

The Effect of Light-yield Reduction in the ND280 Electromagnetic Calorimeter on the Reconstruction of Charged-Current Events

Lucas Hipwell

Supervisor: Dr Laura Kormos

June 26, 2022

University of Lancaster,
Department of Physics



Abstract

The T2K experiment started operation in 2010, where the ND280 scintillator systems experience an average 1% light yield reduction due to aging. This paper will discuss how the aging of Electromagnetic Calorimeter has affected the reconstruction of charged-current events and trends expected to develop up to the year 2040. Due to the Electromagnetic Calorimeter being primarily comprised of lead it has a high density, meaning a majority of neutrino events happening within its fiducial volume therefore it would be of benefit to understand and counteract the effects of aging. The ND280 Electromagnetic Calorimeter has an increased rate of light yield loss (2% reduction per year) when compared to accompanying detectors. The ND280 neutrino interactions form a ‘baseline’ for the far detector to compare against; if the reconstructions begin to reduce in purity and efficiency the neutrino oscillation results from the T2K experiment may become inaccurate, leading to inaccurate results about potential CP violations, the leading explanation for matters domination pver anti-matter in the universe. When a 50% light yield reduction was simulated, the number of ‘bad’ quality ECal objects decreased by 78%; the remaining reconstructed objects shifted away from appearing as electromagnetic showers towards presenting as a minimally ionising particle track or highly ionising particle track. Good quality photon data sample’s purity decreased by 2% and efficiency by 18% within production 7. Charged-current resonance interactions were minimally affected by aging, the largest variation being a 5% purity loss in the 1 pion branch after the events involving photons were removed. Purity and efficiency are minimally affected suggesting that the T2K experiment could remain function as until 2040, significantly later than the upgrades for Hyper-Kamiokande will be implemented.

Contents

1	Introduction	1
2	Particle interactions of interest	2
2.1	Quasi-elastic charged-current scattering	3
2.2	Charged-current resonance	3
3	T2K and ND280 Detectors	4
3.1	The T2K experiment	5
3.2	Scintillation	6
3.3	Electromagnetic calorimeter	6
3.4	Fine-grained detector	8
3.5	Time projection chamber	8
3.6	Side range muon detector	9
4	Current method of data processing	9
5	Aging	12
5.1	Methods of simulating aging in the ECal	12
5.2	Predicted aging vs light-yield	14
6	Selection variables	15
6.1	Total clusters by particle	15
6.2	Energy fit	17
6.3	Charge distribution	18
6.4	Likelihood ratios	20
7	Event selection	22
7.1	Muon cuts	23
7.2	Pion specific cuts	24
7.3	CCPhoton cuts	25
8	Results	27
8.1	Sample purity	27
8.2	Sample efficiency	30
8.3	Assessing quality	33
9	Conclusion	33
10	Appendices	37

1 Introduction

Neutrinos are a fringe case of the standard model. They exhibit behaviour that is unique to them; most interesting of which is neutrino oscillation where the particle will change flavour as it propagates. Neutrino mass is not defined by flavour but instead three mass states are probabilistically correlated to flavour[2]. Oscillation parameters need to be better understood to further develop the standard model. Neutrinos are also suspected to violated CP conservation laws [3], as these particles oscillate at different rates than their anti-neutrino counter parts, possibly explaining why the universe is dominated by ‘regular’ matter. The standard model could suggest the existence of heavy ‘sterile neutrinos’ that don’t interact with matter in any conventional mean due to being right handed ???. Obviously it is important to know if there ‘ghost’ particles as this could explain dark matter. Better defining the oscillation parameters of neutrinos could lead to new techniques for astronomical mapping, as neutrinos rarely interact meaning information is carried intact over stellar distances.

T2K [4] is one of the leading neutrino oscillation experiments, including a near and far detector. A ‘baseline’ of neutrino interactions is formed at the near detector which is compared to the data from the far detector to measure differences which are expected to arise from neutrino oscillations. Background interactions are determined from this ‘baseline’, for example interactions involving neutral pions. To correctly identify background interactions the reconstruction software should be robust against the effects of aging.

This project aims to understand the trends in particle reconstruction due to the current rate of aging in the ND280 Electromagnetic CALorimeter(ECal). The majority of the interactions that occur in the ND280 are charged-current events, so understanding how the reduction in the light yield (covered in section 5.1) will effect variables such as reconstructed momentum and *likelihood* ratios will aid in correcting future experimental data. Charged-current interactions don’t explicitly require the ECal for pion reconstructions so is ideal for this dissertation as secondary aging effects, due to global tracking algorithms, should be measurable. A main aim of this paper is to show how the branch structure of the highland reconstruction software will handle the reduced light yield, understand any shifts in the *quality* of the clusters leading to sample purity and efficiency variation. Simulated data in this report will exclusively use the Monte-Carlo techniques, as the aging effects can be simulated immediately and to utilise the truth directory (discussed in section 5.1).

In current neutrino experiments there are two limiting factors, one being the very small cross section of the neutrino interactions, the second being the systematic uncertainty in the event selection and detection. By understanding how the aging process will affect the future events reconstructions, the data sample’s purity and efficiency can be maximised, decreasing statistical and systematic uncertainties. A secondary aims of this project is to determine if the current reconstruction software will be adequate until the ND280 gets upgraded as part of the Hyper-k [5] experiment which is expected to start experimentation in 2027 [6], the 50% light yield reductions isn’t expected to occur until 2040 (shown in figure 5).

2 Particle interactions of interest

For this report, only interactions in the global tracker region (from section 3) of the ND280 with a primary vertex containing an incoming muon neutrino will be considered; a primary vertex is the vertex furthest upstream within a single reconstruction, which should have the highest reconstructed energy. Given that the primary aim of the T2K experiment is to measure neutrino oscillation parameters, interactions that don't 'start' with a neutrino (there is no neutrino in the primary vertex) will be discarded; however, this should be a negligible quantity as neutrinos are the principal component of the beam composition. Neutral-current interactions satisfy this condition as well, but are difficult to detect as neutrino's into the vertex are identical to the products except for energy transference meaning that the interaction would be reconstructed as an event without a track/vertex (defined in section 4). Figure 1 shows the cross sections of neutrinos within the T2K experiment, Charged-Current Quasi-Elastic (CCQE) and RESonance (CCRES) interactions are the principle components in the vicinity of the peak flux, therefore make up a majority of the reconstructed events. The simulated particle interaction type break down shows that 72% of interactions are charged-current with quasi-elastic interactions being 32% and resonance being 29% of the total [7]. Each type of interaction is often called a topology, this is the case when using the highland files in section 4 e.g. charged-current quasi-elastic interactions can be denoted as topology 0. Meson Exchange Current (MEC) (shown on figure 1) is a low level uniform source of interactions but currently there isn't a way to reconstruct these interactions efficiently, they have been ignored for this report.

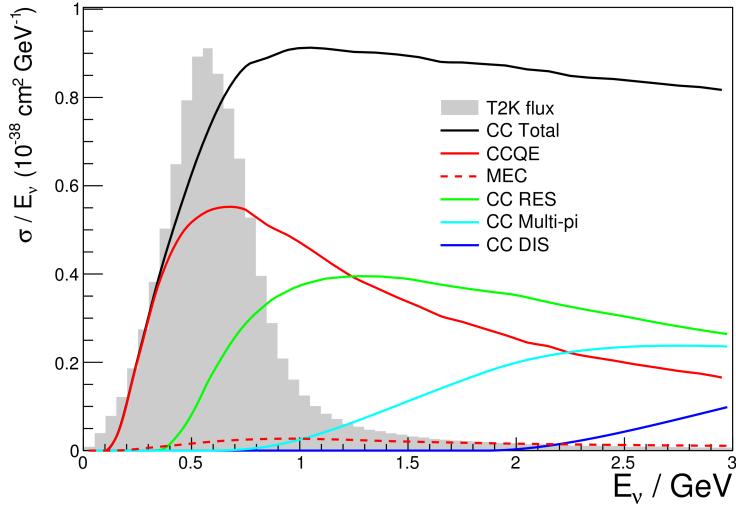


Figure 1: The distribution of event topologies (in the ND280) with respect to the interaction cross section over event energy. This applies only when the beam line is predominantly muon neutrinos (within the range shown), this is called 'Neutrino Mode'. The peak flux is ≈ 600 MeV [8].

2.1 Quasi-elastic charged-current scattering

Quasi-elastic charged-current (CCQE) scattering is a prominent neutrino interaction type for the T2K experiment, as all interactions must start with a neutrino of any flavour; the beam composition is $\approx 92\%$ muon neutrinos when in neutrino mode, which it is for this project. When a neutrino interacts with a nucleon from the detector material a charged lepton of the initial neutrino flavor is produced; nucleons produced may change identity to conserve charge, a process that can be observed in equation 1 and 2.

The rest masses of leptons is larger than current standard model neutrinos so this process will have a minimum neutrino momentum threshold (the nucleons are assumed to be stationary). The cut off energy is many magnitudes smaller than the average energy of beam line neutrinos; however, this does allow of a period of low energy where only neutral-current interactions can occur. These events don't create charged particles so are difficult to detect in many of the detectors. The minimum energy required depends on the flavour of neutrino, with electron rest mass being the smallest at 0.51 MeV but the majority of interactions start with a muon neutrino which explains why the CCQE interactions begin to increase in quantity at ≈ 105 MeV on figure 1 [9].

$$\nu_l + n \rightarrow l^- + p \quad (1)$$

$$\tilde{\nu}_l + p \rightarrow l^+ + n \quad (2)$$

In equation 1 and 2 n is a neutron, p is proton, ν is a neutrino of lepton flavour l . Due to only two daughter particles produced in this process the kinematics of the interaction can be fully reconstructed, making it very useful for profiling the beam neutrinos. As the events in question are quasi-elastic, meaning that energy within the system is conserved, the initial energy is redistributed as mass or momentum in the daughter particles; all mechanisms of 'losing energy' are negligible, such as increased vibrational or rotational energy.

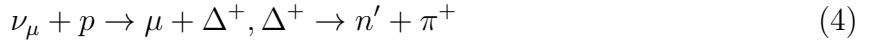
2.2 Charged-current resonance

The production of pions in the ND280 is primarily facilitated through delta baryons (represented as Δ) [2]. These are typically nucleons that have interacted with a high energy neutrino, the neutrino would deposit energy into the baryon to excite it into a higher energy level. Neutral-current interactions can achieve this excitation by scattering particles, only transferring energy, or as a charged-current interaction, producing daughter particles. The delta particle is actually a class of excited baryons (comprised of only up and down quarks) with a range of masses and charges possible, including double charged particles such as Δ^{++} , allowing for multiple permutations depending on initial nucleons [7].

Equation 3 shows the definition of a Charged-current resonance (CCRES) interaction, including the second step where the delta particle decays into pion and a recoil nucleon. CCRES is the only production of pions that will be considered in this paper as all other topologies require energies that significantly deviate from the peak energy shown on figure 1.

$$\nu_l + N \rightarrow l + \Delta, \Delta \rightarrow N' + \pi \quad (3)$$

In equation 3 N is any nucleon, l is any flavour of lepton, Δ is the resonating nucleon (a delta particle) and π is a pion. The charges involved can vary depending on the specific interaction, for example charged pions can be produced or a lepton with the flavour of the incoming neutrino. A recoil nucleon (N') will be in a higher energy level than the original nucleon, which will usually de-excite by emitting a photon; it is possible to produce secondary pions, this does require delta particles of higher masses such as $\Delta(1700)$. These events are referenced as CC Multi-pi within the highland software (section 4) and shown on figure 1. Equation 4 shows a specific example of a CC resonance interaction, one of the branches expected to occur and be measured frequency; due to the beam being $\approx 92\%$ muon neutrino and the pion being charged, allowing for multiple detectors to be able to track its motion (section 3).



In equation 4 the n' refers specifically to a recoil neutron. CCRES is possible from initial energies as low as ≈ 0.17 GeV however as figure 1 shows, there are more incidence at higher energies. For this interaction to happen at 0.17 GeV the neutrino must transfer all its energy which is unusual. The decrease in CCRES events after ≈ 1.25 GeV is due to the increase CC Multi-pi as more pions are created during delta decay on average. For section 7.2 all CC 1-pion and CC other will therefore be considered CC resonance interactions; likewise all CC 0-pion will be considered CCQE interactions.

3 T2K and ND280 Detectors

This report will focus on data from Monte-Carlo simulations, which will simulate the ND280 facility. The near detector facility is 280 m downstream from the neutrino production point, where the neutrino beam is sampled for properties such as energy profile, composition of beam neutrino flavours and beam flux. The location of the facility, relative to the production point, is important as the short distant between them would indicate that a negligible amount of neutrinos will have oscillated flavours, limiting the contaminates in the beam. The ND280 (figure 2) is a collection of five categories of detectors: one electromagnetic calorimeter, three time projection chambers, two fine-grained detectors, one side muon range detector and one Pi-Zero detector [10]. For the purposes of this report the pi-zero detector won't be considered as it aims to detect neutral current interactions so won't be influential in the reconstruction algorithm.

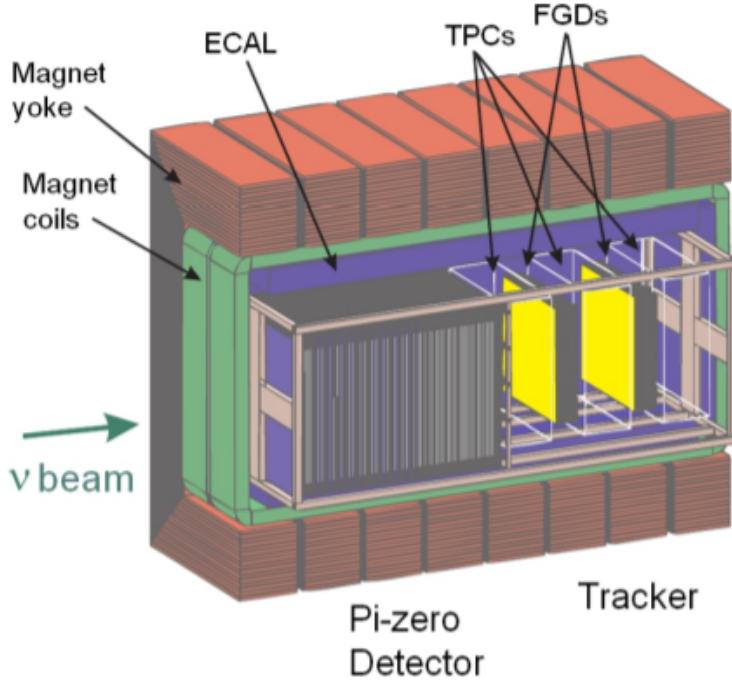


Figure 2: A cross section of the ND280 without the supports or surrounding electronics. This view shows the tracker while it is ‘open’, as the whole system can separate down the center for inspections and maintenance. While the side range muon detectors haven’t been labeled in this diagram they are positioned in between the magnet yokes [11].

3.1 The T2K experiment

The neutrino beam is produced within the J-PARC facility where a synchrotron projects 30 GeV protons onto graphite bars [12]; the synchrotron injects these protons in pulses, where each pulse (or ‘spill’) is comprised of eight smaller pulses called ‘bunches’. Protons that interact with the graphite target produce π mesons, which are funneled into a beam using magnetic fields. Depending on the π meson’s energy the decay can favor either neutrino or anti-neutrino production leading to the beam being in ‘neutrino mode’ or ‘anti-neutrino mode’ [13]. As the proton beam has a structure of eight bunches per spill so does the neutrino beam produced. The branching fraction of these decays aren’t 1 so there are beam contaminates such as anti-neutrinos in the neutrino mode beam or variations in the flavour of neutrino.

280 m downstream from the decay volume is the ND280, a series of detectors that profile the neutrino beam. The characteristic of interest are the beams flux, composition and energy profile. Neutrino interaction cross sections are also measured so that the interaction rates at the far detector Super-Kamiokande can be compared; allowing for oscillation parameters to be determined over the 295 km, since the neutrino production [4].

3.2 Scintillation

Many of the ND280 detectors contain scintillator bars which produce pulses of photons as charged particles travel through these mediums in a process called scintillation. Charged-Current neutrino interactions (section 2) predominately produce charged particles so by measuring the energy and positions of these daughter particles the properties of the original neutrino can be calculated. As a charged particle travels through a scintillator material some of the kinetic energy can be transferred to the surrounding electrons in the material (through the electromagnetic force). For a positively charged particle on approach to an atom the electron will be weakly attracted but while being bound to the nucleus it will not displace far; when the particle is near the displacement will be maximised, which will lessen as the charge retreats. Notably the initial and final displacement would be in opposing directions, giving the electron greater angular momentum so that it can excite to a higher energy level. These electrons will then quickly de-excite to produce a photon pulse which is coupled to a wavelength shifting fiber to capture these photons; both the decay time and light yield profile are material dependant. The key characteristics for a scintillator is transparency in the emission spectrum, high density to encourage more interactions, quick decay time for accurate timing and high quantum efficiency so that the light yield is clearly measurable [14].

3.3 Electromagnetic calorimeter

Surrounding all the ND280 inner detectors (POD, FGD and TPC) is the Electromagnetic Calorimeter detector (ECal) [15]. This detector is comprised of alternating layers of scintillator bars and lead, where the lead is the ‘passive layer’. These lead layers are dense so have an increased chance of initiating an interaction with particles entering the detector, these interactions form ElectroMagnetic (EM) showers which pass into the scintillator layers. These scintillator layers are made up of bars which are perpendicular to the direction of neutrino beam. Each bar is made of plastic scintillator EJ-200 [14] with an elliptical co-axial hole where a wavelength shifting fiber runs. As the EM shower propagates into this scintillator layer they are likely to pair produce leptons, these particles will then scintillate with the bar to produce photons that are then captured in the wavelength shifting fiber. It is also possible that the EM shower would be in the photon stage while passing through the scintillator, these photons can be directly captured without the need for scintillation. These attenuated photons are then detected by the Multi-Photon Pixel Counter (MPPC), each of these MPPC units will have a light yield threshold so that noise will be less suppressed however this does mean that some low energy hits are lost. The light yield threshold is 2.5 PE (photon equivalents, which is the charge produced by a single electron in a MPPC) [10].

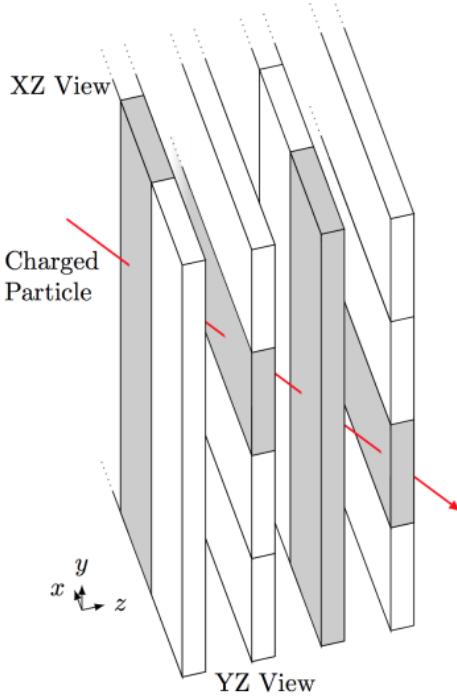


Figure 3: A representation of a particle traveling through the ECal, the shaded scintillator bars are connected to MPPCs that will detect a signal. It is clear to see that the alternating scintillator bar direction in the XY plane will allow for 3D reconstructions using timings to calculate the particle propagation [14].

The amount of light detected at a given location is proportional to the energy of daughter particles as they propagate through that scintillator. Given that the radiation length in lead is 0.54 cm the average EM shower will pass through ten scintillator-lead layers, this provides a large active volume so that the majority of clusters are completely contained with the fiducial volume. Each layer of scintillator bars are orientated perpendicular to the previous layer so that the detector can reconstruct particles trajectories in 3D, see figure 3 for this. The large active region and high number of scintillator layers interacted with per clusters allows for the rate of energy loss per layer to be calculated; this is a key component in determining the particle type and flavour as energy deposited per layer is heavily influence by the species of particle.

In total the ECal is comprised of modules; one downstream detector, six in a ‘barrel’ formation centered on the inner detectors and six surrounding the P0D. Two modules on top and bottom of the inner detectors with a larger with a larger module running along each side comprises the ‘barrel’ formation; this layout is mirrored by the detectors around the P0D. The large modules have increased surface area but a decreased depth of 9.7 radiation lengths due to spacial limitations [16]. For the same reason the modules around the P0D, aiming at measuring photons from EM showers only partially contained, are only 4 radiation lengths thick [15].

3.4 Fine-grained detector

There are two Fine-Grained Detectors (FGD) in the ND280, each positioned between a time projection chamber, as shown in figure 2. These modules are used to detect the recoil of protons and the production of charge particles, characteristics that are useful in the identification of charged-current resonance interactions. The detector needs high resolution as proton recoil happens over short distances, making them undetectable in other detectors [11]. Neutral pions are difficult to identify due to the lack of charge meaning they don't interact with many of the detectors present however FGDs can infer their presence by detecting the electron-positron pair formed from a neutral pion decay. Just like the ECal (section 3.3) this detector can be used for particle identification via measuring the energy deposited per scintillator layer. All of these factors make the FDG very crucial for identifying CC resonance through a Δ particle, as this appears identical to a CCQE interaction in other detectors whereas the FDG would identify both the outgoing lepton and a recoil proton.

Each module is comprised of two layers of polystyrene bars with air gaps between the layers, with a total depth of ≈ 30 cm so that many of particles will penetrate into the Time Projection Chambers. These bars have a cross section of 9.6 mm per side, this provides the high granularity the detector needs. Each bar has a co-axial hole where a wavelength shifting fiber is fed, these fibers are protected by air cushioning with coating of TiO_2 to lessen photon loss. Each of these wavelength shifting fibers leads to a MPPC, however each layer of bars has the fiber leading from either the x or y axis (in the plane of the neutrino beam) in an alternating fashion leading to the name ‘XY modules’. The upstream FDG has fifteen of these modules while the downstream has seven as layers of water (2.5 cm thick) can incorporated into the detector. Allowing for the active medium to be replaced is to measure the relative event rate between water and air so that the event cross section can be determined independently in each medium.

3.5 Time projection chamber

There are three Time Projection Chambers (TPC) in the ND280 tracker system, used for their excellent 3D tracking. A 25 V cathode is suspended in a volume of argon gas to create a nearly uniform electric field, as charged particles travel through this volume ions are created [16]. These ions will be attracted/repelled by the electric field until they fall on the 72 Bulk-Micromegas detectors around the perimeter of the the detector, providing spacial precision of 0.7 mm. These detectors measure the final position of the ions, then with knowledge of the electric field strength and direction the start point can be determined which must be on the original particle’s path. Due to this detector functioning on the ionising capability of the charged particle it does provide more data points for highly ionising particles (HIPs) such as charged pions, which allows for a more precise reconstruction of the particle path compared to leptons and other minimally ionising particles (MIPs).

The charge of the particle can be identified by its path, as the magnetic field provided by the ND280 magnets causes charged particles to be deflected. Charged particles will form circular paths with opposing charges traveling in opposing directions meaning the original neutrinos charge sign can be inferred; a positive charge would mean that the neutrino at the primary vertex was an anti-neutrino and vice versa. The radius of these paths will also mean the particles momentum can be calculated, this is shown in equation 5.

$$r = \frac{p}{qB} \quad (5)$$

Where r is radius of curvature, p is charged particle momentum, q is the charge of the particle and B is the magnetic field strength (0.2 T provided by the ND280 magnet, discussed in figure 2).

3.6 Side range muon detector

Figure 2 shows eight magnetic return yokes comprised of 16 iron layers each, in between these layers is where the Side Range Muon Detectors (SRMD) are housed. The objective of these SRMDs is to measure the angle of projection (with respect to the beam line) and the momenta of particles escaping the tracker volume. The former is to identify back ground radiation such as cosmic rays while the latter is to analyse the distribution of particles ejected from the tracker. This detector can aid particle identification however due to the SRMD being the ‘outer bound’ of the tracker many of the particles will leave the active region, meaning they will get excluded from reconstructions.

Each module is made from layers of polystyrene-based scintillator counters, with wavelength shifting fibers interwoven in a serpentine fashion. By weaving fibers through multiple scintillator bars uniform sensitivity through out the detector should be achieved with minimal fibers; these fibers lead to MPPC units. SMRD modules are positioned on all sides of the tracker, however they are not uniformly constructed; the trackers upstream only have the three inner most layers of the yoke equipped with modules while downstream this increases to the inner most seven layers. The ‘left’ and ‘right’ sides have more modules as this is where the highest-energy particles are usually ejected from.

4 Current method of data processing

This section will outline the process for the Monte-Carlo simulation and ND280 data to being files that can be analysed. Figure 4 shows that the first few steps for simulated and experimental data differ as they need to be compiled into a format that will be effective for both data sets so that the later steps are uniform. It is important that both data sets are comparable as the data collected from the ND280 is used to better model the simulation parameters which in turn is used to explain interactions happening in the detector. Experimental corrections, such as field and distortion corrections, are provided by the tpcCalib and subsequent steps, so will not be considered as this report as exclusively simulated data is considered.

A ‘hit’ is a signal from a MPPC showing a depositing of charge by a particle. Many of these signals can be reconstructed to form ‘clusters’ or ‘ECal objects’, signifying how a particle has interacted with the detector, e.g forming a track as it propagates or decaying into an EM shower. ‘Events’ show the full particle interaction, for this project that required a neutrino to produce a lepton through a charged-current interaction; these leptons may create multiple tracks or showers.

The reconstruction algorithms are constantly being updated, where a batch of updates are known as a ‘production’. The most recent production is production 7 while the most widely used is production 6.

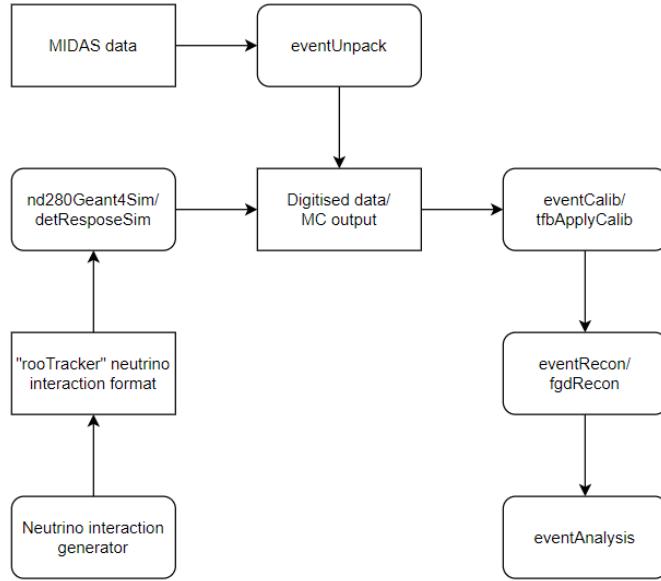


Figure 4: This flowchart shows the step by step process of analysing and formatting data for the ND280 production 7 software. Both Monte-Carlo simulations and data collected from the ND280 follow a similar process to ensure that the conclusions drawn from either data set can be directly compared. Differences would be due to parameters in the simulation and not that data handling methodology is at fault [17].

Neutrino interaction generation & rooTracker - These stages simulate the neutrino interactions; these being incoming particles, the interaction and the daughter particles. By simulating the neutrino beam and the material of the detector, interactions within the detector are simulated but not the detector response.

nd280geant4sim/detResponceSim - This is a two step process, the first part uses the GEANT 4 library to simulate the energy deposited into the detector. The second step simulates all electronics of the detector in reaction to the GEANT simulation, which is broken into three sub steps; how the energy deposited translates into energy in the sensor, the signal generated by the sensor and how the signal is recorded by the supporting electronics. The output of this process is an oaEvents formatted file, oaEvents format is based on Root [18] which is structured to save single ‘raw data’ files as trees, with branches being categories of

data and leaves being variables that describe the data. An important branch is the ‘truth’ branch or directory that contains what ‘actually happened’ in the simulation compared to what was reconstructed, this branch isn’t used for any other later steps but is useful for analysis at the end [19].

MIDAS & eventunpack - Instead of simulating the starting data the analysis can start with real data from the ND280 or any T2K affiliated detectors. In this case the data is stored in a MIDAS formatted file [20].

eventCalib/tfbApplyCalib - This library applies experimental corrections; while the eventCalib sub-step calculates the values required for corrections, ApplyCalib would apply these corrections. Each detector has a specific program for that detectors fiducial volume which is then applied in unison as a ‘global’ operation but for the purposes of this report it is not applied as all data is simulated.

EventRecon/fdgRecon - When supplied with a file of recorded ‘hits’ the reconstruction can begin, becoming a series of tracks and clusters. There are four object types that can be reconstructed; three of these are very similar, a cluster is series of hits spatially close together in a non-linear fashion, whereas a track is a series of hits that are linear and a shower is a series of hits in a conical shape. A vertex is a point where two or more objects meet, these are the locations of suspected particle interactions. Each ECal object will also be generated a particle ID, this is the type of particle creating the object.

eventAnalysis - By unpacking the reconstructed data from each individual detector a ‘global’ image can be formed, as each previous step has used an oaEvents format that is flexible but contains unimportant information such as the hits that are ‘noise’. eventAnalysis reformats the file to a pure-ROOT format, which makes the file quicker to process allowing formal higher level of analysis.

For the EventRecon step there is a parameter, called ‘*quality*’, that determines which reconstruction software is applied. *Quality* is an important attribute that can be used to determine how well the ECal object is defined; a version of this attribute is outlined in section 7.1 (cut 3). For low energy clusters, usually fewer than 7 hits, a separate reconstruction software is used as the distribution of events changes, for example more neutron scattering is observed; these clusters are ‘bad’ *quality*. ‘Good’ *quality* clusters are reconstructed by convectional software.

After the eventAnalysis stage there is an additional step that isn’t shown in figure 4, where the file is converted into a highland formatted file [21]. Highland files are constructed in two stages; firstly creating a flat tree then a micro tree based on the flattree. The files produced by the eventAnalysis use data from a whole beam spill whereas highland will separate out this data into each of its eight bunches. The main benefit is the reduced data size as experiments can use thousands of files which would slow down processing. To make a flat tree each value is saved as a float or integer (or array of each) and each bunch will only have a single entry per variable, to limit file size. Each analysis has its own program to create microtrees, for example there are programs centered on neutral-current interactions or events with neutrons in the final state.

For the initial wave of simulated data in this project nine ‘g4mc’ files were produced, these are files from the nd280Geant4Sim stage. Section 5 outlines how the aging of the ECal was simulated, which has affected the processing at the detResponseSim step as the efficiency of the MPPC is a constituent of the detector response. Each file was processed six times, the first having no aging effects then incrementally reducing light yield until the light yield per hit was halved. It is worth noting that file produced from the detResponseSim (also called ‘elmc’ files) that started from the same ‘g4mc’ file will be ‘duplicates’ in terms of events occurring the detector; these files are not different simulations so are directly comparable. Only nine starting files were selected due to hardware and time restrictions; oaEvent formatted files are large (compared to highland formatted files) which allows for specific analysis but isn’t feasible for a large number of files. These files did provide the data for section 6 and how the event selection variables are affected but to understand how reconstruction of topologies (discussed in section 2) are affected, highland formatted files are required. 2005 simulations were ran to form these highland files, with each ‘g4mc’ being processed twice (once with no aging effects and one with 50% light yield reduction). When using highland all 4010 files were compiled into 2 flat trees (separated into aged and unaged data) which were transitioned into 4 micro trees. Each flat tree was once processed with the ‘regular’ program to detect pions (referred to as the CCPion reconstruction); the second processing adds a photon branch called the CCPhoton reconstruction.

5 Aging

The electromagnetic calorimeter encircles many of the beam line detectors, meaning that events in the inner detectors will often pass into the ECal fiducial volume. The increased rate of hits means the scintillator aging in the ECal is well documented through out the whole volume (two barrel formations and the downstream module show section 3.3)[14]. While the local reconstruction (reconstructions of clusters within single detectors) doesn’t depend on the ECal, the global tracking (implemented in the highland process) calculates attributes of initial neutrino momentum and particle identifications using information from all detectors, these variables are discussed in section 6. Using this concept to see the change on reconstructions that don’t heavily depend on the ECal (such as pions) and the reconstructions that do (such as photons) would help to illustrate primary and secondary aging effects.

5.1 Methods of simulating aging in the ECal

The main criterion of any simulated aging method is that the data produced matches aging already observed in the ND280 ECal. Obviously there are two over-arching strategies: decrease the light yield or decrease the number of hits observed. There is important interplay between these categories, by decreasing the light yield the reconstructed energy of the particles is decreased (section 6.3). To limit the effect of ‘noise’ in the system there is cut off for the minimum light observed to recognise a hit, as discussed in section 3.3. This threshold would stop previously recognised events from being reconstructed as the aged data may lose hits. By reducing the number of hits, the energy spectrum is largely unchanged so hits shouldn’t shift under the threshold therefore this reduced light yield effect is not symmetric.

One example of a method that would decrease the number of recognised hits is to simulate certain bars as faulty. Instead of using a process that universally affects the system, this would be local to certain bars. A faulty bar could either be a bar with a defect or completely broken; the former would allow some light from interaction to be detected by the MPPC; however, a broken scintillator will produce no detected signal. By removing entire bars many events would not be reconstructed due to the sections of the ECal objects being missing; this likely to either make a single object reconstruct as two smaller objects (where the middle would be ‘missing’) or dramatically reduce the energy of the event by either shortening or removing hits from a side of the object. As a single cluster is expected to intersect ≈ 10 scintillator layers the first effect is less likely, as reconstructions could have ‘holes’ but this is verifiable through the average number of ECal objects observed per spill/bunch; decreasing mean charge per hit for the purposes of event selection variables (section 6.3) but the range of total charge per cluster should be unaffected. By not detecting the peripheral hits of the cluster there is an increased chance of it being reconstructed as a track, likely leading to an apparent increase in highly ionising particles (HIPS) as there should still be more hits than expected for a minimally ionising particle (MIPS). Similar effects would be seen in a malfunctioning bar but less severe. Faults in the bar could be that the Wavelength-Shifting Fiber (WSF) has developed partially faulty coupling or the MPPC has debris blocking it. The problem with this method is that this would be an obvious issue, it would be fairly easy to tell when reconstructing the events if sections of an interaction are at different energies with a border in between; this can also be said of hits that have higher energies down stream once they have left the faulty bar. If the bars were too damaged this would also create a ‘dead zone’ where no events are occurring; comparing this to the previous rate of interaction for the area would show a fault. Due to the modular design of the bars it would be easy to remove this section from the reconstructed volume so wouldn’t have a lasting effect on the data. The surrounding electronics could become defective (for example the connected Trip-t Front end board [22]) but this would stop data recording for multiple bars, if an entire sample of scintillator became inactive the issue would be identified therefore this method was discarded as it could not produce results matching the observed aging.

A similar method would be the occurrence of random minor defects suggesting this method would be positional within the detector, so instead of a whole scintillator bar being affected, it would just be one section of a bar. If this defect were due to a break in a WSF then the whole bar would be a ‘dead zone’ as captured photons would feed back into the scintillator bar instead of the MPPC. If a section of surface of the WSF were to oxidise leading to sub-optimal photon capture then this would decrease the light yield of the hits happening within this region but not affect hits further up the bar. Due to the modular nature of the bars it seem unlikely that any fault could spread , so this simulation would lead to a series of ‘small’ defects. Any data provided by this method would be highly situational as both the position of the defect and any hits are pseudo-random. As stated in section 5 the ECal aging is well defined due to high hit rate (compared to the other detectors) and the effects of aging are uniform between all modules so this positional effect was discounted.

After scintillation, the light is delivered along the wavelength shifting fiber to the Multi-Pixel Photon Counter. The wavelength produced by the interaction varies only slightly dependent on the energy of the particle, whereas the amount of photons produced varies significantly; it is for this reason that the amount of photons detected is used to infer the energy of a single hit. For the WSF the peak absorption and emitting wavelengths are not the same, where peak absorption is 430 nm but emits at 470 nm leading to light yield loss. If these values were to deviate this would lead to a further decrease in light yield, not only due to the decrease in absorption into the MPPC but also due to conversion into phonons within the fiber. This process would be uniform for all scintillator bars, which matches observed data; however, this effect wouldn't be uniform for all wavelengths. Higher energy particles scintillate into higher energy photon's, so if the peak emission wavelength increases then the photons measured energy will lessen. In this example the mean charge per hit would shift to lower particle momenta but the bad momentum clusters wouldn't be affected, leading to a bunching effect; this shouldn't lead to a significant increase in bad *quality* objects. Due to the emission profile being asymmetric the wavelengths aren't propagated uniformly, leading to certain reconstructed particle momenta becoming favoured. What is actually being measured is a gradual $\approx 1\%$ decrease in light yield yearly across all detectors [14], which could be due to emission shifting except for it happening uniformly across the spectrum which isn't explained by this method and was therefore discarded as a simulation method.

Data shows that the average light yield has gradually decreased over the years (for the ECal $\approx 2\%$ per year [14]), this light yield loss was uniform throughout the ECal modules. To match the degradation of the real ND280 the simulated detector will have the ‘final’ light yield reduced, this is the light yield recorded from the MPPC. Likely due to the aging of the MPPCs themselves, as the scintillator bars and WSF have been found to be in good condition, through effects similar to gate-oxidation breakdown the performance of the MPPC units could have decreased over time. The most pronounced effect this should have on the system is by lowering the energy of each hit, shifting low energy hits below the light yield threshold, meaning that more objects will have bad *quality*. Low energy testing and reconstruction would be particularly interesting as many more of the poor *quality* clusters will be close to the upper limit of this quality category meaning that the transition between ‘normal’ reconstruction and low energy reconstruction can examined in greater detail. Real world detector behaviour matches the expected data presented by this methodology, while being general enough that the data presented can be easily applicable, giving the most appropriate technique to simulate the aging of the ECal.

5.2 Predicted aging vs light-yield

To determine the amount of light yield reduction that should be simulated two things must be considered, what is the real world timeline for that light percentage loss and is the outcome measurably different. Preliminary investigations involved decreasing the light yield in 10% increments down to 50% of the ‘original’ light yield (the light yield expected of the ECal in 2010). These results can be seen in section 6, some of the event selection variables show variation by 30%; for variables that have a significant role in reconstruction (such as *likelihood* ratios) a 50% light yield loss is when these variations occur. This project would

represent a worst case scenario where the detector has continued to deteriorate until 2040 (after the expected start date of the Hyper-K experiment [5]), without slowing of the aging process, so this report should be considered a test of the reconstruction software trends expected to appear in the future. Using the data provided from this report, the eventCalib stage (presented in 4) could be improved to better correct for light yield loss.

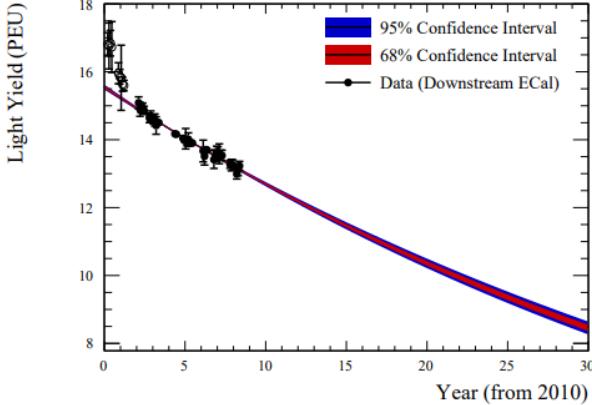


Figure 5: This represents the light yield decrease per year of the downstream electromagnetic calorimeter, which is $\approx 2\%$ [14]. The light yield will be approximately halve the 2010 value by 2040.

6 Selection variables

Selection variables are detected directly from the detector and are used in the reconstruction process [23]. This section will discuss the variables of interest and uses data from the nine oaEvents formatted files (defined in section 5.1), as for this stage it was beneficial to have the ‘noise’ data that would be rejected from the highland files, additionally the reconstructions weren’t necessary for this stage. The errors in this section are estimates, as the variables are highly correlated; the ‘true’ uncertainties are expected to be smaller than the ones represented.

6.1 Total clusters by particle

The definition of ‘*quality*’ used here is separate from the ‘*quality*’ used in the highland software but they are correlated. These *quality* values are generated in the eventRecon stage of section 4, being highly dependant on the number of hits the event had. Figure 6 shows the number of ECal objects separated by truth particle type, not the Particle IDentification (*PID*) generated by the eventRecon but the ‘actual’ particle type simulated.

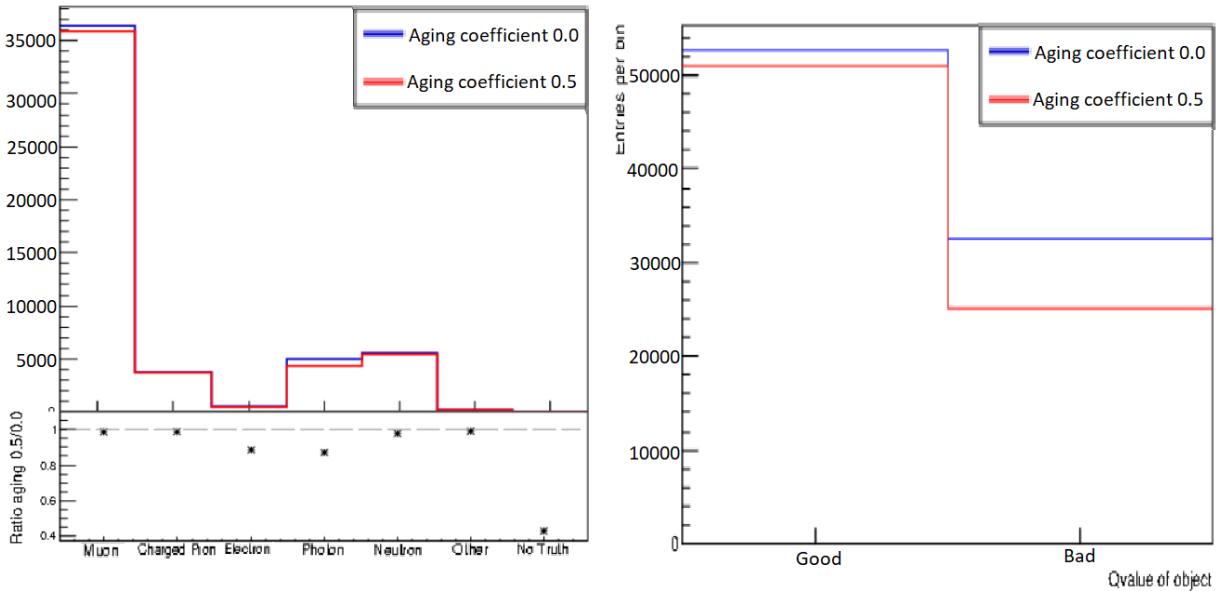


Figure 6: (Left) The total number of ECal objects separated by true particle identical before (blue) and after (red) aging. All particles here are good *quality*, so are used for ‘conventional’ reconstructions. The ratio of aged to unaged objects are shown on the panel below.

Figure 7: (Right) Due to aging reducing the light yield, many ECal objects have dropped below the minimum hits required for the good *quality* tag. Bad *quality* objects that drop under 7 hits in the ECal can’t be reconstructed as an ECal object.

Three ‘particles’ were measurably affected: electrons, photons and ‘no truth’. The no truth category is for objects that were reconstructed but were signals not generated by particles, this can be attributed to reconstructing noise. Usually signals that are generated by noise are low energy, so the light yield loss meant many hits no longer passed the threshold hence why the number was reduced by 0.58%. Electrons and photons are primarily reconstructed in the ECal so were expected to be affected heavily, this is the reason they were selected as particles of interest (the reason that the CCPHoton reconstruction was conducted). Figure 7 shows a shift into the bad *quality* category as hits are lost from most objects but many of the previously bad events are undetected. For a good *quality* event to not be reconstructed at least four hits must be removed whereas a single hit from a low energy event can get it eliminated. These electron and photon events in figure 6 have progressed into bad *quality* so were no longer included, this is unsurprising as electrons and photons produce short showers.

6.2 Energy fit

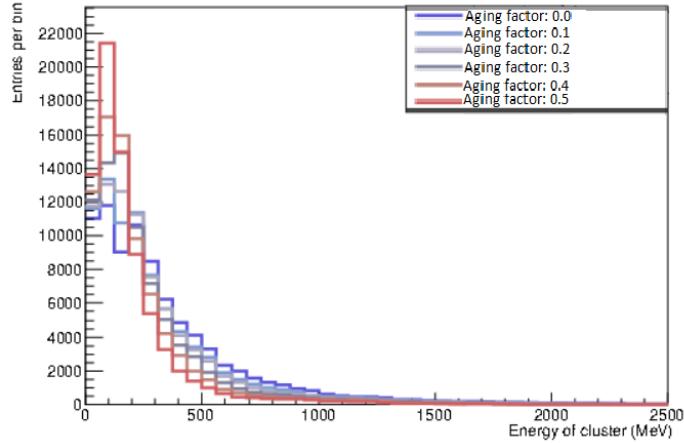


Figure 8: When reducing the oaEvents formatted files light yield in 10% increments the ‘EMEnergy_results’ variable presented the above distribution; often referred to as the ‘energy profile’.

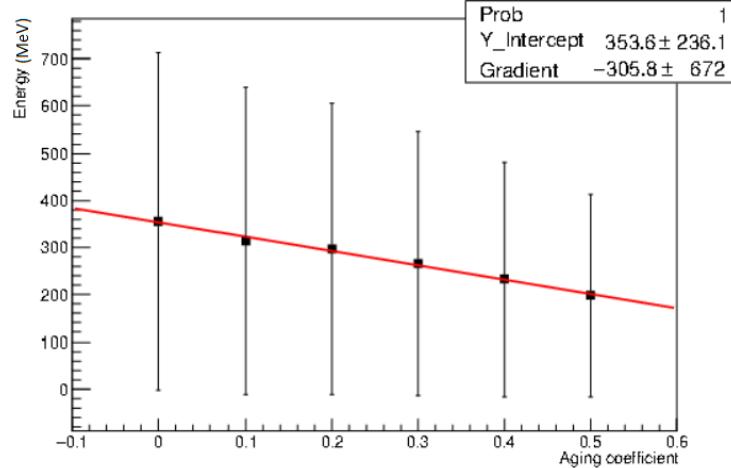


Figure 9: This is the mean energy of an ECal object at each aging increment from figure 8. A linear fit was applied as there is a direct correlation between light yield loss and energy of reconstructed object. The errors represented are an over-estimation due to the high correlation between ECal variables.

When reducing the light yield recorded by the MPPC the energy of the hits are expected to decrease (discussed in section 3.3) which will lower the total energy of the ECal object. Figure 9 shows a incremental lowering of mean energy per object, combined with minimal event loss (as shown by section 6.1) leads to a ‘bunching’ affect at low values. Figure 8 uses both ‘good’ and ‘bad’ *quality* ECal objects leading to loss events being rejected than what

was presented in section 6.1. Figure 8 has not been normalised to account for the event loss. By the time the Hyper-K experiment is expected to be upgraded the ND280, a $\approx 15\%$ light yield reduction should be observed; leading to the mean energy per cluster being 306 MeV as opposed to the current 357 MeV. Post aging data mean energy interaction is 510 MeV while the mean ‘real’ neutrino energy is 600 MeV. All particle reconstructions were affected equally, the particle breakdown of the energy profile is constant (see appendix A).

6.3 Charge distribution

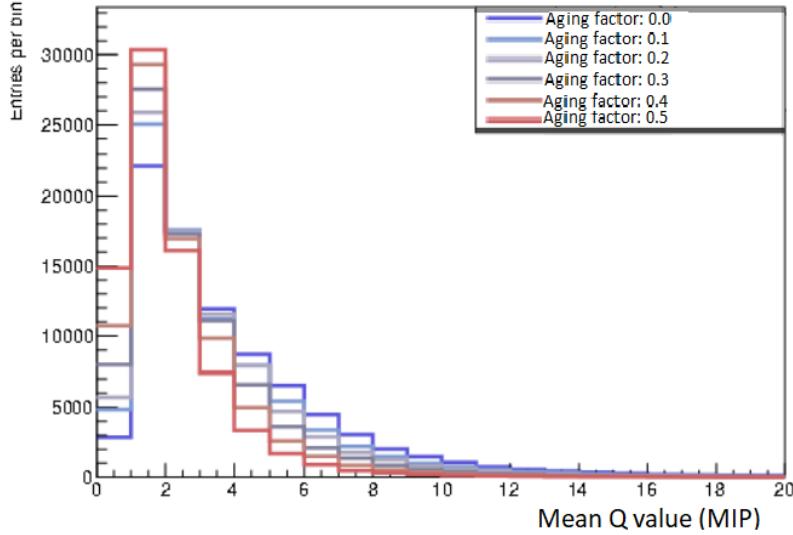


Figure 10: This is the 10% incremental variation of light yield on the variable ‘QMean’ defined as mean charge of hits in the ECal object. The unit MIPs is defined as ‘the MIP energy is defined as the energy deposited by a muon passing at normal incidence through 1 cm of scintillator.’

The ‘Q mean’ (figure 10) is the quantity of charge collected by the MPPC (as the charge is proportional to the light produced) for a single ECal object, in units of the quantity of charge deposited by a MIP. The mean energy diminishes by 0.3 MIPs per 10% light yield reduction (see appendix A) which is expected and reinforced by trends in figure 11. The total charge per object is reduced by 4.6 fundamental charges per increment.

The units MIPs were implemented as a unit of energy measurement that was comparable to particle behaviour. 1 MEU (MIP Energy Unit) is the mean peak energy of muon particles (primary MIP contribution), which is ≈ 30 photons measured by the MPPC; allowing for easy calibration of low and high gain energy deposits.

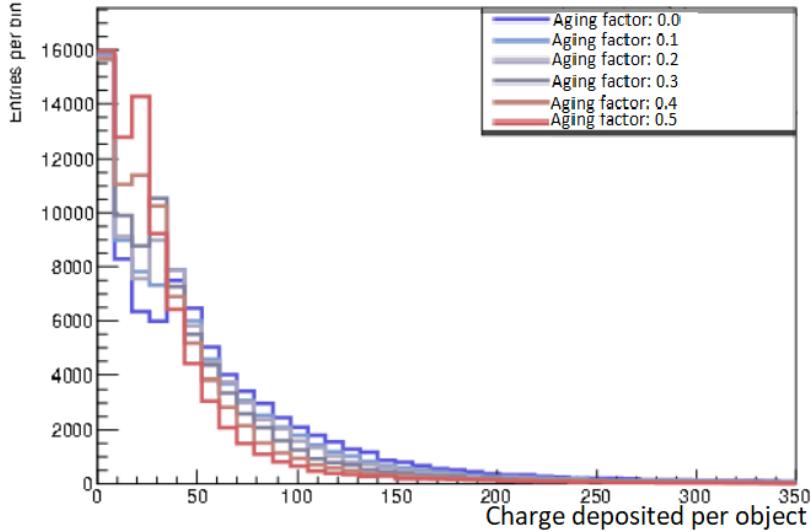


Figure 11: Heavily correlated to figure 10 is the total charge deposited by an object; acting as an analogue to the total light yield produced by a object.

Each aging coefficient peaks at 1 MIP on figure 10, representing a track from a MIP-like particle, usually a muon. The values less than 1 MIP may be due to particles absorbed into the detector material or some of the ‘no truth’ events seen in section 6.1. Pions are highly ionising particles, losing energy as ions are formed until at low energies they decay into an EM shower, these are called lollipop events. A two peak distribution can be expected when considering energy deposited per layer per unit distance; the ‘secondary’ peak of a low energy pion may fall into this region. Normalising figure 10 would be interesting so that the affects on the 3rd bin can directly compared, these would represents events such as high energy MIPs, low energy HIPs or pion decays.

Figure 12 shows the number of events (neutrino interactions primarily) per spill, where there are eight bunchs per beam spill. As to be expected, the mean number of events decreases with lower light yield; objects are required to have at least three hits to be reconstructed, as the light yield is decreased some of these hits get cut off due to being below threshold (0.08 MEU). The gradient of mean objects per spill and aging coefficient is -0.3 (see appendix A).

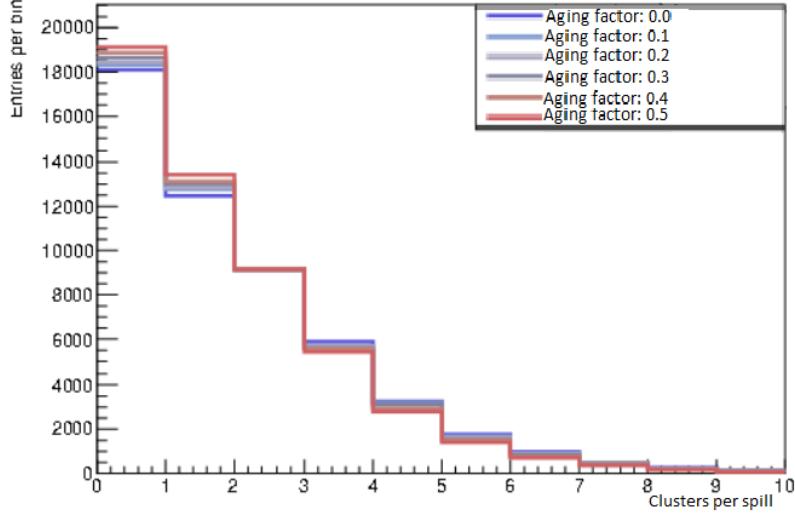


Figure 12: This shows the number ECal objects per spill, these objects may not be used in an event (get rejected by cuts in the highland process) but shows that the detector efficiency wasn't badly affected.

6.4 Likelihood ratios

Likelihood ratios are the precursor to *PID*, they indicate the type of particles that ECal objects will be reconstructed as. For example the LLR_MIP_EM variable leads to reconstructed as a MIP tracks for values less than 0 and EM showers for values greater than 0. The main factor is the object type (track, shower, cluster) and how well the object matches the key characterises of the variable. If an ECal object is a series of hits in a semi-linear fashion but with large spacial gaps it would be reconstructed as a track however would have a low MIP_EM ratio as the track is not well defined.

Figure 13 shows a trend of reconstructions favouring MIP-like tracks opposed to EM-like showers as aging progresses. As a shower structure tends to have lower energy particles on the peripherals, these are the hits that would be lost. These objects have therefore shifted to appearing as linear tracks, this means the reconstruction would be biased to MIP or HIP; however, figure 13 is a comparison of only MIP against EM. As figure 14 shows one of the linear objects most favoured is MIP. Muon reconstructions were affected more than other particle species, due to the minimal number of hits attributed to a MIP. The reconstruction algorithm may have had difficulty with an ‘xy’ plane deviation (with respect to the beam line); as hits are reduced so is the uncertainty of a linear fit. These ‘new’ MIP-like objects would appear as high energy muons that become absorbed as the track would be dense and short.

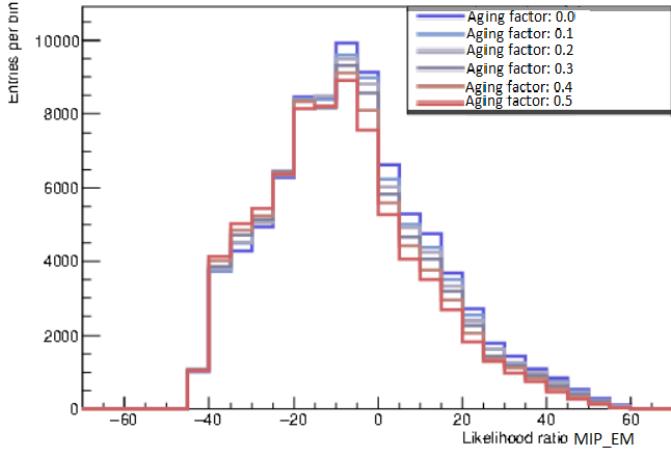


Figure 13: (Left) *likelihood* ratio of ECal objects being MIP-like (abscissa < 0) against EM-like (abscissa > 0). The gradient of MIP_EM *likelihood* ratio value against the aging coefficient is -5.015 ± 45 suggesting that previously EM-like objects are becoming MIP-like.

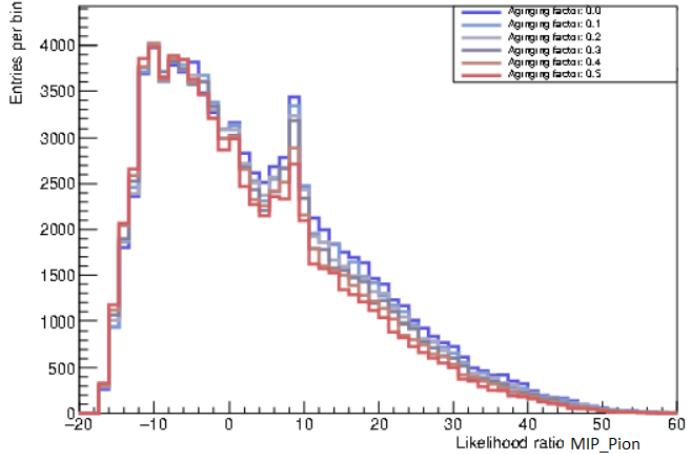


Figure 14: (right) *likelihood* ratio of ECal objects being MIP-like (left) against Pion-like (right), where Pions are HIPs. The gradient of MIP_Pion *likelihood* ratio value against aging coefficient is -3.95 ± 32 suggesting that previously Pion-like objects are becoming MIP-like.

Pions are highly ionising particles, leaving a linear track of a high number of hits, in comparison to MIPs. As expected figure 14 shows an increase in MIP-like objects due to the reduction in light yield. HIP tracks tend to be ‘shorter’ as the particles lose energy to ionising particles with the detector, leading to a comparatively large number of data points. MIP tracks are long with few hits; for the purposes of reconstruction the data point density is weighted more than the track length. The hit loss would decrease the hit density but would also shorten the track lengths, leading to an increase in MIPs. The trend affected all particles equally (see appendix B).

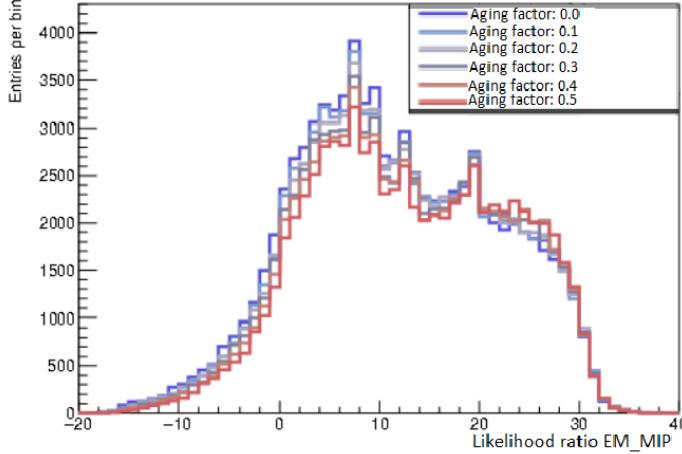


Figure 15: *likelihood* ratio of ECal objects being EM-like (abscissa < 0) against MIP-like (abscissa > 0). The gradient of EM_MIP *likelihood* ratio against coefficient is 2.687 ± 23 suggesting that previously EM-like objects are becoming MIP-like.

As previously discussed many shower-like objects have transitioned into track-like objects, which is further supported by figure 15. Protons were particularly affected by the aging, due to a large proportion of the HIPs being protons this variation is more pronounced. Protons are HIPs but if they are particularly energetic the spacial spread of hits could lead to a difficulty in assigning a linear fit, reducing the hits may have helped the algorithm recognise the track.

7 Event selection

To analyse the charged-current interactions occurring in the detector the highland reconstruction software is applied to ‘base’ oaEvents data, reviewed in section 4. This process was applied to 2005 files twice, once using a reconstruction only looking for how many pions are present in the final state (called CCPion reconstruction); the second time all interactions that ended with a photons were put on an independent branch then the events left over are sorted by the quantity of pions (called CCPhoton reconstruction). It is worth mentioning that while final states with 1 pion or 2 pions are both due to charged-current resonance interaction they have separate topologies within highland, being CC 1-pion and CC multi-pi respectively.

The highland input files are a series of reconstructed objects with attributes such as particle identification (*PID*) and light yield deposited per layer, the highland software will then reconstruct the suspected topology of the objects. Each reconstruction will have a set number of branches, these are a series of cuts used to infer the final state or topology. The process is conducted in a ‘trail and error’ fashion where each event is ran through the cuts of a branch, if the event fails a cut then its compared to the next branch’s cuts. A cut is a specific attribute that is tested to be within an expected range for a given particle, for example the amount of energy deposited per layer. The aim of the cuts is to maximise purity

and efficiency of the final sample, leaving only the signal events in each branch. Background events for a given branch are the events with truth topologies that don't match the current branch; signal events truth topologies do match the branch, so should pass the cuts.

There are two main steps to selecting an event for analysis, one is identifying that an outgoing particle from the primary vertex is a muon and the second is categorising the final state of the event by particles present. All charged current interactions that this dissertation is interested require a muon neutrino propagating into the primary vertex meaning that the muon specific cuts (section 7.1) will be applied to all events[24]. After the first 6 muon related cuts are implemented then the reconstruction specific cuts are applied (CCPhoton or CCPion), so cuts from section 7.3 and section 7.2 will not be applied to a singular data sample consecutively.

7.1 Muon cuts

These cuts are to determine if there was an outgoing muon in the primary vertex. For the charged-current interactions this dissertation is focusing on all events start with muon neutrinos so these cuts are universally applied in event selection. The cuts below are in order of application to the reconstructed events [24].

1. **Event quality** - The neutrino beam line sends pulses of neutrinos, these are called spills which can be separated into 8 distinct bunches. Each bunch is 15 ns ‘wide’, the primary lepton must interact within 4σ of the beam trigger to classify as good *quality*. This means that the standard deviation and mean are calculated on a run-by-run basis, this process uses a Gaussian fit often leading the slow neutrinos in the ‘tail’ to be cut. The lepton with the most energy in FGD is considered the primary lepton [25]. Don’t confuse this with the object *quality*, this is specific to the outgoing lepton not subsequent clusters.
2. **> 0 track** - Only events with at least a single track are accepted.
3. **Quality and fiducial volume** - Firstly this determines if the events primary vertex is inside the upstream FGD fiducial volume, as this is used to calculate the start point of the track. For low angle tracks, with respect to the beamline, the track start point is defined as the having the XY coordinates (perpendicular to beamline) fitted to the furthest upstream hit within the FGD; For large angle tracks the coordinate is rotated to fit the track length. This stage also includes a cut to remove all events that start in the five scintillator bars closest to the edge and in the first XY module as these may have started out of the fiducial volume. The *quality* of the track is also assessed, if the track intersects a Tracker-ECal module then 18 TPC hits are required, otherwise 36 hits are required. This is to remove less reliable/short tracks often indicating very low momentum particles.
4. **Upstream veto** - Scattering events can break one track into multiple, leading to a muon starting further upstream than the reconstruction shows it did. To counteract this, if the second highest momentum track’s start point is less than 150 mm upstream

of the current muon candidate track then that muon is likely the ‘real’ starting track so the event is rejected.

5. **External fdg1** - If the tracks start position is within the last two layers of the FGD and a track is present only in the FDG, the event is rejected. These are usually mis-reconstructed events as if an event spans both the FDG1 and TPC the reconstruction algorithms tends to create two events, one that spans FDG1 only and the second which is detectable in both. In which case the second is considered to be the muon candidate so the first is cut.
6. **Muon pid** - This step uses the TPC particle identification to only contain muon samples at the primary vertex. As there are multiple charged particles that form tracks this compares the energy deposited per distance to expected values for each particle species, using equations 6, 7 and 8. This stage assumes that only protons, electrons, muons and pions are present in the sample.

$$\mathcal{L}_{MIP} = \frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_{Proton}} \quad (6)$$

\mathcal{L} is the *likelihood* value of a given particle, which is defined in equation 7.

Electrons are removed from the sample using equation 6 for $\mathcal{L}_{MIP} > 0.8$ when $p < 500 \frac{MeV}{c}$. Muons are the only minimally ionising particles from the four species presented, while proton and pions are highly ionising particles so can be rejected using the condition $\mathcal{L}_\mu > 0.05$.

$$\mathcal{L}_i = \frac{e^{-Pull_i^2}}{\sum_l e^{-Pull_l^2}} \quad (7)$$

Equation 7 defines the *likelihood* ratio for any particle i. Pull is defined in equation 8 where $\frac{dE}{dx}$ is the energy deposited per unit distance in the TPC. $\frac{dE}{dx \text{ measured}}$ is the energy deposited per unit length measured compared to $\frac{dE}{dx \text{ expected}}$ which represents the average energy deposited per unit length for that species of particle. In this example standard deviation is calculated on a run by run basis.

$$Pull_i = \frac{\frac{dE}{dx \text{ measured}} - \frac{dE}{dx \text{ expected},i}}{\sigma(\frac{dE}{dx \text{ measured}}) - \frac{dE}{dx \text{ expected},i}} \quad (8)$$

7.2 Pion specific cuts

After the event has been verified to contain an outgoing muon in the primary vertex, these cuts aim to identify the clusters/showers/tracks within the event. For the CCPion reconstruction the branch structure isolates 0 pion interactions (CCQE) then 1 pion interactions (CCResonance) and finally other events (CCMULTipi). The ‘other’ events mainly consist of high energy resonance interactions producing multiple pions however the composition is shown in section 8; this branch is referred to as CCOther within the highland software. Below is the structure for the 1 pion branch as the specific cut change per candidate topology.

7. **CCPion1** - This cut is comprised of several steps, first the objects must have the same time stamp as the bunch containing the muon candidate; the objects must also start in the same FDG fiducial volume. Similar to the 6th muon cut the objects must undergo particle identification as pion, if the particle is negative then the hypotheses tests electrons against pions. For negative pions the exact condition is : $\mathcal{L}_{MIP} > 0.8$ when $p < 500 \frac{MeV}{c}$. Neutral pions are more difficult to identify as they are charge-less, only if the pion decays into the pair production of photons or electron-positrons pairs. Low energy/short objects are assigned the Michel electron tag, otherwise the FDG *PID* technique (equation 6, 7 and 8) is used determine the species of particle. If the Michel electron is out of the bunch time window but within a muon lifetime it must have seven hits in FGD1 or six in FGD2. Michel electrons are generated in higher quantities by positive pions opposed to negative pions which are more likely to be absorbed.

8. **ECal pi0 veto** - This cut is a remnant of production 6, this cut doesn't alter the sample data.

For the 0 pion branch, the cut 7 is altered to verify that all ‘out-going’ objects don’t represent pions, called the CCPion 0 cut. The CCOther branch could be considered all interactions left over, the only cut in this branch is that it must have failed all previous branch hence why it is the final branch.

7.3 CCPhoton cuts

The CCPhoton reconstructions primary aim is to identify final states with at least 1 photon. This new branch is the ‘lead’ branch making the branch order: CCPhoton, CC0Pion, CC1Pion then CCOther. The CCPhoton cuts use a new technique called principle component analysis where the dimensionality of a data set is reduced, in practical terms this limits the number of variables that affect the reconstruction of photon showers. The key concept involves weighting the first principle component over the second e.i a change in one dimension is more ‘important’ then a change in the second [26]. Photons are primarily detected in the ECal so are expected to be effected heavily by the aging effects, the most common photon branches are that of pair production shown below:

$$\begin{aligned}\pi^0 &\rightarrow \gamma + \gamma \\ \eta &\rightarrow \gamma + \gamma \\ \eta &\rightarrow \pi^0 + X \rightarrow \gamma + \gamma + X\end{aligned}$$

where X is a variety of particles depending on the specific decay and η is a flavour of meson.

7. **CC photon** - This stage requires two criterion, the first is to be identified as a photon using the EM *likelihood* ratio. The second is confirm that there is no ‘pile up event’, due to the comparatively high neutrino interaction rate of the edges of the detector classifying multiple interactions as one.

Circularity was introduced in production 5 to compare the width of an object to the length. As photons are expected to form short-wide showers instead of long-thin tracks *circularity* can identify the type of reconstructed object that should pass the cut. Using principle component analysis, tracks-like objects have values of 1, see equations 9 and 10 [27].

$$\text{Circularity}_i = (2 \times (\text{2nd principle component})) - 1 \quad (9)$$

Equation 9 defines *circularity* in axis i, assuming a two dimensional space. Equation 10 represents the combined variable used in photon *PID*.

$$\text{Circularity} = \text{Circularity}_x \times \text{Circularity}_y \quad (10)$$

Truncationratio compares the maximum charge deposited into the scintillator layer/bar to the minimum charge. Before calculating the ratio, the 20% highest and lowest charge hits per layer are removed to reduce the effect of noise. Using information, like the Braggs peak, allows this stage to be optimised to differentiate between muons and electrons.

The *hit charge distribution* calculates the total charge deposited by a shower, allowing for particles such as MIPS to be excluded. Equation 11 shows the dimensionless quantity that doesn't depend on the charge scale.

$$q_{rms} = \frac{1}{q} \sqrt{\sum_i^n \frac{(q - \tilde{q})^2}{N}} \quad (11)$$

N is the number of hits in the object, q_i is the hit charge and \tilde{q} is the mean hit charge. Showering particles are expected to have larger q_{rms} than track-like particles.

The final component of the photon *PID* is the *Front-back ratio* also known as the rate of charge loss over a distance ($\frac{dE}{dx}$). When a charged particle slows down the energy loss to the ionisation process increases so MIP has a consistent rate of loss (a value around 1) whereas EM showers are front heavy (a value less than 1). *Front-back ratio* only considers the front and back quarter along a principle axis (x, y or z with respect to the beam line).

Lastly the cut aims to reject ‘pile up interactions’ in ECal isolated events. As the photons in the tracking region of the detector have a higher probability to shower in the passive first layer of ECal, there is an increased chance of secondary neutrino interactions occurring. To reject this, interactions must have at least a single hit in the 6th most inner scintillator layer.

The CCPHoton cut is performed after the muon cuts (section 7.1), separating the photon branch before considering the pion related branches. Once the data with photons have been removed from the sample the 0 pion branch is processed, similar to the outline in section 7.2 except an inverse photon cut is inserted at 7th, where the event has to be rejected from the photon sample then confirmed to have no pions. This framework is applied to all subsequent branches. These reconstructions don't have a ECal pi0 veto, but this doesn't change the data

handling as it was a null cut. For example the 0 pion branch in the CCPhoton reconstruction is the 6 muon related cuts followed by a cut to confirm no photons are present (confirming the event failed the photon cut); next the pion cut (labeled 7th cut in section 7.1) is applied.

8 Results

The most useful variable to measure the effects of aging on the highland reconstructions is purity and efficiency. Purity is the number of signal events that should be in a given branch (as defined by the truth directories topology) compared to how many total events were assigned to that branch. Efficiency is true number of signal events of that topology compared to the true number of signal events in the given branch. Efficiency is a measurement of how well signal events are accepted while purity represents the algorithms ability to reject background events. Sections 8.1 and 8.2 both don't filter by *quality* of the events to highlight what is happening at high and low energies. The uncertainties in this section are overestimated due to the high correlation in reconstruction variables.

8.1 Sample purity

The 0 pion branch has optimised the purity at the peak flux (shown in figure 16), as MIP tracks are linear with few hits but as momentum increases so does the mean hits per cluster appearing more HIP-like. This also leads to the decrease in purity for the 1 pion branch at higher energies. CCOther's purity is minimised at the peak flux (see appendix C) due to the increase in events leading to ‘unidentified’ events going to the CCOther branch. Shower-like objects are difficult to isolate against short HIP-like tracks at low energies as attributes like *truncationratio* and *hit charge distribution* are more accurate given more entries per cluster leading to the decreased purity at the peak flux, this is further supported by the section 6.4.

Figure 16 shows the purity of the 0 pion data sample, as expected the reconstructions of the CCPion were not measurably effected. The average purity of this branch is 0.731 ± 0.003 (shown in table 1), where both the purity and topological structure remained unchanged through aging (see append D). This was expected as the reconstruction of pions relies on the fine grain detector (section 3.4) and a condition of reconstruction in the highland process was the primary vertex being positioned in the FDG1 fiducial volume. Figure 17 reflects this, showing no change in topological breakdown of the branch.

Global tracking software reconstructs some attributes using multiple detectors, such as momentum. While the change in momentum can't be observed in figure 16 due to using true momentum, particles with *likelihood* values close to 0 could be affected leading to small variations in bin sizes. There is a secondary effect changing the number of events reconstructed, when converting OAevents into flattrees the aged files tended to have less entries afterwards. Likely during the process of converting, the separating of the beam slip into bunches would cause some events to occur in-between bunches so were removed, expecting to be noise which is biased towards the aged file as ‘slow’ neutrino interactions would seem more noise-like.

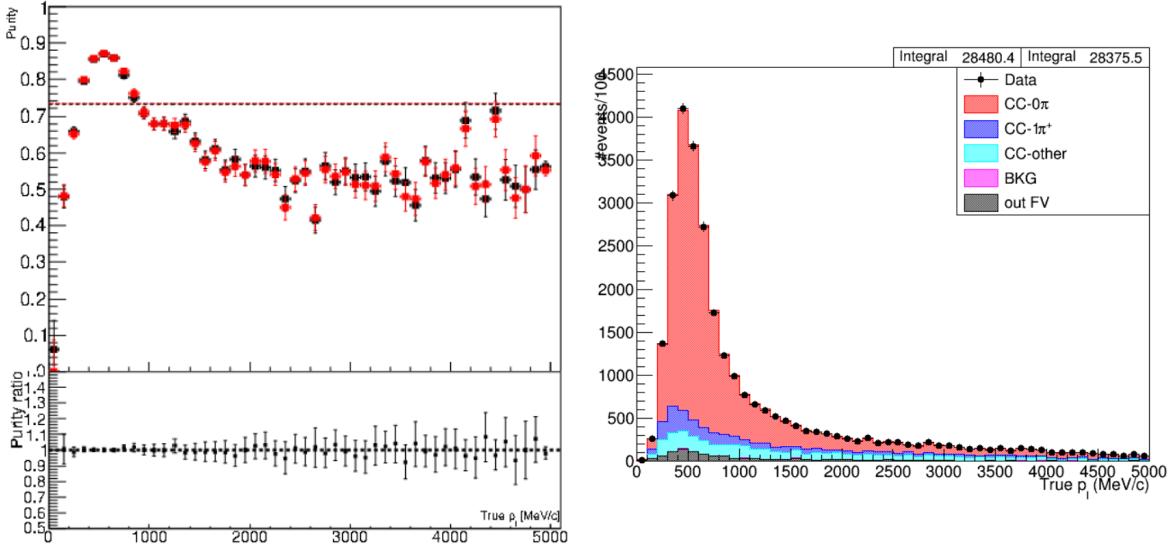


Figure 16: (left) The data in black is the purity of the 0 pion branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged purity over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

Figure 17: (right) A topological graph of the 0 pion branch before aging (from the CCPion reconstruction), the topologies represented are from the truth directory. The black data points are the entries per bin post aging. The left integral is the unaged data while the right is aged; these integrals are weighted hence are not integers. ‘out FV’ is out of fiducial volume and ‘BKG’ stand for background meaning any topology that doesn’t fit in the branch structure (for example an electron neutrino at the primary vertex).

Figure 16 shows a mean purity higher than the average data point as the purity is weighted; the peak flux is at ≈ 600 MeV so these bins are weighted heavier. Due to this branch representing charged-current quasi-elastic (section 2.1), primarily producing MIPs and section 6.4 showing an increase in MIP-like object behaviour the constant purity is a reassurance that the cuts are working as intended.

As photons are primary constructed in the ECal and are the first branch (in the CCPhoton reconstruction) there expected to be a ‘knock on’ effects on the branches ‘in common’ with the CCPion reconstruction. As shown from section 6.4 the both EM *likelihood* ratios (electron and photon) show a decrease in the number of expected photon reconstructions, these events were moved to the 1 pion and CCOther branches decreasing these branches purity (appendix C). The migration towards the 1-pion branch can be explained by EM showers appearing HIP-like, leading one event to possibly being represented as a HIP track then shower. This is referred to as a lollipop interaction and behaviour often exhibited by pions.

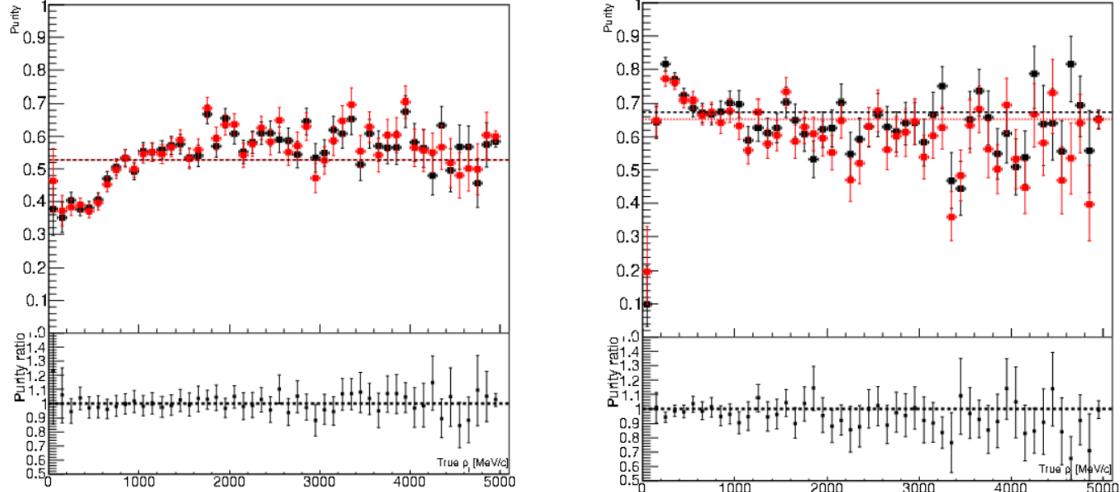


Figure 18: (left) The data in black is the purity of the photon branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged purity over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

Figure 19: (right) The data in black is the purity of the 1 pion branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged purity over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

Table 1 and figure 18 shows that while ≈ 1000 photon events changed topology the purity wasn't affected; these events were primarily 0 pion events so combined with the few 'true' photon events lost make no net change (see appendix B). These events would have specifically been rejected by the *Front-back ratio* (as a majority of the lost entries will be the low energy hits at the 'back') and *circularity* cuts. As the CCQE events in the photon branch move to 1 pion branch, the purity is reduced; these photon topology events would usually have bad *quality* so would be rejected by the 1 pion cut if the *quality* cut was altered to only accept good quality. The purity would be expected to decrease at low reconstructed momentum's as photon reconstructions have lost total charge however figure 19 shows true momentum so would be unaffected. There could be more transitions occurring between each branch but they would be difficult to isolate as only the total number can be reconstructed in production 7.

For a 50% light yield reduction the purity of the branch samples are minimally affected. The largest change was in the 1 pion branch in the CCPHoton where research can recover some purity by comparing to the CCPion reconstruction. This suggests that purity-wise the experimentation can continue until Hyper-K.

Branch (Reconstruction)	Purity before aging	Purity after aging	Ratio
0 Pions (CCPion)	0.731 ± 0.003	0.732 ± 0.003	1.001 ± 0.01
1 Pions (CCPion)	0.563 ± 0.005	0.561 ± 0.005	0.998 ± 0.03
CC Other (CCPion)	0.795 ± 0.005	0.790 ± 0.005	0.992 ± 0.04
Photon (CCPhoton)	0.528 ± 0.005	0.528 ± 0.005	1.000 ± 0.09
0 Pions (CCPhoton)	0.805 ± 0.002	0.795 ± 0.003	0.986 ± 0.09
1 Pions (CCPhoton)	0.674 ± 0.007	0.653 ± 0.006	0.969 ± 0.03
CC Other (CCPhoton)	0.570 ± 0.002	0.566 ± 0.002	0.992 ± 0.07

Table 1: Mean purity of each branch before and after simulated aging. The ratio column is defined as purity after aging over purity before aging.

8.2 Sample efficiency

Efficiency of a branch can be used to understand how many events that should be in a branch (as determined by the truth directory) have been rejected. As to be expected the CCPion reconstruction was unaffected by the aging process where the variation in efficiency can be attributed to changing in the flattree bunching parameters.

The structure of each branch efficiency starts at the origin, increasing to the peak flux (≈ 600 MeV) then a plateau for all higher momentum's. Deriving the ECal object attributes from very few points is currently ineffective, due to the relatively high purity's in the same range (momentum's below 600 MeV) suggests that the muon cuts are rejecting these events. The muon tracks *quality* are typically bad *quality* in the lower energy ranges leading to rejection (appendix C). The plateau coincides with the decrease in purity for each branch, showing that the signal events are now passing all the muon related cuts but are some assigned to incorrect branches, as discussed in section 6.4.

The efficiency of the photon branch (figure 20) is structured the same as the branches of the CCPion reconstruction except the aging gave a measurable effect. If each event was allocated to the same branch after the aging process the efficiency would be expected to increase; section 6.1 shows true photons/electrons events decrease in quantity. The efficiency of the photon branch decreased by 8.5% (shown in table 2) indicating that low energy events are being removed from the branch, evident in figure 21. The events removed from the photon branch must be spread across all topologies otherwise the efficiency in this region would increase, instead figure 20 shows all bins affected equivalently. Given the reduced efficiency across the entire range, the most likely cut responsible is *truncation ratio*. As the *truncation ratio* compares most charge deposited into a layer against least the reduced light yield would mean the least charge can't decrease below the light threshold but the maximum is reduced, meaning for the given size of the cluster the ratio would be too low, rejecting the event.

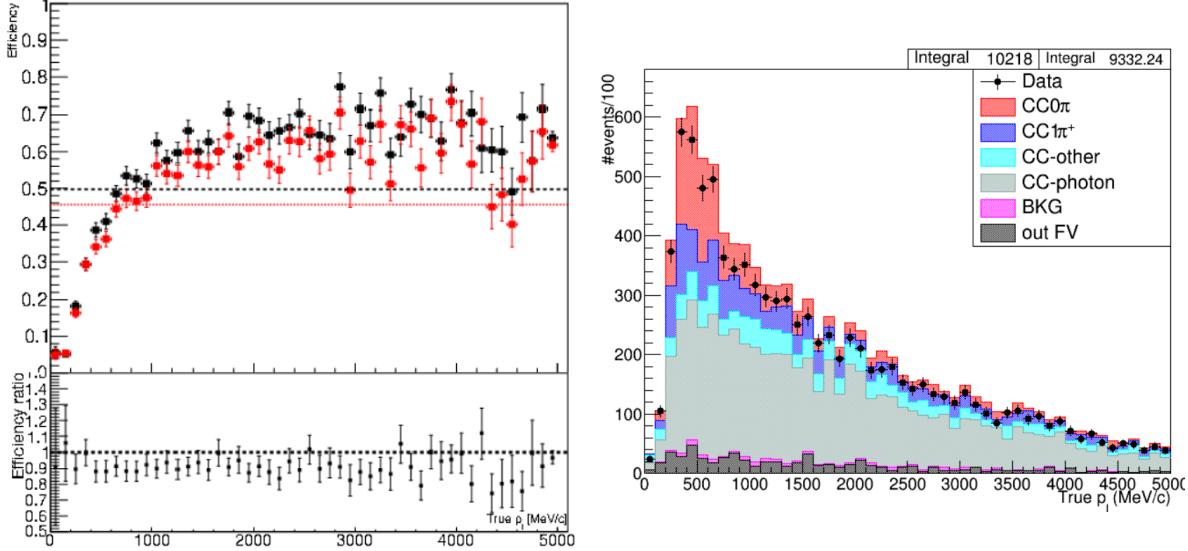


Figure 20: (left) The data in black is the efficiency of the photon branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

Figure 21: (right) A topological graph of the photon branch before aging (from the CCPHoton reconstruction), the topologies represented are from the truth directory. The black data points are the entries per bin post aging. The left integral is the unaged data while the right is aged; these integrals are weighted hence are not integers. ‘out FV’ is out of fiducial volume and ‘BKG’ stand for background meaning any topology that doesn’t fit in the branch structure (for example an electron neutrino at the primary vertex).

CCOther efficiency (figure 22) increased by 14%, suggesting that true CCOther events previously in the CCPHoton branch have transitioned to the CCOther; the algorithm got better at identifying non-specified charged-current interactions. Most interactions introduced to CCOther are below 1500 MeV were pions will often form showers, the increase in track-like *likelihood* ratios would mean interactions with multiple pions (that could have been reconstructed as EM pair production events) would be correctly identified. Section 6.4 shows EM-like tracks become more HIP-like, this aids the identification of charged-current resonant multi-pi events. Figure 23 has a large statistical error due to the sub-optimal number of events, meaning identifying underlying effects is difficult so increasing the number of files used would benefit future work.

The photon branch was expected to be affected most negatively, with a 8.5% reduction in efficiency (shown in table 2) when reducing light yield by 50% is a manageable change that can be corrected in the eventCalib stage of data processing.

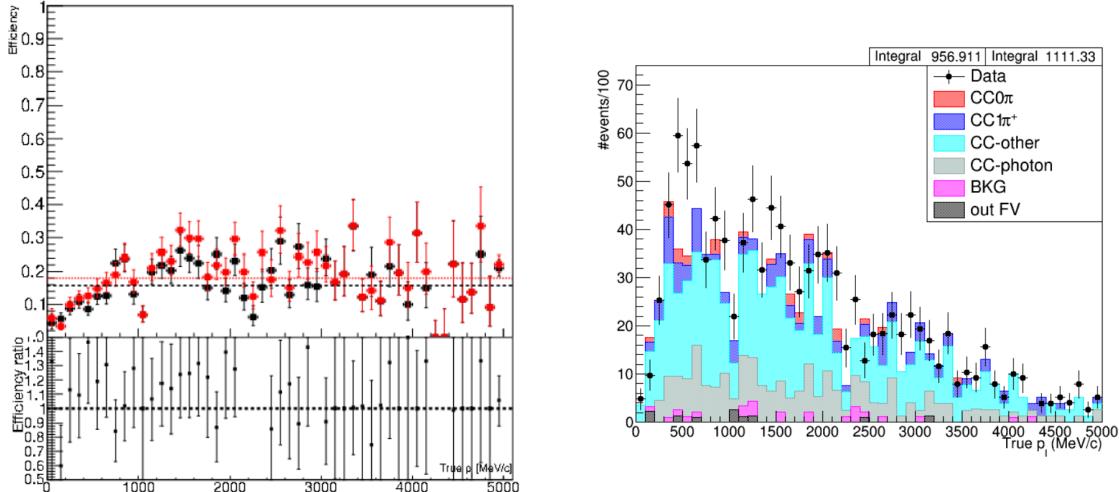


Figure 22: (left) The data in black is the efficiency of the CCOther branch of the CCPion reconstruction, the red data points are the aged data samples. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

Figure 23: (right) A topological graph of the CCOther branch before aging (from the CCPion reconstruction), the topologies represented are from the truth directory. The black data points are the entries per bin post aging. The left integral is the unaged data while the right is aged; these integrals are weighted hence are not integers. ‘out FV’ is out of fiducial volume and ‘BKG’ stand for background meaning any topology that doesn’t fit in the branch structure (for example an electron neutrino at the primary vertex).

Branch (Reconstruction)	Efficiency before aging	Efficiency after aging	Ratio
0 Pions (CCPion)	0.489 ± 0.002	0.488 ± 0.002	0.997 ± 0.01
1 Pions (CCPion)	0.291 ± 0.004	0.288 ± 0.004	0.992 ± 0.04
CC Other (CCPion)	0.295 ± 0.005	0.295 ± 0.005	0.999 ± 0.04
Photon (CCPhoton)	0.498 ± 0.005	0.456 ± 0.005	0.915 ± 0.03
0 Pions (CCPhoton)	0.460 ± 0.002	0.463 ± 0.002	1.000 ± 0.04
1 Pions (CCPhoton)	0.250 ± 0.004	0.251 ± 0.004	1.007 ± 0.04
CC Other (CCPhoton)	0.157 ± 0.006	0.180 ± 0.006	1.140 ± 0.12

Table 2: Mean efficiency of each branch before and after simulated aging. The ratio column is defined as efficiency after aging over efficiency before aging.

8.3 Assessing quality

Production 7 is optimised for low energy reconstructions however these highland cuts are not, so the good *quality* events are used in research. This is achieved in cut 3, the ‘quality and fiducial’ cut is altered, which for this project decreased the amount of events passing this stage by an average 13%.

The CCPion reconstruction was unaffected, due to the fluctuations in purity and efficiency to be less than 0.001%. The efficiency ratio of the 1 pion branch of the CCPhoton reconstruction decreased by 1.2% while photon and CCOther reduced by 9.8% and 13.1% respectively. It was unexpected that using good *quality* events increased the aging affects on the algorithms ability to identify the topology of an interaction. One possible explanation is that as there is an increased range of hits which will get the cut rejected (increased from two hits to six hits). Good *quality* events tend to be higher energy, providing a higher efficiencies and purity per branch before aging but being less resilient to the aging effects. Aging decreases the total reconstruction energy of high momentum events more severely than low momentum events (discussed in section 6.2) causing photon events to fail the *hit charge distribution* sub stage of the CCPhoton cut. *Truncation ratio* and *front-back ratio* may also affect the success rate of event selection.

In the CCPion data, the 1 pion and CCOther branches have reduced efficiencies due to the shower-like cluster becoming HIP-like, as discussed in section 6.4. This effect can also be attributed to the decrease in the CCOther purity ratio (0.973 ± 0.06 for good *quality* objects). EM/HIP blurring seems to be symmetrical as the purity ratio of photon branch became 0.989 ± 0.02 after a rise in CC resonance contaminates.

9 Conclusion

This dissertation aimed to observe the affect of aging the ND280 electromagnetic calorimeter, this was achieved by comparing 2005 simulated data files before and after a 50% reduction in light yield. These files were processed through the highland software using two separate reconstructions. The CCPion reconstruction focused only on the presents of pions while CCPhoton contained a branch for final states with a photon before considering pions.

By predicting the effects of aging on the reconstruction algorithm countermeasures can be enacted to correct any loss in sample purity or efficiency. If the T2K experiment can better define the neutrino oscillation parameters then the standard model could need to accommodate additional particles, such as sterile neutrinos, or reevaluate fundamental postulates of particle physics in CP symmetry. Both of these discoveries would answer cosmic questions, where CP violation would explain matters domination over anti-matter; sterile neutrinos could be a source of dark matter.

After aging the clusters lost on average 150 ± 700 MeV of momentum due to the light yield reduction, increasing the number of hits rejected by the light threshold and the decreasing the charge deposited by the remaining hits. The mean number of objects per beam spill was reduced by 0.15 ± 4 . As only the ECal was aged the electrons and photon reconstructions were affected most, reduced by 12% and 13% respectively, which was expected. The ECal object types shifted away from EM-like showers towards track-like, with a bias towards MIP reconstructions. These uncertainty were overestimated due to the high correlation in reconstruction variables.

Both the purity and efficiency of the CCPion data sample was minimally affected, any variation was within 2σ ; this can be attributed to pion reconstructions not explicitly using the ECal but secondary affects did affect the data, such as global tracker momentum variation leading to changes in *PID*. Aging reallocated events from the photon branch (in the CCPhoton reconstruction) into the 1 pion and CCOther branches; leading to a reduced photon efficiency of 0.456 ± 0.005 (compared to pre-aging 0.498 ± 0.005). This change can be explained by the reduced number of electron/photon *PID* combined with the *likelihood* ratios making shower-like objects reconstruct as track-like. CCPhoton branch purity is mostly unchanged suggesting that the cuts are sufficient. CCOther's efficiency increase could be due to correcting previously identified EM pair production as multiple pion interactions. Another possibility is the CC resonance interactions producing multiple pions become more easily detectable as the pions become more track-like than shower-like.

One area that could be further researched is starting with a single interaction topology to further explore the transition effects in purity and efficiency that maybe overshadowed, incorporating additional detectors or interaction topologies would also be beneficial. Many event selection parameters would need to be adjusted, such as lowering the expected charge deposited per unit distance for each particle species. By lowering the light yield threshold, it may be possible to reverse some effects of aging; this shouldn't affect noise hits on a large scale as they are also being suppressed by the light yield reduction. It would be possible to incorporate low-energy reconstruction software within the ‘eventCalib’ stage of data processing to correct light yield loss, recovering the pre-aged event energy spectrum. Lost hits wouldn't be recovered so wouldn't aid in correcting ECal object types; true shower-like objects would still appear as track-like.

This study has shown that reconstruction in the ECal is minimally affected by aging so can continue taking data until 2040, significantly past the Hyper-K completion date.

References

- [1] J. Kopp et al. *Sterile neutrino oscillations: the global picture*. High Energy. Phys. 2013, 50 (2013). [https://doi.org/10.1007/JHEP05\(2013\)050](https://doi.org/10.1007/JHEP05(2013)050)
- [2] P. Zyla et al. *Review of Particle Physics*. PTEP, 2020(8):083C01, 2020.
- [3] K. Abe et al (T2K Collaboration). *Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations*. Nature, 580:339–344, April 2020.
- [4] K. Abe et al (T2K Collaboration). *The T2K experiment*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 659(1):106–135, December 2011.
- [5] K. Abe et al (Hyper-Kamiokande Proto-Collaboration). *Hyper-kamiokande design report*. arxiv.org, May 2018. 333 pages.
- [6] Hyper-K collaboration. *Hyper-Kamiokande news*. <http://www.hyper-k.org/en/>. Accessed: 2022-03-22
- [7] A. Rodriguez et al. *Measurement of single charged pion production in the charged-current interactions of neutrinos in a 1.3 GeV wide band beam*. Physical Review D, 78(3):032003, 2008.
- [8] L. Kormos (2019). *Recent T2K Neutrino Oscillation Results* EPS-HEPS 2019, European Physical Society Conference on High Energy Physics, Belgium, 10-17 July.
- [9] L. Southwell, *Selecting electron anti-neutrino charged-current interactions at the T2K ND280*, Ph.D. thesis, University of Lancaster, U.K. (2016) [<https://t2k.org/docs/thesis/T2K-THESIS-075>].
- [10] The T2K ND280 Upgrade Working Group. *ND280 Upgrade Project: Addendum to T2K ND280 Upgrade Technical Design Report*. CERN, 2020.
- [11] P. Amaudruz et al. *The T2K fine-grained detectors*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 696:1–31, 2012.
- [12] JPARC collaboration. *What is J-PARC neutrino experimental facility?*. <https://j-parc.jp/c/en/facilities/nuclear-and-particle-physics/neutrino.html>. Accessed: 2022-03-22
- [13] M. Friend. *J-PARC accelerator and neutrino beamline upgrade program*. Journal of Physics: Conference Series, volume 888, page 012042. IOP Publishing, 2017.
- [14] K. Abe et al. *Scintillator ageing of the T2K near detectors from 2010 to 2021* (in preparation for publication). T2K collaboration (private communication), 2022.
- [15] L. Kormos et al (T2K collaboration). *The electromagnetic calorimeter for the T2K near detector ND280*. Journal of Instrumentation, 8(10), October 2013.

- [16] M. Scott, *Measuring charged current neutrino interactions in the electromagnetic calorimeters of the ND280 detector*, Ph.D. thesis, Imperial College London, U.K. (2013) [<https://t2k.org/docs/thesis/T2K-TESIS-036>].
- [17] ND280 collaboration (private communication). *ND280 software - description of nd280 packages*. <https://nd280.lancs.ac.uk/stable/invariant/nd280Doc/workbook/T2KSoftwareExplained.html>. Accessed: 2022-03-22
- [18] R. Brun and F. Rademakers. *Root — an object oriented data analysis framework*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 389(1):81–86, 1997. New Computing Techniques in Physics Research V.
- [19] S. Agostinelli et al. *Geant4 — a simulation toolkit*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3):250–303, 2003.
- [20] P.A. Amaudruz. *Midas data format*. http://lmu.web.psi.ch/docu/manuals/bulk_manuals/software/midas195/html/AppendixA.html. Accessed: 2022-03-22
- [21] ND280 collaboration (private communication). *The flat tree: how to save time and space*. https://nd280.lancs.ac.uk/highland/2.62/highlandND280Doc_2.5/dox/flattree.html. Accessed: 2022-03-22
- [22] Vacheret et al. *The front end readout system for the T2K-ND280 detectors*. In 2007 IEEE Nuclear Science Symposium Conference Record, volume 3, pages 1984–1991. IEEE, 2007.
- [23] T2K collaboration (private communication). *ND280 offline software*. https://nd280.lancs.ac.uk/lxr/source/eventAnalysis_7.0/inc/TTrackerECALReconModule.hxx#0079. Accessed: 2022-03-22
- [24] T. Doyle et al. *ND280 FHC samples with Photon and Proton information for the 2021 Oscillation Analysis*. Lancaster University, United Kingdom. Institute for nuclear research - INR RAS, Russia. NCBJ Warsaw, Poland. Stony Brook University, USA, 2021.
- [25] S. Valder, *Measurement of the νe CC π^+ cross-section using the ND280 tracker and development of optical diffuser calibration systems for Hyper-Kamiokande*, Ph.D. thesis, University of Warwick, U.K. (2020) [https://warwick.ac.uk/fac/sci/physics/staff/academic/boyd/stuff/thesis_110820.pdf].
- [26] T. Jolliffe and J. Cadima. *Principal component analysis: a review and recent developments*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065):20150202, 2016
- [27] G. Barker et al. *Implementation of the second generation PID for the ND280 tracker ECals*. T2K internal technical note TN, 111, 2012.

10 Appendices

A Event selection variables

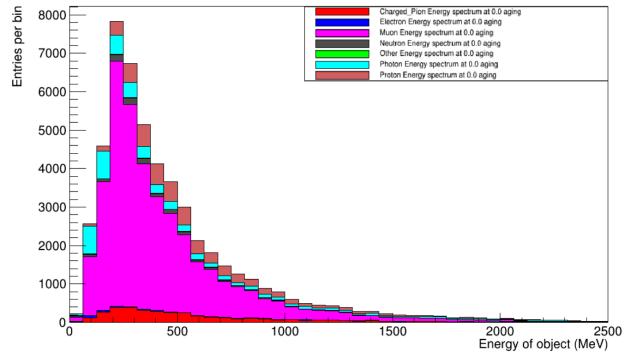


Figure 24: The energy distribution of reconstructed ECal objects separated by PID using unaged oaEvents files.

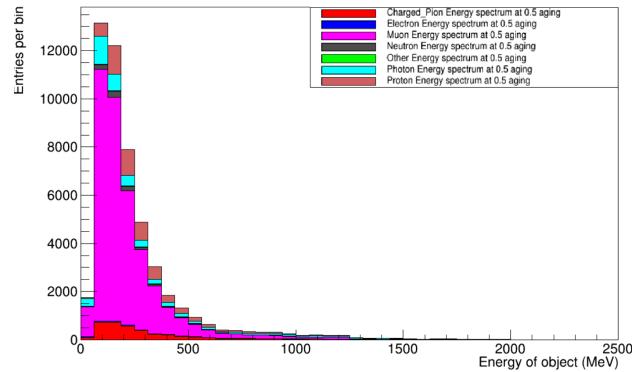


Figure 25: The energy distribution of reconstructed ECal objects separated by PID using aged oaEvents files.

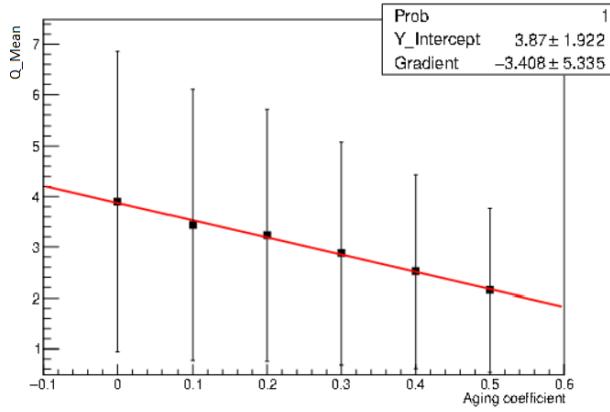


Figure 26: The mean Q value of each ECal object against aging coefficient .

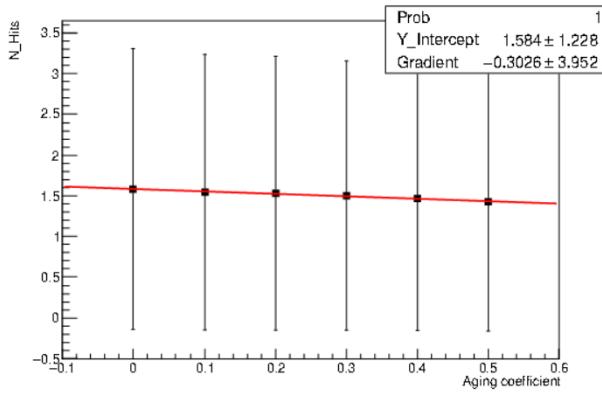


Figure 27: The mean charge of hits in the ECal object against aging coefficient.

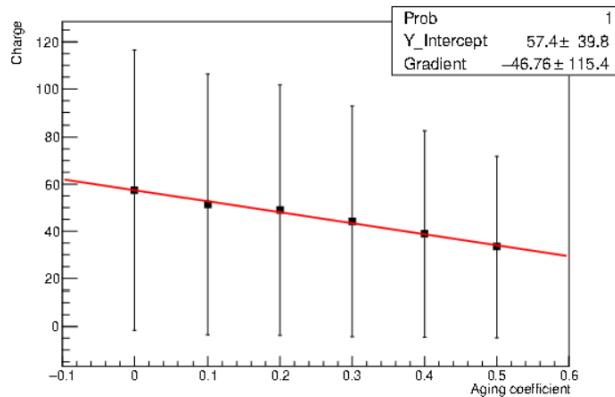


Figure 28: The mean charge deposited per ECal object against aging coefficient.

B Likelihood ratios in event selection

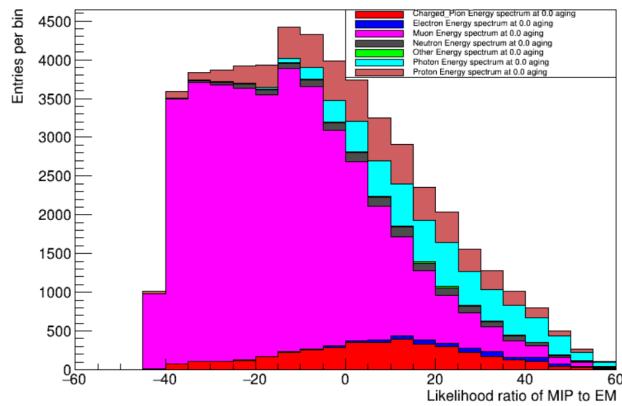


Figure 29: The likelihood_MIP_EM distribution of reconstructed ECal objects separated by PID using unaged oaEvents files.

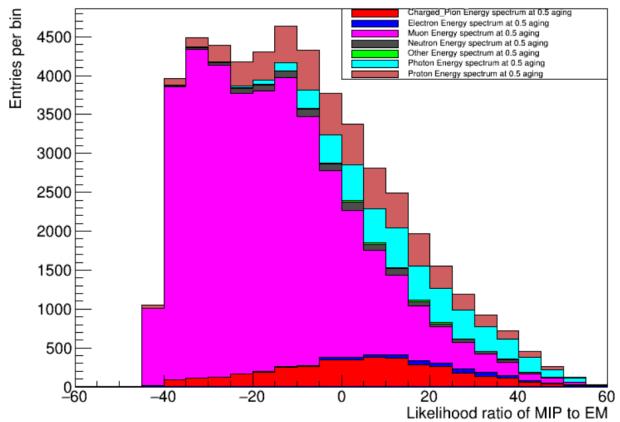


Figure 30: The likelihood_MIP_EM distribution of reconstructed ECal objects separated by PID using aged oaEvents files.

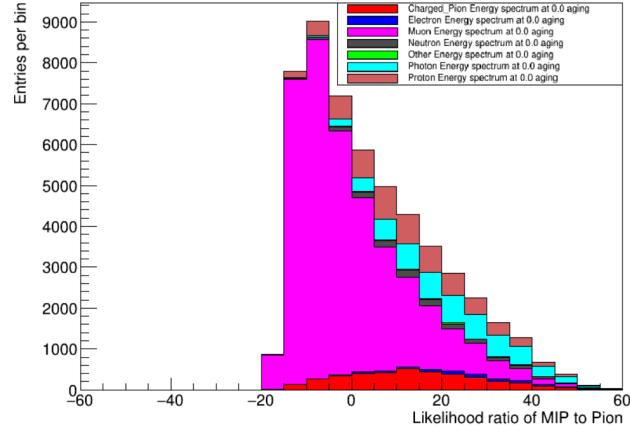


Figure 31: The likelihood_MIP_Pion distribution of reconstructed ECal objects separated by PID using unaged oaEvents files.

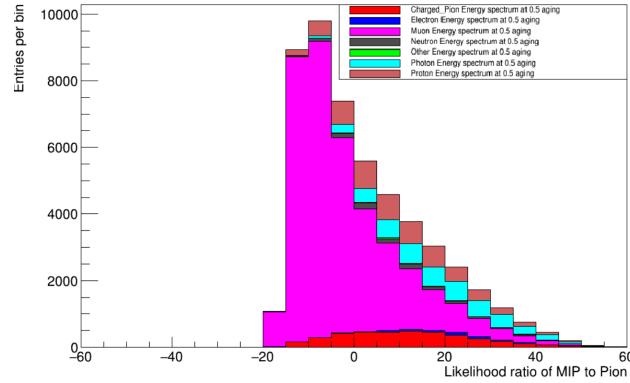


Figure 32: The likelihood_MIP_Pion distribution of reconstructed ECal objects separated by PID using aged oaEvents files.

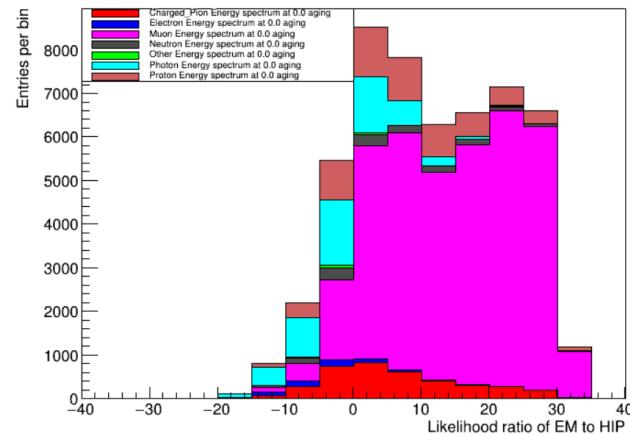


Figure 33: The likelihood_EM_HIP distribution of reconstructed ECal objects separated by PID using unaged oaEvents files.

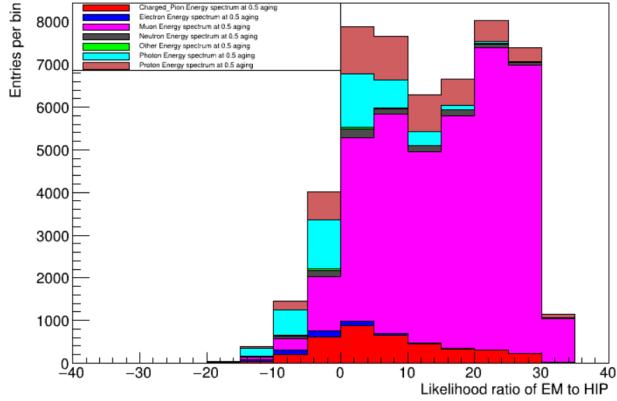


Figure 34: The likelihood_EM_HIP distribution of reconstructed ECal objects separated by PID using aged oaEvents files.

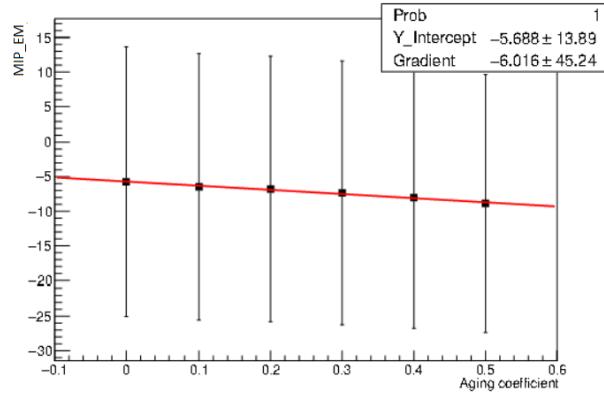


Figure 35: The mean likelihood_MIP_EM of reconstructed ECal objects against aging coefficient.

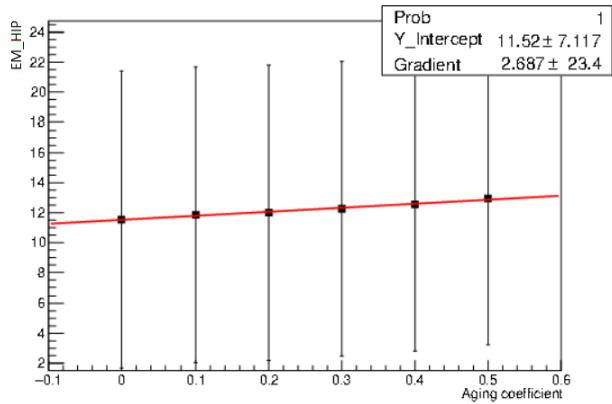


Figure 36: The mean likelihood_EM_HIP of reconstructed ECal objects against aging coefficient.

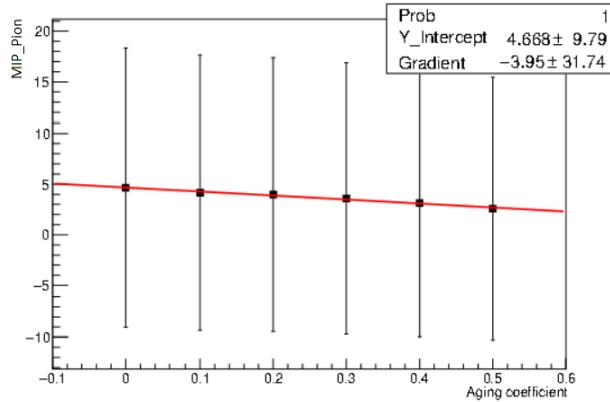


Figure 37: The mean likelihood_MIP_Pion of reconstructed ECal objects against aging coefficient.

C Branch efficiency and purity

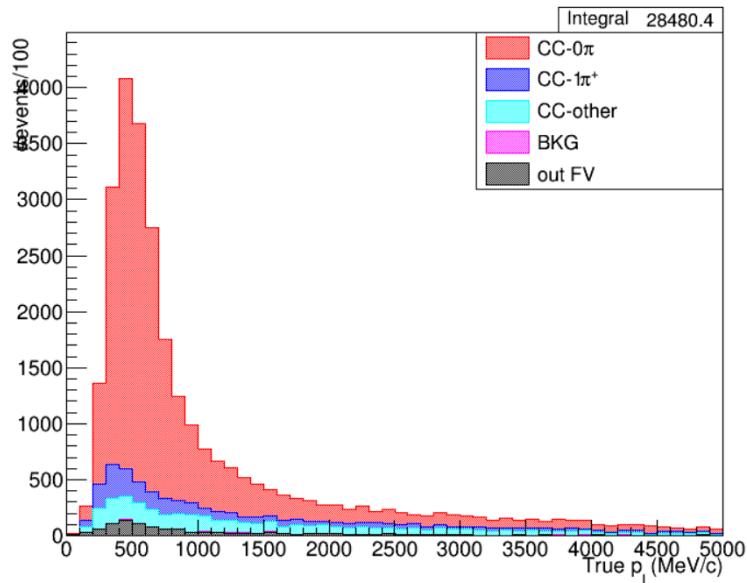


Figure 38: 0 pion branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

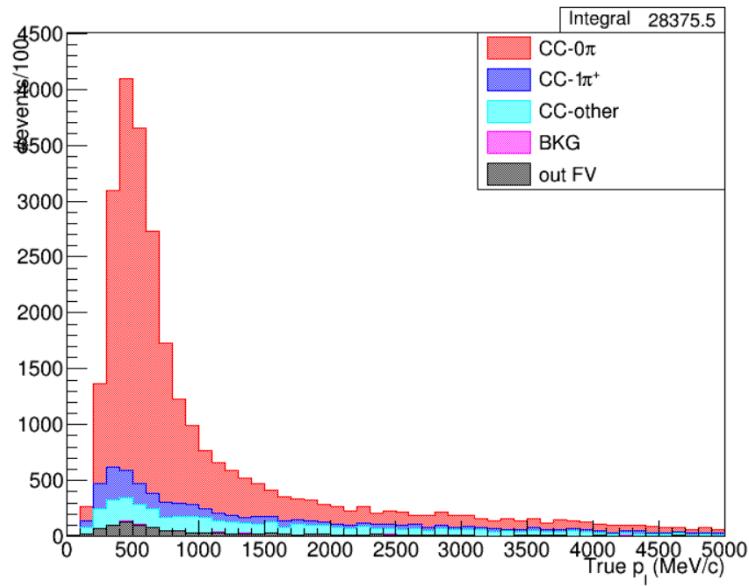


Figure 39: 0 pion branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

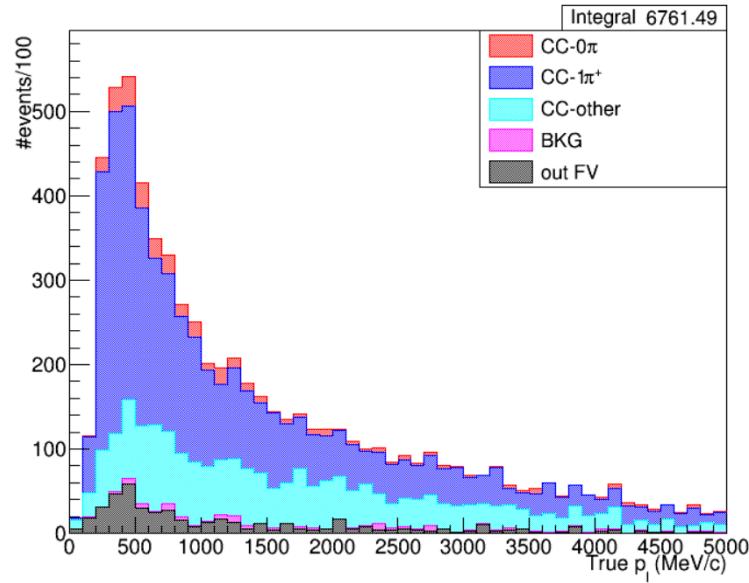


Figure 40: 1 pion branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

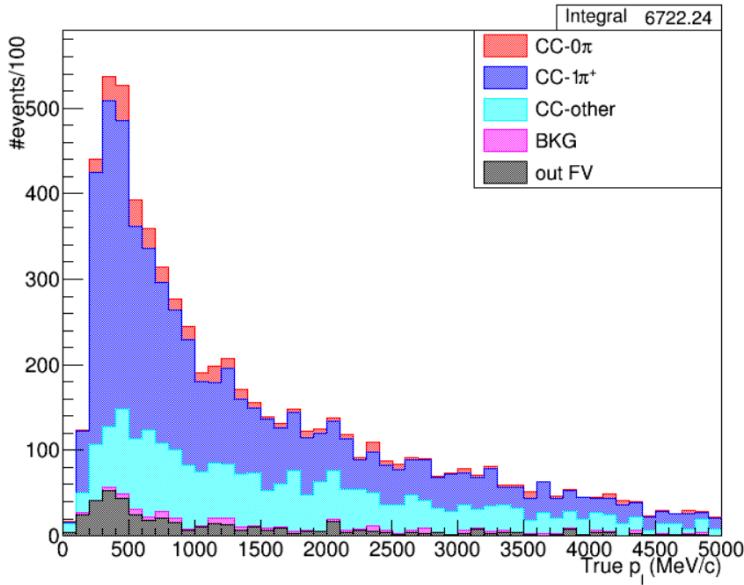


Figure 41: 1 pion branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

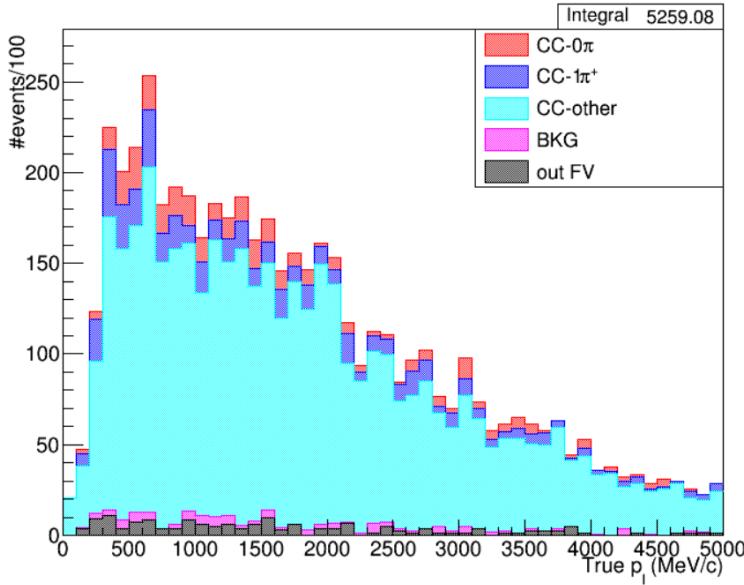


Figure 42: CCOther branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

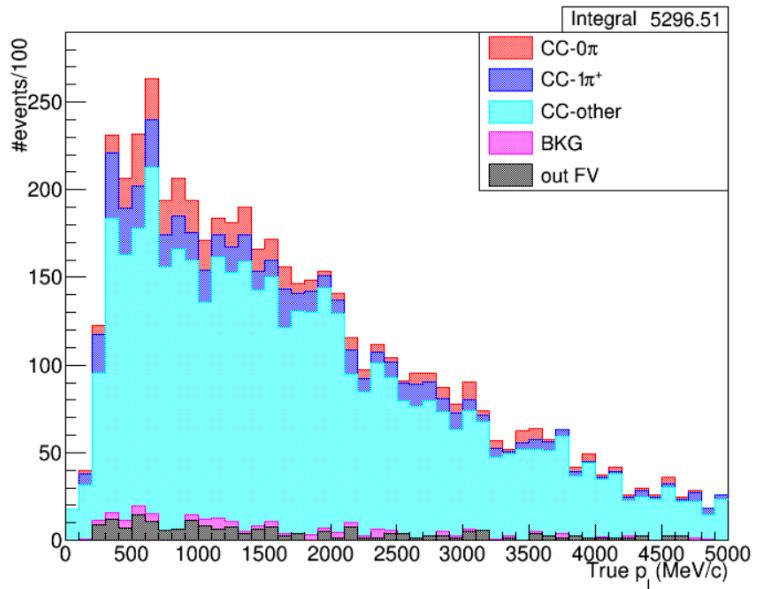


Figure 43: CCOther branch (in the CCPion reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

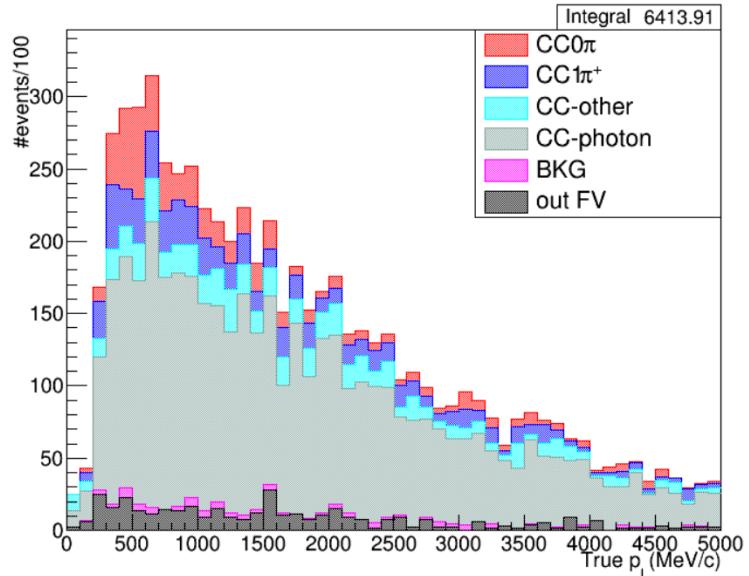


Figure 44: Photon branch (in the CCPhoton reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

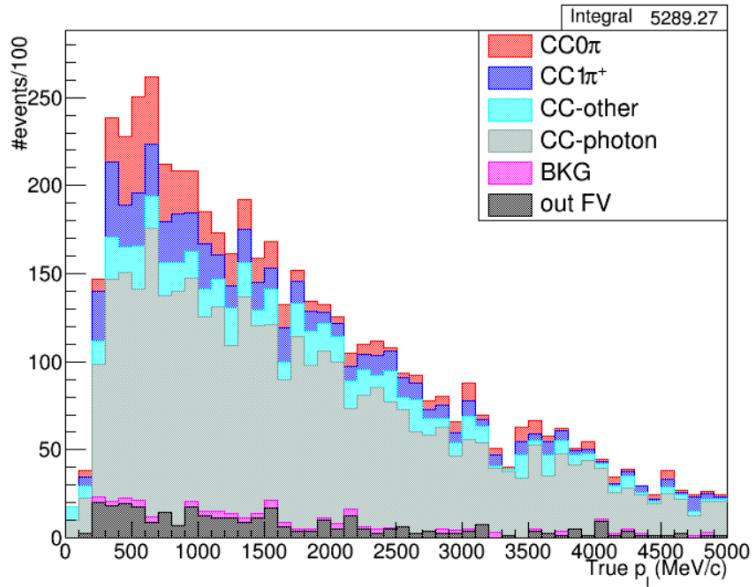


Figure 45: Photon branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

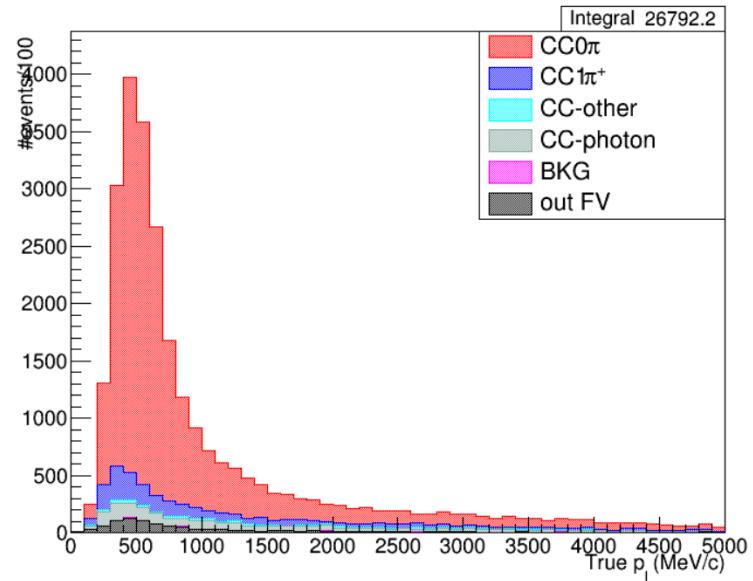


Figure 46: 0 pion branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

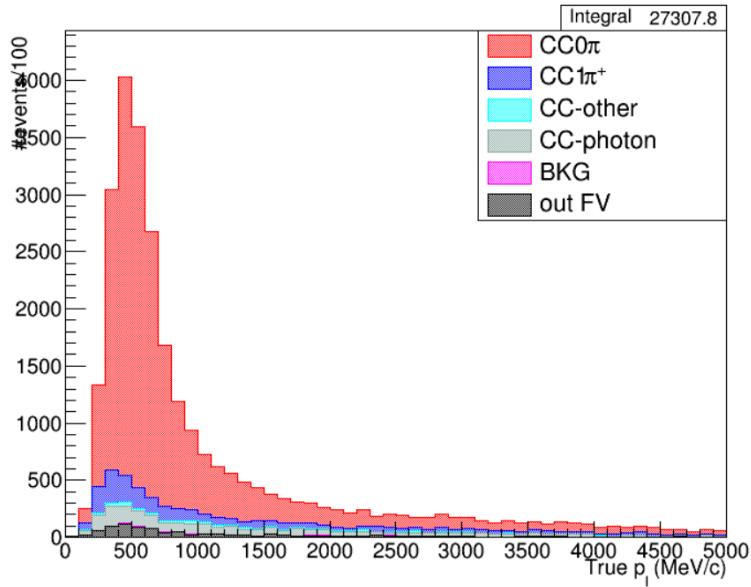


Figure 47: 0 pion branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

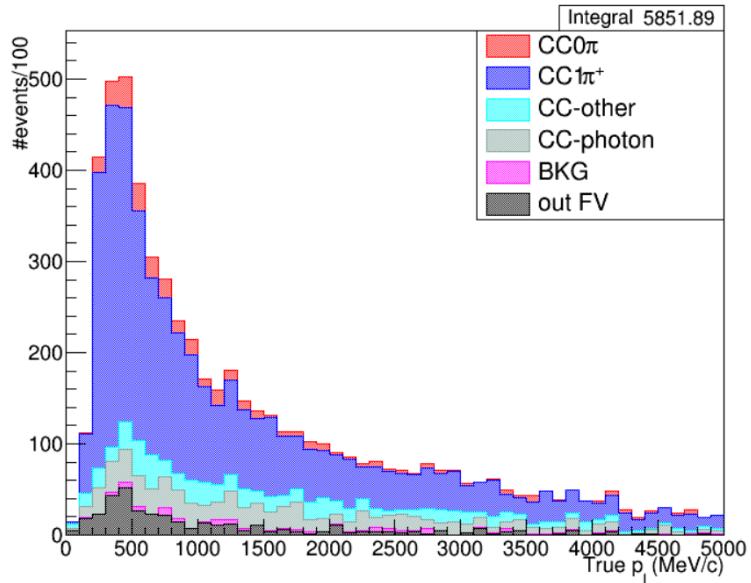


Figure 48: 1 pion branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

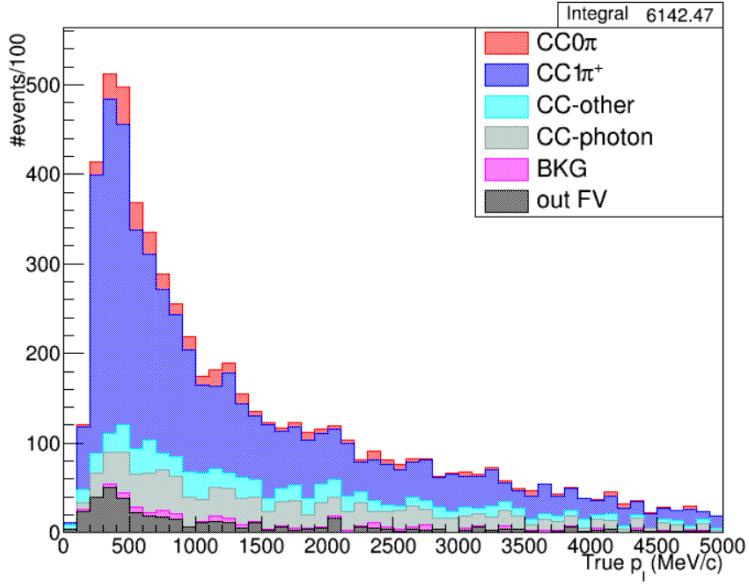


Figure 49: 1 pion branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

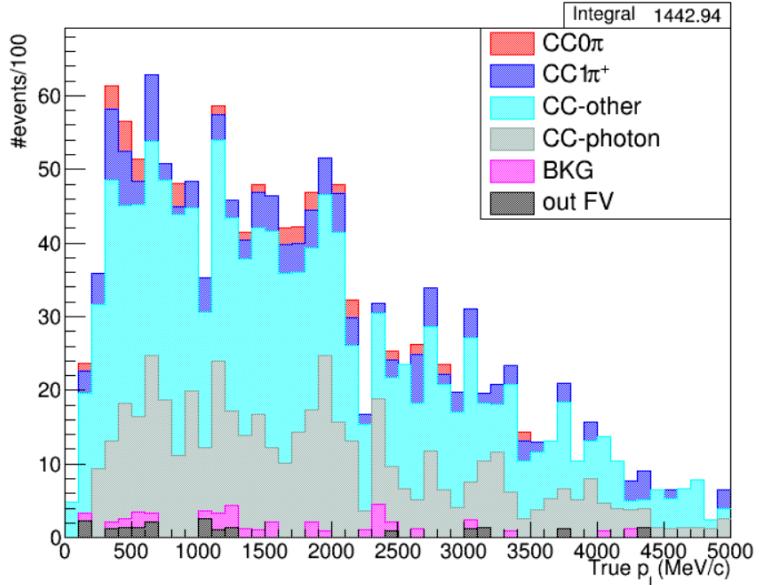


Figure 50: CCOther branch (in the CCPHoton reconstruction) broken down by true interaction topology (using the true directory) using the unaged highland files.

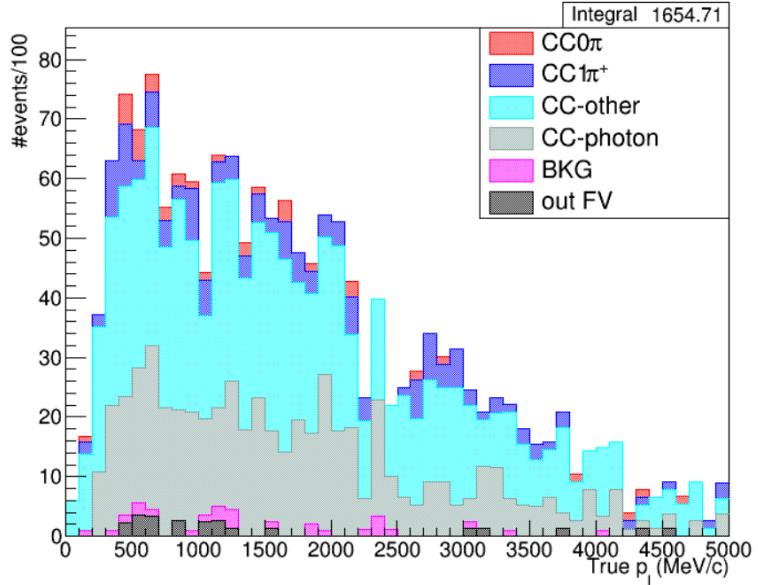


Figure 51: CCOther branch (in the CCPhoton reconstruction) broken down by true interaction topology (using the true directory) using the aged highland files.

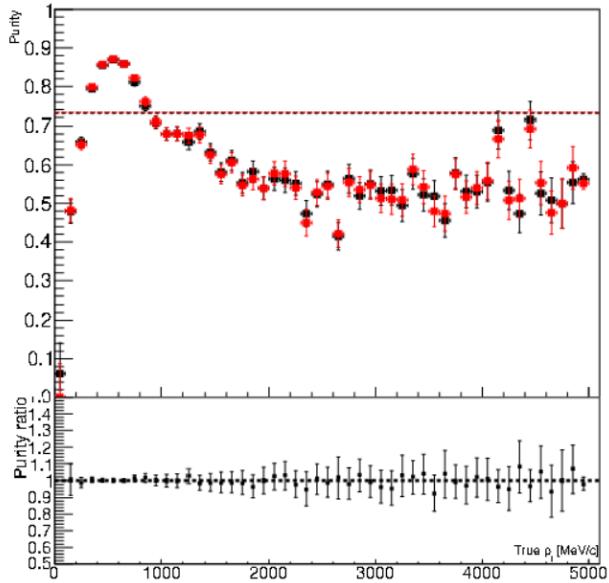


Figure 52: The data in black is the purity of the 0pion branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

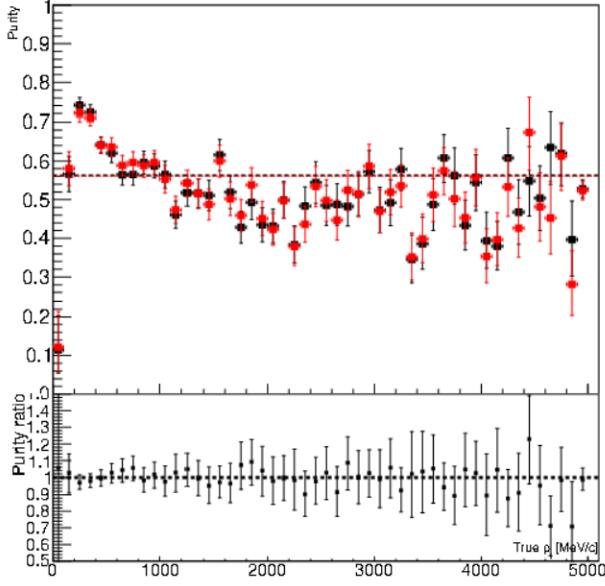


Figure 53: The data in black is the purity of the 1pion branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

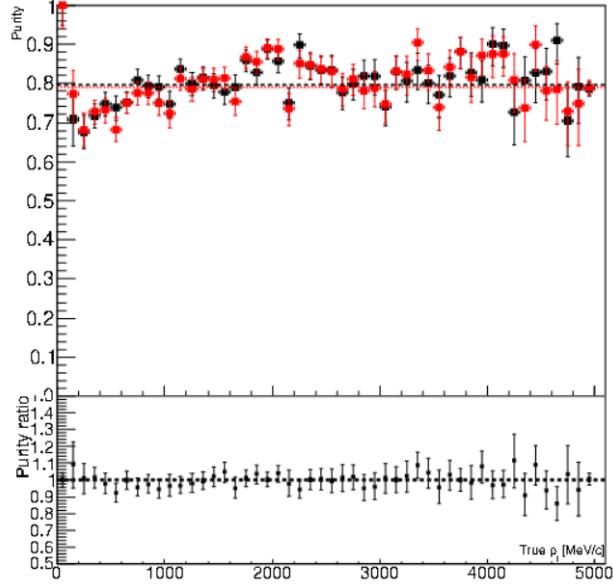


Figure 54: The data in black is the purity of the CCOther branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

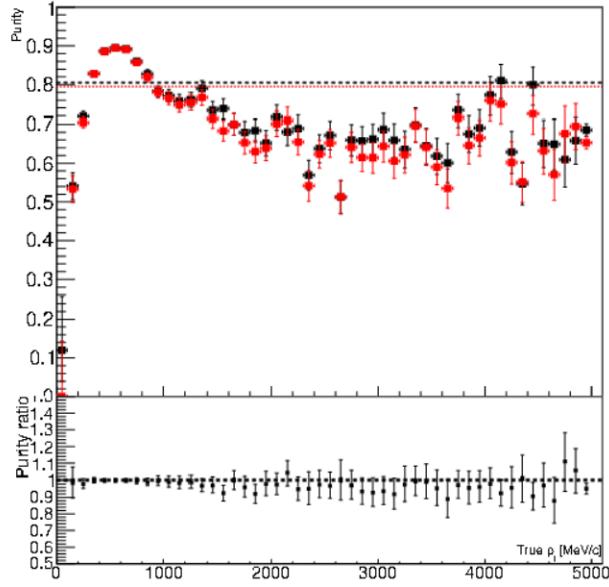


Figure 56: The data in black is the purity of the 0pi branch of the CCPhoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

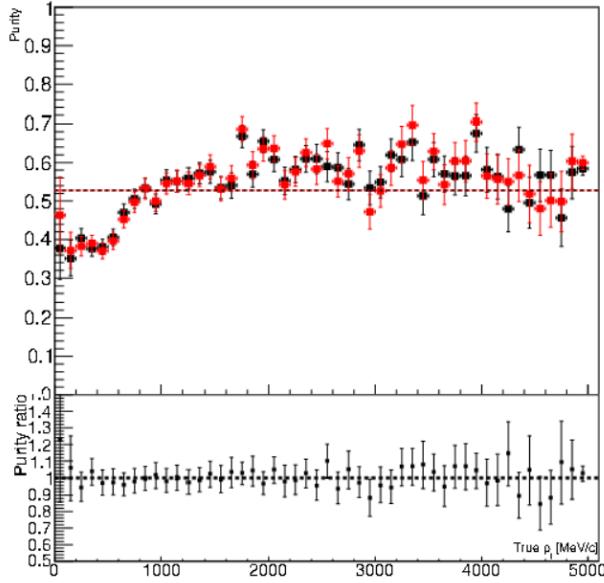


Figure 55: The data in black is the purity of the photon branch of the CCPhoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

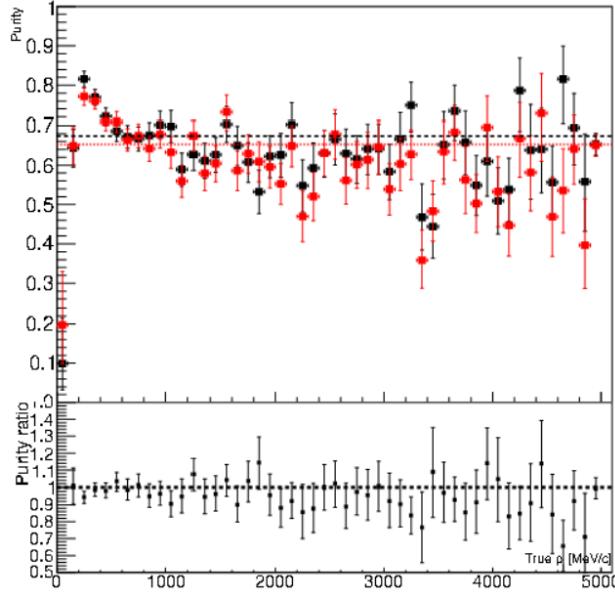


Figure 57: The data in black is the purity of the 1pion branch of the CCPhoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

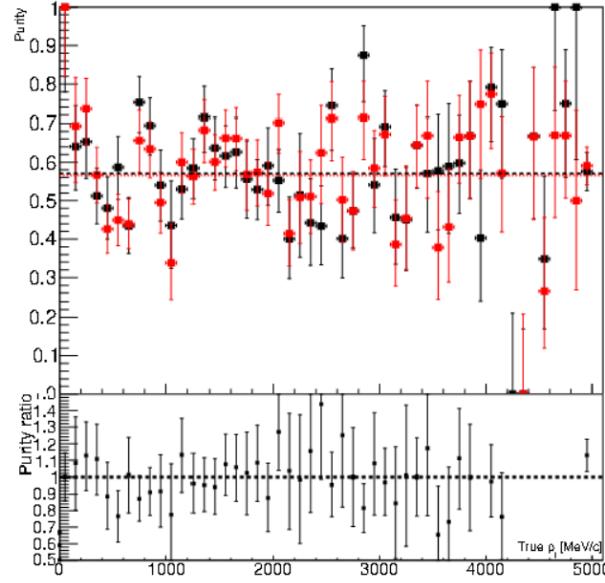


Figure 58: The data in black is the purity of the CCOther branch of the CCPhoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

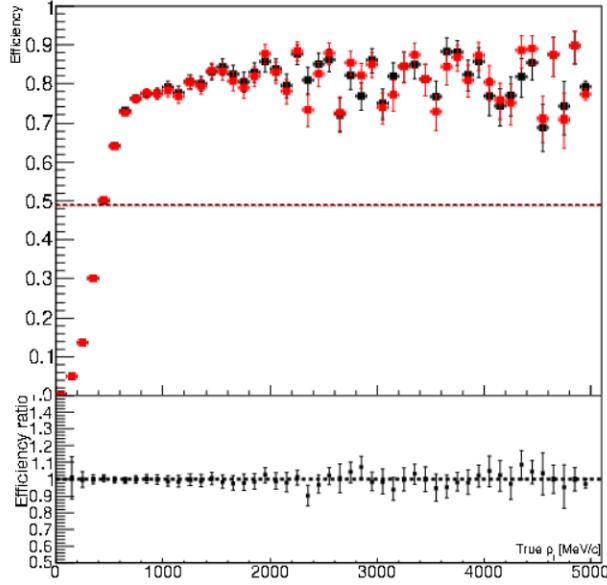


Figure 59: The data in black is the efficiency of the 0 pion branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

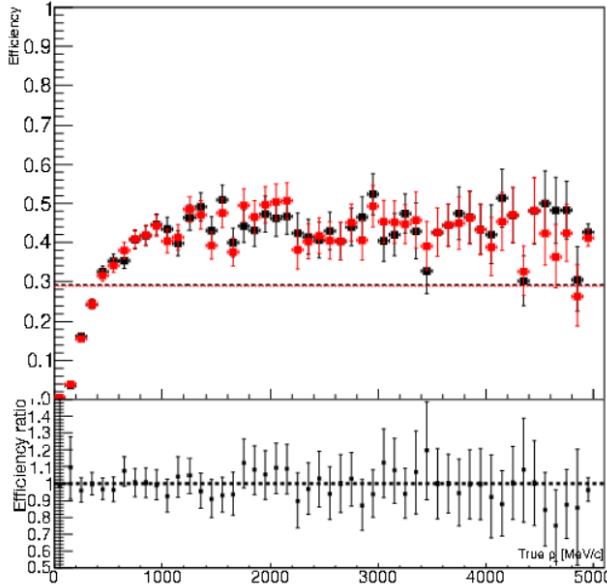


Figure 60: The data in black is the efficiency of the 1 pion branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

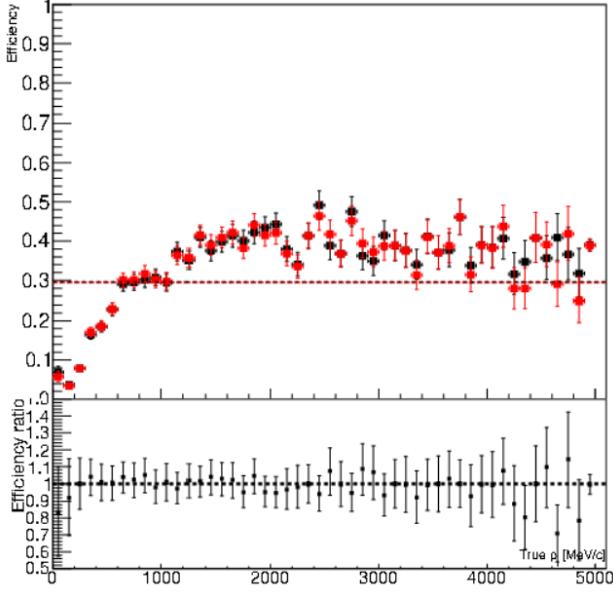


Figure 61: The data in black is the efficiency of the CCOther branch of the CCPion reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

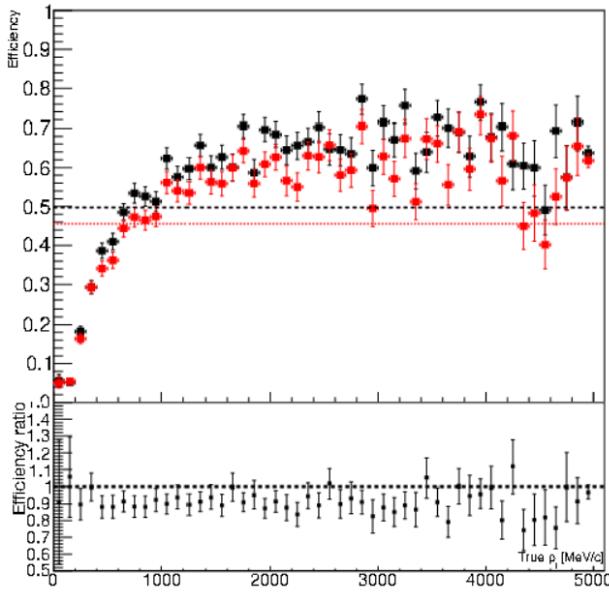


Figure 62: The data in black is the efficiency of the photon branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

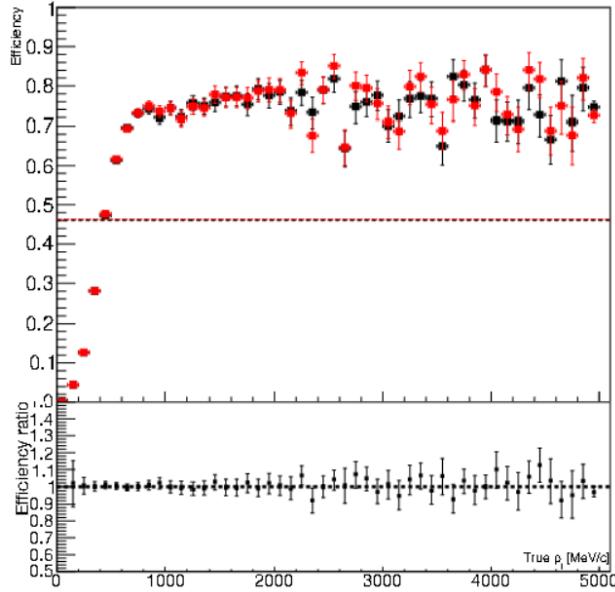


Figure 63: The data in black is the efficiency of the 0 pion branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

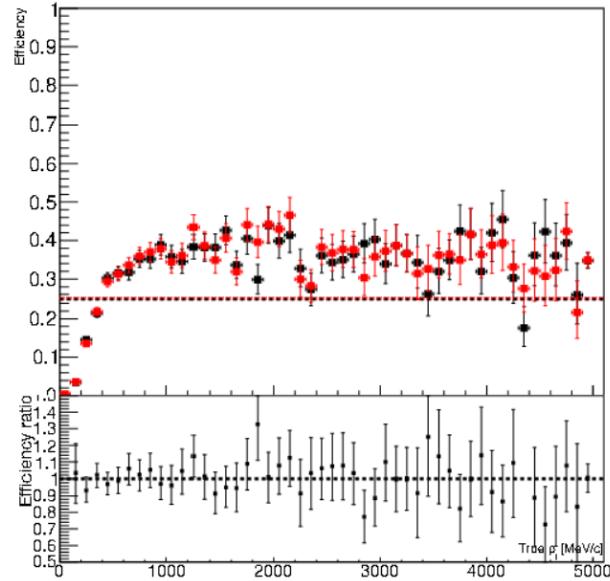


Figure 64: The data in black is the efficiency of the 1 pion branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.

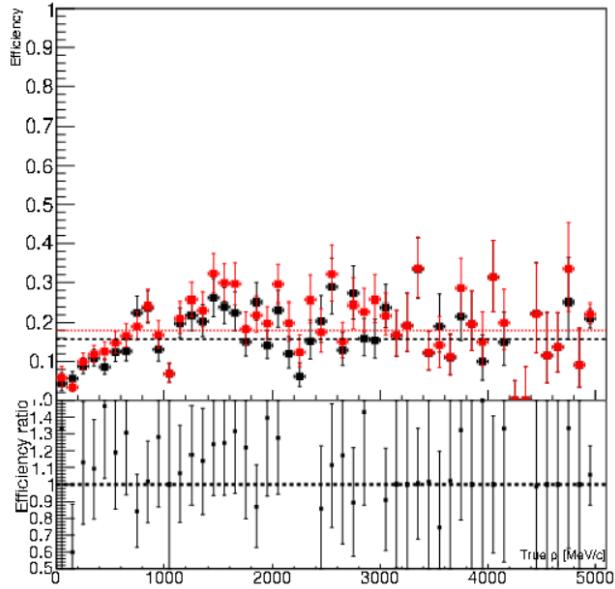


Figure 65: The data in black is the efficiency of the CCOther branch of the CCPHoton reconstruction, the red data points are from aged data. The horizontal guideline line is the mean value across all bins. Below that is the ratio of aged efficiency over unaged purity per bin. The momentum presented is the true lepton (at the primary vertex) momentum.