the rise of the PMNS paradigm*, (2017)

- [14] W. Buchmüller, R.D. Peccei, and T. Yanagida, Annual Review of Nuclear and Particle Science **55**, (2005), p.311-355
- [15] S. Davidson, E. Nardi, Y. Nir, Physics Report **466**, (2008), p.105-177
- [16] B. Abi *et al.*, JINST **15**, (2020)
- [17] M. Tanimoto, Phys. Rev. D **55** 322, (1997)
- [18] S. M. Bilenky, C. Giunti, W. Grimus, Phys. Rev. D **58**, (1998)
- [19] M. V. Diwan *et al.*, Phys. Rev. D **68**, (2003)
- [20] P. Laycock, EPJ Web Conf. **251** 03041, (2021)
- [21] Guinot, Miquel Nebot. "DUNE experiment physics." PoS, vol. NuFact2021, Mar. 2022
- [22] Fermilab, DUNE at LBNF Neutrino beam
- [23] R. Acciarri *et al.*, Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report, Volume 4 The DUNE Detectors at LBNF, (2016)
- [24] V. Barger *et al.*, International Journal of Modern Physics A **31** 07, (2016)
- [25] X. Qian, P. Vogel, Progress in Particle and Nuclear Physics **83** 1-30, (2015)
- [26] B. Abi *et al.*, JINST **15** T08009, (2020)
- [27] C. Rubbia, The Liquid Argon Time Projection Chamber: A New Concept for Neutrino Detectors, (1977)
- [28] R. Acciari *et al.*, [arXiv:1612.05824v2 [physics.ins-det]], (2017)
- [29] C. Andreopoulos *et al.*, "The GENIE Neutrino Monte Carlo Generator," Nucl. Instrum. Meth. **A614** 87–104 (2010)
- [30] S. Agostinelli *et al.*, GEANT4: a simulation toolkit. Nucl. Instrum. Methods A **506**, p250–303 (2003).
- [31] J. S. Marshall and M. A. Thomson, Eur. Phys. J. **C75** no. 9, (2015) 439
- [32] Y. Fukuda *et al.*, Phys. Rev. Lett. **81**, (1998)
- [33] Q. R. Ahmad *et al.*, Phys. Rev. Lett. **89**, (2002)
- [34] Maki, Z., Nakagawa, M., Sakata, S., Progress of Theoretical Physics, **28** (5), 870–880, (1962)
- [35] Mehedi Masud *et al.*, J. Phys. G: Nucl. Part. Phys. **43** (2016)
- [36] Bethe, H. Bremsformel für Elektronen relativistischer Geschwindigkeit. Z. Physik **76**, 293–299 (1932)
- [37] Zyla P *et al* (2020) Review of particle physics. PTEP 2020(8):083C01; and 2021 update
- [38] Bragg, W. H., Kleeman, R., The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **8** (48), 726–738, (1904)
- [39] R. Acciarri *et al.*, JINST **12** P09014, (2017)
- [40] Mistry, K.V.J, MicroBooNE as a LArTPC Detector, (Springer Theses, 2023), p.33-49
- [41] Rocca, P., Riggi, Francesco. The Use of Avalanche Photodiodes in High Energy Electromagnetic Calorimetry. 10.5772/14574. (2011).

- [42] F. Scheck, Physics Reports **44** 4, (1978), p.187-248
- [43] E. Haug, W. Nakel, *The Elementary Process Of Bremsstrahlung*, (2004)
- [44] P. Abratenko, Phys. Rev. D **104**, (2021)
- [45] A. K. Bhattacharyya, Bulletin of the Calcutta Mathematical Society **35** 99-110, (1943)
- [46] Peterson, W. W. T. G., T. Birdsall, W. Fox, Transactions of the IRE professional group on information theory, **4.4**, p171-212, (1954)
- [47] C.J. Van Rijsbergen, Information Retrieval, 2nd edn. (Butterworths, London, 1979)
- [48] Maron, M. E., Journal of the ACM, 8(3), 404–417, (1961)
- [49] K. Albertsson, *et al.*, TMVA 4 Users Guide, arXiv:physics/0703039 [Data Analysis, Statistics and Probability], (2020)
- [?] Y. Freund and R. E. Schapire, Journal of Computer and System Sciences **55**, p119-139, (1997)
- [50] Y. Freund and R.E. Schapire, J. of Computer and System Science 55, 119 (1997)
- [51] Zou, J., Han, Y., So, SS. (2008). Overview of Artificial Neural Networks. In: Livingstone, D.J. (eds) Artificial Neural Networks. Methods in Molecular Biology™, vol 458. Humana Press.
- [52] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, arXiv:1412.6980, (2017)
- [53] O’shea, Keiron, and Ryan Nash. arXiv preprint arXiv:1511.08458 (2015).
- [54] A. Krizhevsky, I. Sutskever, G. E. Hinton, Communications of the ACM. **60** (6): 84–90, (2012)
- [55] A. Chappell, L. H. Whitehead, Eur. Phys. J. C (2022) 82:1099
- [56] B. Abi, *et al.*, Phys. Rev. D **102**, 092003, (2020)