Study of the neutral current interaction with a π^0 in the final state in the ND280 detector of the T2K experiment

Elżbieta Popławska

Thesis submitted to the Queen Mary, University of London
for the degree of Doctor of Philosophy

Particle Physics Research Centre

Queen Mary
University of London

2015

Abstract

The Tokai to Kamioka (T2K) experiment is designed to detect $\nu_\mu \rightarrow \nu_e$ appearance for the first time from accelerator generated neutrinos and measure neutrino oscillation parameters. The most intense accelerator muon neutrino beam ever built is produced at the J-PARC facility on the East coast of Japan, and sent 2.5° off-axis toward the Super-Kamiokande water Cherenkov detector. The main background for the recent appearance analysis comes from neutral current $1\pi^0$ interaction, accounting for 23% of the background events. In addition, the near detector (ND280) has been built to serve the requirement to provide high accuracy beam characteristics and background studies by measuring a number of neutrino cross-sections.

In this thesis the NC$1\pi^0$ channel studies at the near detector ND280 is performed. The studies aim to develop the NC$1\pi^0$ event selection to single out the muon neutrino induced NC$1\pi^0$ interactions in the PØD (π^0 Detector) sub-detector, where one π^0 decay gamma converts in the PØD and the second gamma converts in the PØD-ECal (π^0 Electromagnetic Calorimeter) sub-detector. The used data sample was collected during Runs II,III and IV, consisting of a total of 6.2×10^{20} POT. Simulated events are generated with the NEUT MC to evaluate backgrounds and estimate the selection efficiency. The studies have been performed for the two PØD configurations: filled with water and without water.

The first approach to the usage of PØD-ECal sub-detector in the NC$1\pi^0$ channel studies showed that further improvements on the PØD-ECal reconstruction and matching between the PØD and PØD-ECal is required. However, as a result of the performed analysis the total efficiency of 3.0% (4.0%) and purity of 17.7% (16.4%) is obtained for the PØD filled with water (without water) configuration. Finally, after selection, 68 ± 8.2 (stat) (107 ± 10.3 (stat)) events passed all the requirements for the PØD filled with (without) water. The developed selection lay the groundwork for the future measurement of the NC$1\pi^0$ cross-section for the events induced by neutrinos interacting in the PØD sub-detector and systematic errors evaluation, using larger data sets. As for now, the systematic uncertainties are a subject of a discussion in this thesis.

I, Elzbieta Poplawska, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third partys copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Elzbieta Poplawska

London, 2015

Dedicated to my sister Ula,
with love.

And to all the girl geeks.

E.P.

Acknowledgment

Firstly, I would like to thankfully acknowledge guidance and shared knowledge from my supervisor, Prof Francesca di Lodovico. I would especially like to express my thanks to her for the generous level of support and encouragement I received towards the end of my PhD.

Next, I would like to thank Dr Ben Still, who have introduced me to the work on the π^0, for the countless helpful discussions we had that contributed to the development of this analysis. Next, I would like to thank Dr Jeanne Wilson and Dr Teppei Katori, who answered endless number of my spontaneous random neutrino physics questions which was enormously beneficial, despite their busy schedule. I am grateful to Dr Ryan Terri who has contributed significantly to my motivation over the course of my doctorate and Dr Eram Rizvi for his genuine support and encouragement.

I am very thankful for the feedback and suggestions I received from Dr Phillip Litchfield and Dr Martin Haigh during my work on calibration and other service task matters. I would also like to express my gratitude to the π^0 working group members with whom I had valuable discussions, in particular I would like to acknowledge suggestions and clarifications I received from Dr Helen O'Keeffe. In addition I would like to express my thanks to Dr Sam Short who shared with me all her knowledge and experience in π^0 analysis.

Further, I thank the QMUL IT group: Dr Alex Owen, Dr Cozmin Timis and Dr Chris Walker who kept me sane while dealing with computing matter. The data production would not be possible without the expertise and support from Dr Jon Perkin.

I must give special thanks to all the students and staff at QMUL, in particular the neutrino group, as well as T2K UK members and the T2K collaboration. You have enabled me to have a great research training along with the British, Japanese and Italian culture experience.

As part of the research I had a great pleasure to live and work in Tokai-mura. I would like to express my gratitude to the residents of Tokai for their honorable attitude and courage after the Great East Japan Earthquake.

Finally, I would like to thank all of my family and friends for their dedicated understanding over the last years.

Contents

List of Figures

List of Tables

Abbreviations list

ADC	Analog to Digital Converter
BANFF	Beam And ND280 Flux extrapolation task Force
BDT	Boosted Decision Tree
BDTcat	Boosted Decision Tree with Categories
BrECal	Barrel Electromagnetic Calorimeter
CC	Charged Current
CECal	Central Electromagnetic Calorimeter
CMT	Configuration Management Tool
COH	Coherent
CP	Charge Parity
CTM	Cosmic Trigger Module
CVS	Concurrent Versioning System
DAQ	Data Acquisition
DIS	Deep Inelastic Scattering
DPT	Data Processing Task
DQ	Data Quality
DsECal	Downstream Electromagnetic Calorimeter
ECal	Electromagnetic Calorimeter
EM	Electromagnetic
ES	Elastic Scattering
FGD	Fine Grained Detector
FSI	Final State Interactions
GPS	Global Positioning System

GSC	Global Slow Control
ID	Inner Detector
IH	Inverted Hierarchy
INGRID	Interactive Neutrino GRID
J-PARC	Japan Proton Accelerator Research Complex
LINAC	Linear Accelerator
MC	Monte Carlo
MCM	Master Clock Module
MEU	MIP Equivalent Units
MIP	Minimally Ionising Particle
MLP	Multi Layered Perception
MLPcat	Multi Layered Perception with Categories
MPPC	Multi-Pixel Photon Counter
MSW	Mikheyev-Smirnov-Wolfenstein effect
MVA	Multivariate Analysis
NC	Neutral Current
NH	Normal Hierarchy
OD	Outer Detector
P.E.	photon equivalent unit
PØD	π^0 Detector
PØD-ECal	π^0 Electromagnetic Calorimeter
PDG	Particle Data Group
PEU	Pixel-Equivalent Units
PID	Particle Identification

PMNS	B. Pontecorvo, Z. Maki, M. Nakagawa and S. Sakata mixing matrix
PMT	Photomultiplier Tube
POT	Protons On Target
PS	Proton Synchrotron
QES	Quasielastic Scattering
RCS	Rapid Cycling Synchrotron
RES	Resonance Scattering
RMM	Readout Merger Module
ROC	Receiver Operating Characteristic
SK	Super-Kamiokande
SM	Standard Model
SMRD	Side Muon Range Detector
SNO	Sudbury Neutrino Observatory
SSM	Standard Solar Model
T2K	Tokai to Kamioka
TDC	Time to Digital Converter
TFB	Trip-t Front End Board
TOF	Time Of Flight
TPC	Time Projection Chamber
USECal	Upstream Electromagnetic Calorimeter
WLS	Wavelength Shifting
nuMSM	Neutrino Minimal Standard Model

Chapter 1

Introduction

This thesis describes the study of neutrino induced neutral current interactions with a π^0 in the final state and includes for the first time the PØD-ECal sub-detector in the analysis.

In the first two chapters neutrino physics is briefly presented. First, in chapter 2, the three main themes in contemporary neutrino physics are discussed, i.e. neutrino masses and their hierarchy, oscillations and CP violation phase. Each of them is a challenging subject for the current and future experimental measurements. In order to answer some of the questions in neutrino physics, it is required to reach the best possible knowledge about neutrino interactions with matter. These interactions are described in chapter 3, with emphasis on the neutral current π^0 production, which is the main subject in this thesis. The current knowledge of the neutral current π^0, NC1π^0, and cross section measurements from a number of experiments are summarised.

Chapter 4 is an introduction to the T2K experiment. A detailed description of the near detector ND280 and far detector Super-Kamiokande is provided; this includes a review of the PØD and PØD-ECal sub-detectors that are used to make the measurement in this thesis.

The following chapter, chapter 5, summarises the information about the data collected during Runs 1,2,3,4 and 5 in addition to a description of the data quality requirements at ND280.

The ND280 software structure is introduced and explained in chapter 6. First, the Monte Carlo simulation stages are described followed by the calibration and reconstruction explanation. The PØD and PØD-ECal reconstruction software structure is emphasised as a preparation and introduction towards the NC1π^0 analysis.

In chapter 7 the work on the initial NC1π^0 analysis in the PØD and PØD-ECal

sub-detectors is presented. This analysis lay the basis for the future NC1π^0 cross section measurement for neutrinos interacting in the PØD. The development of a PØD-PØD-ECal inclusive NC1π^0 event selection is described, including the usage of a multivariate analysis method with a boosted decision trees algorithm. The final brief review of the current readiness of the PØD-ECal for the analysis is given along with the proposal of the future improvements.

Chapter 8 contains a discussion on the systematic errors followed by chapter 9 which concludes the work presented in this thesis.

In addition, there are appendices which contain a detailed description of the MVA input variables for the signal and background as well as the Boosted Decision Trees classifier for signal and background discrimination distributions for the different fiducial volume values.

Chapter 2

Neutrinos

Neutrinos were created at the beginning of the Universe and they fill the entire space, similarly to the microwave background. Additionally, they are produced in the celestial bodies, the Earth's atmosphere and man-made machines such as nuclear reactors or accelerators. The neutrino cross sections are distinctively small and as a consequence they can travel long distances through matter without interacting, unlike other particles. Therefore, the message they carry contains important information about the sources that produce them, even the most distant. Furthermore, neutrino physics can be used to probe new theories beyond the Standard Model (SM) of particle physics. This can be achieved by examining neutrino phenomena such as the oscillation of neutrino flavours or the charge parity (CP) violating phase in the leptonic sector, which is thought to help explain the matter-antimatter asymmetry in the Universe. In this chapter, a brief introduction to the main interest topics for neutrino physics is given.

2.1 Neutrino Mass

The Standard Model (SM) theory was developed in the middle of the 20th century and it accommodates three types of the fundamental interactions: electromagnetic, weak and strong nuclear between the elementary particles, fermions and bosons. The SM includes 6 lepton fermion fields which have been arranged in SU(2) group doublets (left-handed) and singlets (right-handed) of definite handedness

$$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, e_R \qquad \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \mu_R \qquad \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}, \tau_R.$$

By construction the right handed neutrinos are not included in the SM, as only left-handed ν have been observed so far in nature and allowed by theory. In the SM neutrinos are massless, however neutrino masses can be introduced into the theory by the use of two mechanisms: the Dirac and Majorana mechanisms.

In the *Dirac mechanism*, the neutrino mass is introduced to the SM via the Higgs mechanism [13], [14], [15], identically to the quark and charged lepton mass origin process[1], where the right handed component is added to the SM Higgs-lepton Yukawa Lagrangian in the neutrino field element. This is called a Neutrino Minimal Standard Model (nuMSM). The right handed neutrinos have different properties to their left handed analogs, i.e. invariance under SM symmetry, hypercharge $Y = 0$, deficiency of quantum anomalies cancellation and interaction through the gravitational force only. Therefore, right handed neutrinos are called *sterile*. In the final outcome there are two spin states which can be doubled by an independent particle and antiparticle state.

In the *Majorana mechanism*, the neutrinos can be identified with its own antiparticle since they have no charge. Only neutral fermions can be described by the Majorana field due to the simplifications on the charge parity. In this case, two spin components are incorporated into one wave function. The left and right handed fermion fields are not independent and the Dirac equation for each type of handedness is the same equation written in two different ways. As a result, out of all fermions only neutrinos can be considered as Majorana particles.

The question about the neutrino's nature, Majorana or Dirac, is currently a subject of an intense experimental investigation. Both mechanisms diverge only for the case of massive neutrinos, which has been proven by the measurement and discovery of ν oscillation phenomenon that is described in the next section. However the kinematic effects for Dirac and Majorana neutrinos are the same and thus, as a consequence, can not be searched for through the neutrino oscillations directly.

The neutrino mass eigenvalues are significantly smaller than those of other elementary particles including the other fermions from the same generation. Though the origin of this phenomenon is still unknown, and not explained by either of the discussed mechanisms, there are existing theories which provide an explanation of the low value of neutrino masses, eg. the *see-saw mechanism* [17].

The absolute values of neutrino masses have not been measured yet, however the upper limits have been determined and are shown in Table 2.1. These are the

[1]The full mathematical formalism of neutrino mass introduction can be found in Ref. [16]. It is omitted due to its irrelevance to the main subject of this thesis.

direct measurements obtained by studying well understood decays of particles. The electron neutrino mass limit was determined using β-decay and gave the result of $m_{\nu_e} < 2$ eV at a 95% confidence level. The new KATRIN [18] experiment plans to start taking data soon in order to decrease this limit to 0.19 eV. For the case where electron neutrino mass would be above 0.35 eV, KATRIN can make a measurement with 5σ precision assuming 5 years of running. The muon neutrino mass was measured from charged pion decay to be $m_{\nu_\mu} < 190$ keV at a 90% confidence level. The lowest precision up to date has been achieved in assessing the tau neutrino mass. This was done by colliding e^+e^- which produces taus decaying to 5 hadron states and ν_τ. The tau neutrino mass is currently less than 18.2 MeV at a 95% confidence level. The difficulties in a direct ν mass

Neutrino Flavour	Upper Mass Limit	Experimental Method
$\overline{\nu}_e$	2 eV	Tritium
ν_μ	0.19 MeV	π decay at rest
ν_τ	18.2 MeV	τ decay

Table 2.1: Recent direct measurements of neutrino upper mass limits from the Particle Data Group (PDG) 2014 [10].

measurement lead to the alternative method based on the double beta decay studies, which concerns Majorana neutrinos. There are existing isotopes which are energetically allowed for the double beta decay according to

$$(Z, A) \rightarrow (Z + 2, A) + 2e^- + 2\overline{\nu}_e. \tag{2.1}$$

However, in the Majorana mechanism, where the neutrino is its own antiparticle, the two produced neutrinos can annihilate giving the final state of $0\nu\beta\beta$. A measurement of the half life of this process will allow the lightest mass eigenstate to be determined. There is a substantial number of such experiments and they aim to improve the lower bound of a ν mass measurement below 0.01 eV or measure the effective mass. An example of such experiment is the SuperNemo [19] experiment which, by design, has a sensitivity of an effective mass $m_{ee} < 0.07 - 0.12$ eV. Also the SNO+ [20] experiment has the ability to determine the neutrino mass limit via $0\nu\beta\beta$. The sensitivity of the mass depends on the radioactive addition load

to the scintillator in the detector tank. A 3% loading, which corresponds to 8 tons of 130 Te isotope in the SNO+ detector, would give the effective potential to probe the majority of this interesting mass range with high sensitivity including the region of the "inverted hierarchy". In order to determine the neutrino mass from the $0\nu\beta\beta$ experiments one has to combine the measurement with oscillation experiment results which are described in the following section.

2.2 Neutrino Oscillation Phenomenology

Neutrinos are unique particles, quantum mechanically speaking, as they are defined by a flavour state in contrary to the other Standard Model particles. The majority of elementary particles have a mass state simultaneous with their flavour state, which makes it straightforward to determine their mass from experiments. However, this cannot be applied to neutrinos, where their flavour state is a superposition of Hamiltonian states. In a vacuum these Hamiltonian states correspond to the neutrino mass states. This exceptional feature results in a neutrino flavour time, t, evolution equation given by

$$\mid \nu_\alpha (t) \rangle = \sum_k U^*_{\alpha k} e^{-i\lambda_k t} \mid \nu_k \rangle, \tag{2.2}$$

where $\mid \nu_\alpha \rangle$ are flavour eigenstates, λ_k is an k^{th} eigenvalue of the Hamiltonian eigenstate $\mid \nu_k \rangle$, and $U^*_{\alpha k}$ is a unitary mixing matrix. When this matrix is not diagonal, it means that neutrinos are mixed. Using the unitarity of the above relation, the massive states can be expressed by flavour states

$$\mid \nu_k \rangle = \sum_\alpha U_{k\alpha} \mid \nu_\alpha \rangle. \tag{2.3}$$

Hence, the superposition of massive neutrino states $\mid \nu_\alpha (t) \rangle$ can be expressed through pure flavour states $\mid \nu_\alpha \rangle$. Then the transition of $\nu_\alpha \rightarrow \nu_\beta$ amplitude in time t is given by

$$A_{\nu_\alpha \rightarrow \nu_\beta} (t) \equiv \langle \nu_\beta \mid \nu_\alpha (t) \rangle = \sum_k U^*_{\alpha k} U_{\beta k} e^{-i\lambda_k t}, \tag{2.4}$$

and the transition probability from one neutrino flavour to an other can be written as

$$P_{\nu_\alpha \rightarrow \nu_\beta} (t) = |A_{\nu_\alpha \rightarrow \nu_\beta} (t)|^2 = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} e^{-i(\lambda_k - \lambda_j)t}. \tag{2.5}$$

For ultrarelativistic neutrinos the following approximation can be used

$$\lambda_k - \lambda_j \simeq E_k - E_j \simeq \frac{\Delta m^2_{kj}}{2E}, \tag{2.6}$$

where Δm^2_{kj} is a square mass difference $(m^2_k - m^2_j)$, and $E = |\overrightarrow{p}|$. Since neutrinos propagate with almost the speed of light, the propagation time t can be approximated by the distance L between the neutrino source and detector. Finally, the oscillation formula can be written as

$$P_{\nu_\alpha \to \nu_\beta}(L, E) = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \exp\left(-i \frac{\Delta m^2_{kj} L}{2E}\right). \tag{2.7}$$

In terms of natural units the oscillation probability of one neutrino flavour into an other can be written as

$$P_{\nu_\alpha \to \nu_\beta}(L, E) = \sin^2 2\theta \sin^2\left(1.27 \Delta m^2_{ij} \frac{L}{E}\right), \tag{2.8}$$

and from there it is trivial to write a formula for the flavour surviving probability (disappearance)

$$P_{\nu_\alpha \to \nu_\alpha}(L, E) = 1 - P_{\nu_\alpha \to \nu_\beta}(L, E) = 1 - \sin^2 2\theta \sin^2\left(1.27 \Delta m^2_{ij} \frac{L}{E}\right), \tag{2.9}$$

where θ is a mixing angle between two flavour neutrino states, defining how different neutrino flavour states depend on the neutrino mass states. The relation between three neutrino mass states and three flavour states can be represented spatially as shown in Figure 2.1. The three neutrino flavour mixing matrix U

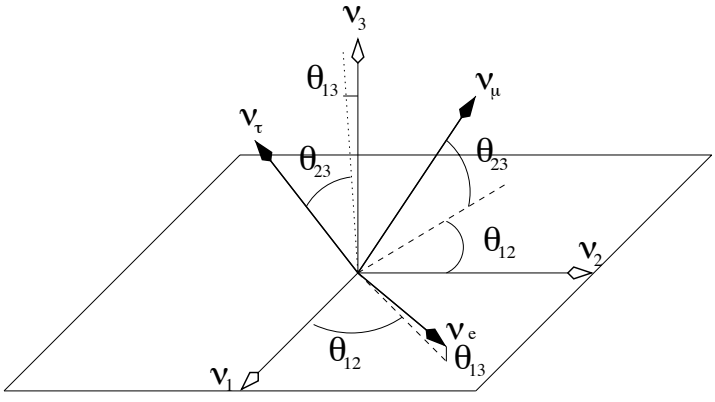


Figure 2.1: Spatial drawing demonstrating the orientation of the neutrino flavour axes (ν_e, ν_μ, ν_τ) with respect to neutrino mass axes (ν_1, ν_2, ν_3), related by the three mixing angles (θ_{12}, θ_{13}, θ_{23}) [1].

is parametrised by the three mixing angles (θ_{12}, θ_{13}, θ_{23}) and one phase (δ_{CP}) which is a CP violating phase. In the absence of interactions (vacuum) this is called the PMNS[2] matrix and is expressed by the oscillation parameters

$$
\begin{aligned}
U_{PMNS} &= R_{32}(\theta_{32})\, U_{\delta}^{\dagger}\, R_{13}(\theta_{13})\, U_{\delta}\, R_{21}(\theta_{21}) \qquad (2.10)\\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} c_{21}c_{13} & s_{21}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{21}c_{32} - c_{21}s_{32}s_{13}e^{i\delta_{CP}} & c_{21}c_{32} - s_{21}s_{32}s_{13}e^{i\delta_{CP}} & s_{32}c_{13} \\ s_{21}s_{32} - c_{21}c_{32}s_{13}e^{i\delta_{CP}} & -c_{21}s_{32} - s_{21}c_{32}s_{13}e^{i\delta_{CP}} & c_{32}c_{13} \end{pmatrix}.
\end{aligned}
$$

Here, $R_{jk}(\theta_{jk})$ describes a rotation in the jk-plane through angle θ_{jk}, $U_{\delta} = \mathrm{diag}(e^{i\delta_{CP}/2},\, 1,\, e^{-i\delta_{CP}/2})$, and $s_{jk} \equiv \sin\theta_{jk}$, $c_{jk} \equiv \cos\theta_{jk}$. Note, that the phase factor $e^{\pm i\delta_{CP}}$ only occurs in the PMNS matrix elements associated with $\sin\theta_{13}$. This means that δ_{CP} can be established experimentally only under the condition of non zero $\sin\theta_{13}$. The phase factor is associated with two discrete symmetries: charge conjugation C that changes particles to antiparticles and parity P which reverses the spatial components of the wavefunction. P flips the direction of the particles but not the spin, hence the particle handedness is transformed. Since CP symmetry changes the sign of the neutrinos and direction of the oscillation, the invariance under CP implies that

$$P_{\nu_\alpha \to \nu_\beta} = P_{\overline{\nu}_\alpha \to \overline{\nu}_\beta}. \qquad (2.11)$$

The inequality of the above probabilities, as a result of the phase δ_{CP}, has the potential to explain the matter-antimatter asymmetry in the lepton sector.

Neutrino oscillations are a quantum mechanical phenomenon that may be observed over large distances, even astronomically large. In such cases, neutrinos pass through a significant amount of medium, electrons and nucleons, which results in a modification of the oscillation probability. This is called the *matter effect* and it occurs due to the neutrino neutral current (NC) and charged current (CC) coherent forward elastic weak scattering. The NC scattering can appear for every flavour, however the CC scattering affects only the electron flavour as matter consists of e, but not μ and τ, see Figure 2.2. The matter effect arises in

[2]The neutrinos mixing matrix was first introduced by B. Pontecorvo, Z. Maki, M. Nakagawa and S. Sakata (PMNS matrix) in Ref. [21], [22], [23] and [24].

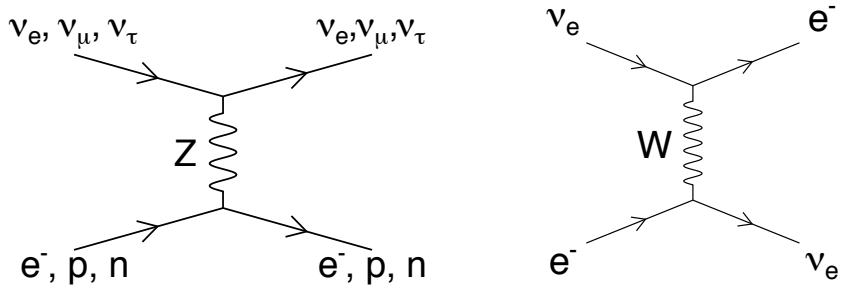


Figure 2.2: Feynman diagrams of the neutrino forward elastic scattering for the neutral current interaction (left) through Z exchange, and charged current interaction (right) through W exchange.

a mathematical formalism as an additional effective potential, V_α, in the vacuum Hamiltonian given by

$$V_\alpha = V_{CC}\delta_{\alpha e} + V_{NC}, \tag{2.12}$$

where the CC and NC potentials are equal to

$$V_{CC} = \sqrt{2}G_F N_e, \qquad V_{NC} = -\frac{1}{2}\sqrt{2}G_F N_n, \tag{2.13}$$

where G_F is a Fermi interaction constant, and $N_e(N_n)$ is the electron (neutron) number density in the medium. Neutrinos and anti-neutrinos experience opposite signs of the potential. This potential can be applied to the time evolution equation of the neutrino flavour state. As a result, the equation is composed with two separated terms: one term accommodates neutral current potential and second term consists of charged current potential. The term that includes V_{NC} is irrelevant for the neutrino flavour transition since it generates a phase common to all flavours and can be eliminated by the phase shift. Therefore, the time evolution leads to the new form of the probability of neutrino oscillation that depends only on V_{CC}. Considering the first order of the matter effect, the probability P_M of $\nu_e \to \nu_\mu$ can be written as

$$P_M = \sin^2 2\theta_M \sin^2\left(\Delta m^2_M \frac{L}{4E}\right), \tag{2.14}$$

where

$$\sin^2 2\theta_M \equiv \frac{\sin^2 2\theta}{\sin^2 2\theta + \left(\cos 2\theta - \frac{2EV_{CC}}{\Delta m^2}\right)^2}, \tag{2.15}$$

and

$$\Delta m^2_M \equiv \Delta m^2 \sqrt{\sin^2 2\theta + \left(\cos 2\theta - \frac{2EV_{CC}}{\Delta m^2}\right)^2}. \tag{2.16}$$

The resonance occurs when $2EV_{CC} = \Delta m^2$ and then mixing becomes maximal. This is known as the Mikheyev-Smirnov-Wolfenstein (MSW) effect [16].

2.3 Neutrino Oscillation Measurements

Neutrino oscillation experiments are called interference experiments and are designed to measure the oscillation parameters. To date, all three mixing angles have been deciphered from the ν oscillation data. Each of the angles, θ_{kj}, was measured from a type of neutrino experiment designed according to the nature of the neutrino source. Along with the angles, all the absolute values of square mass differences Δm^2_{kj} have been determined. The only oscillation parameter which is yet to be measured experimentally is δ_{CP} together with the sign of Δm^2_{23}. The summary of the most recent measurements is shown in Table 2.2.

PMNS parameter	Best Fit
Δm^2_{12}	$(7.53 \pm 0.18) \times 10^{-5}\text{eV}^2$
$\mid \Delta m^2_{23} \mid$	$(2.44 \pm 0.06) \times 10^{-3}\text{eV}^2$
$\sin^2 2\theta_{12}$	0.846 ± 0.021
$\sin^2 2\theta_{23}$	> 0.981
$\sin^2 2\theta_{13}$	0.093 ± 0.008
δ^{CP}	unknown

Table 2.2: Best fit values of the neutrino oscillation parameters as reported by the Particle Data Group (PDG) 2014 [10].

The θ_{12} and Δm^2_{12} parameters were originally measured using neutrinos produced in the Sun and therefore can be referred to as a "solar" mixing angle and "solar" mass difference. Here, electron neutrinos are produced in the nuclear fusion reaction

$$4p + 2e^- \rightarrow^2 He + 2\nu_e + 26.7\text{MeV}. \tag{2.17}$$

The Standard Solar Model (SSM) [25] predicts that the solar flux contains neutrinos with energies up to 30 MeV, which majority is below 0.4 MeV. Initially the solar ν flux observation was based on the detection of electron neutrino via CC interactions. The measurement was made by the Homestake [26] and Super-Kamiokande (SK) [27] experiments which observed a deficiency in the measured solar neutrino rate of about two thirds, which became known as the Solar Neutrino Problem. This problem was explained by the Sudbury Neutrino Observatory (SNO) [28] experiment that detected solar neutrinos via three different interactions: elastic scattering (ES), charged current and neutral current interactions. The latter reaction enabled SNO to measure the entire flux $\phi(\nu_e) + \phi(\nu_\mu) + \phi(\nu_\tau)$. The result was in very good agreement with the SSM prediction leading to the conclusion that the electron neutrinos were changing flavour on their way from the Sun. In addition, the solar oscillation parameters were determined from nuclear fusion reactors. The reactors are a powerful source of $\overline{\nu}_e$ with energies of order $\sim$MeV and a baseline of order $\sim$ 100 km, significantly shorter than in the Sun, thus the matter effect is negligible. An example of a reactor experiment is the KamLAND [29] neutrino experiment, which detects neutrinos from approximately 50 nuclear reactors distributed from a distance of 80 km to 800 km. The $\overline{\nu}_e$ flux is measured as a function of the energy and average baseline over all reactors. A combined average from KamLAND and solar neutrino experiments resulted in $\Delta m^2_{12} = 7.50^{+0.19}_{-0.20} \times 10^{-5}\text{eV}^2$ and $\sin^2 2\theta_{12} = 0.857^{+0.023}_{-0.025}$ ($\theta_{12} \approx 32^\circ$), which means that the solar mixing angle is not maximal.

The θ_{23} and $|\Delta m^2_{23}|$ parameters are called the "atmospheric" parameters as initially they were determined from the observation of the neutrinos produced by cosmic rays in the Earth's atmosphere. A proton entering the atmosphere produces a shower of hadrons which produce pions, which decay to produce neutrinos via the following reactions

$$\pi^+ \to \mu^+ + \nu_\mu \quad \Longrightarrow \quad \mu^+ \to e^+ + \nu_e + \overline{\nu}_\mu, \tag{2.18}$$

$$\pi^- \to \mu^- + \nu_\mu \quad \Longrightarrow \quad \mu^- \to e^- + \overline{\nu}_e + \nu_\mu. \tag{2.19}$$

The ratio consequential to the above reactions can be written as

$$R = \frac{\nu_\mu + \overline{\nu}_\mu}{\nu_e + \overline{\nu}_e}, \tag{2.20}$$

and is expected to be 2 at moderate energies (the spectrum peak is at 1 GeV).

However, the Super-Kamiokande experiment, along with other experiments, measured a lower value which was known as the Atmospheric Neutrino Anomaly. Additionally, SK analysed the direction of incoming neutrinos and observed a large deficit of up-going muons, produced by neutrinos with a longer flight length through the Earth, in comparison to down-going muons. This was explained in the context of neutrino oscillations hypothesis and was an evidence of ν_μ disappearance. Later, this measurement was confirmed by the accelerator generated neutrino experiments such as K2K [30] and MINOS [31]. In accelerator experiments neutrinos are produced via the same reactions as in the atmosphere, see Equations 2.18 and 2.19. The main difference is that the energy and kinematics of the particles within the focused beam can be controlled in the accelerator experiments. The typical energy of accelerator neutrinos is of the order $\sim$ GeV. According to Equation 2.9, the E/L ratio needs to be tuned to probe the required sensitivity to determine the oscillation parameters. Therefore, the desired baseline is typically from 10 km to 10000 km. The most recent measurement of the atmospheric parameters was performed by the T2K experiment [32], which provided the best fit values of $\sin^2\theta_{23} > 0.95$ and $\Delta m^2_{23} = 2.32^{+0.12}_{-0.08} \times 10^{-3}\text{eV}^2$ (for $m_1^2 < m_2^2 < m_3^2$). The sign of Δm^2_{23} has not yet been determined and therefore the hierarchy of the mass eigenstates is not known, see Figure 2.3. There are two

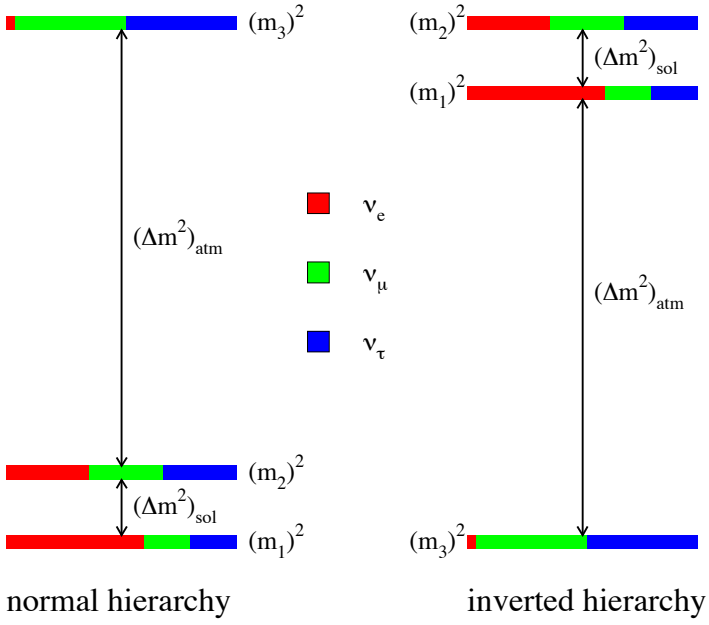


Figure 2.3: The neutrino square masses shown in a representation of the flavour states (colours). As oscillation experiments measure the absolute value of Δm^2, there are two possible orders of neutrino masses: $m_1^2 < m_2^2 < m_3^2$ called a normal hierarchy (left) or $m_3^2 < m_1^2 < m_2^2$ known as an inverted hierarchy (right) [2].

possible neutrino mass state arrangements with respect to its value. By convention, the configuration where ν_3 is the heaviest neutrino is called normal hierarchy (NH) and when ν_3 is the lightest it is called inverted hierarchy (IH). This hierarchy can be established using long baseline oscillation experiments with large matter effects as well as the double beta decay experiments described in section 2.1. The latter need to combine their measurement with oscillation experiment results to determine whether there are two heavy and one light, or one light and two heavy neutrino mass states.

The latest measured oscillation parameter is $\theta_{\mathbf{13}}$. An indication of a non zero value was first reported by the long baseline neutrino accelerator T2K experiment in 2011 [33] from the ν_e appearance channel. Next $\theta_{\mathbf{13}}$ was precisely determined in 2012 by the Daya Bay [34] and RENO [35] reactor experiments from the observation of $\overline{\nu}_e$ disappearance. In 2013, T2K measured θ_{13} to be $\sin^2 2\theta_{13} = 0.150^{+0.039}_{-0.034}$ with a 7.3σ C.L. [36] from the ν_e appearance channel.

It is difficult to measure a δ_{CP} phase as it requires the mixing angles to be determined with a high precision. Otherwise, the current uncertainties associated with the oscillation parameters measurement would cover the effects of δ_{CP}. However, the nearest future experiments have the relevant sensitivity and required accuracy to evaluate the phase factor based on the groundwork of current generation experiments. For example, the T2K experiment has collected $\sim 8\%$ data of design and has an ability to constrain the mixing angles values with much better precision.

To achieve the best possible accuracy of the neutrino oscillation parameters, a good knowledge of the neutrino interactions with matter is essential. The recent theoretical models of such interactions and measurements of neutrino cross sections are discussed in the next chapter.

Chapter 3

Neutrino-Nucleus Interactions

In this chapter, the neutrino interactions at the neutrino oscillation experiments are briefly described with emphasis on the neutral current single π^0 produced in the final state. Along with a more theoretical description of the process, the explanation of the neutral pion's role in the neutrino oscillation experiments is also given. Finally, recent results of the NCπ^0 cross section measurements from neutrino experiments are summarised.

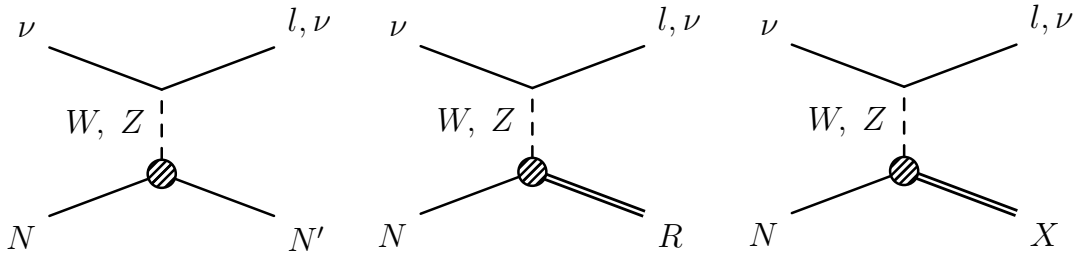
3.1 Neutrino Interactions in the Oscillation Experiments

In the Standard Model neutrinos interact only weakly, exchanging either a W boson in the charged current, or a Z boson in the neutral current interactions. The neutrino interaction cross section can be decomposed into

$$\sigma = \sigma^{CC} + \sigma^{NC}, \tag{3.1}$$

where σ^{CC} and σ^{NC} correspond to the CC and NC cross sections, respectively. When a neutrino scatters on the nucleon, one has to take into account the kinematics of the target inner structure. This structure and the neutrino energy determine the complexity of the interaction and as a result the scattering processes can be categorised. Therefore, for both CC and NC interactions, there are three processes: quasielastic (QES) or elastic (ES) scattering, resonance (RES) pion production and deep inelastic scattering (DIS). Feynman diagrams of these processes are shown in Figure 3.1.

In the neutrino energy region below 1 GeV, ES and QES are the dominating



(a) QES/ES (b) RES (c) DIS

Figure 3.1: The Feynman diagrams of neutrino-nucleon processes for charged and neutral current interactions.

processes. In this process the ν interacts with the whole nucleon from the target nucleus. In the higher energy region of a few GeV, the main process is RES where the nucleon is excited to the resonant state. This nucleon next decays into a variety of mesons. For an energy above 5 GeV, the dominant cross section is DIS where the neutrino scatters on the individual quark within nucleon, transferring energy to the nucleon. This leads to a hadronic shower in the final state. All these processes add up to the total cross section. Although there are theoretical models describing each type of interaction, there is no universal model integrating all the processes. The outline of all the processes is also summarised in Table 3.1 for CC and Table 3.2 for NC interactions.

Contemporary detectors are made of heavy nuclei material which increases the complexity of neutrino interactions. This is manifested by the nuclear effects which occur when the neutrino scatters on the nucleon bounded in the nucleus and can modify the final product results. Therefore, the impact of the nuclear effects on the cross section measurement can vary depending on the type of nuclei incorporated in the interaction with the neutrino. Both, the kinematics and interaction itself are affected mainly in the low energy region.

The current knowledge of theoretical models and recent measurements of neutrino cross sections are shown in Figure 3.2. From there, it can be seen that the cross sections are well measured at neutrino energies above 10 GeV but show large uncertainties below the 1 GeV region. The neutrino oscillation experimental energy ranges vary from $\sim$ 200 MeV to $\sim$ 10 GeV which results in different dominating processes for each experiment. Therefore, there is a number of neutrino generators which require different criteria to be satisfied for using different theoretical neutrino interaction models. As a result, the cross sections measurement depends on the type of nuclei and the theoretical model used in the Monte

CC Neutrino Interaction	Equation	Characteristics
Quasi-Elastic Scattering CCQE	$\nu_\mu + n \rightarrow \mu^- + p$	target changes but does not break up
Nuclear Resonances Production CCπ	$\nu_\mu + n \rightarrow \mu^- + p + \pi^0 (N^* or \Delta)$ $\nu_\mu + n \rightarrow \mu^- + n + \pi^+$	target goes to excited state
Deep Inelastic Scattering CC DIS	$\nu_\mu + quark \rightarrow \mu^- + quark'$	nucleon breaks up

Table 3.1: The neutrino - nucleon processes in the charged current interactions.

NC Neutrino Interaction	Equation	Characteristics
Elastic Scattering NCQE	$\nu_\mu + N \rightarrow \nu_\mu + N$	target unchanged
Nuclear Resonances Production NCπ	$\nu_\mu + N \rightarrow \nu_\mu + N + \pi (N^* or \Delta)$	target goes to excited state
Deep Inelastic Scattering NC DIS	$\nu_\mu + quark \rightarrow \nu_\mu + quark'$	nucleon breaks up

Table 3.2: The neutrino - nucleon processes in the neutral current interactions.

Carlo (MC) simulation. The long baseline oscillation experiments are composed of the far and near detectors. The latter is constructed to also address the shortage of neutrino interaction data, particularly in the energy region around 1 GeV. The T2K experiment has the neutrino energy peak around $\sim$ 0.6 GeV, where the main contribution to the cross section is dominated by quasi-elastic interactions and resonant processes with a pion in the final state, see Figure 3.3. Along with neutrino oscillation experiments there is a number of existing cross section dedicated experiments such as MINERνA [37] or ArgoNeuT [38], which study neutrino interactions and verify theoretical models at a wide range of energies. There is also a number of the planned experiments in the near future, eg. MicroBooNE [39].

From the experimental point of view, the CC interactions are more straightforward to be detected due to the charged lepton presence and neutrino absence

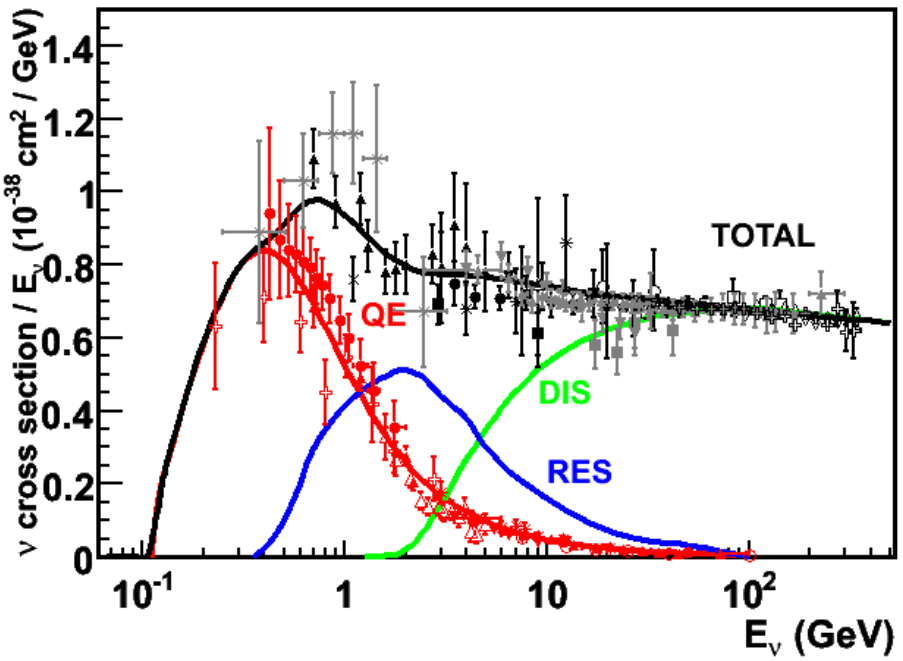


Figure 3.2: Summary of the recent knowledge of ν_μ charged current cross sections [3].

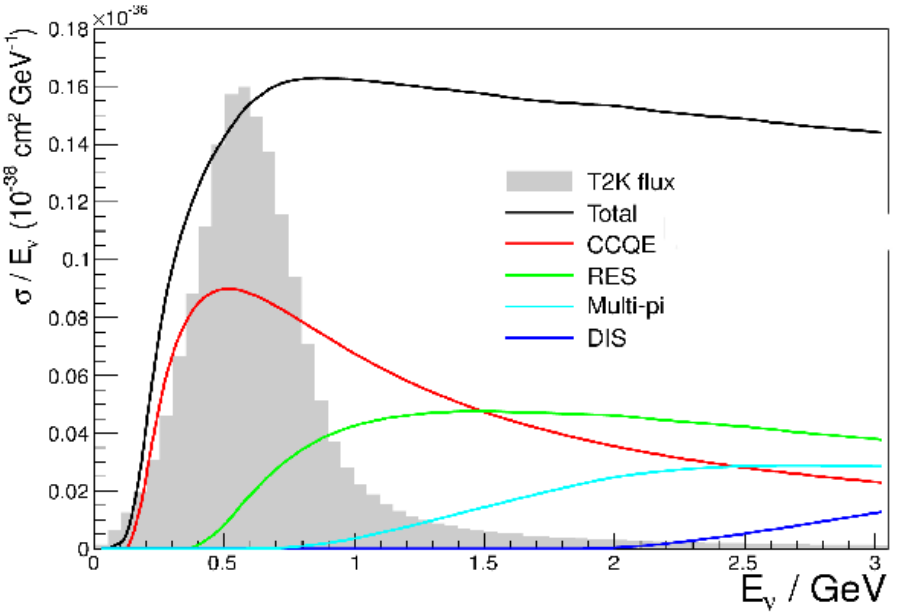


Figure 3.3: Summary of the recent knowledge of ν_μ charged current cross sections with the most recent addition from the T2K experiment [4]. Multi-pi interaction (blue line) is a type of resonance production as described in Table 3.1 and Table 3.2.

in the final state. This makes the energy reconstruction considerably less complicated in comparison to the NC interactions. Therefore, the CC interactions are used as a signal channel in the neutrino oscillation measurements. The neutral current interactions, with the emphasis on the π^0 production, are however extremely important as they constitute a major background to the electron neutrino appearance, as described in the next section.

3.2 NCπ^0 Production in Neutrino Oscillation Experiments

The neutral current π^0 production process is a main background in the Čerenkov detectors for both disappearance and appearance neutrino oscillations analysis. It is the least understood channel, hence the NCπ^0 cross section is a subject of ongoing and future measurements. The π^0 is a neutral particle and while produced in a neutral current interaction it a difficult object to detect. The flavour state of the π^0 in terms of quark states is given by

$$|\pi^0\rangle = \frac{1}{2}\left(|u\bar{u}\rangle - |d\bar{d}\rangle\right). \tag{3.2}$$

The decay product of a π^0 is constituted by two γ particles at 99% of the time, see Figure 3.4. The remaining fraction is a Dalitz decay mode $\pi^0 \to e^+e^-\gamma$. The π^0 decays quickly with mean lifetime $(8.4 \pm 0.6) \times 10^{-17}$ s [10]. The decay γs are

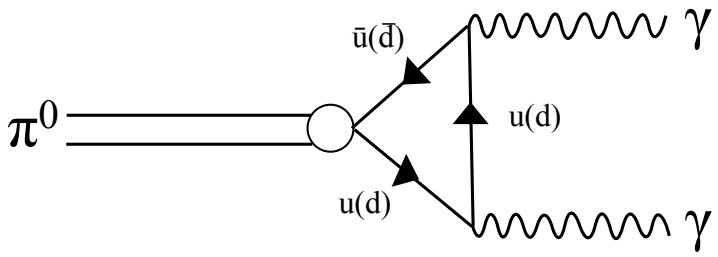


Figure 3.4: The leading order Feynman diagram for the neutral pion decay into two photons due to quark antiquark annihilation. The circle represents quark - antiquark pair in a bound state.

Lorentz boosted in the lab frame and can be produced with a various opening angle between them. This determines the energy difference of both photons. The γs can shower electromagnetically in the Čerenkov detector giving a two rings signature. However, the NCπ^0 can be misidentified with a CCQE ν_e reaction (signal

channel) in two situations. First, when decay gammas have highly asymmetric energies, the lower energy photon might be below the Čerenkov threshold and only one Čerenkov ring will be formed. This implies that the energy threshold for the π^0 detection has to be very low. Second case concerns a small opening angle between two decay gammas which leads to the Čerenkov rings overlapping. As a result one fuzzy ring is identified that mimics an electron neutrino signature. Together with CCπ^0 this is the second largest background for the ν_e appearance in oscillation experiments after the beam intrinsic ν_e induced events. Therefore, the direct measurement of NC1π^0 and determination of its absolute cross section are crucial for the future applications in a long baseline neutrino oscillation experiments. The π^0 can be produced through two dominating modes, coherent (COH) and resonant, which are described in the next section.

3.2.1 Resonant and Coherent π^0 Production

A single π^0 production in the energy region below 2 GeV, where the neutrino interacts within the nucleus (N), is dominated by a Δ *resonance reaction*

$$\nu + N \rightarrow \nu + \Delta, \tag{3.3}$$

$$\Delta \rightarrow N' + \pi^0, \tag{3.4}$$

where N' is a final state nucleus. It is an incoherent reaction where the nuclear target changes its state without breaking up. Resonant pion production mainly comes from the $\Delta\,(1232)$ with minor contributions from higher resonances and non resonant background. An example of baryon resonance models is the Rein and Sehgal model [40]. There are two main processes contributing to the resonant π^0 signal. The first one is a NC1π^0 as a primary interaction where final state interactions (FSI) do not affect the π^0 leaving the nucleus. The second is a $NC1\pi^+$ primary interaction where a π^+ experiences FSI charge exchange reaction and gets transformed into a π^0 which leaves the nucleus. There are several physical factors such as the Fermi motion, final state interactions, Pauli blocking and nuclear potential which cause difficulties in modelling the Δ production in a ν-nucleus interaction. The produced mesons and baryons can interact within the nucleus until they escape (FSI) which can change the number, momenta or direction of the outgoing particle. The presence of higher mass resonances and deep inelastic events are the reason for additional complications in the predictions of the baryon resonances. Understanding the influence of this effect on the production mechanism requires measurements of the emitted π^0s kinematics. The T2K

experiment has a good potential for improvement in understanding the neutrino resonant cross sections.

Coherent π^0 production happens when the neutrino interacts with the target nucleus that is left in its ground state which leads to an enhancement in the cross section. The outgoing particle does not experience any secondary effects from the nucleus. There is no isospin or charge exchange between the ν and the target nucleus. For the case of zero momentum transfer the νN cross section can be related to the πN cross section as

$$\left[\frac{d^3\sigma\left(\nu N \to \nu N \pi^0\right)}{dxdydt}\right]_{Q^2=0} = \frac{G^2 M E_\nu}{\pi^2}\frac{1}{2}f_\pi^2\left(1-y\right)\left[\frac{d\sigma\left(\pi N \to \pi N\right)}{dt}\right]_{yE_\nu=E_\pi}, \tag{3.5}$$

where G is the Fermi coupling constant, M is the nucleon mass, the standard scaling variables are $x = Q^2/2M\nu$ and $y = \nu/E_\nu$, ν is an energy of the hadronic system in the final state, $f_\pi = 0.93 m_\pi$ is the pion decay constant, the $t = p_T^2 = (q - P_\pi)^2$ variable quantifies the coherence (forwardness). In the coherent reaction case, where the momentum transfer is small, the produced pion is almost collinear with the incident neutrino. The coherent channel is still poorly understood and therefore the experiments often assume a high uncertainty on this cross section.

3.2.2 Current Measurements of NCπ^0 Cross Section

The most recent measurements on the NCπ^0 production are summarised in Table 3.3. For these measurements three different fluxes were used: J-PARC neutrino beam and Fermilab neutrino and antineutrino beams. Furthermore, the neutrino reactions took place on three different targets H_20, CH_2 and C_8H_8 [41].

In the resonance region, K2K and SciBooNE performed a NCπ^0 measurement as a ratio to the CC cross section ($\sigma_{NC1\pi^0}/\ \sigma_{CC}$). This means that the result describes a process with only one π^0 in a final state excluding all additional mesons. The SciBooNE signal definition includes at least one π^0 in the final state which means that the signal includes additional occurrences such as $1\pi^0 1\pi^\pm$ and more than 2π with more than one π^0 in the final state. The contamination fraction is estimated to be $\sim 14\%$. Additionally, MiniBooNE has measured the incoherent exclusive π^0 absolute cross section for the first time for the neutrino ν_μ flux with average energy ~ 1 GeV to be $5.71 \pm 0.08\,(\text{stat}) \pm 1.45\,(\text{sys}) \times 10^{-40}\,\text{cm}^2$/nucleon and for the $\overline{\nu}_\mu$ to be $1.18 \pm 0.07\,(\text{stat}) \pm 0.35\,(\text{sys}) \times 10^{-40}\text{cm}^2$/nucleon. In the coherent region, SciBooNE searched for the ratio of the COH NCπ^0 production to the total CC cross section and measured it to be $(1.16 \pm 0.24)\%$. Next, MiniBooNE has

Experiment	Detector Medium	Energy	Type	Measurement stat. ± syst.
K2K [42]	1kT Water	1.3 GeV	ratio to the total CC cross section	6.4×10^{-2} $\pm 0.2 \pm 0.7 \times 10^{-2}$
SciBooNE [43] [44]	Polysteryne C_8H_8	1.1 GeV	inclusive ratio	7.7×10^{-2} $\pm 0.5 \pm 0.5 \times 10^{-2}$
		0.8 GeV	coherent ratio	0.14×10^{-2} $\pm 0.3 \pm 0.5 \times 10^{-2}$
MiniBooNE [45]	Mineral oil CH_2	0.8 GeV	absolute inclusive	$4.76 \pm 0.05 \pm 0.76$ $\times 10^{-40}\,\mathrm{cm}^2$/nucleon
			absolute incoherent	$5.71 \pm 0.08 \pm 1.45$ $\times 10^{-40}\,\mathrm{cm}^2$/nucleon

Table 3.3: Summary of the recent NC1π^0 cross section measurements.

performed two measurements. The first measurement is an inclusive NCπ^0 ratio to all CC interaction at the average energy ~ 1 GeV which is found to be 19.5%. The second measurement is the overall flux averaged NC1π^0 cross sections, which for ν_μ interactions at mean energy 808 MeV is $(4.76 \pm 0.05\mathrm{stat} \pm 0.76\mathrm{syst}) \times 10^{-40}$ cm^2/nucleon and for $\overline{\nu}_\mu$ interactions at mean energy 664 MeV the cross section has value $(1.47 \pm 0.05\mathrm{stat} \pm 0.23\mathrm{syst}) \times 10^{-40}\,\mathrm{cm}^2$/nucleon. All the above measurements are complementary. The interpretation of each result strongly depends on the neutrino generator used in the data analysis.

In the near future, a NCπ^0 measurement is expected from the MINERνA experiment using a number of targets such as carbon, iron, lead and water. Their beam is planned to run at the low neutrino energy region ~ 3 GeV and medium ν energy ~ 6 GeV. In the T2K experiment the π^0 production is examined measured in the near detector ND280 on water and carbon. Considering the high intensity beam and the ability to separate CC from NC interactions and pion production processes, T2K is capable of contributing to the global data set. The T2K experimental setup is described in the next chapter.

Chapter 4

T2K Experimental Setup

This chapter describes the T2K experiment. First, the two facilities placed at the Japanese Proton Accelerator Research Complex (J-PARC), beamline and near detector ND280, are characterised. Next, the Super-Kamiokande far detector is described. Additional explanation is given to the description of the PØD and ECal sub-detectors that are used in this thesis analysis.

4.1 Introduction

The experimental setup was designed to meet the physics goals of the T2K experiment [8]. An accelerator generated neutrino super beam is used and is initially characterised by two detectors that form the near detector suite. This is followed by the beam detection at the Super-Kamiokande detector, located 295 km away from the source, see Figure 4.1.

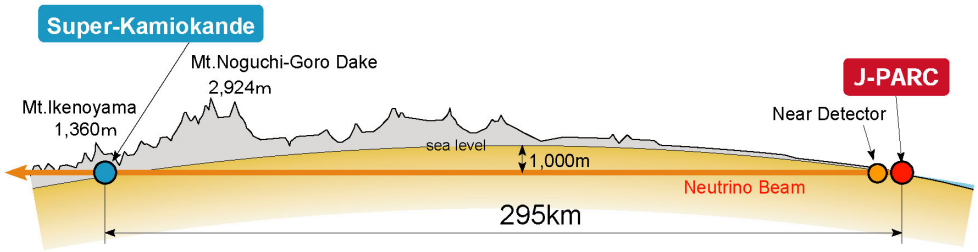


Figure 4.1: Schematic of the T2K experiment.

T2K adopts an off-axis method to generate a narrow band neutrino beam created using the new MW-class proton synchrotron at J-PARC. The neutrino beam is directed at an angle of 2.5° with respect to the baseline between the proton target and Super-Kamiokande, thus the neutrino beam at SK has an

energy peak at $\sim$ 0.6 GeV [5], see Figure 4.2. By selecting this energy and assuming the knowledge of the atmospheric parameters, the maximum probability for the disappearance oscillation at the T2K baseline is achieved whilst minimising the backgrounds.

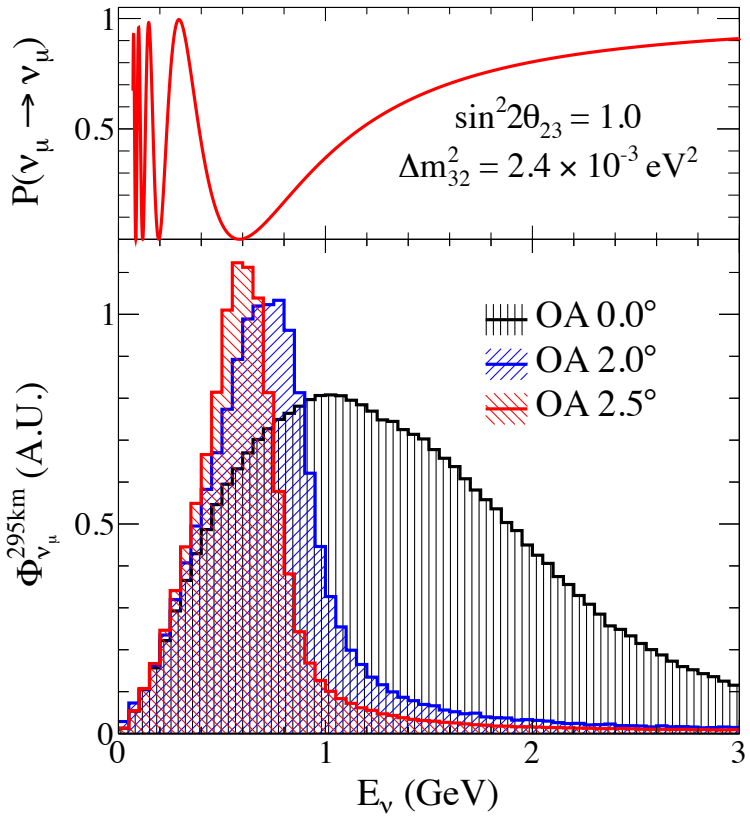


Figure 4.2: Muon neutrino survival probability at 295 km at 2.5° off-axis angle and neutrino fluxes for different off-axis angles [5].

4.2 J-PARC Neutrino Beamline

The beam is generated at the J-PARC facility situated on the East coast of Japan in Tokai, Ibaraki. It is a complex of 3 accelerators: the linear accelerator (LINAC), a rapid cycling synchrotron (RCS) and the proton synchrotron (PS), which is a main ring providing 30 GeV protons, see Figure 4.3. A linear accelerator accelerates a H^- beam up to 400 MeV. Next the beam is converted in the RCS to a H^+ beam and accelerated in the small ring to an energy of 3 GeV. The third section of the accelerator complex is a large proton ring which accelerates protons up to 30 GeV energy. The proton beam is delivered in spills, where each spill consists of 8 proton bunches. It enters the neutrino beamline which is di-

vided in two regions: primary and secondary. In the primary neutrino beamline,

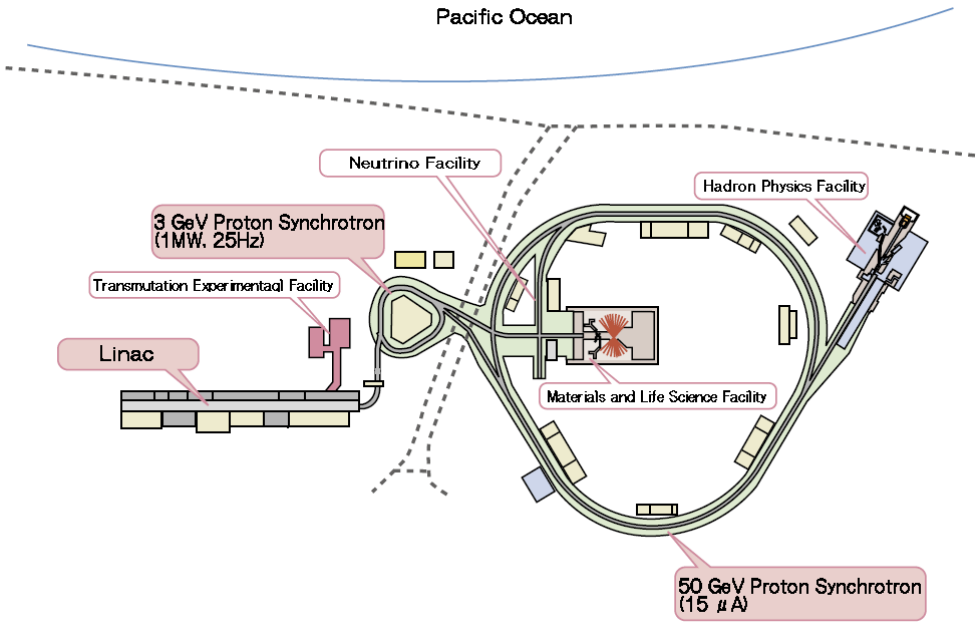


Figure 4.3: Schematic drawing of J-PARC.

the protons are directed towards Super–Kamiokande. This is achieved using a variety of magnets which steer the protons towards a graphite target. The protons are tuned in the preparation section. Their path is curved with a 104 m radius by 80.7° towards SK. In the final stage, the protons are focused into a narrow elliptical shape. The primary neutrino beamline uses a number of monitors to control each phase at the different operational sections. The precise description of the beamline can be found in [8]. At the secondary neutrino beamline, protons impinge on the target to produce charged pions with kaon contamination. Pions promptly decay producing a ν in the final state that constitutes the neutrino beam. The target is a graphite cylinder of 3 m diameter and 92 cm length, which accounts for around two interaction lengths. The targeting precision and proton beam direction is measured with a number of monitors. The charged pions from the proton interaction within the target are focused with a system of three coaxial magnetic horns powered with a 250 kA current. The first horn surrounds the target and is 2 m away from the second horn which is separated by 7.5 m space from the third horn. All horns have different radius, 0.4 m, 1 m, 1.4 m, and length, 1.2 m, 2 m, 2.5 m, starting with the most upstream. The aluminium

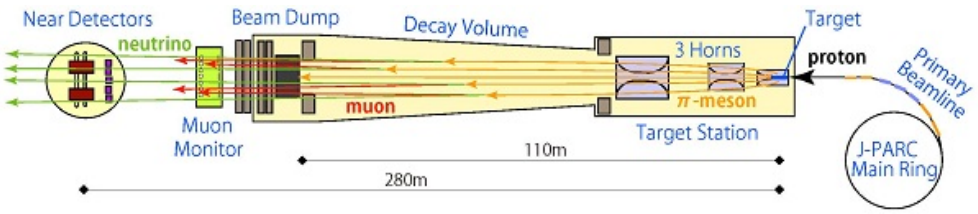


Figure 4.4: Neutrino J-PARC beamline facility.

horns produce a toroidal magnetic field which is precisely monitored due to its large effect on the neutrino beam flux.

Next, the produced and focused positively charged pions enter the 96 m long decay volume filled with helium. Pions decay mainly into muons and muon neutrinos ν_μ, which are the final product. This forms a muon neutrino beam contaminated with additional antineutrinos and electron neutrinos, which originate from kaon and other pion parents. All the hadrons and muons below 5 GeV/c are stopped by the beam dump placed 110 m downstream from the target. The beam dump is made of 75 t of graphite and absorbs the remaining pions and muons from the beam. Muons with momenta greater than 5 GeV/c penetrate the dump area and are measured with a muon monitor, located beyond the beam dump, in order to determine the beam profile.

The integrated number of neutrinos which pass through both the near detector ND280 and far detector Super-Kamiokande is proportional to the number of protons on target (POT). The generated beam power can be expressed by the statistical measure unit, POT, and beam energy

$$Beam\ Power = \frac{(POT\ in\ Spill) \times (Beam\ Energy)}{Spill\ Period}. \tag{4.1}$$

The source of neutrino flux uncertainties that are related to the beam direction and horn currents can be determined from the beamline and muon monitoring system. The expectation is to deliver 7.8×10^{21} protons on target in total for the T2K data physics analysis. The beam produced at J-PARC is characterised by a near suite of coarse fine grained detectors which aim to characterise the content and shape of the beam and estimate the systematic errors. A description of the near detector suite is given in the following section.

4.3 Near Detector Suite

The near detector suite plays a major role in the determination of the neutrino spectrum before oscillation and in the systematic error determination and background studies for the T2K analyses. It consists of two separate sets of sub–detectors: an on axis Interactive Neutrino Grid (INGRID) and an off axis near detector ND280 detector. Both detectors are placed in a pit 280 m away from the ν_μ beam source.

4.3.1 INGRID

The INGRID [6] is designed to measure the on-axis neutrino beam flux and profile. This sub-detector measures with high statistics the beam direction and beam intensity through neutrino interactions on iron. It contains 16 identical cubic modules spatially composed in the way where 14 modules form a cross shape and the remaining two are used as cross checks with addition of the Proton Module between these modules, as shown in Figure 4.5. The Proton Module

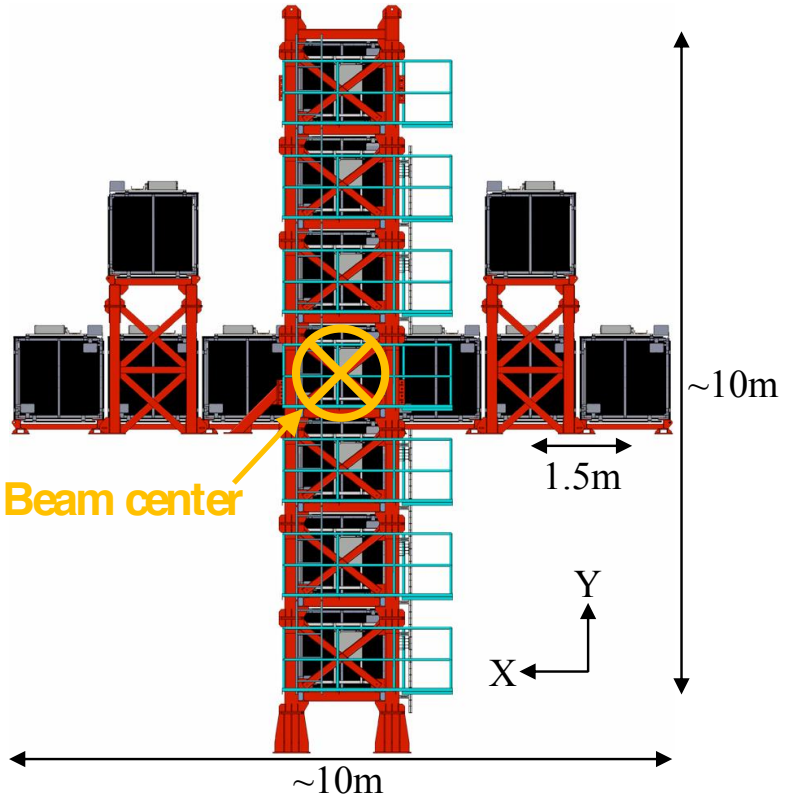


Figure 4.5: INGRID on-axis detector [6].

detects the muons together with the protons produces by the neutrino beam. Its goal is to identify CCQE channel for comparison with Monte Carlo simulations of beamline and neutrino interactions. The Proton Moduleis a fully active tracking detector located between the horizontal and vertical INGRID modules. It consists of 34 tracking planes surrounded by six veto planes, where each tracking plane is an array of 14 two types of scintillator bars. The tracking planes are therefore placed alternately in the x and y directions so that 3D tracks can be reconstructed. These tracking planes also serve as the neutrino interaction target as explained in Ref. [46]. In the INGRID architecture the two middle INGRID modules from the vertical and horizontal rows that overlap exactly in the beam centre position. The entire INGRID setup enables the analysis of the beam in a space 10 m by 10 m. Each of the modules has 11 scintillator tracking layers and 9 iron planes surrounded by 4 veto planes on the sides which helps to eliminate interactions from outside of the module. Each module contains 7.1 t total iron mass target. The tracking planes consist of 8,448 scintillator bars arranged in vertical and horizontal directions. INGRID has 9,592 channels in total, including the veto, which reveals when a cosmic ray passes into the primary detector, allowing the signal to be ignored ("vetoed") or recorded as a cosmic event. Each INGRID channel has a 3.2 ns time resolution. The calibration was performed with cosmic ray data. INGRID has the ability to measure the beam shape and direction with a precision of 0.4 mrad along with a 1 mm accuracy in the shift at the target. This means that the beam centre position is monitored with more than 10 cm precision. Figures 4.6 and 4.7 show T2K beam measurement made by INGRID during Run 5 using 1.45×10^{17} POT data. The beam centre in the x and y directions are measured on a monthly basis and demonstrate a stability within 22 cm.

4.3.2 ND280

The ND280 detector is located 2.5° off axis and is more complex with regard to its design purpose, which is a measurement of the relevant neutrino cross sections. ND280 is a tracking detector and has the ability to distinguish showers and tracks which are the signatures left by passing particles.

The ND280 is composed of a magnetised complex of fine grained sub - detectors as shown in Figure 4.8. With respect to the detector geometry, the direction of the beam is conventionally set as the z direction. Each individual sub detector output is integrated to the global reconstruction which can be viewed with an event display as shown in Figure 4.9. Each fundamental component of the ND280

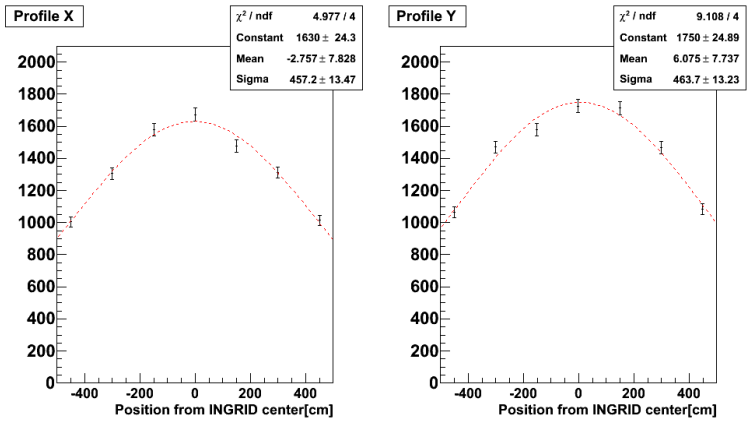


Figure 4.6: Neutrino beam profiles for x (left) and y (right) directions on 1st November 2014 [7].

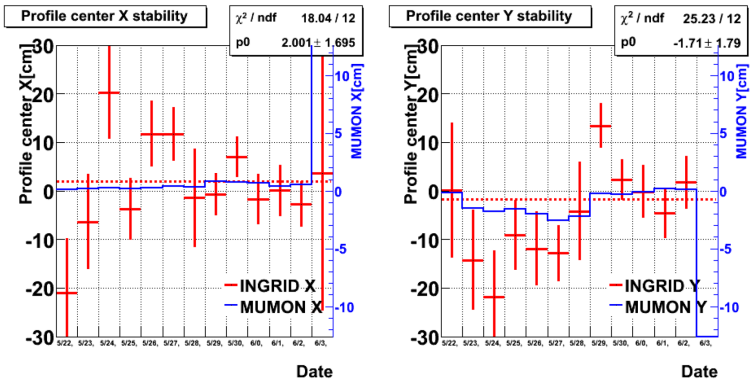


Figure 4.7: Measured neutrino beam centre position in Run 5 [7].

detector is briefly described in the following sections.

Magnet

The ND280 magnet originally comes from CERN, more specifically from the UA1 and NOMAD experiments. Its presence enables the measurement of the momenta of tracked particles with high precision and to determine the sign of their charge. The magnet provides a 0.2 T horizontal uniform dipole magnetic field from a power supply of 6 MW.

The block of magnet contains a set of 4 coils and flux return magnet yoke that is split into 16 C shaped parts made of low-carbon steel plates. The magnet has

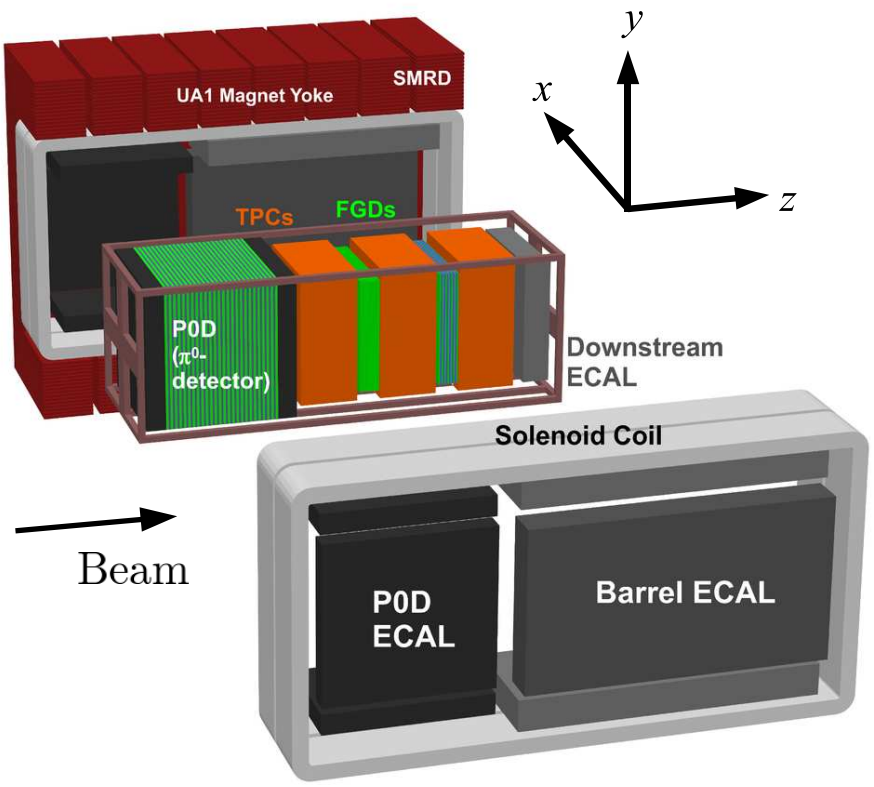


Figure 4.8: View of the ND280 sub-detectors with a global coordinate system convention [8].

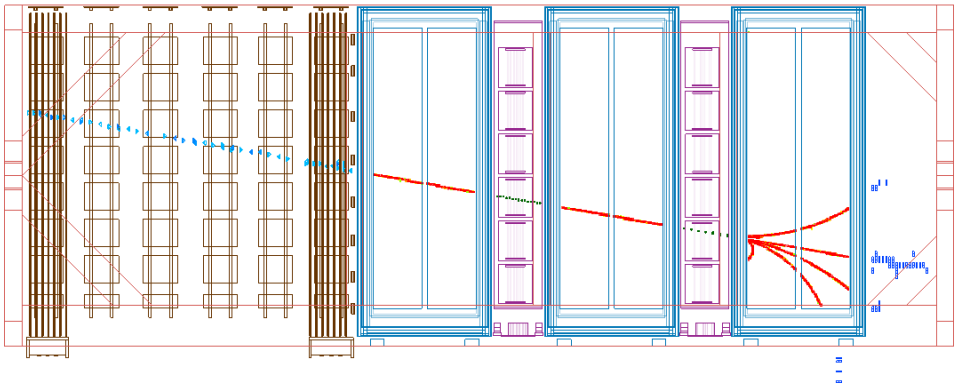


Figure 4.9: Event display showing a muon entering ND280 at the upstream PØD sub-detector and propagating to the tracker region producing secondary particles stopped in the ECAL modules.

an internal volume of 3.5 m × 3.6 m × 7.0 m where an aluminium basket containing other sub-detectors surrounded by electromagnetic calorimeters (ECals) is located. The total weight of the magnet is 850 tonnes. The coils are made of aluminium bars with 5.45 cm × 5.45 cm square cross sections, with a central 23 mm diameter slot for water to flow. The magnet is formed by two parts, right and left clams, which stand on movable carriages and can be opened independently to the sides with the split along z direction, allowing access to the basket.

4.3.3 SMRD

The Side Muon Range Detector (SMRD) [47] is located in the magnet yoke gaps and it measures muon momentum along with using the trigger of cosmic muons. It utilises the iron to measure the energy and direction of the muons produced at high angles to the neutrino beam direction in the CCQE interactions in the basket region.

The SMRD is a system of 440 scintillator paddles which are scintillator bars of dimensions: 0.7 cm thick, 16.7 cm (side segment) or 17.5 cm (top and bottom segments) wide and 87.5 cm length along beam direction (see Figure 4.10). The scintillator is extruded polystyrene and dimethylac-etamide with admixtures of POPOP and para-terphenyl, coated with titanium oxide. A wavelength shifting (WLS) fibre Y11 (1 mm diameter) is placed in the bar as shown in Figure 4.10.

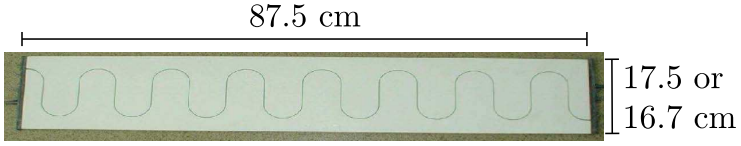


Figure 4.10: Picture of a single SMRD paddle with a visible WLS fibre.

The spatial distribution of the SMRD paddles in the air gaps between the layers of the magnet yoke is asymmetrical. There are more paddles in the downstream end of the magnet than in the upstream due to the greater flux at lower angles with respect to the beam direction. Individual gaps in the side magnet section are filled with five paddles and the top/bottom sections with four paddles.

4.3.4 Tracker Region

The tracker region is located downstream but within the ND280 basket. It consists of two sets of detectors: three Time Projection Chambers (TPC) and two

Fine Grained Detectors (FGD). The tracker performs the reconstruction of the interaction vertices and particle trajectories. It is designed to primarily study the neutrino charged current interactions.

FGD

The FGDs [48] make up the target mass for the neutrino interactions in the tracker region. They are designed to reconstruct the particle tracks and determine 3D vertices. The FGD modules are designed to co-operate with the Time Projection Chambers in order the fully reconstruct passing tracks from the multiple interactions.

Both FGDs are constructed from extruded polystyrene scintillator bars, doped with 1% PPO and 0.03% POPOP, with 1 mm diameter WLS fibre in the centre leading to the electronics. The bars are coated with titanium oxide, to reflect light inside the bars, and their dimensions are 0.96 cm wide, 0.96 cm deep and have a length of 1.8 m. The bars are arranged in layers in both the X and Y directions in the XY planes perpendicular to the beam direction.

Each FGD module contains 1.1 tonnes of target material. The most upstream module is built with 15 horizontal and 15 vertical layers. The second FGD has 7 vertical and 7 horizontal modules. The most downstream FGD has additional water target layers which allow the determination of both the carbon and water cross sections.

TPC

The TPCs [49] play a major role in the 3D charged particle track reconstruction without interfering with the particle path. Each TPC has an outer aluminium box 2.4 m high, 2.5 m wide and 1 m long covering the inner copper-clad G10 box as shown in Figure 4.11. It is filled with a mixture of gases $\mathrm{Ar}, \mathrm{CF}_4, \mathrm{C}_4\mathrm{H}_{10}$ in an active volume which is split in half by a central cathode plane with attached MICROMEGAS modules. Given a nominal voltage of 25 kV, the cathode provides an electric field of 250 V/cm parallel to the magnetic field. Each of these modules provides the signal caused by drifting charged particles with an average velocity around 75 mm/μs, which is ionising the gas. The timing of the received signal allows for the calculation of the particle track position. Figure 4.12 shows the tracking performance of the TPC. Particle identification in the TPC uses a measurements of energy loss of charged particles in the gas and the distributions of the energy loss as a function of the momentum are shown in Figure 4.13.

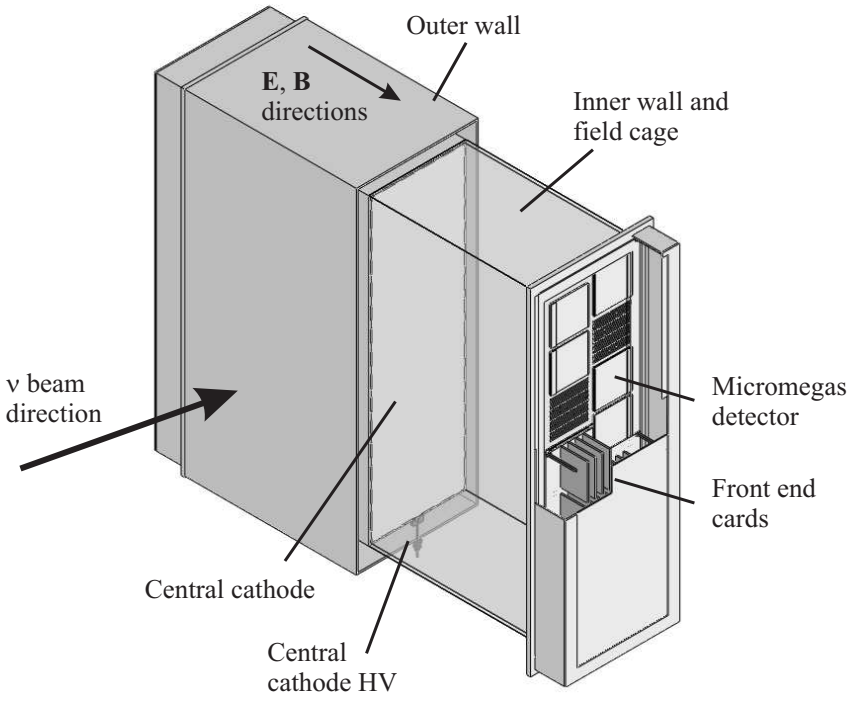


Figure 4.11: Schematic drawing of the TPC module.

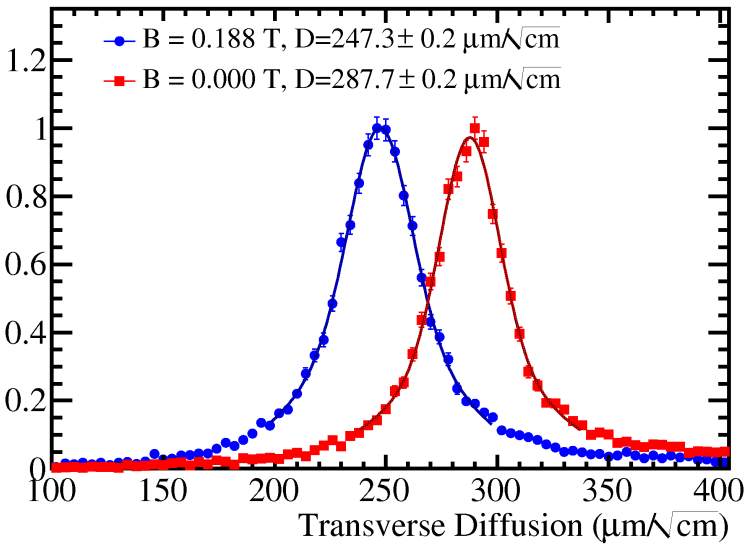
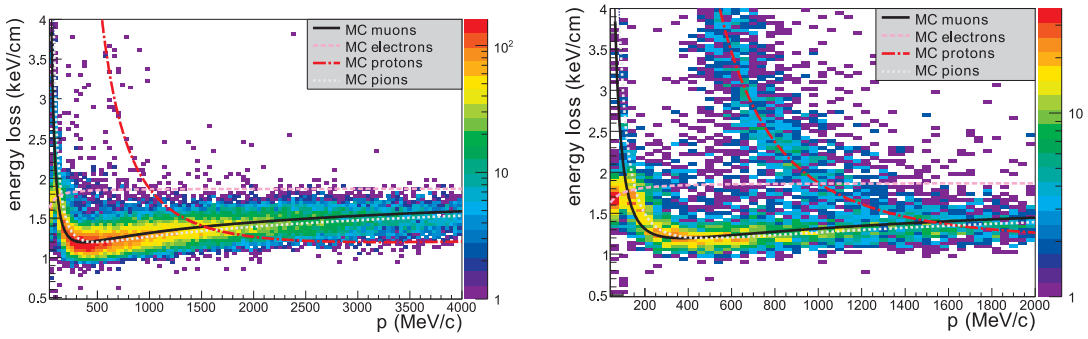


Figure 4.12: Distribution of the diffusion constant estimates from samples of cosmic rays with mean drift distance of more than 30 cm with magnetic field on and off. The quoted uncertainties are statistical only.



(a) MC (b) Data

Figure 4.13: Distribution of the energy loss as a function of the momentum for (a) negatively and (b) positively charged particles produced in neutrino interactions. Lines represent expected curves for muons, electrons, protons and pions.

4.3.5 PØD (π^0 sub-detector)

The PØD sub-detector [50] is located furthermost upstream in the ND280 basket. It is specifically designed to measure the π^0 produced in the neutral current process on water. The measurement is done by looking at the di-photon signal from its decay and therefore each individual PØD unit is optimized to detect the maximum photon energy.

The PØD is built as a sampling tracking calorimeter consisting of water target placed between two electromagnetic calorimeters (ECals). Its total mass is around 15 tonnes where the central region makes up a fiducial mass for π^0 measurements, of which $\sim$ 2 tonnes account for water and $\sim$ 3 tonnes for other material. The PØD consists of four Super-PØDules composed by 40 smaller units, XY plane layers called PØDules. A schematic design of the PØD sub-detector is shown in Figure 4.14. The PØDul layout is significant for a spatial resolution and pointing of showers direction. Each of these PØDule includes an active plastic scintillator made of triangular shaped bars arranged in two perpendicular planes, water target bags and lead and brass sheets. When a scintillating particle produces a signal in at least two bars the result has a better reconstruction resolution.

Scintillator planes have bars aligned along the X or Y direction. The triangular bars are 17 mm in height and 32.5 mm in base. Each bar has a WLS fibre, going trough the centre and along the bar, attached at the end to the MPPC. The PØDule contains two types of inactive layers: targets and radiators. The targets are semi-flexible pillow bladders of size $0.03 \times 1.8 \times 2.1\text{m}^3$ which contain $\sim$ 100 kg of water each. The PØD has the ability to run optionally with water pillows

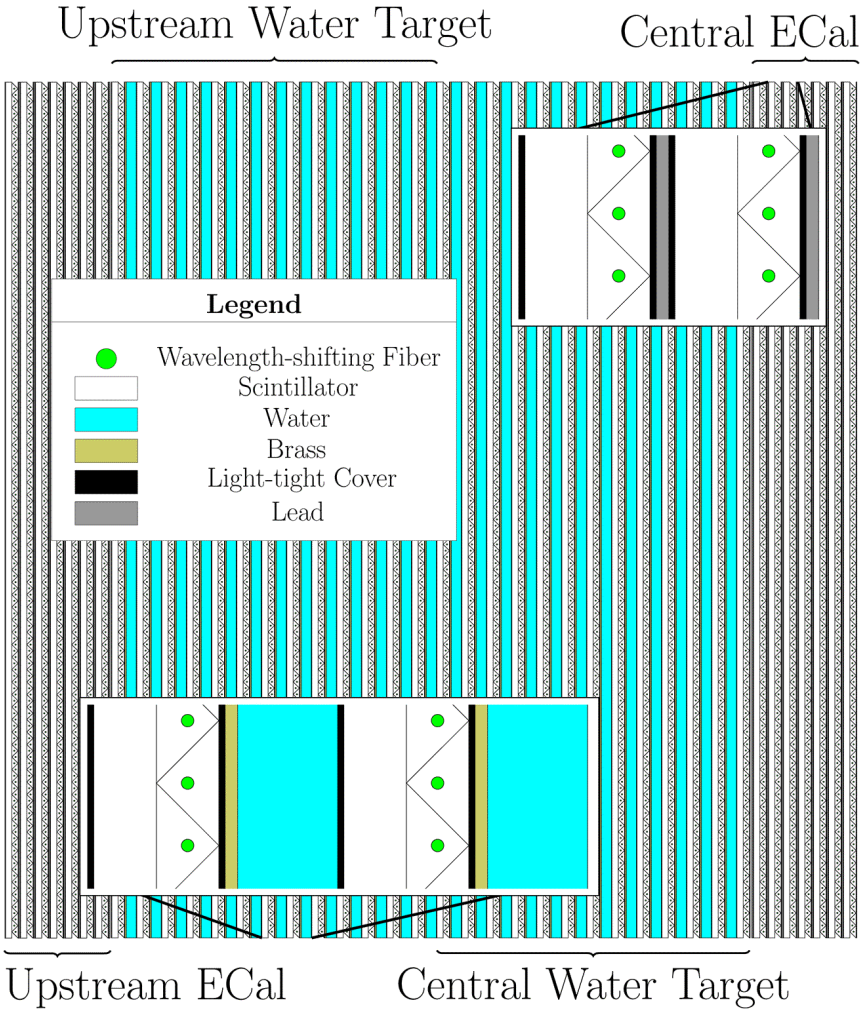


Figure 4.14: The schematic drawing of the PØD sub-detector.

filled or emptied. This will enable the neutrino cross section on oxygen to be determined which is required by Super-Kamiokande. The second type of inactive material is a radiator: brass and lead layers stimulate photons to convert and produce a shower in the detector. Lead layers stop particles more quickly due to a shorter radiation length than those built of brass. However, it also limits the information used by the PØD reconstruction modules and therefore the PØD has a multilayer structure composed accordingly to the following sequence: PØDule, 1.6 mm brass or lead radiator and water target, all repeated 25 times and finished with another PØDule.

The design is optimized to detect the maximum photon energy hence there are two types of ECals, one upstream (USECal) and one in the central region

(CECal). The lead layers have 4.0 mm thick bars in the USECal and 4.4 mm thick bars in the CECal which corresponds to 5.5 X_0 of lead. A limitation in the PØD ECals structure is in the bars layout as they are arranged in the XY plane only. Therefore photons that move in the direction perpendicular to the beam direction will escape the trigger.

To disentangle this situation the PØD is surrounded by the six PØD-ECal modules at four sides. The PØD sub-detector operation is supported by the PØD-ECal to ensure the detection of all the produced photons.

4.3.6 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECal) [51] surrounds the other ND280 sub detectors contained in the basket volume (PØD, TPC, FGD). The role of the ECal is to track down outgoing particles, produced in charged and neutral current interactions, which escaped from the detection in the basket region. For example, the ECal can select interactions producing a π^0 which decays to two photons by converting them into electromagnetic showers or catching partially escaping showers and providing information about their energy and kinematics. The ECal also aims to identify the charged particles passing through, i.e. muons, electrons and pions.

The ECal is the UK contribution to the ND280 detector where all modules were built. ECal modules consist of plastic scintillator (active material) and lead absorbers (dead component). Plastic scintillator layers are composed of bars made of polystyrene with 1% PPO and 0.03% POPOP addition. Each bar has a 4 cm $\times$1 cm area cross section with an elliptical hole going throughout the middle of the bar. Each layer surface is covered with 0.25 mm of TiO_2 in order to isolate and reflect the light. The readout of the signal is transferred with WLS fibres which run along the hole in the centre of each bar. The signal is detected by an MPPC secured on one or both ends of the WLS fibre. Plastic scintillator layers are alternated with lead layers.

There are 13 ECal modules inside the UA1 magnet, 6 modules surrounding the PØD on the top, bottom and side (PØD-ECals) parallel to the beam direction, 6 modules surrounding the tracker region (barrel ECals) with arrangement similar to the PØD-ECals and one module downstream within the tracker region located transversally with respect to the beam (DsECal).

Barrel and Downstream ECal

The barrel (BrECal) and downstream (DsECal) modules of the electromagnetic calorimeter are used for the energy measurements of the particles leaving the tracker region. Their aim is also to improve the systematics of the events that contain a FGD vertex. All 7 modules are composed of layers with bars aligned in planes with two different orientations to enable 3D track and shower reconstruction. A detailed composition of each module is shown in Table 4.1.

The DsEcal is a single module placed at the very end of the basket volume and was commissioned in 2009. The DsECal bars are 2 m long and are arranged in planes with xz and yz orientation. Prior to commission, the DsECal was tested at the CERN T9 PS with electron, muon and hadron beams.

ECal module	Number	Layers	Bars per Layer	Lead (mm)
Downstream	1	34	50	1.75
Barrel Top/Bottom	4	30	96/38	1.75
Barrel Side	2	30	96/38	1.75
PØD Top/Bottom	4	6	38	4
PØD Side	2	6	69	4

Table 4.1: Composition of the ECal modules.

PØD-ECal

The aim of the PØD-ECal is to complement the PØD reconstruction for escaping energy by catching partially contained PØD showers, identifying passing tracks, distinguishing minimally ionising particles (MIPs) and operating as a veto for surrounding backgrounds. The PØD-ECal has the ability to identify the MIPs and therefore can distinguish between CC and NC events. The PØD-ECal consists of 6 scintillating layers with an effective depth of $3.6X_0$ and 5 layers of lead. Bars are arranged into one direction only, parallel to the beam (see Table 4.1). This restricts the objects reconstruction to only 2D dimensions. The PØD-ECal modules were commissioned in Summer 2010 together with the BrECals.

4.3.7 Electronics

The INGRID, SMRD, PØD, FGD and all ECal modules have the same readout electronics system and therefore are called TripT detectors. The near detector complex uses a wavelength shifting fibre for readout. This is a 1 mm diameter Y11 Kuraray fiber connected to the photosensors. This photosensor is a multi-pixel avalanche photodiode (MPPC), customed by Hamamatsu. It is well suited to the WLS fibre emission spectrum and is insensitive to the magnetic field. The sensor is a compact device with a square matrix consisting of 667 sensitive pixels, each measuring 50 μm$\times$50 μm and acting as a Geiger micro counter (see Figure 4.15). The cells form 26×26 array with 3×3 pixel area missing in one corner. The light collection border size is 1.3 mm $\times$ 1.3 mm which is slightly wider than the WLS fibre to catch the light diffused at the fibre exiting.

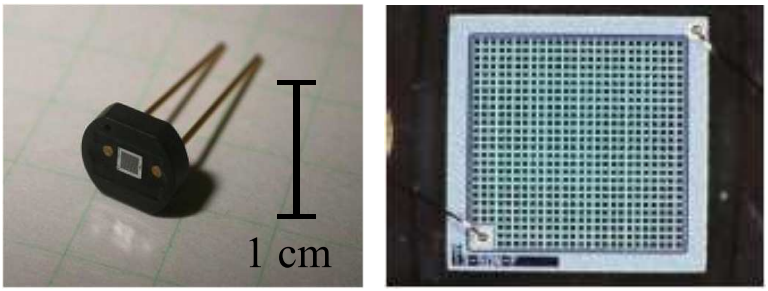


Figure 4.15: Photograph of the MPPC (left) and zoom to the sensitive detection area formed with 667 pixels in a 26×26 array (right).

The MPPC gain is determined by the charge accumulated in a pixel capacitance C_{pixel} : $Q_{pixel} = C_{pixel} \times \Delta V$ where ΔV is the difference between the applied and breakdown voltage of the photodiode with value $\sim$ 70 V. The pixel capacitance is 90 fF, which gives a gain in the range $0.5 - 1.5 \times 10^6$. The photodiode signal is a sum of the fired pixels. Each pixel is a binary device, but the multipixel photodiode sends an analogue signal.

There are three sources of noise coming from the MPPC. The first source is the dark noise caused by the thermal noise triggering avalanche. The dark noise depends on the temperature and its rate has a value around 500 Hz. The second type is called cross-talk as it makes the triggered avalanches crossing to the neighbouring cell for a new avalanche to be initialised. The last source of MPPC noise is the afterpulse occurring when the subsequently trapped electron causes more than one avalanche within the same pixel cell but is delayed in time.

The TripT Frontend Board (TFB) is used to readout the signal from the

MPPC. There are two channel types, low and high gain, to enable a wider range and higher readout precision. Each TFB is composed of four TripT chips, where each TripT has 16 low and 16 high gain channels to which the MPPCs are connected, producing analogue voltages. The signal is transferred to the digital converter which outputs 10-bit ADC counts. There are 23 capacitors used for charge integration in related time, each in 480 ns time window separated by 100 ns reset period.

4.4 Super-Kamiokande

Super-Kamiokande is the largest man made water Čerenkov detector in the world situated 2.5° off the beam axis and 295 km away from the T2K beam source in western Japan. It has been running since 1996 with several improvements and the current run period is called SK-IV. Super-Kamiokande is a 50 kt volume pure water tank of cylindrical shape, 42 m high and 36 m in diameter, located under 1000 m rock (see Figure 4.16).

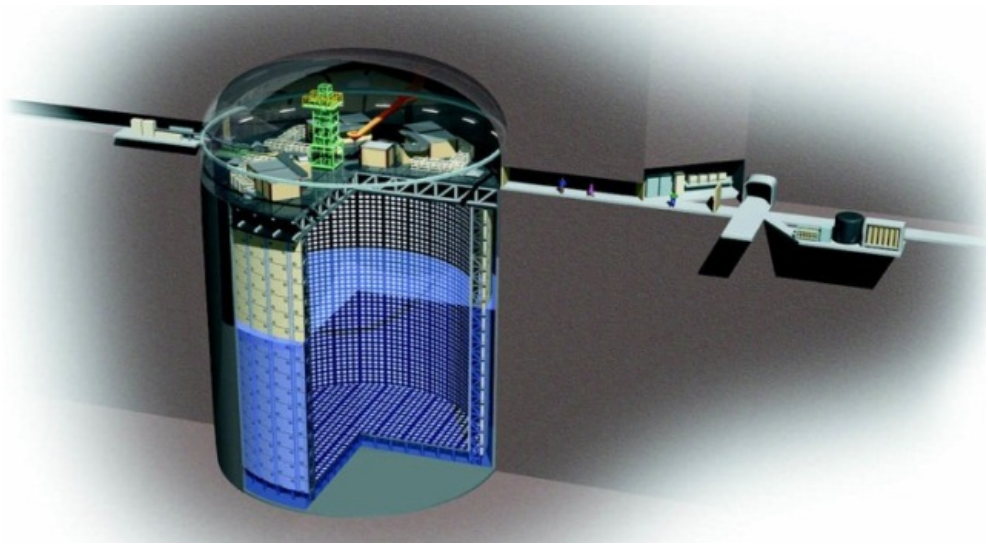


Figure 4.16: Drawing of Super-Kamiokande showing different composition elements and their placement in the mountain.

The detector uses the Čerenkov radiation phenomena to identify the neutrino interactions using 13,000 photomultiplier tubes (PMTs) which detect the light produced by the charged particles. It is divided into an inner detector (ID) and

an outer detector (OD) separated by a cylindrical stainless steel structure.

The ID cylindrical shape has a 33.8 m diameter and 36.2 m high. Its inner walls are covered with 11,129 inward-facing 50 cm diameter PMTs which accounts for 40 % of the inner detector surface. PMTs are able to operate with a sensitivity in the energy range from 4 MeV to 1 TeV. The detected signal by PMTs is used for the particle identification by looking at the Čerenkov ring shape.

The OD is optically separated and its role is to isolate all external events initiated outside the detector except for the highest energy cosmic muons. It has a cylindrical space about 2 m thick and its inner walls are covered with 1,885 outward-facing 20 cm diameter PMTs.

Data is read out using a GPS triggering to synchronise timing information. All hit information in a 1 ms window around the expected beam time is read out, with a reduction of the data happening in the offline analysis.

The operational performance of Super-Kamiokande is stable and both calibration and software are well understood with an accuracy to the percent level.

The T2K experimental setup described above has been designed to achieve a number of scientific goals. To meet all the objectives, a good performance of the beamline and all the detectors is crucial. The T2K experiment operation status during data taking is described in the next chapter.

Chapter 5

Data

This chapter summarises the beam data collected from March 2010 until June 2014. The broad characteristics of the data collected from November 2010 until April 2013 is given as it concerns the analysis presented in this thesis. First, a brief description of the beam data taking is presented. Next, the criteria of a good data selection is described in the data quality section. This is followed by a comprehensive description of the ECal data quality for the data used in the analysis carried out in this thesis. The ECal data quality assessment was performed by the author of this thesis.

5.1 Data Taking Summary

When T2K began taking data in November 2009, both near detectors were under completion with the last few designed modules to be installed. During the Summer of 2010, the detectors were fully commissioned: ND280 with the remaining ECals (PØD-ECal and barrel ECal modules) and INGRID with the two off-diagonal modules. T2K has been taking data with the full experimental setup since October 2010. Until September 2014 the T2K experiment accomplished five periods of data collection, during which all three detectors (ND280, INGRID and Super-Kamiokande) were taking beam and cosmic data. Table 5.1 shows the summary of the completed beam runs [52].

Due to unforeseen natural occurrences, i.e. the 2011 earthquake, data taking had to be suspended for the recovery period from March 2011 until January 2012. Run 3a was a test run taken after the recovery of the beam line and near detectors. These test runs were done with a partially unfocused beam due to one horn not

[1]The off-axis near detector ND280 took data from RUN31.

Period of data collection	Beam Run Number	Period	POT		Comments
			Delivered	Recorded	
Run 1[1]	29	23 Jan 2010 - 5 Feb 2010	0.400×10^{18}	-	ECal consisted of DsECal module only at ND280
	30	24.02.2010 - 28.02.2010	1.100×10^{18}	-	
	31	19.03.2010 - 25.03.2010	1.974×10^{18}	1.952×10^{18}	
	32	14.04.2010 - 1.05.2010	7.652×10^{18}	7.616×10^{18}	
	33	9.05.2010 - 1.06.2010	1.219×10^{19}	1.149×10^{19}	
	34	7.06.010 - 26.06.2010	9.304×10^{19}	8.797×10^{19}	
Run 2	36	18.11.2010 - 24.12.2010	4.188×10^{19}	4.035×10^{19}	Fully commissioned
	37	19.01.2011 - 28.01.2011	5.751×10^{19}	5.561×10^{19}	
	38	4.03.2011 - 11.03.2011	1.263×10^{19}	1.256×10^{19}	
Run 3	41	27.02.2012 - 23.03.2012	2.185×10^{19}	2.155×10^{19}	
	42	8.04.2012 - 25.05.2012	1.130×10^{19}	1.124×10^{19}	
	43	27.05.2012 - 9.06.2012	2.427×10^{19}	2.400×10^{19}	
Run 4	44	19.10.2012 - 12.11.2012	5.181×10^{19}	5.119×10^{19}	
	45	21.11.2012 - 12.12.2012	5.947×10^{19}	5.871×10^{19}	
	46	18.01.2013 - 22.02.2013	1.153×10^{20}	1.143×10^{20}	
	47	25.02.2013 - 1.04.2013	7.486×10^{19}	7.308×10^{19}	
	48	1.04.2013 - 12.04.2013	4.065×10^{19}	4.053×10^{19}	
	49	2.05.2013 - 8.05.2013	1.775×10^{19}	1.770×10^{19}	
Run 5	55	21.05.2014 - 03.06.2014	2.016×10^{19}	1.987×10^{19}	
	56 $\overline{\nu}_\mu$	4.06.2014 - 24.06.2014	5.087×10^{19}	4.898×10^{19}	
	56 ν_μ	24.06.2014 - 26.06.2014	4.206×10^{18}	4.178×10^{18}	

Table 5.1: Summary of the beam data taking runs at the T2K experiment near detector ND280 from March 2010 until June 2014. During the neutrino run (that is a sub-sample of the Run data collection periods) RUN56, an antineutrino beam was delivered for the first time.

powered on. The horn power supply device was under repair until February 2012.

According to the T2K 5 year data taking plan, the collaboration aims to collect 5×10^{21} POT assuming 750 kW $\times 10^7$ s or 10^{21} POT $\times$ 50 GeV/30 GeV per one year beam delivery. This plan includes annual Summer shut-downs and seasonal engineering works for the beam line tuning and improvements. Until June 2014, the maximum achieved value of the beam power was 270 kW with 30 GeV protons. The delivered number of protons on target for Runs 1-5 is shown in Figure 5.1. As of writing of this thesis, Run 6 is undergoing processing through the software chain.

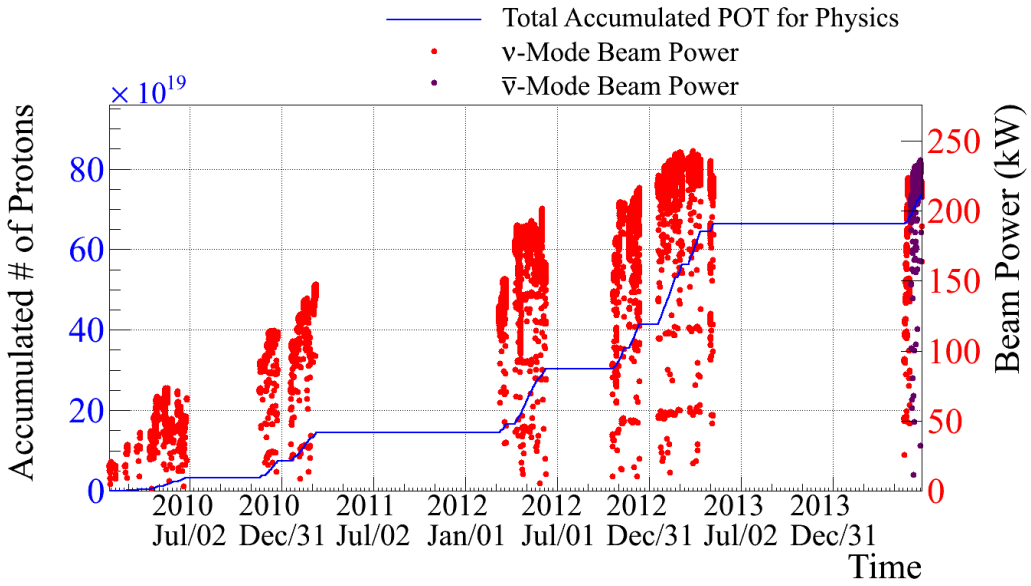


Figure 5.1: Delivered number of protons on target for Runs 1-5 (blue) and corresponding beam power (red).

5.2 Data Quality

Collected data used in physics analyses must be deemed to be of good quality. This requirement accounts for two factors: the beam status and detector condition which are explained in this section. Only the beam provided in physics runs will be considered for the analysis studies. A physics run excludes spills which were generated during beam tuning or monitor studies. Regarding the detector physical state, the data quality group provides a flag associated with the quality of data per each time interval for each detector and sub-detector.

Beam Quality

The quality of the beam delivered in a physics run is checked for Super-Kamiokande and the near detector suite by the application of selection criteria. In order to get good spills, the following cuts are applied to the near off-axis detector, ND280:

1. **Physics run:** it excludes beam tuning and monitoring studies spills.
2. **Beam trigger:** it eliminates test dummy spills.
3. **Good GPS status:** the difference between GPS1 and GPS2 has to be less than 200 ns at the J-PARC site.
4. **Spill flag:** it excludes empty spills.
5. **Nominal beam condition:** it requires a good condition for all beam elements.
6. **Horn current:** each horn current variation has to be within ± 5 kA from its nominal value.
7. **Muon measurement:**
    a. the muon profile centre has to be within ± 10 cm in x and y direction.
    b. the deviation of the charge of the total number of muons collected, normalised by the POT per spill, has to be less than 5%.

The average spill number loss caused by the status of the beam line components is 0.7%. The good spill selection used for the Super-Kamiokande data analysis is based on its DAQ and can be found in Ref. [53].

Collected Data Quality

Each of the three T2K detectors has a separate strategy for data quality evaluation. The Super-Kamiokande detector data quality criteria description can be found in a technical note [53]. The data quality strategy for both near detectors is described in detail in Ref. [54] while the data quality assessment for Run 1,Run 2, Run 3 and Run 4 is presented in the technical notes [55], [9], [56] and [57], accordingly. This strategy was applied to the beam data starting from the neutrino RUN34, see Table 5.1 for further information. The data quality objectives include checking the detector hardware status, checking low level variables and providing a status update for each sub-detector. For this purpose a specific software framework for the data quality analysis was developed. The flow chart of this

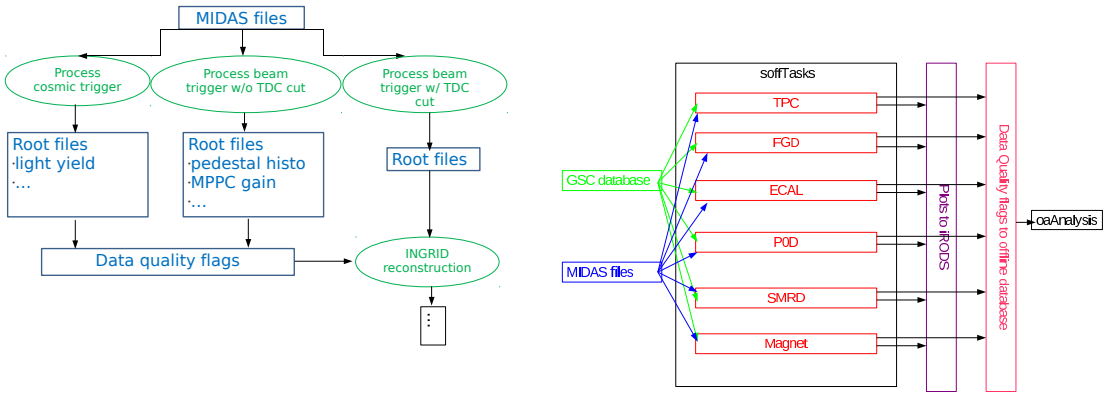


Figure 5.2: Data Quality framework for the (left) INGRID and (right) ND280 detectors.

framework is shown in Fig 5.2. Data taken by both near detectors are written on one of the three semi-offline machines located in the pit. Regarding the ND280 detector, three phases of the data quality framework can be distinguished.

First, a global slow control (GSC) database is used for monitoring the hardware status. Second, the data are processed with the calibration and reconstruction packages to estimate the quality of the physical variables by assigning flags. This is partially done with the software package soffTasks that runs on the semi-offline machines. This package performs the conversion of the data from a MIDAS format to a ROOT format using the RECPACK package, and hosts the data quality framework analysis on the J-PARC cluster. It can run jobs specific for the different sub-detectors depending on a given trigger type in data. At the end, the summary plots and tables of the output of submitted jobs are produced. Figure 5.3 shows an example of the number of completed jobs. The data quality jobs are based on the ND280 software library using the most recent stable release at the time.

At this point each sub-detector performance is represented by a flag which indicates per each time interval whether the data has been assessed good or bad. All the flags of the ND280 sub-detectors used in the analysis are required to have a good value in order to use the off-axis detector data for the analysis. In the next stage, the flags are written to the off-line database. This is done using the oaDataQuality software package which creates data quality tables, loads the flags and queries them when the analysis ntuples are created. The tables contain information about the flags as a function of the Unix time, the person who assessed the data, the flag description and the software version used. The

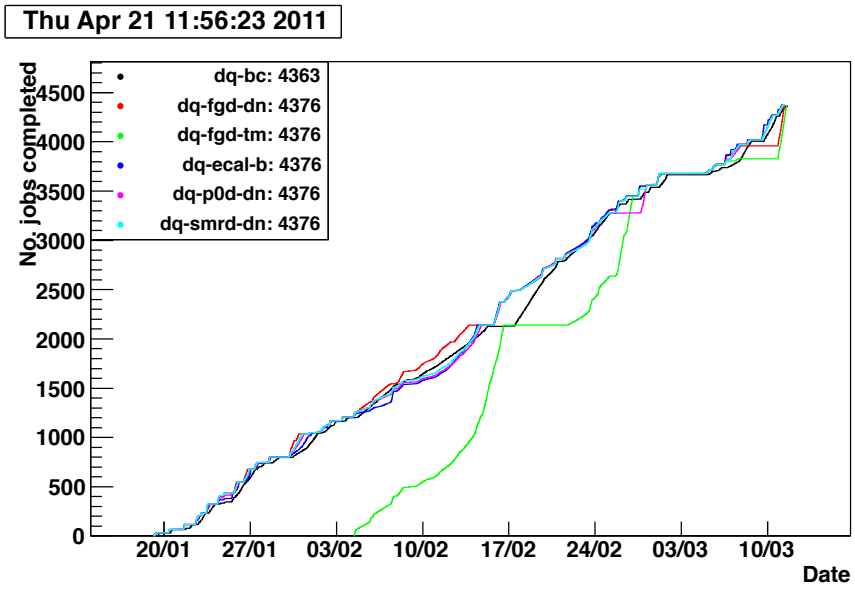


Figure 5.3: Completed jobs history submitted by the semi-offline tasks package for neutrino runs RUN37 and RUN38.

data quality files created during the assessment are stored and accessed remotely with the data grid software called the Integrated Rule-Oriented Data System (iRODS) [58]. Copying files from the semi-offline machines to iRODS is managed by the oaDataQuality package.

The strategy to assess the recorded data by each sub-detector can be found in the referred technical notes. This includes the ECal data quality on which the author of this thesis worked on is described in detail in the next section. The INGRID data quality assessment can also be found in the referred documentation.

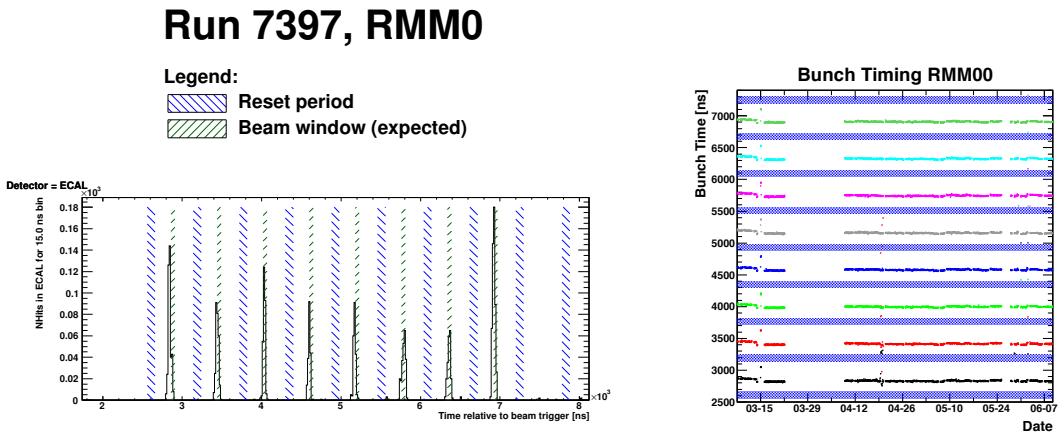
5.3 ECal Data Quality

The ECal data quality assessment evolved in time during the data collection from a manual to a more automated system. The ECal DQ concerns the analysis of raw data as well as visual monitoring of the global slow control. The latter was automated during Run 3. The ECal was fully commissioned prior to the start of RUN34. The readout is done by 12 RRMs, which are connected to a total of 366 TFBs. The ECal data quality checks are reported on a weekly basis during the data taking. All the monitored quantities for different runs are described below. Parts of the electronics readout became temporary or permanently inoperable: RMM5 was down during Run 2 until 22 November 2010 due to a PS trouble. Additionally, TFB13 and TFB25 on RMM9 went completely silent after

the shutdown due to the March 11th 2011 earthquake.

5.3.1 Beam Timing

The beam timing monitors the bunch position within the TFB readout cycle and shows the hit time relative to the beam trigger time. While the beam is configured to have 8 bunches per spill, the trigger offsets are set in such a way that the first bunch appears in the centre of the 5th cycle. There are 23 cycles and each of them has a 480 ns time window separated by a 100 ns reset period. The beam hits appear in the centre of the readout window as shown in Figure 5.4. According to the beam group report, the beam timing fluctuations are expected to be up to 100 ns, and the observed variation is up to 40 ns.



(a) (b)

Figure 5.4: Bunch timing separated by the 100 ns reset period (blue dashed line). (a) Bunch timing for RMM0 during Run 2 period. Green dashed line represents beam trigger time. (b) Hit time relative to the beam trigger time for ECal hits for RMM0 during Run 3.

5.3.2 Dead and Drifting Channels

In order to find dead or bad channels, the DPT histograms are analysed, see Figure 5.5. The DPT histogram represents the ADC spectra generated for each MPPC in the detector, where the nominal channel pedestal value is subtracted from each spectrum. These histograms are an input to two software packages: mppcCalib and tfbCalib. The gain of an individual channel is calculated for

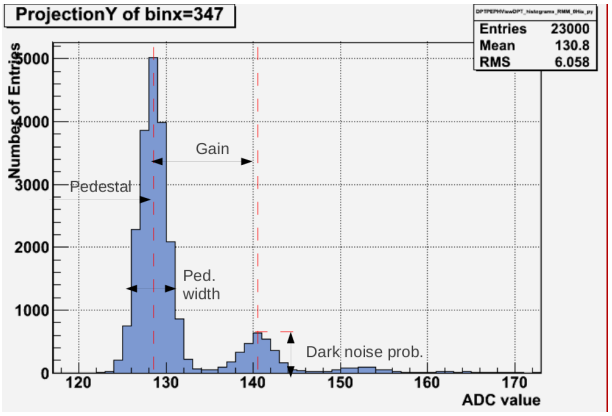


Figure 5.5: One channel DPT histogram.

each sub-run using the DPT histograms. Every DPT histogram accumulates 500 events for both pedestal and beam data. A bad channel is defined as a channel where the PE peak can not be found by the fitting algorithm but it can be recovered in time. A dead channel is a channel in which the PE peak can not be found by the fitting algorithm at any time. Figure 5.6 shows the number of dead and bad channels in Run 2. The full list of dead channels can be found in

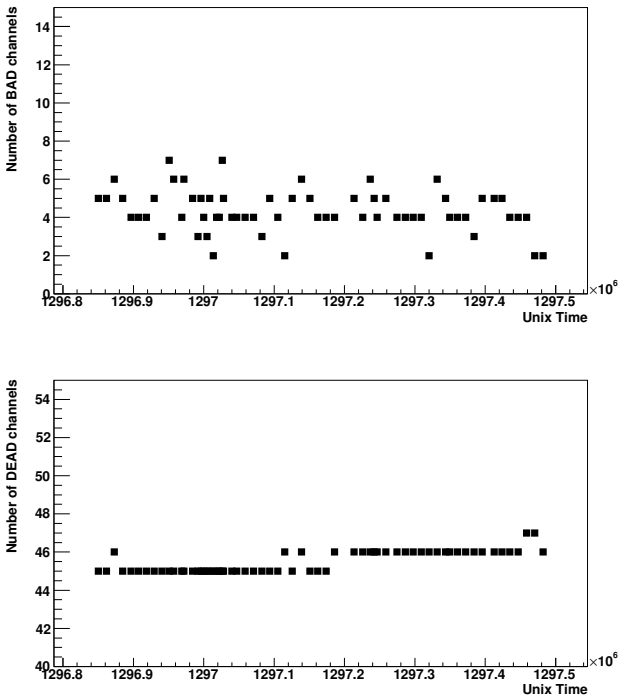


Figure 5.6: An example of dead (top) and bad (bottom) channels check during Run 2 data taking.

the above referred DQ technical notes. At the end of Run 4 there are 53 dead channels in total, and taking into account the silent[1] TFBs this increases to a total of 181 dead channels. Any sudden significant increase of a drifting or dead channels number (more than 10) is a subject to investigation and affects the quality of data.

5.3.3 Gain Variation

The gain (in ADC) of each ECal channel (MPPC) is calculated using the tfbCalib and mppcCalib packages. The data quality group monitors the gain variation to control its stability over time. Instability can affect the efficiency and resolution of the detector. The first sub-run gain values are used as a reference in a given data run. Each RMM is analysed separately every 3 hours, see Figure 5.7.

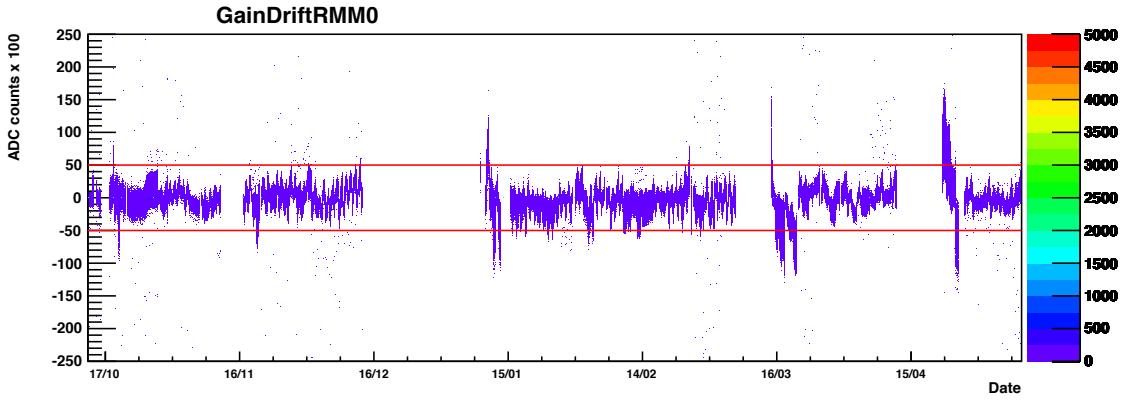


Figure 5.7: Gain variation versus time during Run 4 for RMM0. Red lines correspond to the 0.5 ADC interval limit.

5.3.4 Pedestal Variation

The electronics pedestal's stability is important for the resolution of the detector. Instability can appear due to temperature variations or voltage supply issues. Therefore, it is important to monitor all changes on a regular basis. The pedestal variation is calculated from the ADC counts in the channel without any light signal for both high and low voltage supplied electronics. Figure 5.8 shows the pedestal variation of the RMM0 calculated with reference to the first file from the beginning of a given run period. During Run 2 a pedestal's extensive instability appeared on the RMM4 TFB43. The observed fluctuations were caused by the

[1]Silent channel (TFB) is uninstrumented or without any signal.

2.5 V line on the TFB itself, see Figure 5.9. These fluctuations can be calibrated out, hence they do not affect the data quality assessment.

5.3.5 Trip-T Occupancy

The TripT chip status is controlled by monitoring the validity of the beam triggered hit time stamps. The number of chips without a valid timestamp is calculated for every 10 spills. Figure 5.10 shows the TripT occupancy during Run 3, where up to three Trip-T chips were recording hits without a valid timestamp at a given time unit, aside from the 8 known chips placed on the 2 dead TFBs. This demonstrates that the frontend board electronics is working accurately. Any sudden significant increase of number of invalid Trip-T chips performance is a subject to investigation and affects the quality of data.

Cluster Ratio

The tracker ECal and DsECal modules are checked for the neutrino event rate. This is done using the ecalRecon package which provides objects created by the clustering algorithm. The selected hits from the beam spills are clustered and compared to the average rate value as a ratio of the recorded POT in the 24 hour time unit, see Figure 5.11.

5.3.6 Cosmics

Cosmics data are recorded either as inter-spill or dedicated standalone runs. This data is used to check the ECal status by making hit maps and calculating the RMM offsets. Initially, cosmic muon tracks are reconstructed by a dedicated reconstruction algorithm called the Simple Track Fitter. Next, their position is extrapolated to the inner and outer face of the ECal modules. The final output is used to perform the following checks.

Hit Maps

The reconstructed cosmic muon track hit position for the DsECal and BrECal are shown in Figure 5.12 and 5.13, respectively. Looking at the hit maps there is a visible empty area caused by the broken TFB25 on RMM9. Also there is a notable scattering of white mini dots which represents a distribution of dead channels. At the time of Run 4 these checks were not performed for the PØD-ECal modules.

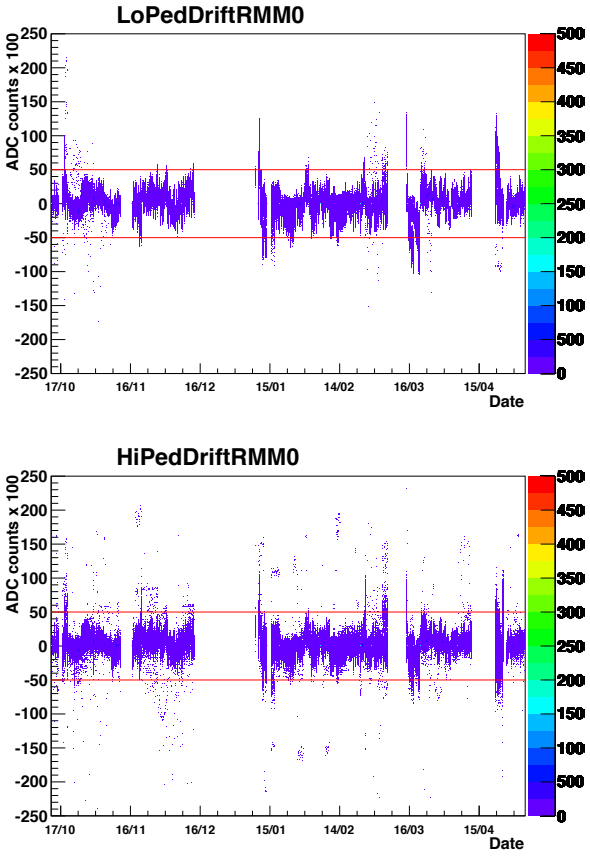


Figure 5.8: Pedestal variation in time for RMM0 during Run 4. The red lines correspond to the 0.5 ADC interval limit.

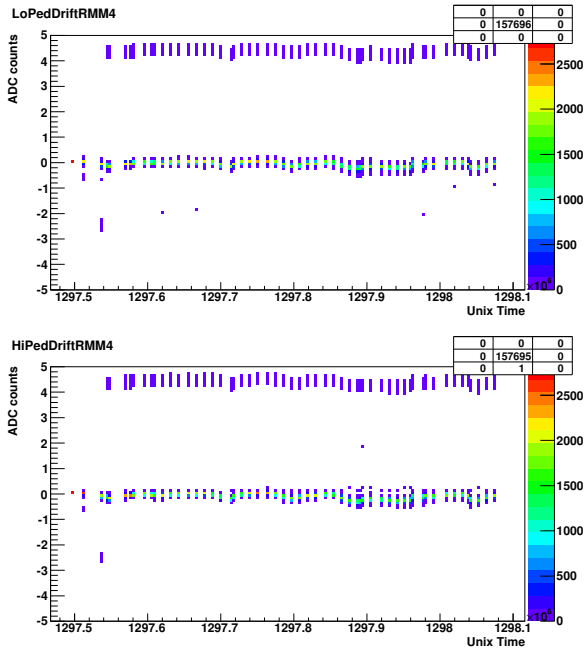


Figure 5.9: Large pedestal variation for RMM4 during Run 2. The fluctuations are caused by TFB43.

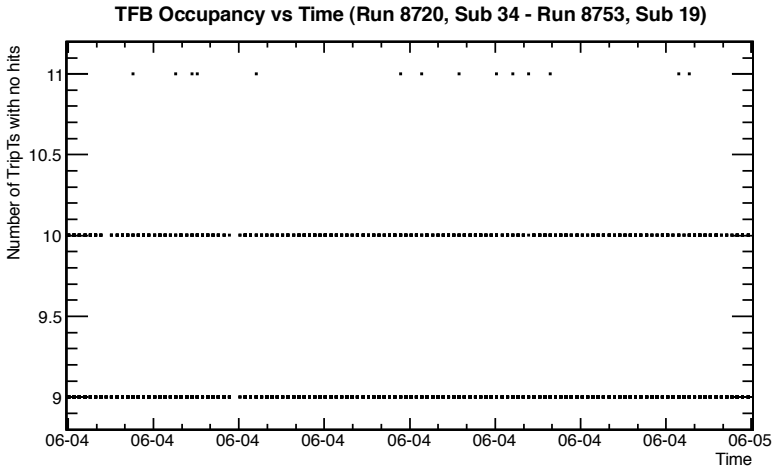


Figure 5.10: Number of the TripT chips with the invalid time stamps during Run 3.

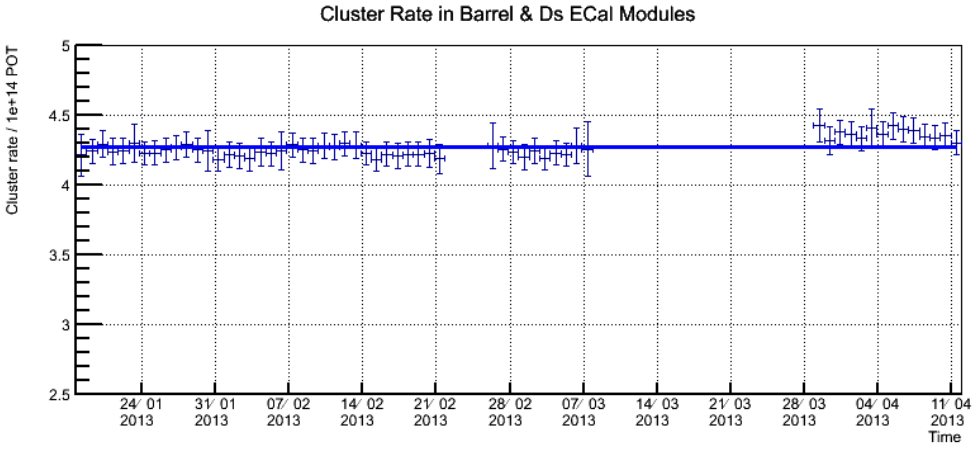


Figure 5.11: Clustering rate in 24 hour bins in the Run 4. There is an observed excess of the ECal clusters from March 2013 with respect to the average rate value, referenced by the line, caused by the changes in a reconstruction software algorithm.

RMM Offsets

The ECal RMM time offsets have a fairly unstable behaviour during data taking as shown in Figure 5.14. This inconsistent behaviour can be calibrated out, hence the offsets do not have a direct impact on the DQ flag. However, the observation of the rapid jumps in the offset values supports the detection of other problems and the starting of a further investigation.

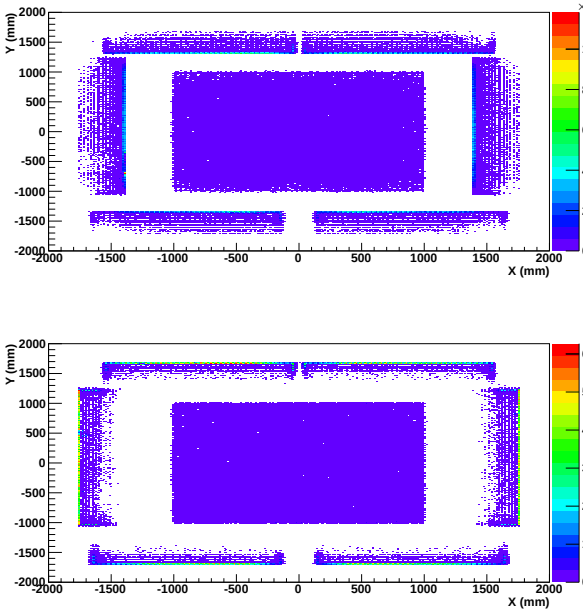


Figure 5.12: Distribution of the reconstructed tracks in the Downstream ECal (Run 4).

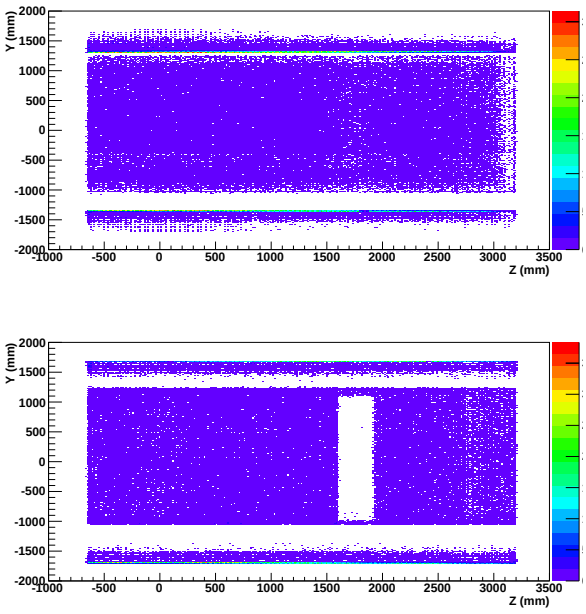


Figure 5.13: Distribution of the reconstructed tracks in the South part (top) and North part (bottom) of BrECal (Run 4).

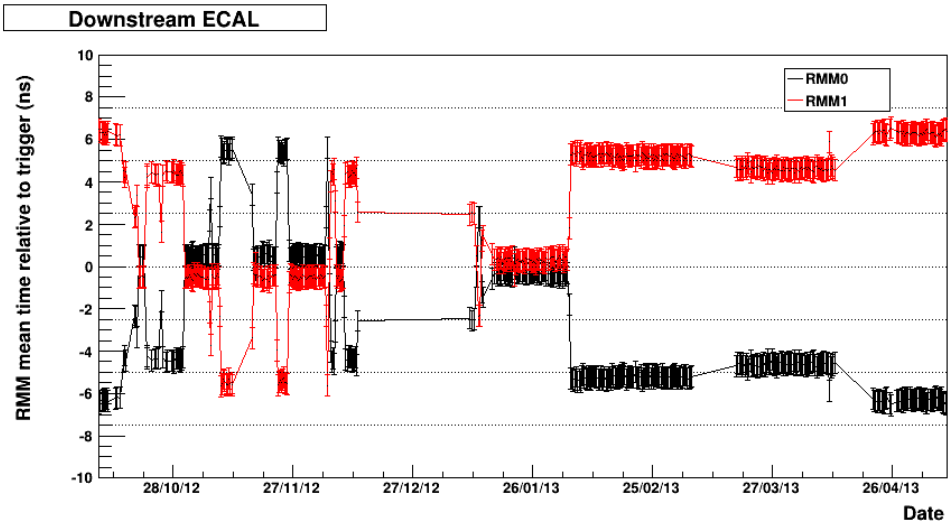


Figure 5.14: DsECal time offsets during Run 4 data taking. Here, the time offsets for RMM0 and RMM1 are calculated with the reference to their average value.

5.3.7 GSC Monitoring

The slow Control Monitoring is a visual inspection of the temperature, voltage, current and cooling water flow. Each of the monitored quantities does not have a direct influence on the DQ flag value, however it constitutes an integral part of the detector performance.

Temperature

Each ECal TFB has 2 temperature sensors located on the inner board and exterior. There are 732 temperature sensors monitored in total. The temperature variation depends on several factors related to the magnet status and weather conditions. At a low temperature condition in the pit, a diurnal cycle is observed as expected. During warmer weather such relations do not occur and the records follow the external temperature variations. Unfortunately, there is a number of broken TFB temperature sensors which are listed in the DQ technical notes for each data taking run.

Voltages

Four voltage lines of values 2.2 V, 3.2 V, 3.8 V and 5.5 V are monitored for each ECal TFB. There are several anomalies in the voltage reading such as widely oscillating readout or flawed readout for well operating TFBs. Hence, the current drawn by the bias voltage supply was examined to support the voltage monitoring.

Cooling Water

The water cooling flow is checked with water meters fitted on the cooling loops. The flow was stable for all the data taking runs with an average value of 4 to 5 l/min with minor annual variations, which are expected.

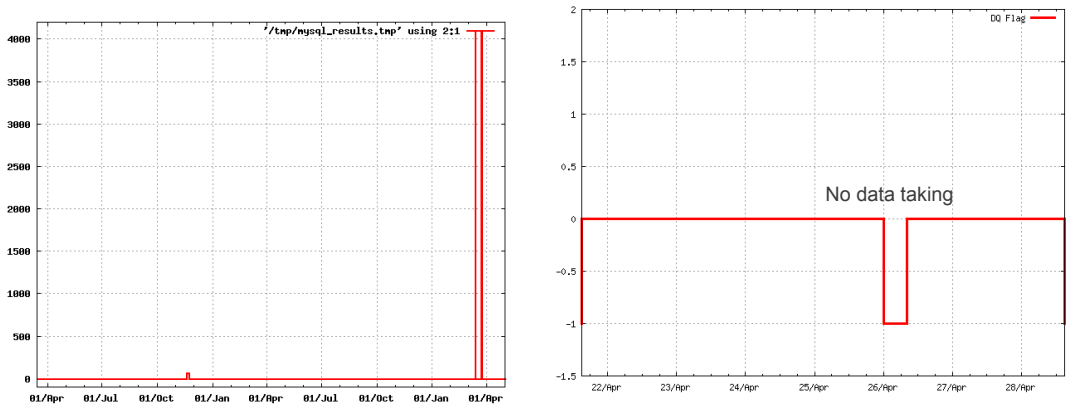
ECAL DQ Flag

The monitored quantities described in this chapter allow the quality of collected data to be assessed in a direct or indirect way. Variables such as the gain and pedestal variation, cluster ratio, cosmics checks and GSC monitored parameters do not change the quality of data result but they allow for any variations to be recognized and calibrated out as well as support any detector operation problem to be identified. The remaining ECal parameters, i.e. beam timing, number of dead and drifting channels and TripT chip status have a direct impact on assigning the data quality flag value to each separate ECal RMM. The overall full ECal flag is also determined as a result of each RMM status. This flag is a 12 bit word where each bit corresponds to the flag value for an individual RMM. The bit values have the following meanings

- 0 is good,
- 1 is bad.

If the decimal value of the bit number is set to zero, than the full global ECal status is good. If it is greater then zero, the global flag is bad. For an undefined status the global flag value is set to be -1. The ECal flag value during the Run 1, Run 2 and Run 3 data taking periods is shown in Figure 5.15.

During the Run 1, Run 2, Run 3 and Run 4 data taking, 6.43×10^{20} POT was recorded at the near detector ND280 with an efficiency of 98.45%. Data with a good ND280 DQ flag is equivalent to 5.98×10^{20} POT, giving a 91.53% efficiency. On the grounds that T2K is a multi purpose experiment, there are a number of analysis groups performing physics measurements and each group uses data sets relevant to the specified requirements, i.e. sub-detectors used in their analysis. In this thesis, the data from Run 2, Run 3 and Run 4 are studied. This is equivalent to 2.28×10^{20} POT when the PØD is filled with water giving a 81.62% efficiency (recorded plus good DQ flag) and 3.49×10^{20} POT when the PØD water bags are empty with a 96.84% efficiency (recorded plus good DQ flag).



(a) (b)

Figure 5.15: The ECAL flag history shown for Run 1, Run 2 and Run 3 (a), and ECal Run 3 (April 2012) flags zoomed to the lower values range (b).

All the data sets are processed with the ND280 software for a number of data productions. Each production is prepared with the most recent and stable software release. The ND280 software is described in the next chapter.

Chapter 6

ND280 Software

The ND280 detector software is coded in C++ and Python with significant use of ROOT [59] and existing high energy physics software libraries. The ND280 software is composed of independent but interconnected packages. Each package is designed for an individual purpose such as simulation, reconstruction, calibration, analysis etc. All the nd280 packages are managed by the Configuration Management Tool (CMT) which is itself built upon the Concurrent Versioning System (CVS). CMT is used to manage the build of each software release, and it operates in conjunction with CVS in order to keep track of all the changes in the set of the package files. This allows several developers to collaborate with continuous code improvement. The first stable nd280 software release came out on 14th Feb 2008 (v4r1). The software version used in this thesis analysis is v10r11p31, and is a sub-production of the MC Production 5.

A diagram showing the nd280 package structure is presented in Figure 6.1. Both data and MC files have an oaEvent format, which is based on the standard ROOT classes. As the data is recorded in the MIDAS format, the event configuration is formed by the oaUnpack package. The fundamental unit of an event is a "hit", which represents the scintillated light from the bar for tracking sub-detectors. All the events pass through the software reconstruction chain that corresponds to calibration, reconstruction and final data reduction to the analysis package oaAnalysis. The author of this thesis contributed to both the time calibration and reconstruction of the PØD-ECal.

6.1 Monte Carlo Simulation

The Monte Carlo events are generated using a specifically developed package, nd280Control, which provides the tools for different types of MC simulations.

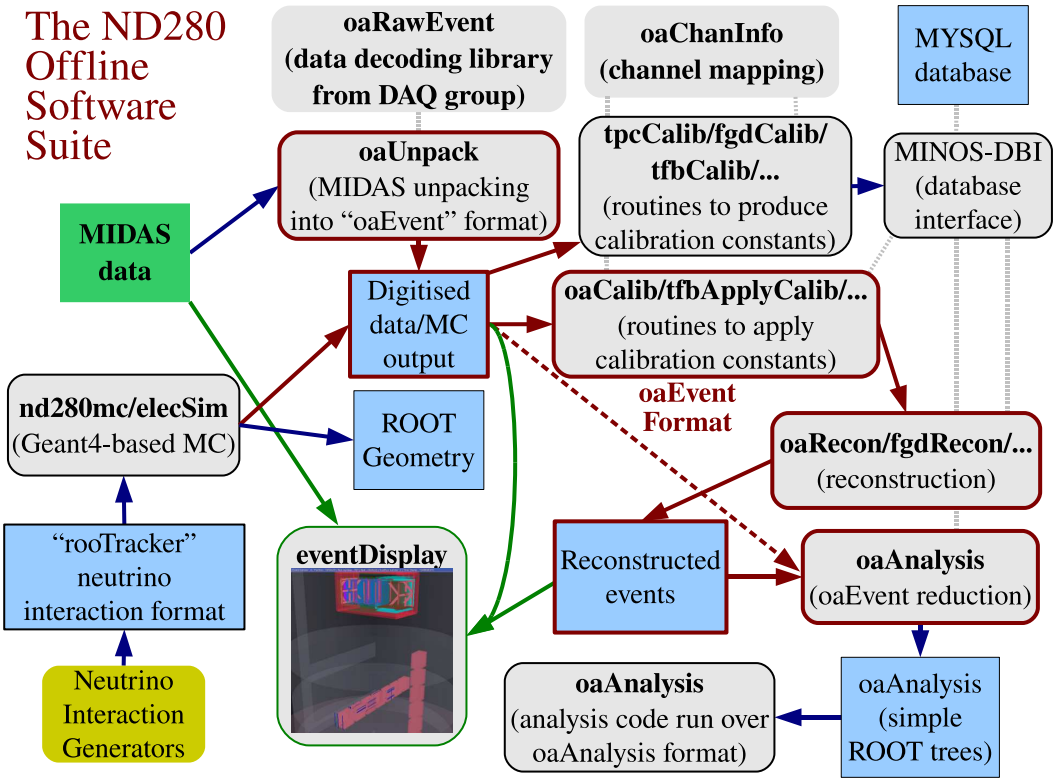


Figure 6.1: A diagram showing the package structure of the ND280 software suite [8].

This enables one to choose the most convenient option for the specific studies. All different options can be specified by a configuration file. The Monte Carlo simulation is carried out in four separated steps as described below.

6.1.1 Beam Flux Simulation

The first stage of the MC production concerns the beam flux simulation. This is a simulation of the physical processes involved in the neutrino production and the neutrino energy spectrum, presented in Ref. [5]. The interaction of the primary proton beam and secondary particles is simulated by FLUKA2008 [60], which appears to be in the best agreement with the external hadron data. The proton beam monitor measurements are used as inputs into the neutrino flux simulation.

The main difficulty in a flux prediction arises from the hadron production process uncertainties appearing from the proton nucleus scattering. Therefore, a simulation of such hadronic interactions at T2K is re-weighted using external data from the NA61/SHINE [61] experiment. The energy spectrum used at the

NA61 experiment covers the entire phase space region for the T2K flux. Both, the pion and kaon, cross section measurements are applied during the simulation to obtain further flux uncertainties reduction.

The kinematic information for particles exiting from the target and propagating through the beamline are simulated with GEANT3 [62]. All simulated particles are tracked to the point where they interact, decay or drop below an energy cut-off. The generated neutrinos are directed towards the per–Kamiokande plane.

A good understanding of the neutrino flux prediction is crucial as all T2K measurements are highly dependant upon this prediction. The ND280 and Super–Kamiokande flux ratio uncertainty is found to be less than 2% near the flux peak, see Figure 6.2. The overall uncertainties on the flux prediction are evaluated to be below 15% near the flux peak.

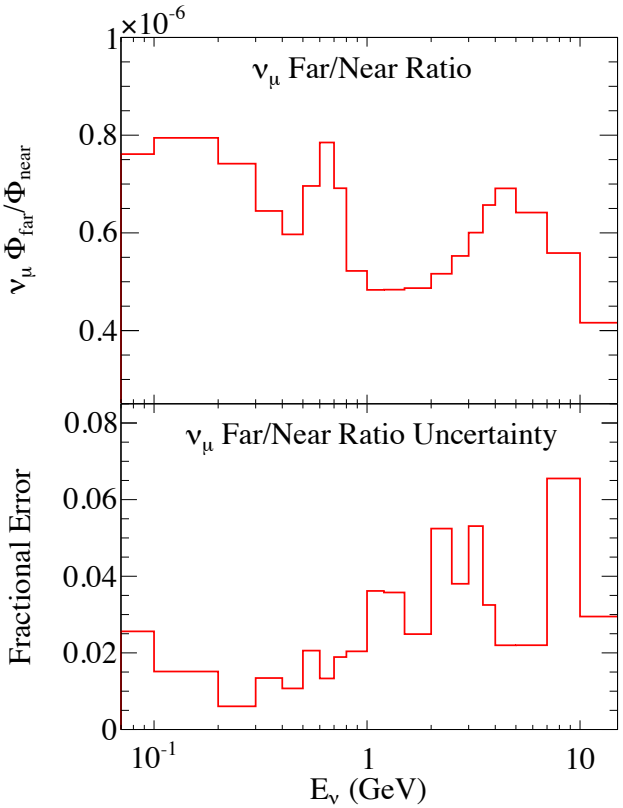


Figure 6.2: The far to near detector ratio for the ν_μ flux prediction (top) and the uncertainty on the ratio (bottom) [5].

6.1.2 Neutrino Interactions Generator

During the second stage of the MC simulation, an event generator simulates the neutrino interactions with the detector materials followed by the movement of the primary produced particles within nuclear matter. This simulation covers all the possible interactions within the entire kinematical region using the chosen theoretical models. At the T2K experiment, three neutrino generators are applied and tested: NUANCE [63], GENIE [64] and NEUT [65]. Each of them has different simplifications and model assumptions in the actual implementation. NEUT has been chosen as the official generator and is used for the MC studies presented in this thesis.

There are two external inputs passed to the generator: the neutrino vector file produced by the beam MC, and a ROOT-based simulation of the ND280 detector geometry. NEUT produces events against a single target and therefore is convenient and successful in the homogeneous detector studies such as Super-Kamiokande. The main target nuclei are protons, oxygen and carbon. However, other types of nuclei are also supported in the code by interpolation techniques. NEUT covers the neutrino energy region from a few MeV to the TeV range.

Within the T2K experiment energy region, 1-10 GeV, each of the primary interactions, characterised in section 3.1, is simulated according to the chosen theoretical model with fixed parameter values. The charged current quasi elastic interaction is simulated according to the Llewellyn-Smith, Smith and Moniz model [66], [67]. Resonance production is simulated in agreement with the Rein and Sehgal model [40]. The coherent pion production has been applied according to the Rein and Sehgal model [68]. The deep inelastic scattering cross section is calculated beyond an invariant mass energy with nucleon structure functions from GRV98 [69] with Bodek-Yang corrections [70]. During the next stage NEUT simulates the secondary interactions of the primary particles within the nuclei environment, e.g. recognizing mesons and recoiling nucleons. These interactions are generated by a particle cascade routine which includes the effects of the Fermi motion and Pauli blocking. This is followed by the final state interactions simulated with a microscopic cascade model. Each individual model's uncertainties contribute to the NEUT associated systematic errors. These errors are determined by a validation of the generator against a broad range of published data sources.

NEUT outputs the final stage particles exiting the nuclei, along with their kinematic information i.e. 4-momenta of target nucleons, outgoing leptons, all

the hadrons and gammas exiting from the target nucleus. Information on the primary interactions is also provided.

6.1.3 Detector and Electronics Simulation

The third step in the MC production is the detector simulation, which models the propagation of particles leaving the nuclei and travelling through the ND280 detector. This procedure is based on the GEANT4 [71] package and uses a realistic detector materials simulation. The outcome is an energy deposit loss in the detector active material such as scintillator bars or TPCs. The output is saved in a hit format for every particle interaction for each spill. Information concerning the position, timing and deposited energy is stored in the hit object.

The last stage of the Monte Carlo generation is the detector electronics simulation. The elecSim package models the conversion of the deposited energy into the actual detector response as of real experimental data. This includes: the attenuation of light along the bars or WLS fibre, MPPC response, and existing electronics noise. The output is a list of digits forwarded to the calibration stage, which has the same format as the raw collected data.

6.2 Calibration

The calibration procedure is carried out by a number of packages managed by the oaCalib package. It gets as an input of raw data or MC simulation, which is a digital information, and converts it into the hits format. This procedure includes calibrating a number of quantities to correct for measurement uncertainties associated with devices and geometry. The application of the calibration constants can be done on a channel-to-channel basis, spill-to-spill basis or during reconstruction procedure.

For all Trip-T type sub-detectors the calibration procedure is as follows. The calibration constants are determined using the inter-spill cosmic data or specific dedicated runs e.g. charge injection. The first step is to correct for the effect of TFB electronics. The charge accumulated from the noise in the electronics is subtracted from the ADC values of the digits. This is the pedestal correction which is constantly monitored during the data acquisition procedure. Next, the ADC format is converted into the charge format. Then, the MPPC related effects such as noise, efficiency, saturation effects and gain drift caused by temperature and voltage variations are calibrated. Subsequently, a normalisation of the bar-to-bar variation across each sub-detector module is performed. Finally, compensation is

made for the time walk effect. This concerns the time of the offsets associated with the different cable lengths connecting the devices and the charge dependency on the time stamps. For the FGD modules, there is an additional step which concerns the assistance from the timing markers assembled in the front end boards. This additional effort to decrease the time fluctuation is essential due to its application to the TPC modules.

A separate calibration procedure is applied to the TPCs due to their different electronics readout system. The calibration accounts for the correction of the micromegas pad-by-pad response, gas density, drift velocity, micromegas-by-micromegas variation and time offsets.

6.3 Reconstruction

The reconstruction phase uses the calibrated hits as input, and creates basic objects for physics analysis through pattern recognition algorithms. Each subdetector has an individual reconstruction package. At the final stage, locally reconstructed objects are passed to the global fit which performs a matching across the ND280 detector, using the RECPACK [72] toolkit. In principle, the global objects are used for the physical studies. However, as of Production 5, the output from the packages, pØdRecon and pØdecalRecon, were not fully integrated into the global fit, and therefore the implementation of the global objects into this thesis analysis has a limited application. In this thesis we use the standalone PØD and PØD-ECal packages reconstruction, instead of the global reconstruction, and those are described in the next section.

6.3.1 PØD Reconstruction

The pØdRecon package is designed to reconstruct objects in the PØD and it has a specific tuning to allow for π^0 reconstruction and particle identification. The reconstruction algorithms are described in Ref. [73]. For the analysis in this thesis, the final pØdRecon output objects are used, such as track, shower, vertex and node. The basic steps of these objects reconstruction are described below.

Initially hits are segregated into 23 cycles, according to their timestamp. The cycle concept has been explained in detail in section 5.3.1. The interaction hits can be contained within one or more of the eight cycles, from five to twelve. If there is more than one time window with the hit information, it means that the secondary or external interactions appeared. Next, the cleaning algorithm is applied. This removes singular hits on the low charge cut basis. For hits

below the 2 PE threshold, the algorithm penetrates the nearest neighbourhood for other hits, and if any hits has not been found, the hit is removed. Additionally, the upper threshold timing cut is applied, which means that hit time has to be within 50 ns of the mean of all hits. This eliminates hot channels, channels with continuously high PE value or electronic distortions.

The remaining hits are reconstructed into larger objects in two distinctive steps. Initially, all tracks are identified and removed from the hit collection, then a shower clustering algorithm assembles remaining hits into the shower objects.

The track reconstruction in the PØD is carried out in the following sequence of steps. A single hit carries 2D spatial information (from the scintillator bar), and therefore the initial track search is made in two planes: xz or yz. First, the Hough Transform algorithm, a robust line detection, is applied, followed by a track extension procedure. The algorithm searches at both ends of a plane for additional hits. Then, when a track is found and removed from the cluster, the procedure is repeated on the remaining hits, until no more tracks can be recognized. All the hits are tested for the spatial and timing requirements, which the hit pattern should satisfy. Otherwise, hits are passed for further clustering.

At this point, the reconstructed two dimensional tracks are examined by the charged particle track matching algorithm to form 3D track objects. It starts with a 2D track pair arrangement, which is tested with the matching criteria. The main principle to join tracks states that the z projection of both tracks can not differ by more than 10 cm. Amongst all possible matching pairs the algorithm will choose the pair which will have the smallest total length difference. In the case of equal lengths, the pair with the smallest charge difference is chosen. This procedure iterates through all the 2D identified tracks.

The following step is a 3D vertex reconstruction based on the iteration of the track crossing points. The vertex point is simply the intersection of the extrapolated tracks and for the one track-only case it is defined as the most upstream track point.

The remaining hits are then clustered together into shower objects. Similarly to the tracks, in the first step clusters are combined into 2D objects. These isolated hits are selected from the hits pattern in the x and y direction by a Delaunay triangulation algorithm. Next, follows a 3D matching procedure and 3D vertex finding. The direction of a shower is determined by the principal component analysis (PCA) concerning the charge distribution.

The last phase is a particle identification (PID) procedure, which classifies tracks and showers. Tracks are separated as muon-like, electron-like or proton-

like. In the case of shower PID, two possible identifications can be assigned. The first one is the EM PID, which concerns photons and electrons, and the second one is Other for not EM like particles. The parameter describing the shower identification is the PID weight. It is calculated as the difference of the log likelihoods of the EM and Other shower PID

$$\log \Delta\mathcal{L} = \sum_i \left(\log \mathcal{L}\left(Q_i \mid \text{EM}\right) - \log \mathcal{L}\left(Q_i \mid \text{Other}\right) \right), \tag{6.1}$$

where Q_i is the corrected charge at the ith scintillator plane. The efficiency for PID studies for all the PØD reconstructed objects can be found in the technical report [11]. The last algorithm in the reconstruction chain is the decay electron muon tagging.

6.3.2 PØD-ECal Reconstruction

The pØdecalRecon package was initially integrated into the ecalRecon package, which carried the reconstruction procedure for all the ECal modules. The original reconstruction studies are described in Ref. [74]. The unique and simple structure of the PØD-ECal leads itself to a separate straightforward 2D analysis. Consequently in February 2011 a standalone pØdecalRecon reconstruction package was created. The new package inherited the basic ECal reconstruction algorithm chain but with modified methods. Figure 6.3 shows a PØD-ECal reconstruction diagram and the PØD-ECal output objects flow with an ND280 software event.

Initially, hits are extracted from the ND280Event by the PØDECALGetHits algorithm. This is followed by the spill and hit preparation stage. The initial hit separation is conducted according to the hit position. Hits are assigned to one of the six PØD-ECal modules: top left, top right, bottom left, bottom right, side left or side right. Next, the hit collection is divided with respect to their time stamp, in order to reconstruct the beam structure. The collected hits, that fill the integrated beam time range, are divided into *bunch* sections when a time difference of 50 ns or more is found between them, and into spill sections when there is a time difference of 100 ns or more.

The PØDECalBasicClustering algorithm is then applied to each spill. It collects the hits induced by a single particle passing through a module, into one group. This is based on a spatial nearest neighbour clustering procedure. First, the hit with the highest charge deposition is used as a seed to integrate other hits into a 2D cluster. The candidate hit first passes a clustering stage on the condi-

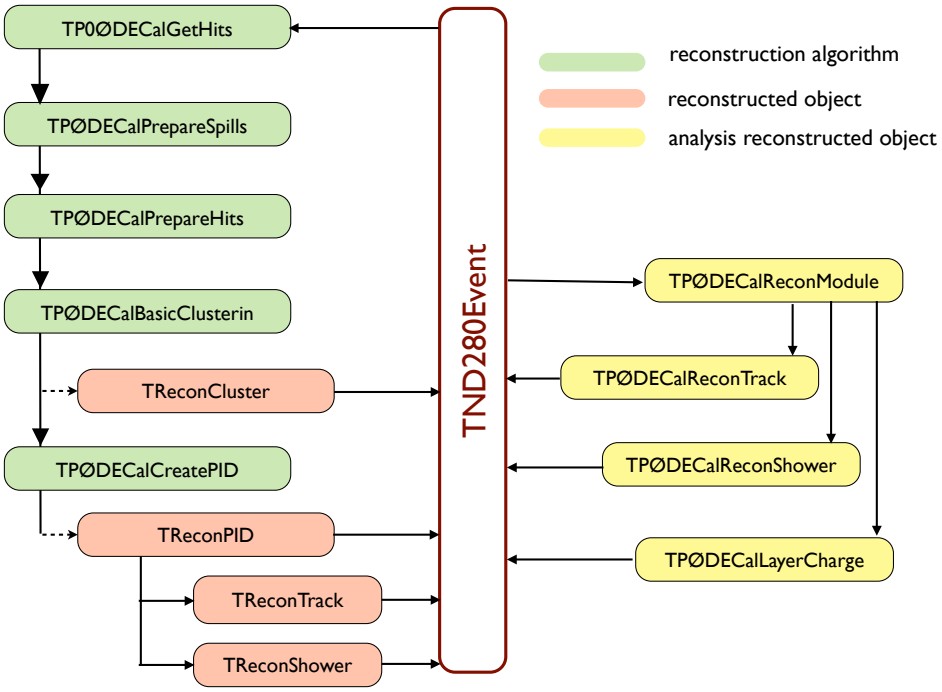


Figure 6.3: Simple flowchart of a PØD-ECal event reconstruction and its flow to the analysis output.

tion that its position is within a 4 layer or 2 bar distance from the seed. Next, a timing cut is applied to eliminate noise. A hit gets associated with a cluster when its time value lies within 15 ns of the seeding hit. Additionally, a minimum of two hits are required to constitute a cluster. When a cluster is formed, the above steps are repeated for all remaining hits, i.e. the new highest charge hit is chosen as a seed for the next cluster formation. This procedure is repeated until the number of remaining hits is below a pre-set minimum. All the unclustered hits are saved at this stage of the reconstruction and the reconstructed clusters are saved as a ReconCluster object in the ND280 Event. The clustering efficiency studies are presented in Ref. [74].

The last reconstruction phase runs the PID algorithms on the formed clusters. Since the clusters are relatively small, ranging in size from two to a dozen or so hits, and they carry only 2D spatial information, the particle identification constitutes only two elementary categories: track and shower. The track-shower discrimination is established by a multivariate analysis technique (MVA), using

the Artificial Neural Network of the Multilayer Perceptron (MLP) class.

All events that pass through the software reconstruction chain for both, data and MC, in the final stage are written into the analysis package where reconstructed informations is stored in an appropriate format for the analysis to be performed. In this thesis we use the standalone PØD and PØD-ECal packages output for the NC1π^0 analysis which are presented in the next chapter.

Chapter 7

NC1π^0 Channel Studies

The goal of the studies, presented in this chapter, is to develop a PØD-PØD-ECal inclusive NC1π^0 event selection which will contribute to the future NCπ^0 cross section measurement for neutrinos interacting in the PØD. It is the first analysis in T2K that attempts to use the PØD-ECal.

The ND280 group is working on a number of complementary NC1π^0 analyses. The sub-division of the analysis is based on the near detector complex topology and depends on the π^0 production area within the ND280. The two main NC1π^0 analyses are the tracker and PØD analyses. This thesis analysis selects events with a vertex in the PØD, where the π^0 is produced within the PØD sub-detector volume. The analysis was performed for the two PØD configurations: filled with water and without water. A global PØD-PØD-ECal NC1π^0 analysis chart flow is shown in Figure 7.1. The PØD standalone, fully contained analysis (green background) is presented in Ref. [75]. This chapter outlines the case where one π^0 decay gamma is converted within the PØD, and the second gamma converts within the surrounding PØD-ECal module (blue background). Other outlined possibilities might be a subject of future studies.

Due to the limited reconstruction information from the PØD-ECal, which is related to its structure as described in the chapters 4 and 6, the studies in this chapter do not include the π^0 mass reconstruction in the final stage as it is not possible.

7.1 Groundwork for the NC1π^0 Analysis

This section presents a study of the NC1π^0 interaction vertex topologies within the PØD sub-detector. The π^0 decay photon topology is shown in Figure 7.2. Out of all the events, 70% (65%) are fully contained in the PØD filled with water

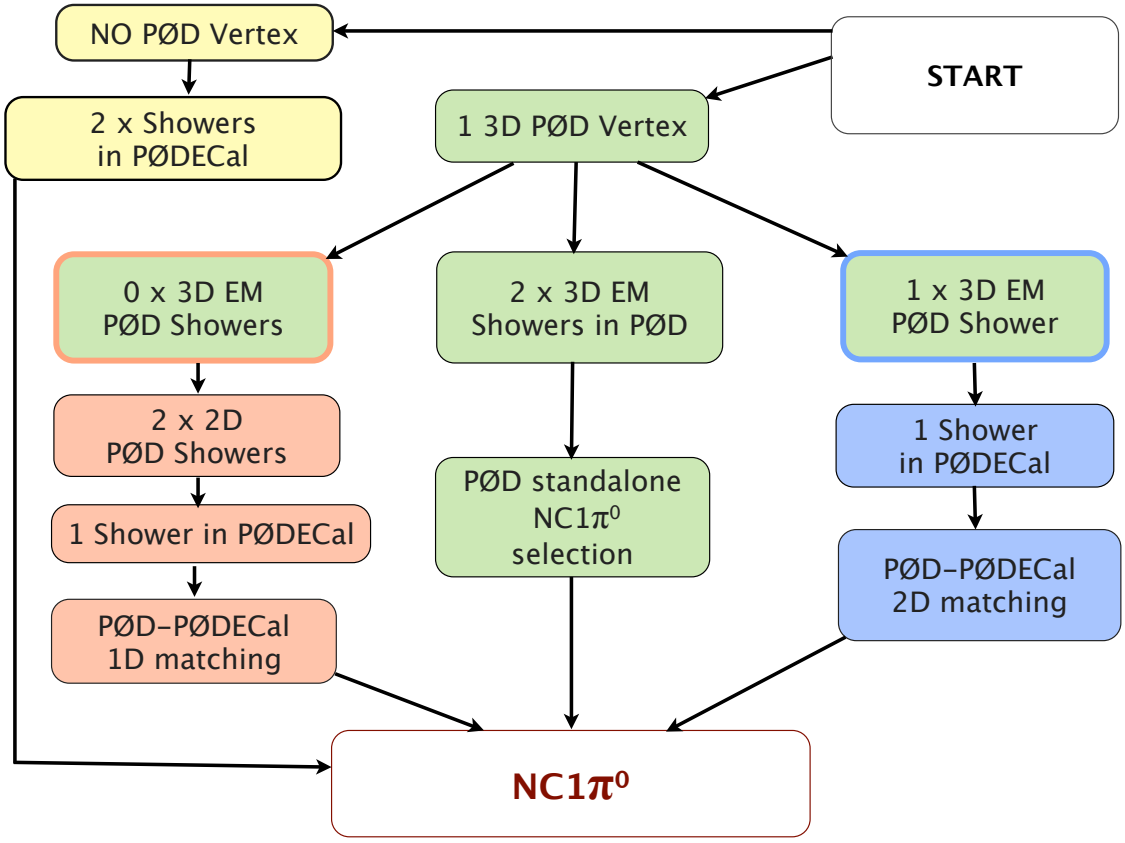
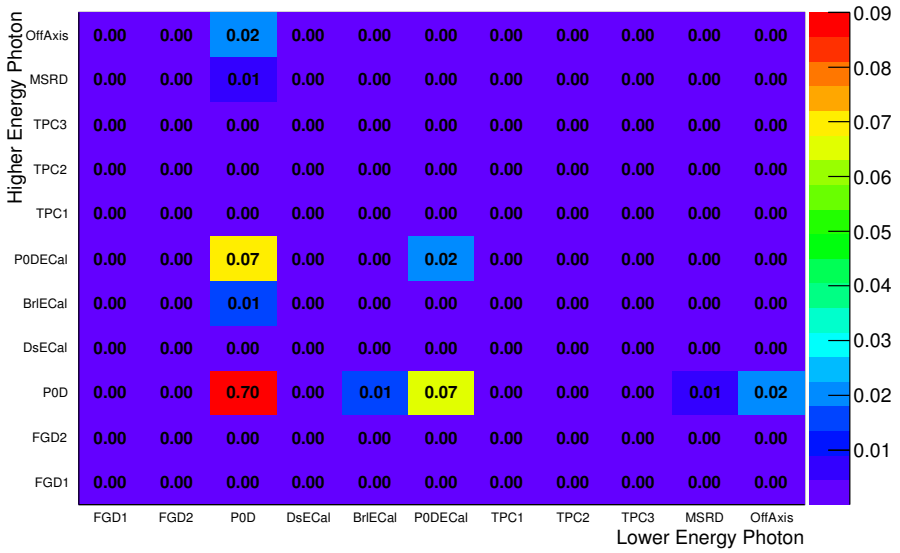


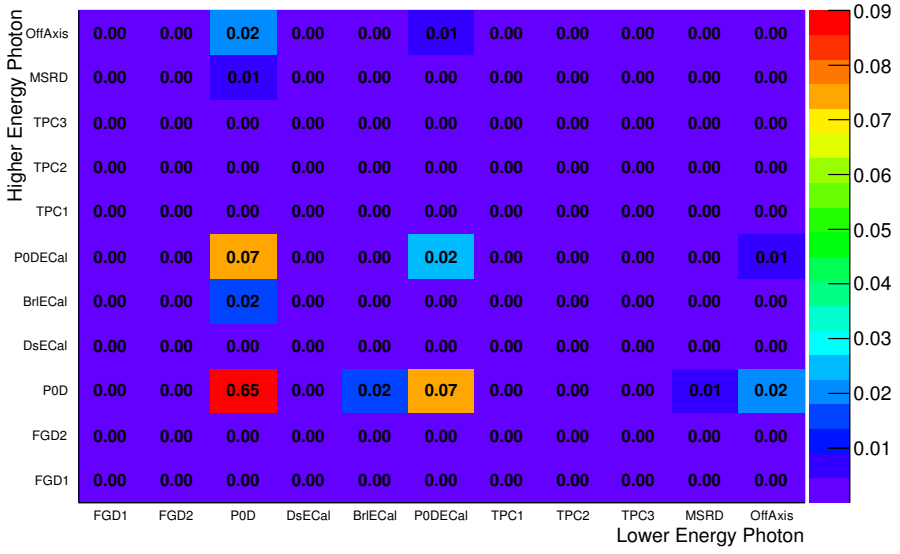
Figure 7.1: Full flow chart for the PØD-PØD-ECal NC1π^0 analysis. The green background boxes represent PØD only analysis flow, and red and blue background boxes show possible PØD-PØD-ECal event analysis. The yellow boxes represent PØD-ECal only analysis flow.

(air). In the remaining 30% (35%) of events at least one decay gamma escapes outside the PØD, where the decay takes place and converts into a shower in another sub-detector. The case where one photon converts within the PØD and the second converts within the PØD-ECal accounts for around 20% of all the PØD fully contained events. Further MC studies showed that the phase space of the π^0 produced by the neutrinos has the distributions as plotted in Figures 7.3 and 7.4 for the PØD filled with and without water, as expected.

The analysis presented in this chapter is based on two types of MC events: *cherry picked* and *magnet* samples from the sub-Production 5F (detailed MC simulations are described in chapter 6). The first one restricts the simulated event vertices to be only in the basket area of the near detector and selects events containing one NC1π^0 neutrino interaction per event. This allows to enhance the number of signal events in the MC studies to meet the statistical requirements.

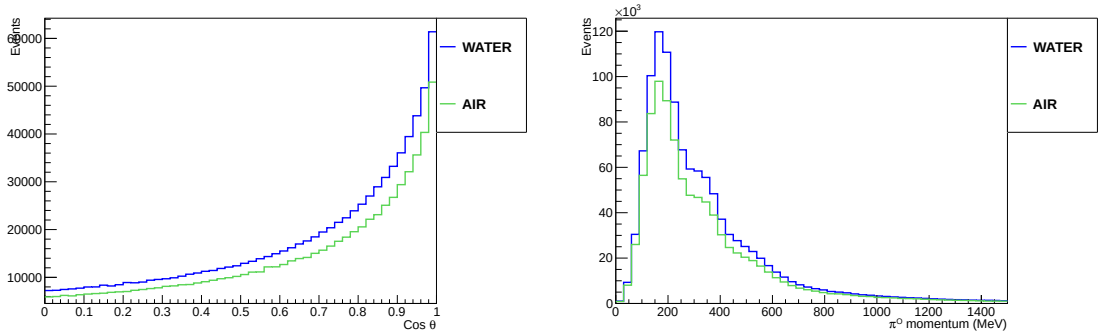


(a) Water



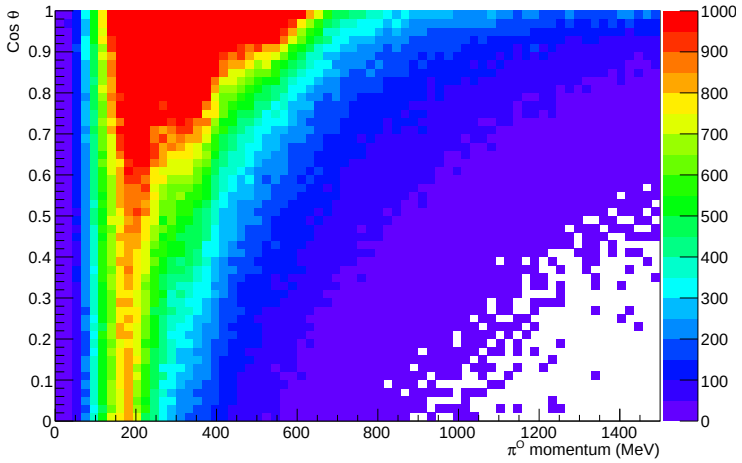
(b) Air

Figure 7.2: The showering topology of the NC1π^0 decay photons for neutrino interactions in the PØD. The conversion sub-detectors of the lower and higher energy photons are showed for the PØD filled with (a) and without (b) water.

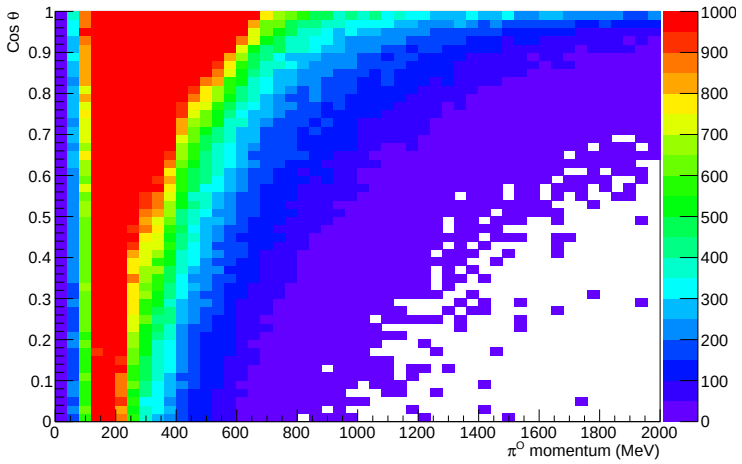


(a) (b)

Figure 7.3: The angular distribution (a) and momentum (b) of the π^0 produced by neutrinos in NC events.



(a) Water



(b) Air

Figure 7.4: The phase space of the π^0 produced by neutrinos interacting in the PØD filled with (a) and without (b) water.

The size of used *cherry picked* files is of 1.8×10^{22} POT. The *magnet* MC sample uses the NEUT generator and includes all the events generated within the magnet. The *magnet* MC summary is presented in Table 7.1. The beam configuration details are presented in Table 7.2.

Run	PØD Configuration	Beam Configuration	P0T
2	water	B	1.134×10^{21} POT
2/3b	air	B	1.002×10^{21} POT
3c	air	C	2.978×10^{21} POT
4	water	C	5.513×10^{21} POT

Table 7.1: The NEUT MC configurations used for the NC1π^0 studies for the *magnet* MC sample. The POT counting has been done after the application of the good beam stability flags requirement.

Beam	Power (kW)	POT/Spill	Repetition [1] period (s)
B	120	7.989×10^{13}	3.20
C	178	9.463×10^{13}	2.56

Table 7.2: The beam configurations used in the MC simulations.

Additionally, the *sand* MC files were accessed to account for the neutrino interactions in the cavern that surrounds the ND280 detector. As of Production 5, the *sand* files are produced separately from the *magnet* MC. However, the sample of *sand* events is statistically small due to the limited availability of produced events at the time of making the analysis in this thesis. All the *sand* events happen then to be rejected after the initial stage of the selection studies.

The data used for the analysis has been recorded during the Run 2, Run 3 and Run 4 data taking periods and processed using Production 5G. Both MC and

[1]Linacs are pulsed accelerators where the beam is delivered to the users in pulses of a given length at a given repetition period.

data productions were locally reprocessed with the most recent version of the pØdecalRecon package integrated into ND280 software v10r11p31 in order to get the best available reconstruction of the PØD-ECal sub-detector. The PØD-ECal time calibration has not been yet applied in the used Productions (at the time of this analysis the Production 6 containing the needed pØdecalRecon package version was being worked on). No global reconstructed objects are used due to a lack of matching algorithms between the PØD and PØD-ECal (as explained in section 6.3.2). The applied flux tuning version is 11b v3.2 and is used for the purpose of a good agreement between data and MC.

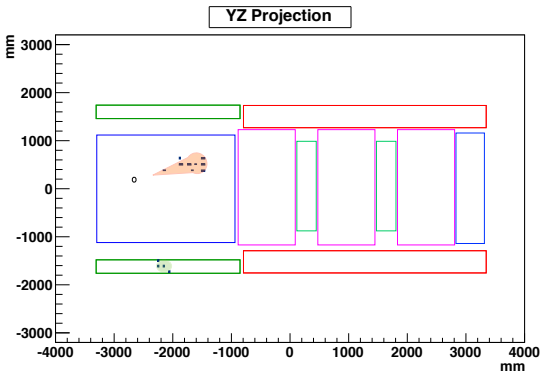
7.1.1 Signal and Background Definition

The NC1π^0 signal is defined by the final state interaction particles. The requirement is for only one π^0 to exit the nuclei without any other leptons or mesons. However, any number of neutrons or protons is allowed to leave the nuclei. This implies that, using the NEUT generator notation code, the signal can be defined by the interaction type with the assigned number 31 and greater. Moreover, every signal event has to be generated by the ν_μ beam component and the interaction to happen within the PØD sub-detector volume. The signature of the π^0 is two decay showering gammas, one gamma in the PØD and the other in the PØD-ECal, which implies the usage of the reconstruction objects from both sub-detectors. The second gamma is a high angle escaping particle, without a specific angular momentum requirement in the signal definition.

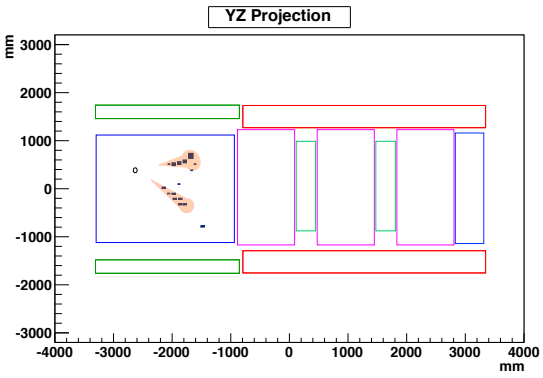
In this analysis all the events are divided into several categories: NC1π^0 (signal), other neutral current, charged current with a charged pion, other charged current and other (backgrounds). The latter category includes events with external vertices, multiple interactions or a reconstructed PØD vertex without a corresponding true vertex. As there is only one defined topology for all the signal events, the NC1π^0 events originating outside the PØD, or leaving their signature elsewhere than within PØD and PØD-ECal, are considered as a background and classified into other NC events. The topology of each type of analysed event is shown in Figure 7.5 as an example on the event display.

7.1.2 PØD Object in the NC1π^0 Event

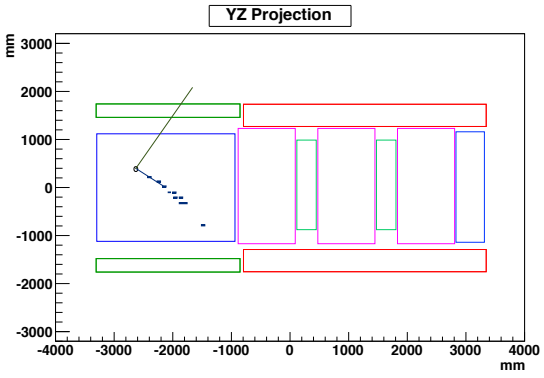
All the PØD reconstructed objects that are used in this chapter analysis, such as 3D vertices, showers and tracks, come from the PØD final reconstruction output. The PØD reconstruction algorithms are briefly described in section 6.3.1 and more



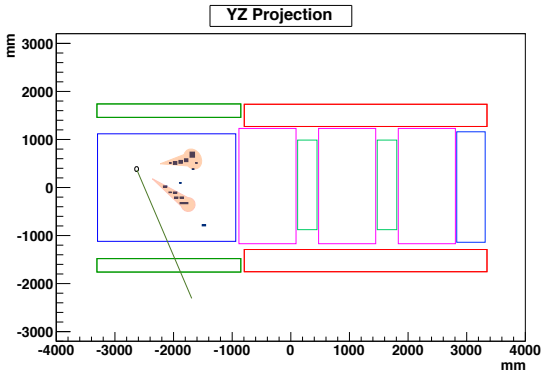
(a) Signal: NC1π^0



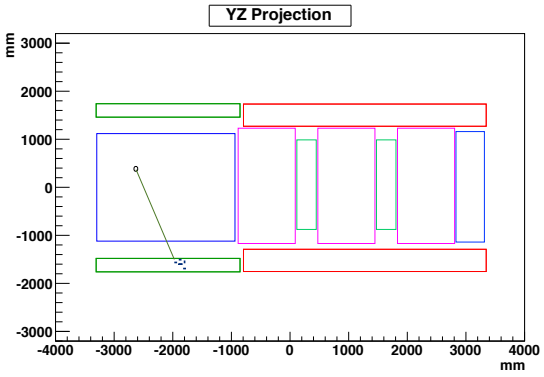
(b) Bcg: NC Other



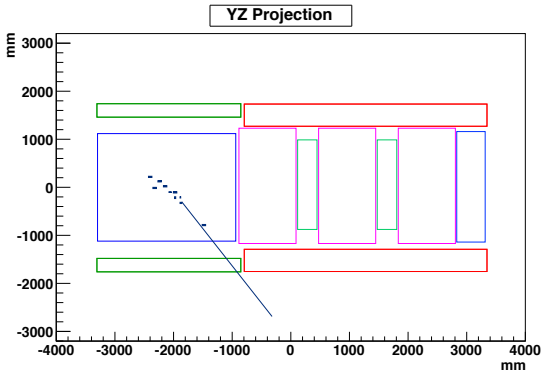
(c) Bcg: CC with $\pi^{\pm}$



(d) Bcg: CC Other



(e) Bcg: CC Other



(f) Bcg: Other

Figure 7.5: Examples of signal and background topology for different event categories.

detailed information with parameters values can be found in Ref. [73], [76]. All the PID reconstruction parameters are described in Ref. [11]. The reconstructed clusters can be categorised as tracks, which have assigned PID as Muon-Like, Proton-Like, EM-Like or Other. The first category (muons) are classified as Light Track and protons are categorised as Heavy Tracks. The last two categories, EM and Other, are a subject to the parametric fit and the shower reconstruction procedure. The output contains EM objects with assigned shower PID weight. The efficiencies of the particle reconstruction as a track or shower are presented in Table 7.3. This shows that the highest reconstruction efficiencies

Reconstruction object	True Muon (%)	True Electron (%)	True Proton (%)
WATER			
Light Track	62.5	1.7	20.3
EM	20.5	96.1	44.8
Heavy Track	17.0	2.3	34.9
AIR			
Light Track	47.2	1.7	12.5
EM	29.7	95.5	45.5
Heavy Track	23.1	2.7	42.0

Table 7.3: The PØD PID efficiencies for three different types of particles signature for the PØD filled with and without water [11].

are achieved for the showering particles, such as electron, 96.1% (95.5%) for the water-in (water-out) configuration, which have significantly better identification than the other types of particles. The highest misidentification as the EM shower object arises from the reconstructed protons. Although they are allowed by the signal definition, in the case where the proton is mimicking a showering particle, it might lead to a misguided number of π^0 decay photons. The photon-alone reconstruction studies in the PØD showed that the difference between the true and reconstructed gamma angle has a distinct peak at $\cos\theta = 1$, which implies a good reconstruction. Additionally, the exiting particles, which are defined as the particles with a hit in the last layer or in the outer two bars in the layer of the PØD, are also a part of the PØD-PØD-ECal study interests. Unfortunately, the PØD PID for exiting particles is not well understood in Production 5 and is

under improvement for Production 6.

As this analysis was developed for the PØD filled with and without water, the energy conversion for each configuration target has to be considered. This is important to accurately predict the energy of photon candidates. The energy scale estimation for the π^0 analysis in the PØD can be found in Ref. [77], where it is reported that the distribution of charge versus energy shows a linear tendency.

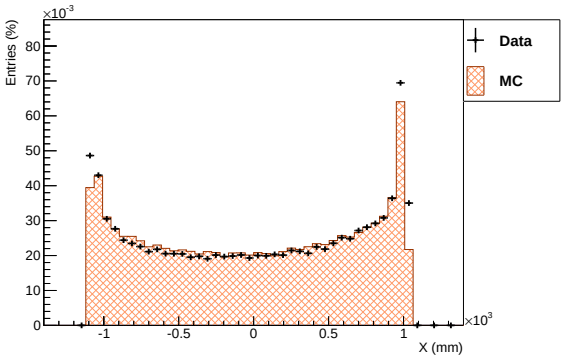
The data and MC comparison for the PØD reconstructed tracks and showers position variables is in good agreement, as shown in Figure 7.6. The comparison of the reconstructed time of reconstructed objects for data and MC is shown in Figure 7.7. Here, the data for the PØD filled with water show a wider peack tail in the distribution which is caused by Run 2, as reported in Ref. [9]. Additionally, a data-MC comparison of the shower reconstructed objects direction is shown in the spherical coordinates, θ and ϕ, in Figure 7.8 and 7.9. Both plots show good data-MC agreement.

Determining the interaction vertex in the PØD is part of the reconstruction algorithm chain, as described in section 6.3.1. The vertex reconstruction performance has been studied in Ref. [77] and presented in Table 7.4. It shows that the highest shift of the position is in the beam direction, z, and the interaction vertices on the water target have a slightly better resolution.

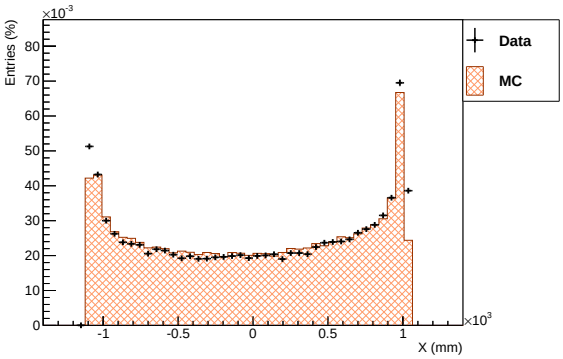
PØD Configuration	$\langle x \rangle$ (cm)	σ_x (cm)	$\langle y \rangle$ (cm)	σ_y (cm)	$\langle z \rangle$ (cm)	σ_z (cm)
Water	-0.06	5.52	0.06	6.06	1.67	8.65
Air	0.08	6.77	0.20	7.95	1.72	11.21

Table 7.4: The PØD reconstructed vertex position resolution [11].

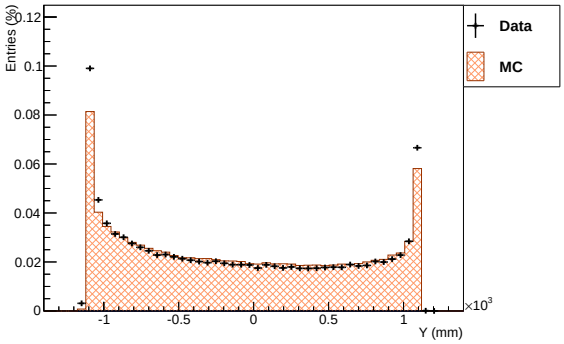
All the PØD standalone studies use only PØD fully contained events. In order to ensure that all the particles from the interaction leave their signature within the PØD, a fiducial volume (FV) cut is applied, see Table 7.5 for the definition of the FV. All the reconstructed interaction vertices are required to be within this FV. However, the surrounding PØD-ECal modules can be used to enlarge teh FV, hence the statistics including events which have a signature in these modules. Therefore, in this thesis, the PØD fiducial volume will be used as a point of reference to the enlargement in order to show the impact of an extended interaction area including the PØD-ECal.



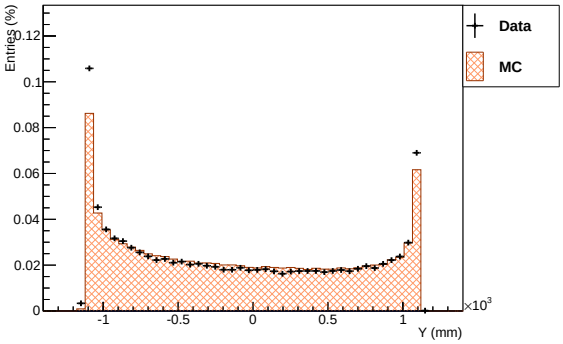
(a) Water



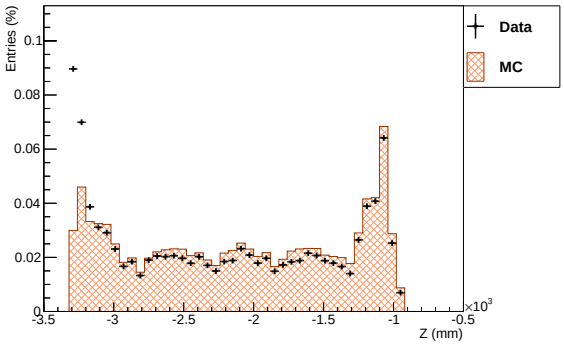
(b) Air



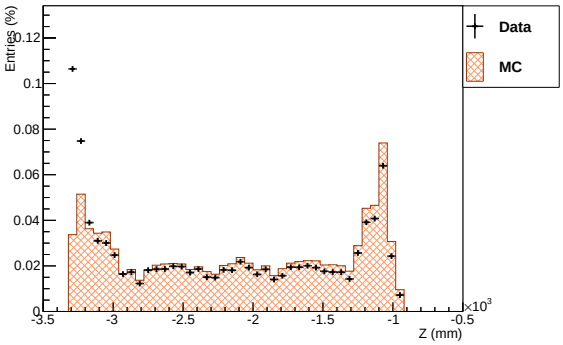
(c) Water



(d) Air

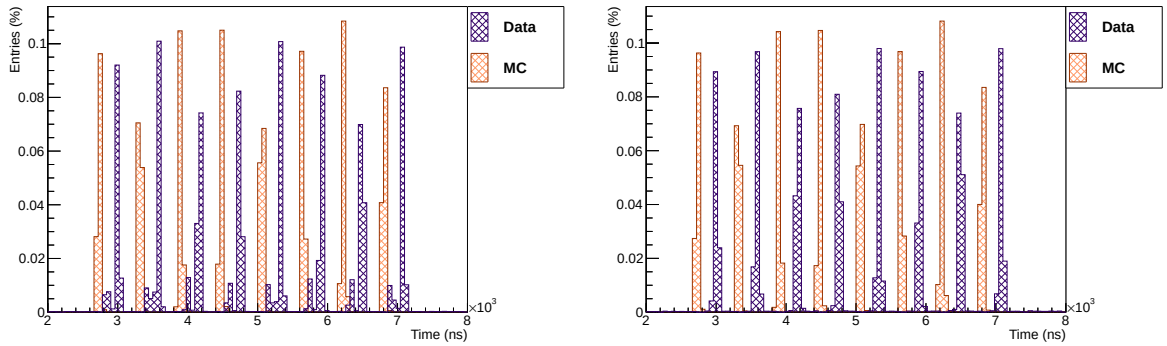


(e) Water



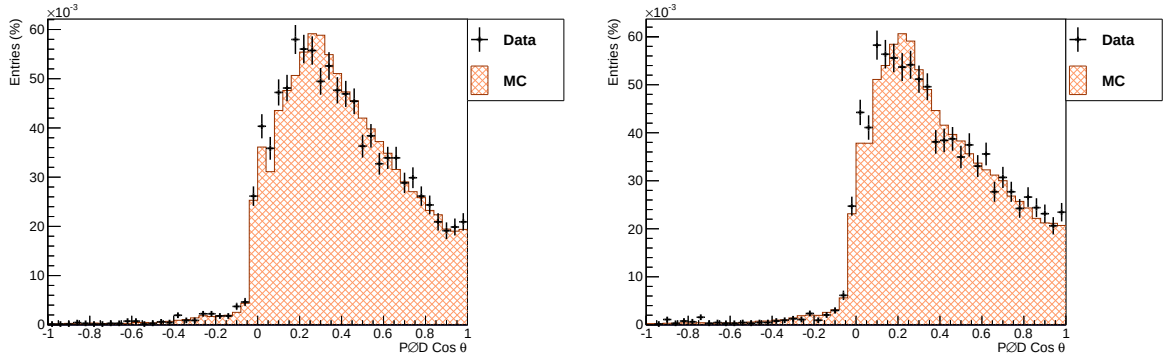
(f) Air

Figure 7.6: Data and MC reconstructed position of the PØD shower objects, area normalised for the PØD filled with (plots on the left) and without (plots on the left) water.



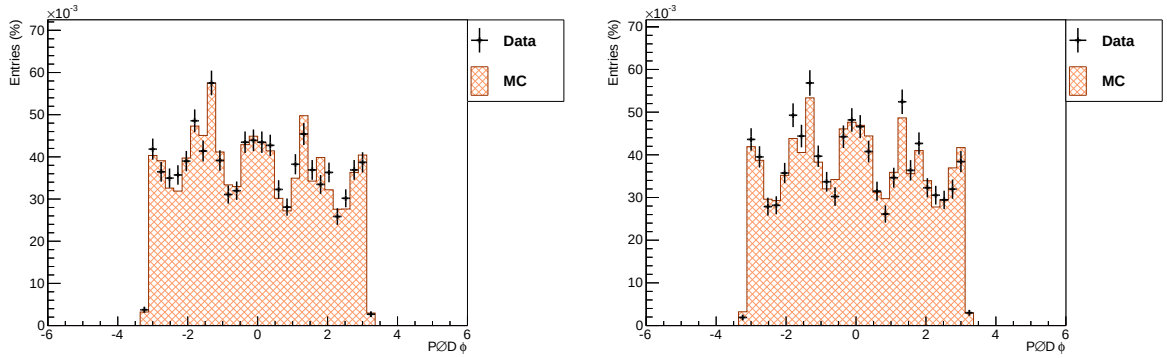
(a) Water (b) Air

Figure 7.7: Data and MC reconstructed time of the PØD shower objects, area normalised for the PØD filled with (a) and without (b) water.



(a) Water (b) Air

Figure 7.8: Data and MC of the reconstructed PØD shower direction in the spherical coordinate θ (the direction in which the angle is increasing from the positive z axis), area normalised for the PØD filled with (a) and without (b) water.



(a) Water (b) Air

Figure 7.9: Data and MC of the reconstructed PØD shower direction in the spherical coordinate ϕ (the direction in which the angle is increasing in the xy plane counterclockwise from the positive x axis), area normalised for the PØD filled with (a) and without (b) water.

Coordinate	Centre (mm)	Half Width (mm)	Min (mm)	Max (mm)
X	-36	800	-836	764
Y	-1	870	-871	869
Z	-2116	852.5	-2969	-1264

Table 7.5: The PØD standalone analysis fiducial volume definition in the global ND280 coordinate system.

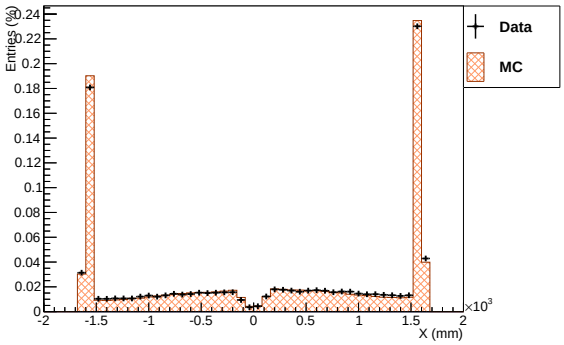
7.1.3 PØD-ECal Object in the NC1π^0 Event

In this thesis, only events with a reconstructed PØD vertex and showers in the PØD and PØD-ECal are considered. All the PØD-ECal reconstructed objects used in the NC1π^0 analysis are clusters, which are identified as either track or shower. Their physical features utilised in this analysis are the position, X and Y, time, T, and the number of clusters, tracks and showers.

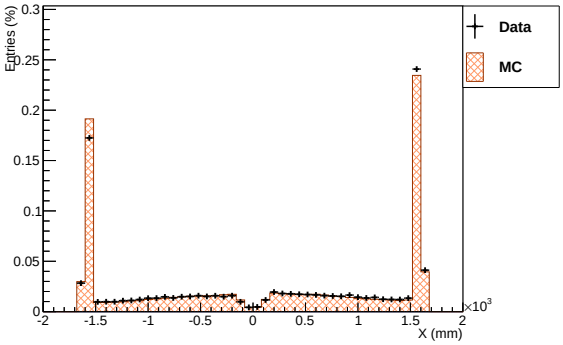
The data and MC comparisons of the position and time variables, area normalised, are presented in Figure 7.10, where on the left hand side are plots for the PØD filled with water and on the right hand side plots for the PØD without water. These plots show a good agreement on the position variables. The timing plots data distribution have a wider distribution at the peak tail. The lack of the time calibration application contributes to the time peaks behaviour. Unfortunately, this indicates that in the consequence of uncalibrated time, the matching accuracy between the PØD-PØD-ECal clusters may be affected when selecting the NC1π^0 events. Furthermore, the data for the PØD water configuration is affected by the Run 2 data, similarly as seen in the PØD.

Figure 7.11 shows the number of reconstructed clusters (tracks and showers) from data and MC files. Unusually, the data-MC ratio changes its behaviour depending on the number of reconstructed clusters. In the case where only one cluster is found, a MC excess can be seen. For the higher number of reconstructed clusters, a MC deficiency occurs with respect to the data, which indicates a cluster fragmentation tendency for the data reconstructed objects. The probable reason of the cluster split is related to the clustering procedure criteria. Regarding the uncalibrated time situation, it is likely due to the time requirement, which demands the clustered hits to be within 15 ns time window from the seeding hit. Therefore, the clustering procedure should be investigated.

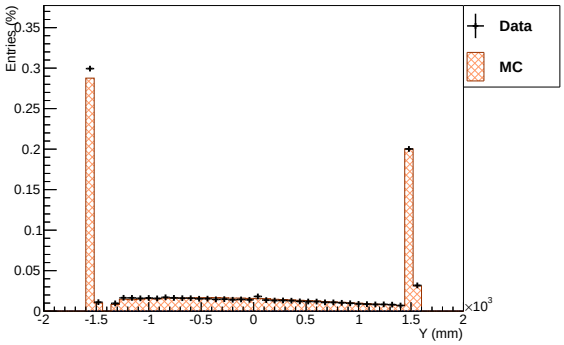
As explained earlier, the particle in the PØD-ECal can be identified as a track or shower. Therefore, in terms of signal definition, there is no possibility



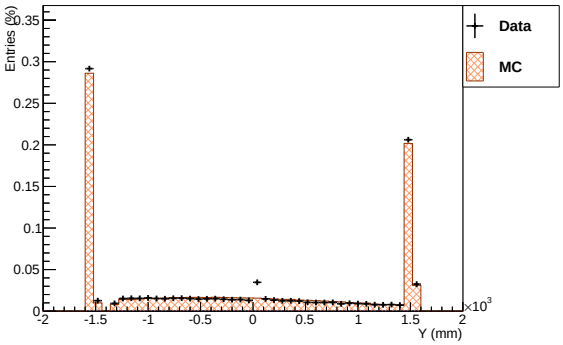
(a) Water



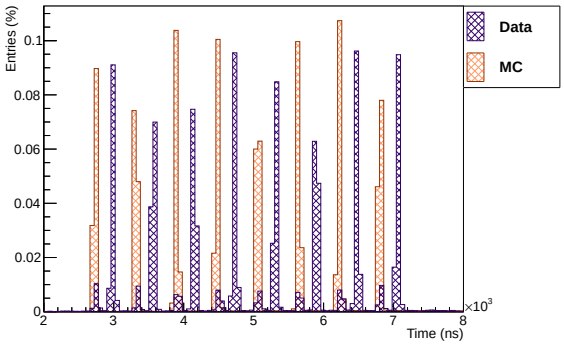
(b) Air



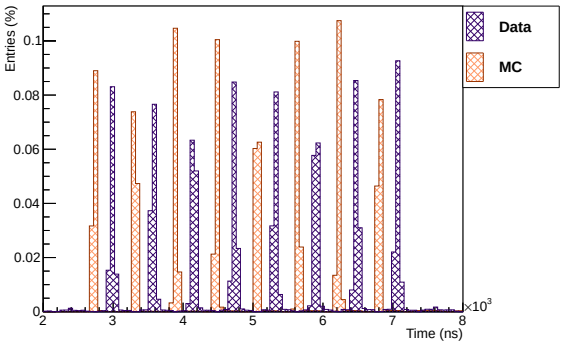
(c) Water



(d) Air

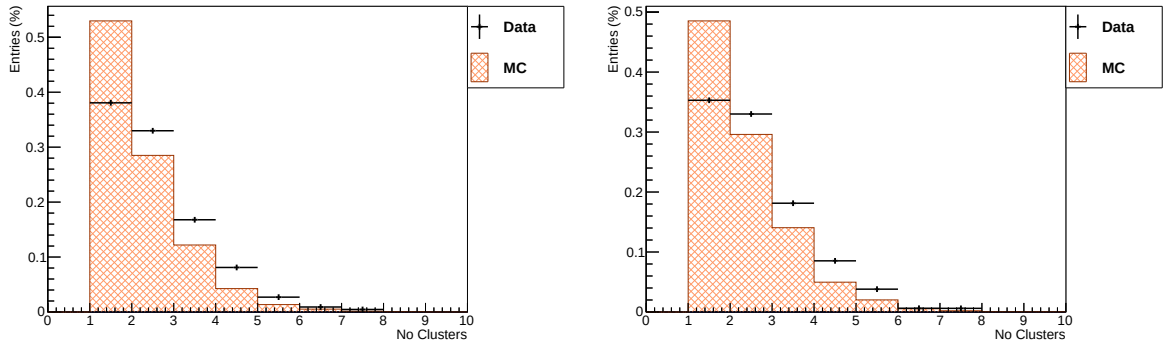


(e) Water



(f) Air

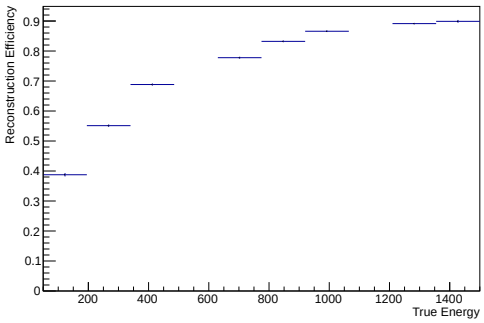
Figure 7.10: Data and MC comparison of the reconstructed X and Y PØD-ECal cluster position and cluster time T for the PØD-ECal objects normalized to the events in the PØD filled with (plots on the lefts) and without (plots on the right) water.



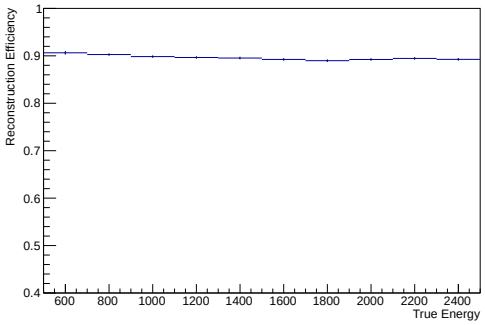
(a) Water (b) Air

Figure 7.11: Data and MC comparison for the number of reconstructed clusters for the PØD-ECal objects normalized to the events in the PØD filled with (plots on the left) and without (plots on the right) water.

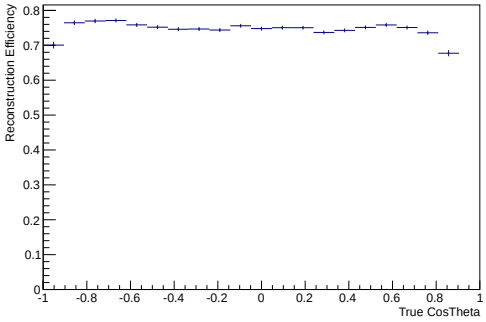
to distinguish an EM shower from another type of shower or a proton particle from other tracking particles. The PØD-ECal PID optimisation and efficiency studies made at the reconstruction level were done using the *particle gun* type of MC. 300 000 true muon and true photon particles were fired in front of each of the 6 PØD-ECal modules and assigned as tracks or showers. The track-like and shower-like PID efficiency against the true energy and true angular distribution are shown in Figure 7.12. It can be seen that the efficiency of the track-like particle reconstruction for the angular distribution is up to 20% higher than for the shower-like particles. However, the efficiency value is above 70% for both types of PID. As for the energy dependence, the efficiency of the shower-like particles reconstruction is increasing from about 40% for the lowest energy bin up to 90% for energies above 1300 MeV. The track-like reconstruction energy dependence has an opposite behaviour and it decreases from above 90% to 88%. However, this good PID efficiency is not transformed into the *magnet* files, where the PID studies of the PØD-ECal reconstructed clusters show entirely different results for the signal NC1π^0 and background events, see Figure 7.13 for the signal (top plots) and background (bottom plots) events. The plotted events were selected in such a way that each of them contains one reconstructed vertex and one reconstructed EM shower within the PØD sub-detector. Therefore, by looking at the signal events, it is expected to find one PØD-ECal reconstructed object with a shower PID. However, it can be seen that the number of showers and tracks for the signal diverges from the expectations. The majority of the signal events, 47%, have one reconstructed track. The next significant events have one reconstructed



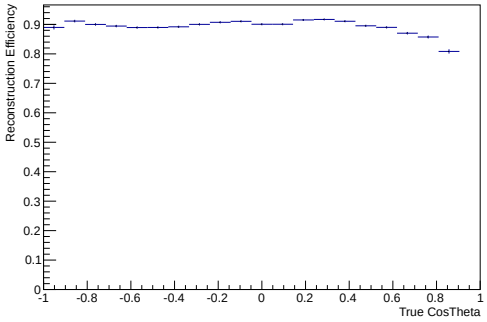
(a) Shower



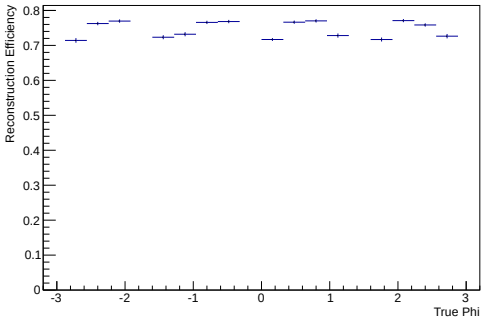
(b) Track



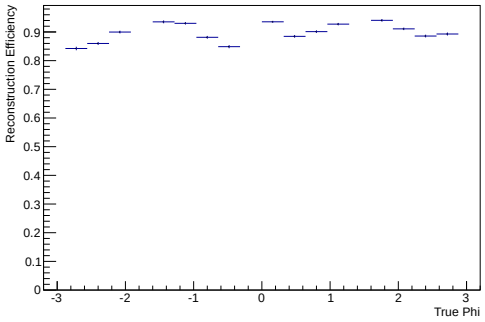
(c) Shower



(d) Track



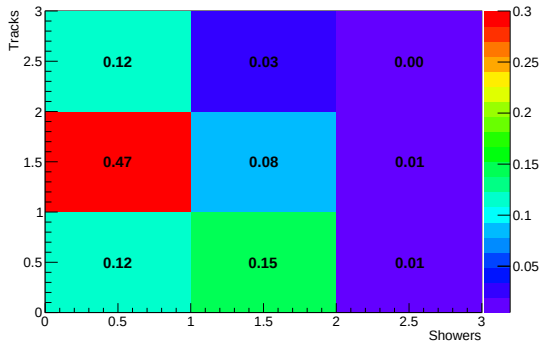
(e) Shower



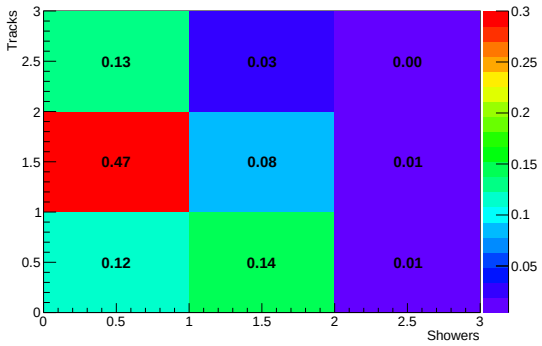
(f) Track

Figure 7.12: The PØD-ECal PID efficiency dependency on the true energy and angles for the track-like and shower-like particles made on the reconstruction level during PID MLP training optimisation.

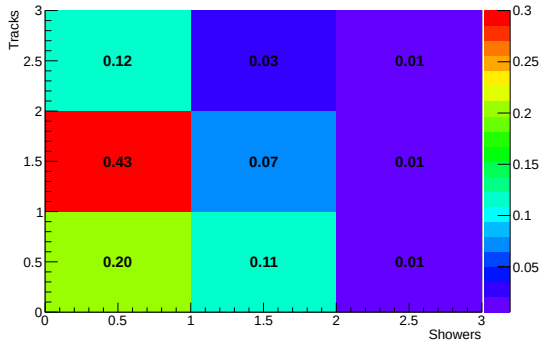
shower, 15%, and ones thereafter contain two reconstructed clusters, $\sim$ 21%, which includes the different types of PID configurations. This shows a significant discrepancy with expectation of the one reconstructed shower and with respect to the PØD-ECal PID efficiency studies made at the reconstruction level. This can be directly related to the different type of MC files used for the PID training and analysis. Therefore, taking into consideration the unsatisfactory PØD-ECal PID input to the analysis, the only requirement used in the selection will be the presence of one or two reconstructed clusters. However, the candidate cluster PID charge and geometric discrimination variables used in the PID training will be used later in the analysis as an input to the multivariate analysis cut for the signal-background separation in the NC1π^0 selection.



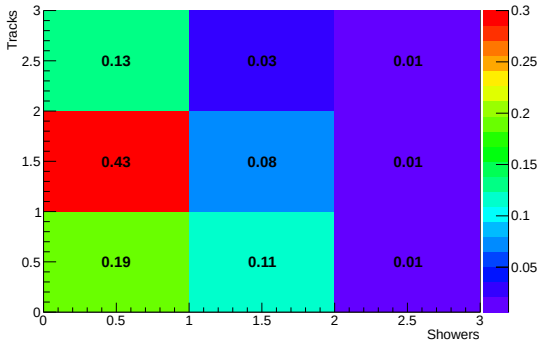
(a) Water



(b) Air



(c) Water



(d) Air

Figure 7.13: The fraction of the PØD-ECal PID reconstructed objects: tracks and showers within the NC1π^0 analysis, for the signal (top plots) and background (bottom plots) events for the PØD filled with (plots on the left) and without (plots on the right) water. All the events were required to have a reconstructed 3D vertex and at least one shower within the PØD sub-detector.

The energy and momentum of the PØD-ECal clusters are not reconstructed due to a limited reconstruction information availability. The attempt to reconstruct the energy with the algorithm developed for the Tracker ECal modules, called likelihood fitter, was unsuccessful.

7.2 NC1π^0 Event Selection

The signature of the signal is two reconstructed EM shower objects, one from the PØD and one from the PØD-ECal. Both showers are candidates of the π^0 decay photons originating from the NC1π^0 interaction in the PØD sub-detector. In order to identify such events, a selection has been developed, where in the final stage, a Boosted Decision Tree (BDT) method is implemented to separate the signal and background events. The NC1π^0 selection flow chart is shown in Figure 7.14. All the cuts are applied on an event-by-event basis, or in the case of the pØdRecon, it can be considered as a cycle-by-cycle basis. The motivation and optimisation of each selection cut are discussed in the following sections.

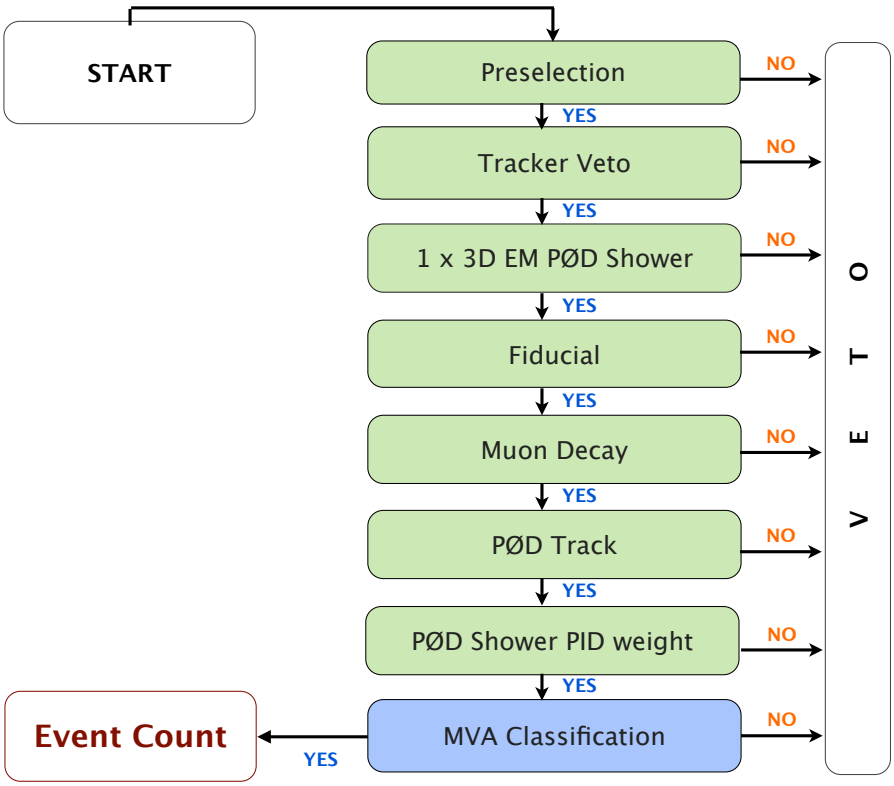


Figure 7.14: Event selection diagram for the PØD-PØD-ECal NC1π^0 analysis. The initial cuts (green background) are using the PØD and tracker sub-detectors information. The last cut (blue background) considers PØD-ECal reconstructed objects.

7.2.1 Pre-selection

The first step is a pre-selection which performs an initial choice of events that contain a 3D reconstructed interaction vertex within the PØD. In order to obtain the cleanest set of reconstructed events, each of them accommodates a single reconstructed vertex only within one timing cycle, with an outgoing shower object. Such vertex distribution is shown in Figures 7.15 for the water-in configuration and 7.16 for the water-out configuration, displaying the MC (plots on the left) and data (plots on the right).

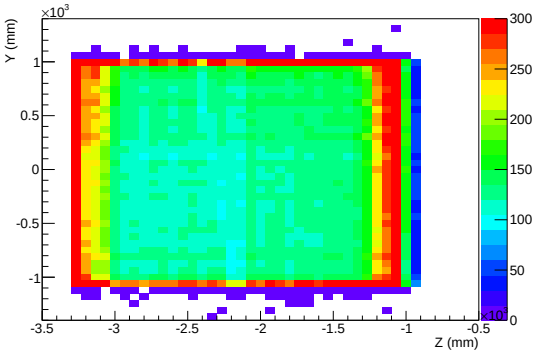
At this stage the additional condition is the presence of the PØD-ECal activity. Further requirement concerns the data, where a good flag of the beam status, PØD and PØD-ECal sub-detectors and Magnet is required. This selects events recorded during a stable mode of running beam and ND280.

After the pre-selection the fraction of reconstructed showers in the PØD and clusters in the PØD-ECal is presented in Table 7.6 for the signal events, and in Table 7.7 for the background events. For the signal interactions the highest fraction of events, 42% (41%), for the water-in (water-out) configuration, has one PØD shower and one PØD-ECal cluster reconstructed. The other significant configuration has one PØD shower and two PØD-ECal clusters, which constitutes around 14% for both, the water-in and water-out PØD configurations. Therefore, the above two cluster number configurations are considered as the π^0 gamma candidates and will be taken into account during NC1π^0 event selection.

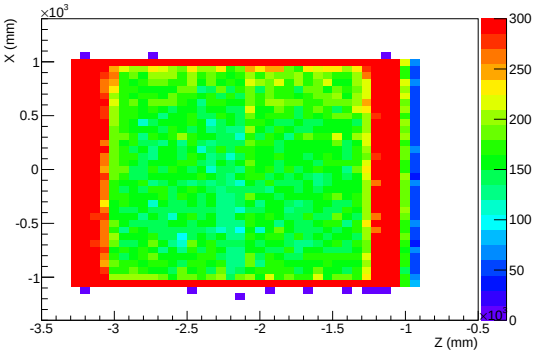
The sand events have been rejected at the pre-selection stage for both the water-in and water-out PØD configurations. This is strongly related to the limited availability of the simulated *sand* files which results in a low statistics.

7.2.2 Veto

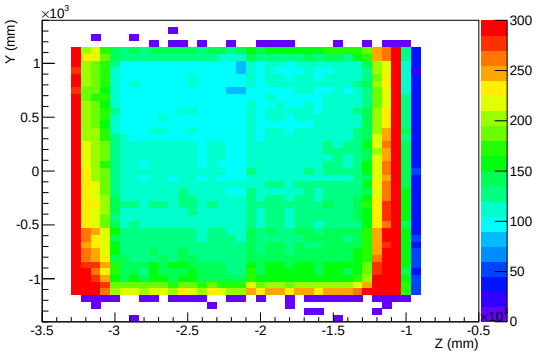
Next, a veto cut is applied using other sub-detectors, mainly from the Tracker region. The considered π^0 candidate topology requires all the events with a detected activity in the TPC1, FGD1, Tracker ECal modules and SMRD to be rejected. This provides the best possible reconstruction performance as no produced particle information is leaked to these sub-detectors. The veto cut removes 33% of the background and 15% of the signal interactions for both PØD water configurations.



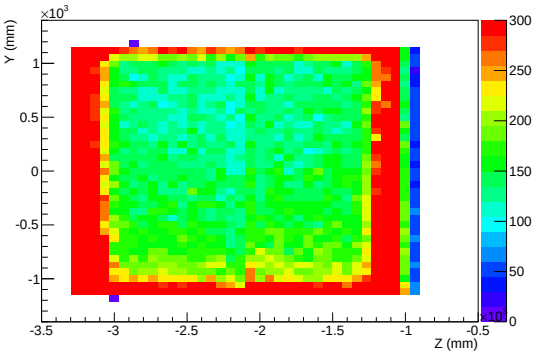
(a) MC



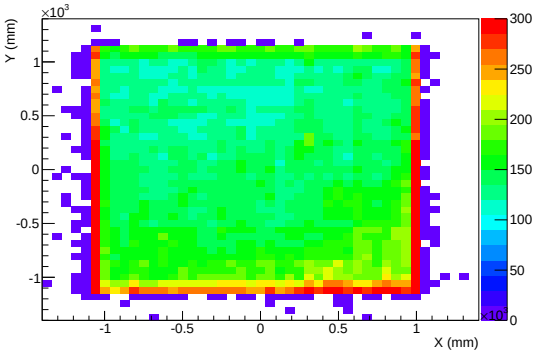
(b) Data



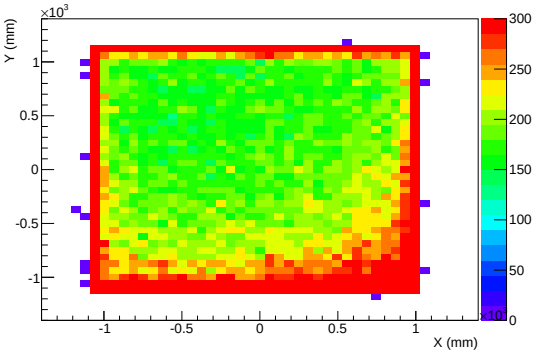
(c) MC



(d) Data

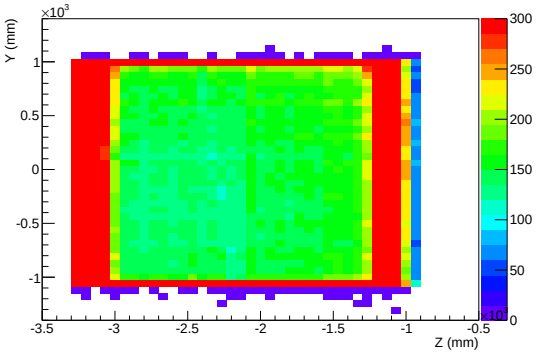


(e) MC

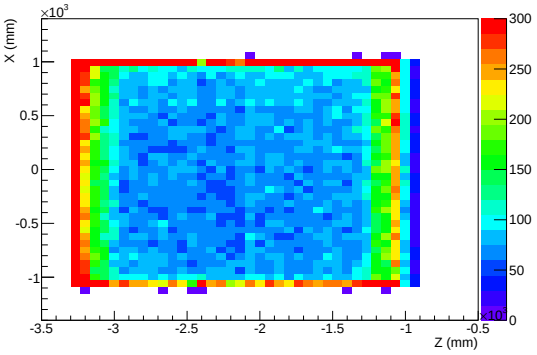


(f) Data

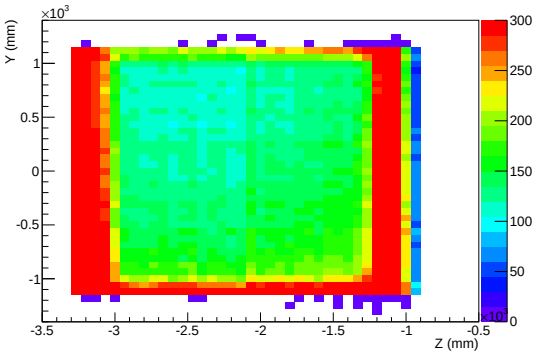
Figure 7.15: The reconstructed vertex position for the Monte Carlo (plots on the left) and data (plots on the right) comparison for the PØD filled with water.



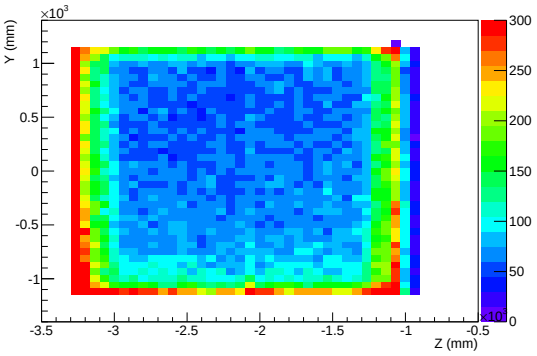
(a) MC



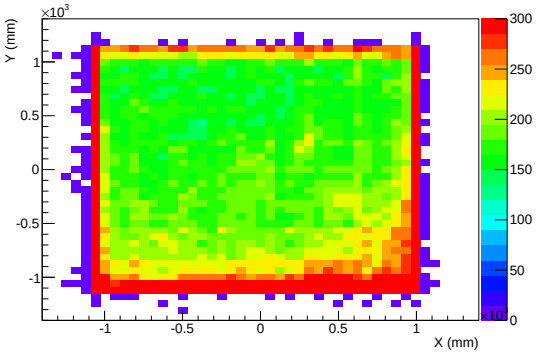
(b) Data



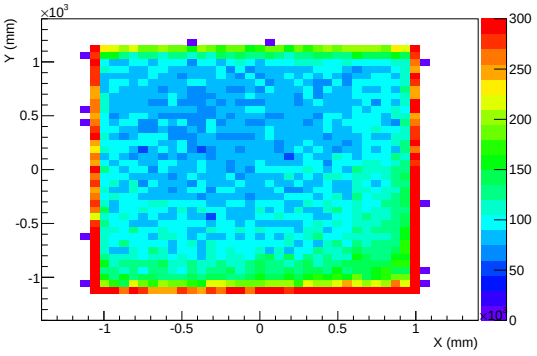
(c) MC



(d) Data



(e) MC



(f) Data

Figure 7.16: The reconstructed vertex position for the Monte Carlo (plots on the left) and data (plots on the right) comparison for the PØD without water.

Reconstructed object		Number of PØD-ECal clusters			
		0	1	2	3+
WATER					
Number of	1	0.21	0.42	0.14	0.05
PØD	2	0.04	0.06	0.03	0.01
Showers	3+	0.01	0.01	0.01	0.00
AIR					
Number of	1	0.19	0.41	0.14	0.06
PØD	2	0.04	0.08	0.03	0.01
Showers	3+	0.01	0.01	0.01	0.00

Table 7.6: Fraction of the number of reconstructed PØD showers and PØD-ECal clusters in the signal after the pre-selection stage of the NC1π^0 event selection.

Reconstructed object		Number of PØD-ECal clusters			
		0	1	2	3+
WATER					
Number of	1	0.34	0.23	0.08	0.04
PØD	2	0.10	0.06	0.03	0.01
Showers	3+	0.04	0.03	0.02	0.01
AIR					
Number of	1	0.32	0.23	0.08	0.04
PØD	2	0.09	0.07	0.03	0.01
Showers	3+	0.04	0.03	0.02	0.01

Table 7.7: Fraction of the number of reconstructed PØD showers and PØD-ECal clusters in the background after the pre-selection stage of the NC1π^0 event selection.

7.2.3 PØD Object Selection Criteria

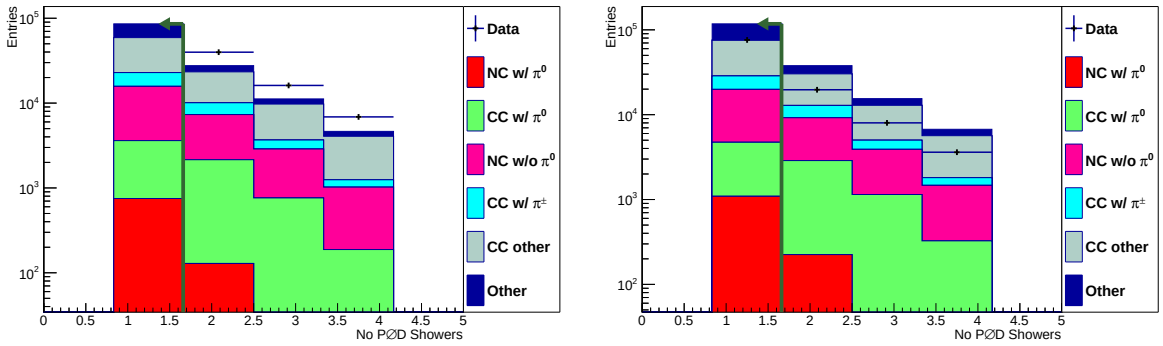
The next five selection criteria are related to the PØD reconstructed objects, where two of the cut values are optimised by and taken from the PØD standalone analysis. Each cut distribution is plotted with previous cuts applied to it. Also, all the selection variable distributions are plotted on the logarithmic scale to emphasize the signal, which is hardly visible otherwise.

The first PØD object requirement applies to the number of 3D EM shower objects reconstructed by the pØdRecon package. All events with more than one PØD shower are rejected, see Figure 7.17. This cut removes 14% (17%) of signal and 23% (24 %) of background events for the PØD filled with (without) water.

Afterwards, the selection criterion on the fiducial volume variable is applied. This cut value has been determined in correlation to the last selection cut, BDT-cat, due to the existing interdependence between them. This correlation is discussed in detail, illustrated and the cutting value choice explained in Appendix B. The distribution of the FV is shown in Figure 7.18 and its value defines the shortest vertex distance to the nearest fiducial volume edge. All the vertices with a negative FV value lay outside the fiducial volume border and those with a positive value are within the fiducial volume. The presence of the surrounding PØD-ECal modules allowed the original PØD fully contained analysis fiducial volume to be extended, as defined in Table 7.5, by 74 mm for the water-in and by 79 mm for the water-out PØD configurations. This enlarged the neutrino interaction volume by 26.6% for the PØD filled with water and by 29.4% for the PØD without water.

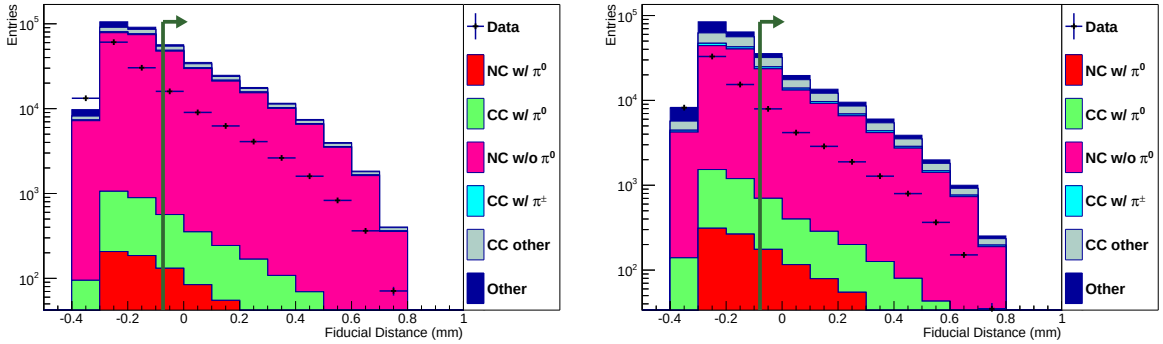
The next selection cut concerns the very low energy muons, which form reconstructed clusters in the way that can be easily misidentified and contribute to the background. The timing cycle difference for the cluster and vertex timing limits the side effect of the TFB electronics simulation. There is a specially developed algorithm which tags muon decay clusters within the PØD. When such decay cluster is identified within the cycle other than the reconstructed vertex, the event is rejected. The number of decay clusters is shown in Figure 7.19. This cut removes 2% (2%) of signal and 4% (3%) of background events for the water in (water out) PØD configuration.

The number of PØD reconstructed tracks in the event at this stage of the selection is shown in the Figure 7.20. Due to the signal definition, tracks originating from the nucleons are also included in the selected sample. Taking into account the track reconstruction efficiency studies presented and discussed in section 7.1.2, all the events with at least one reconstructed track are rejected. The



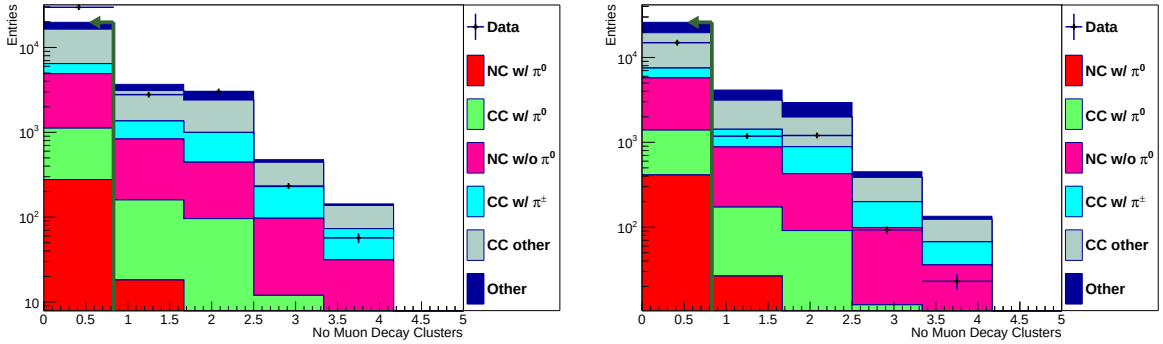
(a) Water (b) Air

Figure 7.17: The distribution of the number of PØD 3D EM reconstructed showers within one event, P0T normalised for PØD filled with (a) and without (b) water. The cut value represents the requirement of one shower for both water configurations.



(a) Water (b) Air

Figure 7.18: The fiducial volume distributions, P0T normalised. All the events placed on the left hand side from zero (negative distance value) are outside the PØD defined FV. To pass the cut the FV value has to be greater than -74 mm for the PØD filled with (a) and greater than -79 mm without (b) water.

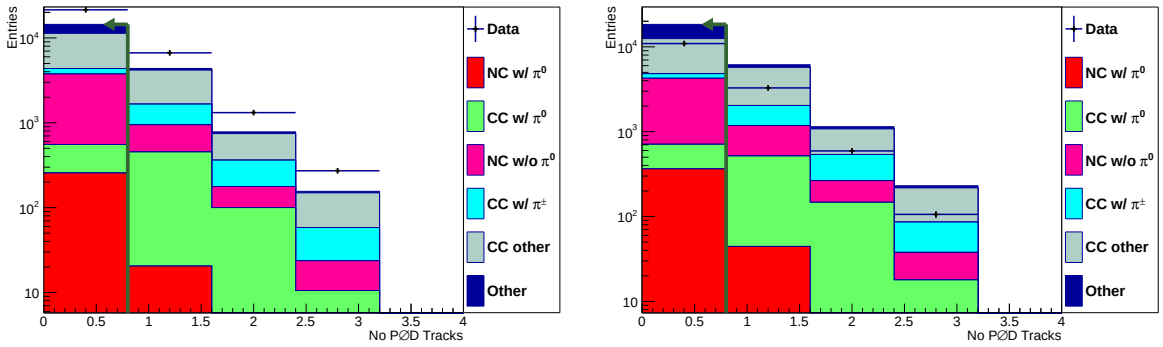


(a) Water (b) Air

Figure 7.19: The number of muon decay clusters for the PØD filled with (a) and without (b) water.

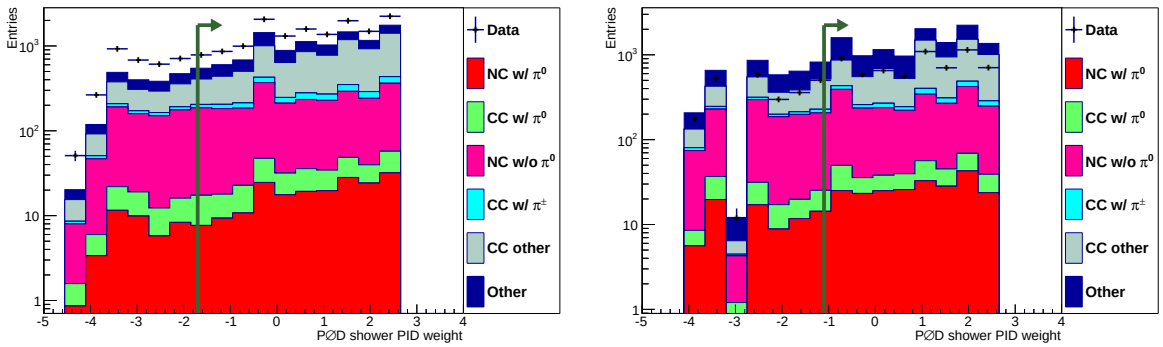
PØD track cut removes 2% (3%) of signal and 3% (3%) of background events for the water in (water out) PØD configuration.

Afterwards, all the remaining PØD reconstructed objects are showers and the PØD shower PID weight cut is applied. The pØdRecon particle identification algorithm assigns to the reconstructed shower two possible PID values: EM and Other. Figure 7.21 shows the distribution representing the differences of the log likelihoods for both shower PIDs. This difference is a cutting parameter where value is set to -1.7 for the PØD filled with water and -1.1 for the PØD without water.



(a) Water (b) Air

Figure 7.20: The number of reconstructed tracks in the PØD, POT normalised, only events without any track pass the selection for the PØD filled with (a) and without (b) water.



(a) Water (b) Air

Figure 7.21: The PØD shower PID weight distributions for the both PØD configurations, water-in and water-out.

At this point all the selected PØD objects can be matched with the PØD-ECal clusters to form a NC1π^0 candidate and separate the signal events from different backgrounds. The variables considered for the matching procedure do not allow a simple cutting value to be determined as these distributions do not have enough separation power if taken individually. However, when considered together, their correlation enables a good discriminating power for the signal and background events to be obtained. The simple timing and spatial relations between the PØD and PØD-ECal clusters turned out to have many challenges given the limited amount and accuracy of the reconstructed information from the pØdecalRecon, i.e. no calibrated time, no spatial Z direction (beam direction) and clusters fragmentation. Consequently, the benefits of matching technique based on the variables cut value appeared to be inadequate. Therefore the task to find an optimal solution for the clusters matching stage required an alternative technique such as the Multivariate Analysis (MVA).

7.2.4 PØD +PØD-ECal Object Selection

The purpose of the MVA method [78] application is to identify and separate the NC1π^0 signal from the whole. There is a number of available classifiers which can be used to separate the signal sample such as the Fisher discriminants, neural networks, decision trees etc. The latter two algorithms require a numerical optimisation called training, followed by the validation of a solution. Each of the methods has various advantages and is useful to separate rare type of events. The neural network is a pattern recognition algorithm that can be a multi-layer perception, which was already used for the PØD-ECal track-shower discrimination. The MLP is an assembly of layers of perceptions, typically with n dimensions (variables) used as an input and one dimension constituting an output, with hidden layers in between. The final output is a discriminant for the signal-background separation. The MLP procedure must be accurately tuned and trained to avoid over training. The alternative good separation method is a boosted decision tree. In this thesis analysis both methods were tested using a TMVA package [79] which is an integral part of ROOT. Additionally, the TMVA package has an option which allows for a supplementary categorisation of events within each technique, which is a convenient solution to account for a cluster fragmentation problem in the PØD-ECal reconstruction. Here, the generic classifier allows to split the training sample into two disjoint categories according to the number of PØD-ECal clusters: one or two, and the categorised methods are called BDTcat and MLPcat. The testing results are shown in Figure 7.22. The

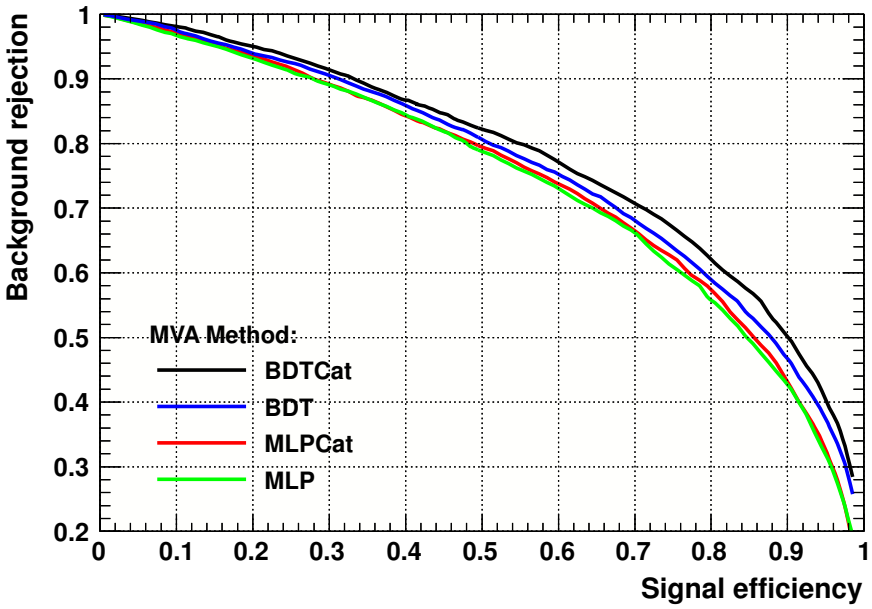


Figure 7.22: The signal efficiency curves for the four MVA methods: MLP, MLPcat with categorisation, BDT and BDTcat with categorisation.

performance of all the methods, MLP and BDT and their categorised versions, is comparable as for the signal efficiencies. However, since the BDTcat technique gives a slightly better separation and is more robust, it has been applied to the NC1π^0 selection.

Boosted Decision Tree

In the BDT [80] method the selection is based on the numerical superiority of the result coming from the number of decision trees. These trees are generated from the training sample by accumulation of the various event weights. The tree like structure is composed of a number of nodes, where each node is using one variable to make a decision. An example of the decision tree is shown in Figure 7.23. All the input events have a weight of 1 and are defined as a signal or background. These events contain a set of variables that can be used to discriminate one from the other. The training procedure starts with a root node which takes as an input the full event and selects the variable and corresponding cut criteria for the best separation. The splitting value is chosen to give the maximum separation between signal and background, which is determined using the Gini index, G,

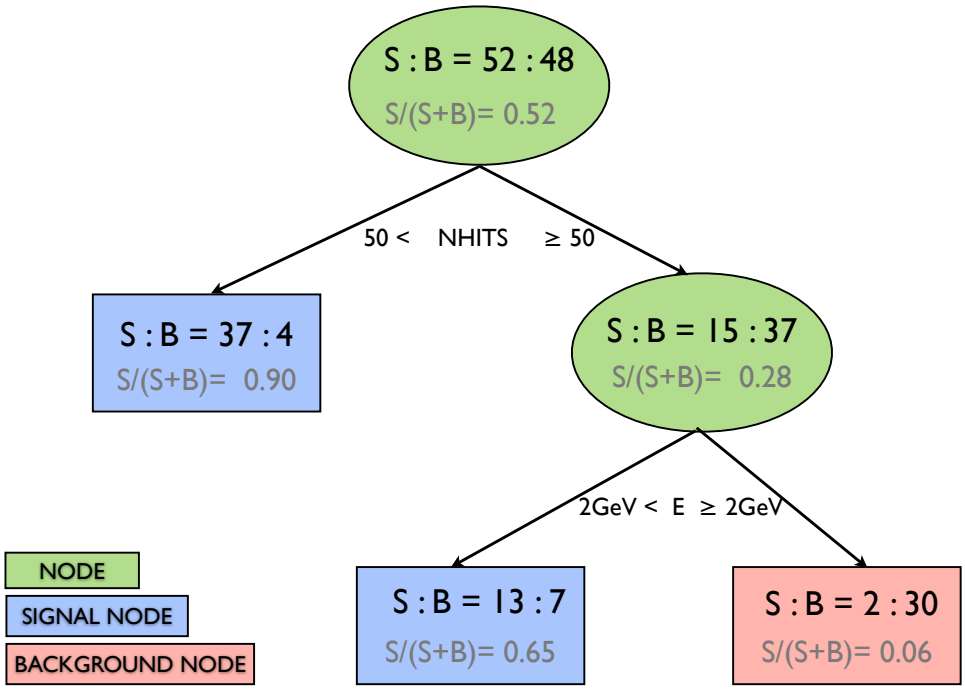


Figure 7.23: An example of a single decision tree structure. The sample events are an input to the root node and then forwarded to the next stage, until they reach the signal (blue) or background (red) nodes. The signal significance is shown for each node along with the cutting value.

defined as

$$G = \left(\sum_{i=1}^{N} W_i \right) P\,(1 - P)\,, \tag{7.1}$$

where W_i is the weight of the event i and P is the purity of the signal events in the current node. Then the sample is divided into two separate sub-samples, which are forwarded to the next nodes. For the maximal splitting the sum of the Gini index of each branch must be minimised. This is achieved by maximising the quantity C given by

$$C = G_{Parent} - (G_{Left} + G_{Right})\,, \tag{7.2}$$

where G_{Parent} is the Gini coefficient of the parent events, G_{Left} is the Gini coefficient of the events that fall to the left node of the splitting value and G_{Right} the Gini coefficient for the right node. The cutting value on the variable for each node is defined during the training process. This procedure can be continued until it ends up at the leaving node where the event is finally classified as a signal-like or background-like. This node is an end-point which contains a minimum number

of events or the maximum of the minimum signal purity. Finally, the node with the higher number of signal events is called the signal node and the node with the majority of background events is called the background node. Unfortunately, the decision trees are unstable due to the fact that a small fluctuations in the training sample or a small change in a setting parameters can create big differences in the decision tree result.

The decision tree stability can be enhanced by implementing the adaptive boosting. In this procedure all the signal events which ended in the final background sample are given a larger event weight value than those from the signal node. This can be done by the AddTree algorithm [80], which defines the error E_m of a m tree as

$$E_m = \frac{\sum_{i=1}^{N_{missclassified}} W_i}{\sum_{i=1}^{N_{total}} W_i,}, \tag{7.3}$$

where W_i is an event weight. The result is called a score of the tree α_m and is given by

$$\alpha_m = \beta \ln \frac{1 - E_m}{E_m,}, \tag{7.4}$$

where β is a constant defined by the user. As a result an event has a new weight of $W_i e^{\alpha_m}$. Next, the new decision tree training is repeated for the re-weighted events sample. This procedure can be iterated a number of times. The result is a set of decision trees called a forest. The gradient boosting works like a function expansion approach, where the parameters for each tree are determined by the minimization of a error function which is a binomial log likelihood for classification. The boosting algorithm is greedy, which means that only one tree is modified at a time and the final BDT classifier output T(x_i) takes into account all the boosting stages, which is given by

$$T\left(x_i\right) = \frac{1}{M} \sum_{m=1}^{M} \alpha_m T_m\left(x_i\right), \tag{7.5}$$

where x_i is the ith event, M is the number of trees, and $T_m(x_i)$ is equal to 1 for a signal node or -1 for a background node.

In this thesis the BDT algorithm parameters such as number of trees, nodes and β value were optimised to give the best outcome with a minimised overtraining effect. As a result, the number of trees was set to 800 with a maximum of 7 final nodes and $\beta = 0.5$. Additionally, the BDT algorithm was categorised for events with one PØD-ECal cluster and with two PØD-ECal clusters, which

are called BDTcat. The BDTcat performance was run and monitored with the TMVA package.

BDT Input Variables

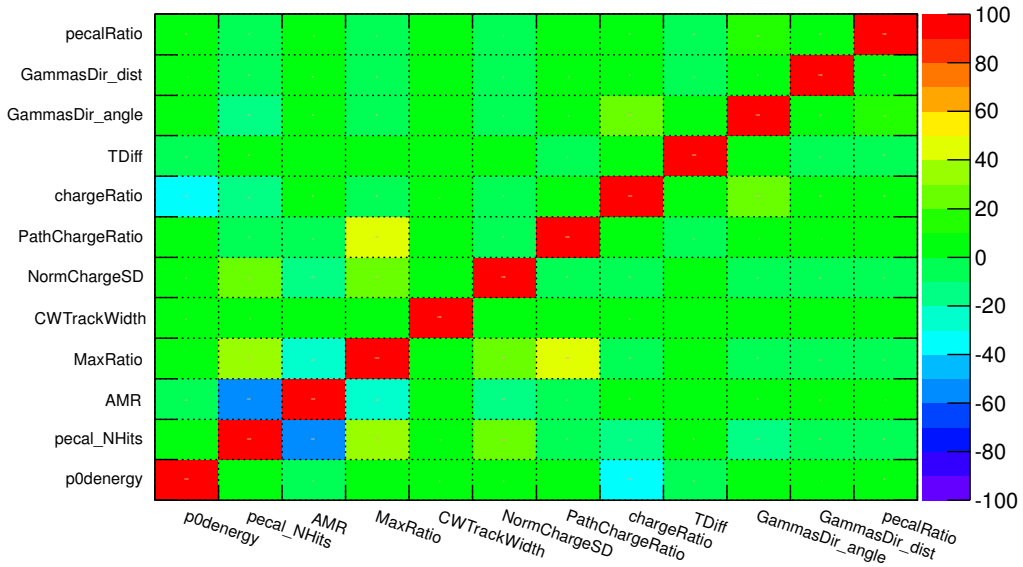
The input variables for the BDTcat classification were chosen on the basis of their power on the signal-background separation. The majority of the distributions varied for signal and background in a different way due to the correlation to other variables. The list of the input variables for the BDTcat training is shown in Table 7.8. There are no pre-MVA cuts applied to any of the input variables. In general the input variables can be classified into four categories: PØD cluster, PØD-ECal PID discrimination, PØD and PØD-ECal cluster relation and categorisation variables. The PØD -PØD-ECal matching variables are based on the timing, charge and geometrical relationship of the two clusters. After the previous selection cuts, only events with one or two PØD-ECal clusters passed to the last selection cut. In the case where the PØD-ECal has two reconstructed objects, the cluster with the highest contained charge is chosen as a space and time reference. The correlation matrix between all the MVA input variables for the background and signal events can be found in Figure 7.24 for the PØD filled with water and in Figure 7.25 for the PØD without water. It can be seen that there is typically a low correlation between the input variables for both PØD water configurations.

It is very important to have a good agreement between data and Monte Carlo generated events due to the large impact of any discrepancies on the systematic error. The area normalised plots of the BDTcat input variables can be found in Figures 7.26, 7.27, 7.28 and 7.29.

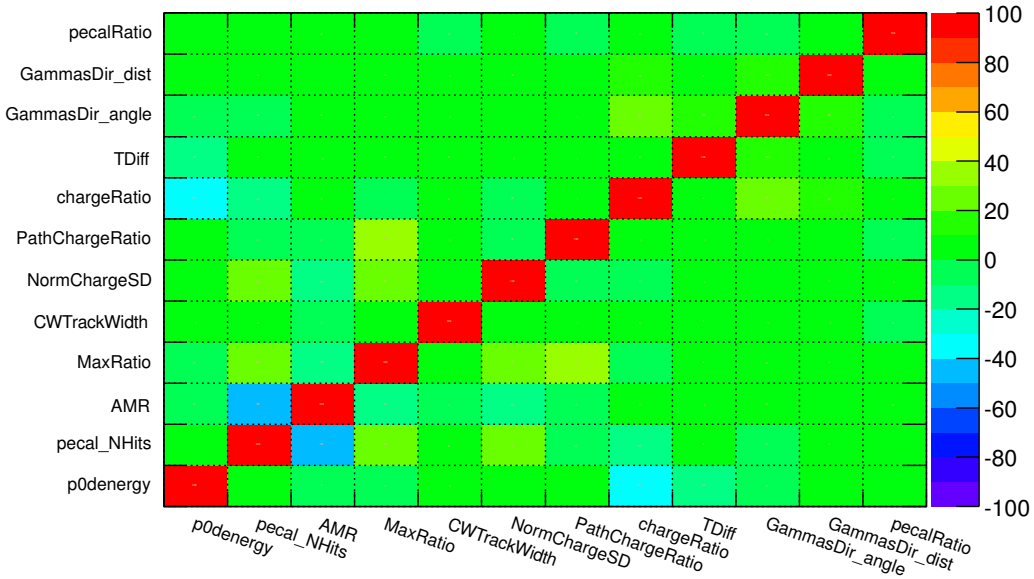
It can be seen that the majority of the distributions show an acceptable data-MC agreement. However, there are two discrimination variables, called gamma clusters direction angle GammasDir_angle (Figure 7.29 (a) and (b)) and gamma clusters direction distance GammasDir_dist (Figure 7.29 (c) and (d)), which show disagreement between data and MC, with the higher divergence for the PØD filled with water events. Both these variables represent the time-space relation of the PØD and PØD-ECal reconstructed clusters, serving as a cluster matching procedure and contribute to the signal background separation with a significant strength. Unfortunately, such discrepancy is expected to contribute significant addition to the systematic errors. The data MC discrepancy of these variables arises from the individual spatial and time components in their definition. An illustration of the geometry of two showering decay photons in the ND280 detector is shown in Figure 7.30.

No	Variable	Description
PØD cluster variable		
1	p0denergy	Reconstructed energy of the PØD 3D shower
PØD-ECal PID discriminating variables		
2a	pecal_NHits	Number of hits in the PØD-ECal cluster
2b	AMR	Ratio of major and minor axis of the cluster of hits, measures sphericity of the cluster
2c	MaxRatio	Ratio of charge in the layer with the largest and the smallest deposited charge
2d	CWTrackWidth	Charge weighted track width of a cluster is a number of bars in each layer multiplied by the charge deposited in the layer, summed up and divided by the total charge in the cluster
2e	NormChargeSD	Combination of two variables: Charge Standard Deviation divided by Mean Charge
2f	PathChargeRatio	Ratio of a charge deposited in a first two layers to the charge deposited in a two last layers
PØD -PØD-ECal clusters relation		
3a	chargeRatio	Sum of the PØD-ECal deposited charge and PØD deposited charge divided by the PØD shower reconstructed energy
3b	TDiff	Absolute time difference between the PØD shower and PØD-ECal cluster
3c	GammasDir_angle	Cosine angle between PØD -PØD-ECal cluster distance direction and the PØD shower direction
3d	GammasDir_dist	Angle between 2D projection (xy plane) of the PØD-PØD-ECal clusters distance and the PØD shower direction
BDT category variable		
4a	NClusters	Number of the PØD-ECal reconstructed clusters
4b	charge_Ratio	Charge ratio of the two PØD-ECal clusters

Table 7.8: Input variables for the BDTcat method.

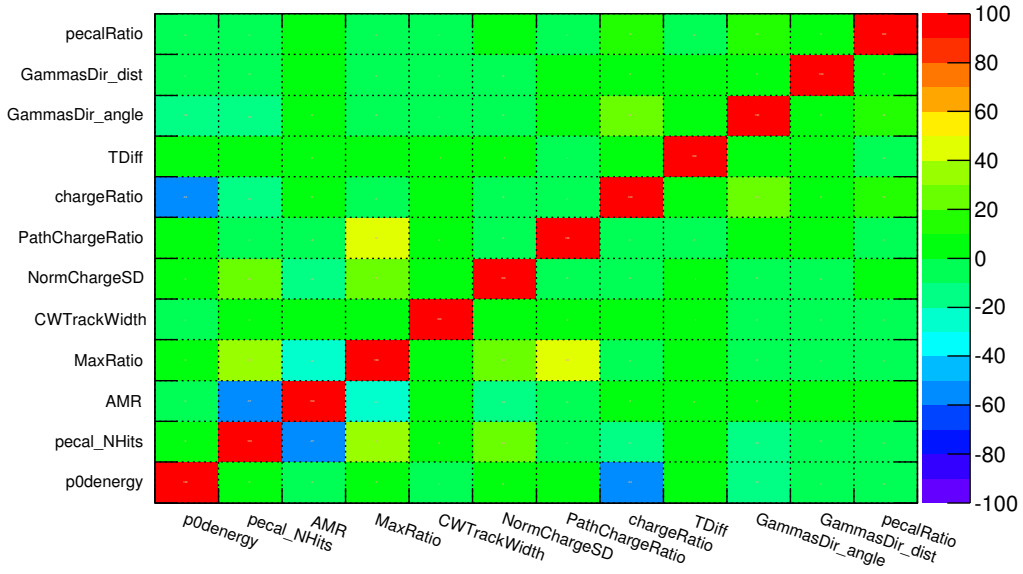


(a) Signal

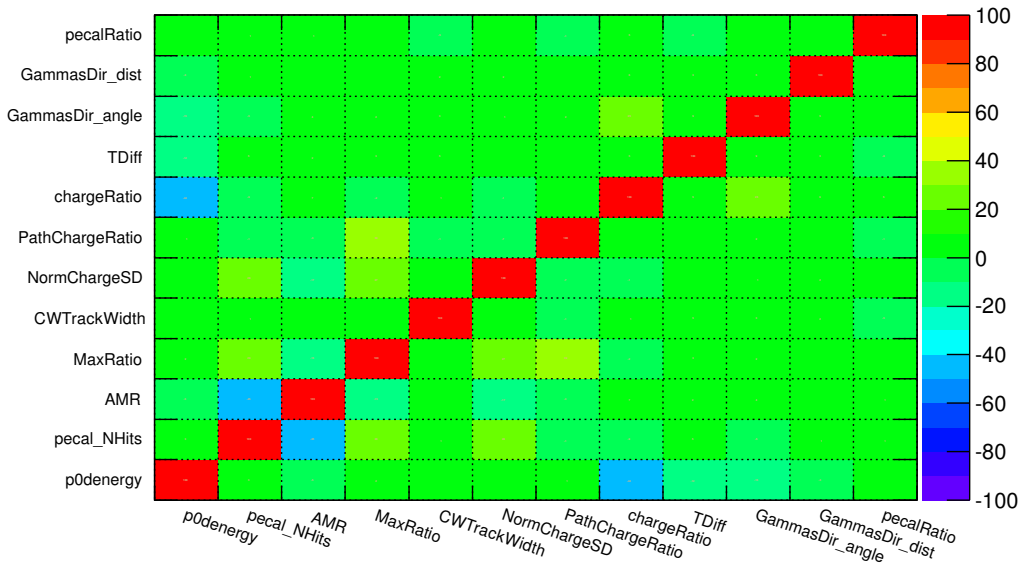


(b) Background

Figure 7.24: The correlation matrices of the BDTcat training variables for signal (a) and background (b) events for the PØD filled with water.

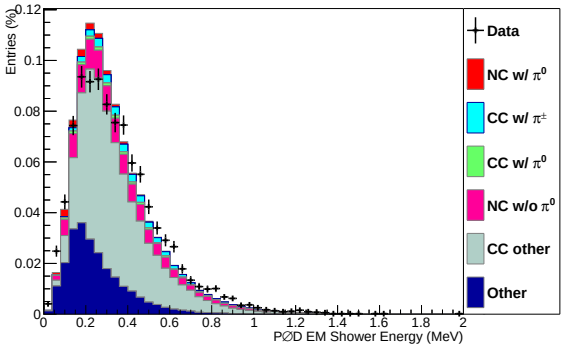


(a) Signal

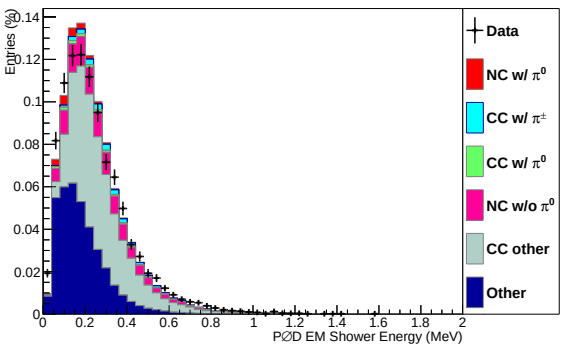


(b) Background

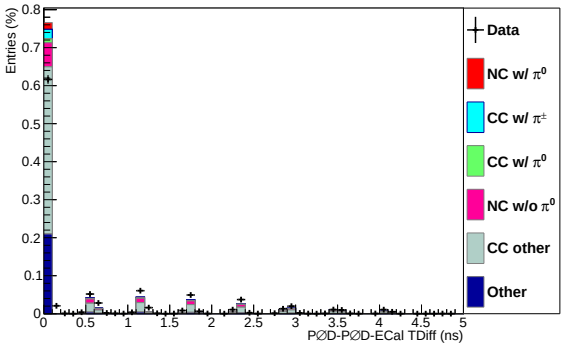
Figure 7.25: The correlation matrices of the BDTcat training variables for signal (a) and background (b) events for the PØD without water.



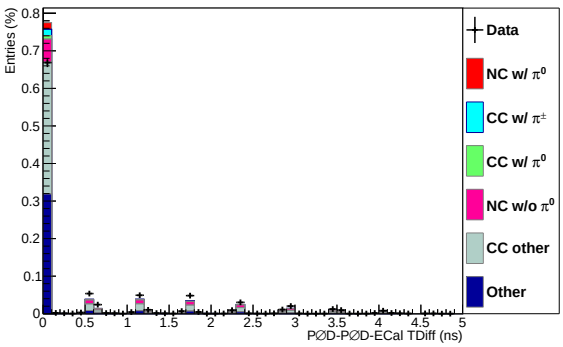
(a) Water



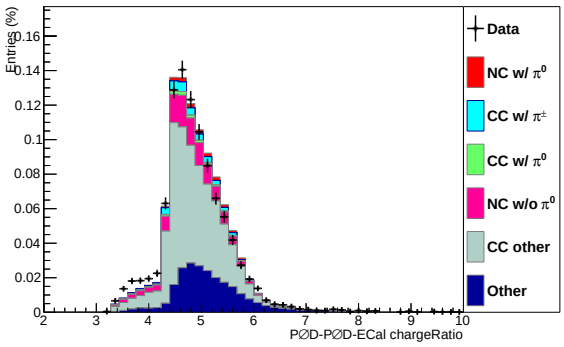
(b) Air



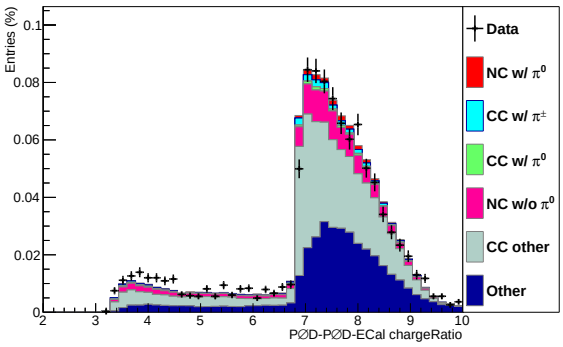
(c) Water



(d) Air

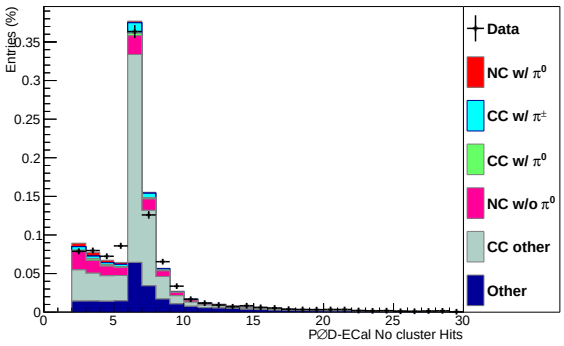


(e) Water

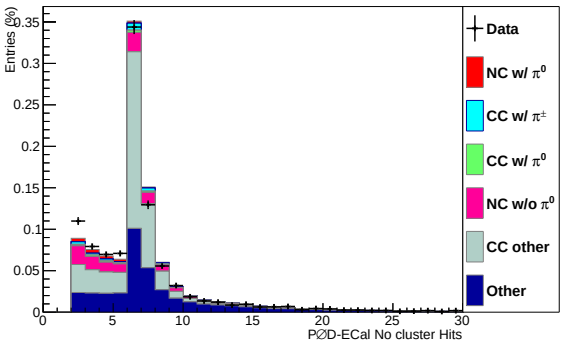


(f) Air

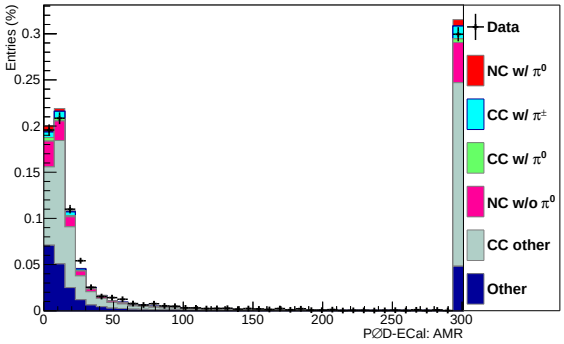
Figure 7.26: Data MC comparison of the BDTcat classification input variables for the PØD filled with (plots on the left) and without (plots on the right) water.



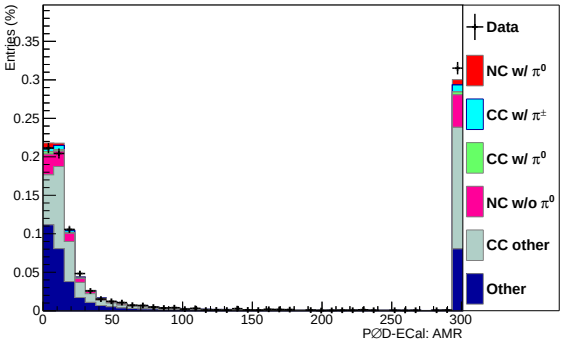
(a) Water



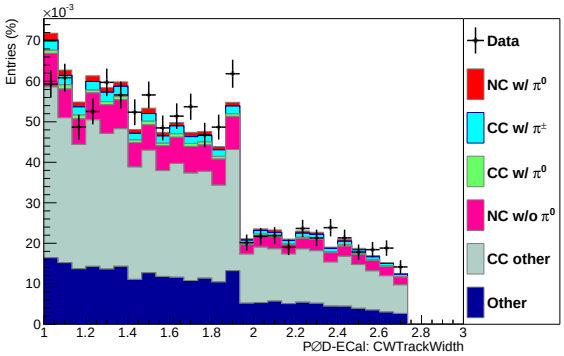
(b) Air



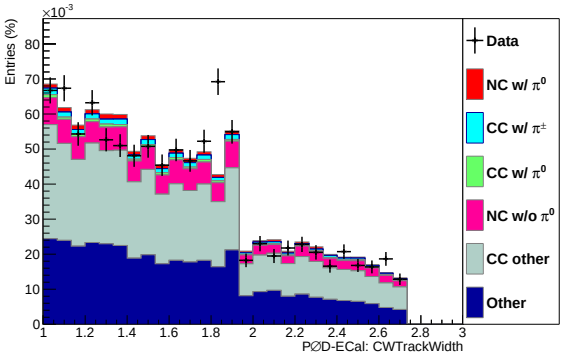
(c) Water



(d) Air

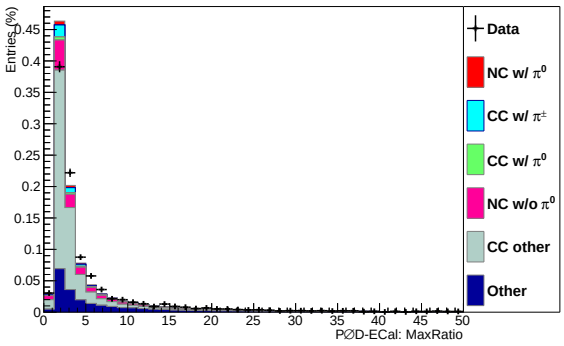


(e) Water

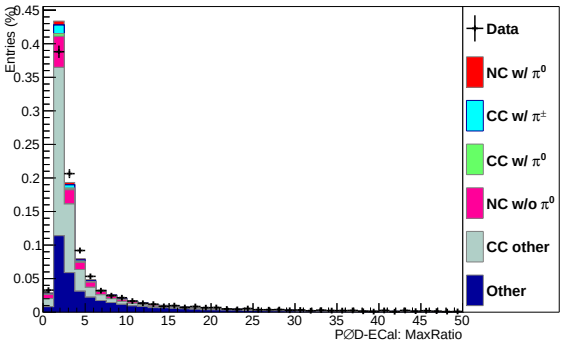


(f) Air

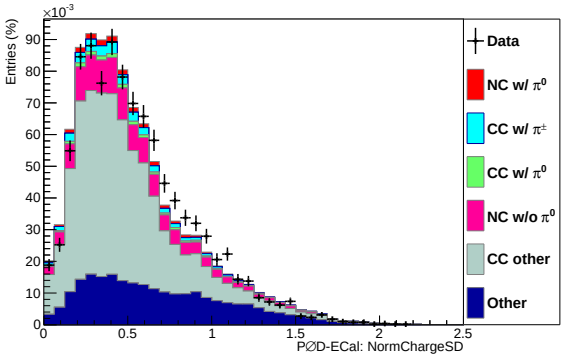
Figure 7.27: Data MC comparison of the BDTcat classification input variables for the PØD filled with (plots on the left) and without (plots on the right) water.



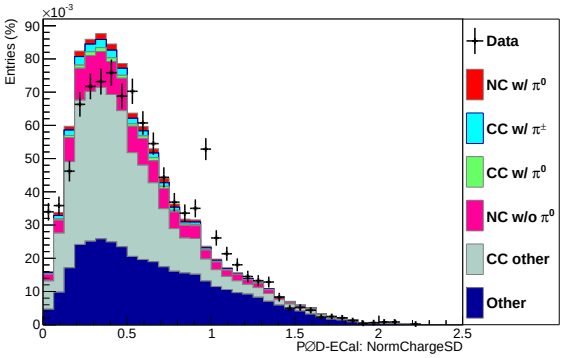
(a) Water



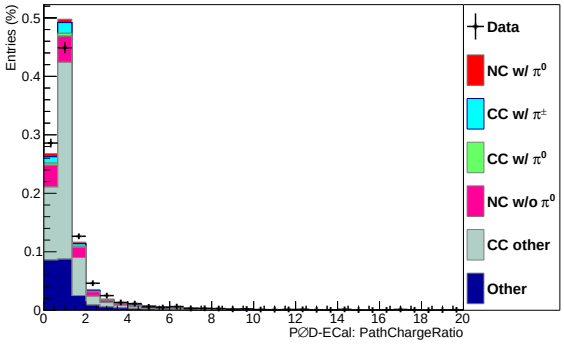
(b) Air



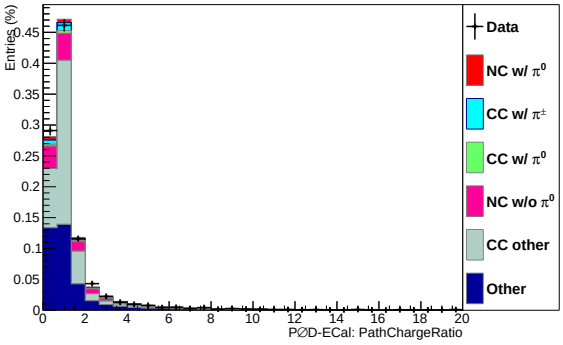
(c) Water



(d) Air

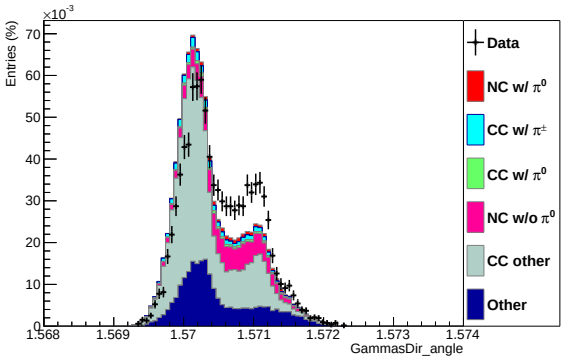


(e) Water

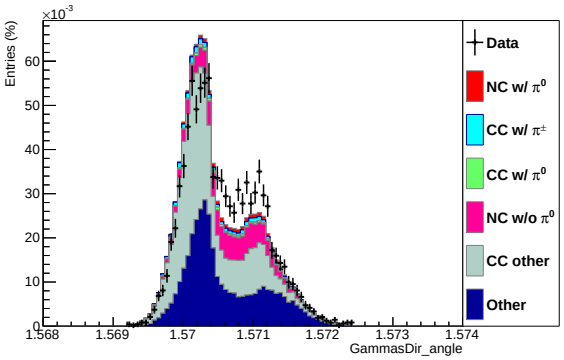


(f) Air

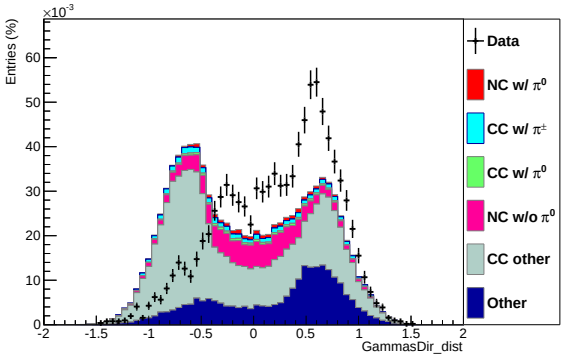
Figure 7.28: Data MC comparison of the BDTcat classification input variables for the PØD filled with (plots on the left) and without (plots on the right) water.



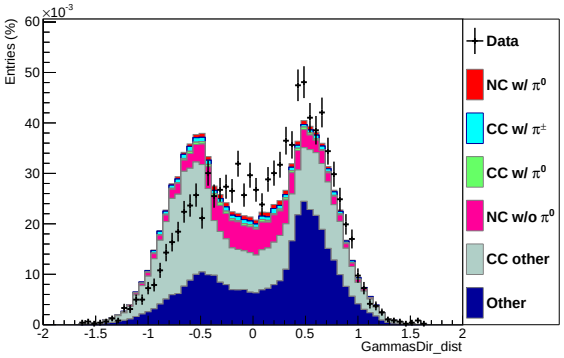
(a) Water



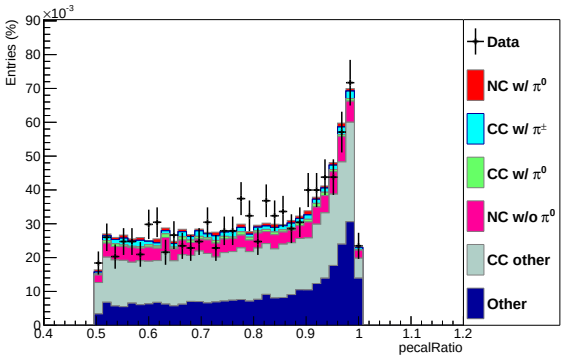
(b) Air



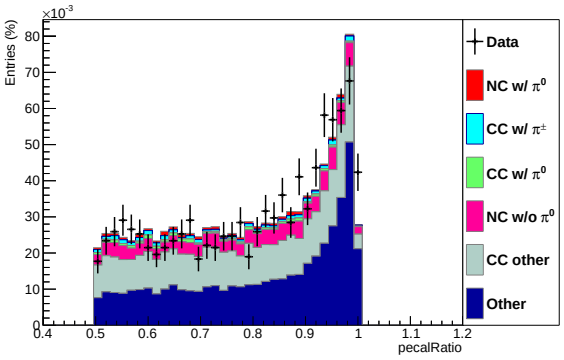
(c) Water



(d) Air



(e) Water



(f) Air

Figure 7.29: Data MC comparison of the BDTcat classification input variables for the PØD filled with (plots on the left) and without (plots on the right) water.

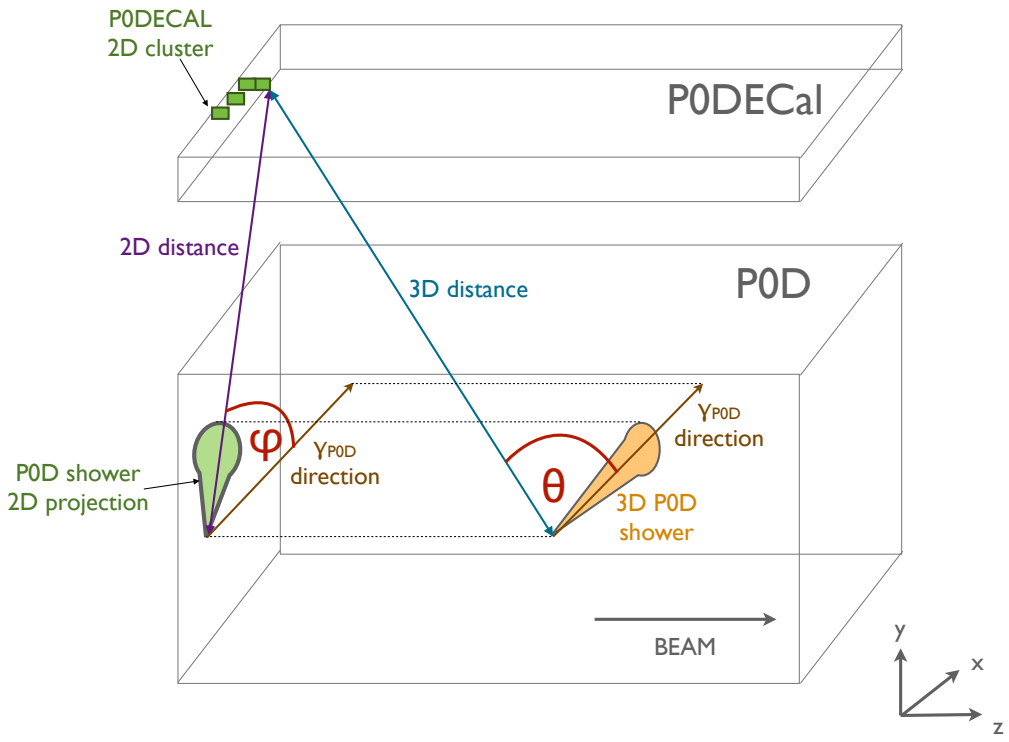


Figure 7.30: Geometry of the NC1π^0 event in the PØD and PØD-ECal sub-detectors. Distance between two reconstructed clusters can be determined in 3D (blue) and 2D (purple). The latter is calculated between 2D PØD-ECal object and 2D xy projection of the PØD object. Highlighted are two angles, θ and φ, between 3D and 2D distances and PØD reconstructed shower direction.

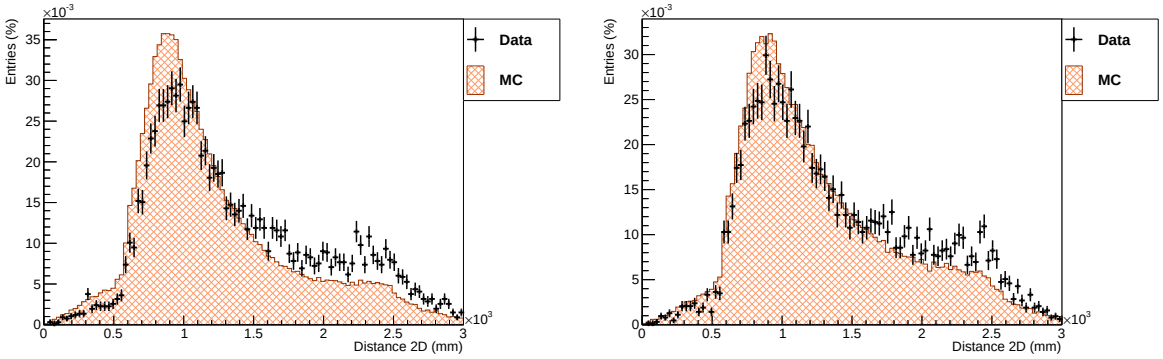
The variable GammasDir_angle is defined as

$$\text{GammasDir\_angle} = \frac{\Delta t}{|\Delta t|} \cos\theta, \tag{7.6}$$

where Δt is the time difference between the PØD and PØD-ECal clusters and θ is the angle between the PØD and PØD-ECal cluster distance direction and the PØD shower reconstructed direction. The second variable, GammasDir_dist, is the angle, φ, between 2D distance of the PØD and PØD-ECal clusters in the xy plane[1] and the 3D reconstructed PØD shower direction. Spatial components of these variables are shown in Figures, 7.31 and 7.32, which are two dimensional distance in the xy plane, and a three dimensional distance between the PØD and

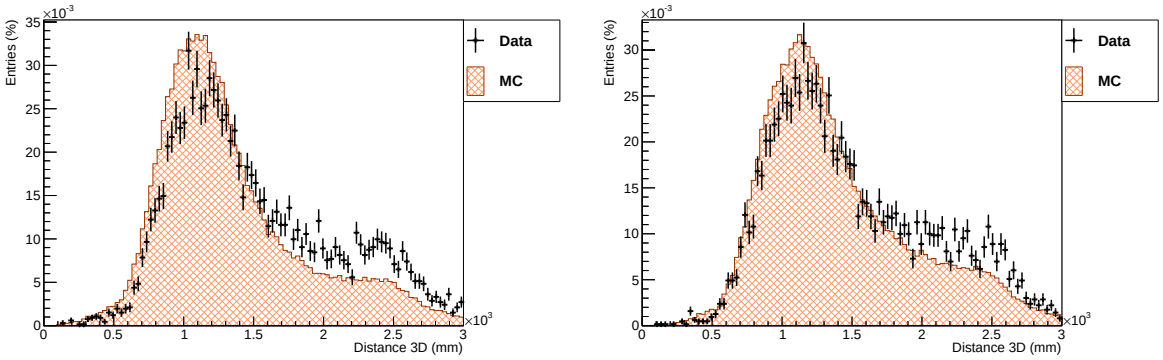
[1]Projection of the PØD 3D cluster on the 2D xy plane, where the 2D PØD-ECal reconstructed objects reside.

PØD-ECal clusters, respectively. It can be seen that the data MC disagreement



(a) Water (b) Air

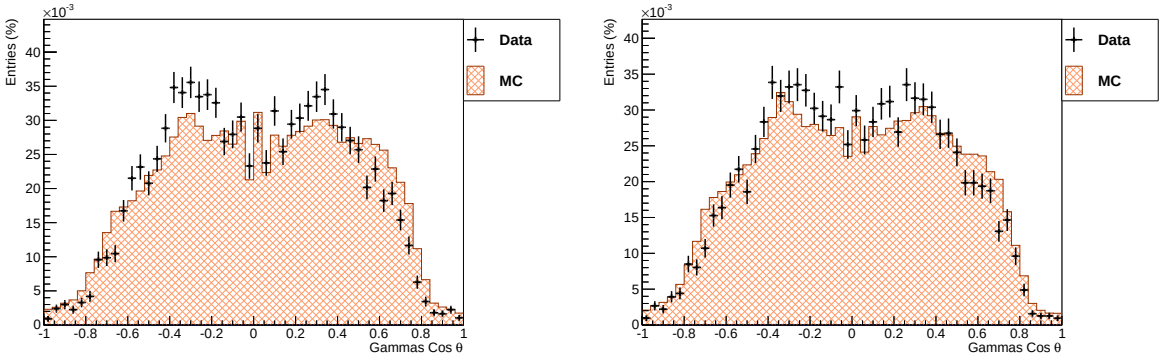
Figure 7.31: Data and MC comparison of the 2D distance in the xy plane between the PØD and PØD-ECal reconstructed clusters, area normalised to the PØD filled with (a) and without (b) water.



(a) Water (b) Air

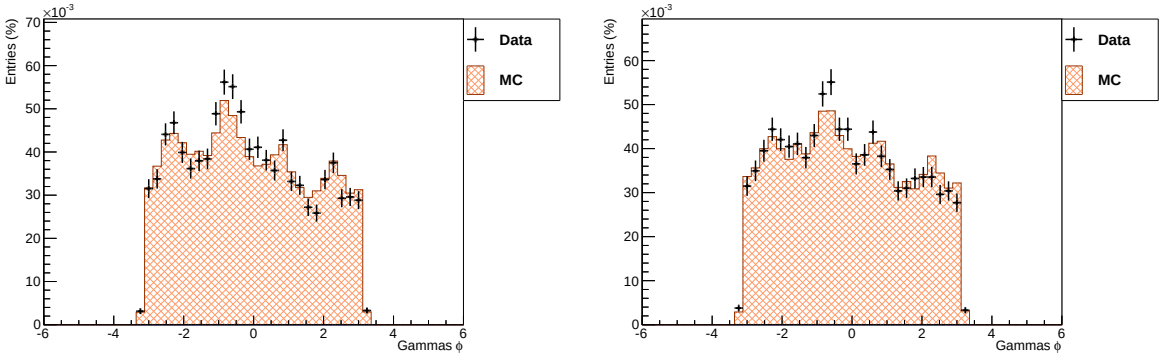
Figure 7.32: Data and MC comparison of the 3D distance between the PØD and PØD-ECal reconstructed clusters, area normalised for the PØD filled with (a) and without (b) water.

is significantly higher for the distance above 2 m. Repeatedly, data and MC show higher discrepancy for the PØD filled with water events.The next two variables concern the direction between the two candidate clusters. Such direction is calculated from the PØD cluster towards the PØD-ECal cluster and is shown in the spherical coordinates in Figure 7.33 and 7.34. The first figure shows a data-MC disagreement in the angle distribution, θ, which indicates that the angular propagation is increasing with the respect to the positive z axis. This discrepancy occurs for values of the angle between $-\pi/4$ and $\pi/4$. The second distribution



(a) Water (b) Air

Figure 7.33: Data and MC of the calculated direction between the PØD and PØD-ECal clusters in the spherical variable θ (the direction in which the angle is increasing from the positive z axis), area normalised for the PØD filled with (a) and without (b) water.



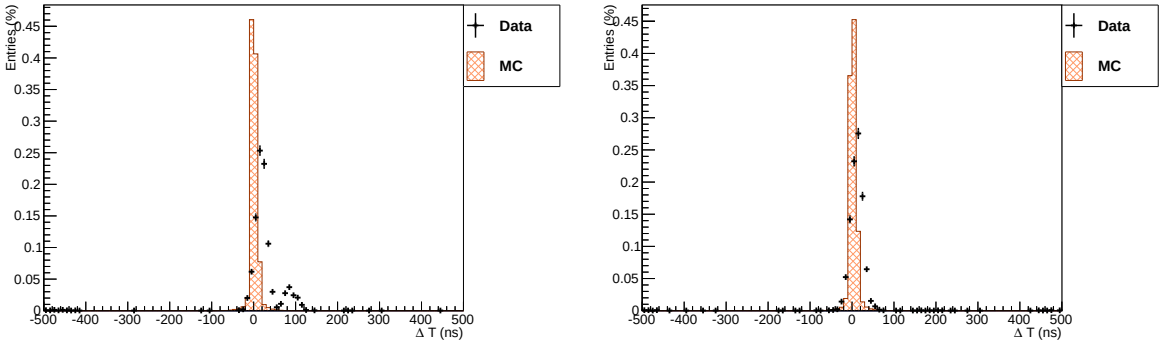
(a) Water (b) Air

Figure 7.34: Data and MC comparison of the calculated direction between the PØD and PØD-ECal clusters in the spherical variable ϕ (the direction in which the angle is increasing in the xy plane counterclockwise from the positive x axis), area normalised for the PØD filled with (a) and without (b) water.

represents the data-MC comparison of the direction of angular distribution, ϕ, where the angle is increasing in the xy plane counterclockwise from the positive x axis. The ϕ distribution plot shows a good agreement between data and MC.

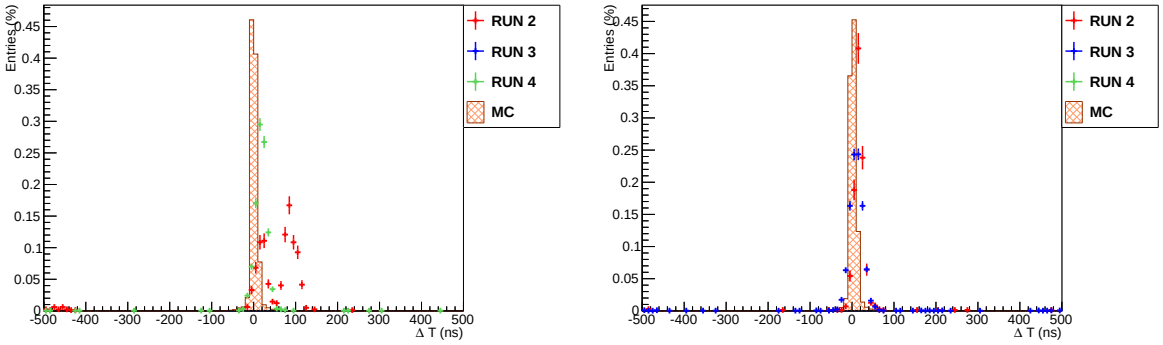
The next important component that contributes to the data-MC discrepancy is time difference between two reconstructed gamma candidate clusters. The time difference between the PØD and PØD-ECal candidate clusters is shown in Figure 7.35. The large disagreement between data and MC can be seen for the PØD filled with water, where the additional small crest has appeared next

to the main peak. This is caused by the Run 2 data, as shown in Figure 7.36, and explained in the data quality report [9]. During the first part of Run 2,



(a) Water (b) Air

Figure 7.35: Data and MC of the time difference between the PØD and PØD-ECal reconstructed clusters, area normalised for the PØD filled with (a) and without (b) water.

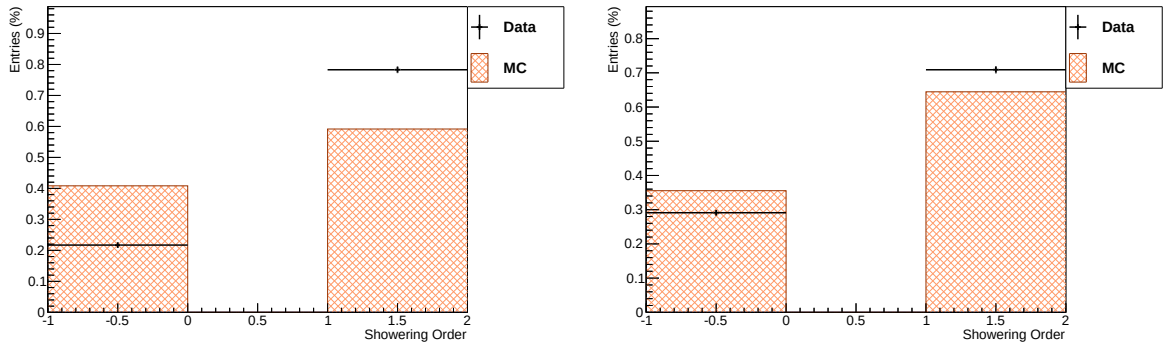


(a) Water (b) Air

Figure 7.36: Data for specified Runs and MC of the time difference between the PØD and PØD-ECal reconstructed clusters, area normalised for the PØD filled with (a) and without (b) water. The data points shift for the water-in configuration is caused by the Run 2 as reported in in Ref. [9].

between November and December 2010, the offsets for the ECal modules were not properly optimized and as a consequence the timing distributions were peaking near the beginning of the integration window. In January 2011 the offsets were adjusted. Unfortunately, this issue was not calibrated out. Additionally, the beam group reported various fluctuations for the Run 2 and the observed jitter for the PØD and PØD-ECal modules was up to 40 ns.

Next, the showering order of the two photon candidates, calculated as $\Delta T/|\Delta T|$, is shown Figure 7.37. It can be seen that the photon converting in the PØD is favoured to shower before the PØD-ECal photon for data events, which in a given notation convention has value +1, and PØD-ECal photon to shower as a first for MC events, with value −1. This causes data-MC asymmetry that influences GammasDir_angle variable where $\Delta T/|\Delta T|$ is a component of its definition in Equation 7.6. Therefore all data events are shifted towards positive values of the distribution and MC events are moved in the opposite direction. As a result, the data and MC shapes in Figures 7.29 (c) and (d) are shifted relative to each other.



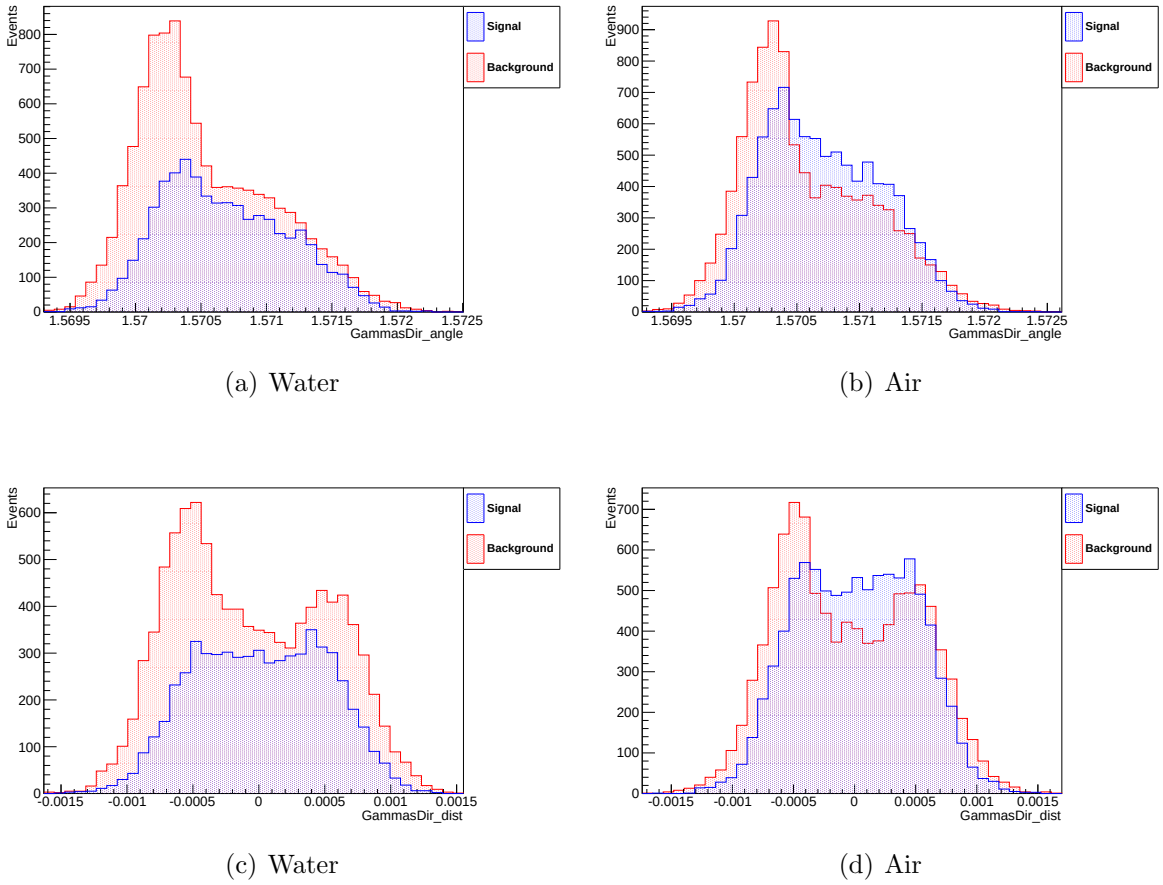
(a) Water (b) Air

Figure 7.37: Data and MC comparison for the π^0 decay photon candidates showering order ($\Delta t/|\Delta t|$), where the value of 1 is assigned when the PØD photon was converted first, and −1 when it showered second. Plots are area normalised for the PØD filled with (a) and without (b) water.

Despite the fact that GammasDir_angle and GammasDir_dist variables are not so well described by MC they have the strongest separation power to single out NC1π^0 signal from all types of background events, see Figure 7.38. Therefore, it is important to use both these variables as an input for the BDTcat method. All the signal and background distributions which are input variables used during the BDTcat training can be found in appendix A.

BDTcat Performance

The receiver operating characteristic curve (ROC) is shown in Figure 7.39 and the BDTcat output is shown in Figure 7.40. In the latter figure, each plot displays



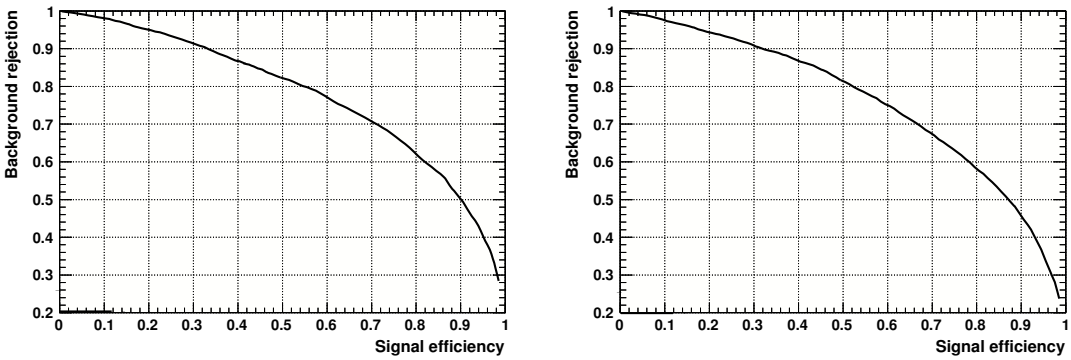
(a) Water (b) Air

(c) Water (d) Air

Figure 7.38: Two the most powerful BDTcat input discrimination variables, GammasDir_angle (top) and GammasDir_dist (bottom), plotted as the signal and background distributions by the TMVA package tools. Both variables are shown for the PØD filled with (left) and without (right) water.

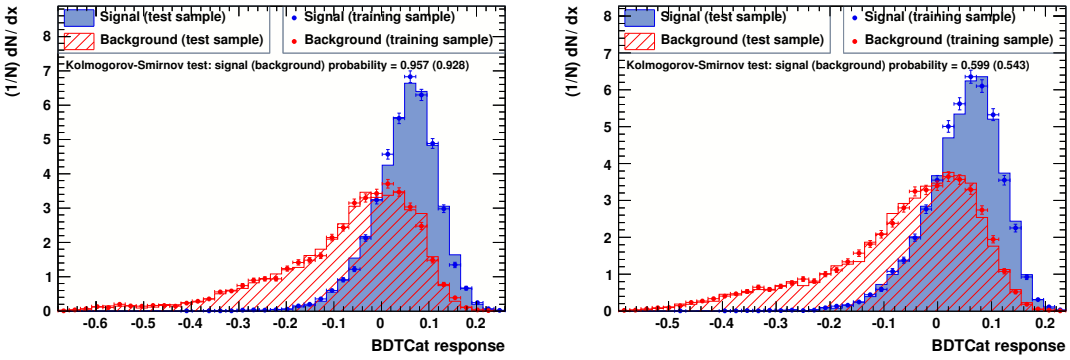
both, the signal (blue) and background (red) distributions, where the points represent a training sample response and the histogram represents a MC testing sample response. It can be seen, that the testing sample distribution shows a good agreement with the training sample, which indicates that the sample overtraining was minimised. As mentioned earlier in this chapter, the final result of the separation was strongly dependent on the fiducial volume cut value. The separation distributions for the different values of FV, but the same configuration of training input variables, are shown in appendix B. The final FV cutting value was chosen from the distribution of the maximum $S/\sqrt{S+B}$ as well as agreement training between the training sample and testing sample. The signal significance dependence on the fiducial volume value is shown in Figure 7.41. Considering the previously discussed PØD-ECal related disadvantages, i.e. the geometry related

limitations of detector ability and particles reconstruction status, the achieved separation of the signal and background distributions is currently the best possible and relatively satisfying.



(a) Water (b) Air

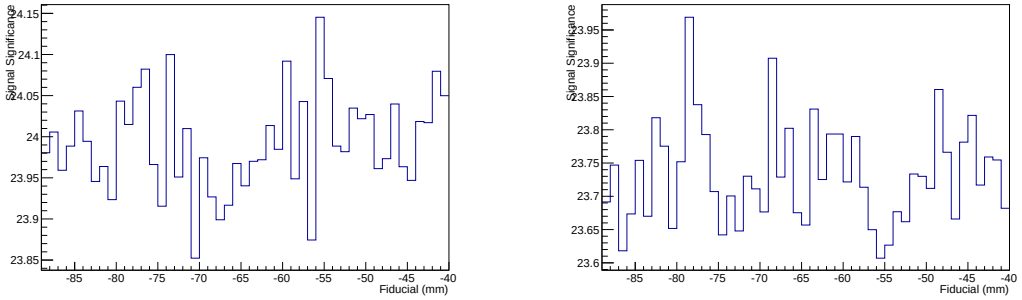
Figure 7.39: The ROC curve illustrating the performance of the BDTcat classifier for the PØD filled with (a) and without (b) water.



(a) Water (b) Air

Figure 7.40: MVA response for signal and background separation using the categorised BDT method, BDTcat, for PØD filled with (a) and without (b) water. Note the good agreement between training and test sample.

Next, the cut on the BDTcat response distribution has to be optimised and the most common choice to optimise the cutting figure is to calculate its value according to the expression



(a) Water (b) Air

Figure 7.41: Signal significance for BDTcat response trained with different values of fiducial volume for the PØD filled with (a) and without (b) water.

$$\frac{\text{Number of signal events after cuts}}{\sqrt{(\text{Total number of SIGNAL events}) + (\text{Total number of BACKGROUND events})}}. \tag{7.7}$$

This method can be determined using tools in the TMVA package, which give the best possible ratio for the efficiency and purity as shown in Figure 7.42 and explained in appendix B.

However, in the case of rare events such as the NC1π^0, it is common to choose alternative method in order to maximise the purity of the sample. Therefore, to determine the BDTcat cutting value, a different approach has been applied in this analysis. The discriminator value is calculated from the multiplication of efficiency and purity

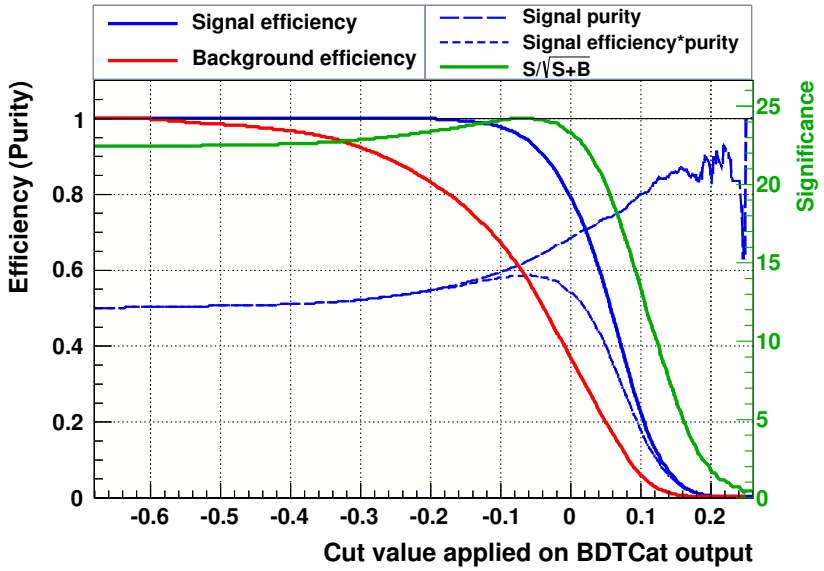
$$M_{EP} = \text{efficiency} \times \text{purity} \qquad \text{and} \tag{7.8a}$$

$$M_{EPP} = \text{efficiency} \times \text{purity}^{purity}, \tag{7.8b}$$

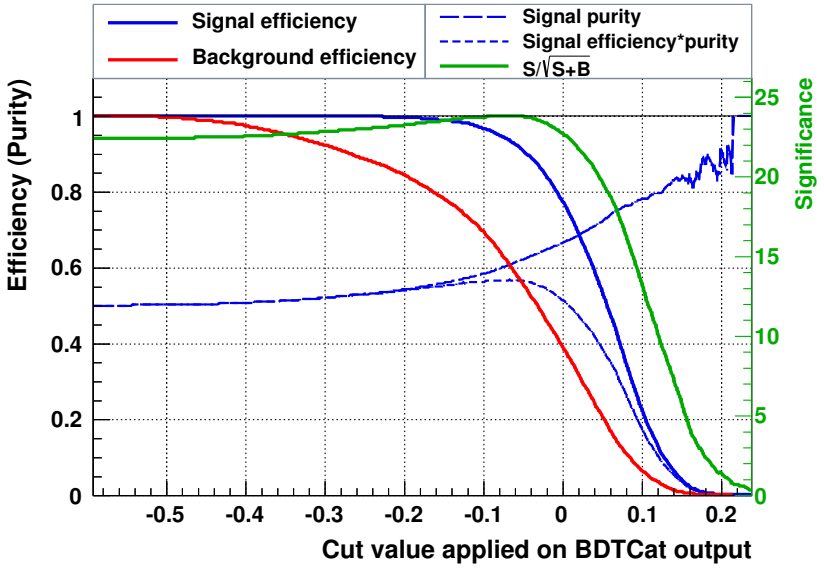
where the efficiency is defined as

$$\text{EFFICIENCY} = \frac{\text{Number of reconstructed events from true signal events sample}}{\text{Number of true signal events}} \tag{7.9}$$

and the purity is given by



(a) Water



(b) Air

Figure 7.42: TMVA NC1π^0 selection cut classifier (BDTcat) tuning of the signal and background ratio for the PØD filled with (a) and without (b) water.

$$\text{PURITY} = \frac{\text{Number of reconstructed events from true signal events sample}}{\text{Number of reconstructed events}}. \tag{7.10}$$

Figure 7.43 shows the distributions of the quantities from Equations 7.8, 7.9 and 7.10, where both discriminators, M_{EP} and M_{EPP}, indicate the same value of a cutting BDTcat point. Table 7.9 summarises the BDTcat cut values chosen by the different optimisation methods.

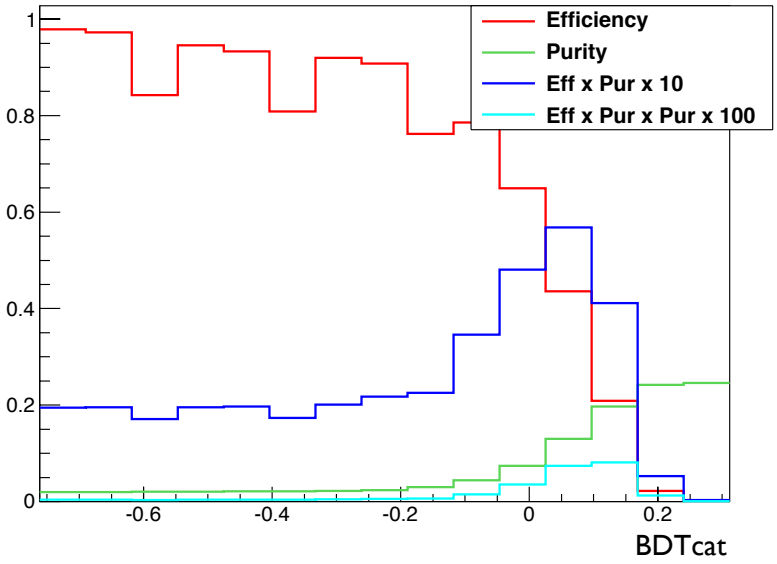
Configuration	$S/\sqrt{S+B}$		$M_{EP} \times 10$		$M_{EPP} \times 100$	
	Value	BDTcat Cut	Value	BDTcat Cut	Value	BDTcat Cut
WATER	24.1	-0.06	0.57	0.07	0.08	0.07
AIR	23.9	-0.07	0.65	0.08	0.11	0.08

Table 7.9: The cutting value of the BDTcat discriminator for the three different optimisation methods.

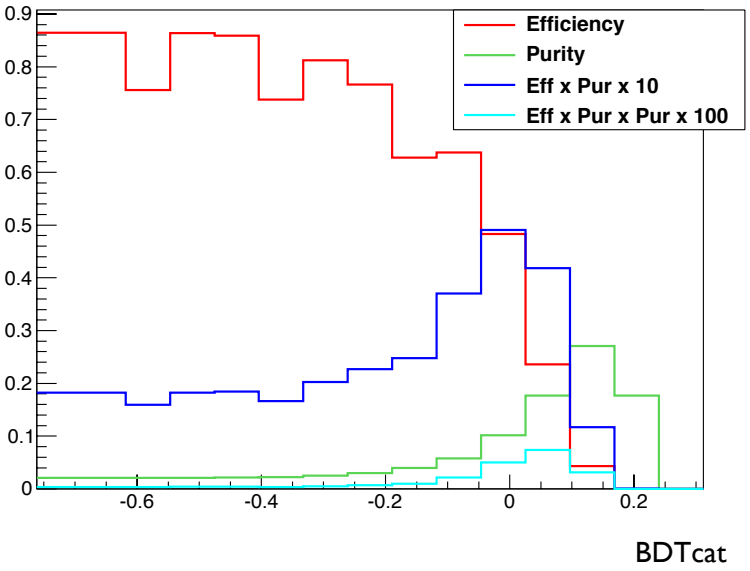
Based on the above, the BDTcat cutting value was chosen to be 0.07 for the water-in PØD configuration and 0.08 for the water-out configuration. Figure 7.44 shows the distribution of the BDTcut variable for the analysed MC and data events for both PØD water configurations. It can be seen, that distributions of BDTcat variable for data and MC differ with the significant excess of the MC events. This disagreement varies for both PØD water configurations. The water-in data points are "twisted" with respect to the MC events in such way that events selected after the BDTcat cut application show an excess of the MC. The rejected events show a deficiency of the MC events. This is related to the Run 4 data calibration status, the first pass production, which means that not a full calibration was applied as of this thesis writing. The water-out configuration data and MC events have a consistent behaviour and the MC events excess can be seen throughout all the BDTcat value events.

7.2.5 Efficiency and Purity of NC1π^0 Event Selection

After the selection, an efficiency of 3.0% and purity of 17.7% is obtained for the PØD filled with water, and an efficiency of 4.0% and purity of 16.4% is obtained

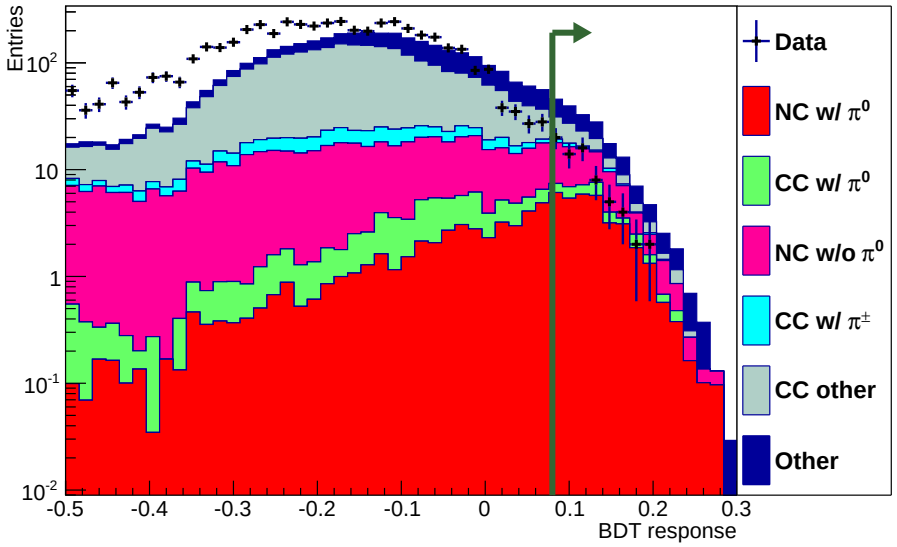


(a) Water

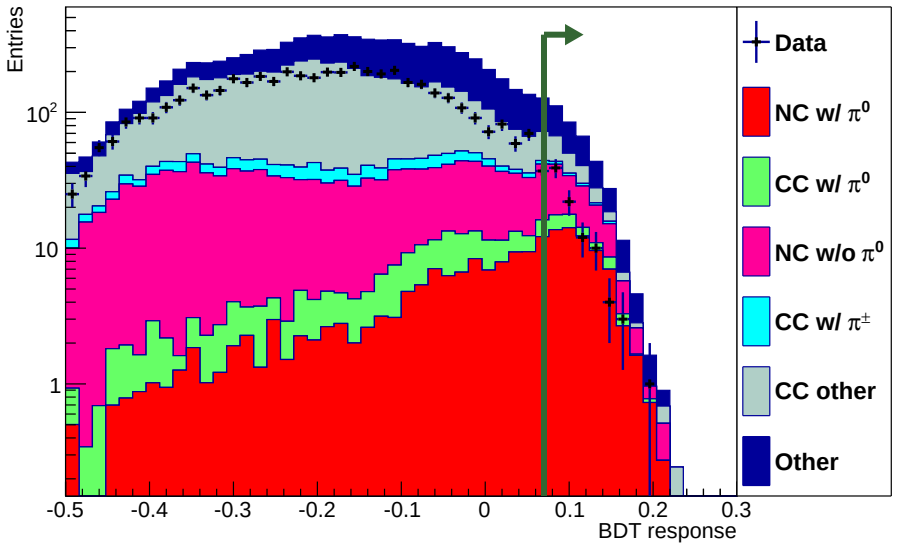


(b) Air

Figure 7.43: NC1π^0 efficiency and purity for different BDTcat response values for the PØD filled with (a) and without (b) water.



(a) Water



(b) Air

Figure 7.44: The BDTcat discriminator distributions for the PØD filled with (a) and without (b) water. The arrow indicates the chosen cut value in the NC1π^0 evetns selection.

for the PØD without water. The summary of the signal and background events with the neutral current π^0 candidates passing each selection cut in shown in Table 7.10 for the water-in and in Table 7.11 for the water-out PØD configuration.

The MC for the combined Run 2 and Run 4 (Run 2 and Run 3) analysis for the PØD filled with water (without water), predicts a total 181 ± 13.4 (stat) (405 ± 20.1 (stat)) NC1π^0 candidate events passing all the cuts for the PØD filled with (without) water. The selected events include 32 ± 5.7 (stat) (66 ± 8.1 (stat)) signal events and 149 ± 12.2 (stat) (339 ± 18.4 (stat)) background events.

The vertex distribution of the selected events for data and MC is shown in Figure 7.45. As expected, both data and MC events have vertices located towards the edge of the PØD FV for x and y direction, and leaning upstream beam direction for the z coordinate for water-in PØD configuration. Events for water-out configuration have vertices scattered more evenly.

The energy of the parent neutrino for the MC events which passed all the selection cuts are showed in Figure 7.46. As expected, the main peak is slightly above 0.6 GeV for both PØD water configurations. It also shows the distribution of different types of background which is discussed in more detail in the next section.

Background

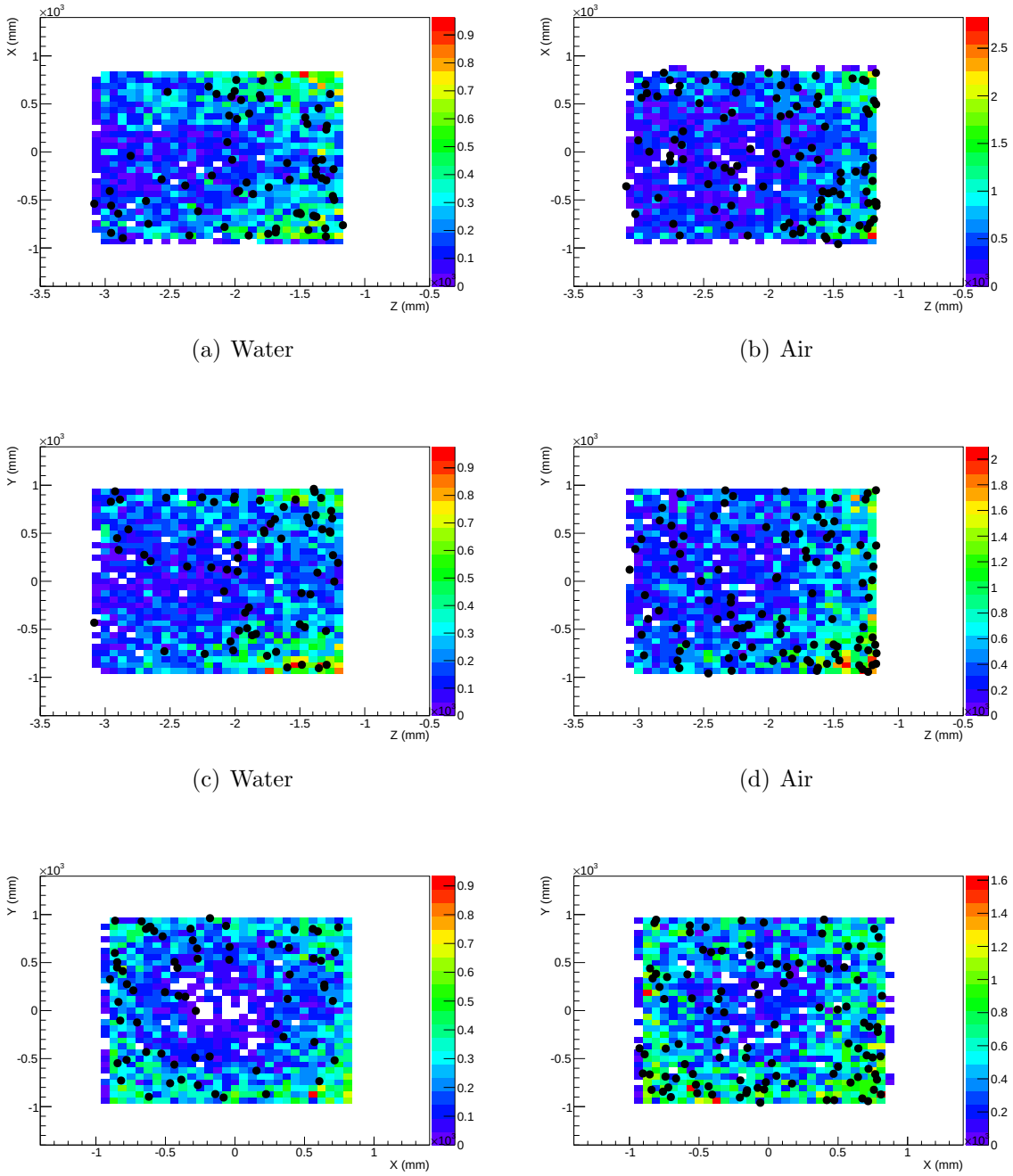
Tables 7.12 and 7.13 show the breakdown of the background events which passed all the selection cuts for the PØD filled with and without water, accordingly. The largest background is constituted by the *other* type of events and it accounts for 30% (39%) of all selected events for water-in (water-out) PØD configuration ("other" type background events are not included in Figure 7.46 as there is no ν energy information in MC). The second largest background originates in other charged current events, which initially after the pre-selection stage, established the highest NC1π^0 background addition. This indicates that the good understanding of external and multiple neutrino interactions is essential for the next stage of analysis and therefore the *sand* MC files integration into the NC1π^0 studies with the PØD and PØD-ECal sub-detectors is a crucial requirement. The CCπ^0 events account for 16% of all selected CC interactions, for both water configurations. The neutral current interactions, other than PØD-PØD-ECal NC1π^0, constitute 25% (22%) of the total background for the PØD filled with (without) water.

No_{cut}	CUT	Total	Signal	Efficiency (%)	Purity (%)
1	Preselection	197 264	1080	100	0.5
2	Tracker Veto	131 468	910	84.2	0.7
3	One 3D EM PØD Shower	86 093	750	69.3	0.9
4	Fiducial_{water}	27 133	303	28.0	1.1
5	Muon Cut	19 786	277	25.6	1.4
6	PØD Track	14 467	255	23.6	1.8
7	$\text{PØD Shower PID weight}_{water}$	5 147	99	9.2	1.9
8	BDT_{water}	181	32	3.0	17.7

Table 7.10: Efficiency and purity for each NC1π^0 selection cut for the PØD filled with water.

No_{cut}	CUT	Total	Signal	Efficiency (%)	Purity (%)
1	Preselection	273 978	1623	100	0.5
2	Tracker Veto	184 228	1383	85.2	0.8
3	One 3D EM PØD Shower	118 673	1101	67.8	0.9
4	Fiducial_{air}	33 940	455	28.0	1.3
5	Muon Cut	26 155	455	25.6	1.6
6	PØD Track	18 528	367	22.6	2.0
7	$\text{PØD Shower PID weight}_{air}$	11 803	229	14.1	1.9
8	BDT_{air}	405	66	4.0	16.4

Table 7.11: Efficiency and purity for each NC1π^0 selection cut for the PØD without water.

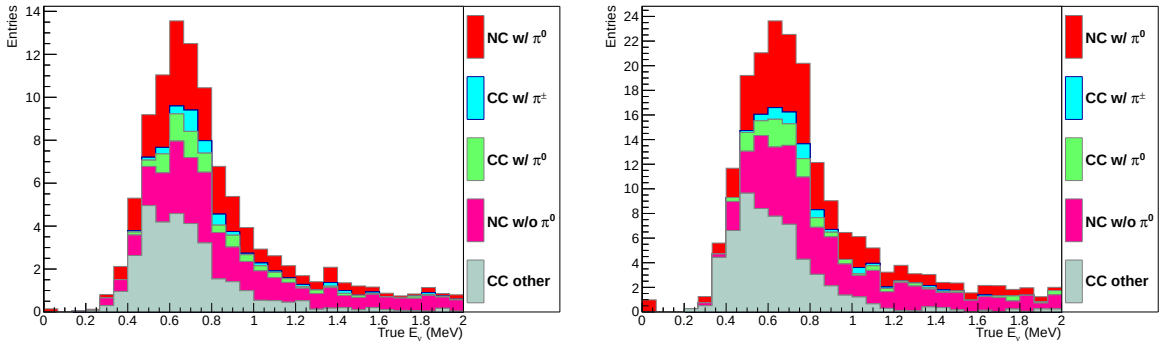


(a) Water (b) Air

(c) Water (d) Air

(e) Water (f) Air

Figure 7.45: The reconstructed vertex position of the selected events for data and Monte Carlo comparison for the PØD filled with (a) and without (b) water.



(a) Water (b) Air

Figure 7.46: The energy of the parent neutrino for the selected MC events for the PØD filled with (a) and without (b) water.

Background	Total		
Other Neutral Current	46		
Charge Current	48	CC with π^0	8
		CC with $\pi^\pm$	5
		CC other	35
Other	55		

Table 7.12: A summary of the selected background events. The number of expected background events is broken down into the contributions from the different event categories for the PØD filled with water.

Background	Total		
Other Neutral Current	87		
Charge Current	93	CC with π^0	16
		CC with $\pi^\pm$	8
		CC other	69
Other	159		

Table 7.13: A summary of the selected background events. The number of expected background events is broken down into the contributions from the different event categories for the PØD without water.

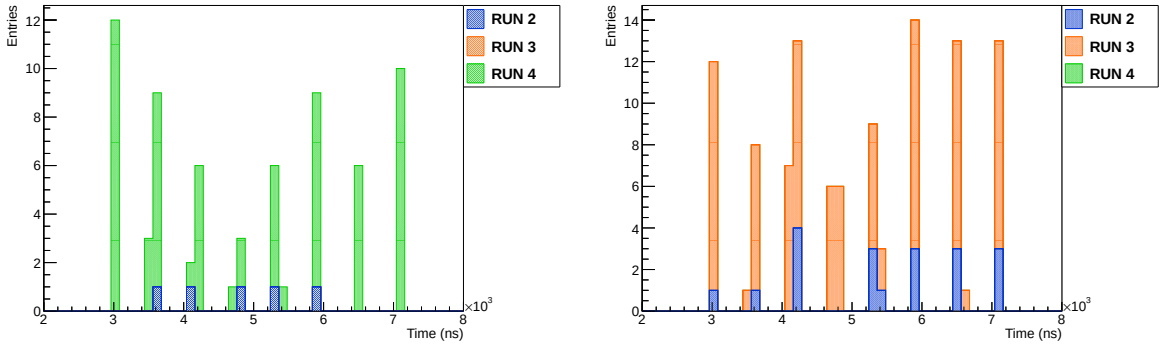
7.2.6 Data Outcome

The observed number of data events for data combined Run 2 and Run 4 (Run 2 and Run 3) for the PØD filled with water (without water) is 68 (107). The selected events for each run number are presented in Table 7.14. The timing dis-

Configuration	Total	Run 2	Run 3	Run 4
WATER	68	5	NA	63
AIR	107	19	88	NA

Table 7.14: A summary of the observed events that passed all the selection cuts and split according to the individual run number for the both PØD configurations, water-in and water-out.

tribution of the observed events in the beam is shown in Figure 7.47 which details the observed events form NC1π^0 interactions within the beam bunch structure.



(a) MC (b) Data

Figure 7.47: The timing distribution of the observed events for each run number for the PØD filled with (a) and without (b) water.

In order to extract the total number of signal events from the observed events, commonly the fit to the variable with the distinctive signal and background distributions is used. In the PØD standalone NC1π^0 analysis such variable is chosen to be the reconstructed mass , $M_{\gamma\gamma}$, of the candidate π^0, which is determined according to the expression

$$M_{\gamma\gamma} = \sqrt{2E_{\gamma 1}E_{\gamma 2}\,(1 - cos\theta_{\gamma\gamma})}, \tag{7.11}$$

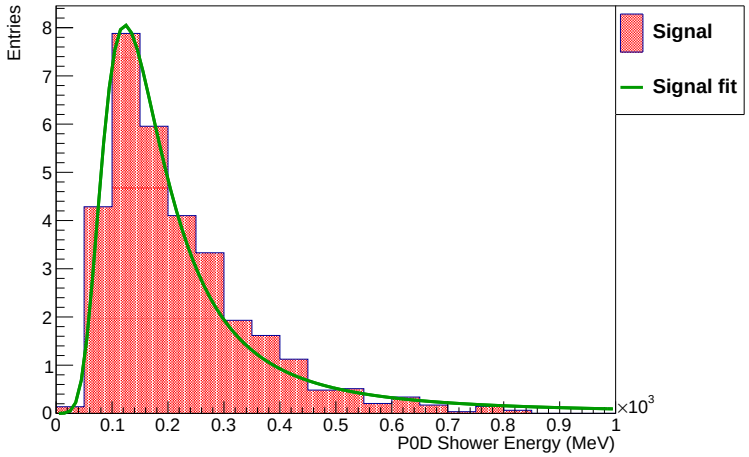
where $E_{\gamma 1}$ and $E_{\gamma 2}$ are the reconstructed energies of the candidate showers and θ is the angle between the two γ reconstructed clusters. However, in this thesis analysis the lack of the energy reconstruction of the PØD-ECal cluster and "fractional" information about the reconstructed angle between the decay photons (no z spatial interaction information in the PØD-ECal) does not allow the invariant mass of the candidate π^0 to be reconstructed. Therefore, the other reconstructed variables have been considered and analysed, mainly the BDTcat input variables, eg. the PØD shower energy variable as shown in Figure 7.48 (7.49) for the PØD filled with (without) water.

The Landau fit has been applied to each distribution: signal, background and data points for both PØD water configurations. Unfortunately, non of the examined variables has enough distinctive shape and significantly large statistics of the signal events to extract them from the whole observed events set.

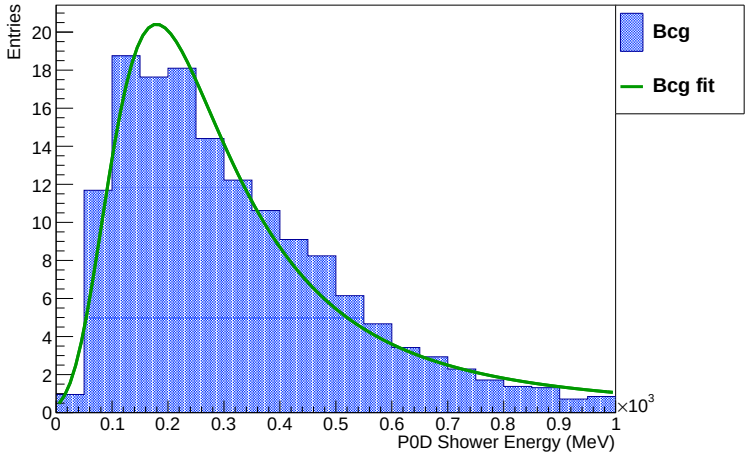
7.3 Discussion on the PØD-ECal Status in the NC1π^0 Analysis and Future Improvements

The first use of the PØD-ECal sub-detector in the physics analysis has been presented in this chapter. Based on the obtained results, it can be concluded that the readiness of PØD-ECal sub-detector, at the current stage of development, is constrained and requires more improvements and statistics to provide a contribution to the cross section physics studies. The initial *particle gun* MC examination have shown that the PØD-ECal can provide a good addition to the planned measurement. This was also confirmed on the low level data checks. However, using the full beam analysis with applied *magnet* MC files, the performance of the PØD-ECal has been verified and showed a need for further developments in the very aspects of the reconstruction and analysis.

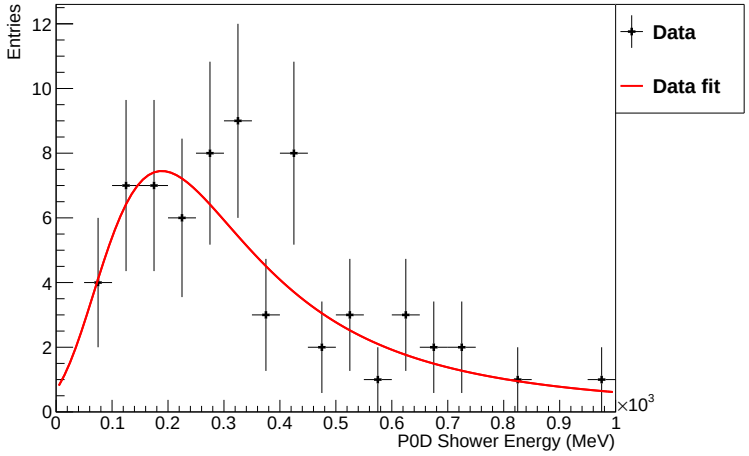
The first notable insufficiency concerns the reconstruction stage of particles passing through the PØD-ECal. It has been shown, that amongst the selected NC1π^0 events, the PØD-ECal cluster objects are fragmented which results in the high number of clusters with 2 or 3 hits. This makes the PID assignment very challenging. As a consequence, the PID efficiency was too low to be applied in the analysis which led to a decision of using cluster objects only. However, the particle identification was handled by the MVA technique and all the PID discrimination variables were utilised as an input during matching stage of reconstructed objects between the PØD and PØD-ECal sub-detectors. It is shown, that the MVA technique is a powerful tool to increase the efficiency and purity of NC1π^0 event



(a) Signal

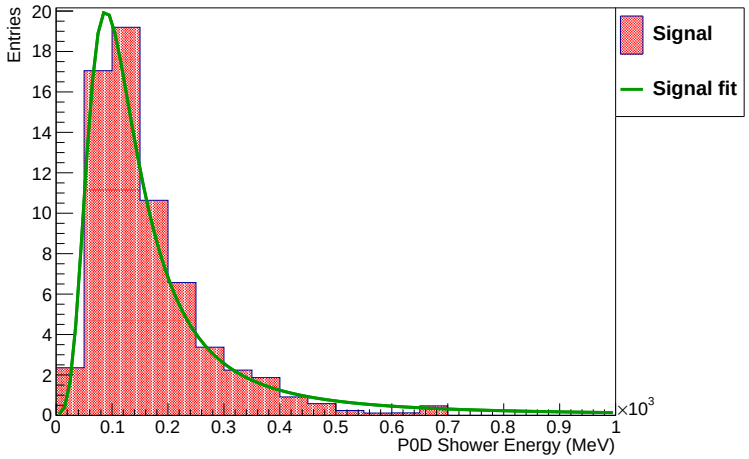


(b) Background

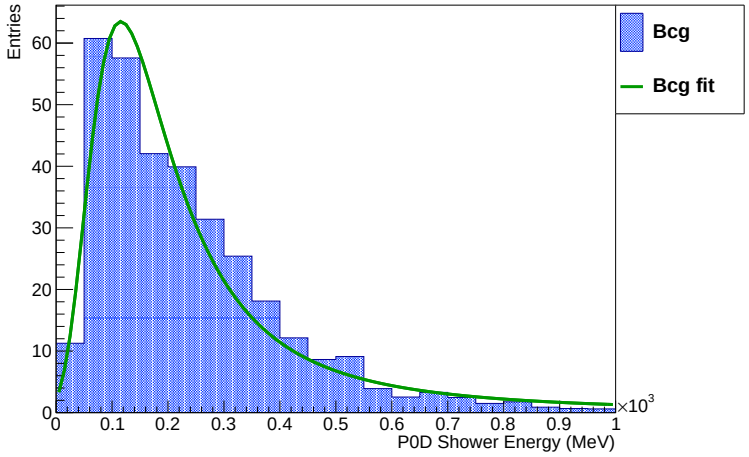


(c) Data

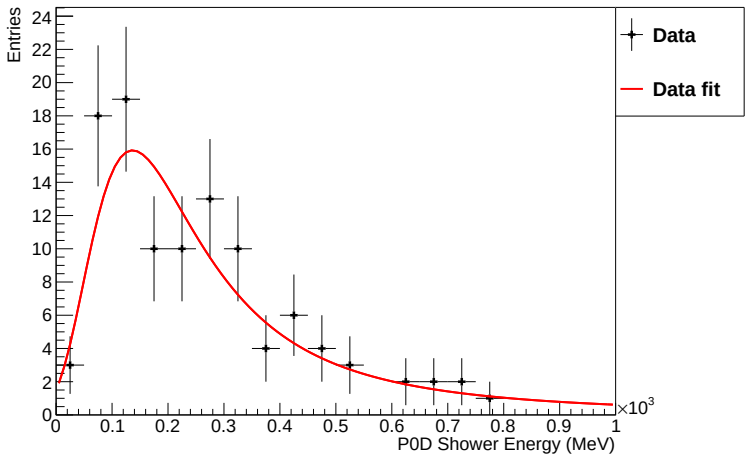
Figure 7.48: The landau fit to the signal (a), background (b) and data (c) events that passed all the cuts for the PØD filled with water.



(a) Signal



(b) Background



(c) Data

Figure 7.49: The Landau fit to the signal (a), background (b) and data (c) events that passed all the cuts for the PØD without water.

selection.

Since the PØD-ECal has no ability to reconstruct object directions, it does not distinguish between events where a particle comes from the interactions appearing inside ND280 detector, or from the surrounding ND280 detector pit. Therefore, the lack of sand files in the simulation studies had a significant impact on the background estimation. The sand files are tuned to match the collected data and it should be closely examined in the future analysis.

The other factor that was an obstacle in the background and uncertainties estimation, is low statistics. The accumulated POT collected by the T2K near detector ND280 in the Run 1,2,3 and 4 represents less than 10% that expected from the full designed nominal collection. The increased statistics would widen the exploration areas for the side band studies and signal extraction. There is an indication that with higher statistics the selection can be successfully applied to the data sample, and when interactions in the pit and the relationship between the PØD and PØD-ECal is better understood, the PØD-ECal can contribute to the cross section measurement.

The PØD-ECal based NC1π^0 event selection developed in this thesis demonstrated the use of PØD-ECal sub-detector to catch high angle escaping π^0 decay photons from neutrino induced interactions in the PØD sub-detector. In principle, the PØD-ECal can contribute also to the CC analysis, as high angle tracks could be selected in a similar way. Moreover, the track matching might be more successful as the PØD reconstructed track can be extrapolated to the PØD-ECal volume and as a result the z position (beam direction) can be estimated. This also can contribute to the reduction of CC background in the NC1π^0 studies. The next option would be to examine the partially escaping gamma showers from the PØD, which can be matched with the PØD-ECal objects. Obviously and unfortunately, the amount of escaped energy cannot be calculated due to the coarse nature of the PØD-ECal.

In general, the PØD-ECal by design plays a supportive role for the PØD sub-detector analysis. However, there is a limitation to the PØD-PØD-ECal event selection related to the PØD-ECal basic construction and geometry, dictated by the limited budget. The simplicity of the crude nature of the PØD-ECal, described in sections 4.3.6 and 6.3.2, determines current PØD-ECal performance and imposes limits on the possible improvements. The budget PØD-ECal construction was a trade off for the physics potential that can be provided by the PØD-ECal. The ideal electromagnetic calorimeter should hermetically surround the tracking detector with the target interaction area and have the ability to stan-

dalone and simultaneously reconstruct particles passing from the active region. It should distinguish the passing charged particles sign in presence of surrounding magnetic field. The energy and direction reconstruction are crucial in order to have kinematic information. Moreover, the particle identification that can distinguish between photons, muons and protons is essential to search for the events with neutral particle in the neutral current. In particular the structure should allow to get a reconstructed object position along the sub-detector beam direction to be obtained in order to identify the interaction point. However, the two dimensional layout of the PØD-ECal bars and the limited number of the layers, do not allow to the above requirements to be satisfied. Therefore, the critical evaluation of the PØD-ECal design, based on budget construction, concludes that it does not perform as an electromagnetic calorimeter but rather as a shower converting, tagging and tracking detector. This has influenced seriously the performance of the PØD-ECal sub-detector and the final analysis result performed in this chapter.

Based on the achieved results in which PØD-ECal was used for the first time, the potential PØD-ECal contribution looks promising and the aspects discussed above provide a number of indications for the future improvements that can enhance PØD-ECal performance in the analysis. The primary function of the PØD-ECal was shower tagging. The achievement of this goal requires further particle reconstruction algorithm improvements in order to have an important contribution to the NC1π^0 cross section measurement. This includes already ongoing work, which is planned to be implemented in data and MC production 6 and 7, as well as a number of suggestions for advancement work that would be beneficial for a future analysis.

Planned Improvements

The first important improvement is to increase the accuracy of time readout from the PØD-ECal which will benefit for the time relation between the PØD and PØD-ECal clusters. This includes the application of the time calibration which is essential for effective cluster matching between the two sub-detectors. The high accuracy of timing readout is crucial for analysers to obtain good results. The PØD-ECal time calibration algorithms are already developed, tested and ready to be applied in the next data and MC sub-production 6.

There is also an ongoing work on the integration of the PØD-ECal reconstructed clusters to the global reconstruction which performs matching between objects reconstructed in multiple sub-detectors. The global tracks algorithm is

currently working on the identification of tracks that pass through the PØD, PØD-ECal and SMRD sub-detectors, although it has to be re-optimised due to the recently diagnosed PØD-ECal PID problem. As for the shower-like reconstructed objects, the development of the global matching method is in progress. This can be particularly beneficial for the events where the showering photon is partially leaving the PØD sub-detector and continuous to shower in the PØD-ECal. The ongoing work on the integration of the PØD-ECal objects to the global reconstruction is of high importance to improve usiblility of the PØD-ECal in the physics analysis.

Suggested Improvements

The number of suggestions and thoughts arose during the analysis work followed by conclusions from the obtained results. The first recommended change concerns clustering algorithm and particle identification assignment. This requires stepping back to the development of the PØD-ECal reconstuction software. At the beginning, the clustering algorithm should be revised in order to reduce cluster fragmentation problem. The clustering criteria should be reoptimised and as a result it is expected to reduce the number of clusters that contain the lowest, 2 or 3, number of hits within the multicluster events. Therefore, the particular event type might contain the expected number of reconstructed clusters helping to achieve improvement in NC1π^0 selection efficiency. Moreover, the reduction of the low hits clusters has a direct impact on the PID algorithm performance as it is rather difficult to distinguish between the track-like and shower-like particles for the 2 or 3 hits clusters. As shown in section 7.1.3, the PØD-ECal PID needs to be changed. There is a number of options how the new PID assignment algorithm can be improved, e.g the track-shower discrimination could be only performed on the 4 or more clustered hits. The other possibility is to apply multivariate analysis to the PØD-ECal clusters on the analysis level, where the output PID will be assign to the particular type of event such as CCπ^0 or NCπ^0, as the MVA performs only signal-background discrimination. This way PID assignment could be trained on the more complex simulation type such as *magnet* MC. Finally, the evaluation of the PØD-ECal cluster identification can be implemented at the global reconstruction stage where the PID can be determined in coordination with objects from the other sub-detectors but the same event, ie. the PØD reconstructed object could be a seed to the global reconstruction event topology. Also this can be done with the usage of the SMRD sub-detector where the tracks going through the PØD, PØD-ECal and SMRD could be matched. This would

be highly beneficial in the CC event rejection.

The next possible improvement could be a development of the cross PØD and PØD-ECal sub-detector time calibration. This would enable both sub-detectors time readout to be correlated and enhance the PØD-PØD-ECal cluster time difference accuracy which is crucial for the NC1π^0 event selection.

The above changes hopefully will lead to an improved PØD-ECal performance and therefore at this point the NC1π^0 event selection should be revised. In particular, the PØD-PØD-ECal matching algorithm can be upgraded with the new input variables representing more precise spatial and time relation between the two sub-detectors. The next analysis development could use events division into smaller sub-samples with respect to the vertex interaction point along the beam direction to constrain NC1π^0 studies on the angular dependence of the interaction vertex and the PØD-ECal reconstructed cluster position.

The analysis and discussion made in this chapter provide a good insights into the most recent status of the PØD-ECal sub-detector. The NC1π^0 event selection results indicated the crucial changes that need to be done at the next stage of data analysis. The developed event selection presented in this chapter has not given high enough statistics to evaluate the systematics, which currently are expected to be large. Having applied the described improvements and larger data set, the new NC1π^0 analysis and systematic error evaluation can be performed. In this thesis, the systematic uncertainties are a subject to discussion in the next chapter.

Chapter 8

A Discussion on the Systematic Errors

In this chapter the systematic errors of the NC1π^0 event selection are discussed. The systematic uncertainties have a number of different sources and their good understanding is important for interpretation of the result. These errors originate from the cross section models defined by the parameters in the neutrino generator, the final state interactions, neutrino flux simulation, the boosted decision trees technique from the MVA method and the detector systematics. There is also an uncertainty associated with the PØD and PØD-ECal objects reconstruction. As of the time of writing this thesis the full set of uncertainties is not developed or possible to apply and therefore only a brief description of errors is described.

8.1 Flux Systematic Error

The good understanding of the T2K neutrino flux is crucial for the neutrino interactions and oscillations measurements. The flux was studied from a number of external and internal data, that enable uncertainties to be constrained. The first studied beam data are provided by the NA61 collaboration from the NA61/SHINE experiment for hadron production measurements. The next data sources are the proton beam monitors placed at the beam production terminal and INGRID detector. The used information concerns pion, kaon and secondary nucleon production, off-axis uncertainties, proton beam uncertainties, horn and target alignment, absolute horn current and the MC statistical uncertainty on the flux samples. All the information is then combined with uncertainties from differences in the production of secondary nucleons between the beam and MC. The beam working group evaluates all the considered elements and provides final

knowledge about the flux uncertainties and their errors for each of the three T2K detectors. Additionally, there is a constraint from the ND280 which is developed by the Beam And ND280 Flux extrapolation task Force (BANFF), called the BANFF fit [81]. The ND280 or SK other sample predictions are reweighted with the fitted parameters, and data versus MC prediction comparisons are made as a cross check. The outputs of the BANFF fit are a vector of the fitted flux and cross section parameters, and the corresponding covariance matrix. The neutrino flux parameters values prior to and after the BANFF fit are shown in Figure 8.1.

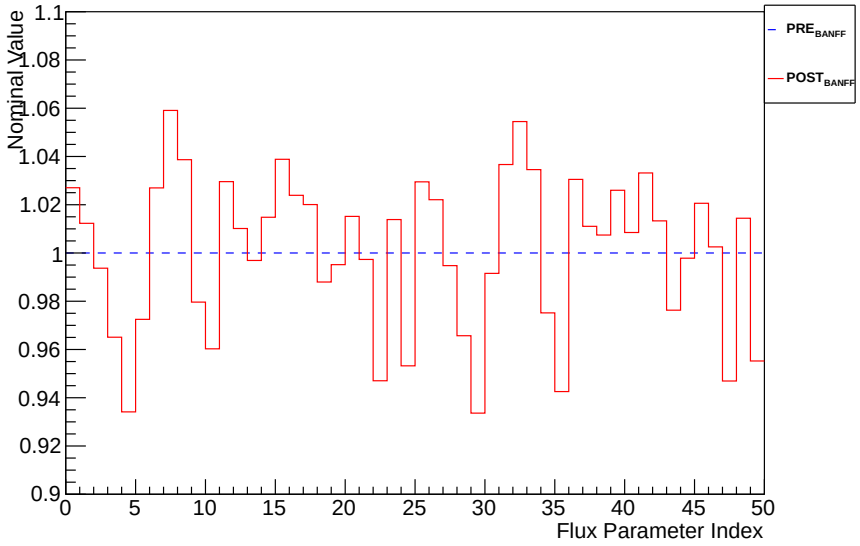
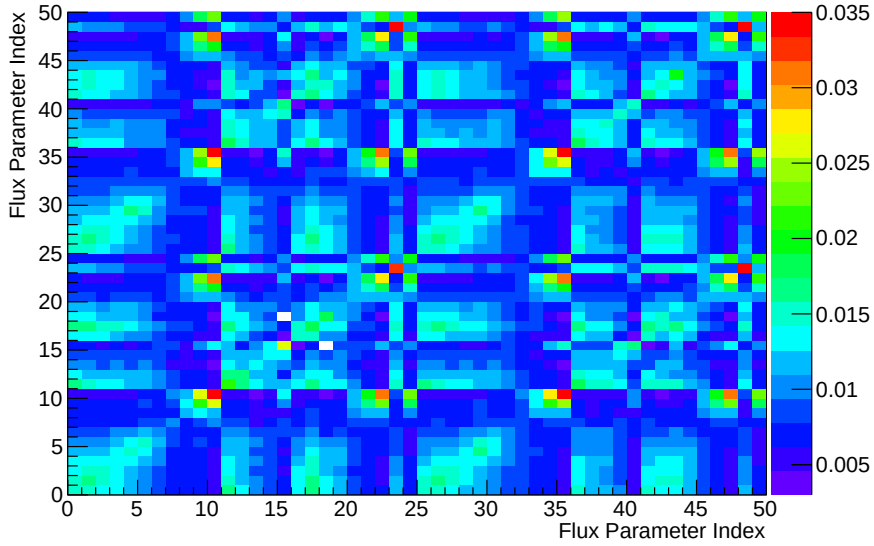


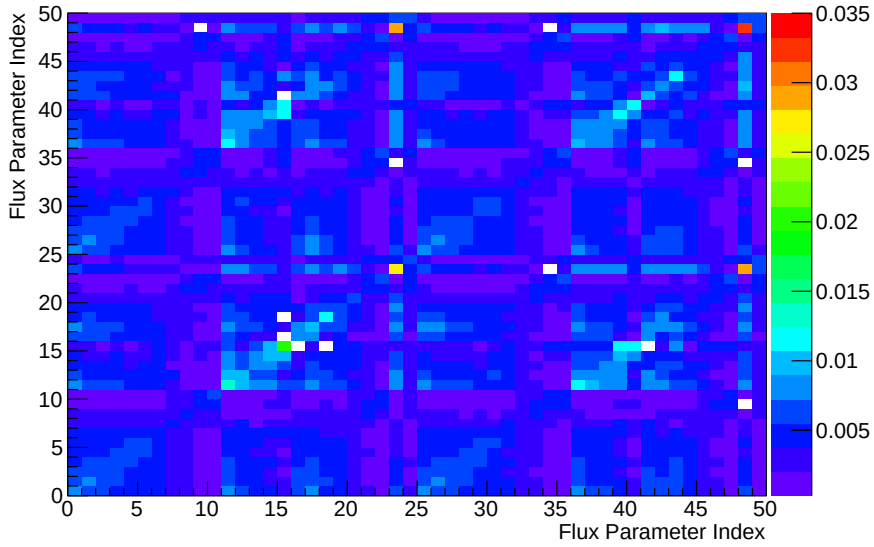
Figure 8.1: Nominal values of flux parameters prior to and after the BANFF fit. The first 25 parameters are related to the ND280 detector, and the remaining 25 parameters describe the Super-Kamiokande detector flux.

The T2K beam working group provides event weights related to the neutrino flux prediction. The flux inputs to the ND280 detector for Runs 1-4 are described in the technical note [12]. In total there are 25 flux parameters at ND280 which are summarised with their errors in Table 8.1. The weights have a form of a covariant matrix that is a function of neutrino flavours, true neutrino energy and type of detector: ND280 or Super-Kamiokande. Figure 8.2 shows the input flux parameters correlations prior to and after the BANFF fit, where the first 25 parameters are related to the ND280 detector, and the remaining 25 parameters

describe the Super-Kamiokande detector flux.



(a) pre-BANFF



(b) post-BANFF

Figure 8.2: Input covariance matrices for neutrino flux parameters at the ND280 detector prior to (a) and after (b) the BANFF fit.

There are two alternative methods for applying the flux systematic uncertainty on the analysis results: reweighting and the multiple flux simulations with a variation of the parameters. The first method is based on the random throws of

Parameter Index	Parameter Name	pre-BANFF fit fractional error	post-BANFF fit fractional error	Energy bin (GeV)
ND280 ν_μ Flux				
1	E1	0.1221	0.0829	0.0-0.4
2	E2	0.1284	0.0848	0.4-0.5
3	E3	0.1202	0.0798	0.5-0.6
4	E4	0.1180	0.0809	0.6-0.7
5	E5	0.1242	0.0870	0.7-1.0
6	E6	0.1209	0.0813	1.0-1.5
7	E7	0.1025	0.0670	1.5-2.5
8	E8	0.1002	0.0673	2.5-3.5
9	E9	0.1067	0.0653	3.5-5.0
10	E10	0.1473	0.0744	5.0-7.0
11	E11	0.1958	0.0788	7.0-30.0
ND280 $\overline{\nu}_\mu$ Flux				
12	E12	0.1446	0.1112	0.0-0.7
13	E13	0.1258	0.0968	0.7-1.0
14	E14	0.1146	0.0945	1.0-1.5
15	E15	0.1152	0.0949	1.5-2.5
16	E16	0.1609	0.1345	2.5-30.0
ND280 ν_e Flux				
17	E17	0.1242	0.0921	0.0-0.5
18	E18	0.1353	0.0941	0.5-0.7
19	E19	0.1377	0.1080	0.7-0.8
20	E20	0.1093	0.0783	0.8-1.5
21	E21	0.1086	0.0741	1.5-2.5
22	E22	0.1211	0.0660	2.5-4.0
23	E23	0.1675	0.0791	4.0-30.0
ND280 $\overline{\nu}_e$ Flux				
24	E24	0.1816	0.1645	0.0-2.5
25	E25	0.1393	0.0817	2.5-30.0

Table 8.1: Beam flux systematic fractional errors provided by the T2K beam group [12].

the flux nominal parameters and by constructing a covariance matrix, the flux prediction is reweighted. In the second method, each individual flux parameter is changed, which is followed by the flux re-simulation within one standard variation range of the flux parameters variation. As a result of any of these two methods, the final covariance matrix is formed which contains all combined uncertainties from different sources on the flux prediction.

8.2 Neutrino Generator Systematic Error

The neutrino MC generator, NEUT, systematic uncertainties originate in the modelling of the neutrino-nucleus interactions. Due to the limited knowledge about these interactions from measured data, the associated error is expected to be relatively large. The systematic errors related to the generator are associated with the neutrino cross section model parameters and final state interactions, which concerns nuclear matter. The generator parameters are constrained by the external up to date published data from other experiments. The cross section results used to constrain inputs at the T2K NEUT come from MiniBooNE, SciBooNE and NOMAD [82]. The list of parameters considered in generator simulation of theoretical models and their calculated fractional errors are presented in Table 8.2. The first six parameters from the table are describing final state interactions parameters that characterise used nuclear model:

- FSI with the low energy QE scattering parameter and with a single charge exchange
- FSI with the high energy QE scattering parameter
- FSI with the pion production parameter
- FSI with the pion absorption parameter
- FSI with the low energy single charge exchange branching fraction parameter
- FSI with the high energy single charge exchange branching fraction parameter

The next 15 parameters concern cross section modeling in the NEUT generator and their brief explanation is as follows

Index	Parameter Name	pre-BANFF fit fractional error	post-BANFF fit fractional error	Energy bin (GeV)
1	FSI inelastic low	0.000	0.118	0 - 30.0
2	FSI inelastic high	0.000	0.445	0 - 30.0
3	FSI Pion Production	0.000	-0.685	0 - 30.0
4	FSI Pion Absorption	0.000	-0.270	0 - 30.0
5	FSI Charge Exchange low	0.000	0.360	0 - 30.0
6	FSI Charge Exchange high	0.000	-0.381	0 - 30.0
7	M_A^{QE}	1.000	1.025	0 - 30.0
8	M_A^{RES}	1.163	0.797	0 - 30.0
9	CC other	0.000	0.225	0 - 30.0
10	$SF^{12}C$	0.000	0.240	0 - 30.0
11	E_b ^{12}C	1.000	1.236	0 - 30.0
12	p_F ^{12}C	1.000	1.227	0 - 30.0
13	π less Δ decay	0.000	0.006	0 - 30.0
14	CCQE E1	1.000	0.966	0 - 1.5
15	CCQE E2	1.000	0.931	1.5 - 3.5
16	CCQE E3	1.000	0.0.852	3.5 - 30.0
17	CC1π E1	1.154	1.265	0 - 2.5
18	CC1π E2	1.000	1.122	2.5 - 30.0
19	CC coherent	1.000	0.449	0 - 30.0
20	NC other	1.000	1.410	0 - 30.0
21	NC1π^0	0.963	1.135	0 - 30.0

Table 8.2: The neutrino-nucleus interaction and final state interactions models fractional errors provided by the T2K neutrino interaction working group (NIWG) [12].

- $\mathbf{M_A^{QE}}$
  The axial mass in the axial vector form factor, M_A^{QE}, has a fitted value of 1.64 ± 0.03 GeV, which is obtained from the MiniBooNE CCQE data. The uncertainty on this value is calculated as a difference between the fitted value of M_A^{QE} and the nominal NEUT M_A^{QE} value that is 1.21 GeV.

- $\mathbf{M_A^{RES}}$
  The axial mass in the resonant interactions, M_A^{RES}, is estimated from various MiniMooNE data sets.

- **CC other**
  CC other is a shape parameter that modifies a combination of CC cross section channels such as multi pion, DIS and resonant production as a function of neutrino energy. The cross section results from the MINOS collaboration imply an error for this parameter of 10% at the energy of 4 GeV.

- **SF**
  The spectral function, SF, is a function that describes a nuclear potential and defines a probability distribution of nucleon momenta and energies.

- $\mathbf{E_B}$ and $\mathbf{p_F}$
  The CCQE interactions simulation utilizes a relativistic Fermi gas model to recreate a nucleus. The Fermi gas model assumes a uniform distribution of nucleons within the nucleus bounded by the constant energy. There are two parameters: the Fermi momentum, p_F, and the nuclear binding energy, E_B that describe this theoretical model. Their values are extracted from the various electron scattering data.

- π **less** $\mathbf{\Delta}$ **decay**
  A pionless Δ decay accounts for 20% of all Δ decay events and constitutes irreducible background to the CCQE fits.

- **CCQE** and **CC1**π series
  The CCQE E1 is the low energy CCQE normalisation factor included in the uncertainty of the MiniMooNE flux, where it has an error of 11% with the nominal value of 1.0. The low energy $CC1\pi$ E1 normalisation is also estimated from the MiniBooNE data. The high energy normalisation parameters CCQE E2, CCQE E3 and CC1π are not constrained by any experimental data set and have an assigned uncertainty of 30% to account for the discrepancy in results from MiniBooNE and NOMAD experiments.

- **CC coherent**, **NC other** and **NCπ^0**
  The $NC\pi^0$ normalisation and CC coherent normalisation factors are obtained from the fit of different MiniBooNE data sets. The NC coherent parameter has to be constrained in association with $NC\pi^0$ due to the negligible difference in shape for the absolute π^0 momentum spectrum.

Similarly as for the flux parameters, the constraint from the ND280 detector can be added to the NEUT generator parameters as shown in Figure 8.3. The

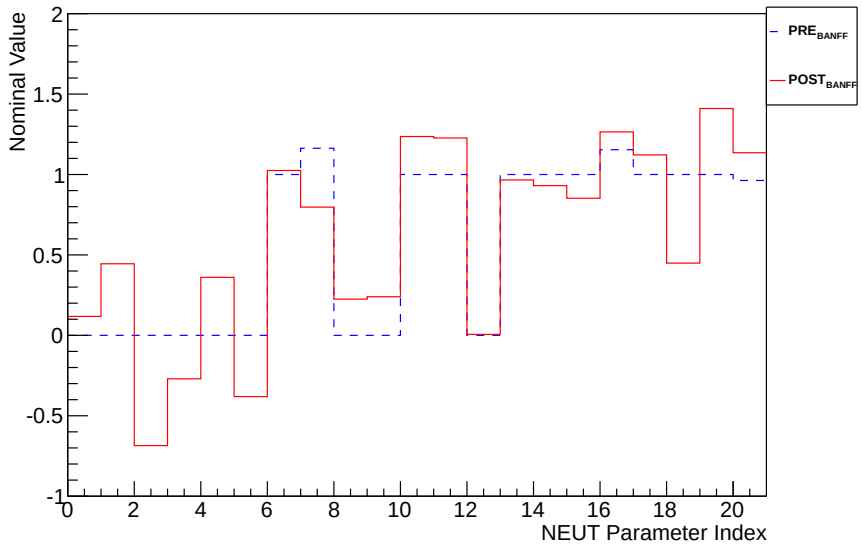
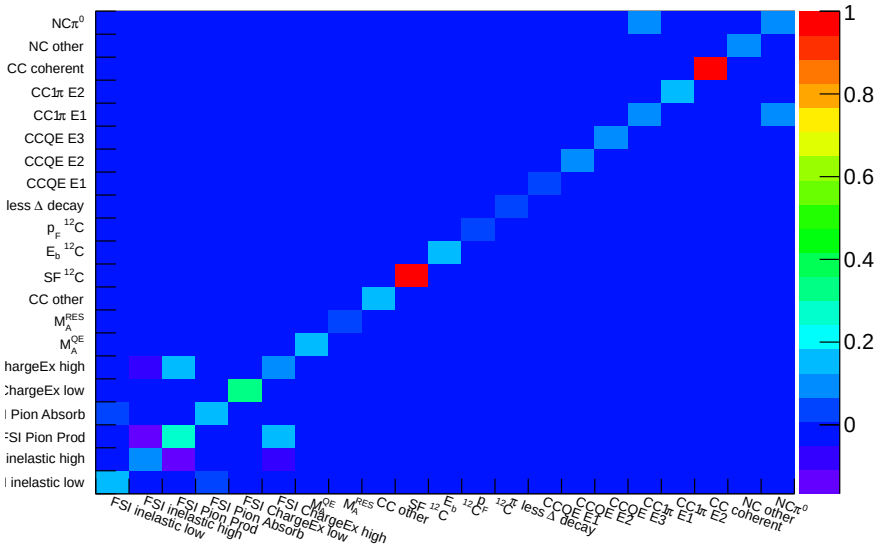


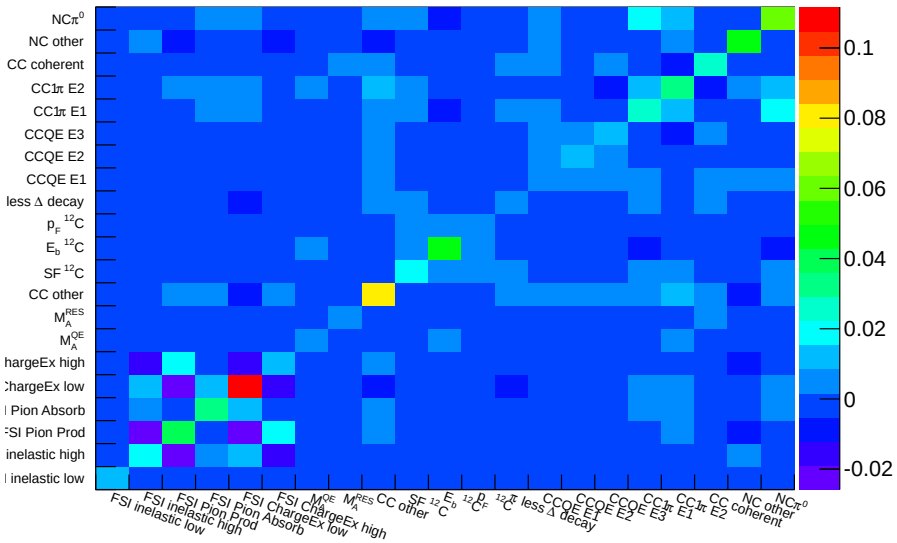
Figure 8.3: Nominal values of NEUT parameters prior to and after the BANFF fit. The first 6 parameters are related to the final state interaction parameters, and remaining 15 parameters concern cross section modelling in the NEUT generator.

correlation between FSI and cross section parameters have been formed into the matrix and illustrated in Figure 8.4.

Before the BANFF fit application there was no correlation between the flux and cross section parameters. After the BANFF fit such correlation appears and is shown in Figure 8.5. The implementation of the calculated flux, cross section and FSI uncertainties into analysis can be done using a specifically dedicated package, called T2KReWeight, which is a convenient tool within the ND280 software.



(a) pre-BANFF



(b) post-BANFF

Figure 8.4: Input covariance matrices for the NEUT generator FSI and cross section modelling parameters.

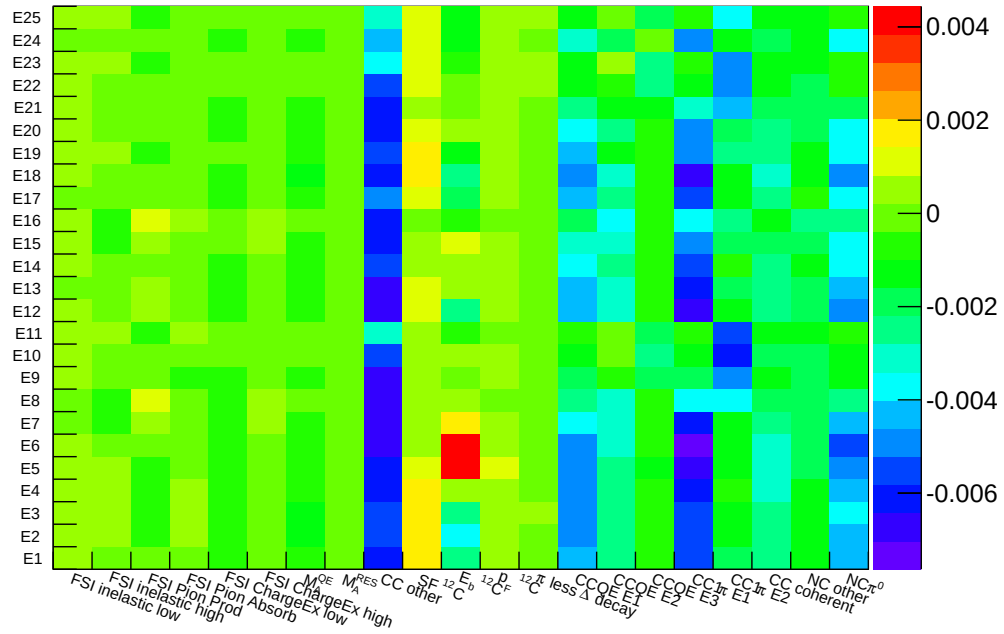


Figure 8.5: Correlation matrix of flux and NEUT parameters for post BANFF fit.

8.3 MVA Uncertainties

The main contribution to the uncertainties, that are related to the applied multivariate analysis technique, arises from data-MC disagreement of the input discriminating variables. These discrepancies consistently cause disproportions between events that pass over the nodes in the forest of decision trees. Therefore, it is important to estimate the impact of such an effect, which can be done in a sequence of steps. Initially, the data-MC ratio has to be calculated for each variable that is forwarded to the BDTcat training. The ratios are determined on the area normalised histograms and they constitute the set of weights. Next, the variable with the distinctive signal and background distribution shapes is expressed as a function of the input variables and used to apply the relevant calculated weights. In the NC1π^0 analysis usually such variable is chosen to be a π^0 reconstructed mass however, as explained is the previous chapter, it is not possible for the mass to be determined in this thesis study. The reweighted histogram shows the difference between the rescaled and nominal MC, which gives a systematic error arising from the input variable. The sum of a quadrature of individual variables

errors constitutes a total MVA addition to the systematic error.

The next factor to consider is related to the separation power of each input variable, which influences MVA uncertainties. The calculation of such errors can be achieved by including the scaling of MC with the weights from the initial data over MC ratios. This scaling factor is proportional to each input variable importance in the BDTcat method.

As shown in section 7.2.4, the majority of the BDTcat input variable distributions demonstrate a good data-MC agreement, with the exception of two variables. These variables represent the spatial and timing relation between the PØD and PØD-ECal clusters and the origin of the discrepancy is explained in section 7.2.4.

In this thesis analysis the discrimination variables are chosen to have a small correlation as shown in Figures 7.24 and 7.25. Therefore, each variable can be considered as an independent and the correlation factor can be omitted. This obviously might underestimate the MVA contribution to the systematic error, however the possible effect is expected to be minimal.

8.4 Detector Systematics

The uncertainties related to the detector concerning PØD and PØD-ECal sub-detectors and are briefly discussed in this section. The systematics for the PØD have been developed by the PØD working group on the number of analysis. The PØD-ECal systematic errors are not established yet and will be developed at the next stage of PØD-ECal studies.

8.4.1 PØD

The PØD sub-detector systematic uncertainties account for a number of variations related to its structure and readout. As shown in Ref. [76], channel to channel variations have a negligible effect to the overall systematics and are estimated to be less than 1%. The next important detector variation is a charge deposit response that fluctuates over time. For the majority of the recorded data, i.e. Runs 1, 2 and 3, this effect is fully calibrated out. Run 4 calibration is not fully processed in Production 5 and each peak gets a separate calibration constant. However, this variation effect is small and can be studied on the cosmic data sample. The systematic uncertainty related to the detector response over time in the PØD standalone analysis is found to be 1.8%. The other PØD sub-detector related uncertainties have been studied by the PØD working group and for fully

contained NC1π^0 analysis and are described in Ref. [75]. A short summary of these uncertainties is presented below.

Target Energy Scale

The energy scale effects can influence the analysis outcome in the three ways. The first element concerns geometry difference caused by two different targets: water and air. This concerns changes of sub-detector density in the water configurations as well as the minor changes in its dry mass. The change in the PØD mass is found to be approximately 155.5%. The next energy scale uncertainty effect comes from the PE peak uncertainty. In Production 5 the photo-electron peak has been imprecisely simulated for both water configurations. The average of the data to Monte Carlo discrepancy is giving the systematic of 0.6% (0,4%) for the PØD water-in (water-out) configuration. The last element is the energy scale fit performed with Gaussian function. As a result it is expected for the two PØD water configurations to have distinct energy resolutions followed by different systematic errors values.

Alignment and Fiducial Volume

Studies on the PØD sub-detector alignment are presented in Ref. [83] and are reported to be less than 2 mm. Alignment has a negligible effect on the shift of the fiducial volume in "z" direction. This is related to the fact that FV boundaries are located in the middle of the PØDule. However, the change of FV related to the alignment, in the "x" and "y" directions is found to be 0.31%. Additionally, there is another element that contributes to the systematic error related to the fiducial volume per se. It concerns FV scaling of data and Monte Carlo. This error accounts for the migration of selected vertices into and out of the fiducial volume if the volume definition changes within its uncertainty, e.g. vertex that is created at the upstream FV edge is more likely to be included within fiducial volume than one at the downstream edge. This reflects the reconstruction bias between the data and MC. In the NC1π^0 PØD standalone analysis these biases are found to be relatively small. As the PØD-PØD-ECal studies are shifting the PØD defined FV boundaries, the new FV systematic error estimation is required.

8.4.2 PØD-ECal

There is an uncertainty associated with the PØD-ECal performance and detection quality that influences final analysis results. The main source of this type of

error originates in a difference between the MC simulation and real data. The set of such systematic errors has not been developed at the time of writing the thesis. The possible factors related to the PØD-ECal sub-detector performance are shortly explained below.

MPPC Response

The MPPC simulation response diverges from the response during real data taking by around 10%. This difference is currently not calibrated out in Production 5 and contributes to the systematic error. The uncertainty can be estimated by comparing the number of selected events with and without the MPPC response correction.

Noise

The next uncertainty is related to the difference between the number of noise hits simulated in MC and these present in data. The convenient way to estimate an error from such effect is to compare the number of cosmic MC true noise hits to the measured number of noise hits from the cosmic muon data.

PØD-ECal Geometry

The specifications of ECAL module geometry that are used in the MC simulation originate from the construction geometry data sheets. However, the uncertainties related to the PØD-ECal moduls components such as scintillator bar dimensions, amount of dead material and coating material can differ. Additionally, the response of active material, dependent on the WLS fibre position within the bar, can vary from bar to bar.

PØD-PØD-ECal Matching Uncertainty

The matching uncertainty of a particle footprint, which passed through the PØD and PØD-ECal sub-detectors, depends on the time and position readout accuracy and sub-detectors alignment. The timing is a subject to a cross detector time calibration. These two elements are not been examined yet and are to be developed in a future analysis. The accuracy of the signal synchronisation between two sub-detectors determines the matching efficiency between the PØD and PØD-ECal reconstructed objects.

Chapter 9

Conclusions

Neutrino experimental physics is progressing rapidly since neutrino oscillations have been experimentally confirmed. Currently the majority of experiments are focusing their studies on the precision measurements since more accurate understanding of neutrino interactions in crucial to determine the neutrinos mass hierarchy and the CP-violating phase δ_{CP}. The most intense neutrino beam at the T2K experiment enables the interactions to be studied at the near detector ND280, including NC1π^0.

The main goal of the analysis performed in this thesis was a development of the NC1π^0 event selection at the near detector ND280 using the the implementation of the PØD-ECal sub-detector information. The presented analysis constitute the first look at neutrino neutral current interactions that produce a π^0 whose decay photons leave a trace in the PØD-ECal sub-detector. The studies were performed for the two PØD configurations: filled with and without water. This also gave the insight to the most recent PØD-ECal sub-detector status.

The simplicity of the PØD-ECal structure and its coarse nature results in the lack of reconstructed energy and momentum of the passing particles which constitutes a limiting factor to the event selection final result. The lack of the beam direction position variable upsets the spatial information between reconstructed objects and the limited accuracy in geometry. It was shown, that at the reconstruction stage the clustering and PID algorithms require further improvements, which is partially related to the lack of the time calibration which is to be implemented in the next data and MC production. The main conclusion that arises from the analysis, is that despite the need for further PØD-ECal preparation work, it has the ability to contribute to the physics analysis. However, the current PØD-ECal status restricts it from providing the full information about the underlying physics. There is a number of challenges that are being faced

and the work is done to overcome these challenges. In particular the improvements in calibration and object reconstruction algorithms is required to help to achieve the goals. However, despite the above disadvantages, the addition of the PØD-ECal allowed the PØD fiducial volume to be increased hence the area of neutrinos interaction expanded by 26.6% (29.4%) for the PØD filled with (without) water. Therefore, since NC1π^0 measurement is a counting experiment, the PØD-ECal information can boost the detection of NC1π^0 events for neutrinos interacting in the PØD.

The NC1π^0 event selection was developed with particular focus on the spatial and time relation between the PØD and PØD-ECal reconstructed clusters in order to match two showering gamma candidates. The matching algorithm was conducted using the BDT multi-variate analysis which selected the decay photons converted in the PØD and PØD-ECal. It was done using a number of input variables representing the time and geometric relation between the PØD and PØD-ECal clusters. Considering the lack of the reconstructed kinematics in the PØD-ECal, the correlations between position, direction and deposited charge was maximised. It has been shown, that the output of the BDT performance constituted the most powerful NC1π^0 selection criterion for the purity increase. The demonstration of the matching analysis presented here also highlights how useful it would be to see the effect of PØD-ECal time calibration. Therefore, further progress on the matching algorithm between the PØD and PØD-ECal clusters might increase the power of BDT separation technique and improve the analysis event selection.

In general, the NC1π^0 measurements are characterised with low purity in comparison to the other type of studied events. The lack of accurate position, PID and event kinematics in this thesis analysis leads to the expectation of relatively low purity outcome. The event selection described in the section 7.2 showed that it is possible to select neutrino interactions with PØD-ECal input, however, with accordingly limited accuracy. The NC1π^0 event selection total efficiency of 3.0% (4.0%) and purity of 17.7% (16.4%) is obtained for the PØD filled with water (without water) configuration. Finally, after selection, 68 ± 8.2 (stat) (107 ± 10.3 (stat)) events passed all the requirements for the PØD filled with (without) water. The first look into data shows a significant discrepancy between data and MC which needs to be investigated. The discrepancies between MC and data events in the final result lead to the conclusion that revision of understanding of the MC and reconstruction is required to correctly select and describe events using PØD-ECal information. In particular, data MC disagreement in the BDT

input discrimination variables are expected to lead towards the large systematic uncertainties. During the current stage of the T2K experiment, the collected data constitutes still a limited statistics for the PØD-ECal events. The higher statistics will allow for side band studies and systematic error evaluation. In chapter 8, the effects of the neutrino generator and flux uncertainties are discussed together with detector and matching uncertainties.

The first application of the PØD-ECal sub-detector has shown, that despite of the simplicity in the sub-detector construction and limited reconstruction information, it is capable of contributing to the physics analysis. The performed studies provide a good insights into the current status of the PØD-ECal sub-detector and constitute the groundwork for the future measurement of the NC1π^0 cross section with neutrinos interacting in the PØD sub-detector. The presented NC1π^0 selection become a good groundwork that can be used as a tool for the future analysis. The focus should be on the improvement of the NC1π^0 event selection efficiency and purity which can be achieved with greater data samples.

There is an indication that the implementation of the suggested PØD-ECal improvements will help to achieve the expectations for the PØD-ECal contribution to the future NC1π^0 measurement. The PØD-ECal contribution can extend beyond the analysis presented in this thesis. In particular, it can help in the PØD-PØD-ECal charged current analysis where matching between the reconstructed track might be more beneficial. This would also contribute to remove the CC interactions from the NC analysis, where it constitutes a main background. It is shown, that the PØD-ECal can be used to identify the π^0, which indicates that the analysis with PØD-ECal usage will provide a significant contribution to the inclusive NC1π^0 cross section measurement.

Bibliography

[1] S.F. King et al. Neutrino mass and mixing with discrete symmetry. *Rept.Prog.Phys.*, 76(056201), 2013. (Cited on pages 9 and 30.)

[2] J.L. Hewett et al. Fundamental physics at the intensity frontier. *arXiv:1205.2671*, 2012. (Cited on pages 9 and 35.)

[3] G.P. Zeller J.A. Formaggio. From eV to EeV: Neutrino cross sections across energy scales. *Rev. Mod. Phys.*, 84(1307), 2012. (Cited on pages 9 and 40.)

[4] The plot is a courtesy of Andrew Furmanski (T2K) from the University of Warwick (2015). (Cited on pages 9 and 40.)

[5] T2K Collaboration. The T2K neutrino flux prediction. *Phys. Rev. D*, 87 (012001), 2013. (Cited on pages 9, 11, 46, 80 and 81.)

[6] K. Abe et al. Measurements of the T2K neutrino beam properties using the INGRID on-axis near detector. *arXiv:1111.3119*, 2012. (Cited on pages 9 and 49.)

[7] The plot is a courtesy of the T2K INGRID Collaboration group (2014 - 2015). (Cited on pages 10 and 51.)

[8] T2K Collaboration. The T2K experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectometers, Detectors and Assosciated Equipment*, 659(106), 2011. (Cited on pages 10, 11, 45, 47, 52 and 80.)

[9] F. Di Lodovico et al. Data quality for the near detectors for Run 2. Technical Report T2K-TN-050, 2011. (Cited on pages 15, 66, 97 and 128.)

[10] Particle data group. `http://pdg.lbl.gov`. (Cited on pages 18, 27, 33 and 41.)

[11] K. Gilje. Track PID efficiency in the π^0 detector in the ND280 basket. Technical Report T2K-TN-200, , 2014. (Cited on pages 18, 86, 96 and 97.)

[12] A. Kaboth M. Hartz and K. Mahn. Constraining the flux and cross section models with data from the ND280 detector for the 2013 oscillation analysis. Technical Report T2K-TN-166, 2013. (Cited on pages 19, 150, 152 and 154.)

[13] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett*, 12(132), 1964. (Cited on page 26.)

[14] P.W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Lett*, 13(508), 1964. (Cited on page 26.)

[15] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13(321), 1964. (Cited on page 26.)

[16] C. W. Kim C. Giunti. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 2007. (Cited on pages 26 and 32.)

[17] Jose W. F. Valle J. Schechter. Neutrino masses in SU(2) $\otimes$ U(1) theories. *Phys. Rev.*, 22(9), 1980. (Cited on page 26.)

[18] KATRIN Collaboration. KATRIN, a next generation tritium decay experiment in search for the absolute neutrino mass scale. *Progress in Particle and Nuclear Physics*, 48(1), 202. (Cited on page 27.)

[19] Alexander Barabash for the SuperNEMO Collaboration. SuperNEMO double beta decay experiment. *arXiv:1112.1784 [nucl-ex]*, 2011. (Cited on page 27.)

[20] M.C.Chen. The SNO liquid scintillator project. *Nucl. Phys. B (Proc. Suppl.)*, 145(65-68), 2005. (Cited on page 27.)

[21] B. Pontecorvo. Mesonium and anti-mesonium. *Zh. Eksp. Teor. Fiz. (JETP)*, 33(549), 1957. (Cited on page 30.)

[22] B. Pontecorvo. Inverse beta processes and nonconservation of lepton charge. *Zh. Eksp. Teor. Fiz. (JETP)*, 34(247), 1958. (Cited on page 30.)

[23] B. Pontecorvo. Neutrino experiments and the problem of conservation of leptonic charge. *Zh. Eksp. Teor. Fiz. (JETP)*, 53(1717), 1967. (Cited on page 30.)

[24] S. Sakata Z. Maki, M. Nakagawa. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28(870), 1962. (Cited on page 30.)

[25] M. H. Pinsonneault J. N. Bahcall and S. Basu. Solar models: current epoch and time dependences, neutrinos, and helioseismological properties. *The Astrophysical Journal*, 555(p.990-1012), 2001. (Cited on page 32.)

[26] B. T. Cleveland et al. Measurement of the solar electron neutrino flux with the Homestake chlorine detector. *The Astrophysical Journal*, 496 no. 1(505), 1998. (Cited on page 32.)

[27] S. Fukuda et al. Determination of solar neutrino oscillation parameters using 1496 days of Super-Kamiokande I data. *Physics Letters*, B 39(34), 2002. (Cited on page 32.)

[28] SNO Collaboration. Measurement of the rate of ν_e+d to p+p+e^- intractions produced by 8b solar neutrinos at the Sudbury neutrino observatory. *Phys. Rev. Lett.*, 87, 2001. (Cited on page 33.)

[29] KamLAND Collaboration. Precision measurement of neutrino oscillation parameters with Kamland. *Phys. Rev. Lett.*, 100(221803), 2008. (Cited on page 33.)

[30] K2K Collaboration. Measurement of neutrino oscillation by the K2K experiment. *Phys. Rev.*, D 74(072003), 2006. (Cited on page 34.)

[31] MINOS Collaboration. Observation of muon neutrino disappearance with the Minos detectors in the ν_μ neutrino beam. *Phys. Rev. Lett.*, 97(191801), 2006. (Cited on page 34.)

[32] T2K Collaboration. Precise measurement of the neutrino mixing parameter θ_{23} from muon neutrino disappearance in an off-axis beam. *Phys. Rev.*, D 89(092003), 2014. (Cited on page 34.)

[33] T2K Collaboration. Indication of electron neutrino appearance from an accelerator-produced off-axis muon neutrino beam. *Phys.Rev.Lett.*, 107 (041801), 2011. (Cited on page 35.)

[34] F. P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 108(171803), 2012. (Cited on page 35.)

[35] RENO Collaboration. Observation of reactor electron antineutrinos disappearance in the RENO experiment. *Phys. Rev. Lett.*, 108(191802), 2012. (Cited on page 35.)

[36] T2K Collaboration. Observation of electron neutrino appearance in a muon neutrino beam. *Phys. Rev. Lett.*, 112(061802), 2014. (Cited on page 35.)

[37] MINERvA Collaboration. Design, calibration, and performance of the MINERvA detector. *Nucl. Inst. and Meth.*, A743(130), 2014. (Cited on page 39.)

[38] C. Anderson et al. The ArgoNeuT detector in the NuMI low-energy beam line at Fermilab. *JINST*, 7 P10019, 2012. (Cited on page 39.)

[39] MicroBooNE Collaboration. The MicroBooNE technical design report. *TDR from CD3b review*, 2012. (Cited on page 39.)

[40] D. Rein and L. M. Sehgal. Neutrino excitation of baryon resonances and single pion production. *Annals Phys.*, 133(79), 1981. (Cited on pages 42 and 82.)

[41] J.T. Sobczyk J.G. Morfin, Juan Nieves. Recent developments in neutrino/antineutrino - nucleus interactions. *arXiv:1209.6586v1*, 2012. (Cited on page 43.)

[42] K2K Collaboration. Measurement of single π^0 production in neutral current neutrino interactions with water by a 1.3-GeV wide band muon neutrino beam. *Phys.Lett.B*, 619(255-262), 2005. (Cited on page 44.)

[43] SciBooNE Collaboration. Measurement of inclusive neutral current neutral pion production on carbon in a Few-GeV neutrino beam. *Phys.Rev.D*, 81 (033004), 2010. (Cited on page 44.)

[44] SciBooNE Collaboration. Measurement of inclusive charged current interactions on carbon in a few-GeV neutrino beam. *Phys.Rev.D*, 83(111102(R)), 2010. (Cited on page 44.)

[45] MiniBooNE Collaboration. Measurement of ν_μ and $\bar{\nu}_\mu$ induced neutral current single π^0 production cross sections on mineral oil at E_νO(1 GeV). *Phys.Rev.D*, 81(013005), 2010. (Cited on page 44.)

[46] INGRID working group. Measurement of the flux averaged inclusive ν_μ charged current cross section with INGRID and Proton Module. Technical Report T2K-TN-160, 2014. (Cited on page 50.)

[47] S. Aoki et al. The T2K side muon range detector (SMRD). *arXiv:1206.3553*, 2012. (Cited on page 53.)

[48] T2K ND280 FGD Collaboration. The T2K fine-grained detectors. *Nucl. Instrum. Meth.*, A 696(1), 2012. (Cited on page 54.)

[49] T2K ND280 TPC Collaboration. Time projection chambers for the T2K near detectors. *arXiv:1012.0865*, 2010. (Cited on page 54.)

[50] S. Assylbekov et al. The T2K ND280 off-axis π^0 detector. *Nucl. Instrum. Meth.*, A 686(48), 2012. (Cited on page 56.)

[51] D. Allan et al. The electromagnetic calorimeter for the T2K near detector ND280. *JINST*, 8(P10019), 2013. (Cited on page 58.)

[52] Presentation at:. `http://www.t2k.org/nd280/beam/beam_summary/`. (Cited on page 63.)

[53] K. Iyogi et al. T2K data acquisition and FC event selection at Super-Kamiokande. Technical Report T2K-TN-027, 2011. (Cited on page 66.)

[54] F. Di Lodovico et al. Quality assessment strategy of the 2010b data set at ND280. Technical Report T2K-TN-021, 2011. (Cited on page 66.)

[55] F. Di Lodovico et al. Quality assessment of the 2010a data set at ND280. Technical Report T2K-TN-013, 2010. (Cited on page 66.)

[56] F. Di Lodovico et al. Data quality at the near detectors for Run 3. Technical Report T2K-TN-127, 2012. (Cited on page 66.)

[57] F. Di Lodovico et al. Data quality at the newr detectors for Run 4. Technical Report T2K-TN-168, 2013. (Cited on page 66.)

[58] `http://www.irods.org`. (Cited on page 68.)

[59] `http://root.cern.ch`. (Cited on page 79.)

[60] A. Fasso A. Ferrari, P. R. Sala and J. Ranft. Fluka. A multi-particle transport code. (Cited on page 80.)

[61] N. Abgrall et al. Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c. *Phys. Rev.*, C84(034604), 2011. (Cited on page 80.)

[62] F. Carminati R. Brun and S. Giani. CERN-W5013. 1994. (Cited on page 81.)

[63] D. Casper. The NUANCE neutrino simulation, and the future. *Nucl. Phys. (Proc. Supp.)*, 112(161), 2002. hep-ph/0208030. (Cited on page 82.)

[64] C. Andreopoulos et al. The GENIE neutrino monte carlo generator. *Nucl. Instrum. Methods*, A614(87), 2010. hep-ph/09052517. (Cited on page 82.)

[65] Y. Hayato. NEUT. *Nucl. Phys. (Proc. Supp.)*, 112(171), 2002. (Cited on page 82.)

[66] C.H. Llewellyn Smith. Neutrino reactions at accelerator energies. *Phys.Rept.*, 3(261), 1972. (Cited on page 82.)

[67] R.A. Smith and E.J. Moniz. Neutrino reactions on nuclear targets. *Nucl.Phys.B*, 43(605), 1972. (Cited on page 82.)

[68] D. Rein and L. M. Sehgal. Coherent π^0 production in neutrino reactions. *Nucl.Phys.B*, 223(29), 1983. (Cited on page 82.)

[69] A. Vogt. M. Gluck, E. Reya. Dynamical parton distributions revisited. *Eur. Phys. J.*, C5:461470, 1998. (Cited on page 82.)

[70] A. Bodek and U.Yang. Axial and vector structure functions for electron and neutrino nucleon scattering cross sections at all Q^2 using effective leading order parton distribution functions. hep-ph/1011.6592. (Cited on page 82.)

[71] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science,*, 53:270–278, 2006. (Cited on page 83.)

[72] J. Gmez-Cadenas A. Cervera-Villanueva and J. Hernando. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 534(180), 2004. Proceedings of the IXth International Workshop on Advanced Computing and Analysis Techniques in Physics Research. (Cited on page 84.)

[73] PØD Working Group. PØD reconstruction. Technical Report T2K-TN-072, 2011. (Cited on pages 84 and 96.)

[74] B. Still. *T2K ND280 π^0 Electromagnetic Calorimeter*. PhD thesis, University of Sheffield, 2009. (Cited on pages 86 and 87.)

[75] C. McGrew H. Okeeffe K. Gilje, J. Adam. Measurement of the NC1π^0 rate on water in the PØD. Technical Report T2K-TN-144, 2014. (Cited on pages 89 and 160.)

[76] G. Lopez et al. Measurement of NC1π^0 production with the T2K π^0 detector (PØD). Technical Report T2K-TN-56, 2012. (Cited on pages 96 and 159.)

[77] K. Gilje. Photon energy scale of the π^0 detector in the ND280 basket. Technical Report T2K-TN-202, 2013. (Cited on page 97.)

[78] Multivariate analysis journal:. `http://www.journals.elsevier.com/journal-of-multivariate-analysis/`. (Cited on page 113.)

[79] J. Stelzer J. Therhaag E. von Toerne H. Voss A. Hoecker, P. Speckmayer. TMVA 4 toolkit for multivariate data analysis with ROOT, users guide. *arXiv:physics/0703039*, TMVA version 4.2.0, 2013. (Cited on page 113.)

[80] J. Zhu Y. Liu I. Stancu G. McGregor B.P.Roe, H. Yang. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl.Instrum.Meth.*, A543(577-584), 2005. (Cited on pages 114 and 116.)

[81] K. Mahn P. de Perio, Mark Hartz1 and S. Oser. Constraining the flux and cross section models with data from the ND280 detector for the 2012a oscillation analysis. Technical Report T2K-TN-106, 2012. (Cited on page 150.)

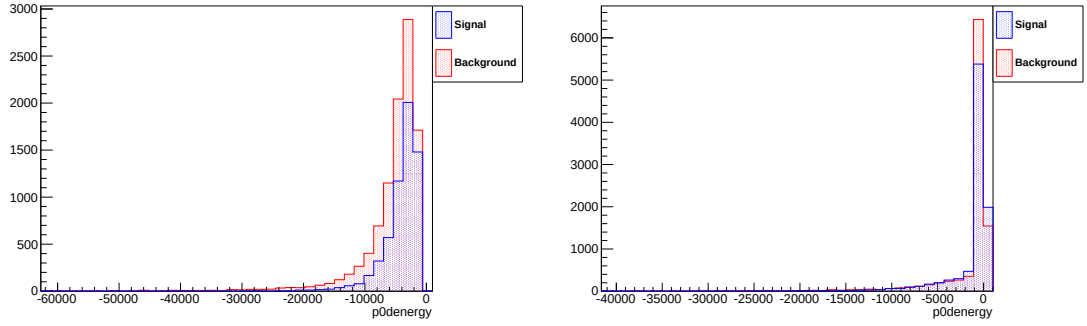
[82] V. Lyubushkin et al. A study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment. *Eur. Phys. J.*, C63, 2009. (Cited on page 153.)

[83] K. Gilje et al. Geometry and mass of the π^0 detector in the ND280 basket. Technical Report T2K-TN-073, 2012. (Cited on page 160.)

Appendix A

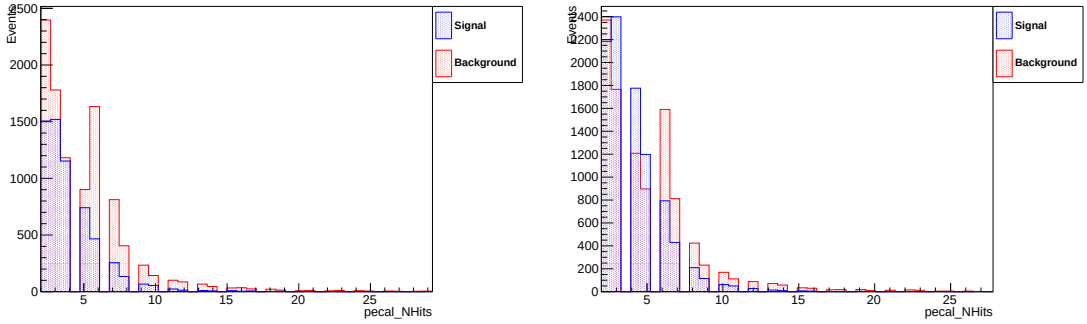
TMVA Input Variables

Plots demonstrating the distributions of TMVA input variables for signal and background studies of the NC1π^0 events which occur in the PØD and PØD-ECal sub-detectors for the PØD filled with (left-hand side plots) and without (right-hand side plots) water, described in chapter 7.



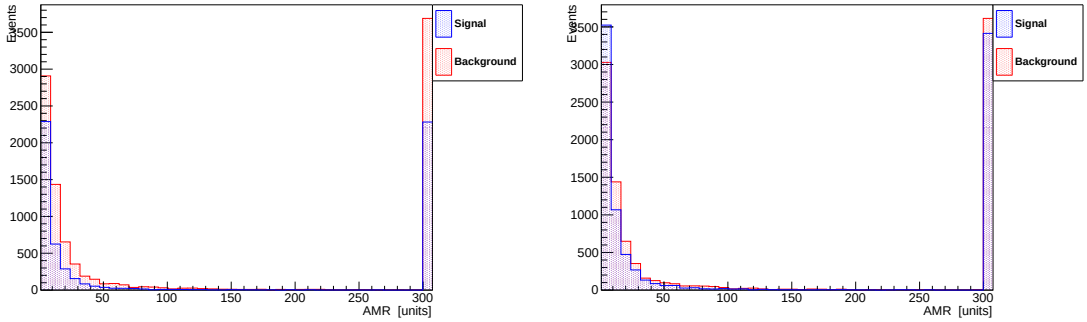
(a) Water (b) Air

Figure A.1: The energy of decay gamma showered in the PØD.



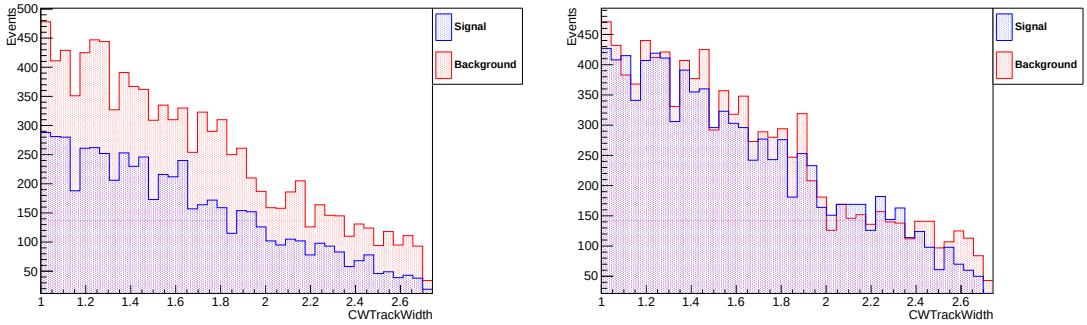
(a) Water (b) Air

Figure A.2: Number of hits in the PØD-ECal cluster.



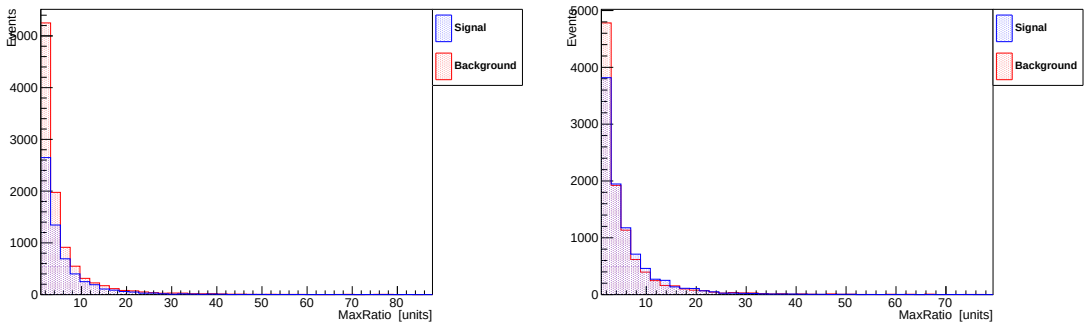
(a) Water (b) Air

Figure A.3: AMR, a PID discrimination variable for the PØD-ECal cluster.



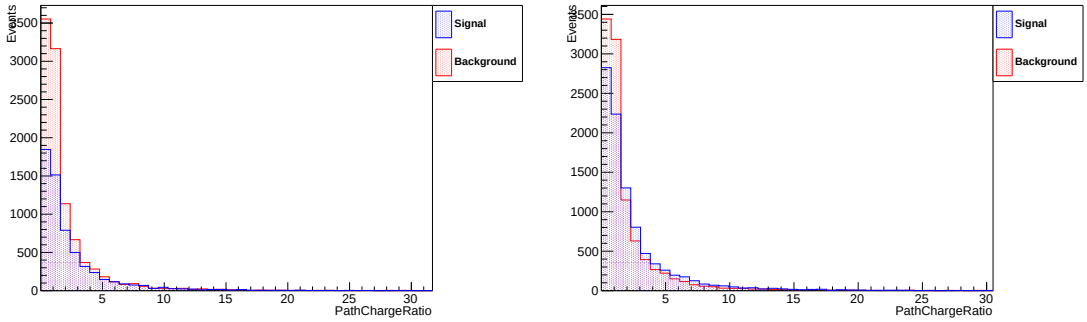
(a) Water (b) Air

Figure A.4: Charge Weighted Cluster Width, a PID discrimination variable for the PØD-ECal cluster.



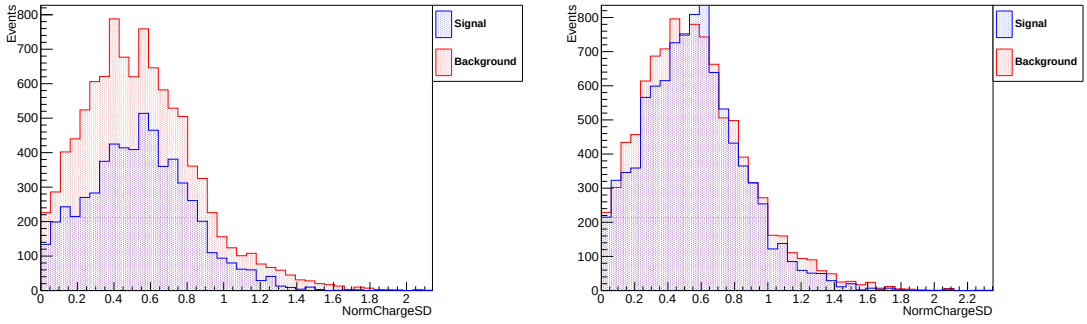
(a) Water (b) Air

Figure A.5: Ratio of PØD-ECal cluster axis, a PID discrimination variable for the PØD-ECal cluster.



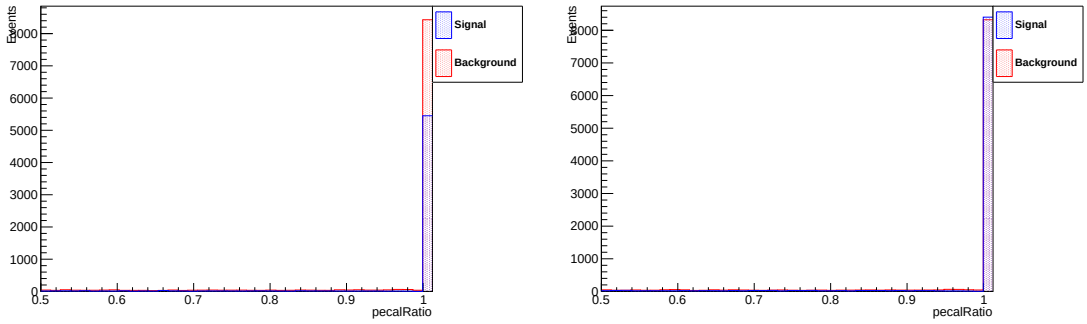
(a) Water (b) Air

Figure A.6: PathChargeRatio, a PID discrimination variable for the PØD-ECal cluster.



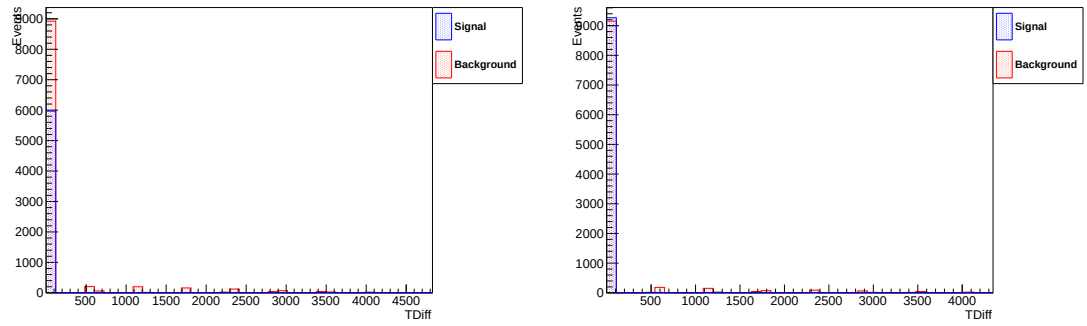
(a) Water (b) Air

Figure A.7: NormChargeSD, a PID discrimination variable for the PØD-ECal cluster.



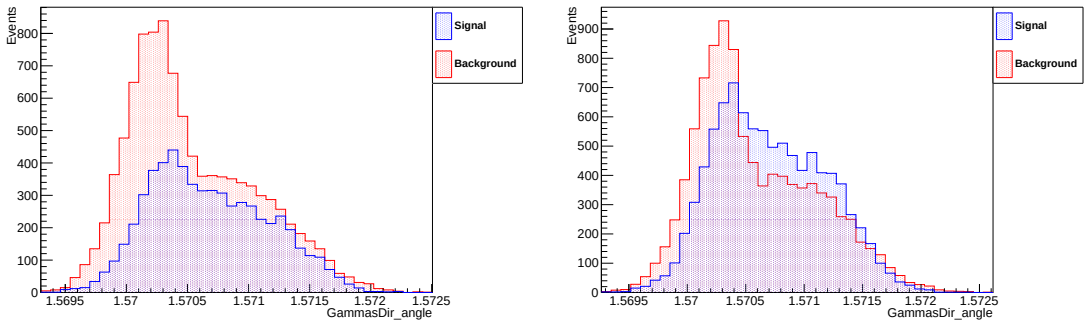
(a) Water (b) Air

Figure A.8: Ratio of the two PØD-ECal clusters charge, considered only for the second category of BDT discrimination.



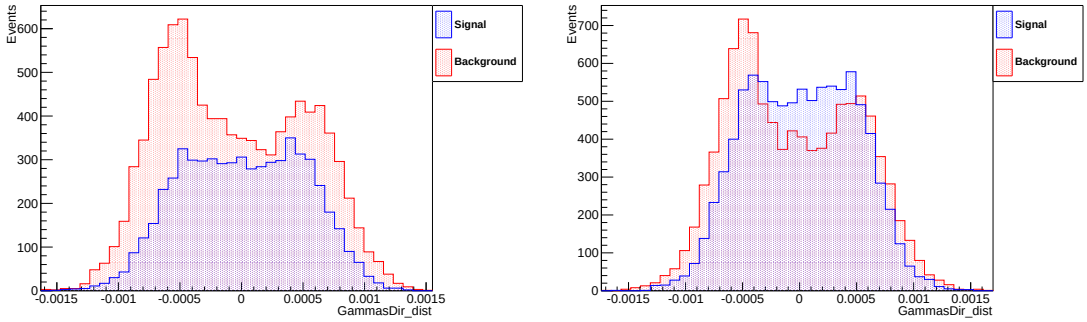
(a) Water (b) Air

Figure A.9: Time difference between the two reconstructed π^0 decay photons, one in the PØD and one in the PØD-ECal sub-detectors.



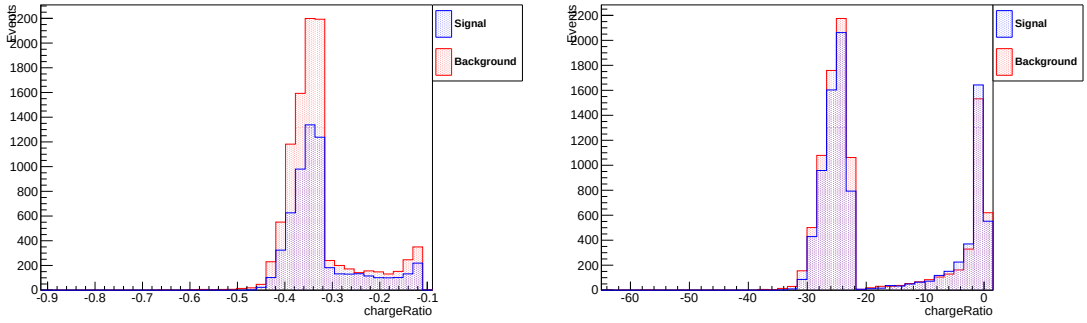
(a) Water (b) Air

Figure A.10: Angle between the two reconstructed π^0 decay photon directions.



(a) Water (b) Air

Figure A.11: Distance between the two reconstructed π^0 decay photon directions.



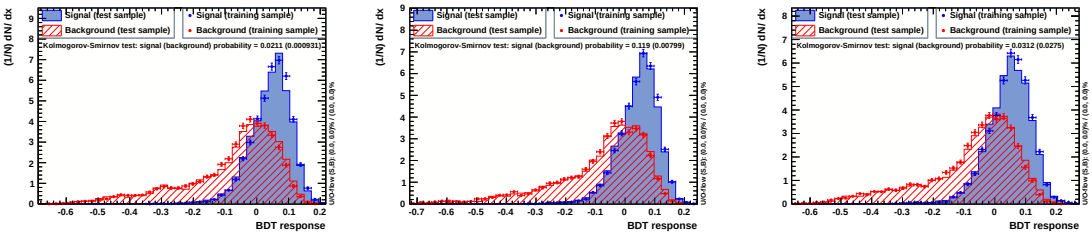
(a) Water (b) Air

Figure A.12: Ratio of the two reconstructed π^0 decay photons charge.

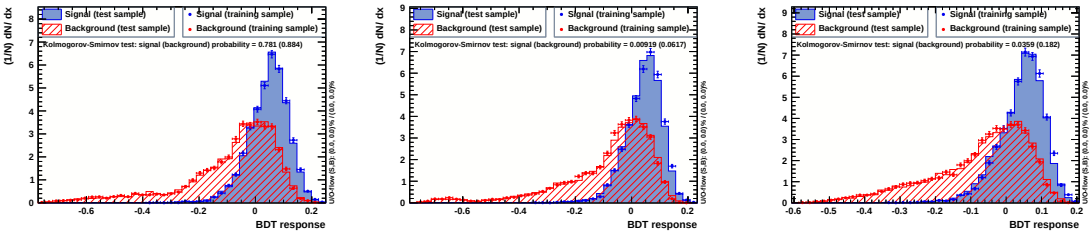
Appendix B

Fiducial Volume Cutting Value Optimisation in Correlation to the BDTcat Response.

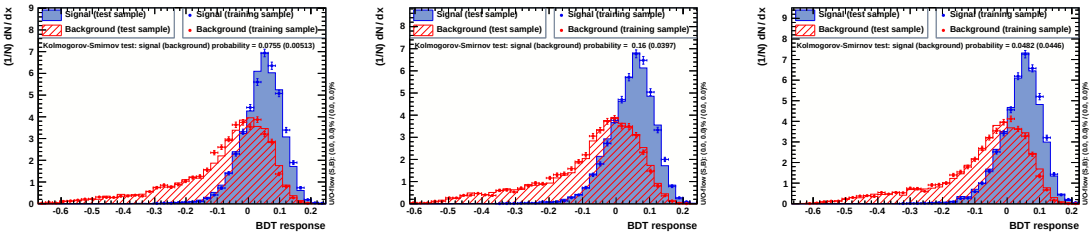
The fiducial volume cutting value has been chosen in correlation to the BDTcat training output. Plots demonstrating the correlation of the BDT cut with the different fiducial volume cutting values are shown in Figures B.1, B.2, B.3 and B.4 (B.5, B.6, B.7 and B.8) for the water-in (water-out) PØD configuration. Each of the plots displays the signal (blue) and background (red) events distributions, where the points represent a training sample response and the histogram represents a MC testing response. The FV value influences two important BDTcat elements: the strength of signal-background separation and agreement of the training and testing event samples. The first element, separation power, is illustrated in Figure 7.41 where the signal significance, calculated according to the Formula 7.7, shows the highest value of maximum $S/\sqrt{S+B}$ to be 24.1413 (23.9541) which corresponds to the FV of -55 mm (-79 mm) for the PØD filled with (without) water. The second element determines the events sample overtraining which should be minimised, otherwise it can defectively affect the final BDTcat response application outcome. Therefore, it is important to maximise the agreement between training and testing samples, which was determined visually. The samples agreement for the water-in PØD configuration was beyond acceptance, and therefore the second large signal significance value, 24.1054, was chosen. As a result of this two factors, the FV cutting value is chosen to be -74 mm (-79 mm) for the water-in (water-out) PØD configuration.



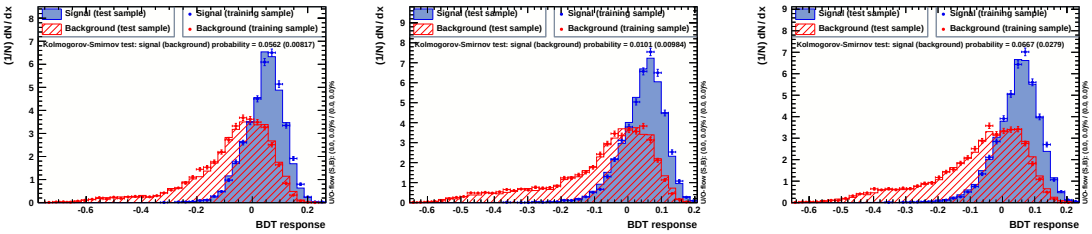
(a) Fiducial = -90 mm (b) Fiducial = -89 mm (c) Fiducial = -88 mm



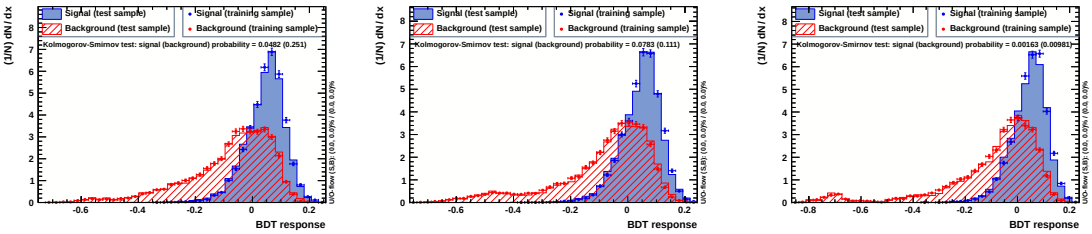
(d) Fiducial = -87 mm (e) Fiducial = -86 mm (f) Fiducial = -85 mm



(g) Fiducial = -84 mm (h) Fiducial = -83 mm (i) Fiducial = -82 mm

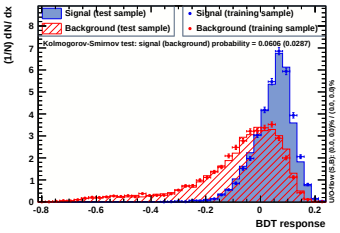


(j) Fiducial = -81 mm (k) Fiducial = -80 mm (l) Fiducial = -79 mm

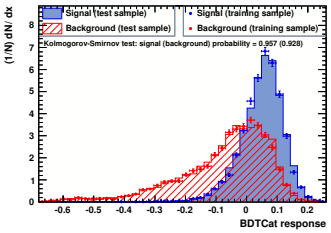


(m) Fiducial = -78 mm (n) Fiducial = -77 mm (o) Fiducial = -76 mm

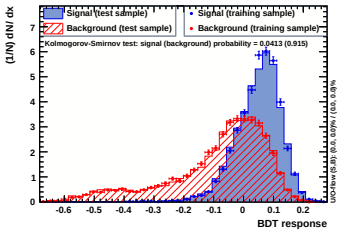
Figure B.1: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD filled with water.



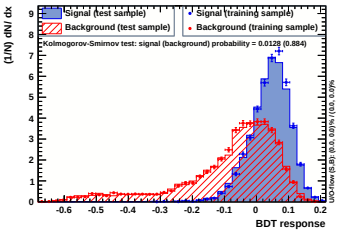
(a) Fiducial = -75 mm



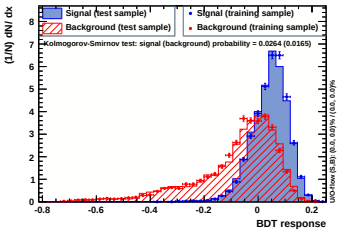
(b) Fiducial = -74 mm



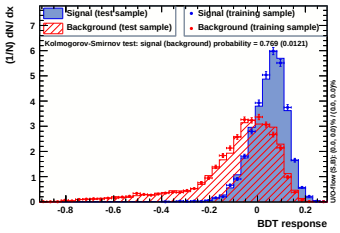
(c) Fiducial = -73 mm



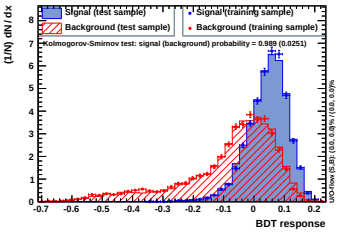
(d) Fiducial = -72 mm



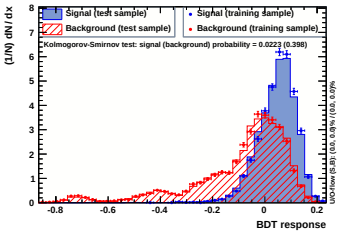
(e) Fiducial = -71 mm



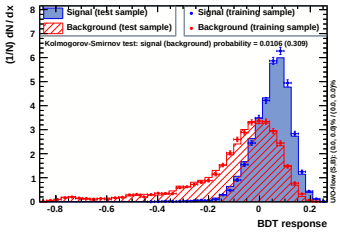
(f) Fiducial = -70 mm



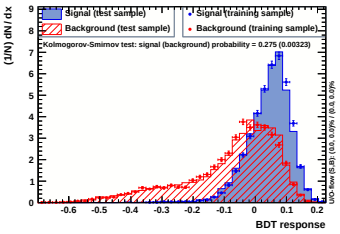
(g) Fiducial = -69 mm



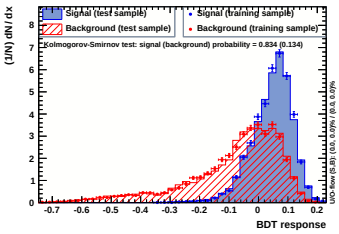
(h) Fiducial = -68 mm



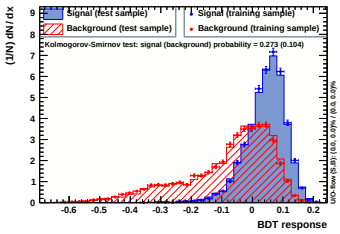
(i) Fiducial = -67 mm



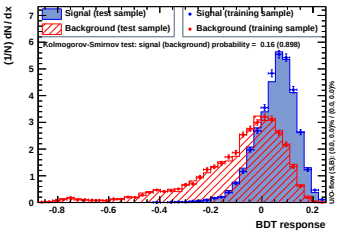
(j) Fiducial = -66 mm



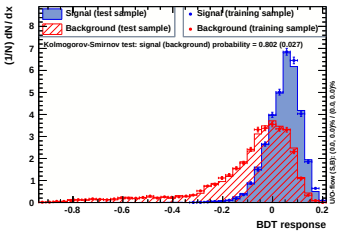
(k) Fiducial = -65 mm



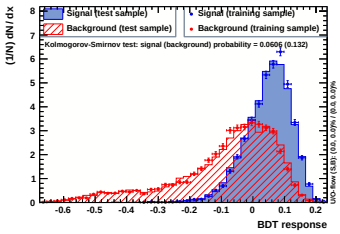
(l) Fiducial = -64 mm



(m) Fiducial = -63 mm

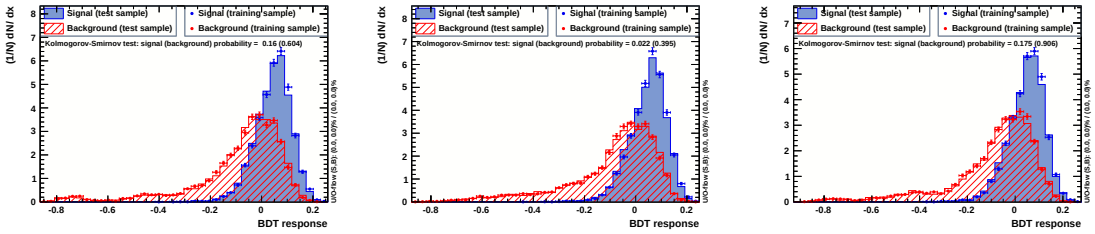


(n) Fiducial = -62 mm

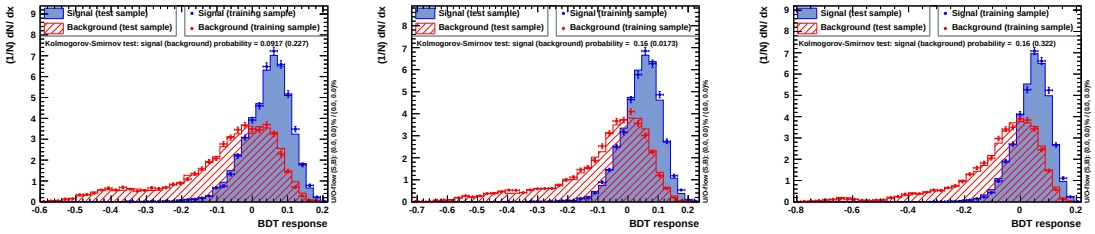


(o) Fiducial = -61 mm

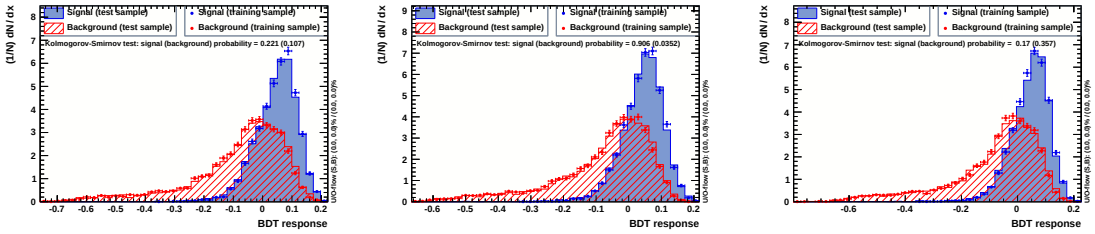
Figure B.2: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD filled with water.



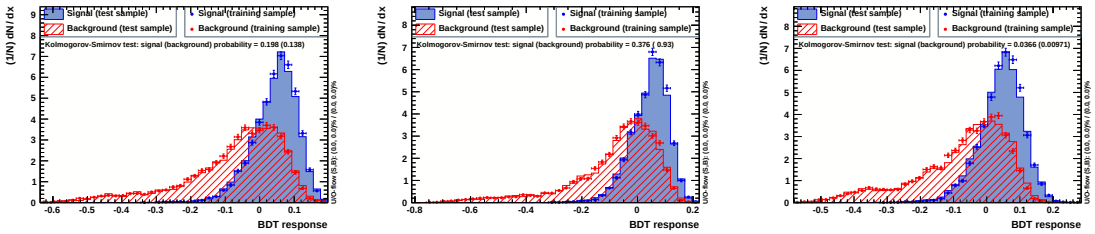
(a) Fiducial = -60 mm (b) Fiducial = -59 mm (c) Fiducial = -58 mm



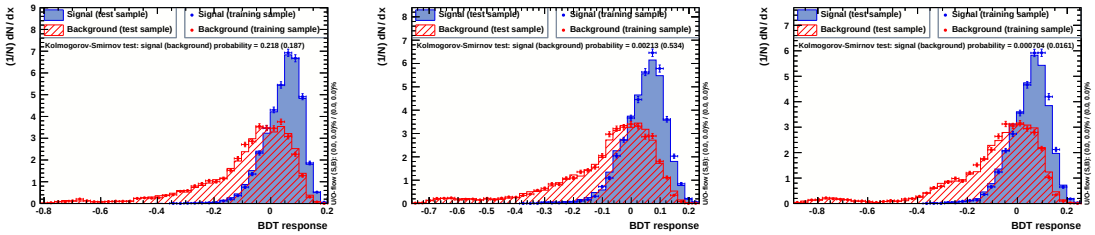
(d) Fiducial = -57 mm (e) Fiducial = -56 mm (f) Fiducial = -55 mm



(g) Fiducial = -54 mm (h) Fiducial = -53 mm (i) Fiducial = -52 mm

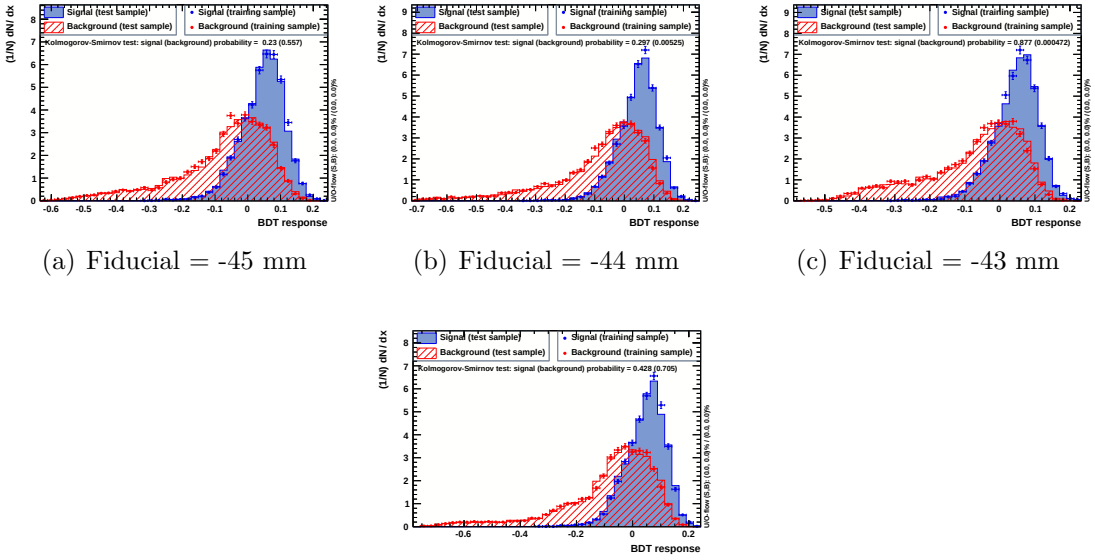


(j) Fiducial = -51 mm (k) Fiducial = -50 mm (l) Fiducial = -49 mm



(m) Fiducial = -48 mm (n) Fiducial = -47 mm (o) Fiducial = -46 mm

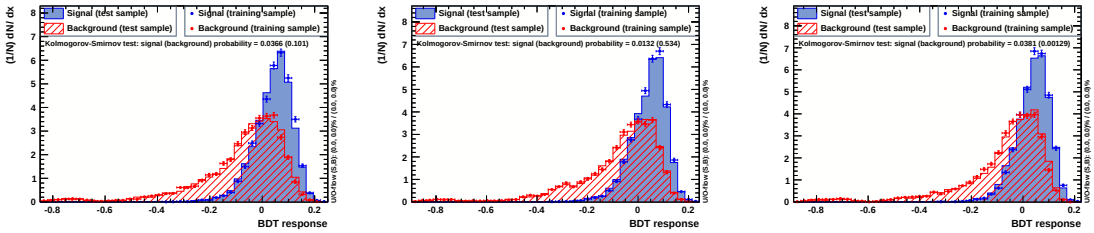
Figure B.3: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD filled with water.



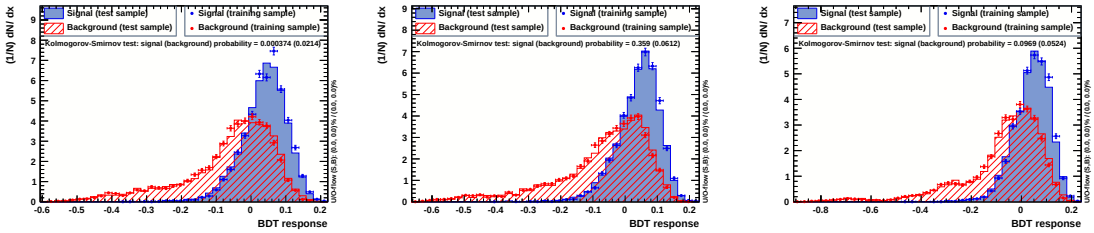
(a) Fiducial = -45 mm (b) Fiducial = -44 mm (c) Fiducial = -43 mm

(d) Fiducial = -42 mm

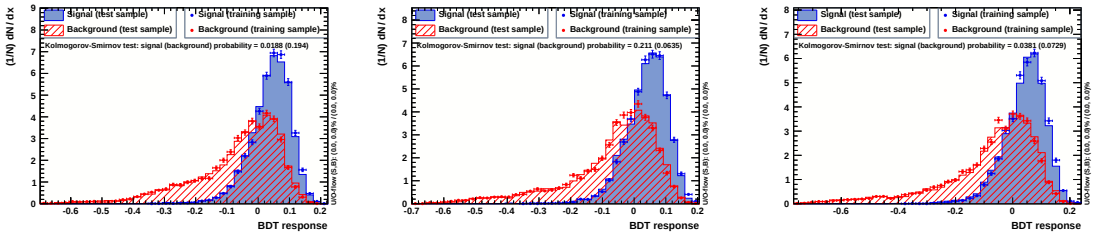
Figure B.4: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD filled with water.



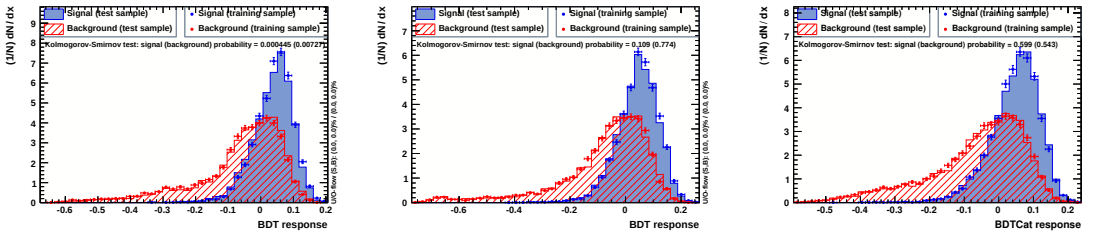
(a) Fiducial = -90 mm (b) Fiducial = -89 mm (c) Fiducial = -88 mm



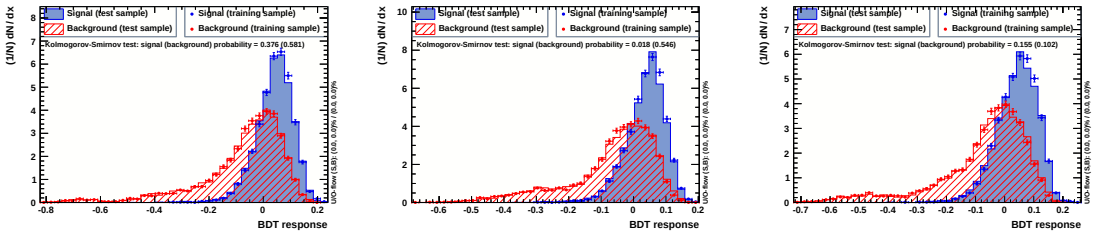
(d) Fiducial = -87 mm (e) Fiducial = -86 mm (f) Fiducial = -85 mm



(g) Fiducial = -84 mm (h) Fiducial = -83 mm (i) Fiducial = -82 mm

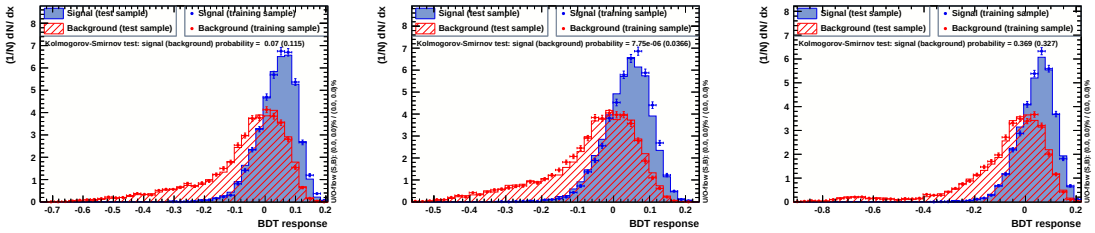


(j) Fiducial = -81 mm (k) Fiducial = -80 mm (l) Fiducial = -79 mm



(m) Fiducial = -78 mm (n) Fiducial = -77 mm (o) Fiducial = -76 mm

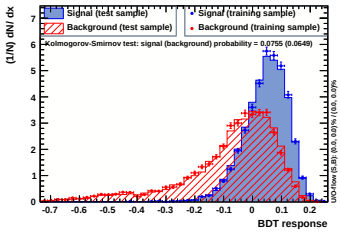
Figure B.5: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD without water.



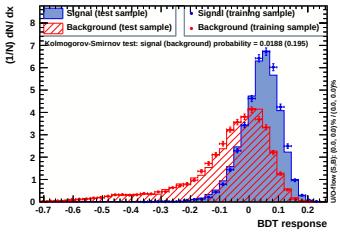
(a) Fiducial = -75 mm

(b) Fiducial = -74 mm

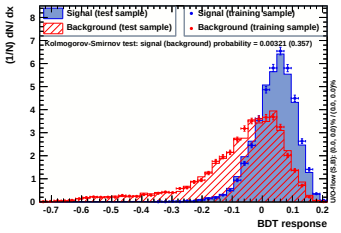
(c) Fiducial = -73 mm



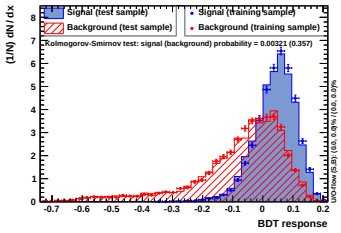
(d) Fiducial = -72 mm



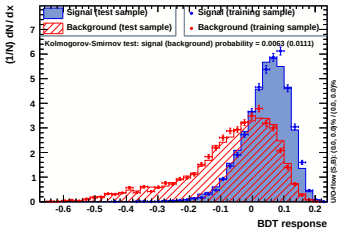
(e) Fiducial = -71 mm



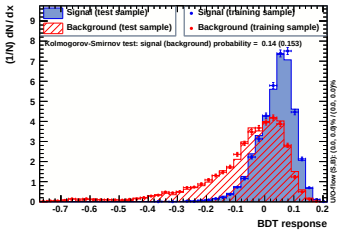
(f) Fiducial = -70 mm



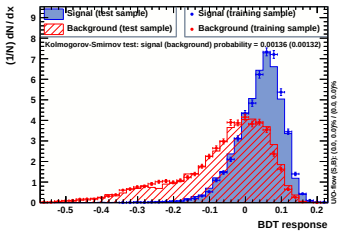
(g) Fiducial = -69 mm



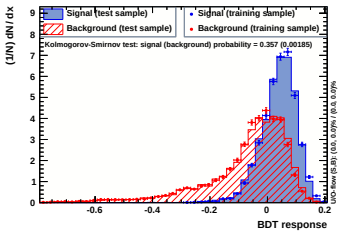
(h) Fiducial = -68 mm



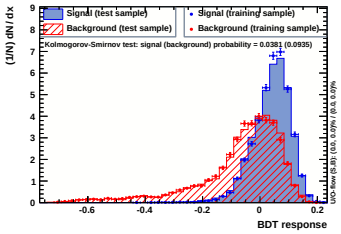
(i) Fiducial = -67 mm



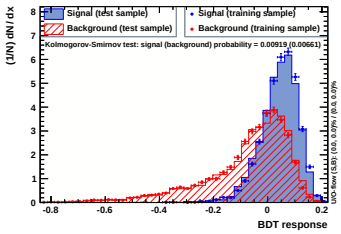
(j) Fiducial = -66 mm



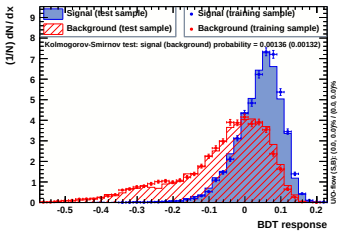
(k) Fiducial = -65 mm



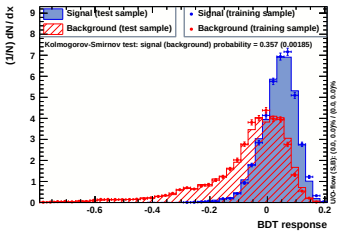
(l) Fiducial = -64 mm



(m) Fiducial = -63 mm

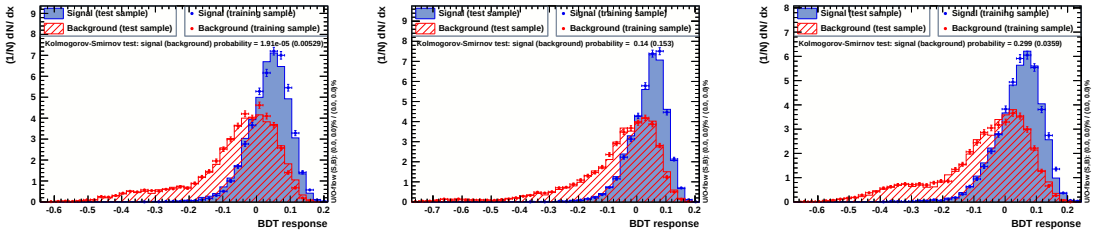


(n) Fiducial = -62 mm

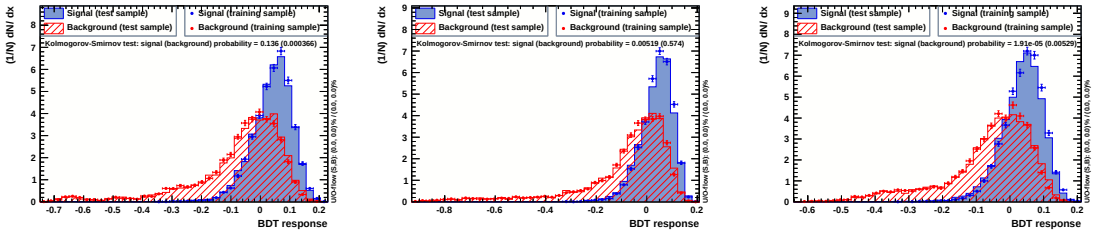


(o) Fiducial = -61 mm

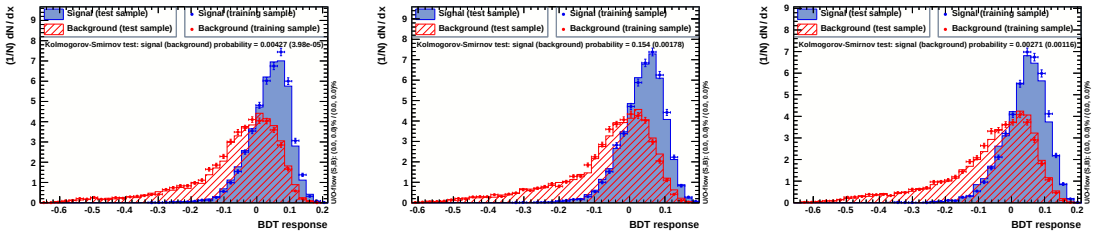
Figure B.6: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD without water.



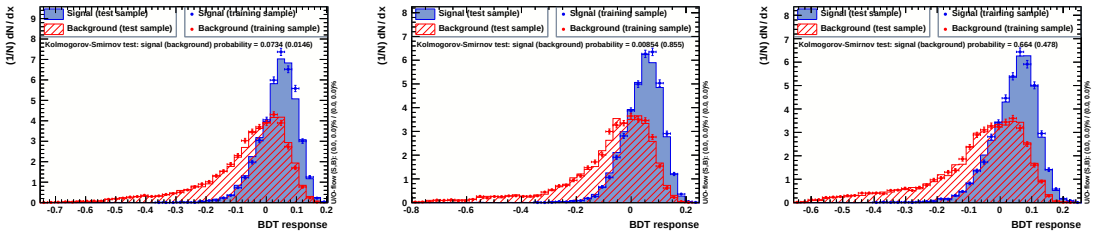
(a) Fiducial = -60 mm (b) Fiducial = -59 mm (c) Fiducial = -58 mm



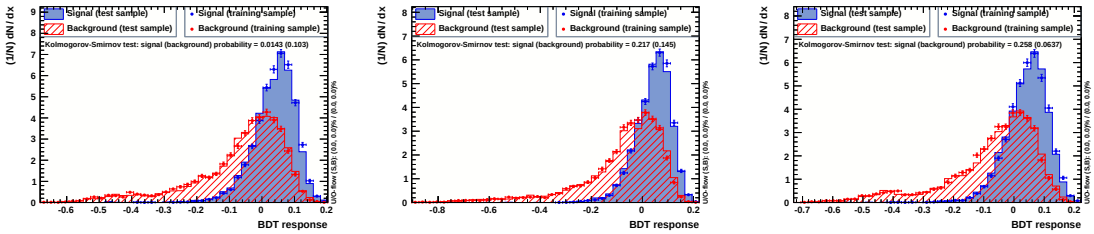
(d) Fiducial = -57 mm (e) Fiducial = -56 mm (f) Fiducial = -55 mm



(g) Fiducial = -54 mm (h) Fiducial = -53 mm (i) Fiducial = -52 mm

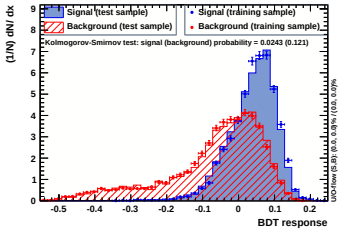


(j) Fiducial = -51 mm (k) Fiducial = -50 mm (l) Fiducial = -49 mm

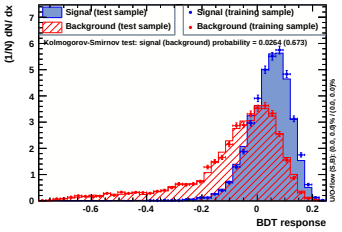


(m) Fiducial = -48 mm (n) Fiducial = -47 mm (o) Fiducial = -46 mm

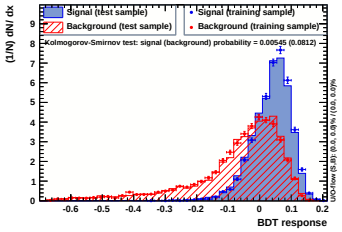
Figure B.7: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD without water.



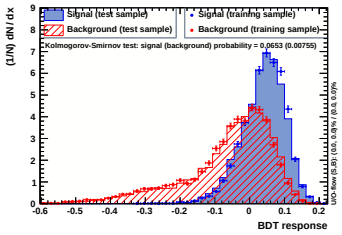
(a) Fiducial = -45 mm



(b) Fiducial = -44 mm



(c) Fiducial = -43 mm



(d) Fiducial = -42 mm

Figure B.8: MVA response for signal and background separation using BDTcat method for different values of fiducial volume for the PØD without water.