Imperial College London
Department of Physics

Measurement of the muon scattering cross-section on Carbon using the T2K Near Detector and characterisation of CERN's T9 beam for the Water Cherenkov Test Experiment

Alie Adeline Laure Craplet

Submitted to Imperial College London in partial fulfilment
of the requirements for the degree of Doctor of Philosophy in Physics

June 2025

Abstract

The field of long-baseline neutrino oscillation physics is entering a precision era. Upcoming experiments such as Hyper-Kamiokande (HK) and the Deep Underground Neutrino Experiment (DUNE), both dominated by systematic uncertainties, aim to begin data-taking in the coming years. A primary goal of these experiments is the discovery of CP violation in the lepton sector, provided it is sufficiently large. Reducing systematic uncertainties—particularly those related to neutrino interactions and detector response—is therefore critical.

The Water Cherenkov Test Experiment (WCTE) and the Tokai-to-Kamioka (T2K) experiment are actively contributing to this effort. This thesis describes the methods developed by WCTE to identify particles and measure their momenta before they enter a water Cherenkov detector, which provide essential truth-level inputs for validating detector response and simulation models. It also presents the first characterisation of CERN's T09 beamline in the sub-GeV momentum range, which was crucial in optimising WCTE's data-taking strategy.

This work also contains the first inclusive measurement of muon scattering on carbon in the 200 MeV/c to 4.5 GeV/c momentum range, using data from the T2K near detector ND280. The analysis is based on nearly one million negatively charged muons identified with over 98% purity. Comparisons with the Geant4 Bertini model reveal that it underestimates the total scattering cross section by (2.90 ± 0.03)%. Discrepancies are also found in differential distributions: the model under-predicts scattering in regions associated with correlated nucleon pairs by 10–20% and over-predicts single pion production by about 15%. Comparisons with an experimental version of the NEUT generator, adapted for muon scattering, also show poor agreement as expected. This work demonstrates the feasibility and value of further charged particle scattering measurements at T2K to constrain systematic uncertainties in future neutrino oscillation experiments.

Statement of Originality

I, Alie Craplet, declare that the work presented in this thesis is my own and that any figures or studies from external sources are appropriately referenced.

In Chapter 4, I developed the method and algorithm for particle identification on the data collected by me and other WCTE collaborators during the 2023 beam test. The peak finding algorithm mentioned in Section 4.3.1 was mainly developed by Nick Prouse, I implemented the rest of the pre-processing algorithm described in Section 4.3 except for the calibration of the digitiser timing offset described in the second half of Section 4.3.2.2 which was made by Arturo Fiorentini. I characterised the beam in terms of particle number and momentum, which is discussed in Sections 4.4 and 4.5. The characterisation of the beam structure presented in Section 4.6.4 was made by Dean Karlen. I used his results to infer the total particle rates in the beam presented in Section 4.6.

In Chapter 6, I derived and implemented in the existing T2K software the event selection described in Section 6.2. I developed the normalisation and unfolding techniques presented in Section 6.3 with valuable input from Mark Scott and ran the systematic uncertainty estimation presented in Section 6.4 using existing ND280 tools, which I adapted to run for the muon scattering analysis. The interpretations of the comparisons presented in Section 6.5 are my own, with insight from Mark Scott and Charlie Naseby. The NEUT version compared with the data in Section 6.6 has been adapted by Jake McKean and me, heavily based on the work by Seisho Abe. The interpretation of the results are my own.

Copyright notice

Acknowledgements

First and foremost, I would like to thank my supervisor, Mark Scott, for his unwavering help throughout these last four years. Thank you for sitting with me every week, providing me with great input and suggestions on seemingly never-ending sets of slides. Thank you for encouraging me to pursue the research topics that I really cared about and helping me get through times of doubt. I could not have wished for a better supervisor, and I am extremely proud to call myself your student.

I would also like to thank Akira Konaka, Mark Hartz and Dean Karlen for sharing their Physics insight with me and for their extremely valuable inputs on my WCTE work. The T09 beamline characterisation would not have been the same if not for their contributions, and it was an honour to work with you on the WCTE experiment. In particular, I would like to thank Akira for teaching me so much so patiently and for sharing with me his love and excitement for neutrino Physics.

I extend my gratitude to all my WCTE family for their hard work and constant dedication, which made such an ambitious project possible. To my beam test crew, thank you for these fun, long, intense, crazy, and oh so rewarding weeks of data taking. Bruno, I've told you many times, but as well as a friend, you are also a real inspiration, and I can't wait to work with you again. Deesha, thank you for keeping me sane throughout the long night and day shifts, you've been a ray of light. Josh, Nick, Diego, Arturo, Luan, Manoor and everyone else involved in the data taking and analysis, thank you for your hard work, it's been a pleasure working with you. To Lauren and Oli, thank you for making of this dream detector a shiny stainless steel reality. To Yassine, thank you for all the laughs and support, you've been the cool older brother I've always wanted.

To my T2K and HK family, Michael and Yashwanth, brothers in crime, thanks for welcoming me so perfectly into the collaboration, Sam and Katarina, for the ski slopes, Alex, for Helsinki, Andres, Carrie and Amy, for bringing some Joyful Honda into shift-week, Abe-san and Jake, for your invaluable NEUT expertise, Clarence and Minoo, for your precious insights and keen interest in my muon scattering adventure.

To my Blackett family, all my thanks for these unforgettable four years. Charlie, first of course, for all your help, so often. With you, I've learned about stats and fits, how one draws a beamline properly, that it can be cool to munch coffee beans and snack on sticks and of course, how a swan sings. Thanks for helping make this workplace my home. Charlotte, for showing me the secret tunnel under the mountain, and always being here to hug and listen when I needed someone, you've helped me through tough times, and I'm forever grateful. To David Colling, for giving me my first taste of research and believing in me all these years ago, I wouldn't be here if not for you. To the Neutrino group, for their continuous input and advice on my work. To all my friends in the department, Rohan and David D, my second parents, you guys are absolute gems, thanks for all the food, all the drawings, all the books and all the laughs, you're the coolest. To my buddies, Tiago and Irene, you two are the kindest and most beautiful souls I've ever met. To Kieren for always being there when I needed, knowing how to make me laugh, you too are an inspiration. To Jay and Naseem, for the emotional roller coasters, to Anežka for being such a great example to follow, to Koustubh for the parties, to Liban for his unbeatable motivational speeches, to Lucas for the heart-to-hearts, to Marie for standing by my side all these

years, you are a true sister, to Yordan, Parth and Jac, for the board games and pizzas, to Robert, Lily, Karol, and all the Acton House lot for the much needed breathers, to Paula B., Paula C. Louise and Sunita for their invaluable admin support throughout the years and the unbeatable HEP parties, to Simon, Ray, Dan, Daniela, Sam, Duncan, Mounir for the all the computing and technical help.

Finally, to my loved ones, Marvin, thank you, for everything. Throughout the last five years, you have been my anchor, and I couldn't have done any of it without you. You've taught me more about myself and the world than I thought I could ever learn, I look forward to many more years by your side. Clara, Noémie, Papa et Maman, vous êtes tous mes modèles, et je vous aime plus que tout. Merci de m'avoir encouragée à poursuivre mes passions, de m'avoir soutenue à chaque tournant et de me suivre toujours, vous êtes les meilleurs.

Contents

List of Tables

List of Figures

Chapter 1

Neutrinos

According to the Standard Model of Particle Physics (SM), neutrinos are the most abundant massive particles in the cosmos. They barely interact with matter and their properties are less well measured than those of most particles [1]. We know that neutrinos are massive, but we only have upper limits on the value of their mass and no widely accepted theory of mass production. Furthermore, recent measurements of neutrino oscillation parameters [2] have shown hints that neutrinos and anti-neutrinos could be propagating differently through space-time, which could contribute to explaining the matter-antimatter asymmetry observed in our Universe [3].

At the time of writing, the field is a few years away from the turn-on of next-generation neutrino oscillation experiments: Hyper-Kamiokande (HK) [4], the Deep Underground Neutrino Experiment (DUNE) [5] and the Jianmen Underground Neutrino Observatory (JUNO) [6], which will measure neutrino oscillation with an unprecedented accuracy, moving the field of oscillation physics away from being dominated by statistical uncertainties into being systematics-dominated. For these experiments to reach their physics goals, it is critical to reduce systematic uncertainties. This is achieved through improving our modelling of the incoming neutrino flux, our understanding of the detector response and our modelling of neutrino interactions with matter.

The work presented in this thesis contributes to two out of these three main challenges using the data collected by two current-generation experiments: the Water Cherenkov Test Experiment (WCTE) [7] and the Tokai to Kamioka (T2K) experiment [8]. After presenting the history, status and prospect of neutrino oscillation physics in the remaining part of this chapter, this thesis discusses the physics of lepton interactions in Chapter 2. Chapter 3 presents the WCTE detector and its physics program, working towards improving our understanding of the HK detectors and of neutrino scattering. Chapter 4 presents the first characterisation of the CERN T09 particle beam used by WCTE, whose entire physics program relies on having an accurate understanding of the incoming particles [7]. Chapter 5 moves on to describe the T2K long baseline neutrino experiment, and Chapter

6 presents the first measurement of muon scattering on hydrocarbon, made using the T2K near detector ND280, and discusses how it can contribute to improving our understanding of neutrino interactions in matter. Chapter 7 concludes this thesis.

1.1 A brief history of neutrino physics

The study of Particle Physics began at the turn of the twentieth century, with the discovery of radioactivity by Becquerel in 1896 [9]. The years that followed saw the first measurements of the energy spectrum of electrons (called beta particles at the time) produced in radioactive decays. Experiments by Meitner and Hanh [10], Danysz [11] and Chadwick [12] in the early 1910s showed that the electrons were produced with a continuous energy spectrum, contrary to the mono-energetic signature that is expected for a two-body $n \rightarrow p + e^-$ decay [13].

A few years later, in 1930, Pauli proposed a solution to this problem, namely that a third particle, electrically neutral, was produced in these radioactive decays but not detected. This particle would carry away part of the momentum [14], thus satisfying the energy conservation principle. He called this particle the neutron, later to be renamed neutrino.

Reines and Cowan's was the first experiment to directly verify the existence of the neutrino in 1956 [15, 16]. Their experiment aimed to measure the inverse beta decay $\bar{\nu}_e + p \rightarrow n + e^+$ process, using the nearby Savannah River nuclear reactor as a source of electron anti-neutrinos. The experiment was composed of two tanks of $CdCl_2$-doped water, sandwiched between layers of scintillation counters, made out of liquid scintillator and read out by photo-multiplier tubes (PMTs). Any anti-neutrino interacting with a proton in the water would produce a positron and a neutron, the former rapidly annihilating with an electron, producing a characteristic pair of 0.511 MeV photons and the latter being absorbed a few microseconds later by a Cadmium atom, emitting a photon. The prompt and delayed photons were observed in the scintillator and the delay between them reduced when the Cadmium concentration was doubled. The experiment took data with the nuclear reactor on and off, and thus proved the existence of the neutrino.

Following the discovery of the neutrino, Goldhaber *et. al.* went on to measure the neutrino helicity (i.e. relative orientation of their spin compared to their direction of travel) by studying the decay via electron capture of ^{152}Eu atoms to ^{152}Sm* which de-excite to the ground state via the emission of a photon whose direction and polarity is directly related to the neutrino helicity [17]. This experiment measured that the neutrino has a purely negative chirality (i.e. left-handed, spin is anti-parallel to the direction of propagation). This result directly confirmed Sudarshan, Markshak, Feynman and Gell-Mann's theory (now part of the SM) [18, 19] that the weak interaction has a vector-axial structure, in agreement with the 1957 measurement by Wu *et. al.* [20]

that parity is violated by the weak force which couples exclusively to left-handed particles and to right-handed anti-particles.

Since the discovery of the muon in 1937 [21], limits on its decay rate to an electron-photon pair ($\mu \to e\gamma$) have shown that the lepton flavour seemed to be a conserved quantity in physical reactions [22, 23]. This meant that neutrinos produced by the decay of pions to muons would not produce an electron when interacting with matter, but instead a muon. This hypothesis was tested by Lederman, Schwartz and Steinberger in 1962 with the first accelerator-based neutrino experiment at Brookhaven National Laboratory [24]. In this experiment, a beam of 15 GeV protons was used to produce a pion-enriched beam that decays to a muon-neutrino pair. The majority of muons were stopped by a thick iron shielding block placed before the detector. Any electron produced in the aluminium spark chamber detector located behind the shielding block would shower and rapidly get absorbed. On the contrary, muons would travel a large distance. The experiment measured a significant excess of muons compared electrons, disproving the one-neutrino-only hypothesis and concluding the existence of a second kind of neutrino, the muon neutrino. It was only much later, in 2001, that the third flavour of neutrino, namely the tau neutrino, was discovered by the DONUT experiment [25]. They were produced in the interaction of 800 GeV/c protons with a beam dump and detected using emulsion targets. High precision measurements of the Z boson width indicate that there are 2.9840 ± 0.0082 flavours of active neutrinos [26].

Ray Davis' Homestake experiment [27], operating from 1967 to 1992 aimed at measuring the rate of neutrinos coming from the Sun through $\nu_e + {}^{37}\mathrm{Cl} \to e^- + {}^{37}\mathrm{Ar}$ interactions. Interestingly, the measured (electron) neutrino spectrum was two to three times smaller than predicted by models of nuclear reactions in the Sun [28]. This so-called solar neutrino problem was further investigated by the Sudbury Solar Observatory (SNO). In its 2001 measurement of the ν_e and total neutrino flux [29], SNO confirmed both the results of Davis et. al. and the solar flux predictions. These seemingly disagreeing results could be well explained by Pontecorvo's theory [30, 31] which postulated that neutrinos change flavour as they propagate, similarly to mesons, i.e. neutrino oscillations. In 1998, the Super-Kamiokande experiment, followed by SNO in 2009 [32], directly observed neutrino oscillations in the comparison of the atmospheric neutrino flux composition as a function of the zenith angle [33], confirming Pontecorvo's theory.

1.2 Neutrino oscillation

The theory of neutrino oscillations describes the three neutrino flavour (also called weak) eigenstates ν_e, ν_μ and ν_τ as linear superpositions of three mass eigenstates ν_1, ν_2 and ν_3. The mass eigenstates ψ are stationary states of the free particle Hamiltonian $\hat{\mathrm{H}}$:

$$\hat{\mathrm{H}}\psi = i\frac{\partial \psi}{\partial t} = E\psi \tag{1.1}$$

where E is the energy of the particle. All particles travel in space as mass eigenstates, which propagate as plane waves with the following time dependence:

$$\psi_i\left(\mathbf{x},t\right) = \phi_i\left(\mathbf{x}\right)e^{-iE_it} = \phi_i(0)e^{i(\mathbf{p}_i\cdot\mathbf{x}-E_it)}, \tag{1.2}$$

where $\mathbf{p}_i$ and E_i are the three-momentum and energy of the mass eigenstate. After it has travelled a distance L in a time T, the mass eigenstate has picked up a phase $p_i \cdot x = p_iL - E_iT$.

The weak force interacts with flavour eigenstates (ν_e, ν_μ or ν_τ) which are a superposition of mass eigenstates (ν_1, ν_2 or ν_3) and described as follows:

$$\nu_\alpha = \sum_i^3 U_{\alpha i}\nu_i \tag{1.3}$$

where $U_{\alpha i}$ is a 3 complex and unitary matrix which is often expressed using the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) parametrisation [30, 31, 34]. In this parametrisation, $U_{\alpha i}$ is a combination of three real parameters (θ_{12}, θ_{13} and θ_{23}) and a single phase parameter (δ_{CP}).[1] The PMNS matrix is often presented as follows:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.4}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$.

If the three mass eigenstates ν_i where $i = 1, 2, 3$ have different masses m_i, they will propagate with a different phase. This means that a neutrino initially produced with a flavour α and energy E has a probability $P(\nu_\alpha \to \nu_\beta)$ to be detected as a β-flavour neutrino after travelling a distance L in vacuum. This probability is given by:

$$P(\nu_\alpha \to \nu_\beta) = \sum_{i,j} U_{\alpha i}U^*_{\beta i}U^*_{\alpha j}U_{\beta j}\exp\left[-i\frac{\Delta m^2_{ji}}{2}\frac{L}{E}\right] \tag{1.5}$$

where $\Delta m^2_{ji} = m^2_j - m^2_i$, is commonly referred to as mass squared splitting.

When neutrinos travel through matter instead of a vacuum, they also experience matter effects, which cause

---

[1] For Majorana neutrinos there are two additional “Majorana phases” which do not affect neutrino oscillations but impact the conservation of lepton flavour, e.g. permitting neutrino-less double beta decay. See [35] for more details.

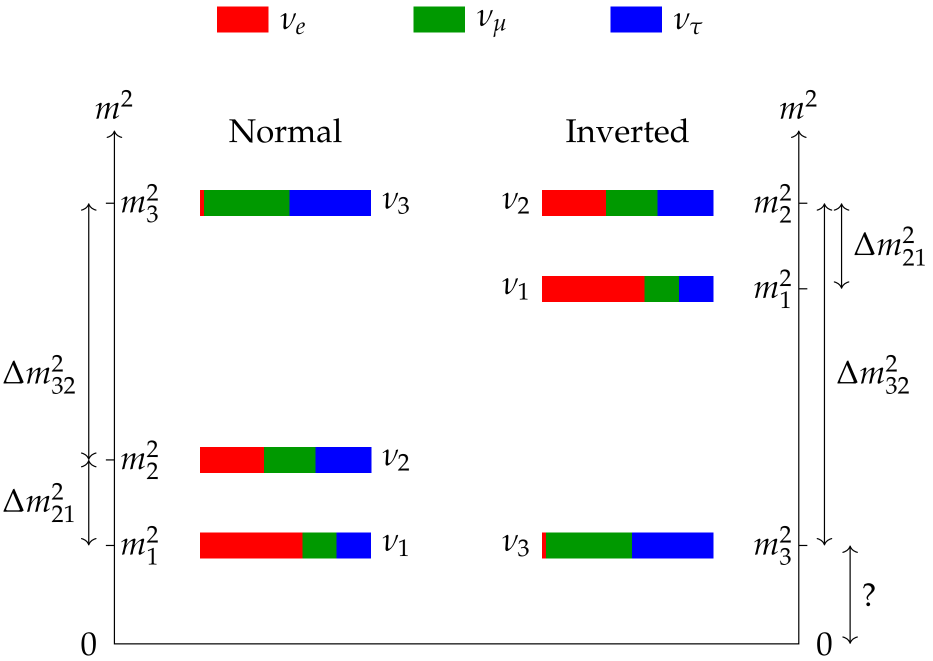


Figure 1.1: Presentation of the approximate composition of the three neutrino mass eigenstates in terms of the flavour eigenstates for the two possible mass orderings. Figure reproduced from [39].

the electrons neutrinos to behave differently from muon and tau neutrinos [36, 37]. These account for the fact that electron neutrinos can undergo an additional weak interaction with electrons in matter, causing them to experience a potential modified compared to the potential experienced by the other flavours by:

$$\Delta V = \pm 2\sqrt{2} G_F N_e \tag{1.6}$$

where G_F is the Fermi constant and N_e is the electron number density and the positive and negative signs affect neutrinos and anti-neutrinos, respectively [38]. The matter effects strongly dependent on the ordering of the neutrino masses, which is currently unknown: either the normal ordering (NO): $m_1 < m_2 < m_3$ or the inverted mass ordering (IO): $m_3 < m_1 < m_2$. Figure 1.1 presents the ordering of the neutrino masses and approximate compositions of each of the neutrino flavours.

Because energies and distances travelled by neutrinos coming from different sources are different, they can be used to probe different parameters of the PMNS matrix. Solar neutrino experiments are primarily sensitive to θ_{12} and Δm^2_{21} parameters, atmospheric neutrino experiments mainly probe θ_{23}, Δm^2_{32}, and δ_{CP} and reactor neutrino experiments are sensitive to δ_{CP} parameters and reactor experiments to θ_{13} and Δm^2_{31}. Depending on the channel(s) studied (disappearance $\overset{(-)}{\nu}_\mu \rightarrow \overset{(-)}{\nu}_\mu$ or appearance $\overset{(-)}{\nu}_\mu \rightarrow \overset{(-)}{\nu}_e$), accelerator based long baseline experiments are sensitive to Δm^2_{32}, θ_{23} and δ_{CP} and to mass ordering [40]. Table 1.1 summarises the current best fit value (bfv) and uncertainty on the parameters of the PMNS matrix obtained from global fits to solar, atmospheric, reactor and accelerator neutrino data available as of September 2024 [41].

Parameter	**Normal Ordering (bfv $\pm 1\sigma$)**	**Inverted Ordering (bfv $\pm 1\sigma$)**
$\sin^2\theta_{12}$	$0.308^{+0.012}_{-0.011}$	0.308 ± 0.011
θ_{12} $(^\circ)$	$33.68^{+0.73}_{-0.70}$	$33.68^{+0.73}_{-0.70}$
$\sin^2\theta_{23}$	$0.470^{+0.017}_{-0.013}$	$0.550^{+0.012}_{-0.015}$
θ_{23} $(^\circ)$	$43.3^{+1.0}_{-0.8}$	$47.9^{+0.9}_{-0.9}$
$\sin^2\theta_{13}$	$0.02215^{+0.00056}_{-0.00058}$	0.02231 ± 0.00056
θ_{13} $(^\circ)$	8.56 ± 0.11	8.59 ± 0.11
δ_{CP} $(^\circ)$	212^{+26}_{-41}	274^{+22}_{-25}
Δm^2_{21} $(10^{-5}\,\mathrm{eV}^2)$	$7.49^{+0.19}_{-0.19}$	$7.49^{+0.19}_{-0.19}$
$\Delta m^2_{3\ell}$ $(10^{-3}\,\mathrm{eV}^2)$	$+2.513^{+0.021}_{-0.019}$	-2.484 ± 0.020

Table 1.1: Current best fit values and uncertainty on the neutrino oscillation parameters from NuFIT 6.0 with SK atmospheric data, assuming NO (left) or IO (right). Note that $l = 1$ for NO and $l = 2$ for IO. When the SK atmospheric data is included a weak statistical preference for normal mass ordering is observed ($\Delta\chi^2 = \chi^2_{\mathrm{IO}} - \chi^2_{\mathrm{NO}} \approx 6.1$), corresponding to about 2.5σ. When the SK atmospheric data is not included no preference for either mass ordering is observed. Values reproduced from [41].

1.3 Future prospects

From Table 1.1 we can see that our current picture of neutrino oscillations is only partially complete. Indeed, we do not yet know the neutrino mass ordering, experiments only slightly favouring NO [41]. Knowing the mass ordering is necessary to constrain models of neutrino masses [42–46] and is one of the main objectives of JUNO a multi-purpose next generation neutrino experiment currently under construction in China [6]. The JUNO experiment benefits from a high flux of neutrinos originating from multiple reactors, a very good detector energy resolution ($3\%/\sqrt{E}$) and an accompanying satellite experiment, the Taishan Antineutrino Observatory (TAO), which will directly measure the reactor antineutrino spectrum [47]. The JUNO experiment is expected to determine the neutrino mass ordering at the $\sim 3\sigma$ level within about 7 years of data taking [48] as well as measuring $\sin^2\theta_{12}$ and Δm^2_{21} to sub-percent precision [6, 48].

We can see from Table 1.1 that existing limits on the the CP-violating phase δ_{CP}, made by the T2K and NuMI Off-axis ν_e Appearance (NoνA) experiments, have a very large uncertainty. This parameter is currently the least well-known parameter of the PMNS matrix. Depending on the true value of δ_{CP}, the two next-generation long baseline neutrino experiments HK and DUNE will be able to exclude the CP-conserving hypothesis in the neutrino sector ($\delta_{\mathrm{CP}} = 0^\circ$ or 180°) within a few years of data taking [49–51]. HK and DUNE will directly measure the value of δ_{CP} and be able to test the validity of various Beyond the Standard Model theories [4, 6, 52].

Both of these experiments are also expected to measure $\sin^2\theta_{23}$, which governs how much ν_μ and ν_τ couple to ν_3, and identify in which octant θ_{23} is, i.e. if $\theta_{23} > 45^\circ$ or $\theta_{23} < 45^\circ$. The maximally mixing value of $\theta_{23} = 45^\circ$ would indicate a fundamental $\mu - \tau$ symmetry and constrain models of neutrino mass generation and quark-lepton unification. Both experiments will also be sensitive to neutrino mass ordering and be able to measure neutrinos from supernovae [4, 52]. Finally, JUNO, HK and DUNE all search for physics beyond the standard model, including searches for proton decay, non-standard neutrino interactions and sterile neutrinos as well as testing the unitarity of the PMNS matrix [4, 6, 53]. The current systematic uncertainties have to be reduced to cope with the increased statistics of next generation experiments and allow them to reach their physics goals.

Chapter 2

Lepton scattering

One of the main sources of systematic uncertainties in neutrino oscillation measurements is linked to the modelling of neutrino interactions [54]. Neutrino oscillation measurements rely on comparing the neutrino interaction rates as a function of the neutrino flavour and energy before and after they have oscillated. T2K's most recent oscillation analysis (OA) reported uncertainties of about 3% on the interaction model, which was the leading or sub-leading source of uncertainty in all samples considered [2], as presented in Table 2.1. These uncertainties affect both the reconstruction of the neutrino's energy and the absolute and relative ($\overset{(-)}{\nu}_e$ to $\overset{(-)}{\nu}_\mu$) cross-section rates.

The cross-section systematics are constrained by fitting the data at the near detector. As an additional step, the T2K collaboration also employs simulated data studies, which test whether the uncertainty model is flexible enough to accommodate variations in the underlying neutrino cross-section model. In these studies, simulated datasets are generated using alternative models and parameter tunes for neutrino-nucleus interactions, which can significantly increase or decrease the predicted number of events in both near and far detectors. Their effects is added in quadrature to previous estimation of the systematic uncertainties. While the overall impact of these simulated data studies on the estimation of $\sin^2\theta_{23}$ and δ_{CP} was found to be small compared to existing systematic and statistical uncertainties [2], they provide valuable insight into the impact of cross-section modelling in oscillation parameter measurements for next-generation neutrino experiments.

In long baseline neutrino experiments, the near detector(s) provide dedicated measurement of the unoscillated neutrino flux and interaction cross-section, which reduces the uncertainty on the measurement of the oscillation parameters. However, some cross-section parameters cannot be measured in the near detectors. For HK these are expected to represent a systematic uncertainty between 0.5% and 1.5% (depending on the channel studied), which is significant compared to the total systematic budget of about 3% [51].

The neutrino cross-section uncertainties impact the oscillation measurements in multiple ways. The muon

Sample	**Flux (%)**	**Interaction (%)**	**FD + SI + PN (%)**	**Flux⊗Interaction (%)**	**Total (%)**
1Rμ ν	2.9	3.1	2.1	2.2	3.0
1Rμ $\bar{\nu}$	2.8	3.0	1.9	3.4	4.0
1Re ν	2.8	3.2	3.1	3.6	4.7
1Re $\bar{\nu}$	2.9	3.1	3.9	4.3	5.9
1Re1De ν	2.8	4.2	13.4	5.0	14.3

Table 2.1: Uncertainties on the number of events in each far detector sample after the fit to the near detector data in the 2023 T2K neutrino oscillation analysis [2]. FD + SI + PN represent the joint error on the far detector response, the modelling of secondary interactions and photo-nuclear effects. The samples correspond to one muon-like ring (1Rμ), one electron like ring (1Re) and one electron-like ring with a Michel electron (1Re1De).

scattering measurement presented in Chapter 6 can contribute to reducing the uncertainty on the ratio of electron neutrino to muon neutrino cross-section ratio, which is a critical source of uncertainty on oscillation measurements as it mimics an oscillation signal (i.e. if electron neutrinos have a cross-section higher than the models the resulting neutrino oscillation probability measured by the experiment will be overestimated) [54, 55]. Additionally, this measurement is a first step towards measuring the radiation of a real photon by a muon which, if collinear with the muon's direction, can lead to misidentification of a muon neutrino as an electron neutrino at the far detector[56]. Furthermore, the accuracy of the neutrino energy reconstruction strongly relies on our models of the initial nuclear state, of final state interactions and secondary interactions taking place within and outside of the struck nucleon [57] which can also be studied using muon scattering data. A more comprehensive discussion of the impact of cross-section uncertainties in T2K can be found in [2].

This chapter presents the different neutrino scattering processes in Section 2.1 followed by a discussion of how charged lepton scattering can contribute to reducing the uncertainty on neutrino-nucleus cross-section models in Section 2.2.1. Section 2.2.2 presents a brief description of the NEUT event generator and the Geant4 simulation tools used in the muon scattering analysis of Chapter 6.

2.1 Neutrino scattering

Neutrinos are detected through their interactions with nuclei, which produce secondary particles, some of them measured by our detectors. Take a simple case of a neutrino with initial four momentum $k = (E, \mathbf{k})$ scattering off a nucleus A at rest, producing an outgoing lepton with four momentum $k' = (E', \mathbf{k}')$. The energy transferred in the scatter is $q_0 = E - E'$ and the momentum transfer is $\mathbf{q}_3 = \mathbf{k} - \mathbf{k}'$. The squared four-momentum transfer Q^2, is given by:

$$Q^2 = -q_0^2 + |\mathbf{q}_3|^2. \tag{2.1}$$

The four-momentum transfer is related to the de Broglie wavelength which determines the scale of the system that the neutrino interacts with. When $Q^2 \approx 1$ GeV^2, which is typically the size of a nucleon, the incoming particle interacts with a single nucleon inside the nucleus, ejecting it from the nucleus alongside a charged lepton or a neutrino of the same flavour, in so-called charge current (CC) and neutral current (NC) quasi-elastic (QE) interactions respectively. These interactions are the dominant interaction modes at T2K and HK as shown on Figure 2.1.

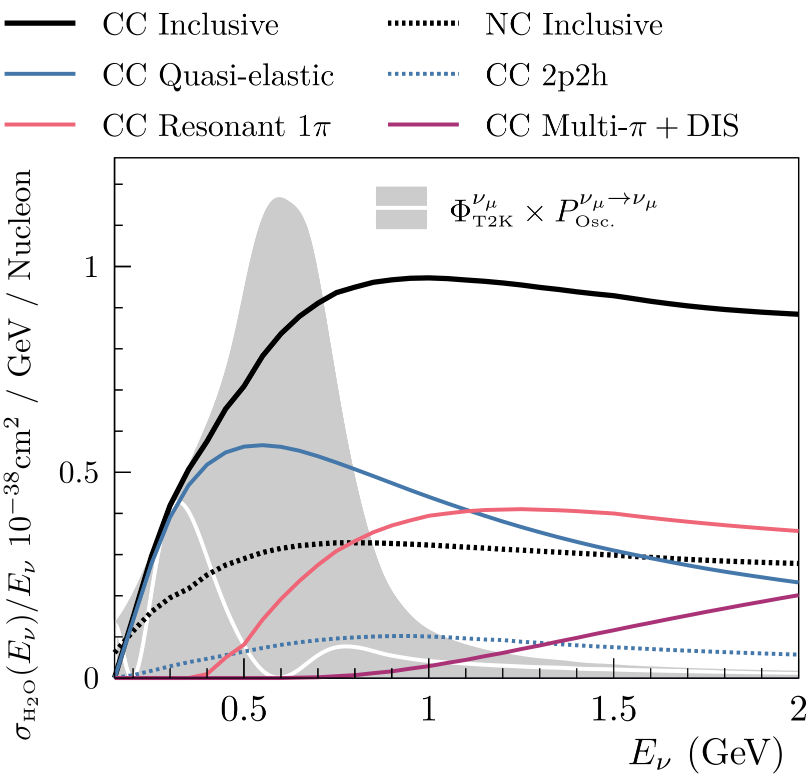
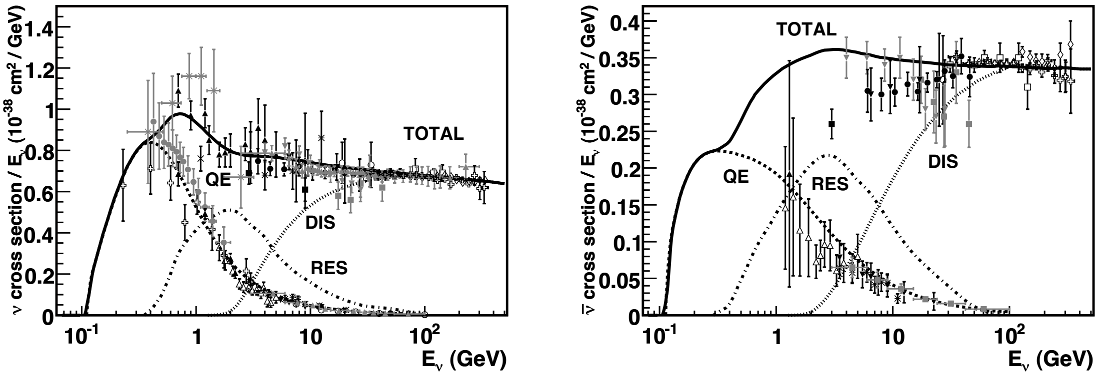


Figure 2.1: NEUT prediction of the cross-section for each of the scattering modes presented in the text overlayed onto the T2K muon neutrino flux. The HK neutrino flux is very similar to T2K's [4]. Figure reproduced from [58].

In this $Q^2 \approx 1$ GeV^2 region, it can also happen that the nucleon that the neutrino interacts with is correlated with another nucleon inside the nucleus. In this case, both nucleons are ejected from the nucleus. These are referred to as 2p-2h (and more generally np-nh) interactions, and are described in Section 2.1.2. At higher energy, the neutrino transfers enough energy to produce one (see Section 2.1.3) or more (see section 2.1.4) hadrons, typically pions. Finally, at very high Q^2 the neutrino is energetic enough to interact with constituent quarks within a nucleon, causing it to break apart and undergo hadronisation. This regime, briefly presented in Section 2.1.4, is called deep inelastic scattering (DIS) and is a subdominant process in T2K/HK. At values of MeV$^2 < Q^2 <$ GeV^2, the neutrino can excite the nucleus in collective vibrations, a fundamental excitation mode called giant resonances, challenging to model where both NC and CC interactions are possible. At even lower Q^2, the neutrino interacts coherently with the whole nucleus. This neutral current interaction has a relatively large cross-section and is mostly forward peaked [59, 60] but is typically not observable in long baseline neutrino experiments. Figure 2.2 presents results from inclusive and semi-inclusive measurements of the neutrino and anti-neutrino cross-section, which have contributed to building our models of neutrino interactions.



(a) Neutrino

(b) Anti-neutrino

Figure 2.2: Inclusive and semi-inclusive cross-section per nucleon divided by energy. The data points represent results from experiments up to 2011 and solid lines correspond to theoretical predictions. Reproduced from [61].

2.1.1 Charged Current Quasi-Elastic interactions

Charged current and neutral current quasi-elastic interactions shown in Figure 2.3 dominate in the T2K and Hyper-K energy regime. To accurately predict the cross-section rate and kinematics of the outgoing particles, one needs to model exactly the initial nuclear state, the interaction vertex, and correctly model the behaviour of produced leptons and hadrons as they escape the nucleus. Instead of trying to treat all of these aspects simultaneously, many frameworks, including event generators like NEUT [58] and GENIE [62] used in neutrino oscillation measurements, use the impulse approximation (IA) and spectral functions (SF) to simplify these calculations following the scheme presented in [63, 64].

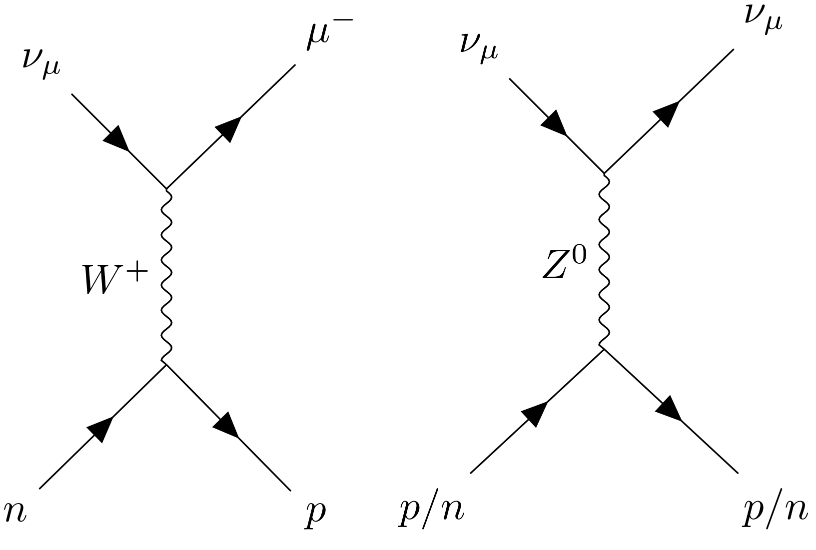


Figure 2.3: Feynman diagram of CCQE (left) and NCQE (right) muon neutrino scattering.

In the IA, the lepton is taken as scattering off a free nucleon, exchanging a single boson (following the Born approximation [65, 66]). The remaining $A-1$ nucleons in the nucleus are only spectators of the interaction. Most event generators use the plane wave impulse approximation (PWIA) which ignores the distortion of the nucleon wave function caused by the nuclear potential, treating it as a plane wave instead [67]. The SF encodes all the information on the nuclear initial state and structure, which is assumed to be independent of Q^2. Figure 2.4 shows the two-dimensional probability density function for the oxygen SF used by NEUT, which represents the probability of finding an initial state nucleon with a given momentum and removal energy E_{rmv} which is the the energy required to remove the nucleon from the nuclear potential. The removal energy is directly related to the nuclear binding energy E_{B} by [68, 69]:

$$E_{\mathrm{B}} = E_{\mathrm{rmv}} + E_N - m_N \tag{2.2}$$

where E_N is the initial nucleon energy and m_N its mass.

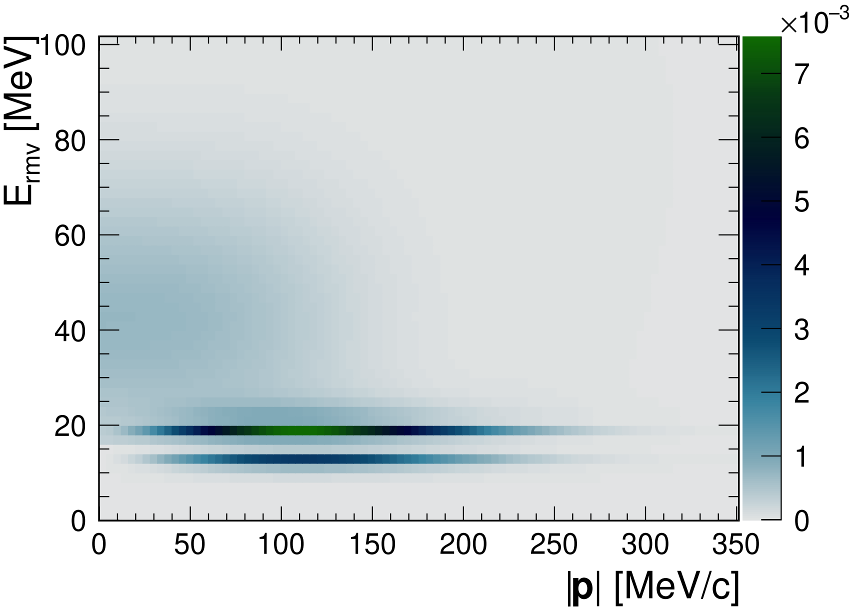


Figure 2.4: Probability for a nucleon in an oxygen nucleus to have a given momentum and removal energy as given by the spectral function [63] in NEUT. The shell structure is clearly visible. Figure reproduced from [2].

The inclusive double differential cross-section for the scattering of a neutrino off a nuclear target A at rest is given by:

$$\left(\frac{d\sigma}{dE'd\Omega'}\right)_{\nu_l/\bar{\nu}_l} = \frac{G^2}{4\pi^2} k' E' L_{\mu\nu} R^{\mu\nu} \tag{2.3}$$

Here, $k = (E, \mathbf{k})$ and $k' = (E', \mathbf{k}')$ denote the four-momenta of the incoming and outgoing leptons, and Ω' is the solid angle of the outgoing lepton. The coupling constant is $G = G_F$ for NC processes and $G = G_F \cos\theta_C$ for CC interactions, where $\cos\theta_C = 0.97425$ [70] and $G_F = 1.1803 \times 10^{-5}$ GeV^{-2} is the Fermi constant.

The leptonic tensor $L_{\mu\nu}$ is entirely determined by lepton kinematics. The hadronic tensor $R^{\mu\nu}$, however, encodes the nuclear response and depends on the transition between initial and final nuclear states via the nuclear current operator $J^{\mu}(\mathbf{q}_3, q_0)$. This operator generally includes both one-body (single nucleon) and two-body (meson exchange or multi-nucleon) contributions.

If we restrict ourselves to the case of scattering off a free nucleon — i.e., neglecting nuclear effects — then the hadronic tensor can be described using the Llewellyn-Smith formalism [71], which expresses the CCQE cross section in terms of four nucleon form factors: two vector form factors $F_1(Q^2)$ and $F_2(Q^2)$, the axial-vector form factor $F_A(Q^2)$, and the pseudo-scalar form factor $F_P(Q^2)$. The vector form factors are constrained by electron scattering data [72], while the axial-vector form factor is typically modeled with a dipole form:

$$F_A(Q^2) = \frac{F_A(0)}{\left(1 + Q^2/M_A^2\right)^2} \tag{2.4}$$

Here, $F_A(0)$ is the axial coupling constant, determined from beta decay experiments [73], and M_A is the axial mass, tuned using neutrino–deuterium data [74–76]. The pseudo-scalar form factor $F_P(Q^2)$ is related to $F_A(Q^2)$ through the PCAC hypothesis [77].

In the more realistic case of a nuclear target like ^{12}C, the situation is more complex: nucleons are bound in the nuclear medium, and one must account for nuclear effects such as Fermi motion, binding energy, Pauli blocking, multi-nucleon (2p2h) correlations, and final-state interactions (FSI). These effects modify both the shape and normalization of the cross section, and are incorporated in modern event generators like NEUT. Nonetheless, the form factor structure from the single-nucleon picture still forms the foundation upon which these nuclear effects are added.

In the case of CCQE interactions, the neutrino energy E_ν is directly measured from the kinematics of the outgoing charged lepton $k' = (E', \mathbf{k}')$. For a lepton emitted at an angle θ_l with respect to the neutrino, the reconstructed neutrino energy is:

$$E_\nu^{\text{rec}} = \frac{2m_N E_B + 2m_N E' - m_l^2}{2(m_n - E' + |\mathbf{k}'|\cos\theta_l)} \tag{2.5}$$

where m_l and m_N are the lepton and nucleon masses respectively and E_B is the binding energy defined in Equation 2.2. In NEUT, E_B is obtained from the SF removal energy distribution, which reflects the shell structure of the nucleus [2].

2.1.2 Two-Particles-Two-Holes interactions

Nucleons can be correlated with one another in the nucleus, either via short range correlations (SRC) or meson exchange currents (MEC), where a meson, typically a pion is exchanged between the two nucleons, as shown in Figure 2.5. When the neutrino interacts with a nucleon in such correlated pair, both nucleons are ejected from the nucleus as shown on Figure 2.6. Alternatively, nucleons can be correlated via long-range correlation which also impact the total cross-section but are much weaker than SRC and MEC contributions. These interactions do not typically produce two nucleons in the final state.

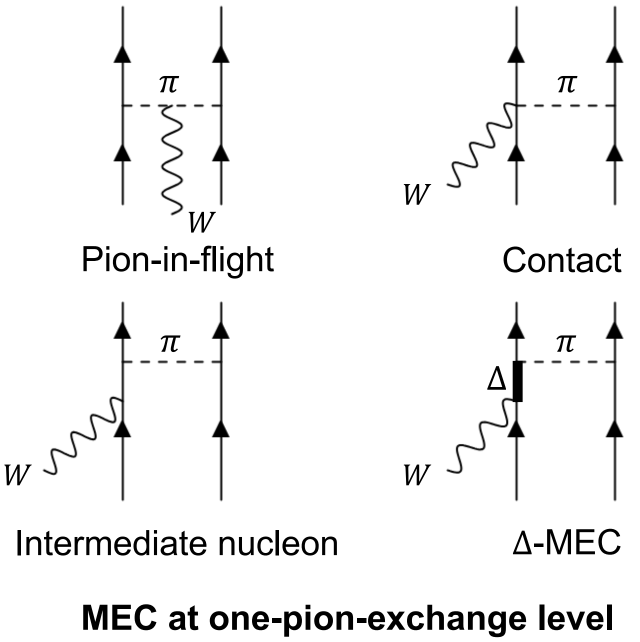


Figure 2.5: Possible meson exchange currents at the one-pion exchange level. Figure based on [78], reproduced from [79].

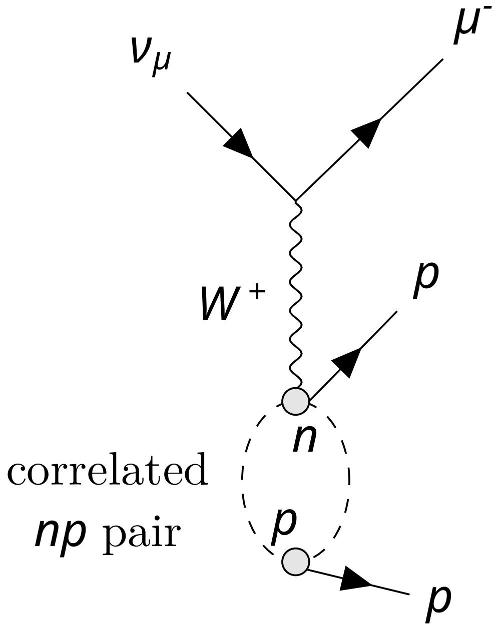


Figure 2.6: Muon neutrino interacting with a pair of correlated nucleons.

The IA presented in Section 2.1.1 does not appropriately deal with correlated nucleon pairs. The importance of multi-nucleon correlations was demonstrated by MiniBooNE's measurement of CCQE neutrino interactions on carbon [80]. This measurement presented significant disagreement with theoretical models, leading to an estimation of the axial mass significantly different from the world's average. This so-called MiniBooNE puzzle was resolved by the inclusion of interactions from correlated nucleons in the model [81]. Since then, many different theoretical models of 2p2h interactions have been developed, the most commonly used are by Martini *et. al.* [82], Nieves *et. al.* (the Valencia model, used in NEUT) [81, 83] and the SuSav2 model [84, 85]). Their predictions vary significantly, indicative that this process is not well understood.

Nucleon-nucleon correlations are a purely nuclear phenomenon and are also observed in electron scattering data [78]. Studies of electron scattering on carbon [86] have been able to measure the fraction of *pn*, *pp* and *nn* correlated pairs by detecting both nucleons in the final state. Ref [86] shows that about 20% of the ^{12}C nucleons were involved in SRC. This study measured that out of the SRC correlated pairs, 90 ± 10% of them were *pn* pairs, 5 ± 1.5% were *pp* pairs, and 5 ± 1.5% were *nn* pairs.

Unless pions are detected in the final state, the T2K neutrino energy estimator assumes that the interaction is CCQE, using equation 2.5 to estimate the neutrino energy from the kinematics of the outgoing charged lepton. In the case of a 2p2h interaction, this introduces a bias in the reconstructed neutrino energy, directly impacting the oscillation analysis. The relative contribution of 2p2h interactions and the energy bias they cause to the reconstructed neutrino energy needs to be correctly modelled in order to correctly measure the neutrino oscillation parameters [54]. Section 2.2.1 will discuss how charged lepton scattering measurements can be used to improve this modelling.

2.1.3 Single pion production

At higher Q^2, the neutrino can transfer enough energy to excite the nucleon to a baryon resonance which then decays back to the ground state, producing additional particles, typically pions as shown on Figure 2.7. This is referred to as resonant pion production (RES). The prediction of RES single pion production (SPP) in NEUT and most neutrino event generators follows the Rein-Sehgal model [87] with improvements to the nucleon axial form factors [88, 89] and inclusion of the non-zero lepton mass [90, 91]. Because of phase-space limitations, the different lepton mass causes a sizeable difference in the electron and muon neutrino cross-sections, especially at low Q^2. The various models of RES SPP predict different cross-section ratios $\sigma_{\nu_e}/\sigma_{\nu_\mu}$, which is a systematic uncertainty in oscillation analysis [54]. Charged lepton scattering data can also test SPP models, as described in Section 2.2.1.

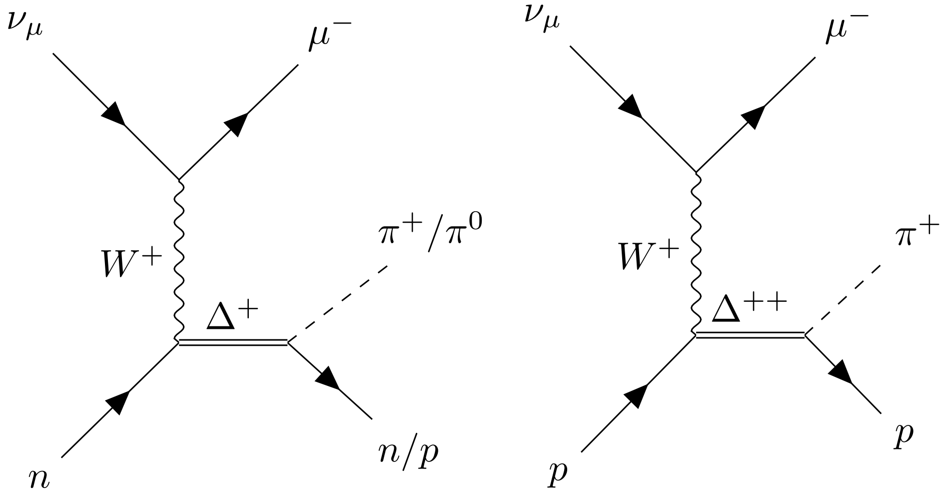


Figure 2.7: Charged current resonant pion production via an intermediate Δ resonance.

The RES SPP model used in NEUT includes 18 baryonic resonances with masses up to 2 GeV and the interference between them. The $\Delta(1232)$ resonance is the dominant contributor to the resonant pion production cross-section. Accounting for the initial nucleon kinematics, the minimum amount of transferred energy required to excite a Δ is $q_0 \approx 200$ MeV. The modelling of resonant pion production is a very active area of research with models like the DCC[92, 93] HNV [94] and MK[95] making their way into event generators.

At Q^2 lower than about 0.05 GeV^2, the neutrino can also produce pions by coherently scattering off the whole nucleus without breaking it apart. In NEUT, this is modelled using the Berger-Seghal model [96]. The experimental signatures of resonant and coherent pion productions are extremely similar, making it in principle a background in Cherenkov detectors [54]. At T2K and HK energies, the coherent pion production has a very small cross-section and is actually negligible.

2.1.4 Multi-pion production and deep inelastic scattering

At four momentum transfer larger than about 1(GeV/c)^2, the neutrino can excite higher mass resonances which can decay to multiple pion states. At still higher energy, the neutrino resolves and interacts with single constituent quarks within the nucleon, which breaks apart and hadronises. Both of these processes produce multiple hadrons in the final state. The region between the single pion production and DIS is sometimes referred to as shallow inelastic scattering (SIS) and is very challenging to model. NEUT, like most other event generators, uses a custom interpolation between the SPP and DIS regions [97] based notably on the Bodek-Yang modifications which extend the DIS models to (relatively) low values of $Q^2 \leq 1.5$(GeV/c)^2 [98, 99]. The hadronic side of DIS is typically modelled using the PYTHIA generator [100], built for modelling much higher energy interactions (e.g. in proton-proton collision experiments). In T2K oscillation analyses, the multi-π and DIS are subdominant processes due to the relatively low neutrino energy and the selection at the far detector.

2.1.5 Nuclear effects

2.1.5.1 Initial nuclear state

As mentioned throughout Section 2.1, nuclei with more than a few nucleons are highly complex systems and the quality of the neutrino cross-section modelling strongly depends on the accuracy and precision with which we understand the nuclear medium. Because nucleons are fermions, they obey the Pauli exclusion principle, which states that two fermions cannot occupy the same momentum-spin state in the nucleus.

The available nuclear quantum states are filled from the ground state up. The momentum of the highest occupied state is the Fermi momentum p_F, meaning that no uncorrelated nucleon can have a momentum larger than p_F. Nucleons have a random Fermi motion in the nucleus. In carbon, this momentum is about 100 MeV/c. The initial Fermi motion of nucleons typically smears the distribution of energy transferred in neutrino-nucleus and electron-nucleus scatterings, due to the nucleons having themselves some initial energy.

Because all the available bound energy states in the carbon nucleus are filled, the struck nucleon needs to have enough energy to escape nucleus; this effect is referred to as Pauli blocking and suppresses CCQE interactions at low Q^2. NEUT uses $p_F = 209$ MeV/c for carbon, which approximately corresponds to a binding energy of 23 MeV/c [58]. Interestingly, models which do not use the plane-wave approximation calculate that the binding energy is linearly dependent on the transferred three-momentum [101, 102]. This is in line with results obtained using electron scattering data modelled with the NEUT event generator [103]. Refs [103] and [101] show that applying momentum-dependent corrections to E_B obtained from electron scattering data improves the accuracy of the neutrino energy reconstruction.

Correlated nucleons, discussed in Section 2.1.2, can individually have $p_{corr} > p_F$ because their momenta are approximately equal and opposite, resulting in a small combined momentum in the centre-of-mass frame. Correlated nucleons represent the high-momentum tail of the spectral function.

2.1.5.2 Final state interactions

Hadrons produced by the interaction of a neutrino inside the nucleus can undergo final-state interactions (FSI) inside the nuclear medium. These FSIs typically alter the kinematics, multiplicity and/or nature of the outgoing hadrons [104]. This can occur through elastic and inelastic scattering, absorption, charge exchange and/or new hadron production, as shown on Figure 2.8. FSIs can charge the multiplicity and type of particles that exit the nucleus, causing different primary interactions to have the same experimental signatures. They represent an important source of uncertainty in neutrino oscillation analyses.

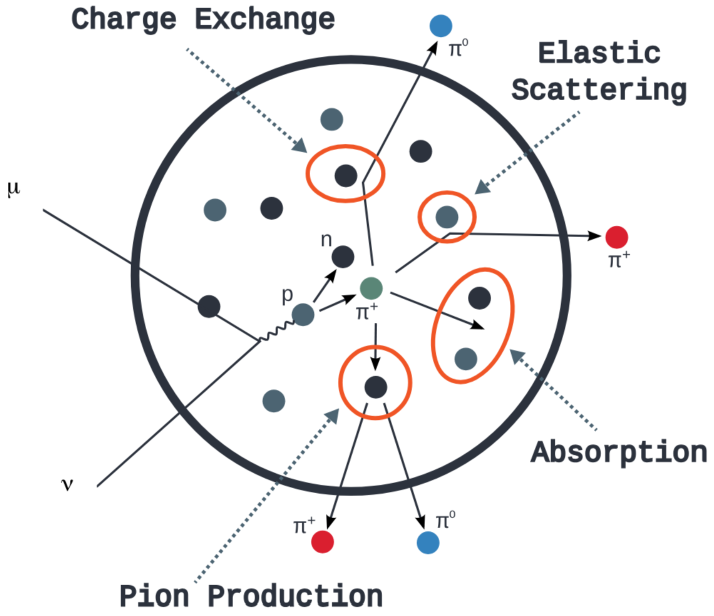


Figure 2.8: Diagram of the final state interactions that a pion can undergo before leaving the nucleus. Figure reproduced from [54].

Event generators typically model final-state interactions by stepping the outgoing hadron(s) through the nucleus, calculating at each step a probability for it to re-interact. This procedure is often referred to as internuclear cascade. NEUT uses the pion and nucleon internuclear cascades presented in Ref [58] and Refs [105, 106], respectively. The probabilities for each interaction to occur are based on external pion scattering data [107]. Any secondary hadron produced in the cascade is also propagated through the nucleus until it escapes.

After they have left the nucleus, outgoing particles can interact with other atoms in the detector, in so-called secondary interactions (SI). These affect the reconstruction of the event topology and neutrino energy. In particular, the scattering of pions and muons in Cherenkov detectors can lead in the case of hard scatter to the identification of a second ring from the scattered particle. Small-angle scatters, typically in the case of pions, can blur the Cherenkov ring, making it look like an electron, as will be discussed in Section 5.4. In T2K's 2020 OA [2], the total uncertainty associated with the far detector, secondary interactions and photonuclear effects is between 1.9% and 13.4%, depending on the sample considered, as shown in Table 2.1. The WCTE experiment is working towards reducing the uncertainty from these three sources, as discussed further in Chapter 3.

Unlike hadrons, outgoing leptons do not interact via the strong force. They are, however, affected by the nuclear Coulomb field. To first order, this field effectively shifts the energy of the (outgoing) charged lepton, increasing (decreasing) the energy of the incoming lepton (anti-lepton). The size of this shift is determined from electron and positron scattering data [75]. NEUT uses a value of 3.1 MeV for the optical potential of the carbon nucleus [103]. This Coulomb correction is applied to both the incoming and outgoing lepton in the case of $e^{\pm} - A$ and $\mu^{\pm} - A$ interactions.

2.1.6 Radiative corrections

In the electromagnetic field of the nucleus, charged leptons can absorb and emit both virtual and real photons — processes collectively known as radiative corrections [108]. These effects alter the kinematics of the outgoing lepton and must be carefully incorporated into theoretical models of neutrino-nucleus interactions, especially when aiming for high-precision measurements. In particular, the probability for radiating a soft (i.e., low-energy) photon depends strongly on the lepton mass, leading to notable differences between muons and electrons. Theoretical calculations estimate that the probability of soft photon emission differs by approximately 10% between muons and electrons [109, 110]. This mass-dependent difference directly affects the ν_μ/ν_e cross-section ratio, which is a critical quantity for long-baseline oscillation experiments aiming to measure CP violation.

To account for this, a 2% systematic uncertainty is applied on this ratio in the official T2K oscillation analysis [111]. This uncertainty propagates into the predicted event rates at the far detector, particularly in appearance channels sensitive to ν_e interactions. Further studies comparing different implementations of charged-current quasi-elastic (CCQE) models — including those with and without radiative effects — have shown that the uncertainty on the ν_e/ν_μ and $\bar{\nu}e/\bar{\nu}\mu$ cross-section ratios remains at a similar level (about 2%) across the T2K and Hyper-Kamiokande energy range [55]. These corrections are therefore essential for ensuring that any observed difference in appearance rates is not misinterpreted as a genuine oscillation signal.

2.2 Charged lepton scattering

2.2.1 Motivation

We have seen throughout Section 2.1 how electron-nucleus scattering data was used for building our understanding of the initial nuclear state, an essential block in our models of neutrino interactions used in the measurement of neutrino oscillations.

Aside from providing external information about the nuclear state, charged lepton scattering data can also be used to directly test the neutrino event generators (e.g. NEUT, GENIE), because of the similarities between the two interactions. The inclusive double differential cross-section for the scattering of a charged lepton off a nucleon, exchanging a virtual photon, is:

$$\left(\frac{d\sigma}{dE'd\Omega'}\right)_e = \frac{\alpha^2}{Q^4}\frac{E'}{E}L_{\mu\nu}R^{\mu\nu} \tag{2.6}$$

where $\alpha \simeq 1/137$ is the fine structure constant. We can see that the charged lepton scattering cross-section is very similar to the neutrino one presented in Equation 2.3 with the notable addition of a $1/Q^4$ factor in the electron cross-section originating from the fact that the photon, unlike the weak bosons, is massless. Figure 2.9 shows the similarities between the electron and neutrino CCQE-like cross-sections. Just as in the neutrino case, the charged lepton form factor $L_{\mu\nu}$ is fully determined by the kinematics and mass of the lepton (m_l), in this case:

$$L_{\mu\nu} = 2\left(k_\mu k'_\nu + k_\mu k_\nu - \left(k \cdot k' - m_l^2\right)\right). \tag{2.7}$$

The hadronic tensor $R^{\mu\nu}$ is a function of the electric and magnetic vector form factors already mentioned in Section 2.1.1.

Recently, the neutrino interaction community has been looking at adapting neutrino event generators to simulate electron scattering events (NEUT [103], GENIE [62], GiBUU [112], NuWro[113]). These use the same nuclear models and cascade treatment of the FSI used in neutrino scattering, with some simple modifications to the hadronic form factors and coupling constants. The predictions made by these generators can then be directly tested against high precision and high statistics electron scattering data [103, 114], pointing out the weaknesses of the different models used. Muons are typically harder to produce and measure than electrons, which explains the low availability of muon scattering datasets, especially at low energy [115]. Having a good understanding of the scattering of low-momentum muons could however help reduce the uncertainty on the ratio of ν_μ to ν_e cross-section as it would directly validate our estimates for the radiative corrections in muons. It is also essential for achieving muon cooling which is a necessary step towards a future muon collider [116, 117].

In the scope of neutrino physics, a high-statistic, high-precision sample of muon scatters could be used to directly test lepton scattering models, similar to what is achieved with electron scattering data, reaching new regions of parameter space. Inclusive $\mu + A \to \mu + A$ and exclusive $\mu + A \to \mu + A + \gamma$ cross-section measurements would give precious validation of the muon radiation correction calculations and help quantify the $\nu_\mu + A \to \mu + \gamma + X$ contamination in the ν_e sample in the far detector.

Other exclusive measurements, in particular those with nucleons and/or pions in the final state, would yield very useful information about the relative contribution of the different interaction modes and FSI presented in Section 2.1.

2.2.2 Modelling

Chapter 6 presents a selection of muon scattering events on carbon using the T2K near detector, and goes on to compare the data with predictions made by the Geant4 version 9.4 Monte-Carlo (MC) simulation tool [118] and a version of NEUT 5.9.0 adapted to generate muon scattering events. The basic algorithm followed by NEUT to generate CC neutrino and charged lepton scattering events is [58]:

1. Select an initial lepton energy E_i based on templates of the incoming lepton flux and pre-calculated cross-section tables, accounting for the Coulomb correction in case of a charged lepton.

2. Randomly select a nucleon of initial momentum and removal energy according to the SF.

3. Calculate the centre of mass energy E_{COM} for the lepton-nucleon pair, and check that it is sufficient to produce a charged lepton of the corresponding flavour. If not go back to step 2.

4. Generate a random outgoing direction (in the COM frame) for the outgoing lepton and nucleon, respecting energy and momentum conservation.

5. Apply Pauli blocking, i.e. verify that the nucleon momentum in the lab frame is greater than p_F, if not, go back to step 2.

6. Calculate the transferred four-momentum and calculate an event weight based on the total cross-section formulae (obtained from integrating over energy and solid angle the double differential cross-sections presented in equations 2.3 and 2.6, for neutrino and charged lepton scattering, respectively).
7. Verify that the event weight is smaller than the maximum cross-section at E_i if not, go back to step 2.
8. Verify if the nucleon selected is in the correlated tail of the spectral function; if yes, produce a second nucleon with equal and opposite momentum
9. Propagate the outgoing hadrons through the nucleus using the intranuclear cascade model.
10. Apply Coulomb correction to the outgoing lepton energy.

The muon scattering analysis presented in Chapter 6 uses a version of the NEUT 5.9.0 originally built for electron scattering [103]. The only modification made was exchanging m_e for m_μ. Figure 2.10 presents the total quasi-elastic electron and muon cross-sections on carbon as calculated by NEUT. The cross-sections are identical at high energy with a slightly different turning-on point at low energy due to the mass difference. Only the quasi-elastic interaction mode is currently available for muon scattering; the inclusion of elastic scattering, nucleon correlation, radiative corrections, single pion production and DIS, necessary for a more complete description of the data, is not yet available. These effects, except for the nucleon correlations, are included in the Geant4 model of low momentum muon scattering [118, 119].

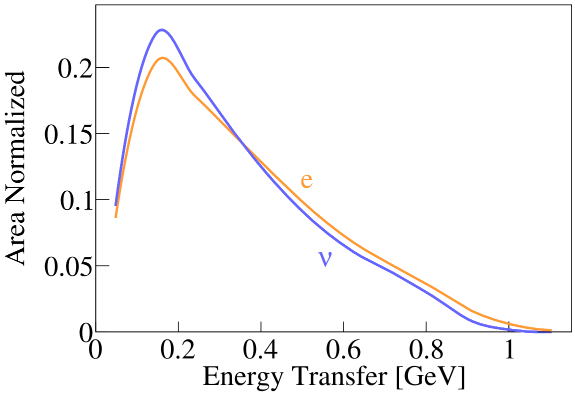


Figure 2.9: Comparison of CCQE-like semi-exclusive lepton-carbon scattering cross-section as modelled by GENIE. Both curves are area normalised, and the number of electron events has been weighted by Q^4. Figure reproduced from [120].

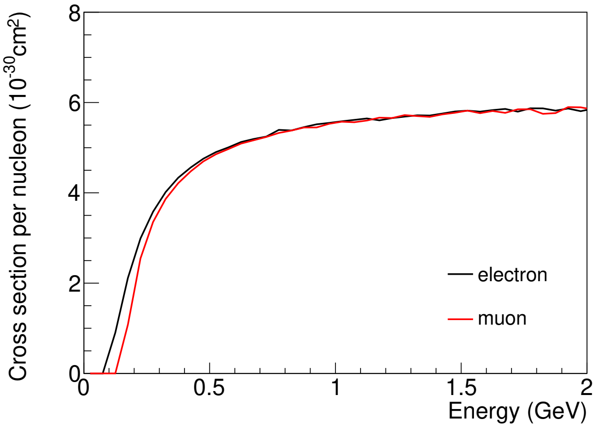


Figure 2.10: Total quasi-elastic scattering cross-section for electrons and muons as calculated by the NEUT event generator.

After particles have been generated by NEUT, they are propagated through the detector using the Geant4 Monte-Carlo simulation tool. The version used in this analysis uses the QGSP_BERT physics list which models the nuclear interactions using the Bertini cascade model [118]. It uses the standard GEANT4 electromagnetic physics as built by the G4EmStandardPhysics. The scattering of electrons (above 100 MeV) and muons follows the Wentzel IV model [121]. This model has been validated and muon scattering data with relatively good agreement, especially at low scattering angles [122].

Chapter 3

The Water Cherenkov Test Experiment

The Water Cherenkov Test Experiment (WCTE) was installed in the CERN T09 (East Area) beamline in the Autumn of 2024. It comprises a small-scale cylindrical water Cherenkov detector and a suite of beam monitoring apparatus. It receives a beam of protons, muons, electrons, pions and photons of energy between 200 and 1200 MeV. This Chapter presents the WCTE experiment, starting with a discussion of the motivations for this experiment in Section 3.1. It goes on to describe the CERN T09 beamline used by WCTE in Section 3.2 and the detector itself in Section 3.3.

3.1 Motivation for the WCTE

3.1.1 Hardware motivation

The WCTE experimentis a technology demonstrator for small-scale water Cherenkov detectors and in particular Hyper-Kamiokande's Intermediate Water Cherenkov Detector (IWCD) [4]. All of the multi-PMTs (mPMTs) and some of the designs (e.g. of the central deployment system (CDS), calibration sources) and methods (e.g. particle identification, event reconstruction, etc.) will be directly re-used in IWCD and HK. The WCTE is also testing four mPMTs built by the Hyper-Kamiokande collaboration for its far detector.

Furthermore, the WCTE collaboration is developing novel techniques for sub-GeV particle identification (PID) and tagged photon production, presented in Chapter 4 and Section 3.4 respectively. The WCTE has demonstrated how these novel tools can be used to build a small-scale multi-purpose sub-GeV beam monitoring facility which can easily be adapted to other beamlines to provide good PID and high-rate production of quasi-mono-energetic photons.

3.1.2 Physics motivation

Alongside its role as a technological test-bed, the WCTE also has a broad physics program which will contribute to reducing the uncertainty of long-baseline neutrino experiments. Thanks to its unique combination of water Cherenkov technology, high incoming particle rate and comprehensive beam monitoring apparatus, WCTE can study the response of Cherenkov water detectors to control samples of particles of known type and momentum [7]. In particular, WCTE will measure the rates of hadronic ($p/n/\pi$) secondary scattering in water, which are essential benchmarking tools for the modelling of secondary interactions [123]. Additionally, data from WCTE can be used to improve the energy reconstruction of particles in water Cherenkov detectors since the incoming momentum of each particle is measured by the beam apparatus before it reaches the tank. It is also improving the calibration of water Cherenkov detectors using new techniques and upgraded hardware. Both of these will help reduce the uncertainty on the energy scale, which is currently a 4% systematic error in Super-Kamiokande [124]. The WCTE data will be used to validate the existing `FitQun` algorithm [125] and novel ML-based `WatchMaL` framework [126]. These reconstruction and particle identification tools are to be used by the IWCD. The WCTE will be competitive with existing measurements of electron [127], pion [128] and muon [115] scattering on Oxygen. Measurements of charged lepton scattering can be used to validate the vector part of the lepton-nucleus interaction model [114], as introduced in Section 2.2.1. WCTE's electron and muon scattering sample will also be used to validate the radiative correction calculation [108].

Thanks to its tagged photon production capabilities, the WCTE is also an ideal environment to study the electron/photon separation capability of water Cherenkov detectors. This is essential to improve the sensitivity to ν_e appearance and reduce the risk of misidentifying neutral current events (e.g. $\mathrm{NC}\pi^0$) as charged current ν_e events. This background represents the second largest systematic error on Super-Kamiokande's single-ring atmospheric ν_e sample [129] and is the dominant background in ν_e detection for CP violation measurement [130].

Following the first period of data taking, where the detector is filled with ultra-pure water, the WCTE will operate with Gadolinium(Gd)-doped water, allowing it to tag neutrons using the 2.2MeV photon that is emitted when a Gd atom absorbs a free neutron. In particular, the $\pi^- \mathrm{p} \rightarrow \pi^0 \mathrm{n}$ interaction provides a control sample of tagged π^0 which can be used to validate Super-Kamiokande's π^0 background rejection method [131]. Furthermore, the Gd-doping allows a study of photonuclear reactions and the production of single photons through the $\Delta \rightarrow \mathrm{n}\gamma$ decay.

All of the Physics studies planned for WCTE rely on the accurate estimation of the incoming particle type and energy and on having an excellent understanding of the detector.

3.2 The T09 beamline at CERN

The WCTE has been approved for a total of 127 days of beam time, between Autumn 2024 (42 days) and Spring 2025 (85 days) in the CERN's East Area T09 beamline. This beamline has undergone a complete renovation during the second long shutdown including the upgrade of its beamline magnets, a renovation of the experimental hall and the addition of new targets [132]. A diagram of the East Area beamlines is shown in Figure 3.1.

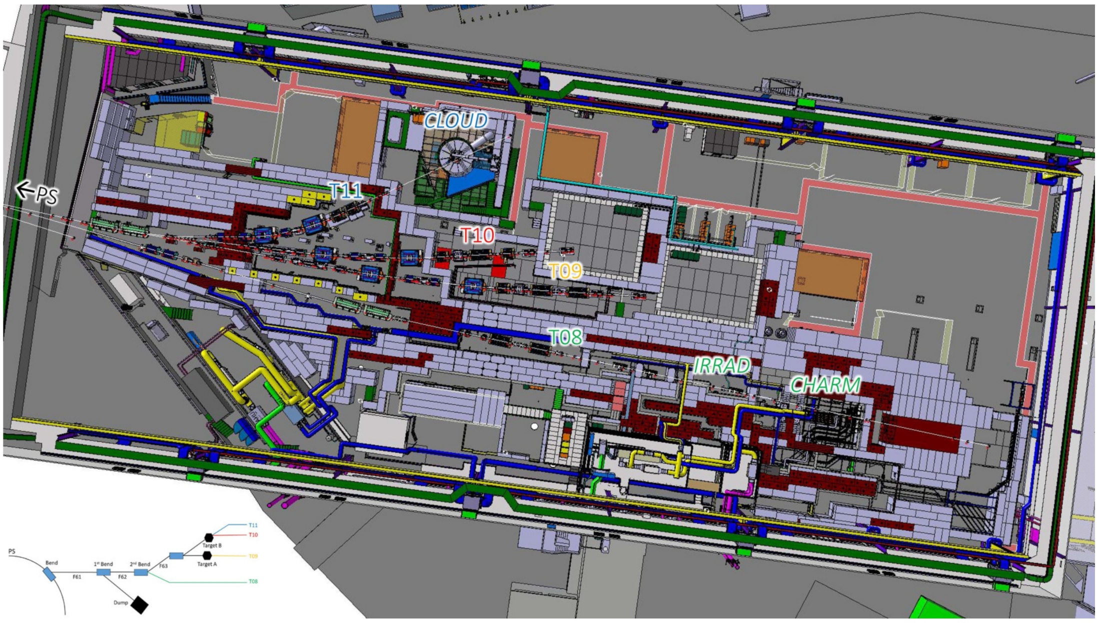


Figure 3.1: Layout of the CERN East Area Experimental Hall where the T09 beamline is located. The primary protons come directly from the PS. Reproduced with permission from [132].

The primary 24 GeV protons are diverted from the Proton Synchrotron (PS) and collide with a target chosen by the T09 user out of five available ones, producing a large number of hadrons, typically protons and pions and leptons, electrons and muons (mostly through the decay of pions). Figure 3.2 presents the T09 secondary beamline. The WCTE uses a 200mm Aluminium (Al) target when running in charged particle mode to benefit from the hadron production rate (and therefore muon production rate). In the tagged gamma configuration, WCTE uses a 200 mm Beryllium target coupled to a 3 mm Tungsten radiator (Be+W target) sheet, which favours electron production. Users can set the power of the bending magnets to change the momentum of the beam and adjust the width of the collimator slit opening to modify the momentum spread of the beam (from 3% down to 1%) and correspondingly increase or decrease the particle rate. Following the renovation, the range of momenta available in the T09 beamline is 0.2 to 15 GeV/c which fits very well with WCTE's need for sub-GeV particles to match the T2K and Hyper-Kamiokande energy range.

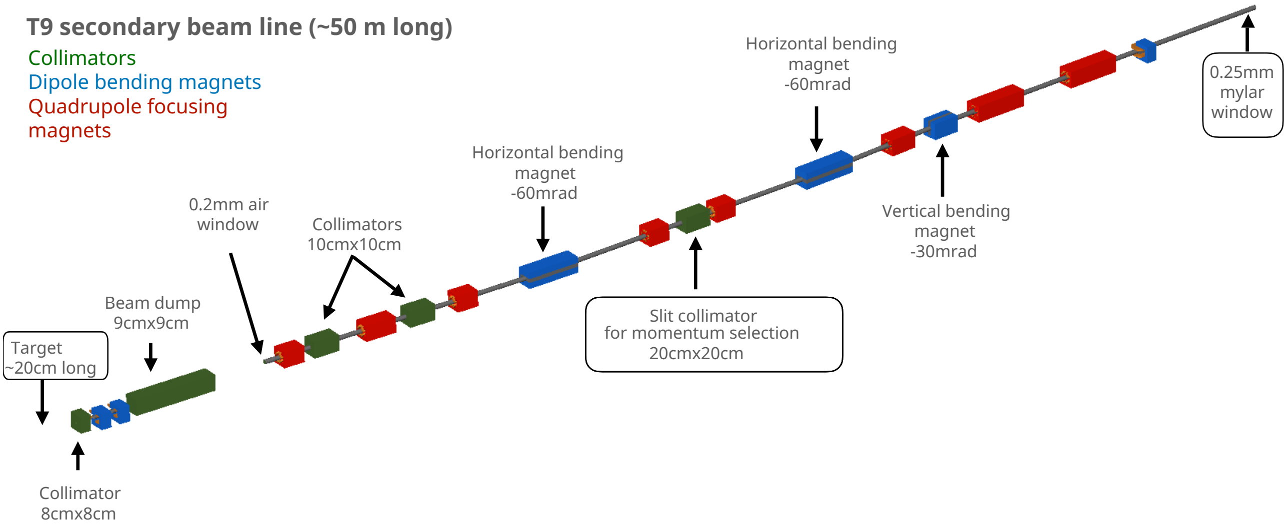


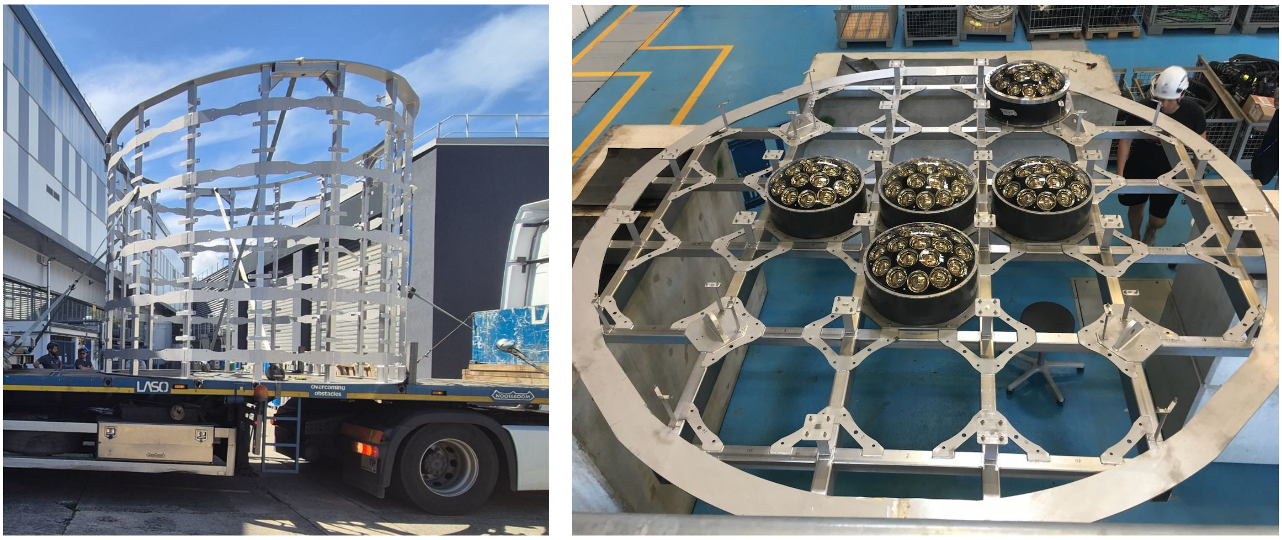
Figure 3.2: Diagram of the T09 secondary beamline. Figure made by A. Fiorentini.

3.3 The Water Cherenkov detector

3.3.1 Mechanical design

The WCTE is contained within a cylindrical stainless steel outer tank of 3.76m (diameter) by 3.52m (height), which acts as a light shield and a container for the ultra-pure water (total volume: $30 \pm 3\text{m}^3$). Its inner structure is an assembly of three mPMT support grids: a cylindrical barrel and two flat endcaps. Pictures of the tank and support structure pieces before and during assembly are presented in Figure 3.3.

A central deployment system (CDS) is fitted to the top end cap of WCTE as shown on Figures 3.4 and 3.5. This system enables the user to position a load with a precision of a few centimetres within the tank. This load can be a calibration source, a sampling device or a light source, depending on the specific needs. Additionally, a white lamp is fitted on each of the end caps, providing light to the eight photogrammetry cameras used to monitor the detector when the mPMTs are off. The last structural element of the water Cherenkov detector is the beam window, which is fitted onto the tank, at beam level and is crossed by particles when entering the detector. The beam window is visible on the outer tank in Figure 3.3c.



(a) Picture of the barrel arriving at CERN

(b) Picture of the bottom end cap with five mPMTs already installed

(c) Outer tank and its lid on the floor of the East Area Experimental Hall

(d) The assembled inner structure (right) filled with mPMTs and fitted with the CDS arm next to the tank (left)

Figure 3.3: Pictures of the WCTE inner structure and outer tank at different stages of the assembly process. Picture credit: WCTE collaboration.

3.3.2 mPMTs

In smaller water Cherenkov detectors like IWCD and WCTE, Cherenkov rings produced by particles have less space to expand compared to larger detectors like Super-Kamiokande and Hyper-Kamiokande. To obtain the same amount of information about these smaller rings, one needs sensors with a higher granularity, i.e. a larger amount of smaller detectors. The mPMTs, designed for IWCD and Hyper-Kamiokande [4, 133], inspired by the KM3NeT design [134, 135] and used in WCTE are examples of such high-granularity photodetectors. In the IWCD and Hyper-Kamiokande designs, the mPMTs are made up of 19 3-inch R14373 Hamamatsu Photonics

Figure 3.4: Picture of the CDS arm attached to the support structure's top endcap. Picture credit: WCTE collaboration.

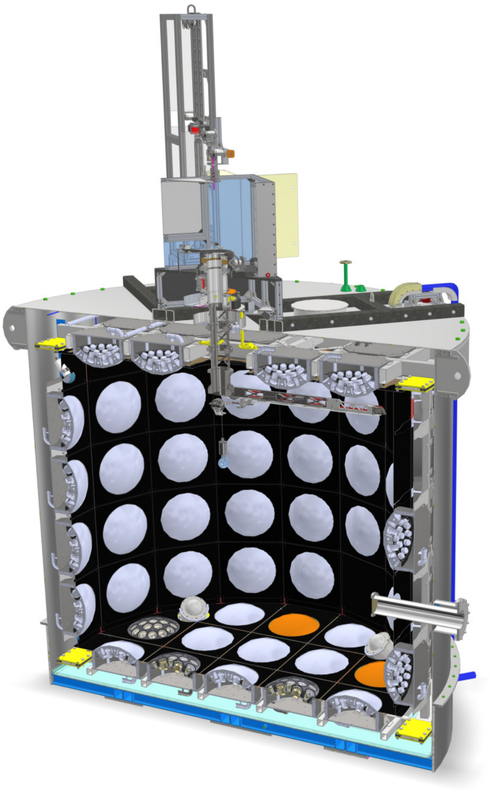


Figure 3.5: Technical drawing of the WCTE detector showing the mPMTs, CDS, photogrammetry cameras (white) and laser diffuser made by O. Jeremy. Empty mPMT slots are shown in orange.

PMTs [1] sharing the same housing and electronics. The 3-inch PMTs have been chosen for their good timing resolution, reduced dark rate and weaker sensitivity to the Earth's magnetic field compared to larger PMTs [133]. A picture of the IWCD-style mPMTs is presented in Figure 3.6a alongside a view of the WCTE barrel with all mPMTs installed in Figure 3.6b.

These PMTs are fitted with a reflective collar to increase their field of view and are arranged in a flower-like manner on a 3D-printed support matrix. They are coated with optical gel before being enclosed in a water-tight housing made of a PVC outer cylinder covered with a transparent acrylic dome. Two different methods have been used for applying optical gel within an mPMT, either the ex-situ or in-situ methods shown in Figure 3.7 a) and b) respectively. For WCTE, a total of 100 mPMTs have been produced, of which 64 and 36 have been produced with the ex-situ (each PMT is coved with optical gel individually) and in-situ (the optical gel is directly poured inside the acrylic dome) methods, respectively. A detailed summary of the WCTE mPMT assembly procedure is available in [136].

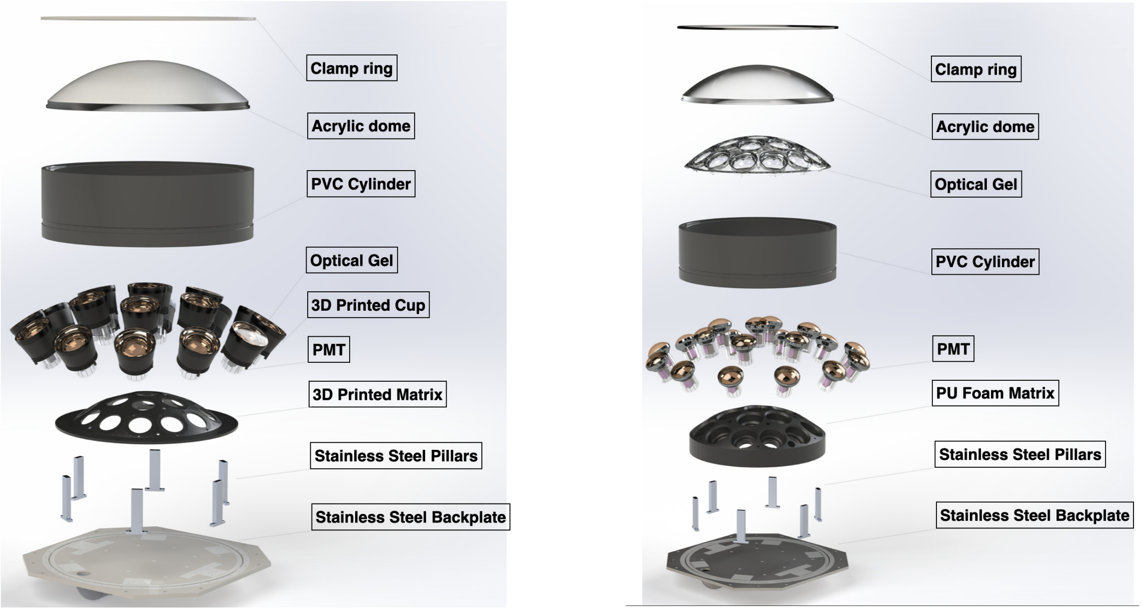
---

[1] Hamamatsu Photonics: `https://www.hamamatsu.com/jp/en/product/optical-sensors/pmt.html`

(a) Picture of WCTE mPMTs made with the in-situ (left) and ex-situ (right) gelling techniques.

(b) All mPMT and some of the black sheet in place in the barrel.

Figure 3.6: Pictures of the WCTE mPMTs. Picture credit: WCTE collaboration.



(a) Ex-situ style mPMT

(b) In-situ style mPMT

Figure 3.7: Diagram of the two IWCD mPMT designs. The mPMT LEDs are not indicated in the diagram but are fixed on the printed matrix between the PMTs and visible in Figure 3.6a. Figure reproduced from [136].

3.3.3 Black Sheet

Because of the relatively small size of the detector, the amount of light reflection from its wall has to be reduced as much as possible. Indeed, depending on the travelling distance, the reflected light can reach the mPMTs so soon after the prompt Cherenkov light that it can be difficult to distinguish from it, which biases the reconstruction.

The mPMT support structure is fitted with black PVC sheet to help limit the amount of reflected light. These sheets come in pre-cut standardised superposable pieces which maximise the coverage and simplify the installation. They are visible in Figure 3.4 and 3.6b.

3.3.4 Water System

Ultra-pure water is the target material in which particles Cherenkov light and potentially interact. It makes up most of the weight of the detector, and it is highly important to ensure that the WCTE water quality, level and transparency are well understood. The height, pressure and temperature of the water are constantly monitored by an ultrasonic level sensor on the top of the tank and a submerged pressure sensor which also measures the temperature and conductivity of the water. All of the hardware components that are expected to be in contact with the water have passed a soak test prior to installation to ensure that they do not reduce the water purity.

3.3.4.1 Ultra-pure water

The WCTE water purification system filters and chills the de-ionised water supplied by CERN ensuring good transparency. The purification system includes a set of filters, an ultra-violet sterilizer, an ultra-filter (which removes particles and micro-organisms larger than about 0.01μm) and a Total Organic Carbon (TOC) reduction unit which reduces the concentration of organic carbon in the water. The temperature of the water is maintained around 12° to limit bacterial growth.

3.3.4.2 Gadolinium-doped water

A portion of the 2025 beam time is dedicated to running with Gadolinium-doped water. About 60kg of $\mathrm{Gd_2\,(SO_4)_3}$ will be dissolved in the WCTE water thus reaching a 0.21% concentration, which is just over twice the final target concentration in the Super-Kamiokande detector's Gd-phase operation [137]. A 0.033% Gd concentration in Super-Kamiokande corresponds to a 75% probability of neutron capture [137]. WCTE expects more than 85% neutron tagging efficiency thanks to its higher Gd concentration. A dedicated 2m^3 mixing tank will be used to dissolve the $\mathrm{Gd_2\,(SO_4)_3}$ powder in batches of 8kg.

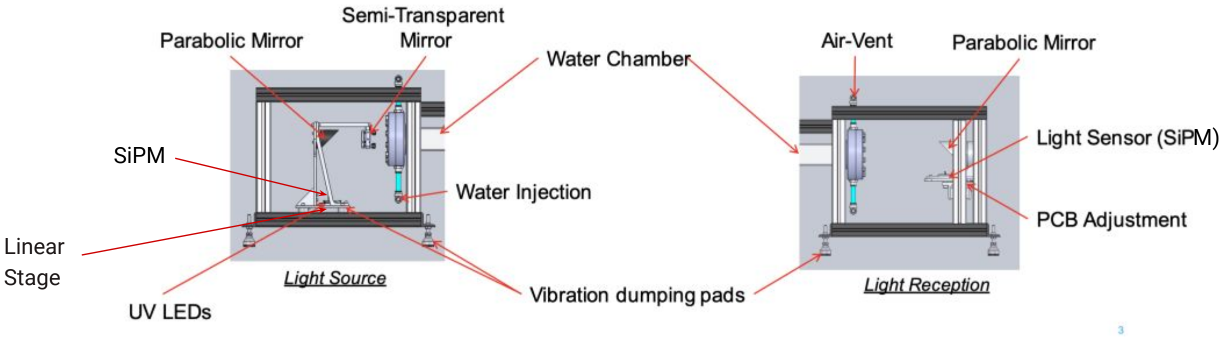


Figure 3.8: Diagram of the WCTE water monitoring system. Figure made by S. Taghayor.

3.3.5 Calibration Systems

3.3.5.1 Water monitoring system

The transparency of the water is continuously monitored using a 10 m long pipe coupled at one end to seven different wavelength LEDs and silicon photomultiplier light sensors at both ends, as shown in Figure 3.8. The light is measured both before and after travelling through the pipe filled with WCTE tank water or a pure water sample, which is produced using a built-in miniature purification system. Specific sources of contamination like fine dust, bacteria, and ions, absorb (and in some cases emit) light differently, revealing their presence in the water.

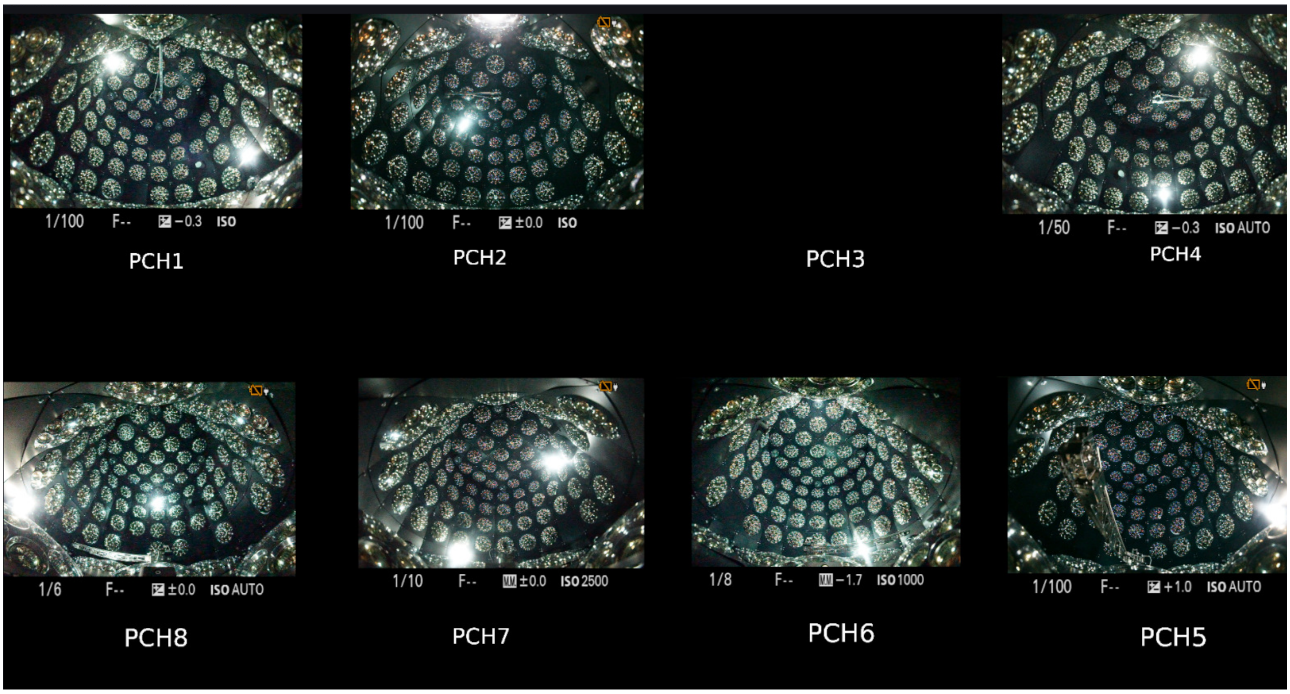
3.3.5.2 mPMT LED

Each of the mPMTs is fitted with two sets of LEDs, one set of three pulsed LEDs used for timing and optical calibration and one set of six continuous LEDs which serves as a target for the photogrammetry system. The pulsed LEDs are fitted between the central PMT and the first ring of PMTs in the mPMT, the continuous ones are located between the first and second ring. The pulsed LEDs have wavelengths of either 365 nm, 405 nm or 470 nm and have pulse widths smaller than 1ns. The mPMT LEDs are visible in Figure 3.6a, some of the LEDs are fitted with an additional plastic diffuser cap such that their light reaches more mPMTs.

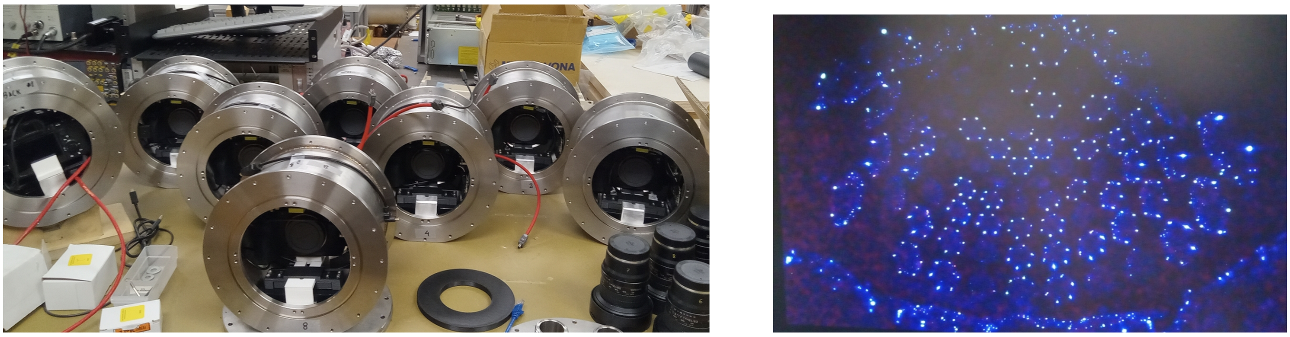
3.3.5.3 Photogrammetry

The water Cherenkov tank is fitted with eight photogrammetry cameras, of which seven are operational. These cameras are equipped with a fisheye lens and kept in water-tight enclosures across the detector's top and bottom endcaps. These cameras are used to measure the relative position of the mPMTs within the tank and verify the position of calibration sources as they are deployed within the tank. They also monitor the tank water filling and emptying procedures. Photogrammetry pictures taken with the white lamps and the mPMT continuous

LEDs are presented in Figure 3.9.



(a) Photogrammetry pictures taken with the white lamps showing the CDS on the top endcap.



(b) Picture of the cameras and their enclosures before assembly.

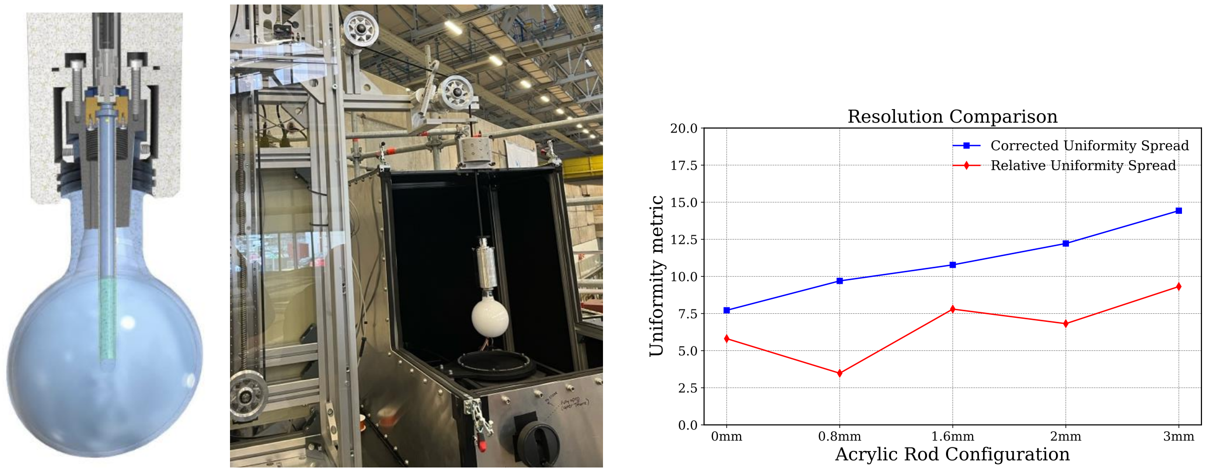
(c) Photogrammetry image with some of the mPMT slow LED used as light sources

Figure 3.9: The WCTE photogrammetry cameras and pictures. Picture credit: WCTE collaboration.

3.3.5.4 The laser diffuser

One of the calibration devices that can be deployed in the water tank using the CDS is the laser diffuser, pictured in Figures 3.10a and 3.10b. The WCTE's laser diffuser is similar in design to what was used by the SNO and SNO+ collaborations [138, 139] and shows similar performance. This optical calibration device is made up of a mixture of glass microspheres suspended in optical gel within a transparent acrylic sphere. An optical fibre connected to a 401.9 nm laser is embedded within the gel such that the laser light is diffused and as uniform as possible. Figure 3.10c shows that the WCTE laser diffuser light non-uniformity is about 10%.

Flashing the laser with the diffuser sphere placed at various positions within the tank permits the calibration of the mPMT timing and an additional measurement of the water quality properties using the relative amount of light seen by each PMT as a function of its distance to the light source.



(a) Technical drawing of the WCTE laser diffuser made by O. Jeremy.

(b) Laser diffuser ready to be deployed in the tank by the CDS.

(c) Characterisation of the laser diffuser showing a ∼ 10% light uniformity across the surface. Figure by K. Joseph.

Figure 3.10: Design, picture and light uniformity of the WCTE laser diffusers.

The variation in the time taken by photons to escape the laser diffuser is 0.33 ± 0.03 ns across the entire surface, which is much shorter than the PMTs' response time. They therefore see the light from the diffuser as an instantaneous burst. Simulation work has shown that the attenuation length of light in water can be measured with a 15% accuracy using the laser ball, despite the small size of the detector. All of the calibration techniques and hardware designs that are developed for WCTE will be re-used for the IWCD detector.

3.3.5.5 UKLI diffuser for Hyper-Kamiokande

WCTE is also used to test the diffuse light injector (UKLI diffuser) prototype produced for Hyper-Kamiokande. A picture of it in WCTE is shown in Figure 3.11. The UKLI is mounted on the support structure in the CDS home slot position. It was successfully operated in 2024, and its light was seen by the bottom endcap mPMTs as expected. Further analysis is underway in preparation for the Hyper-Kamiokande assembly.

3.3.5.6 Radioactive sources

The last item in the WCTE calibration suite is a pair of radioactive sources. Figure 3.12a shows the Nickel-Californium (NiCf) calibration source, which is made up of a 10 μCi (0.037MBq) ^{252}Cf radioactive source encapsulated in a brass rod, which is embedded in a 13.5cm diameter Nickel ball. This ball is made up of a mixture of Nickel oxide (NiO) (42%), epoxy (36%) and high-density polyethylene (22%). The design of this source is greatly inspired by work done by the Super-Kamiokande calibration group [124]. Californium-252 can undergo spontaneous fission, which emits a lot of neutrons. Those thermal neutrons are captured by ^{58}Ni

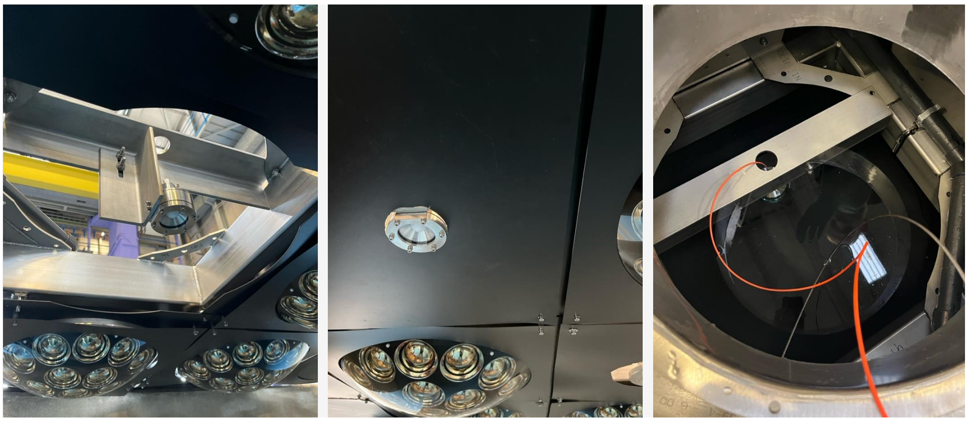


Figure 3.11: Picture of the UKLI diffuser in WCTE (left) during assembly, (centre) after installation of the black sheet and (right) from above the support structure. Picture credit: WCTE collaboration.

atoms to form an excited ^{59}Ni atom, which decays back to its ground state by emitting a photon of about 9 MeV, which emits Cherenkov light in WCTE. This calibration source is used to study the single photoelectron response of the individual PMTs and measure their relative efficiencies.

(a) Picture of the Ni-Cf calibration source about to be lowered into the tank by the CDS. Picture credit: WCTE collaboration.

(b) Technical drawing of the AmBe calibration source made by O. Jeremy. The BGO crystals are shown as a grey cylinder inside the water-tight enclosure.

Figure 3.12: The two radioactive calibration sources used in the WCTE.

The second radioactive calibration is an Americium-Beryllium (AmBe) source used to test the neutron detection capabilities of WCTE. The AmBe calibration device is composed of a 100 μCi AmBe encapsulated source of α-particles fitted inside a block of scintillator Bismuth Germanate (BGO) crystal. The Americium decays predominantly through the emission of an α-particle which can be absorbed by a Beryllium atom to produce an excited Carbon-12 nucleus and a free neutron. The Carbon rapidly decays to its ground state by emitting a 4.4 MeV photon, which causes scintillation in the BGO crystal and acts as a tag for the neutron emission. The neutron thermalises in the Gd-doped water (with a capture time of 204.8 ± 0.4 μs) [137], producing a 2.2 MeV photon which rapidly pair-produces into an electron and a positron which emit Cherenkov light detected by the mPMTs.

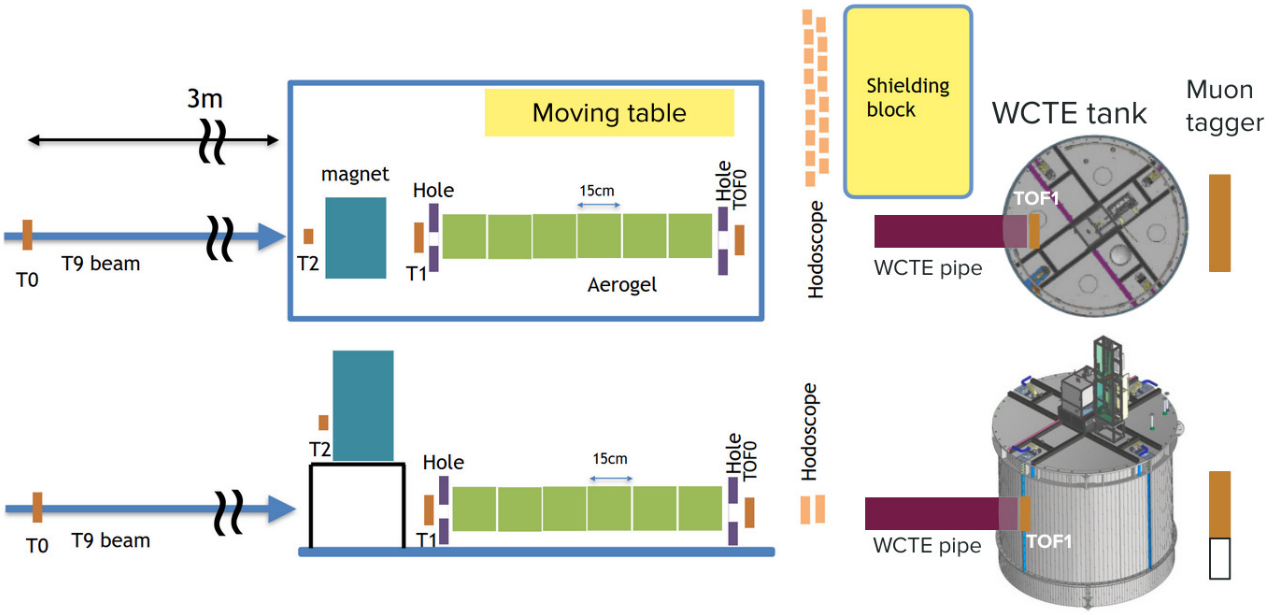


Figure 3.13: Top and side view of the two-story beam monitoring set-up used in 2024 and 2025, which re-uses detectors and expands on the 2022 and 2023 beam tests set-ups. Adapted by the author from a figure made by A. Konaka.

3.4 The beamline monitoring system

Upstream of the WCTE tank lives a suite of detectors which is designed to identify the incoming particle type and provide a measurement of its momentum. The WCTE beam apparatus used in 2024 and 2025 is divided into two systems on two separate levels: the charged particle set-up on the bottom and the tagged photon set-up on the top, as shown in Figure 3.13. Most of the components making up this beam apparatus are re-used from two test beams in 2022 and in 2023, and are presented in detail in Chapter 4 dedicated to the 2023 charged particle beam test.

Two sets of detectors are newly introduced for the 2024 and 2025 WCTE operation: a pair of scintillator slabs placed behind the water tank to tag any through-going muon and the new time of flight (TOF) detectors, which are thin scintillator slabs coupled to 16 ultra-fast MPPCs. The new TOF system provides a better timing resolution than the trigger scintillators (T1 and T0) used in 2023.

The tagged-photon setup is a novel compact facility for producing tagged photons of known energy using a low-momentum, high-intensity electron beam. This set-up uses a compact (16×24×24 cm) hollow 0.23 Tm permanent magnet made up of 16 Neodymium (N52) permanent magnets arranged in a Halbach array configuration, which was originally built for the EMPHATIC experiment [140]. The charged particles are deflected away from the calorimeter by the magnetic field. If electrons radiate one or more Bremsstrahlung photons while traversing the beamline, they lose energy and are bent more strongly by the magnetic field, causing them to strike a different scintillator bar among the 15 that make up the off-axis hodoscope. The position of the bar can be directly related to the energy of the emitted photon, knowing the initial electron energy and assuming that a single photon was produced. The method is summarised in Figure 3.14.

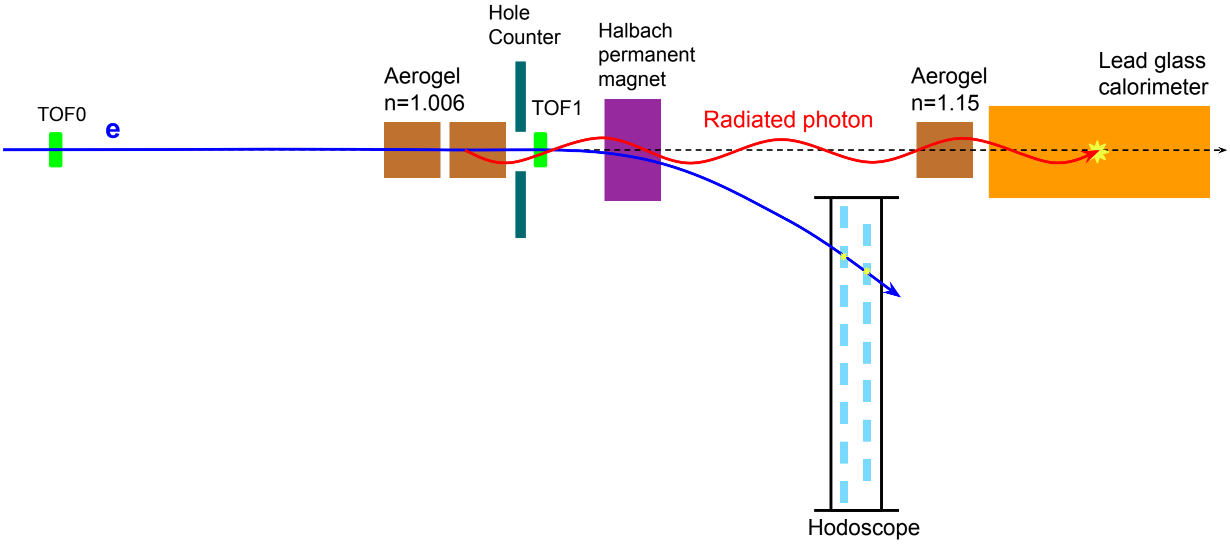


Figure 3.14: Diagram of the method used to produce tagged photons, the calorimeter is used for energy calibration and the other detectors for vetoing.

This method was successfully demonstrated in 2023, where a calorimeter was used in place of the water Cherenkov detector to measure the photon energy. Figure 3.15 shows a linear relationship between the expected photon energy (calculated from the electron bending radius and its initial energy) and the coincident charge measured by the calorimeter. Section 4.5.3 describes the energy calibration method that was devised during the 2023 beam test to convert the measured lead glass charge to photon energy. Most of the 2023 beam data analysis methods are reused for WCTE operation.

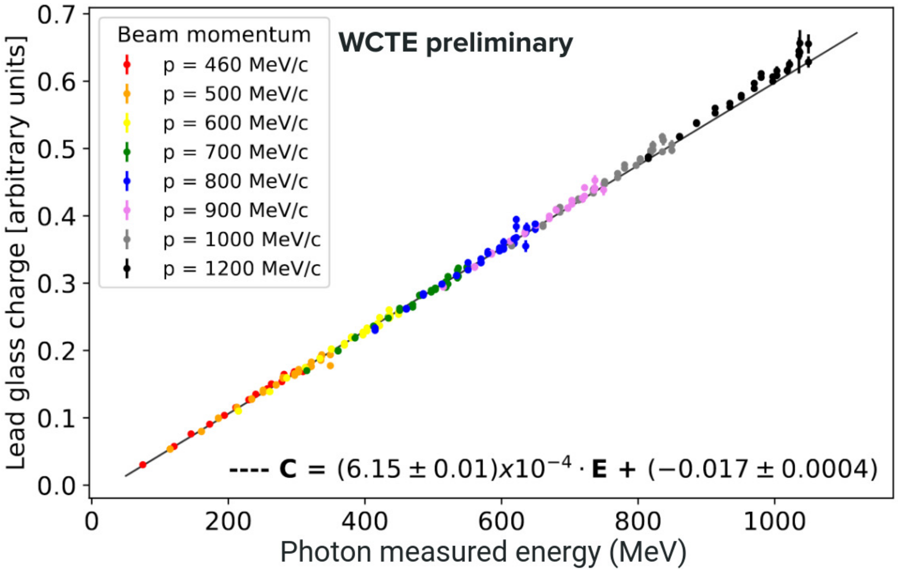


Figure 3.15: Charge deposited by photons in the calorimeter against their energy as measured by the hodoscope during the 2023 beam test. Figure made by J. Renner.

Chapter 4

PID and characterisation of the CERN T09 beam

4.1 Motivation

Physics studies made by WCTE require that the particle type and momentum is measured with high accuracy and precision upstream of the WCTE detector. Table 4.1 summarises the purity that the beam line PID has to reach for each WCTE Physics analysis, which are very high and require a dedicated and robust PID analysis.

This Chapter describes the experimental set-up in Section 4.2 and the methods developed to analyse the data collected during the 2023 beam test in Sections 4.2-4.5. It also presents in Section 4.6 the first characterisation of the T09 beam since its renovation in 2019-2021, which is critical for planning the WCTE data taking campaign and relevant for other users of the beam line. At the time of writing, no characterisation of the T09 beam in the sub-GeV region has been published.

Running Mode	Analysis Name	Particle	Purity Required	Statistics Required
Pure water	Muon/electron separation	μ and e	$\geq$99.99%	400 mis-IDed muons
	Muon scattering	μ	$\geq$99.99%	10^4 scattered muons
	Pion scattering	π	$\geq$99%	10^4 scattered pion
	Photon/electron separation	γ and e	$> 80\%$	$\geq$ 2000 photons
Gd-doped	Neutron production	π	$> 90\%$	10^4 pions
	Photo-nuclear effects	γ	-	2×10^5 photons

Table 4.1: Summary of the sample purity and statistics that need to be reached by the beam line apparatus for a sub-set of the WCTE Physics studies.

4.2 Experimental set-up for the 2023 operation

The WCTE beam monitoring system is composed of nine detectors aligned along the beam direction as presented in Figure 4.1.

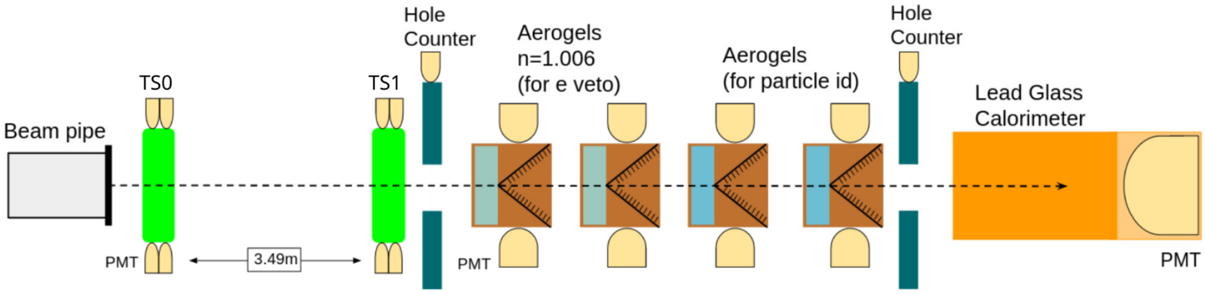


Figure 4.1: Diagram of the experimental setup that was used to identify charged particles and measure their momentum during the July 2023 beam test at CERN.

4.2.1 ACTs

In the 2023 charge beam set-up, there are four Aerogel Cherenkov Threshold detectors (ACTs), which were developed by the WCTE collaboration. A technical drawing and pictures of the ACT detector can be seen in Figure 4.2.

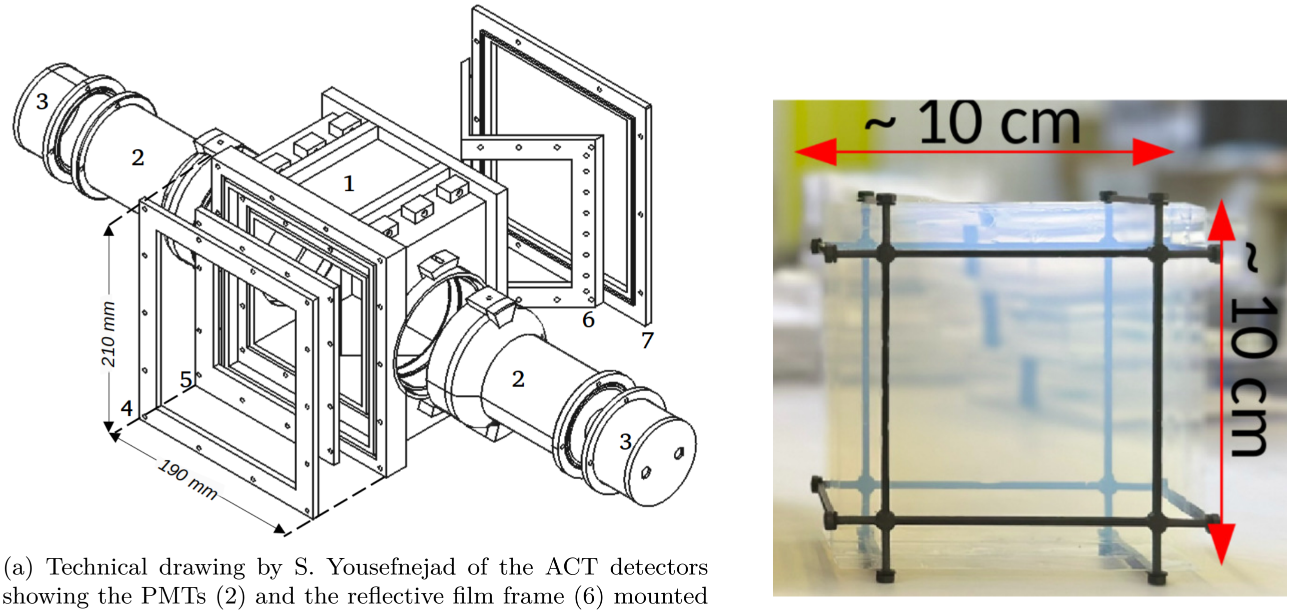
Each of the four ACT boxes holds a block of silica aerogel Cherenkov radiator of a given refractive index, produced for WCTE by the Aerogel Factory[1] in Japan. The aerogel blocks used during the 2023 beam test are presented in Table 4.2. A charged particle of mass m is above Cherenkov threshold, and produces Cherenkov light when crossing a material with refractive index n, if its momentum is:

$$\mathrm{p} \geq \frac{m}{\sqrt{n^2 - 1}}. \tag{4.1}$$

The two upstream ACT detectors, ACT0 and ACT1, have $n = 1.006$, which means that for particles with momenta below 1 GeV/c, only electrons are above the Cherenkov threshold of the aerogel. These detectors provide online and offline electron veto capabilities as described by Sections 4.2.5 and 4.4.2.2, respectively. The two downstream ACTs boxes, ACT2 and ACT3, are chosen out of the available boxes presented in Table 4.2 such that muons are above the Cherenkov threshold and pions below. Figure 4.3 shows how the available aerogels are used to cover the momentum range 220-1200 MeV/c. The two downstream boxes have the same refractive index and are swapped in and out of the beamline depending on the momentum range under investigation.

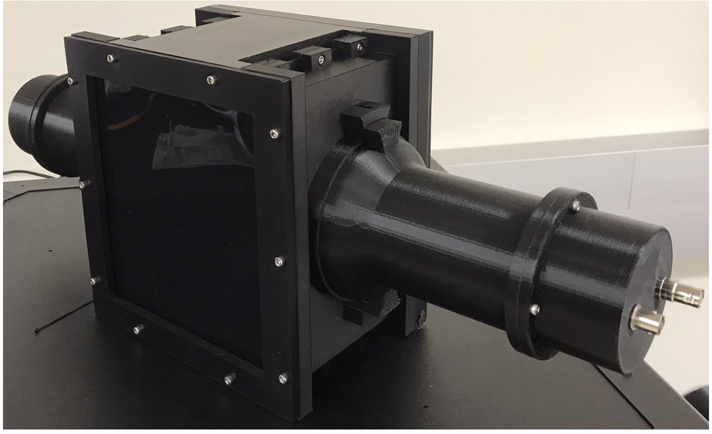
---

[1] Makoto Tabata, Aerogel Factory Co., Ltd, `https://aerogel-factory.jp/`



(a) Technical drawing by S. Yousefnejad of the ACT detectors showing the PMTs (2) and the reflective film frame (6) mounted inside the 3D printed box (1).

(b) Picture of an aerogel block in its holder.



(c) Picture of the assembled ACT detector.

Figure 4.2: Presentation of the ACT detectors used and developed by the WCTE collaboration. Picture credits: WCTE collaboration.

		Threshold momenta (MeV/c)	
n	Thickness (cm)	μ	π
1.006	8+8	962	1258
1.01	6+6	744	973
1.015	6+6	607	803
1.02	6+6	525	687
1.03	4+6	427	559
1.047	8+8	340	445
1.06	4+6	300	393
1.11	2+2	219	286
1.13	2+2	200	262
1.15	2+2	185	243

Table 4.2: WCTE 2023 beam test Aerogel Cherenkov detectors. The electron is above Cherenkov threshold in all of the aerogels.

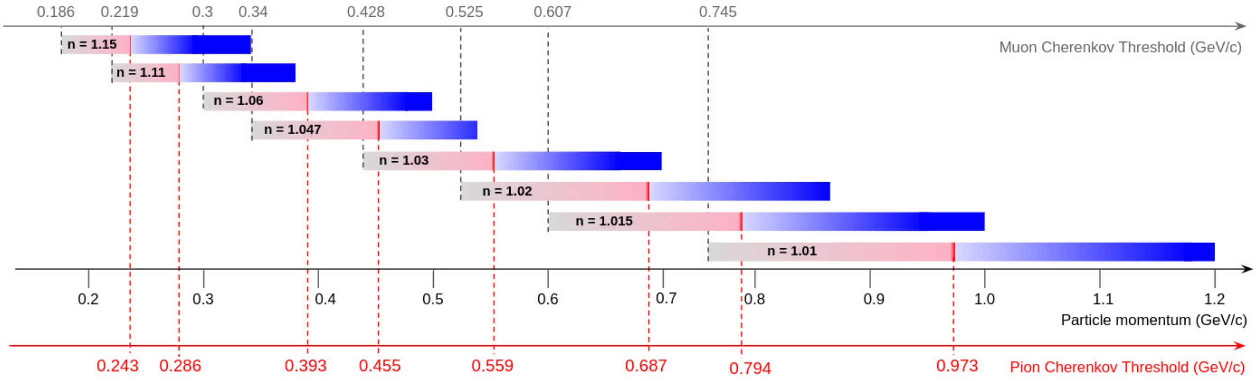


Figure 4.3: Presentation of the aerogels used as a function of the beam momentum.

4.2.2 Trigger scintillators

As they leave the beam pipe, particles cross first a pair of scintillator trigger counters, labelled TS0 and TS1, separated by a distance $L = 3.49 \pm 0.01$ m. Each trigger scintillator detector is composed of a $9 \times 9 \times 0.635$ cm slab of plastic scintillator, optically coupled to two 1-inch Hamamatsu R6427 fast PMTs on either side (four in total). The TS detectors are used to measure the time of flight (TOF) of each particle, as described in Section 4.5.1.

4.2.3 Lead-Glass calorimeter

An electromagnetic Lead-Glass calorimeter (shown in orange on Figure 4.1) is located at the end of the beam line. The $10 \times 12 \times 36$ cm lead-doped crystal of glass is optically coupled to a PMT and wrapped in a reflective Mylar sheet to contain the light produced by particles inside it. Pictures of the Lead-Glass crystal and the assembled calorimeter before it is light shielded are shown in Figure 4.4. In the 2024 and 2025 set-up the lead-glass calorimeter is situated on a height-adjustable table in front of the water Cherenkov detector so it can be placed in the beamline to perform calibration (e.g. for the tagged photon set-up) and lowered out of the way, letting the particles reach the tank, otherwise.

4.2.4 Hole counters

A pair of scintillator bars with a circular hole in their centres, shown in dark green on Figure 4.1 is used to veto particles travelling in the halo surrounding the beam. These particles are expected to have different properties from those of particles belonging to the core of the beam [141]. A disk of 45 mm and 60 mm in diameter has been cut out of the upstream (HC0) and downstream (HC1) scintillators, respectively. The width of these hollowed out portions corresponds to the beam width at the position of the two hole counters. Each scintillator slab is coupled to a PMT R10233-100 Hamamatsu PMT and completely light shielded. A picture of the hole

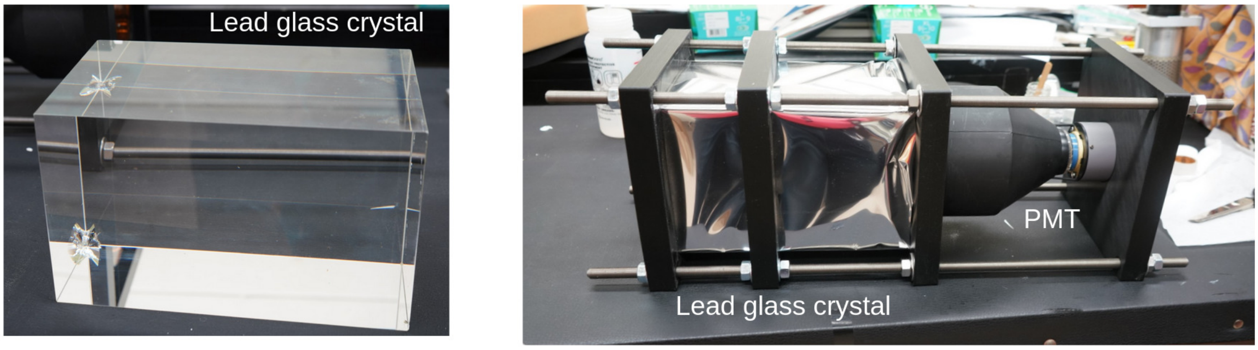


Figure 4.4: Picture of the Lead-doped glass crystal (left) and of the assembled calorimeter prior to it being light shielded with black adhesive tape (right). Pictures by M. Mansoor.

Figure 4.5: Picture of one of the hole counters composed of a hollow scintillator slab (right) coupled to a PMT (left). Picture by D. Divecha.

counters is shown in Figure 4.5. Any particle producing light in the hole counters crossed the scintillator slab instead of travelling through the hole, meaning it was not part of the beam centre. The DAQ system vetoes these particles.

4.2.5 Data acquisition and trigger logic

During the 2023 beam test, the WCTE team used four CAEN DT5730 500MSPS digitisers to collect data from the nineteen PMTs present in the setup. Each digitiser has eight channels and each channel is connected to at most one of the PMTs. Out of the nineteen PMTs used in the setup, all except the ACT0 ones were coupled to custom-made pre-amplifiers, which scaled the amplitude of the signal up by a factor of ten. The two ACT0 PMTs were connected to VT120 200× pre-amplifiers.

The eight PMTs of TS0 and TS1 were connected to the same digitiser. The eight PMTs of the ACT detectors were connected to the second digitiser, leaving the lead-glass calorimeter and HC PMTs connected to the third digitiser. Unfortunately, the clocks on the three digitisers could not be phase locked during the data-taking, which caused some discrepancies in time between the different detectors. Section 4.3.2.2 discusses how this issue was fixed in the data analysis pipeline.

Twelve out of the 19 PMTs are used to trigger the data acquisition (DAQ). These are the ones coupled to TS0, TS1, ACT0, HC0 and HC1. For these PMTs, the output of the amplifier is sent first to a splitter which sends 20% of the amplitude to the digitiser and the remaining 80% to a discriminator. The threshold of the discriminators is set approximately 3σ away from the mean pedestal value of each PMT. The pedestal of a PMT is the voltage it records when no photo-electrons (PEs) are detected. It is the discriminators' output that is used in the trigger logic. The amplified signal corresponding to the seven PMTs that are not used in the trigger is directly sent to the digitisers. When the digitisers receive a trigger, they write to disk the 540 ns of data surrounding the arrival time of the trigger.

We designed the trigger to accept only real particles crossing the beam line, which are not part of the beam halo and (ideally) not electrons, which make up the overwhelming majority of particles in the beam. Dedicated electron-rich samples were taken without the electron veto for studying the beam structure. The pair of trigger scintillators, TS0 and TS1, ensures that a real particle is going through by requiring loose coincidence between all eight of their PMTs. Since only electrons are expected to emit Cherenkov light in ACT0 and ACT1, whose refractive index is $n = 1.006$, we require that the signal from neither of the ACT0 PMTs exceeds the discriminator threshold within a 40 ns-long window opened by the coincident detection of a signal in the four TS0 PMTs. Only ACT0 is used as part of the online electron veto; ACT1 is used offline to tag electrons and study the performance of the online vetoing made with ACT0. Finally, the two hole counters ensure that the particle is not part of the beam halo by requiring that neither of the two PMTs is above threshold within the same 40 ns window. This ensures that the particle went through the hollowed-out part of both detectors and, therefore, that it is part of the beam core. The hole counters were also used to align the beam.

4.3 Data pre-processing

After the data is collected, it needs to be pre-processed before it can be analysed. The pre-processing code applies some preliminary calibrations to the data and removes unusable events.

4.3.1 Peak finding algorithm

Firstly, a simple peak finding algorithm is applied to the data to count and flag the peaks in the waveform. The peaks are defined with respect to the pedestal of the PMT. The mean value and standard deviation of the pedestal are calculated using a portion of the waveform where no signal is expected. Depending on the PMT, this portion is located either before or after the expected signal. Figure 4.6 shows an example of a digitised waveform presenting a peak at 38ns. The periodic fluctuations in the pedestal value are caused by the pre-amplifier.

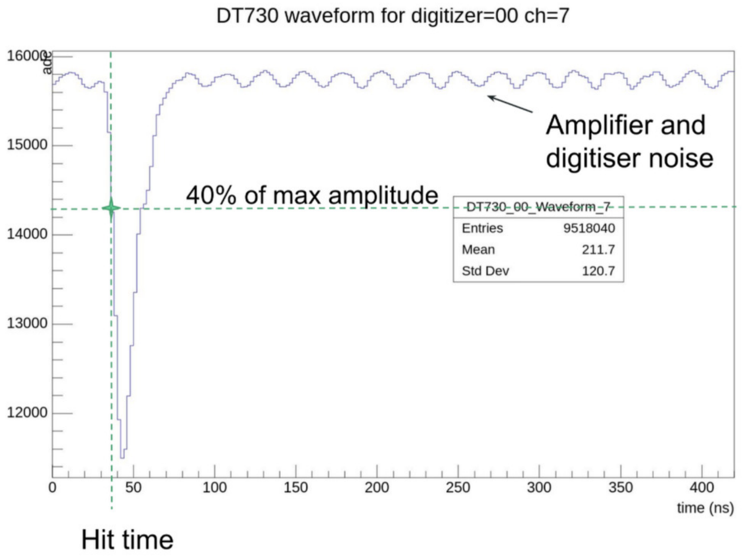


Figure 4.6: Example of a waveform that is saved for one of the ACT PMTs with a clear peak above the noisy pedestal. In this case, the portion of the waveform after 200 ns is used to measure the pedestal.

The peak finding algorithm identifies as part of a peak any portion of the waveform whose amplitude is greater than either three times the pedestal's standard deviation or 20% of the largest amplitude recorded in the waveform, whichever is highest. This ensures that no pedestal fluctuations are flagged as peaks. The hit time associated with a peak corresponds to the middle of the 2 ns-wide time bin where the peak reaches 40% of its maximum amplitude.

Something worth highlighting is that multiple particles are sometimes expected to cross the beam line within the 540 ns of the stored waveform. We know from studies discussed in section 4.6.4 that the delay between two consecutive beam bunches is about 330 ns and that each bunch can contain more than one particle. We observed that the minimal delay between two consecutive peaks in the same waveform decreases as the momentum increases until it plateaus at about 20 ns. From this and looking at individual waveforms, we conclude that the peak finding algorithm cannot resolve peaks that are less than 20 ns apart. Section 4.3.5 discusses how we remove events with two (or more) particles.

4.3.2 Preliminary calibrations

4.3.2.1 Voltage conversion to photoelectron

The digitisers store charges in ADC units, which we convert to volts using a fixed factor. To make meaningful comparisons across detectors, which is particularly important for the ACT detectors, we have to convert the voltage to units of photo-electrons (PE). This conversion factor is called the gain. To perform the gain calibration, a few datasets were recorded using a random trigger whilst the beam was off. Since no particle was going through the beam line, these datasets represent the intrinsic noise of the PMTs. The noise comes from a wide

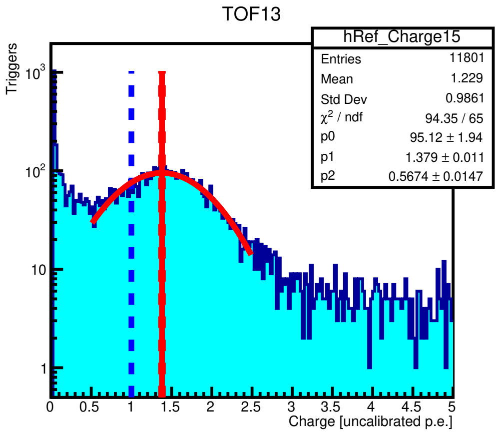


Figure 4.7: Charge spectrum of one of the ACT3 PMTs taken during beam off time. The 1PE peak is fitted with a Gaussian distribution. The units shown here are an arbitrary conversion of ADC to the number of PE before gain calibration; the blue dashed line indicates the 1 PE mark. Figure by J. Kvita.

range of sources, including but not limited to, the thermal production of an electron inside the PMT, charge deposition by a cosmic ray crossing the PMT, radioactivity within and outside it and leakage current [142].

The sources of noise which produce a physical photon or electron appear as a clear peak in the PMT spectrum over the exponentially decaying background corresponding to other processes. Additionally, a pedestal peak, corresponding to the baseline signal when no photon is detected serves as a reference point. The calibration procedure fits a Gaussian distribution to the single photoelectron (PE) peak, which is well-separated from the pedestal, to extract the gain of the PMT. The gain is unique to each PMT and depends on the supplied voltage. This calibration process is illustrated in Figure 4.7.

4.3.2.2 Timing calibration across detectors

As will become apparent in the following sections, the timing recorded by the trigger scintillators is an essential tool when doing PID with the WCTE beam set-up. The pre-processing performs a preliminary timing calibration of these detectors.

Relative offsets in the timing of PMTs belonging to the same detector (mainly due to differences in cable lengths) are removed first. We then use a sample of electron-like events (as identified by the calorimeter) to perform inter-detector calibration. There, we ensure that the difference between the mean arrival time of electrons between the two trigger scintillators was equal to L/c, the expected electron arrival time. This preliminary calibration adds fixed values to the recorded hit times for a coarse agreement between detectors. The finer run-dependent inter-detector timing calibration is presented in Section 4.5.1.

Additionally, the absolute hit times recorded by the three digitisers had to be synchronised since they are not phase-locked and did not all start at the same time. The clock of each digitiser has its own frequency, ticks every 8 ns approximately and reads data out every other tick of the clock (i.e. every 16 ns). This means that for a trigger arriving at the same time in two digitisers, the corresponding recorded trigger times can differ by as much as 16 ns between the two digitisers.

The difference in digitiser periods (T_i where $i = 0, 1, 2$ for the three digitisers) can be measured by fitting the linear relationship that exists between t_0 and $\Delta t_i = t_0 - t_i$ where t_i is the time at which the global trigger is recorded by digitiser i. Because of the slightly different clock periods, the width of the time bins is not exactly 8 ns for all digitisers, causing the position of the global trigger to drift relative to the reference digitiser as time goes by.

We assume for simplicity that the period of one of the digitisers, used as a reference, is $T_0 = 8$ ns. Using this method, we measure that the two other digitisers are offset from T_0 by $+1.63 \times 10^{-5}$ ns and -4.69×10^{-6} ns respectively. Based on these differences in period, for each event we correct the width of time bins recorded by these two digitisers such that they match the time domain of the reference digitiser, allowing us to compare timing information across detectors connected to different digitisers.

4.3.3 Coincidence requirements

It is essential to correctly identify the arrival time of a particle in the two trigger scintillators (TS) to accurately measure its TOF. Furthermore, it is important to flag any accidental hits which would otherwise bias the TOF. To do that, the pre-processing code requires coincidence of hits within and across trigger scintillators. Hits in two PMTs from the same TS are in coincidence if they are no more than $3\sigma_{\mathrm{TS}}$ from each other where $\sigma_{\mathrm{TS}} = 0.368$ ns is the standard deviation of the Gaussian distribution fitted to the delays between hits in any two PMTs belonging to the same detector.

The mean of the coincidence-matched times t_{TS} in the two PMTs is used as the arrival time of the particle in this TS, which slightly improves the timing resolution by removing offsets caused by the photon travelling within the scintillator.

Once a coincidence has been found between a pair of PMTs in each of the TS, these pairs of hits are matched if they are in loose coincidence with each other, that is:

$$t_{\mathrm{TS0}} + \frac{L}{c} - 7 \text{ ns} \leq t_{\mathrm{TS1}} \leq t_{\mathrm{TS0}} + \frac{L}{c} + 45 \text{ ns} \tag{4.2}$$

where the upper bound accounts for the difference in TOF expected for heavier particles. In case two hits satisfy the above condition, we select the earliest one. A preliminary version of the analysis required that there be exactly one peak in all eight of the trigger scintillator PMTs. At 1 GeV/c, that selection accepted 75% of events [143], whereas the coincidence-based selection described above finds that 95.07% of events have exactly one particle going through the beamline within a given window. Two and three coincidences are found in 3.24% and 0.05% of events, respectively, bringing to 98.35% the total fraction of events for which at least one particle is found to cross the beamline.

4.3.4 Integration window

One of the main challenges of this analysis is to correctly match signals that correspond to the same particle across the beamline detectors. Only considering the peaks found by the peak finding algorithm is not sufficient for two reasons:

1. Some particles are not expected to produce any light in some of the ACTs (e.g. pions), meaning that any accidental peak in the ACTs reduces the PID capabilities.
2. The signal produced by the ACT and lead glass detectors tends to be reconstructed as multiple peaks despite having been produced by the same particle, because the light reflects and scatters within the detector.

To overcome both of these challenges, the pre-processing code integrates the waveform over a window centred around the expected arrival time of the particle. The duration of the integration window depends on the detector. It maximises the amount of charge collected for a single particle whilst minimising the window overlap in case multiple particles are present in the waveform. For the ACTs, the optimal window was found to be 61 ns-long, starting 16 ns before the expected time of the particle as presented in [143].

The expected arrival time of the particle in detector x is:

$$t_d = t_{TS0} + (t_{TS1} - t_{TS0})\frac{D_x}{D_{TS1}} \tag{4.3}$$

where $(t_{TS1} - t_{TS0})$ is the time of flight of the particle between TS0 and TS1 and D_x is the distance between detector x and TS0. The distance between each detector and TS0 is presented in Table 4.3.

Detector	TS1	HC0	ACT0	ACT1	ACT2	ACT3	HC1	Lead-Glass
D_x (cm)	349 ± 1	354 ± 5	362 ± 5	377 ± 5	392 ± 5	407 ± 5	417 ± 5	457 ± 10

Table 4.3: Distance between each detector and the upstream trigger scintillator, used to estimate the arrival time of the particle in each detector.

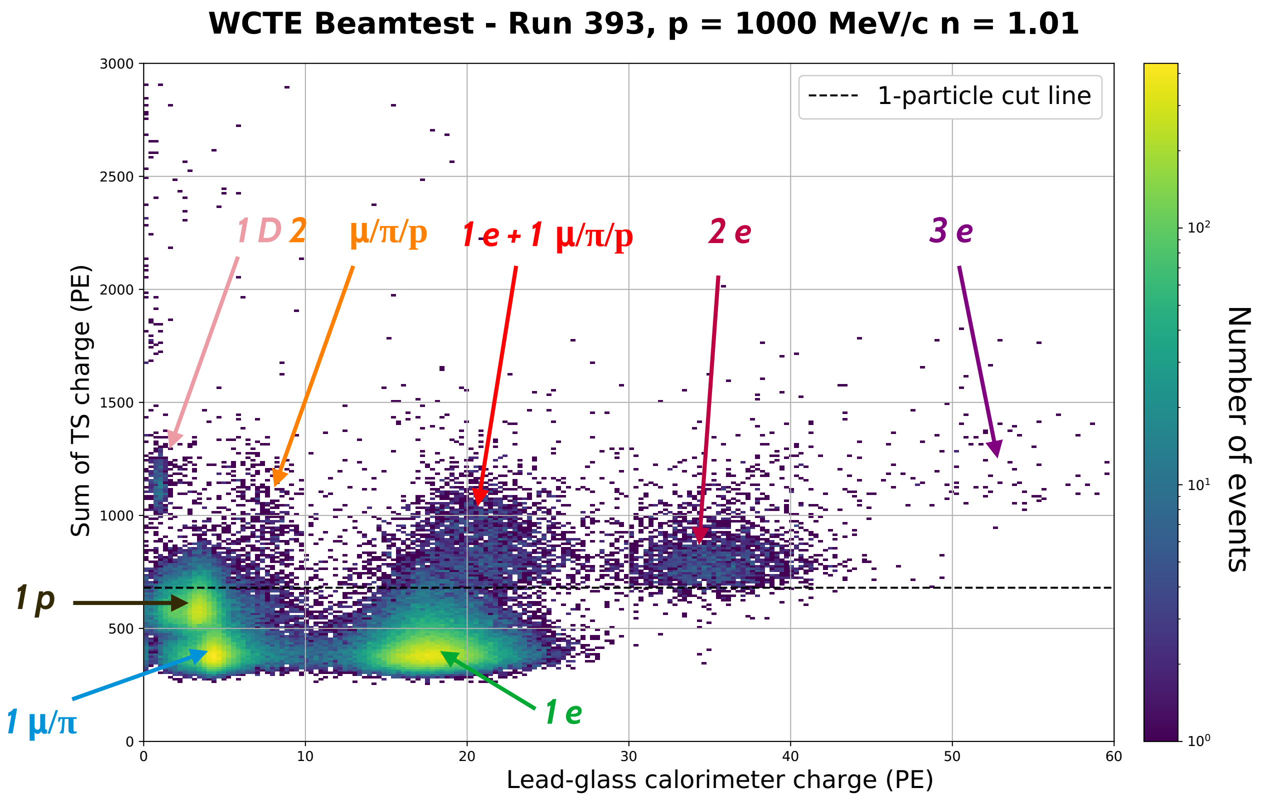


Figure 4.8: Plot of the charge in the trigger scintillators against the charge in the lead glass calorimeter for a run at 1GeV before applying the coincidence requirement described in Section 4.3.3. The cut is only applied to light (electron, muons, pions) particles.

4.3.5 Removal of multiple particle events

After the timing offsets have been calibrated and the correct portion of the waveform has been integrated, the last thing that the pre-processing code has to handle is the case where two particles arrive so soon after one another that they are treated as a single one, as discussed at the end of section 4.3.1.

These events, where two (or more) particles are coincidental, are identified by looking at the energy deposited in the trigger scintillators. In the sub-GeV range, we expect from the Bethe-Bloch equation that protons and deuterium nuclei will deposit significantly more energy in the TS than electrons, muons and pions. The latter two are minimally ionising particles (MIPs), which will produce a roughly constant amount of light in the TS.

The pre-processing algorithm identifies “light” particles (electrons, muons and pions) based on their TOF and removes any of those events for which the total charge collected in TS0 and TS1 is more than 700 PE. Figure 4.8 shows the amount of charge produced by particles in the trigger scintillators against the charge measured by the calorimeter. Events with a short TOF and with TS charge greater than about 700 PE tend to have more than one peak in coincidence, whereas the overwhelming majority of other events have only one, confirming the validity of this method in removing events where more than one particle arrives within the same window.

4.3.6 Removal of noise events

The pre-processing code also removes all events which deposit less than 1 PE in total in ACT2 and ACT3. These events were identified as noise, where either the online triggering or the offline coincidence check failed. Discarding this population increases the purity of the selection without significantly reducing its efficiency since only about 2% of triggers are below this noise threshold.

4.4 Particle identification

4.4.1 Proton and deuteron identification using time of flight

Using the Lead-Glass calorimeter to select electrons, we can test the performance of the TOF detectors. We measured that the standard deviation of the TOF of the electron-like population is at most $\sigma_e^{\mathrm{TOF}} = 0.35$ ns [143]. Figure 4.9 shows the nominal difference in time of flight for different particles. From this, we see that the TOF measured using the TS can be used to separate protons from pions and deuterium nuclei.

An event is identified as proton-like if its TOF is within 3 ns of the nominal proton TOF and no less than $3\sigma_e^{\mathrm{TOF}}$ (i.e. 1.15 ns) away from the nominal pion TOF. This ensures that no pions or deuterium particles are misidentified as protons.

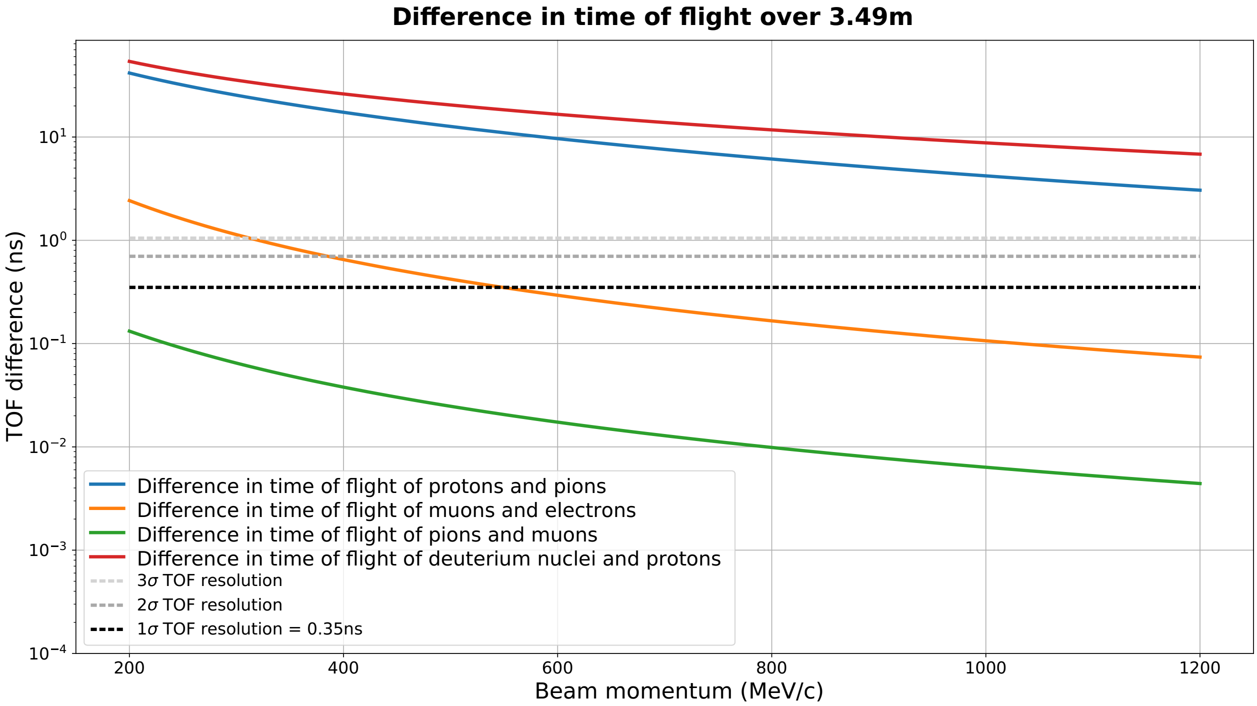


Figure 4.9: Expected difference in the time taken by various particles to travel a 349 cm distance. The experimental resolution on the TOF is about 0.35 ns, which is sufficient to separate protons from pions and deuterium nuclei but not electrons from muons or muons from pions in the 200-1200 MeV/c momentum range.

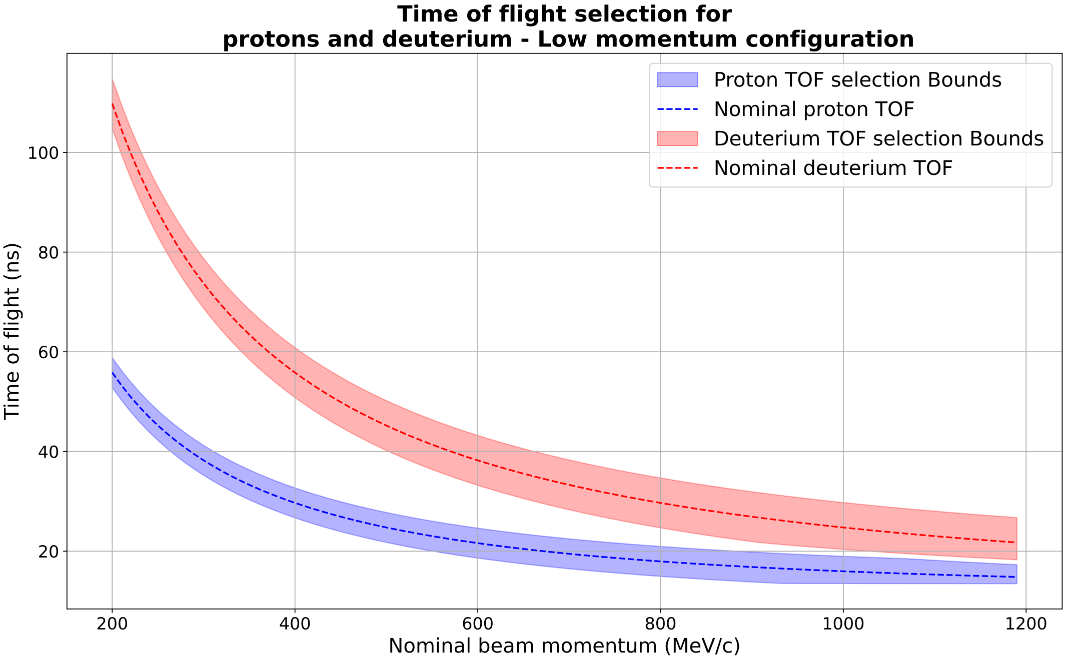


Figure 4.10: Plot of the TOF bounds used to select proton-like and deuterium-like events overlaid onto the calculated TOF for these particles (dashed line) as a function of their momentum.

Events are flagged as deuterium-like if their TOF is within −5 ns and +3 ns of the nominal deuterium TOF and greater than the mean TOF between a deuteron and a proton. The selection bounds for identifying protons and deuteron-like events are presented in Figure 4.10. An example of TOF distribution in data is shown later in Section 4.5.1.

4.4.2 Electron identification

Electron removal is the most difficult part of the PID procedure as they make up the majority of the beam and can contaminate both the muon and pion samples. To remove them we use the calorimeter to provide a tag for the particles (either electron or non-electron) and use this tag to determine the optimal ACT cuts for rejecting electrons.

4.4.2.1 Electron tagging using the lead-glass calorimeter

The Lead-Glass calorimeter is used to perform a preliminary electron tagging. Figure 4.11 shows the charge deposited in the calorimeter by all fast particles (i.e. not protons nor deuterium nuclei) for a 1000 MeV/c beam.

A Gaussian distribution $G(A_x^{\mathrm{LG}}, \mu_x^{\mathrm{LG}}, \sigma_x^{\mathrm{LG}})$ where $x = \{e, \mu\}$ is fitted to the two peaks in the lead glass charge spectrum. The peak that is at about 16 a.u. corresponds to electrons depositing all of their energy in the form of an electromagnetic shower, whilst the peak at 4 a.u. corresponds mainly to minimally ionising muons.

Because pions are hadrons, their energy deposition in the electromagnetic calorimeter is much more complex than that of leptons, depositing between very little and all of their energy in the form of visible electromagnetic showers, depending on their interactions within the lead glass (e.g. absorption, charge exchange...) [144]. Pions make up the majority of the events on the left of and above the muon peak in Figure 4.11. An approximate selection of electron- and muon-like events based on the charge deposited in the ACTs provides prior values for the parameters of the Gaussian fit.

The cuts applied to the calorimeter charge to tag light particles as electron or non-electron-like depend on the separation between the muon and electron populations (and therefore on the beam momentum):

1. If the electrons and muons peaks are $5\sigma_e + 3\sigma_\mu$ away from each other or more, the cut lines are:

$$\mathrm{LG}^{\text{non-e}}_{\text{max}} = \mu_\mu + 3\sigma_\mu \tag{4.4}$$

   and:

$$\mathrm{LG}^{e}_{\text{min}} = \mu_e - 5\sigma_e. \tag{4.5}$$

   where $\mathrm{LG}^{x}_{\text{min/max}}$ is the minimal/maximal charge that a particle can deposit in the calorimeter and still be tagged as $x = e, \text{non-e}$.

2. If the electron and muon peaks are less than $5\sigma_e + 3\sigma_\mu$ away from each other, the cut line is calculated based on a weighted average between the two populations:

$$\mathrm{LG}^{\text{non-e}}_{\text{max}} = \mathrm{LG}^{\text{e}}_{\text{min}} = \frac{(\mu^{\text{LG}}_{\mu} + \sigma^{\text{LG}}_{\mu}) \times A^{\text{LG}}_{e} + (\mu^{\text{LG}}_{e} - \sigma^{\text{LG}}_{e}) \times A^{\text{LG}}_{\mu}}{A^{\text{LG}}_{e} + A^{\text{LG}}_{\mu}}. \tag{4.6}$$

The cut lines presented in Equations 4.4 to 4.6 maximise the product of the sample's purity and efficiency for events tagged as non-electrons, which are used in most of the WCTE Physics analyses. Figures 4.11 and 4.12 show the fitted lead glass charge distribution and the cuts used for tagging non-electron and electron-like events in runs at 1000 MeV/c and 460 MeV/c, respectively.

The purity and efficiency of the electron-like tag are:

$$p^{\text{e}}_{\text{tag}} = \frac{N^{\text{e true}}_{\text{e tag}}}{N_{\text{e tag}}} = 1 - \frac{N^{\text{non-e true}}_{\text{e tag}}}{N_{\text{e tag}}} \text{ and } \epsilon^{\text{e}}_{\text{tag}} = \frac{N^{\text{e true}}_{\text{e tag}}}{N^{\text{e true}}} \tag{4.7}$$

where $N^{x\text{ true}}_{y\text{ tag}}$ is the total number of true particles of type x tagged as y with $x, y = \{e,\ \text{non-e}\}$, $N_{x\text{ tag}}$ is the total number of events tagged as type x and $N_{\alpha\text{ true}} = \int_{\text{- inf}}^{\text{+inf}} G(A_\alpha, \mu_\alpha, \sigma_\alpha) dx$ is the total number of true $\alpha = \{e,\ \mu\}$ particles.

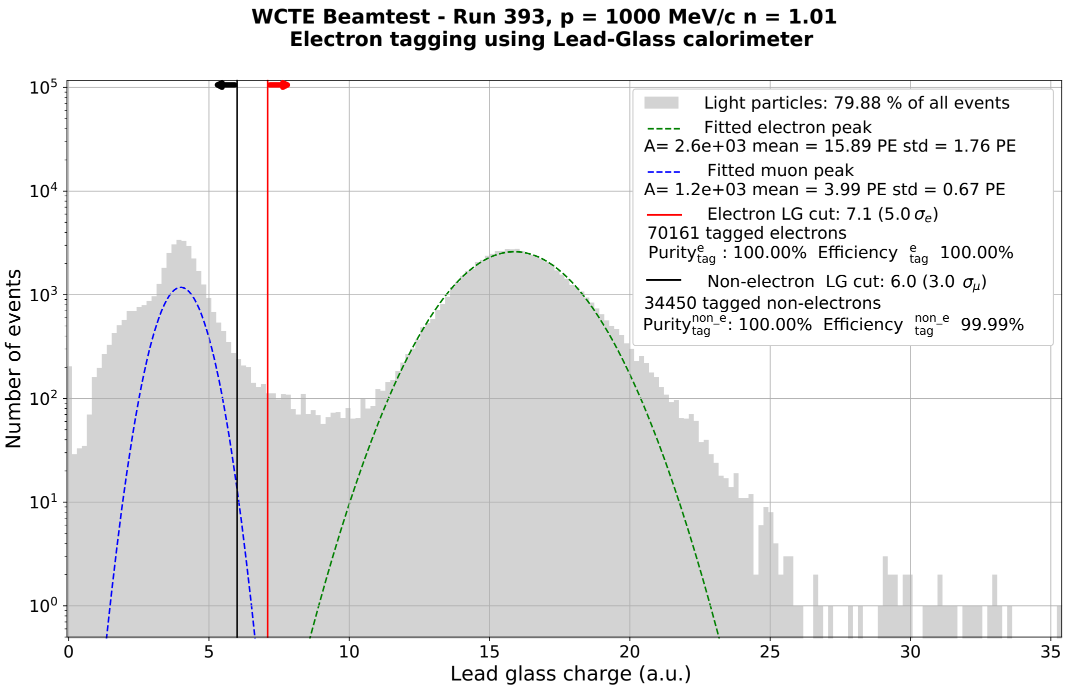


Figure 4.11: Lead glass charge spectrum collected for all light particles in a 1000 MeV/c run, showing the Gaussian distributions fitted to the electron and muon peaks as well as the cut lines applied to tag electron and non-electron-like events.

The contamination of true non-electrons in the electron sample $N^{\text{non-e true}}_{\text{e tag}}$ is estimated by scaling the muon contamination in the electron-like sample ($N^{\mu \text{ true}}_{e \text{ tag}}$) up by a factor $C = N_{\text{non-e tag}}/N^{\mu \text{ true}}$ to account for the pion contribution to the contamination:

$$N^{\text{non-e true}}_{\text{e tag}} = \frac{N_{\text{non-e tag}}}{N^{\mu \text{ true}}} \times N^{\mu \text{ true}}_{\text{e tag}} = \frac{N_{\text{non-e tag}}}{\int_{-\text{ inf}}^{+\text{ inf}} G(A_\mu, \mu_\mu, \sigma_\mu)dx} \times \int_{\text{LG}^{\text{e}}_{\text{min}}}^{+\text{ inf}} G(A_\mu, \mu_\mu, \sigma_\mu)dx. \tag{4.8}$$

Equation 4.8 is a conservative overestimate of the non-electron event contamination in the electron sample. It assumes that pions are just as likely as muons to be mis-tagged as electrons. Pions typically deposit less charge in the lead glass than muons (that is, if they are minimally ionising and don't interact with the nuclei in the calorimeter). They are therefore less likely to contaminate the electron sample.

Similarly, the purity and efficiency of the non-electron-like tag are:

$$p^{\text{non-e}}_{\text{tag}} = \frac{N^{\text{non-e true}}_{\text{non-e tag}}}{N_{\text{non-e tag}}} = 1 - \frac{N^{\text{e true}}_{\text{non-e tag}}}{N_{\text{non-e tag}}} \text{ and } \epsilon^{\text{non-e}}_{\text{tag}} = \frac{N^{\text{non-e true}}_{\text{non-e tag}}}{N^{\text{non-e true}}} = \frac{p^{\text{non-e}}_{\text{tag}} \times N_{\text{non-e tag}}}{N^{\text{non-e true}}_{\text{non-e tag}} + N^{\text{non-e true}}_{\text{e tag}}} \tag{4.9}$$

where we have used $N^{\text{non-e true}} = N^{\text{non-e true}}_{\text{non-e tag}} + N^{\text{non-e true}}_{\text{e tag}}$ which is exactly true when $\text{LG}^{\text{non-e}}_{\text{max}} = \text{LG}^{\text{e}}_{\text{min}}$ and approximately true when $\text{LG}^{\text{non-e}}_{\text{max}} \neq \text{LG}^{\text{e}}_{\text{min}}$ which is the case where the electron and muon populations are more than 3σ away from each other and the purity and efficiency of both selections are almost 100%.

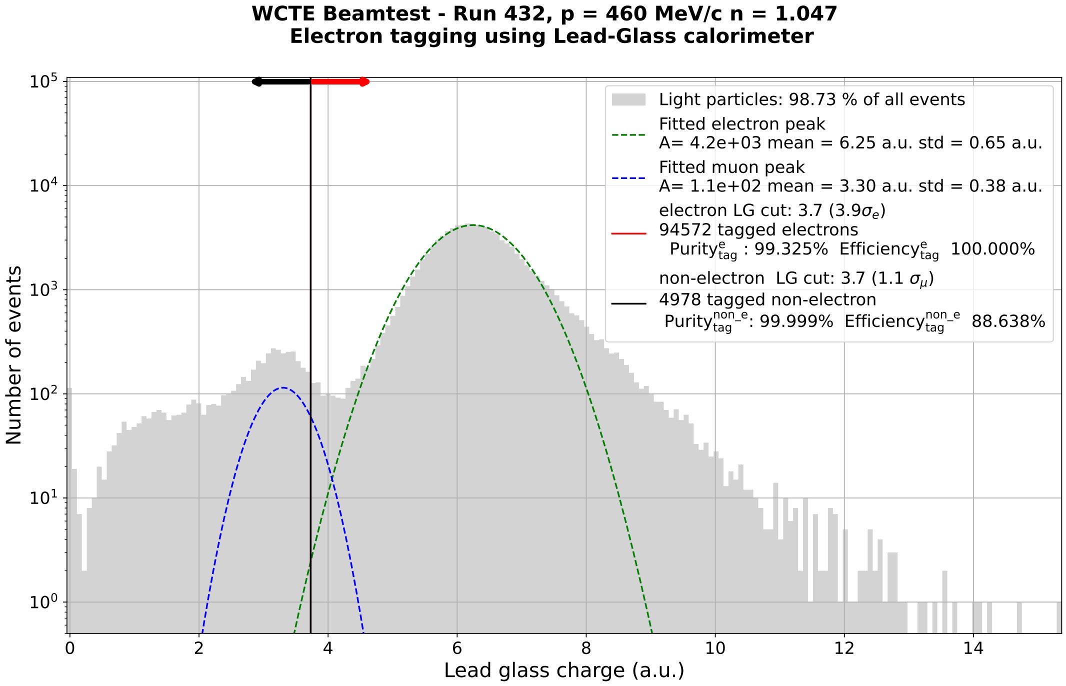


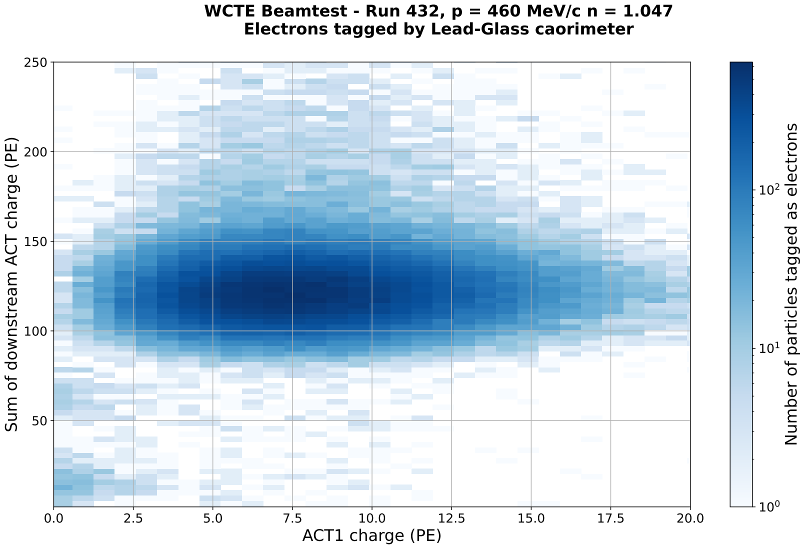
Figure 4.12: Lead glass charge spectrum collected for all light particles in a 460 MeV/c run. Contrary to the high momentum case shown in Figure 4.11, the selections have some inefficiencies and impurities.

4.4.2.2 Electron rejection using the ACTs

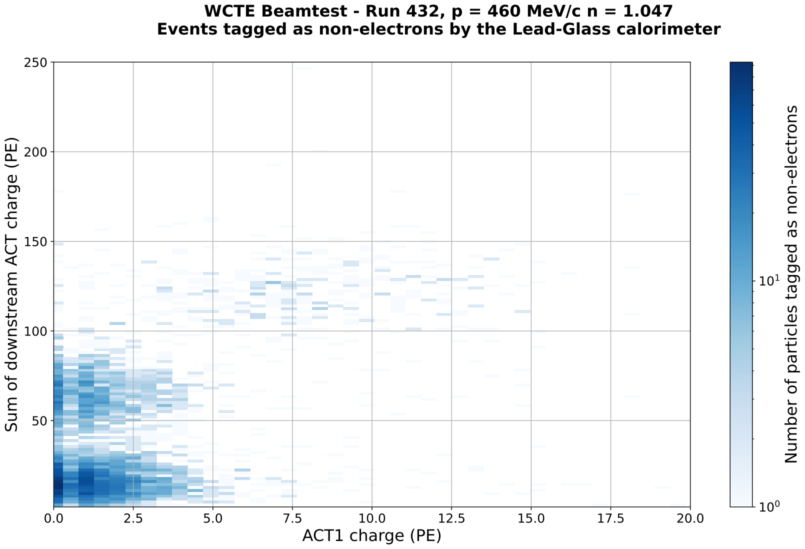
During WCTE operation, the lead-glass calorimeter will not be in the beam line, and we must rely on the other detectors to identify the particles. The 2023 beam test aimed to develop and validate a PID method based exclusively on the ACT and TS detectors.

As discussed in Section 4.2.1, in the sub-GeV regime, only electrons are expected to produce light in all ACT detectors. The charge deposited in ACT1 and the two downstream ACTs (later referred to as ACT23) by particles tagged by the calorimeter as electrons (non-electrons) is shown in Figure 4.13a (Figure 4.13b). The small number of events that are incorrectly tagged by the calorimeter is visible in those figures with the non-electrons producing less than about 75 PE in ACT23 and less than about 5 PE in ACT1, and the electrons producing more charge in both.

Figure 4.14 shows the fraction of light particles that are tagged as electrons by the calorimeter and presents an example of a cut line that can be applied to the ACT distributions to separate the electrons from the non-electrons. In this example of a run at 460 MeV/c, there is clear separation between the electrons and non-electron-dominated regions in the ACTs; the separation worsens at higher and lower momenta, as shown in Figure 4.15. From these we see that the position of the cut line is highly run-dependent and cannot be drawn by hand easily. Instead, we automatically scan over a wide range of cut lines, trying to optimise the electron rejection. The ACT selection identifies any light particle as an electron if it satisfies:



(a) Events tagged as electron-like.



(b) Events tagged as non-electron-like.

Figure 4.13: Charge deposited in the ACTs by particles tagged as electrons (a) and non-electrons (b) for a 460 MeV/c run. The minority of events which has been incorrectly tagged by the calorimeter can be seen by comparing the two figures.

$$C_{\mathrm{ACT23}} > \begin{cases} a \times C_{\mathrm{ACT1}} + b, & \text{if } a \times C_{\mathrm{ACT1}} + b < c \\ c, & \text{otherwise} \end{cases} \tag{4.10}$$

where C_{ACT1} and C_{ACT23} is the charge collected in ACT1 and ACT2 and ACT3 detectors respectively and the a, b, and c parameters are determined by the scan. The 2023 campaign aimed to identify as many of the particles as possible, we therefore identify as non-electrons all light particles that do not satisfy Equation 4.10.

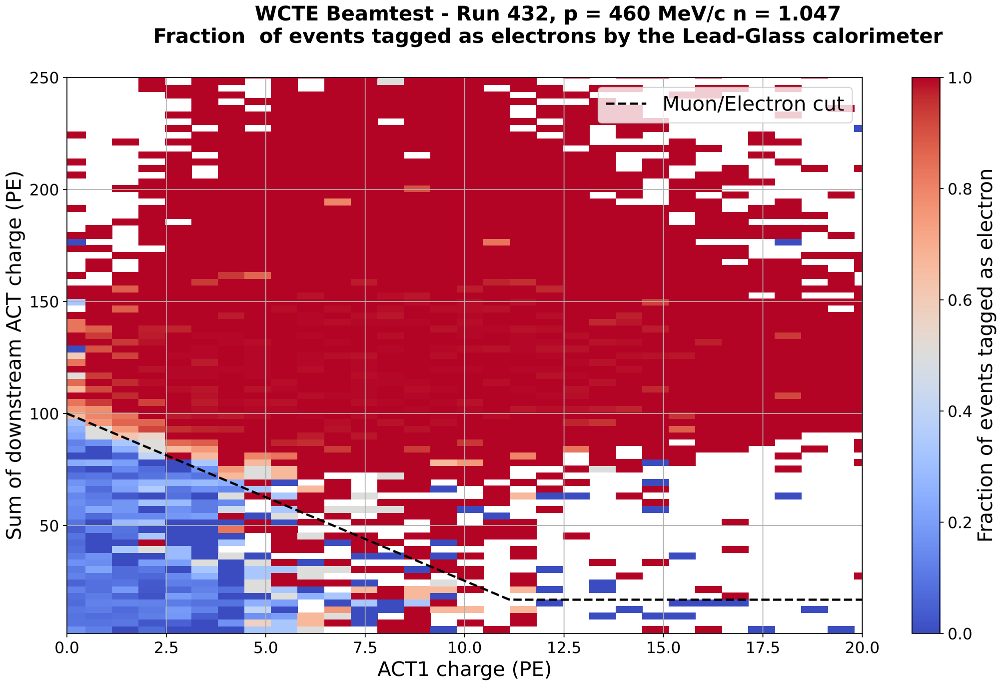


Figure 4.14: Fraction of light particles which are tagged as electrons by the Lead-Glass calorimeter for a run at 460 MeV/c.

The parameter scan maximises the following figure of merit (FOM):

$$\mathrm{FOM} = p_{\mathrm{sel}}^{\text{non-e}} \times p_{\mathrm{sel}}^{\text{non-e}} \times \epsilon_{\mathrm{sel}}^{\text{non-e}} \tag{4.11}$$

where:

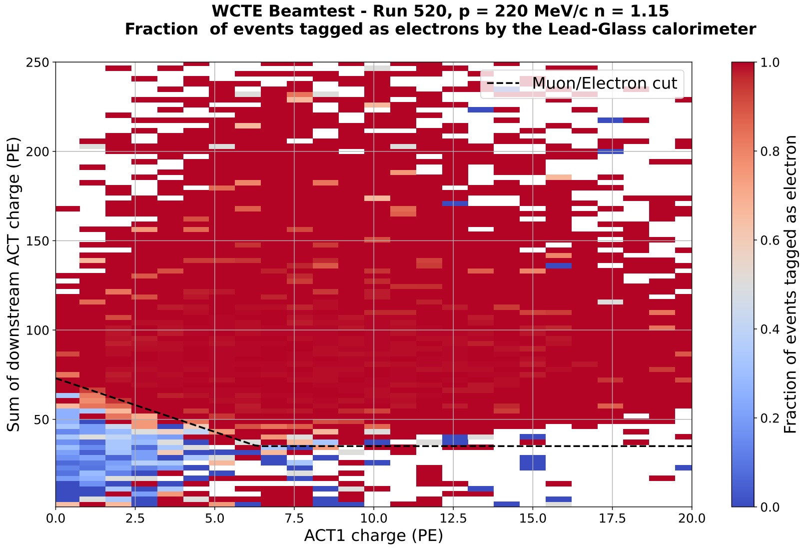
$$p_{\mathrm{sel}}^{\text{non-e}} = \frac{\text{Number of selected non-e events tagged as non-e by the calorimeter}}{\text{Total number of selected non-e events}} \tag{4.12}$$

and

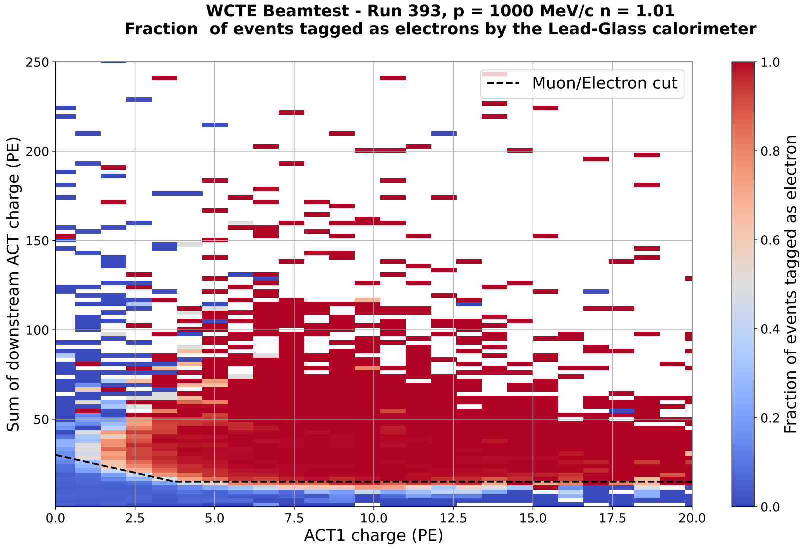
$$\epsilon_{\mathrm{sel}}^{\text{non-e}} = \frac{\text{Number of selected non-e events tagged as non-e by the calorimeter}}{\text{Total number of tagged non-e events}} \tag{4.13}$$

are the purity and efficiency of the non-electron selection based on the calorimeter tag. The FOM presented in Equation 4.11 ensures that the non-electron selection has a very high purity and a relatively high efficiency.

The parameter scan starts with a coarse scan, which tests seven candidate values for each of the three cut parameters a, b and c. These values correspond to $\pm 50\%$, $\pm 33.3\%, \pm 16.6\%$ and $+0\%$ of the prior values manually estimated. This coarse scan returns a', b' and c', the optimal combination of parameters out of the 343 ($7 \times 7 \times 7$) combinations tested. This coarse scan is followed by a finer scan which tests another 343 configurations corresponding to $\pm 20\%, \pm 13.2\%, \pm 6.6\%$ and $+0\%$ of a', b' and c'. We checked the robustness of this method by increasing the number of combination tests and modifying the prior values of a, b and c within reasonable values. We observed no significant difference in the scan results in these cases.



(a) Beam momentum: 220 MeV/c.



(b) Beam momentum: 1000 MeV/c.

Figure 4.15: Fraction of events tagged as electron showing that the separation in the ACTs between the populations is worse both at higher and at lower momenta. The dashed line shows a possible cut line.

Figure 4.16 shows the different configurations tested by the coarse parameter scan for a run at 460 MeV/c, and Figure 4.17 shows the result of the finer scan for the same run. A good way to validate the resulting selection is to look at the charge deposited by each population in the calorimeter. This is shown in Figure 4.18 for the same run. At the time of writing, there is no robust Monte-Carlo simulation of the WCTE beam setup, which would typically be used to validate the purity and efficiency of the selection. Instead, this analysis optimises the selection using the calorimeter tag as a proxy for the Monte-Carlo truth.

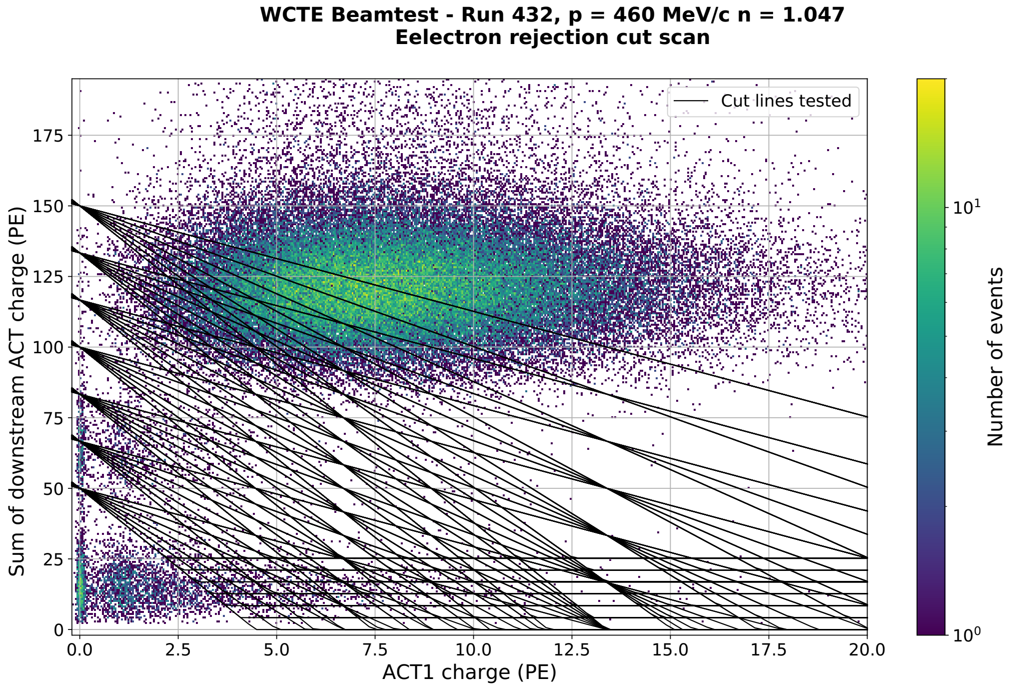


Figure 4.16: Presentation of the 343 ACT selection cuts tested by the coarse parameter scan for a run at 460 MeV/c.

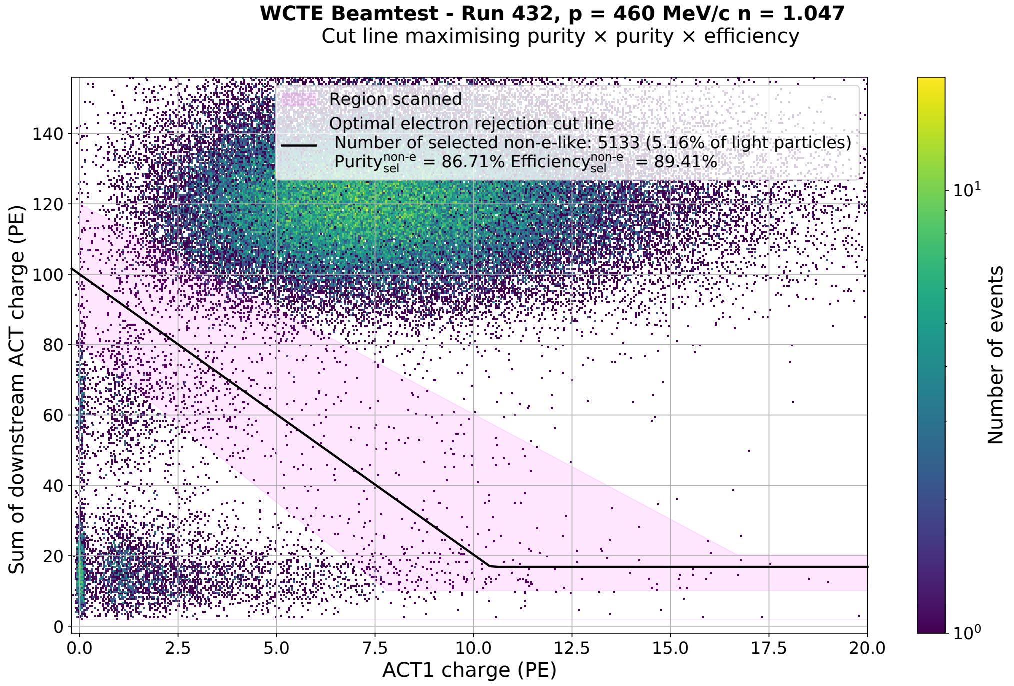


Figure 4.17: Result of the fine parameter scan for a run at 460 MeV/c. The purity and efficiency indicated on the figure are calculated based on the calorimeter tags (as per Equations 4.12 and 4.13).

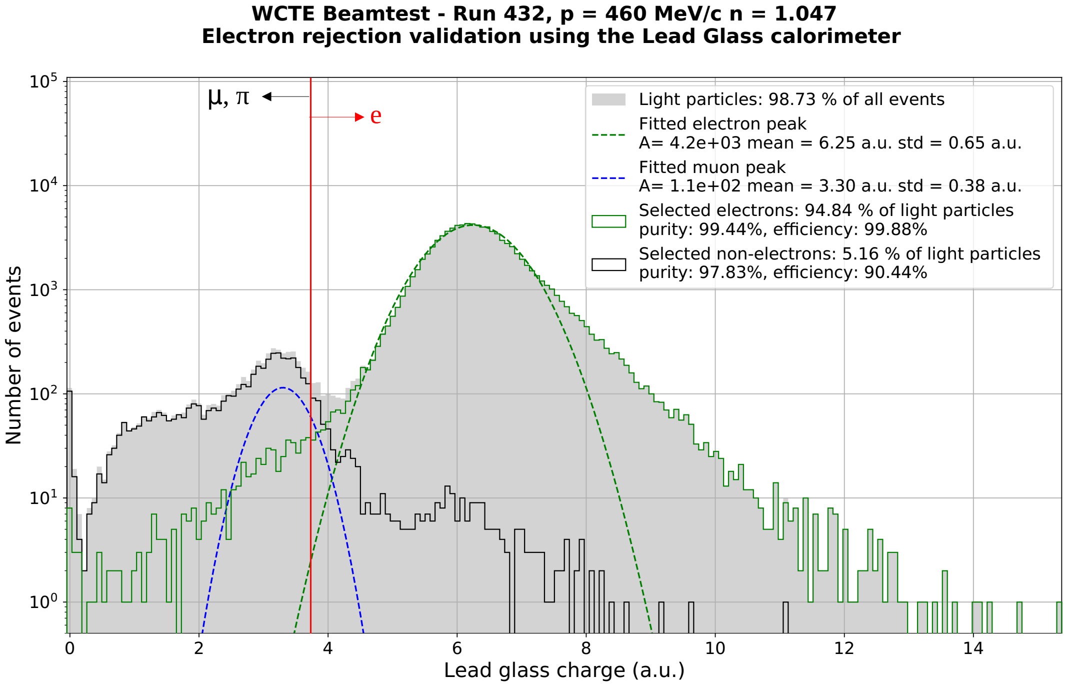


Figure 4.18: Charge deposited in the lead-glass calorimeter by events identified as non-electron (black) and electron (green) particles using the ACT selection for a run at 460 MeV/c. The purity and efficiency of the selection are calculated by assuming that all particles on the right (left) hand side of the lead glass cut line (red line) at 3.7 a.u. are electrons (muons or pions).

4.4.3 Muon/pion separation using the ACTs

After the electrons have been removed using the cuts described in Section 4.4.2.2, the pions and muons can be separated using the charge they deposit in the downstream ACTs. As presented in Section 4.2.1, the refractive index of the ACT2 and ACT3 detectors is chosen such that pions are below the Cherenkov threshold and muons above it.

Figure 4.19 presents the charge deposited in ACT23 by light particles, which have been identified as non-electrons. We see two clear peaks in the distribution corresponding to the pion and muon populations. We fit the charge distribution with a sum of Gaussians: $y_{\mathrm{ACT23}} = G(A_\mu, \mu_\mu, \sigma_\mu) + G(A_\pi, \mu_\pi, \sigma_\pi)$.

The 2023 beam test campaign aims to measure the composition of the T09 beam as precisely as possible. With this in mind, we present here the selection where the product of the pion and muon purities is maximal. The selection cuts can be adapted for each WCTE analysis depending on the purity and efficiency requirements. A summary of the cut lines used for each 2023 beam test sample is presented in Ref [143].

The purity of the muon/pion separation is calculated from the leakage of each of the Gaussians on either side of the cut line, labelled $\mathrm{ACT23}^{\mu/\pi}_{\mathrm{cut}}$. We define the purity of the muon selection as:

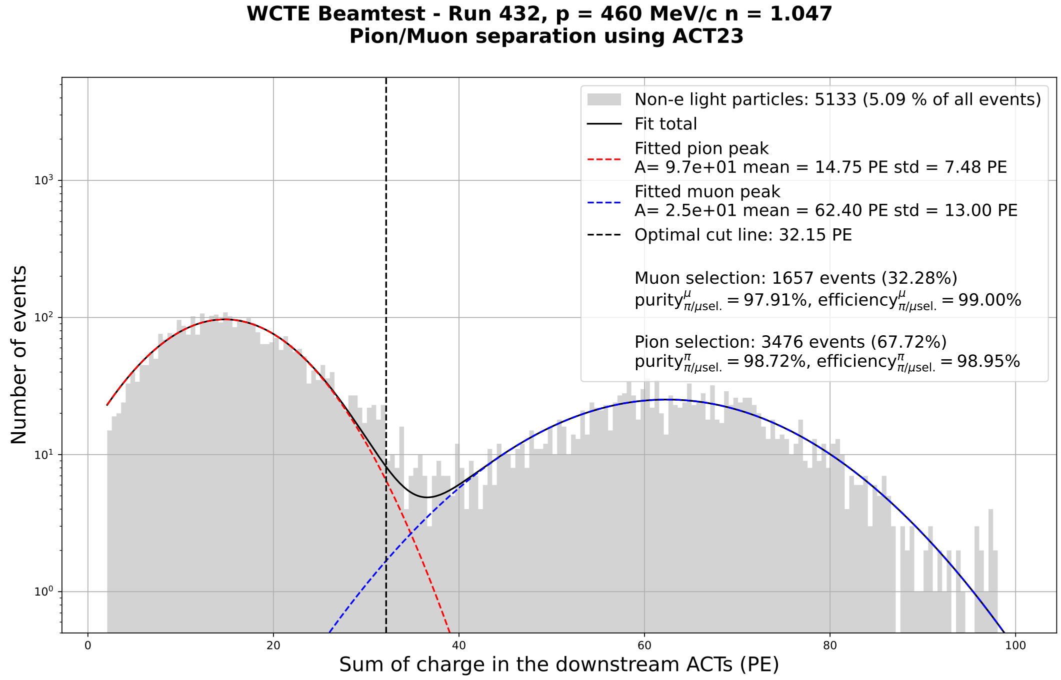


Figure 4.19: Charge deposited in the downstream ACTs by particles in the non-electrons selection showing two populations identified as muon-like (blue) and pion-like (red).

$$p_{\mu\ \text{sel}} = \frac{\int_{\text{ACT23}^{\mu/\pi}_{\text{cut}}}^{+\text{inf}} G(A_\mu, \mu_\mu, \sigma_\mu)dx}{N^{\mu\ \text{sel}}_{\text{non-e sel}}}, \tag{4.14}$$

and its efficiency as:

$$\epsilon_{\mu\ \text{sel}} = \frac{\int_{\text{ACT23}^{\mu/\pi}_{\text{cut}}}^{+\text{inf}} G(A_\mu, \mu_\mu, \sigma_\mu)dx}{\int_0^{+\text{inf}} G(A_\mu, \mu_\mu, \sigma_\mu)dx}. \tag{4.15}$$

The purity and efficiency of the pion selection are calculated in the same way. It can be shown that the position of the cut line that maximises $\text{p}_{\mu\ \text{sel}} \times \text{p}_{\pi\ \text{sel}}$ is:

$$\text{ACT23}^{\mu/\pi}_{\text{cut}} = \frac{\mu_\mu \sigma_\pi^2 + \mu_\pi \sigma_\mu^2}{\sigma_\pi^2 + \sigma_\mu^2} \tag{4.16}$$

4.4.4 Scintillation light

It is interesting to note on Figure 4.19 that events identified as pions produce a significant amount (about 15 PE) of light in the two downstream ACTs despite being nominally below the Cherenkov threshold. Furthermore, we observe in Figure 4.13b that non-electron-like particles still produce light in ACT1 (up to a few PE) despite being well below the Cherenkov threshold in the $n = 1.006$ aerogel. We measured that the light produced in ACT1 by particles below the Cherenkov threshold was delayed by about 20 ns on average compared to the light

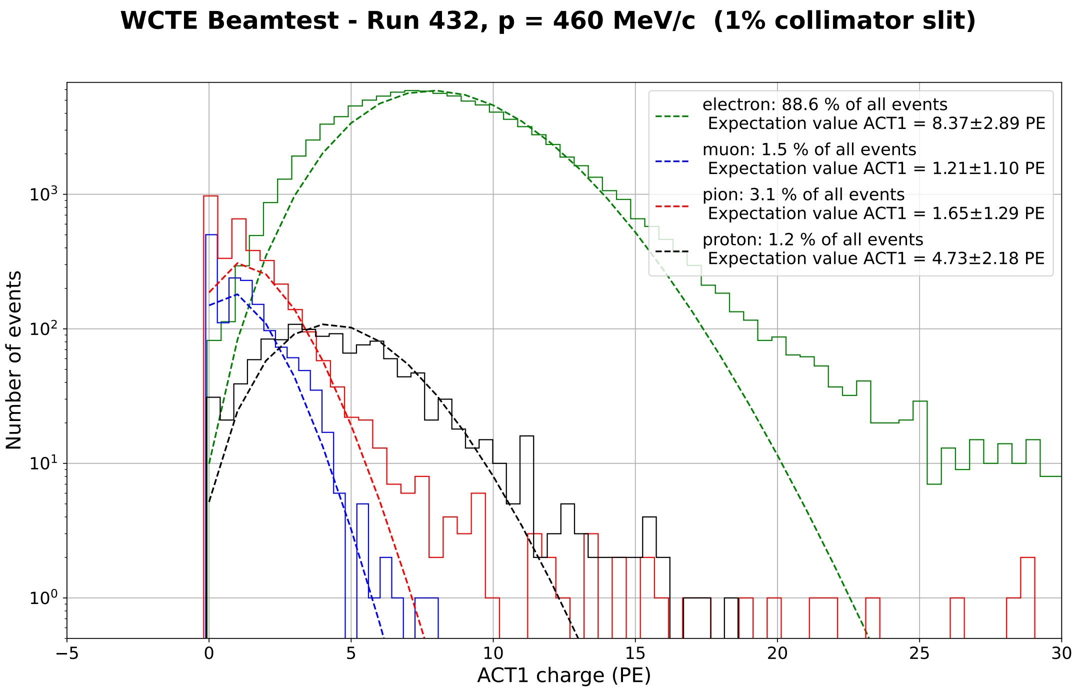


Figure 4.20: Total charge collected in the ACT1 detector for different particle types, fitted with a Poisson distribution. All particles but the electron are below the Cherenkov threshold at this momentum.

emitted by electrons [143]. Figure 4.20 shows the amount of light produced in ACT1 by each particle type (as identified using the TOF and ACT cuts presented previously) for a run at 460 MeV/c.

Figure 4.21 compares the amount of light produced in ACT1 by particles below threshold as a function of the nominal beam momentum. We observed that heavier particles produce more light than MIPs and that the amount of light produced reduces as the momentum increases.

The late arrival of this light, combined with its dependency on particle mass and momentum, lets us suspect that it is scintillation light. The source of this background was later identified as the 3M™reflective film, which is a known scintillator [145]. Following this, we modified the ACT detector design to replace the 3M™film with a non-scintillating Mylar film.

4.5 Momentum measurement

Once all particles have been identified, we obtain their momentum from their TOF, measured by the trigger scintillators.

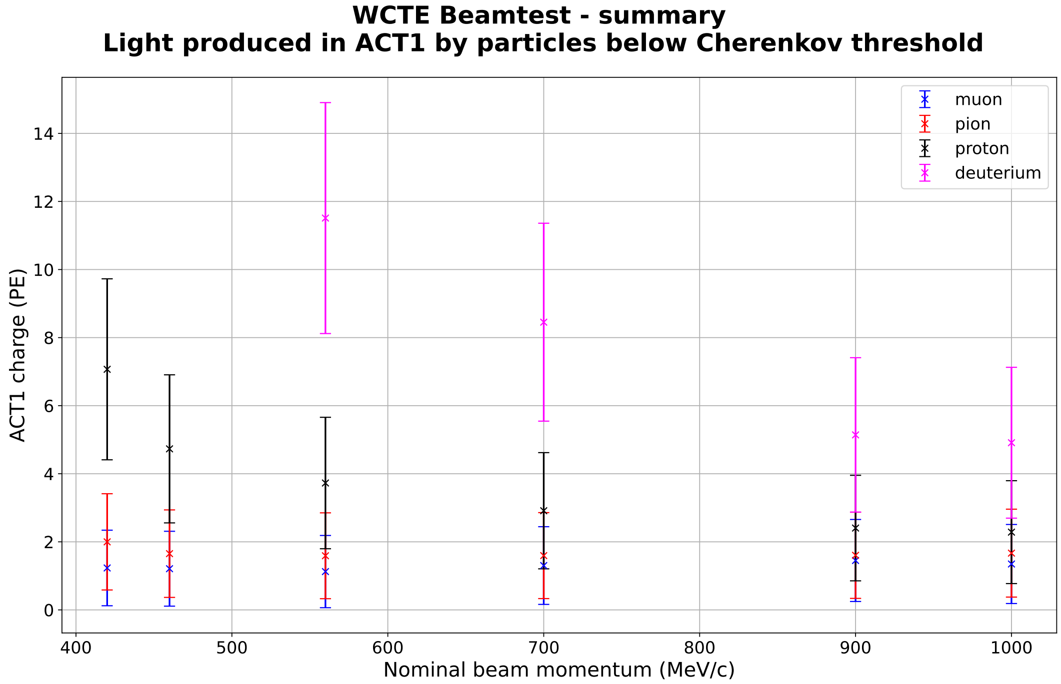


Figure 4.21: Summary of the charge collected in ACT1 ($n = 1.006$) for different particles below Cherenkov threshold, showing that heavier and slower particles, which deposit more energy, produce more light, as expected from scintillation light.

4.5.1 Momentum estimate from the time of flight

After the particles have been identified as each particle type (x = D, p, μ, π) using the methods presented above, their time of flight is used to measure the mean momentum for each particle population. The TOF distribution is fitted with a Gaussian distribution of mean τ_x and standard deviation σ_x^{TOF}. The mean particle momentum is calculated as follows:

$$\mathrm{p}_x = \gamma v_x m_x = \frac{v_x m_x}{\sqrt{1 - \left(\frac{v_x}{c}\right)^2}} \tag{4.17}$$

where m_x is the mass of particle x and v_x is its mean velocity:

$$v_x = \frac{L}{\tau_x + \delta_e} = \frac{L}{\tau_x + \tau_e - \frac{L}{c}} \tag{4.18}$$

where $L = 3.49 \pm 0.01$ m is the distance between the two trigger scintillators. The correction factor $\delta_e = \tau_e - \frac{L}{c}$ is applied to the measured mean TOF to remove any remaining timing offsets between the two trigger scintillators, as these would cause a systematic bias in the estimate of the particle's momentum. Only the mean momentum of the given particle type (as opposed to individual particle momenta) is calculated using Equation 4.17 because of the non-zero width of the TOF distributions (caused by finite detector resolution), which causes some of the particles to have a recorded TOF shorter than L/c. This is not the case for the mean value of the TOF for any of the particle distributions in the datasets collected.

4.5.2 Energy loss correction and error estimation

Particles lose energy as they travel through matter. This loss is described by the Bethe-Bloch equation with an accuracy better than 1% in the sub-GeV region [146]. We use a Geant4 simulation to compute reference tables of the energy lost by particles along the beam line and add the momentum lost upstream of TS1 to the estimate of the momentum calculated with Equation 4.17. Table 4.4 summarises the properties of the materials upstream of TS1 in the beam line, the density is obtained from the Geant4 database [147]. Most of the energy is lost in TS0 whose thickness is known with a ±7.9% uncertainty. We attribute a conservative uncertainty of $\sigma_{\delta_p} = \pm 15\%$ [143] to the estimate of the total energy lost upstream of TS1 (δ_{p_x}) by particle x which also accounts for the approximation $\mathrm{p}_x \approx \mathrm{p}_{\mathrm{nominal}}$ that we make when estimating the energy lost.

Name	Material	Thickness (cm)	Density (g cm^{-3})
Beam pipe window	Mylar	0.025 ± 0.001	1.39
TS0	Plastic scintillator	0.635 ± 0.05	1.032
Air	Air	349 ± 1	1.204

Table 4.4: Summary of the material budget on the path of the particles before they reach TS0.

The momentum of a particle of type x right before it exits the CERN beam pipe is therefore:

$$\mathrm{p}_x^{\mathrm{initial}} = \mathrm{p}_x^{\mathrm{at\ TS1}} + \delta_{\mathrm{p}_x} = \frac{L m_x}{(\tau_x + \tau_e - L/c) \times \sqrt{1 - \left(\frac{L}{\tau_x + \tau_e - L/c}\right)^2}} + \delta_{\mathrm{p}_x}. \tag{4.19}$$

The errors on each of the parameters involved in this equation are summarised in Table 4.5. A derivation of the total error on $\mathrm{p}_x^{\mathrm{initial}}$ is given in Appendix A.

Variable	Error	Comment
Particle x mean time of flight (τ_x)	$\sigma_x = \frac{\sigma_x^{\mathrm{TOF}}}{\sqrt{n_x}}$	$\sigma_x^{\mathrm{TOF}} = 0.24 - 0.70$ ns
Electron mean time of flight (τ_e)	$\sigma_e = \frac{\sigma_e^{\mathrm{TOF}}}{\sqrt{n_e}}$	$\sigma_e^{\mathrm{TOF}} = 0.24 - 0.30$ ns
Distance between TS0 and TS1 (L)	±1 cm	Measured
Momentum upstream of TS1 ($\delta_{\mathrm{p}_x}^{\mathrm{p\ lost}}$)	±15%	See Section 4.5.2.

Table 4.5: Summary of the error on each of the parameters used to measure the initial mean momentum of a given particle population. The leading sources of error are τ_x and τ_e.

Figure 4.22 shows the time of flight measured for each particle type for a run with a nominal beam momentum of 460 ± 5 MeV/c. The momentum estimate and its associated error are calculated as described in Equation 4.17 and in Appendix A. The momentum measured for the muons, pions and protons are respectively $\mathrm{p}_\mu = 426 \pm 41$ MeV/c, $\mathrm{p}_\pi = 457 \pm 30$ MeV/c and $\mathrm{p}_\mathrm{p} = 460 \pm 3$ MeV/c where the quoted uncertainty is the error on the mean. These are all compatible with the nominal beam momentum at the 1σ level. The error on the muon and pion

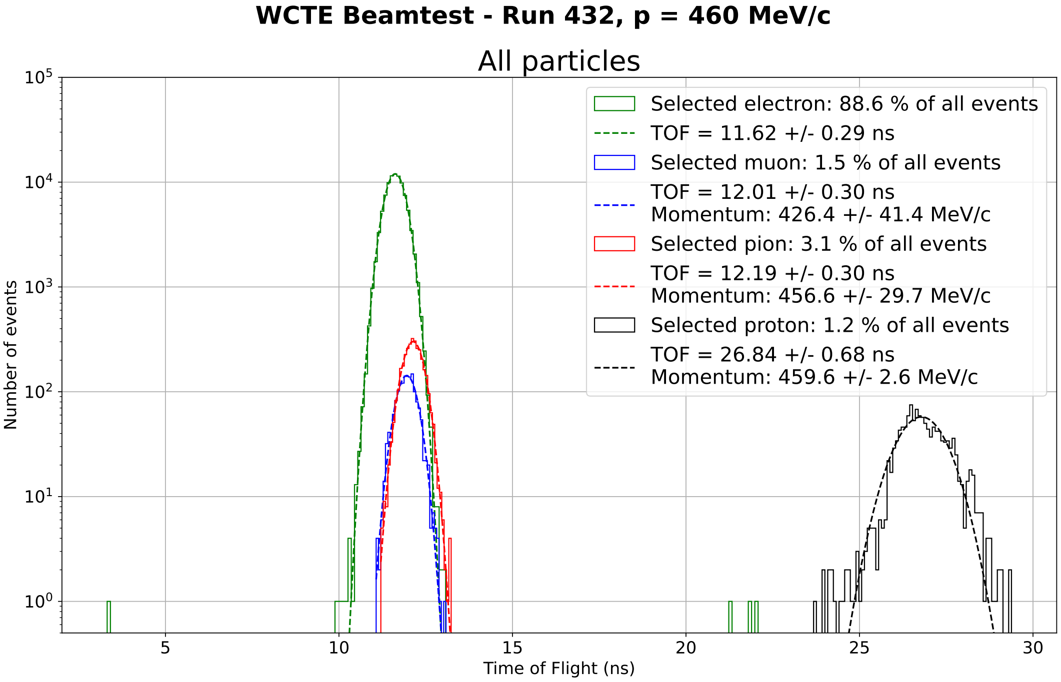


Figure 4.22: Distribution of the time of flight for events identified as electrons, muons, pions and protons for a run at 460MeV/c, indicating in legend the momentum measured for each selection.

momentum is significantly larger than on the proton momentum due to the much higher speed of these particles and contaminations from electrons and pion or muons (as discussed in Section 4.4.3) which shifts the mean value of TOF away from the true muon TOF.

4.5.3 Momentum measurement using the Lead-Glass calorimeter

The charge deposited by electrons in an electromagnetic calorimeter is proportional to their energy. The Lead-Glass calorimeter is more than 15 radiation lengths long, which means that electrons deposit effectively all of their energy in it. Using this and the known energy lost by electrons upstream of the calorimeter, we can convert the calorimeter charge to units of momentum by looking at the electron sample. This conversion is required for the tagged photon set-up described in Section 3.4, where the calorimeter is used to validate the measurement of the photon energy made using the hodoscope.

Figure 4.23 shows the mean charge collected in the calorimeter for the electron sample as a function of the nominal electron momentum $\mathrm{p}^{e}_{\mathrm{nominal}} = \mathrm{p}^{\mathrm{initial}}_{\mathrm{nominal}} - \mathrm{p}^{e}_{\mathrm{lost}}$ where $\mathrm{p}^{e}_{\mathrm{lost}}$ is the total momentum lost by the electron as it travels through the set-up (see Section 4.5.2 for more details).

We observe a linear relationship between the charge collected in the lead glass and the nominal electron momentum, both in negative and positive polarities. The linear lines of best fits are compatible with each other, and both y-intercepts are compatible with zero at the 1σ level. The calibration of the lead glass charge can be

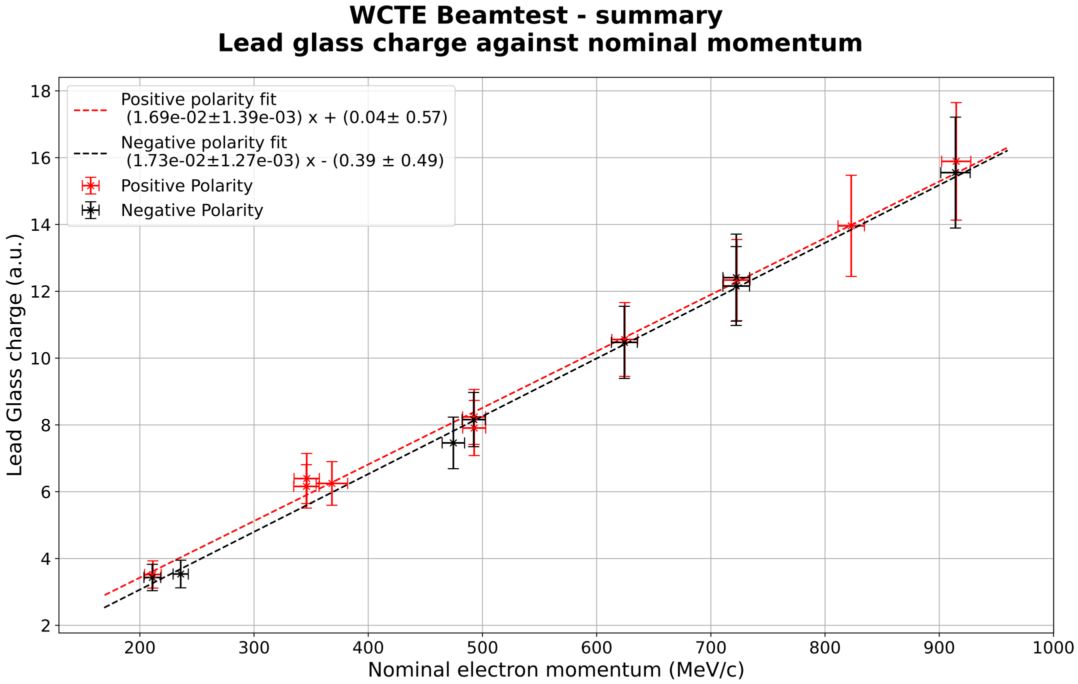


Figure 4.23: Charge collected in the lead glass calorimeter for electrons (black) and positrons (red) as a function of the nominal electron momentum at the lead glass. The error bar indicated the width of the distribution of charge deposited in the lead glass (which increases with electron momentum), not the error on the mean momentum.

done against the momentum calculated using the particles' TOF instead of the nominal one, using the same method.

4.5.4 Measurement of the T09 beam momentum

We measure the momentum of particles in the T09 beam using the method presented in Section 4.5.1 and compare it with the nominal beam momentum in Figures 4.24 and 4.25 for the Be+W and Al secondary beam production targets, respectively. All runs with momenta less than 450 MeV/c were taken with a momentum spread of 3%, whilst runs at higher momenta had a 1% momentum spread.

From these figures, we see that the nominal beam momentum is in good agreement with the measured one for all particles except muons at high momenta. There is good agreement between the momentum of particles and of their anti-particles, as well as between the two targets. Using the more reliable proton momentum as a proxy for the beam momentum, we observe that the line of best fit $p_{\text{protons, Be+W target}} = (1.06 \pm 0.01) p_{\text{nominal}} - (36.1 \pm 5.7)$ MeV/c for the Be+W target and $p_{\text{protons, Al target}} = (1.07 \pm 0.01) p_{\text{nominal}} - (40.32 \pm 8.0)$ MeV/c for the Aluminium target diverge significantly from the $y = x$ line. If confirmed, this would mean that the beam has a slight momentum bias with higher (lower) values than nominal at higher (lower) momenta.

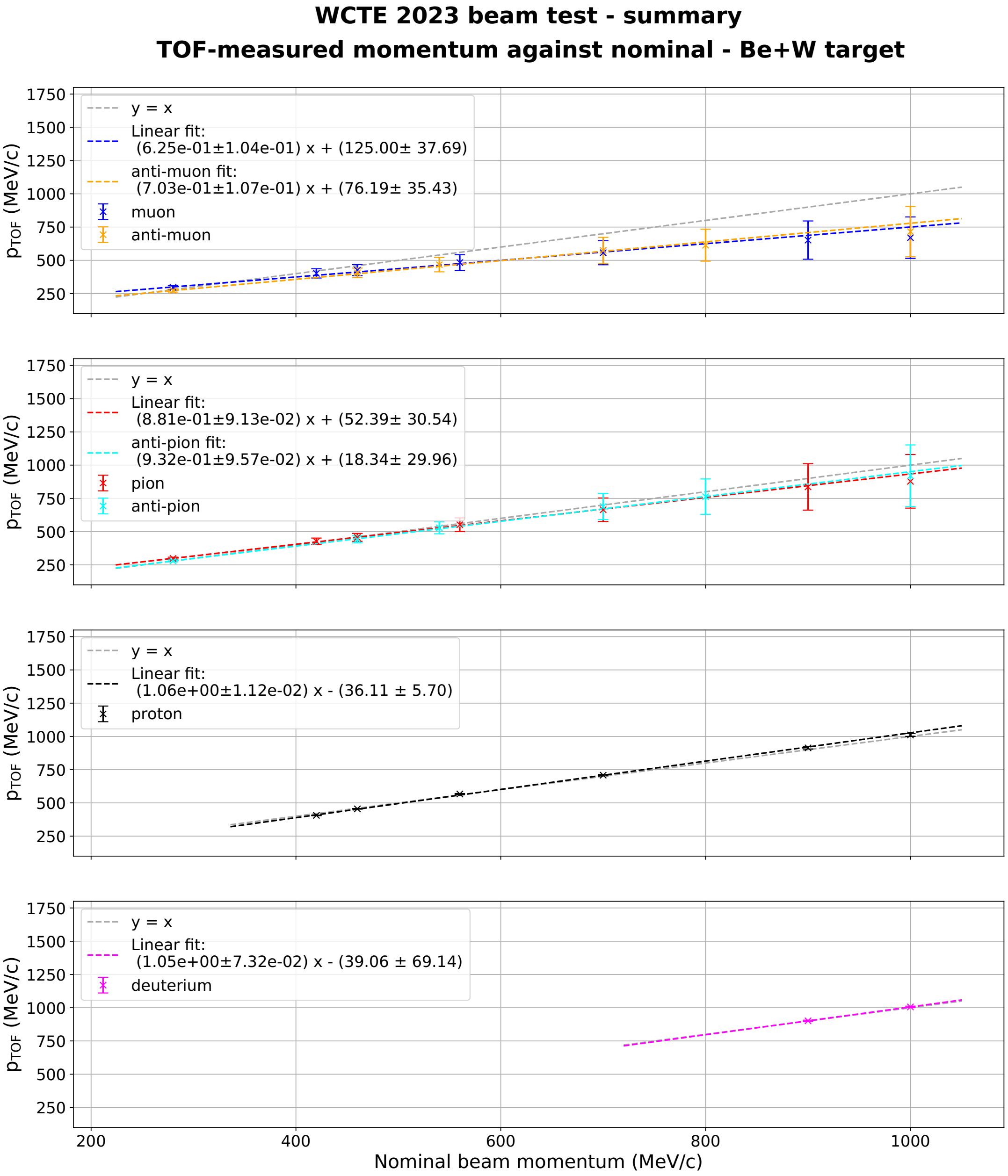


Figure 4.24: Momentum measured using the time of flight of different particles produced with the Be+W target. The error corresponds to the error on the mean calculated as described in Section 4.5.2 and Appendix A, which is larger as β approaches one.

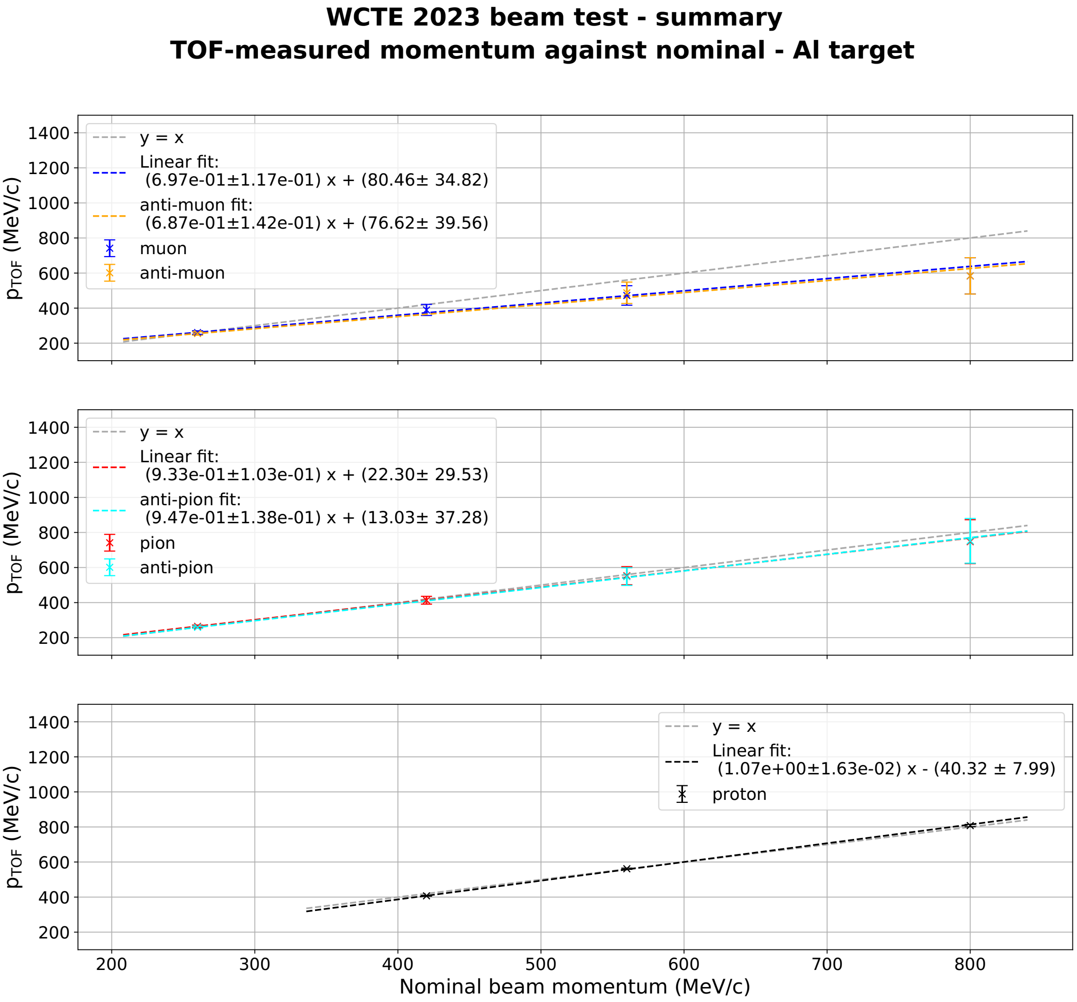


Figure 4.25: Momentum measured using the time of flight of different particles produced with the Al target. Only one dataset made with the Al target corresponding to $p_{\text{nominal}} = (800 \pm 1\%)$ MeV/c contained deuteron events. The mean deuterium momentum was measured to be $p_D = 784.1 \pm 4.0$ MeV/c.

4.6 Composition of the CERN T09 beam

4.6.1 Unfolding to true particle type

4.6.1.1 Protons and deuterium

The time of flight of protons and deuterium nuclei is more than $3\sigma^{\mathrm{TOF}}$ away from each other's and from those of lighter particles over the 200-1200 MeV/c momentum range, as shown in Figure 4.9. We therefore assume that the purity and efficiency of the proton and deuteron selections are 100%.

4.6.1.2 Electrons

For runs with momenta less than about 500 MeV/c, the purity p_{tag} and the efficiency ϵ_{tag} of the electron tag made with the calorimeter are less than 100%. To obtain the true number of electron and non-electron particles in each sample made using the ACT selection described in Section 4.4.2.2, we need to account for the impurity and inefficiency of the electron tag.

The number of events in selection x that are also tagged as type x by the calorimeter ($N_{\mathrm{sel}\ x}^{\mathrm{tag}\ x}$) is related to the number of true particles x in that selection ($N_{\mathrm{sel}\ x}^{\mathrm{true}\ x}$) by the tagging response matrix $\mathbf{R}_{\mathrm{tag}}$:

$$\begin{bmatrix} N_{\mathrm{sel}\ x}^{\mathrm{tag}\ x} \\ N_{\mathrm{sel}\ x}^{\mathrm{tag}\ y} \end{bmatrix} = \mathbf{R}_{\mathrm{tag}} \begin{bmatrix} N_{\mathrm{sel}\ x}^{\mathrm{true}\ x} \\ N_{\mathrm{sel}\ x}^{\mathrm{true}\ y} \end{bmatrix} = \begin{bmatrix} \epsilon_{\mathrm{tag}} & 1-p_{\mathrm{tag}} \\ 1-\epsilon_{\mathrm{tag}} & p_{\mathrm{tag}} \end{bmatrix} \begin{bmatrix} N_{\mathrm{sel}\ x}^{\mathrm{true}\ x} \\ N_{\mathrm{sel}\ x}^{\mathrm{true}\ y} \end{bmatrix}, \tag{4.20}$$

where $x,\ y = \{e,\ \text{non-e}\}$. Inverting the response matrix in Equation 4.20 gives the total number of true and background events in each selection:

$$\begin{bmatrix} N_{\mathrm{sel}\ x}^{\mathrm{true}\ x} \\ N_{\mathrm{sel}\ x}^{\mathrm{true}\ y} \end{bmatrix} = \mathbf{R}_{\mathrm{tag}}^{-1} \begin{bmatrix} N_{\mathrm{sel}\ x}^{\mathrm{tag}\ x} \\ N_{\mathrm{sel}\ x}^{\mathrm{tag}\ y} \end{bmatrix} \tag{4.21}$$

From there, the true purity and efficiency of the non-electron selection made with the ACTs are:

$$p^{\text{non-e}} = \frac{N_{\text{sel non-e}}^{\text{true non-e}}}{N_{\text{sel non-e}}} = \frac{N_{\text{sel non-e}}^{\text{true non-e}}}{N_{\text{sel non-e}}^{\text{true } e} + N_{\text{sel non-e}}^{\text{true non-e}}} \tag{4.22}$$

and

$$\epsilon^{\text{non-e}} = \frac{N_{\text{sel non-e}}^{\text{true non-e}}}{N^{\text{true non-e}}} = \frac{N_{\text{sel non-e}}^{\text{true non-e}}}{N_{\text{sel non-e}}^{\text{true non-e}} + N_{\text{sel } e}^{\text{true non-e}}} \tag{4.23}$$

the purity p^{e} and efficiency ϵ^{e} of the electron selection are calculated in the same way.

Run number	Momentum	$N_{\text{selected}}^{\text{light \& non-e}}/N^{\text{total}}$	Purity	Efficiency
520	220 MeV/c	0.71%	86.28%	44.68%
480	420 MeV/c	2.97%	87.1%	78.8%
432	460 MeV/c	5.10%	97.8%	90.5%
412	560 MeV/c	9.27%	96.0%	89.6%
438	700 MeV/c	15.4%	94.0%	92.7%
393	1000 MeV/c	26.8%	90.9%	93.1%

Table 4.6: Summary of the performance of the electron rejection made using the ACTs calculated based on the calorimenter tags for a few runs. The beam polarity is positive in all the runs presented, and the Be+W radiator target is used. N^{total} is the total number of triggers identified as single particles. The uncertainty on the purity and efficiency is estimated to be between 2 and 5% depending on runs.

Table 4.6 gives the total purity and efficiency of the non-electron selection for a few representative runs showing that the ACTs' capability to reject electrons worsens at lower momenta which is due both to hardware limitations and a low number of muons and pions in the beam compared to the number of electrons.

4.6.1.3 Muons and pions

The total number of electrons in the non-electron sample is given by Equation 4.21. To measure the number of true electrons in the muon sample ($N_{\mu\text{ sel}}^{\text{e true}}$) selected with the ACTs we need to estimate what fraction $f_{\mu\text{ sel}}^{\text{e true}}$ of the electron contamination falls into the muon selection:

$$f_{\mu\text{ sel}}^{\text{e true}} = \frac{N_{\mu\text{ sel}}^{\text{e true}}}{N_{\text{non-e sel}}^{\text{e true}}} \tag{4.24}$$

To estimate this fraction, we use the sample of events selected as non-electrons using the ACTs but tagged as electrons by the lead-glass calorimeter ($N_{\text{e tag, non-e sel}}$). Some of these events will be true electrons, correctly tagged as such ($N_{\text{e tag, non-e sel}}^{\text{e true}}$) and the rest will be true non-electrons mis-tagged as electrons ($N_{\text{e tag, non-e sel}}^{\text{non-e true}}$):

$$N_{\text{e tag, non-e sel}} = N_{\text{e tag, non-e sel}}^{\text{e true}} + N_{\text{e tag, non-e sel}}^{\text{non-e true}}. \tag{4.25}$$

The number of true electrons correctly tagged as electrons by the calorimeter and selected as non-electrons by the ACTs is $N_{\text{e tag, non-e sel}}^{\text{e true}}$:

$$N_{\text{e tag, non-e sel}}^{\text{e true}} = \epsilon_{\text{e tag}} \times N_{\text{non-e sel}}^{\text{e true}}. \tag{4.26}$$

where $\epsilon_{\text{tag}}^{\text{e}}$ is the efficiency of the electron tagging, defined in Section 4.4.2.1 above. We measure in the data the fraction of events selected by the ACTs as non-electrons which are also identified as muons by the ACTs:

$$f_{\mu\text{ sel}}^{\text{non-e sel}} = \frac{N_{\mu\text{ sel}}^{\text{non-e sel}}}{N^{\text{non-e sel}}} \tag{4.27}$$

and assume that it is equal to the fraction of all true non-electron events that we would select as muons: $f^{\text{non-e sel}}_{\mu\text{ sel}} = f^{\text{non-e true}}_{\mu\text{ sel}}$. This approximation is acceptable because of the high purity of the selection. As an example, we see on Figure 4.19 that $f^{\text{non-e sel}}_{\mu\text{ sel}} = 1 - f^{\text{non-e sel}}_{\pi\text{ sel}} = 32.28\%$ for a run at 460 MeV/c.

We also measure the fraction of events that are tagged as electrons by the calorimeter and selected as muons by the ACTs $f^{\text{e tag, non-e sel}}_{\mu\text{ sel}}$:

$$f^{\text{e tag, non-e sel}}_{\mu\text{ sel}} = \frac{N^{\text{e tag, non-e sel}}_{\mu\text{ sel}}}{N^{\text{e tag, non-e sel}}} \tag{4.28}$$

Combining Equations 4.25-4.28 and rearranging for the fraction of true electron contamination identified as muons $f^{\text{e true}}_{\mu\text{ sel}}$ gives:

$$f^{\text{e true}}_{\mu\text{ sel}} = \left(f^{\text{e tag, non-e sel}}_{\mu\text{ sel}} - f^{\text{non-e sel}}_{\mu\text{ sel}}\right)\frac{N_{\text{e tag, non-e sel}}}{N^{\text{e true}}_{\text{e tag, non-e sel}}} + f^{\text{non-e sel}}_{\mu\text{ sel}} \tag{4.29}$$

Using the same method, we calculate the fraction of events selected as electrons which are, in fact, real muons:

$$f^{\mu\text{ true}}_{\text{e sel}} = \left(f^{\text{non-e tag, e sel}}_{\mu\text{ sel}} - f^{\text{e sel}}_{\mu\text{ sel}}\right)\frac{N_{\text{non-e tag, e sel}}}{N^{\text{non-e true}}_{\text{non-e tag, e sel}}} + f^{\text{e sel}}_{\mu\text{ sel}} \tag{4.30}$$

where $N^{\text{non-e true}}_{\text{non-e tag, e sel}} = \epsilon^{\text{non-e}}_{\text{tag}} \times N^{\text{non-e true}}_{\text{e sel}}$ is the number of true non-electron events that are tagged by the calorimeter as non-electron and selected by the ACTs as electrons. The fraction of true electron contamination identified as pions is $f^{\text{e true}}_{\pi\text{ sel}} = 1 - f^{\text{e true}}_{\mu\text{ sel}}$ since the muon and pion selections are complementary.

We need to know the number of true electrons misidentified as pions and muons to measure the purity of the pion and muon selections and the number of pions and muons misidentified as electrons to calculate the efficiency of these selections. Figure 4.26 shows the charge collected in ACT23 for events selected as non-electrons and tagged as electrons. The legend indicates the fraction of those events selected as muons ($f^{\text{e tag, non-e sel}}_{\mu\text{ sel}} = 39.74\%$) and pions ($f^{\text{e tag, non-e sel}}_{\pi\text{ sel}} = 60.26\%$). Similarly, Figure 4.27 shows the charge collected in ACT23 for events selected as electrons and tagged as non-electrons, used to measure $f^{\text{non-e tag, e sel}}_{\mu\text{ sel}} = 97.72\%$ and $f^{\text{non-e tag, e sel}}_{\pi\text{ sel}} = 2.28\%$.

We calculate the number of true non-electron events selected as pions and muons using:

$$\begin{bmatrix} N^{\text{non-e true}}_{\mu\text{ sel}} \\ N^{\text{non-e true}}_{\pi\text{ sel}} \end{bmatrix} = \begin{bmatrix} f^{\text{non-e sel}}_{\mu\text{ sel}} \\ f^{\text{non-e sel}}_{\pi\text{ sel}} \end{bmatrix} N^{\text{non-e true}}_{\text{non-e sel}}. \tag{4.31}$$

where we assume again that the fraction of selected non-e events identified as muons $f^{\text{non-e sel}}_{\mu\text{ sel}}$ is equal to the fraction of true non-electron events that would be identified as muons using the same selection.

From there, using the purity and efficiency of the pion/muon selection in the ACTs given by Equations 4.14-4.15,

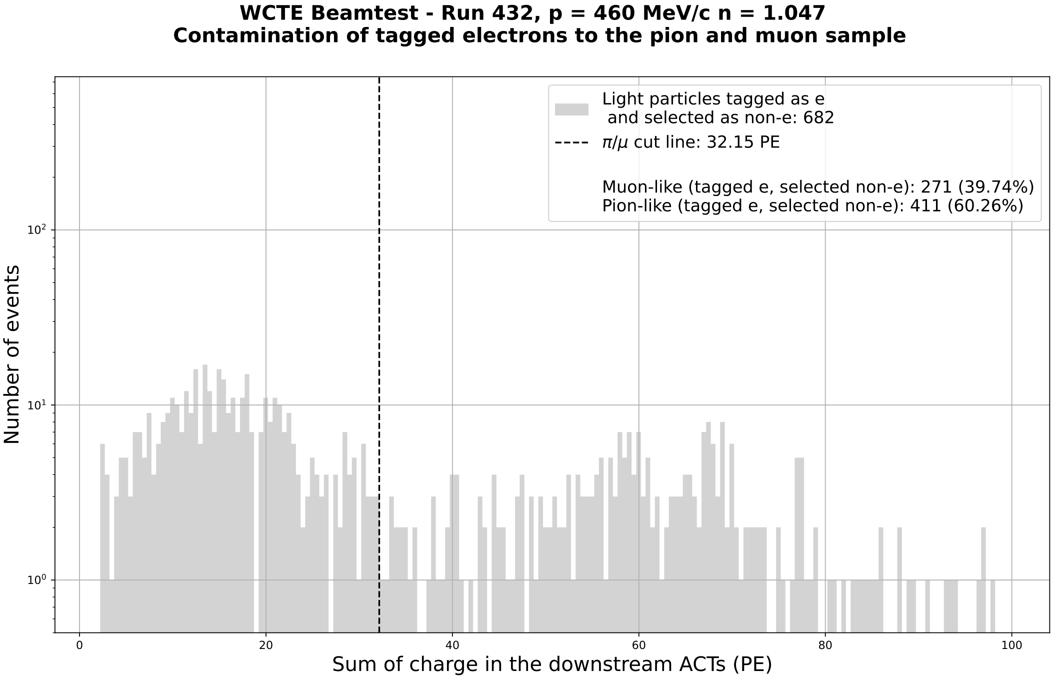


Figure 4.26: Charge deposited in the downstream ACTs by events tagged as electrons by the lead glass calorimeter and selected as non-electrons by the ACTs.

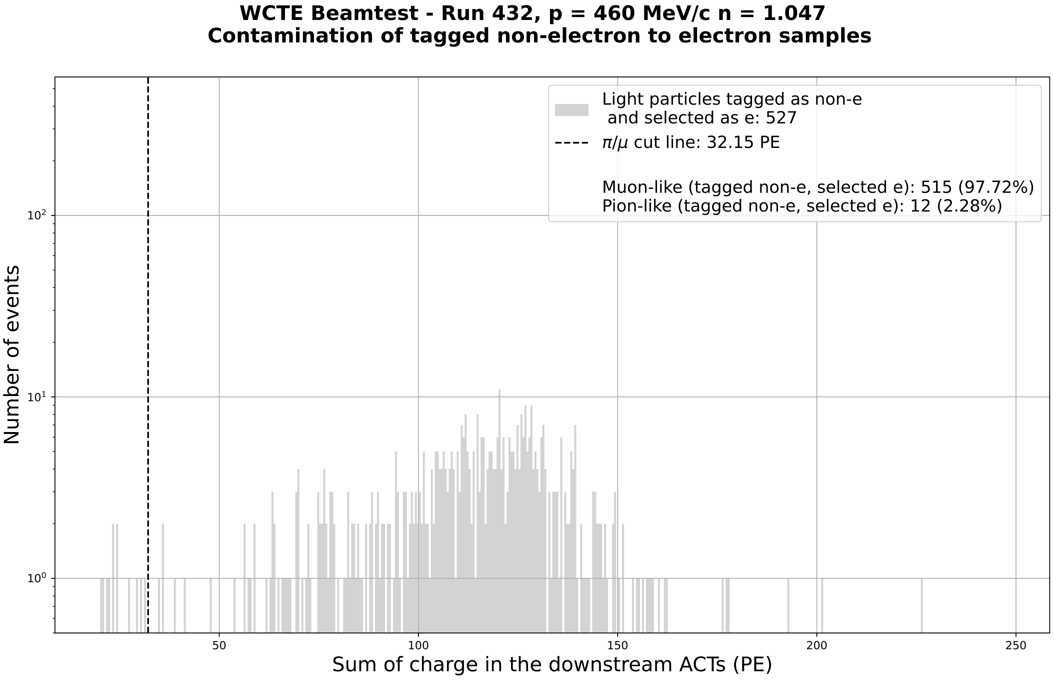


Figure 4.27: Charge deposited in the downstream ACTs by events tagged as non-electrons by the calorimeter and selected as electrons by the ACTs.

	e selection	μ selection	π selection
Fraction of true e	99.4%	5.0%	0.72%
Fraction of true μ	0.54%	90.9%	0.61%
Fraction of true π	0.01%	4.07%	98.7%

Table 4.7: Fraction of true particles in each selection for a run at 460 MeV/c. The uncertainty on the purity of the electron sample is less than 0.5%. The uncertainty on the purity of the muon and pions samples is about 3%.

we calculate the number of true muons and pions selected as muons and pions using:

$$\begin{bmatrix} N^{\mu \text{ true}}_{\mu \text{ sel}} & N^{\pi \text{ true}}_{\mu \text{ sel}} \\ N^{\pi \text{ true}}_{\mu \text{ sel}} & N^{\pi \text{ true}}_{\pi \text{ sel}} \end{bmatrix} = \begin{bmatrix} \mathrm{p}_{\mu \text{ sel}} & 1 - p_{\mu \text{ sel}} \\ 1 - p_{\pi \text{ sel}} & \mathrm{p}_{\pi \text{ sel}} \end{bmatrix} \begin{bmatrix} N^{\text{non-e true}}_{\mu \text{ sel}} & N^{\text{non-e true}}_{\pi \text{ sel}} \end{bmatrix}. \tag{4.32}$$

Similarly, we calculate the number of true electrons selected as muons and pions:

$$\begin{bmatrix} N^{\text{e true}}_{\mu \text{ sel}} \\ N^{\text{e true}}_{\pi \text{ sel}} \end{bmatrix} = \begin{bmatrix} f^{\text{non-e sel, e true}}_{\mu \text{ sel}} \\ f^{\text{non-e sel, e true}}_{\pi \text{ sel}} \end{bmatrix} N^{\text{e true}}_{\text{non-e sel}} \tag{4.33}$$

and the number of true muons and pions selected as electrons:

$$\begin{bmatrix} N^{\mu \text{ true}}_{\text{e sel}} \\ N^{\pi \text{ true}}_{\text{e sel}} \end{bmatrix} = \begin{bmatrix} f^{\text{e sel, non-e true}}_{\mu \text{ sel}} \\ f^{\text{e sel, non-e true}}_{\pi \text{ sel}} \end{bmatrix} N^{\text{non-e true}}_{\text{e sel}} \tag{4.34}$$

From there, we obtain the true number of muons:

$$N^{\mu \text{ true}} = N^{\mu \text{ true}}_{\mu \text{ sel}} + N^{\mu \text{ true}}_{\pi \text{ sel}} + N^{\mu \text{ true}}_{\text{e sel}}, \tag{4.35}$$

And the purity and efficiency of the muon selection:

$$\mathrm{p}^{\mu \text{ true}} = \frac{N^{\mu \text{ true}}_{\mu \text{ sel}}}{N^{\mu \text{ true}}_{\mu \text{ sel}} + N^{\text{e true}}_{\mu \text{ sel}} + N^{\pi \text{ true}}_{\mu \text{ sel}}} \text{ and } \epsilon^{\mu \text{ true}} = \frac{N^{\mu \text{ true}}_{\mu \text{ sel}}}{N^{\mu \text{ true}}_{\mu \text{ sel}} + N^{\mu \text{ true}}_{\pi \text{ sel}} + N^{\mu \text{ true}}_{\text{e sel}}}. \tag{4.36}$$

The purity and efficiency of the pion and electron selections are calculated in the same way. Table 4.7 shows the fraction of each true particle in the different ACT selections calculated using the above unfolding method for a run at 460 MeV/c. We observe that the muon selection has the worst purity of 91% out of the three populations, which is expected because the muon population lies between the other two in both the calorimeter and ACT23 charge distributions.

4.6.2 Summary of the PID using ACTs

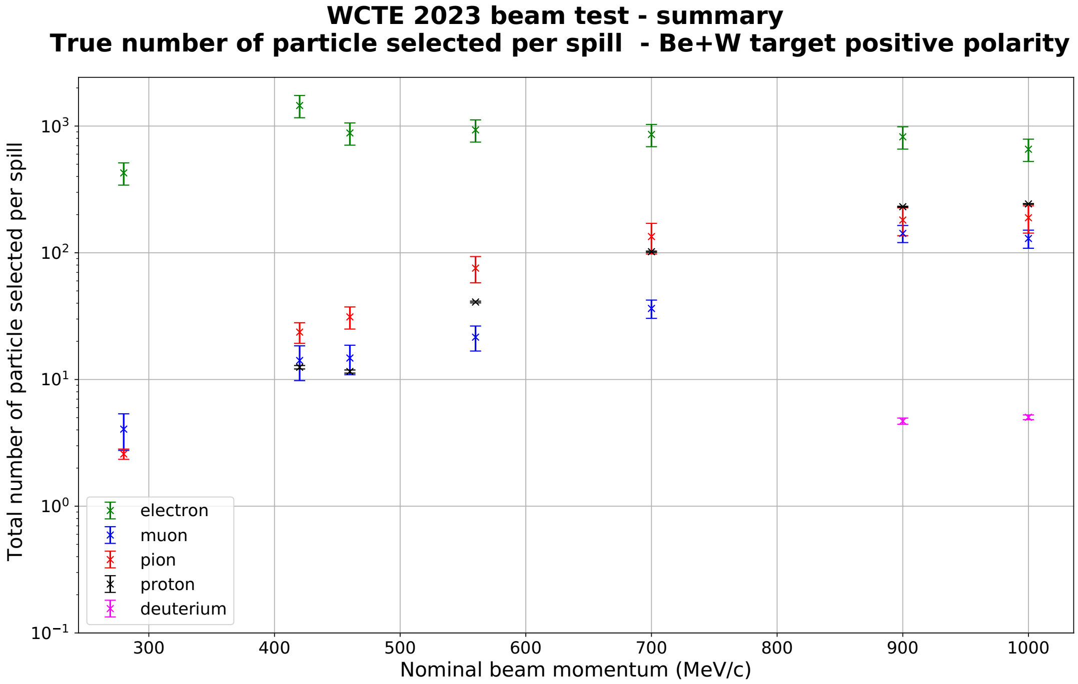
The total number of events falling into each selection (electron-like, muon-like, pion-like, proton-like, deuterium-like) per beam spill is shown in Figure 4.28 for the Be+W target in positive polarity (a) and negative polarity (b), and in Figure 4.29 for the Al target in positive polarity (a) and negative polarity (b). The apparent kink at 420 MeV/c is easily explained by the different slit opening setting (1% or 3%) discussed in Section 4.6.3. In the data corresponding to run 432 at 460 MeV/c, which has been shown as an example in earlier sections, a total of 1657 muons and 3476 pions were identified by the selection (see Figure 4.19) out of 107 spills corresponding to 15.5 muons and 32.5 pions selected per spill on average, as shown on Figure 4.28a.

The 2023 beam test aimed to identify as many of the single particles going through the beam line as possible. Some of the triggers that were collected are not accepted by any selections either because two (or more) particles cross the beam line within the 540 ns window or because the event is identified as noise (see Section 4.3 for more detail).

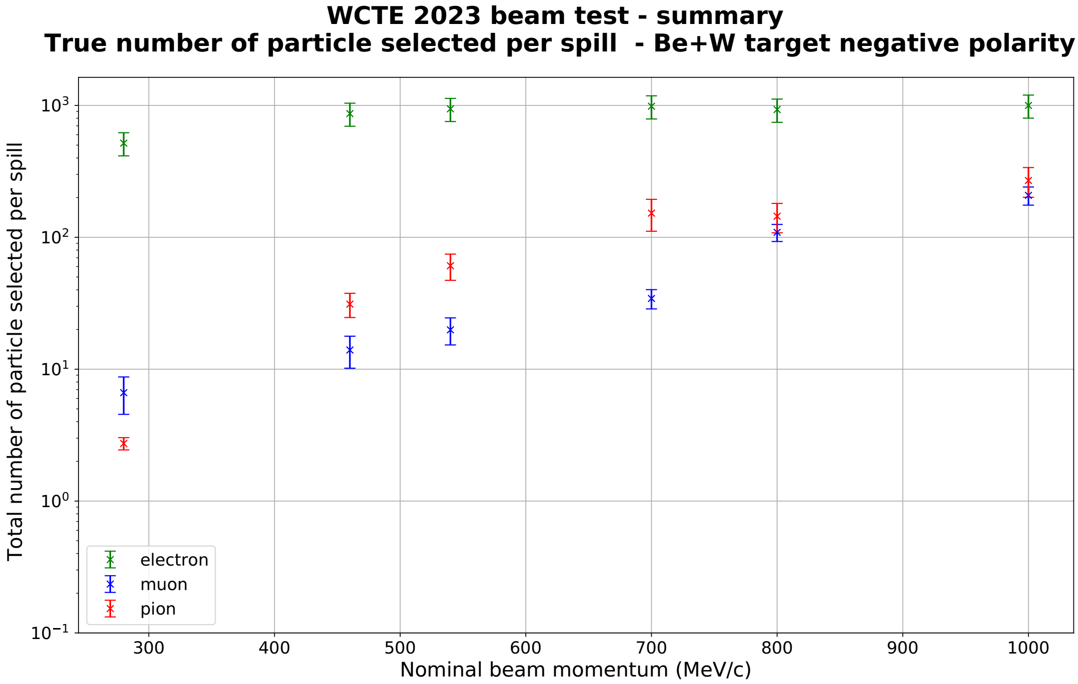
We report in Figure 4.30 the purity of the electron, muon and pion selection made using the 2023 data. During WCTE operation, the selections presented here will be adapted to each analysis to ensure that the purity requirements presented in Table 4.1 are met. With the 2023 set of aerogel refractive indexes, this analysis reaches a purity for the muon selection of about 80% at momenta below 500 MeV/c and about 50% at higher momenta. The purity of the pion selection is roughly constant at about 90%. The purity and efficiency of the selection in positive and negative polarities are approximately equal.

Figures 4.31 show the efficiency of these selections. The efficiency of the muon selection is about 70% on average at momenta higher than 400 MeV/c; it is lower at lower momenta, with a value of about 50%. The efficiency of the pion selection is very high at low momenta, with values close to 100% and decreases to about 70% at momenta higher than 700 MeV/c.

The main source of particle misidentification is the scintillation light discussed in Section 4.4.4. It reduces both the electron rejection and the pion/muon separation capabilities. Figure 4.32 shows the amount of light produced per cm of aerogel by particles above Cherenkov threshold in ACT23 (i.e. electrons and muons), showing the linear relationship between the light yield and (1-n) expected from Cherenkov light. The linear fits do not intercept the origin due to the contribution of scintillation light (about 1.5 PE per box for minimally ionising particles, as can be seen in Figure 4.21). Using these values, we can estimate the total number of Cherenkov photons produced per particle in the ACTs, which can be used to estimate the purity and efficiency of the selections if the scintillating film is removed from the ACTs. For example, a muon above threshold in the 6 cm ACT box with the lowest available refractive index of $n = 1.01$ produces on average 3.6 PE/$\epsilon^{\text{collection}}$ where $\epsilon^{\text{collection}}$ is the light collection efficiency, which we assume to be the same for Cherenkov and scintillation light.



(a) Positive polarity.

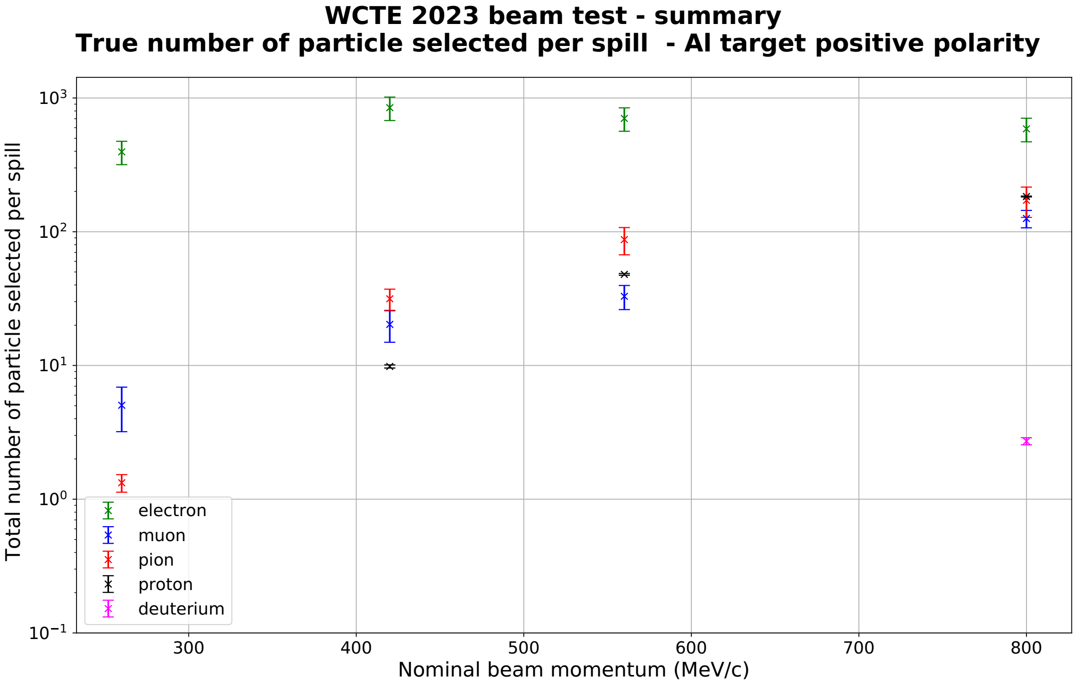


(b) Negative polarity.

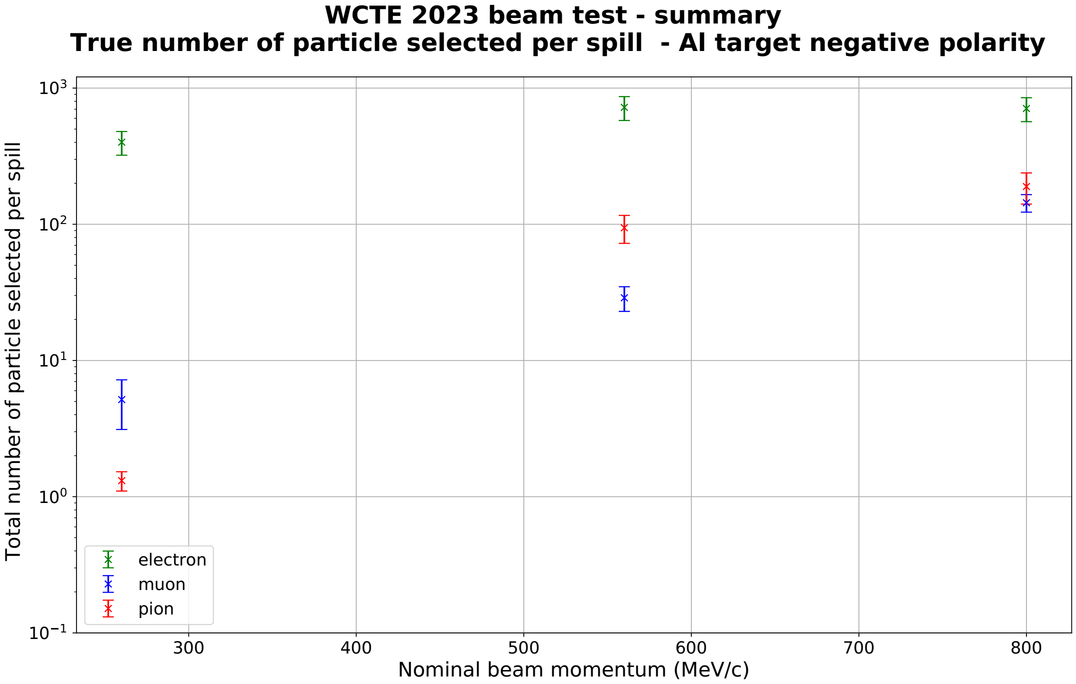
Figure 4.28: Number of triggers accepted by each selection when using the T09 Be+W target.

4.6.3 Particle rates and missed spills

All datasets with nominal beam momenta below 450 MeV/c were collected with a beam momentum spread of 3%, whilst datasets at higher momenta had a 1% momentum spread, which is set by the opening of the



(a) Positive polarity.



(b) Negative polarity.

Figure 4.29: Number of triggers accepted by each selection when using the T09 Al target.

collimator jaws. We expect the overall number of particles to increase when the momentum spread is larger. Table 4.8 shows the number of triggers collected per beam spill and gives a rough indication of the increase in particle rate caused by the opening of the collimator jaws. During data taking, we observed that the DAQ was not capable of coping with more than about 1000 triggers per spill and would crash at higher rates, causing

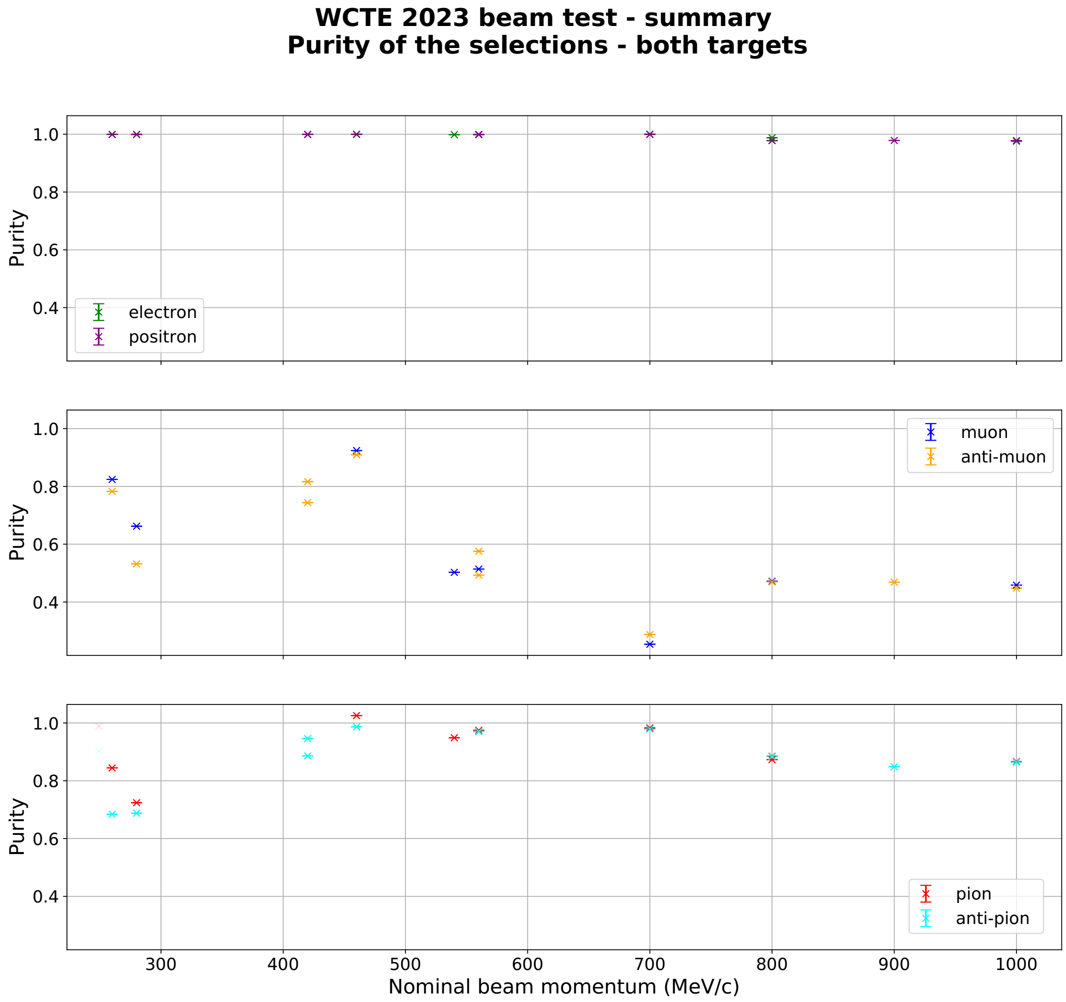


Figure 4.30: Purity of the electron, muon and pion selections. Both Al and Be+W targets are included since no significant differences were observed between them.

the spill to be entirely missed by the DAQ. For example, at 700 ±21 MeV/c the DAQ misses 59% of the spills, whereas it misses 46% of them at 700±7 MeV/c due to the DAQ dead time. We increased the built-in dead-time to reduce the rate of DAQ failure, and modified the collimator jaws opening to maximise the data-taking rate.

4.6.4 Beam structure and composition of the CERN T09 beam

The total number of particles going through the beam line can be estimated from the number of triggers that were identified as each particle type by the ACT selection. Many particles crossing the beam line will not trigger the DAQ either because they are rejected by the online electron veto or because they cross the beam line during the DAQ's dead time of 408 μs. Tables 4.9 and 4.10 present the fraction of triggers (f_e, f_μ, f_π, f_p and f_D) that are accepted by each selection, identified as many-particle events (f_{2+}) or rejected as noise-like (f_noise) for all runs that were used in this analysis. The fraction of events identified as noise reflects the limitation

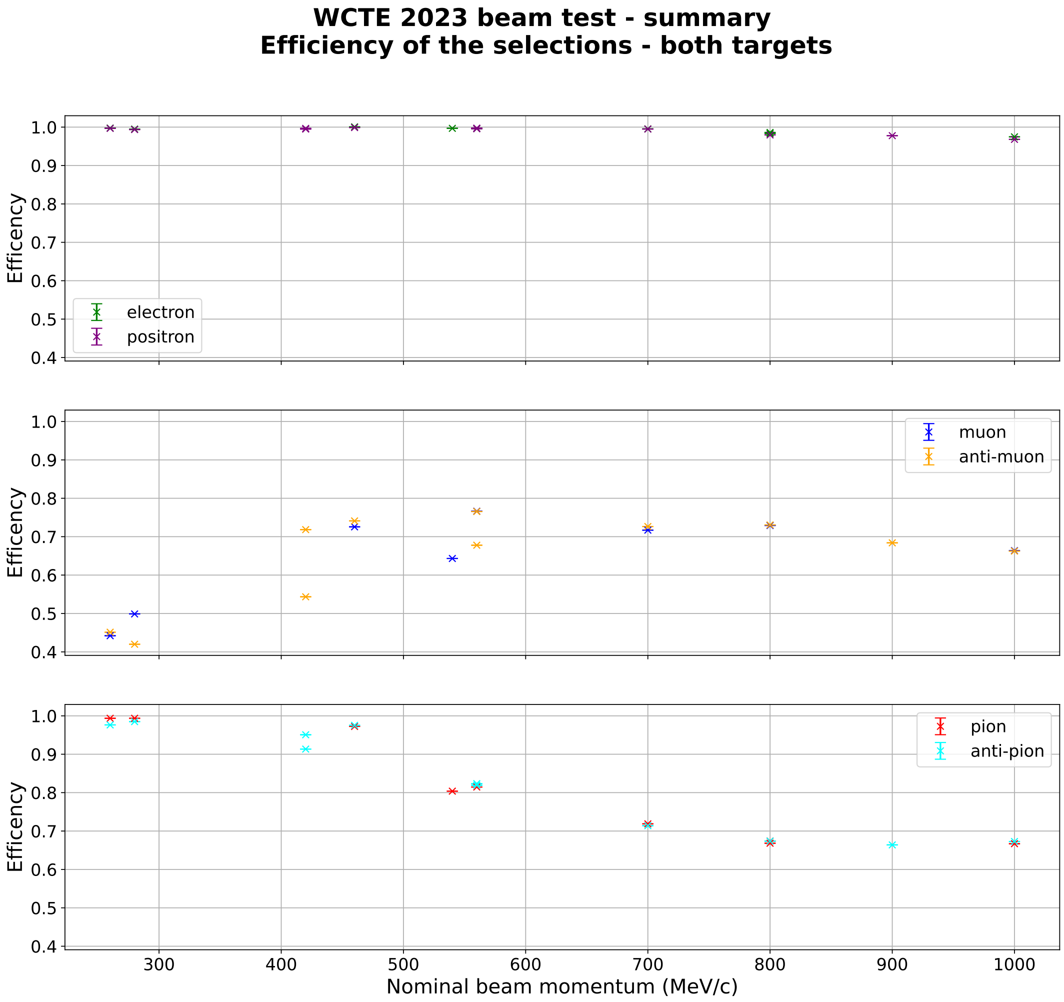


Figure 4.31: Efficiency of the electron, muon and pion selections. Both Al and Be+W targets are included since no significant differences were observed between them.

of the WCTE beam line detectors. As the beam momentum increases, the fraction of triggers identified as multi-particle-like increases due to the overall higher particle rate.

It is interesting to note that runs 342, 352 and 480 have a higher rate of two particle-like and noise-like triggers than other runs at similar momenta. This is because the digitiser timing offset correction failed for these runs. In these cases, the coincidence requirement was loosened such that the two PMTs in TS0 had to be in coincidence with only one in TS1 instead of two. This looser coincidence requirement accepts more accidental hits, which leads to a higher rate of multi-particle events as well as a less accurate placement of the integration window, which causes more triggers to look like noise. These runs are reported for completeness but will not be used for further analysis. We see, for example, that the fraction of identified electrons is lower than the expected 90% or so seen in the Aluminium sample.

A comparison between data taken with and without the online electron veto shows that it removes 1 out of

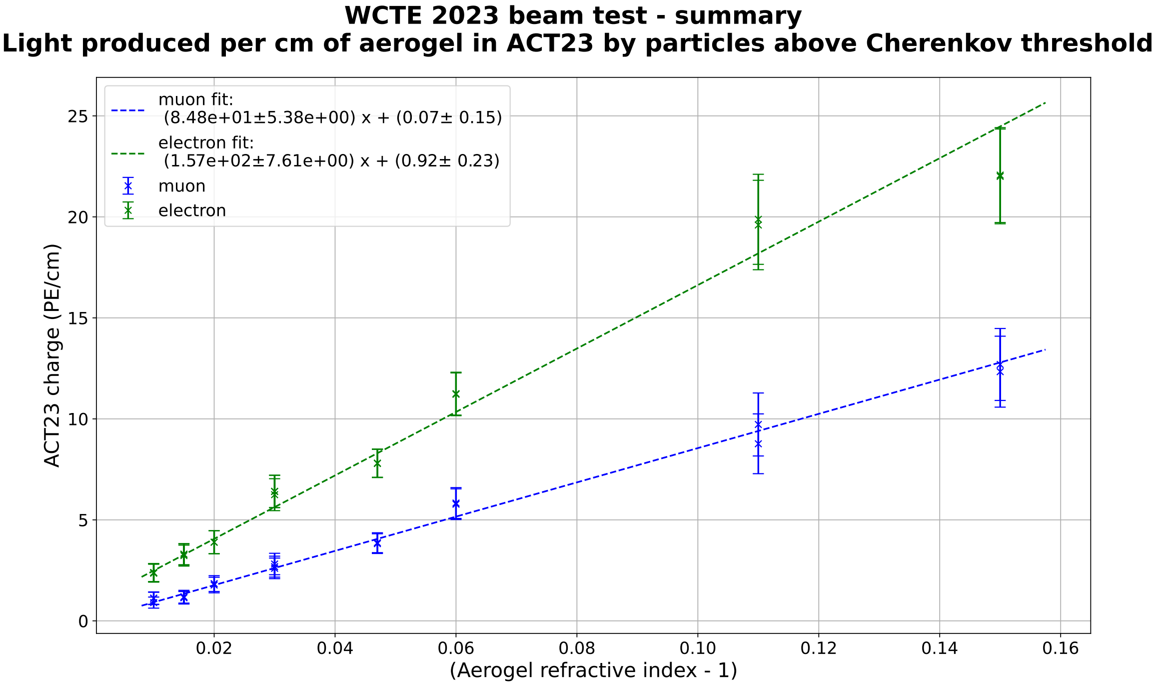


Figure 4.32: Number of PE collected per cm of aerogel by muons and electrons above threshold, showing a clear linear dependency on the refractive index n. The thickness of each aerogel is given in Table 4.2.

Momentum	1% momentum spread	2% momentum spread	3% momentum spread
500 MeV/c	1336.8	-	1380.7
700 MeV/c	909.4	895.4	498.8

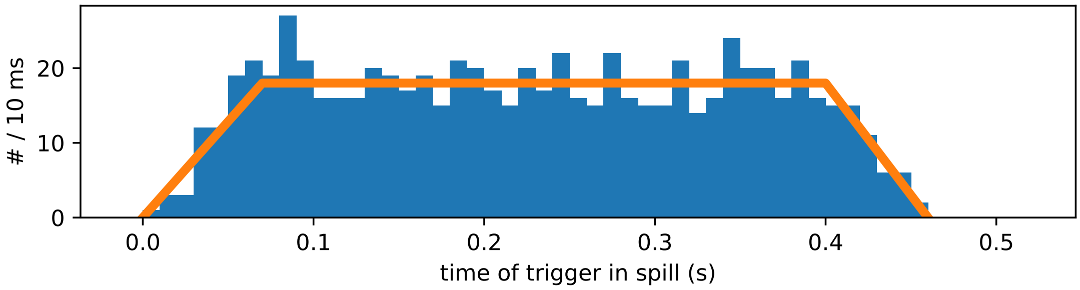
Table 4.8: Number of single-particle-like events selected per spill for a few runs with different collimator jaw settings. The reduction in jaw opening width is observed to significantly reduce the number of particles.

every 7 electrons. In this version of the analysis, we treat the online electron veto as a (very small) contribution to the overall dead time. The beam structure was studied in depth by WCTE collaborators [143], and it was found that each beam spill lasts 0.48 ± 0.02 s during which the rate of incoming particles increases for 0.08 s and remains constant for the remaining 0.40 s after sharply dropping back to zero. Figure 4.33a shows the number of triggers collected as a function of time, highlighting the spill structure described above.

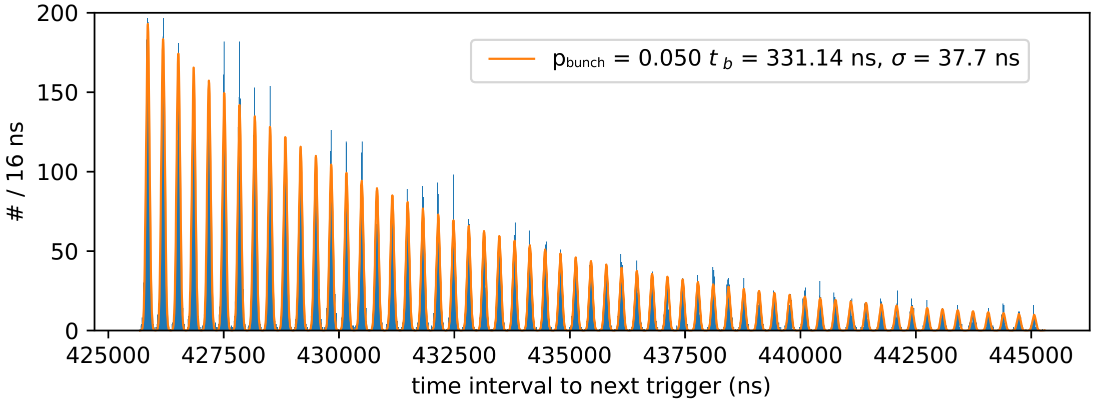
Measuring the time delay between consecutive triggers, as shown in Figure 4.33b, shows that particles arrive in bunches separated by 331 ns on average. The standard deviation of bunches is 26.7 ns. If we assume that each bunch has the same probability ($p_{\text{bunch}} = P(T)$) to contain (at least one) particle that triggers the DAQ, the probability that the next trigger is n bunches away from it is [143]:

$$P(T_n|T_0) = (1 - p_{\text{bunch}})^{n-1} p_{\text{bunch}}. \tag{4.37}$$

We measure p_{bunch} by fitting the decaying amplitude observed in the discrete distribution of delay between triggers. For the run at 1GeV/c during which the electron veto was disabled, the probability for a particular



(a) Number of triggers collected per unit of 8 ns, showing the time structure of one spill.



(b) Distribution of delays between two consecutive DAQ triggers, fitted with an exponentially decaying sine function showing that particles arrive in fixed time intervals, so-called bunches, separated by 331 ns on average. The exponentially decaying amplitude is directly related to the probability for a given bunch to contain a particle, in this case 0.05%.

Figure 4.33: Study of the triggering rate made with the beam at 1 GeV/c when the online electron veto was off showing the beam spill length and bunched structure. The author thanks D. Karlen for providing both figures.

bunch to trigger the DAQ is $p \approx 5\%$. This probability reduces for lower momenta and when the electron veto is applied. The probability p_{bunch} for each bunch to trigger the DAQ is shown in Tables 4.9 and 4.10.

Knowing the probability p that a bunch triggers the DAQ, we can extrapolate the total number of particles of each type x in the beam:

$$n_x = p_{\text{bunch}} \times f_x \times n_{\text{bunch}} \tag{4.38}$$

where f_x is the fraction of triggers identified as each true particle $x = \{e,\ \mu,\ \pi,\ \text{p},\ \text{D})$ and n_{bunch} is the total number of bunches per spill:

$$n_{\text{bunch}} = \frac{\text{duration of spill}}{\text{delay between bunches}} = \frac{0.46 \pm 0.03\ \text{s}}{331 \pm 38\ \text{ns}} = (1.39 \pm 0.18) \times 10^6\ \text{bunches/spill} \tag{4.39}$$

where the total spill duration is reduced and its error increased somewhat arbitrarily to account for the lower particle rate at the beginning and at end of the spill.

Run	p_{nominal} (MeV/c)	n_{ACT23}	n spills	n triggers	p_{bunch} (%)	f_{e} (%)	f_{μ} (%)	f_{π} (%)	f_{p} (%)	f_{D} (%)	f_{noise} (%)	f_{2+} (%)
342	(+) 280 ± 3%	1.11	154	83,285	0.05	79 ± 16	1.2 ± 0.2	0.22 ± 0.04	0.00	0.00	4.23	11.62
352	(-) 280 ± 3%	1.11	124	81,622	0.06	78 ± 16	1.6 ± 0.3	0.02 ± 0.01	N/A	N/A	2.51	14.1
480	(+) 420 ± 3%	1.06	82	140,142	0.07	85 ± 17	1.3 ± 0.3	1.3 ± 0.3	0.7 ± 0.1	0.00	1.75	8.70
432	(+) 460 ± 1%	1.047	107	106,515	0.23	88 ± 18	1.9 ± 0.4	3.1 ± 0.6	1.2 ± 0.1	0.00	1.99	3.34
425	(-) 460 ± 1%	1.047	108	103,827	0.21	89 ± 18	2.0 ± 0.4	3.4 ± 0.7	N/A	N/A	1.78	3.14
412	(+) 560 ± 1%	1.03	101	114,547	0.40	82 ± 16	2.1 ± 0.4	7.8 ± 1.6	3.6 ± 0.1	0.00	1.86	3.43
423	(-) 540 ± 1%	1.03	98	106,118	0.33	87 ± 17	2.1 ± 0.4	6.3 ± 1.3	N/A	N/A	1.06	3.65
438	(+) 700 ± 1%	1.02	84	101,527	0.73	71 ± 14	2.5 ± 0.5	15.0 ± 3.0	8.4 ± 0.1	0.00	1.91	4.04
445	(-) 700 ± 1%	1.02	93	116,068	0.64	79 ± 16	2.3 ± 0.5	16.6 ± 3.3	N/A	N/A	1.09	4.15
469	(-) 800 ± 1%	1.015	91	115,155	0.83	74 ± 15	6.4 ± 1.3	14.4 ± 2.9	N/A	N/A	1.23	4.43
398	(+) 900 ± 1%	1.01	66	98,236	0.99	55 ± 11	7.3 ± 1.5	15.1 ± 3.0	15.5 ± 0.1	0.3 ± 0.1	2.16	4.47
393	(+) 1000 ± 1%	1.01	107	141,960	1.49	49 ± 9.8	7.9 ± 1.6	17.4 ± 3.5	18.3 ± 0.1	0.4 ± 0.1	2.10	4.73

Table 4.9: Summary of the data taken with the Be+W target used for characterising the T09 beam. The beam polarity is indicated in brackets before the nominal momentum. n_{ACT23} is the refractive index of the ACT2 and ACT3 detectors; the other quantities are defined in the text.

Run	p (MeV/c)	n_{ACT23}	n spills	n triggers	p_{bunch} (%)	f_{e} (%)	f_{μ} (%)	f_{π}(%)	f_{p} (%)	f_{D}(%)	f_{noise}(%)	f_{2+}(%)
545	(+) $260 \pm 3\%$	1.150	168	71,254	0.041	92 ± 19	2.1 ± 0.4	0.2 ± 0.1	0.0	0.0	2.59	2.04
549	(-) $260 \pm 3\%$	1.150	176	74,639	0.041	93 ± 19	2.4 ± 0.5	0.2 ± 0.1	N/A	N/A	1.61	2.04
575	(+) $420 \pm 3\%$	1.060	136	132,996	0.24	86 ± 17	2.7 ± 0.6	2.9 ± 0.6	1.0 ± 0.1	0.0	2.10	3.81
566	(+) $560 \pm 1\%$	1.030	92	83,888	0.17	77 ± 15	3.7 ± 0.7	11.0 ± 2.2	5.3 ± 0.1	0.0	1.46	2.19
569	(-) $560 \pm 1\%$	1.030	97	85,430	0.33	82 ± 16	3.4 ± 0.7	12.4 ± 2.5	N/A	N/A	1.14	2.17
563	(+) $800 \pm 1\%$	1.015	99	110,461	0.47	53 ± 11	8.4 ± 1.7	19.6 ± 3.9	16.4 ± 0.1	0.2 ± 0.1	1.41	2.21
564	(-) $800 \pm 1\%$	1.015	91	98,757	0.39	65 ± 13	9.8 ± 2.0	22.5 ± 4.5	N/A	N/A	1.19	2.4

Table 4.10: Summary of the data taken with the Al target used for characterising the T09 beam. The beam polarity is indicated in brackets before the nominal beam momentum.

Figures 4.34b and 4.35b report the total number of true particles per spill that would be selected by the analysis as single event particles, in the absence of dead time. Again, the rates reported use a 3% spread for momenta below 450 MeV/c and a spread of 1% at higher momenta, which explains the kink at 420 MeV/c. The beam line apparatus is not magnetised and therefore blind to particle charge. Here we assume that all particles recorded are negatively (positively) charged when the beam polarity is negative (positive).

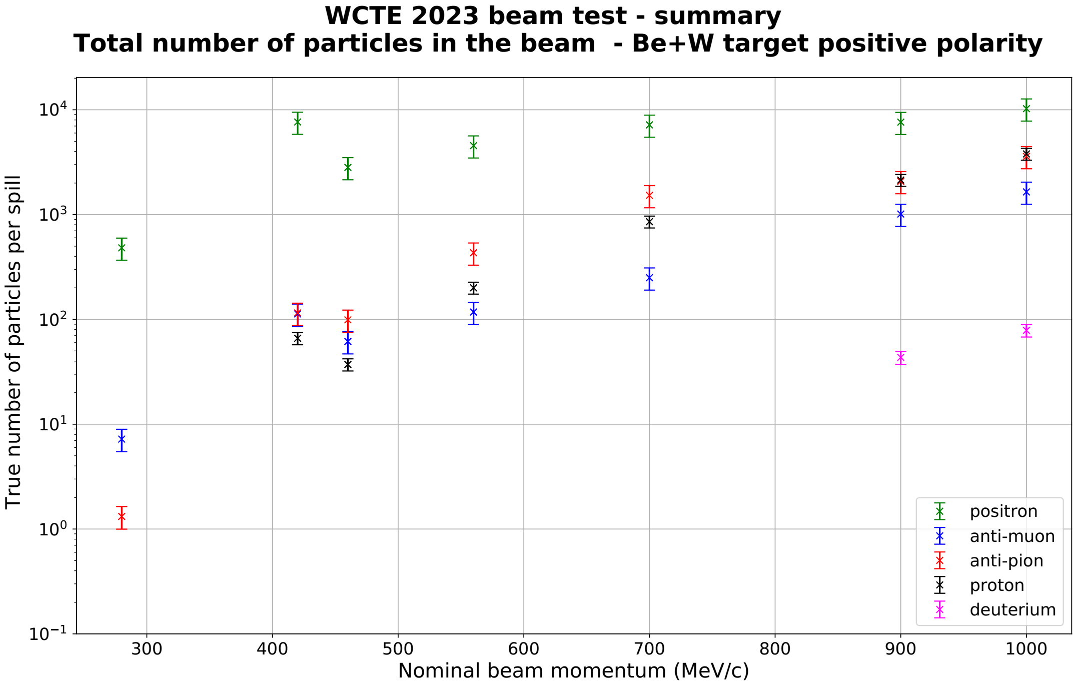
We observe from the particle rates presented here that electrons make up the majority of the beam in the sub-GeV region for both targets. We can see that the fraction of electrons in the beam is larger for the Be+W than for the Al target at the same momenta, which is the expected impact of the tungsten radiator. The overall single-particle rates are similar for the two targets, but the rate of multiple-particle events is significantly lower for the Aluminium target. These results motivated the decision to use the Al target for the WCTE charged set-up operation and the Be+W target in the tagged photon mode.

The results of the 2023 beam test were used to upgrade the design of the beam line hardware and draft a data-taking schedule in preparation for WCTE.

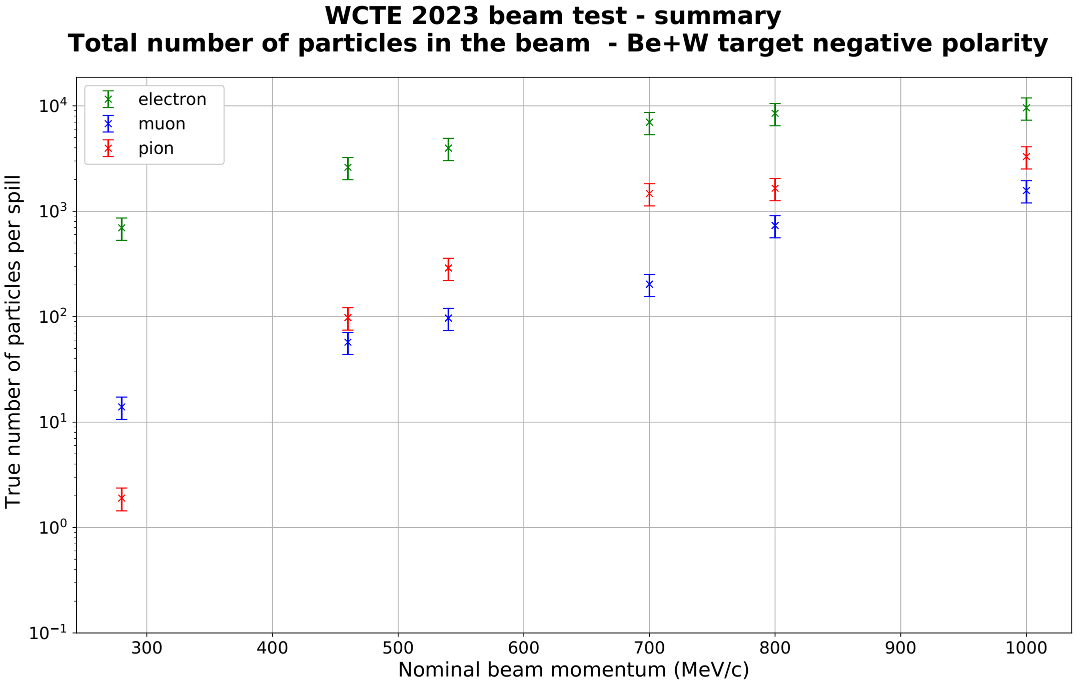
4.7 Summary

After presenting the experimental set-up and detectors used for the 2023 beam test at CERN, this chapter describes the pre-processing and preliminary calibrations applied to the data to make it usable for an analysis. It goes on to discuss how heavier particles like protons and deuterium nuclei are identified using their time of flight before presenting how lighter particles like electrons, muons and pions are identified using the ACTs. Due to the lack of robust beam line simulation, the calorimeter information is used to estimate the true particle type and assess the efficiency and purity of the selection.

From 260 MeV/c to about 600 MeV/c, the muon sample purity is about 60%, it decreases to about 45% at higher momenta. The muon sample efficiency is around 70% at momenta above 400 MeV/c and around 45% at lower momenta, due predominantly to the low number of muons per spill at these low momenta. The purity of the pion sample is consistently above 80%, except at momenta lower than about 300 MeV/c where it is around 70%. Its efficiency is above 90% at momenta above about 500 MeV/c due to the low number of muons in the beam and drops to about 70% after that due to the lower Cherenkov light yield for muons at higher momenta. The electron selection purity and efficiency is close to 100% over the whole momentum range considered because of the high number of electrons in the beam. The main factor limiting the purity of the pion and muon selections is the scintillation light originating from the 3M$^{\text{TM}}$reflective film located inside the ACTs. The problematic film, identified by the present analysis, has been removed from the ACTs in preparation for WCTE data taking, strongly improving the particle separation capabilities. A simple unfolding procedure



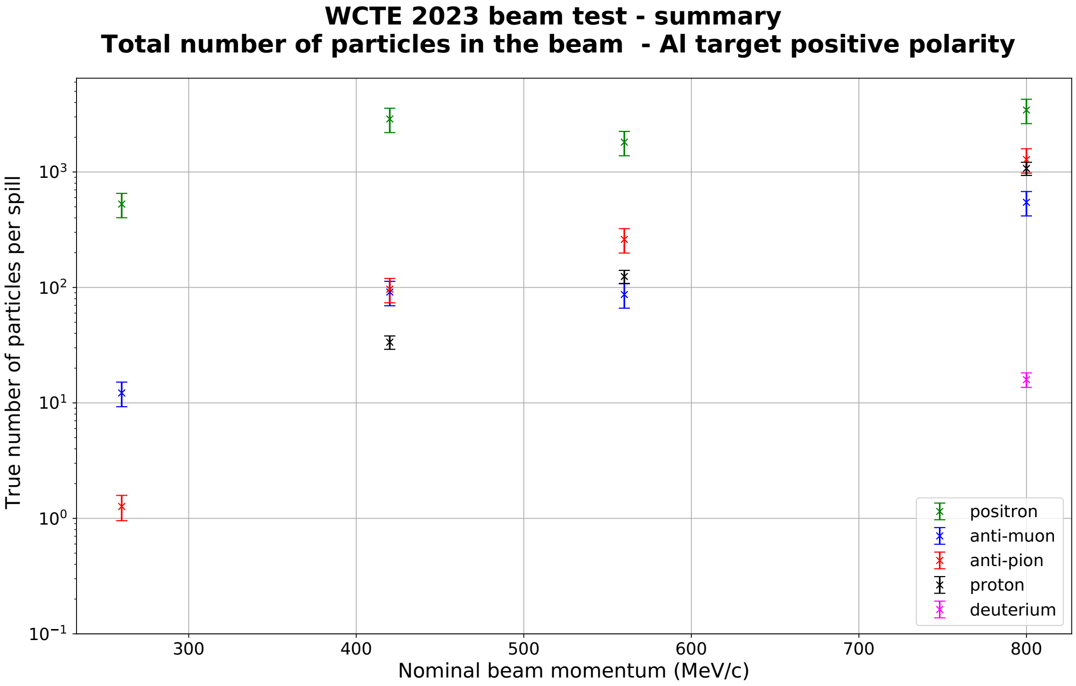
(a) Be+W target in positive polarity.



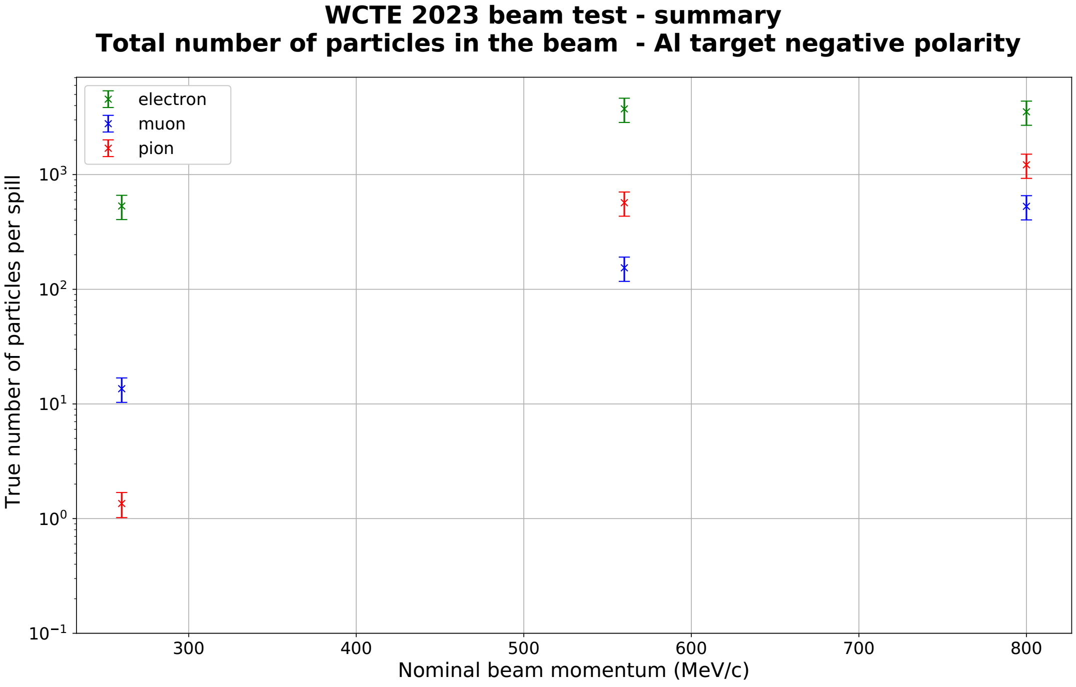
(b) Be+W target in negative polarity.

Figure 4.34: Total number of true particles per spill in the T09 beam, after unfolding to true particle type and accounting for the dead time.

based on calorimeter information was used together with a characterisation of the beam structure, provided by a WCTE collaborator, D. Karlen were used to produce the total number of true particles per spill. We observe that the Be+W target produces about 20% more particles per spill than the Al one in total (e.g. at



(a) Al target in positive polarity.



(b) Al target in negative polarity.

Figure 4.35: Total number of true particles per spill in the T09 beam, after unfolding to true particle type and accounting for the dead time.

560±5.6 MeV/c, about 1130 particles per spill with the Be+W target as opposed to about 910 particles per spill with the Al target). However, we note that the proportion and total number of muons, pions and protons is larger for the Al target, which produces, for example, 33 muons, 100 pions and 48 protons per spill at 560 MeV/c compared to the Be+W target, which produces only 23 muons, 88 pions and 40 protons per spill at the same momenta. This difference is even more prominent at high momenta. For example, at 800 MeV/c the fraction of electrons measured in the beam is 74±15% for the Be+W target compared to 65±13% for the Al target. These results were used by WCTE to decide to run with the Al target for its charged particle studies and with the Be+W target for the tagged photon production studies, it also informed the time that should be dedicated to each momentum point in order to acquire sufficient statistics. In light of the results of this study, the choice was made, for example, to collect a large amount of data, to be used for muon scattering studies, at 700 MeV/c, which gives the highest expected rate of muon scattering events (i.e. finding the optimal balance between number of muon per spill and scattering cross-section).

Finally, this chapter also includes a measurement of the beam momentum using both the time of flight of particles identified with the ACTs and the energy deposited in the lead-glass calorimeter by electrons. The first method was used to estimate the momentum bias of the beamline and will be used in WCTE to measure the momentum of particles before they reach the Cherenkov detector. The second method provides a calibration benchmark necessary for measuring the energy of photons produced in the tagged photon mode and evaluating the precision and accuracy of this novel set-up, to be re-used in WCTE.

Chapter 5

The T2K experiment

5.1 Introduction

The Tokai to Kamioka (T2K) experiment is a long-baseline neutrino experiment located in Japan. It started operating in 2009 [8] and has experienced many successes since then. Its main objective was to observe electron neutrino appearance in a muon neutrino beam [8]. Other experiments, including Super-Kamiokande (which discovered neutrino oscillations in 1998 [33]), had already measured neutrino oscillation via the disappearance of muon neutrinos [29, 148, 149], but T2K was the first to measure evidence of flavour appearance in the $\nu_\mu \rightarrow \nu_e$ channel. It concluded that θ_{13} had a non-zero value at 7.3 σ confidence level [150]. The precise measurement of θ_{13} was later achieved by reactor neutrino experiments [151, 152].

Since achieving its initial goal, the T2K experiment has carried on its measurement of the neutrino oscillations, publishing in 2014 the world's best measurement (at the time) of the θ_{23} neutrino mixing parameter [153]. More recently, the analysis of 3.6×10^{21} protons on target (POT) collected by T2K excluded at more than 90% the CP-conserving values of δ_{CP} : 0 and π [2]. T2K continues its oscillation physics program which requires a precise understanding of neutrino interactions. This motivates the upgraded near detector, which started taking data in 2023 [154].

This chapter describes the T2K experiment, shown on Figure 5.1. The T2K (anti) muon neutrino beam produced at the Japan Proton Accelerator Research Complex (J-PARC) is presented in Section 5.2. The near detector complex used to characterise the neutrino beam before it oscillates is discussed in Section 5.3. T2K's far detector, Super-Kamiokande, where the neutrinos are measured after having travelled 295 km through the Earth, is described in Section 5.4.

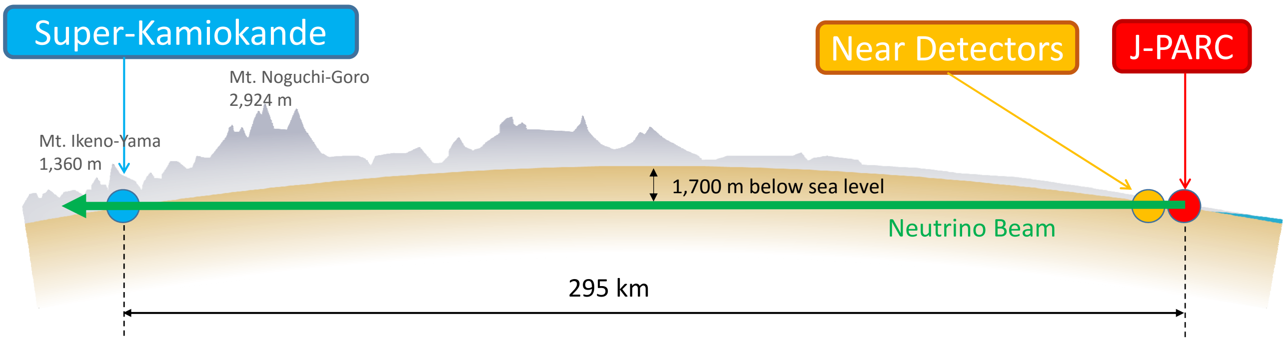


Figure 5.1: The T2K experiment. Figure modified from [8].

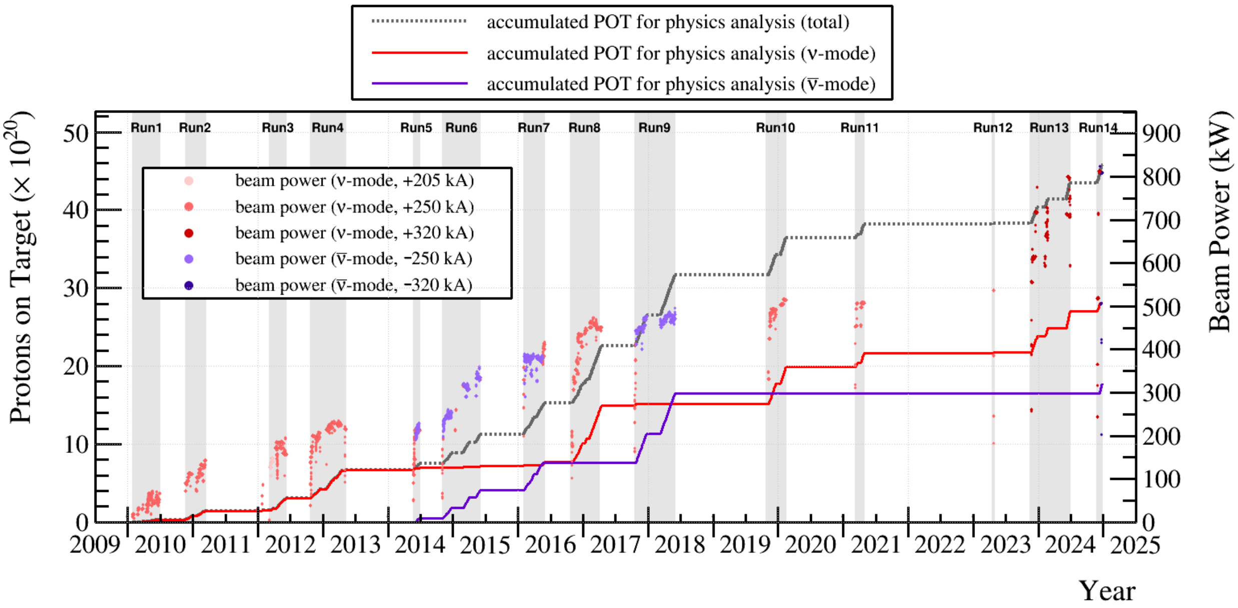


Figure 5.2: Summary of the POT delivered to the T2K experiment and the beam power. The current of the magnetic horns is indicated in the legend. Adapted from [2].

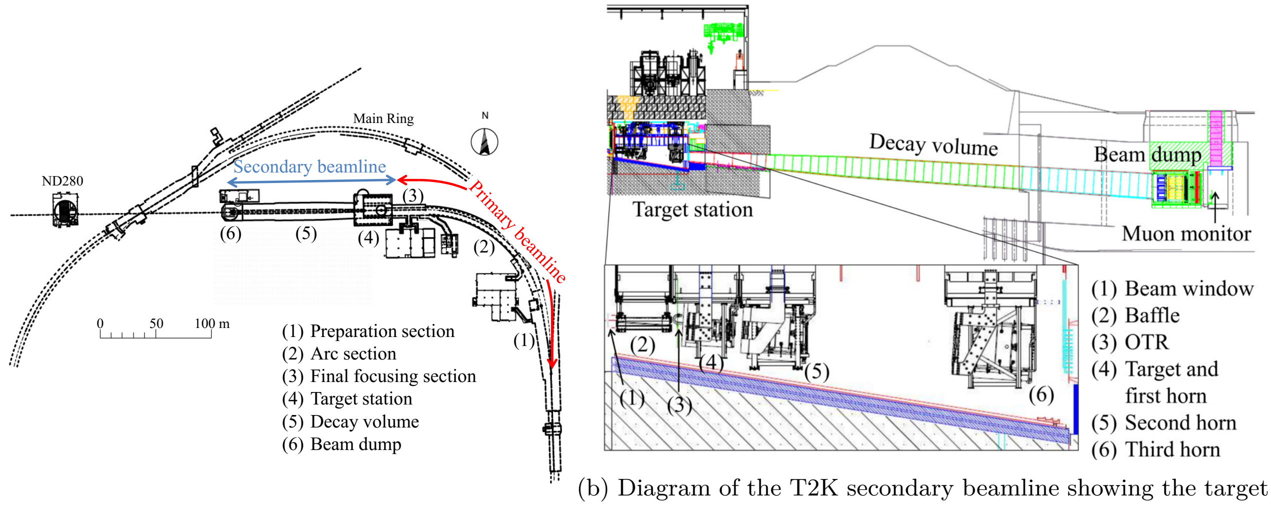
5.2 Neutrino beam

5.2.1 The beam line

Accelerator neutrinos are typically produced by shooting a proton beam onto a target. The T2K experiment measures the amount of data it collects in terms of protons on target (POT). The amount of POT collected since the start of the experiment is shown in Figure 5.2. At the time of writing, the T2K experiment has collected a total of 2.7×10^{21} POT in Forward Horn Current (FHC, i.e. neutrino) mode and 1.6×10^{21} POT in Reverse Horn Current (RHC, i.e. anti-neutrino) mode. A stable beam power of 750 kW has been achieved in December 2023 and exceeded since, in preparation for the Hyper-Kamiokande experiment. The analysis presented in Chapter 6 uses the data collected in the FHC mode with the pre-upgrade T2K near detector ND280, corresponding to runs 2, 3, 4 and 8. The following sections present this configuration of the T2K experiment, the reader is encouraged to consult [154] for more details on the near detector upgrade.

Protons colliding with the target produce many secondary hadrons, which decay to neutrinos. The J-PARC beamline uses three accelerators to produce its proton beam. The first one is a LINAC, which accelerates H^- ions to 181 MeV. The second one is a rapid-cycling synchrotron (RCS). At the injection point of the RCS, the H^- ions are converted to H^+ (i.e. protons) using charge stripping foils [155]. They are then accelerated to 3 GeV with a 25 Hz cycle. The third accelerator is the 30 GeV Main Ring synchrotron (MR). The MR has nine radiofrequency (RF) buckets, eight of which are filled with protons coming from the RCS. All eight filled buckets (together forming one “spill”) are transferred to the neutrino beamlines (primary and secondary) using fast extraction (i.e. within one complete revolution of the synchrotron). The extraction takes about 5 μs. This ensures that the neutrino beam has a clear timing structure, which helps reduce the rate of accidental backgrounds at the near and far detectors. There are approximately 3×10^{14} protons per spill.

In the primary neutrino beamline, shown in Figure 5.3a, the beam is first focused and bent into an arc section and finally focused and directed towards Kamioka. The beam needs to be well-tuned (i.e. focused and guided towards the target) and stable for reliable neutrino production. A suite of primary beamline detectors is used to monitor the beam intensity, profile, position and loss [8] before they reach the secondary beamline.



(a) Diagram of the T2K primary and secondary beamlines reproduced from [8].

(b) Diagram of the T2K secondary beamline showing the target station, the decay volume and the beam dump as well as the muon monitor, reproduced from [156].

Figure 5.3: The T2K primary and secondary beamlines.

The secondary beamline is made up of a target station, which is contained within a 1.5×10^3 m^3 helium vessel, followed by a 96 m-long decay volume and finally a 75 tons, 3.174 m-long beam dump made of graphite. A diagram of the target station is shown in Figure 5.3b. The protons enter the target station through the beam window and go through the baffle (a graphite block with a hollowed-out circular portion, 30 mm in diameter). The baffle acts as a collimator and protects the horns downstream from receiving stray charged particles that could damage them. The position and width of the proton beam is measured one final time by the Optical Transition Radiation monitor (OTR) [157] before it collides with the target. The T2K target is a 91.4cm-long (1.9 radiation length) graphite rod. As the beam collides with the target, about 80% of the protons interact with the target, producing secondary mesons, mostly pions and kaons.

The T2K experiment uses three magnetic horns to collect and focus the charged particles produced in the target, which is located within the first of the three horns. Each horn is run with a positive or negative 320kA pulsed current producing a 1.7 T magnetic field. When the current is positive, the horns focus positively charged particles and defocus negatively charged ones, the FHC mode and reciprocally when the current is negative, i.e. the RHC mode. As they travel through the decay volume, these secondary (and tertiary) particles will decay. The main decay modes of the particles expected in the T2K secondary beamline are shown in Table 5.1. We see that in the FHC (RHC) mode a majority of neutrinos (anti-neutrinos) are produced. All particles except neutrinos and high-energy ($\geq$ 5 GeV) muons are absorbed by the beam dump located at the end of the decay volume.

Particle	**Decay Products**	**Branching Fraction (%)**
π^+	$\rightarrow \mu^+\nu_\mu$	99.9877
	$\rightarrow e^+\nu_e$	1.23×10^{-4}
K^+	$\rightarrow \mu^+\nu_\mu$	63.55
	$\rightarrow \pi^0\mu^+\nu_\mu$	3.353
	$\rightarrow \pi^0 e^+\nu_e$	5.07
K_L^0	$\rightarrow \pi^-\mu^+\nu_\mu$	27.04
	$\rightarrow \pi^- e^+\nu_e$	40.55
μ^+	$\rightarrow e^+\bar{\nu}_\mu\nu_e$	100

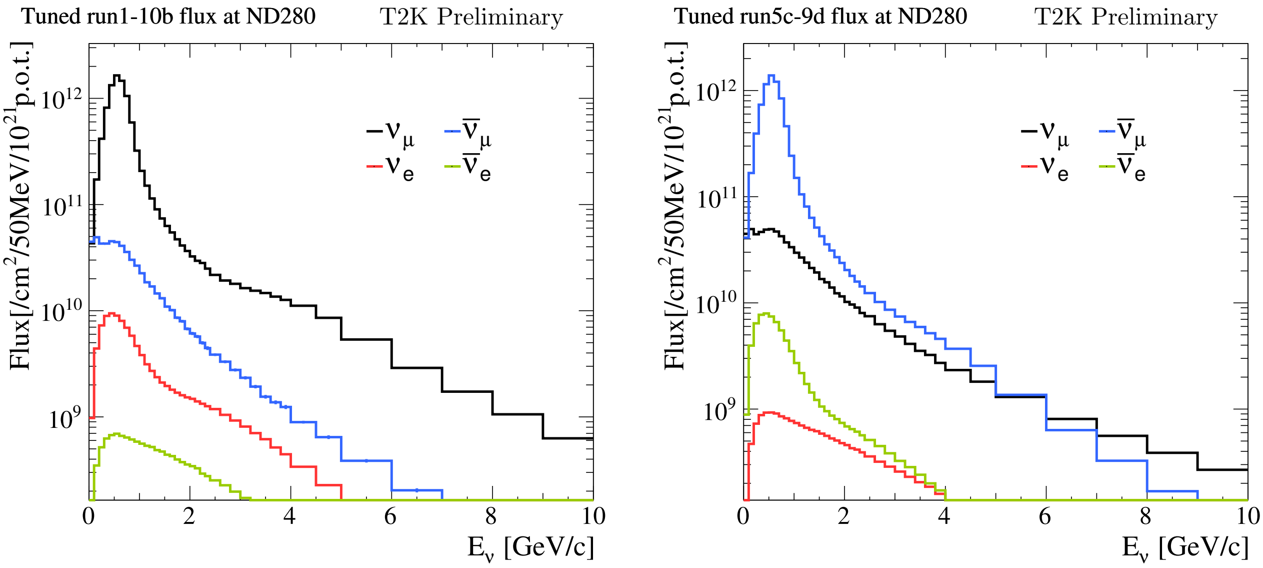
Table 5.1: Table of the decay modes of particles producing neutrinos in FHC mode. Reproduced from [156].

The high-energy muons that escape the beam dump are used to infer the direction and intensity of the neutrino beam using the muon monitor (MUMOM). This detector consists of ionisation chambers and silicon photodiodes. Using MUMOM and the fact that high-energy muons are mostly produced in the two-body decay of pions: $\pi^+ \rightarrow \mu^+ + \nu_\mu$, the direction of the neutrino beam is measured with a precision better than 0.28 mrad, and the stability of its intensity is measured with a precision better than 3% [158]. The direction of the neutrino beam is also measured by the INGRID on-axis detector described later in section 5.3.1.

5.2.2 Beam composition and simulation

An MC simulation is used to estimate the T2K neutrino flux at the near and far detectors. This simulation tracks all secondary particles and any proton that did not interact with the target up to their absorption or decay to daughter particles. The simulation of the primary interaction of 30 GeV protons with the graphite target and of secondary particle production, used to make neutrino flux predictions is tuned using NA61/SHINE data on a replica of the T2K target experiments at CERN [159, 160]. The subsequent interactions within the target is simulated using FLUKA [161]. The particle interactions outside the target are simulated using GEANT3. The predicted neutrino flux at ND280 is shown for the FHC and RHC modes in Figure 5.4.

The uncertainty in the neutrino flux is one of the main sources of uncertainty in T2K. The muon scattering



(a) FHC mode (b) RHC mode

Figure 5.4: Simulated neutrino flux at the T2K near detector. Reproduced from [2].

analysis presented in Chapter 6 makes use of the fact that the incoming muon flux is measured before entering the fiducial volume of interest, allowing us to normalised the MC simulation directly to the data, as discussed in Section 6.3.1.

5.2.3 Off-axis angle concept

The beam axis is offset from the target to the far detector direction by 2.5°. The T2K experiment was the first to use this off-axis method, which provides great advantages in the study of neutrino oscillations and interactions. The majority of muon neutrinos in the beam are produced in the two-body pion decay: $\pi^+ \rightarrow \mu^+ + \nu_\mu$. The energy E_ν of such a neutrino depends on its angle θ with respect to the pion direction:

$$E_\nu = \frac{m_\pi^2 - m_\mu^2}{2(E_\pi - p_\pi \cos\theta)} \tag{5.1}$$

where m_π and m_μ are the pion and muon masses, E_π is the pion's energy and p_π its momentum.

By choosing a value of 2.5° for the off-axis angle, T2K ensures that the neutrino flux peaks close to 600 MeV. This corresponds to the maximum ν_e appearance (and ν_μ disappearance) probability for a 295 km baseline, as shown in Figure 5.5. This figure also shows that a higher off-axis angle reduces the energy spread of the neutrino beam, keeping it in the region dominated by CCQE interactions, for which it is easier to reconstruct the neutrino energy from the kinematics of the observed lepton, as was discussed in Section 2.1.1. It also limits the number of higher-energy unoscillated events at the far detector, in particular NCπ^0 events, which are a significant background to the oscillation measurement [162, 163].

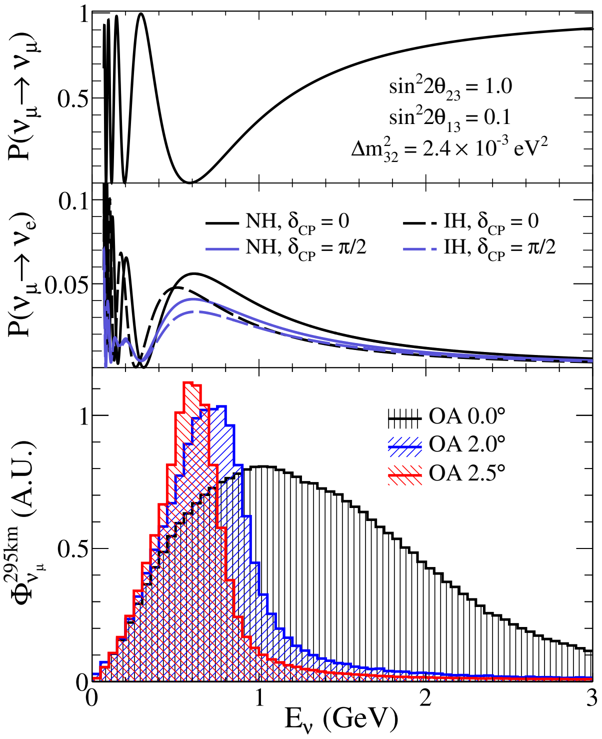


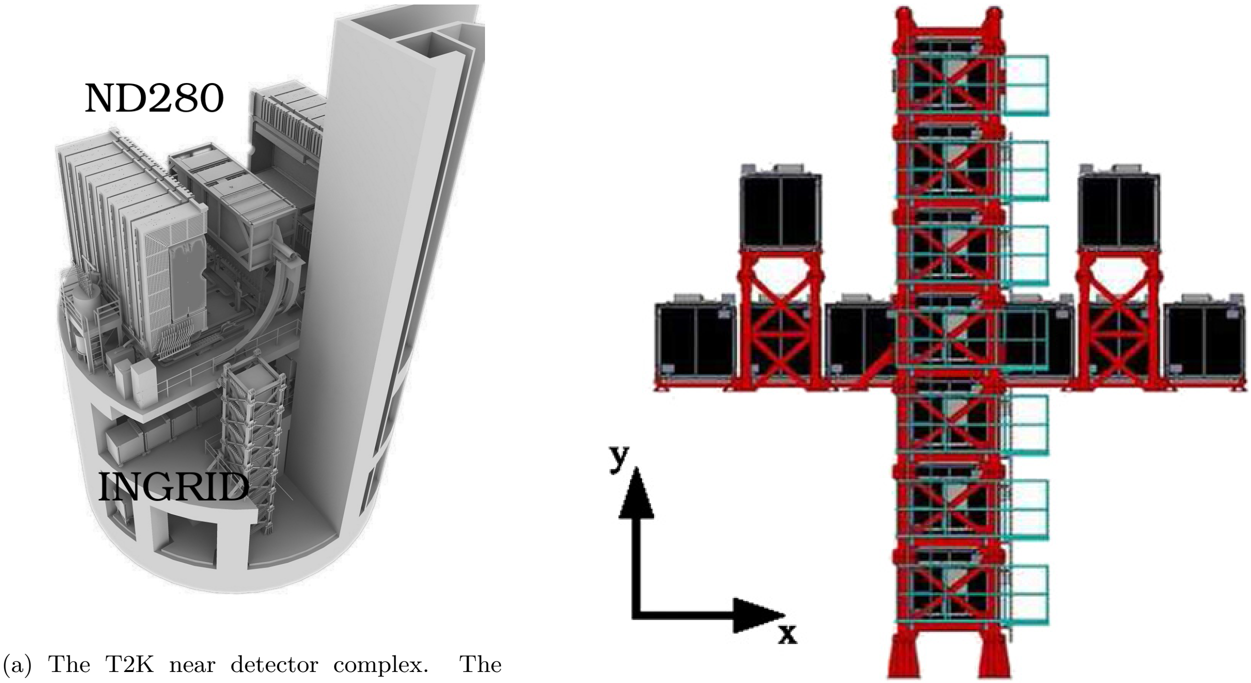
Figure 5.5: Neutrino oscillation probabilities as a function of their energy for a 295 km beamline and presentation of the T2K flux for three different choices of off-axis angle showing that an off-axis angle of 2.5° is optimal. Reproduced from [156].

5.3 Near detectors

The T2K experiment uses a pair of detectors, located 280 m away from the target. These are housed in the near detector complex, shown in Figure 5.6a. Two of them, the on-axis Interactive Neutrino GRID (INGRID) and the off-axis Near Detector at 280m (ND280) are used to characterise the unoscillated neutrino beam and measure its interactions with matter before it travels to the far detector.

5.3.1 INGRID

The T2K experiment uses the INGRID detector to measure the beam profile and rate. This detector is made up of 16 identical modules composed of interleaved layers of iron plates (summing up 7.1 tons of mass per module) and scintillator planes, which measure the particles produced in the interaction of a neutrino with the iron. Each module is surrounded by scintillator planes which veto charged particles coming from outside it. The modules are arranged in a cross-shape, as shown on Figure 5.6b. The cross's centre is centred along the beam axis.



(a) The T2K near detector complex. The ND280 magnet and ECALs are open, showing the inner tracker.

(b) T2K's on-axis neutrino detector: INGRID.

Figure 5.6: T2K's Near detector complex with focus on the INGRID detector [8].

5.3.2 ND280

The off-axis near detector of T2K is called ND280. The z component of the ND280 coordinate system is parallel to the ground; its x and y components are perpendicular to it, in the horizontal and vertical directions, respectively. This detector measures the off-axis neutrino flux and energy spectrum which allows it to put constraints on the combined neutrino flux and cross-section quantities (which are anti-correlated) [2]. ND280 also measures the contamination of electron (anti)neutrino and wrong sign background (e.g. $\bar{\nu}_\mu$ in FHC mode) in the beam. The presence of electron (anti)neutrinos in the beam an irreducible background in T2K's (and, in truth, any long baseline [164, 165]) neutrino oscillation analyses. T2K's 2013 observation of ν_e appearance in a ν_μ beam [150] is a good illustration of this: the irreducible ν_e contamination coming from the beam was predicted to be 3.2 events whilst the total background ($\theta_{13} = 0$) expectation was 4.92 ± 0.55 events. A total of 28 ν_e candidates were observed in the far detector in this analysis.

ND280 is a magnetised 5.6m×6.1m×7.6m (width × height × length) detector. It consists of a tracker region and electromagnetic calorimeters (ECALs) all enclosed within a solenoid magnet. This dipole magnet is re-purposed from the UA1 [166] and NOMAD [167] experiments at CERN. It produces a 0.2 T field (known with a precision of 2×10^{-4}) in the x direction, causing forward-travelling negative (positive) particles to bend downwards (upwards). This is used to infer the particle charge. The gaps between the return yokes of the magnet that are closest to the inner detector are instrumented with scintillator counters. Together, they make up the Side

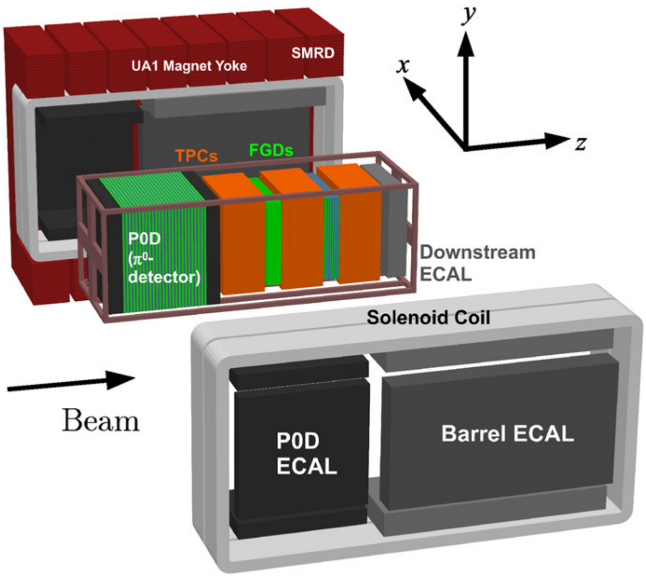


Figure 5.7: Diagram of the ND280 detector composed of the UA1 magnet, the electromagnetic calorimeters and the inner tracker. Reproduced from [2].

Muon Range Detector (SMRD), whose main purpose is to tag muons entering and leaving the ND280 detector [168].

In its pre-upgrade (T2K phase I) configuration, shown in Figure 5.7, the tracker region of ND280 is composed of the π^0 detector (PØD) on its upstream side followed by two Fine Grained Detetors (FGDs) sandwiched between three gaseous Time Projection Chambers (TPCs) in the following arrangement: PØD-TPC1-FGD1-TPC2-FGD2-TPC3.

The tracker is surrounded by three ECALs: the PØD ECAL around the PØD, the barrel ECAL around the tracker and the downstream ECAL placed after TPC3. These calorimeters are made of scintillator material interleaved with lead and provide a quasi-complete coverage of the tracker. They are used to distinguish the track-like (i.e. muons) from shower-like (i.e. electrons, photons) particles. [169].

The PØD is composed of interleaved layers of brass sheets, scintillator planes and thin bags filled either with water or air. This detector is designed to measure neutrino neutral current interactions generating π^0 particles (NCπ^0) [170]. This channel is an important background in neutrino oscillation analyses since it can look like CCν_e in the far detector [171]. The PØD's capability to switch between water and air targets is useful for cross-section comparison; an interesting example using the ν_e component of the beam is discussed in [172].

Both the SMRD, PØD and ECALs are used to identify particles coming from outside the detector, either cosmic-rays, or particles produced by neutrinos interacting upstream of ND280 (these are later referred to as SAND interactions, i.e. happening in the sand before ND280).

5.3.2.1 FGDs

Each of the two ND280 FGDs weighs approximately 1.1 tons, together, they make up the main target material used to measure the unoscillated flux and study neutrino interactions at ND280. Both FGDs are composed of layers of polystyrene scintillator bars of 9.61mm×9.61mm×1864.3mm arranged either along the x or y direction. Each bar has a central hole running through its length and fitted with a wavelength-shifting (WLS) fibre collecting the scintillation light produced in the bar when charged particles cross it. The combination of two layers is called an "XY module", there are fifteen such modules in FGD1. In FGD2 there are only seven XY modules and the six gaps between the modules are filled by bags of water [173].

As well as serving as hydrocarbon (FGD1) and water (FGD2) targets, the FGDs provide particle tracking and some PID information based on the amount of energy deposited by particles in it. Figure 5.8 shows the energy deposited by muons, pions and protons in the FGD1 as a function of the track length, compared with MC prediction. Where possible, the tracks measured in the FGDs are matched by the reconstruction algorithm to those measured in the TPCs. FGD tracks of particles travelling at a high angle compared to the beam axis are matched with ECAL and SMRD tracks. The muon scattering inclusive cross section presented in Chapter 6 uses FGD1 as its target.

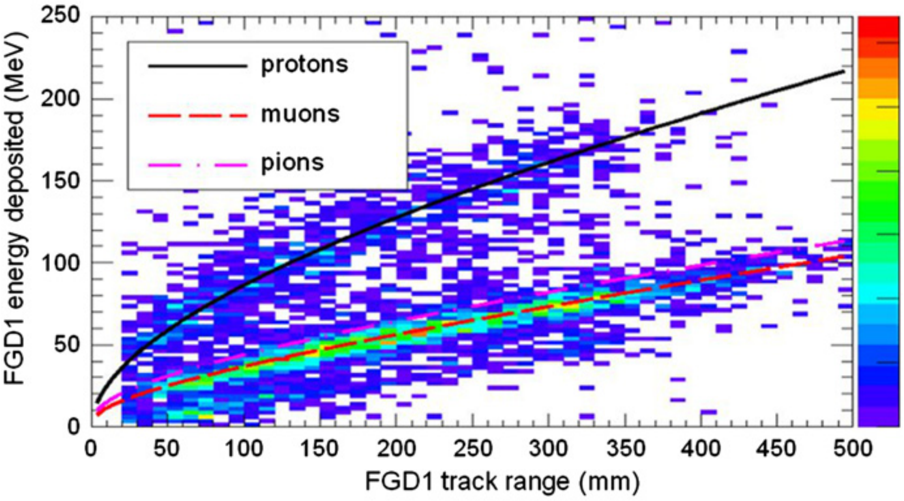
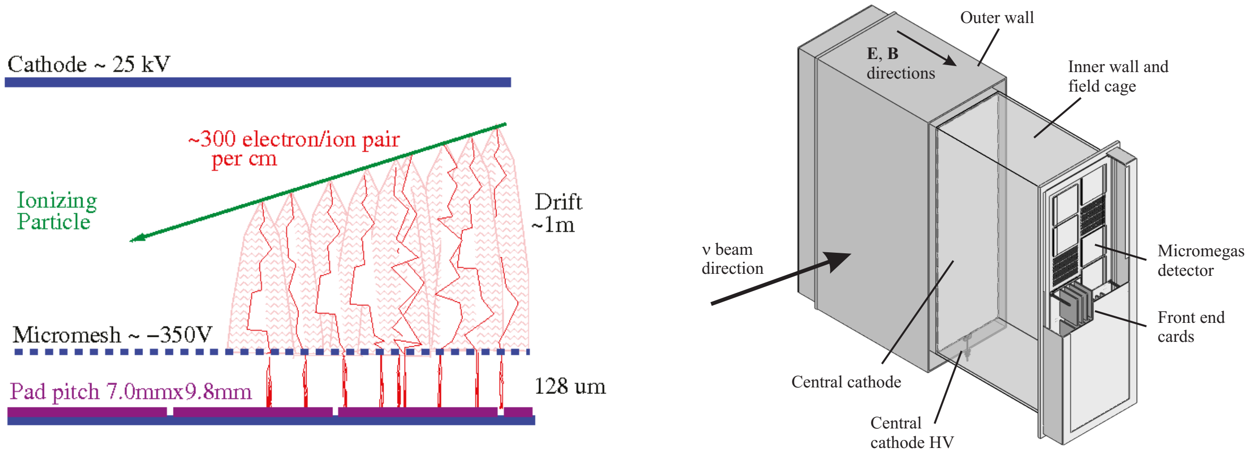


Figure 5.8: Energy deposited by particles stopping in FGD1 as a function of the track length. The MC expectation is shown in the solid lines. The scatter plot shows neutrino beam data. Reproduced from [173].

5.3.2.2 TPCs

There are three time projection chambers (TPCs) interleaved around the two FGDs. These are made up of an outer and an inner airtight "box" whose dimensions are respectively 2302 mm × 2400 mm × 974 mm and 1808 mm × 2230 mm × 854 mm (x × y × z). The inner box is filled with an Ar_{40}:CF_4:C_4H_{10} (95:3:2) gas mixture whilst the gap between the inner and outer boxes is filled with CO_2 [174].



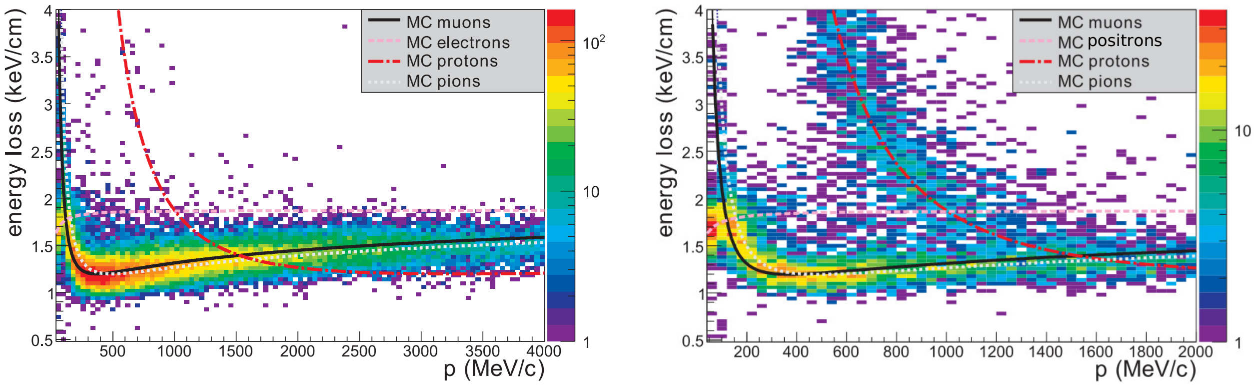
(a) Working principle of the Micromegas, reproduced from [175].

(b) Diagram of the ND280 TPCs, reproduced from [174].

Figure 5.9: Presentation of the ND280 TPCs.

A voltage difference is applied between the inner box's cathode, at −25 kV, located at the centre and the bulk Micromegas modules located on the two sides the TPCs. Their nominal voltage is −350 V. When charged particles cross the inner volume, they ionise the gas mixture. The ionisation electrons drift towards the Micromegas under the influence of the electric field, where they are amplified and read out. This is shown in Figure 5.9a and discussed in [176]. The TPCs provide a 3D reconstruction of the particle's trajectory, using the measured drift time of electrons to estimate the x component of the particle's trajectory.

The particle identification is achieved by comparing the energy deposited in the TPC gas mixture to the expected deposited energy for each particle type. Figure 5.10 shows the energy lost by particles in the gas with MC predictions. The TPC's resolution on the energy deposited by MIPs per unit distance is 7.8% ± 0.2%, which exceeds ND280's requirement of 10%, needed to separate electrons from muons at the 4σ confidence level [174].



(a) Negative particles

(b) Positive particles

Figure 5.10: Energy deposited by various particles in the ND280 TPCs used for particle identification. Figures reproduced from [174].

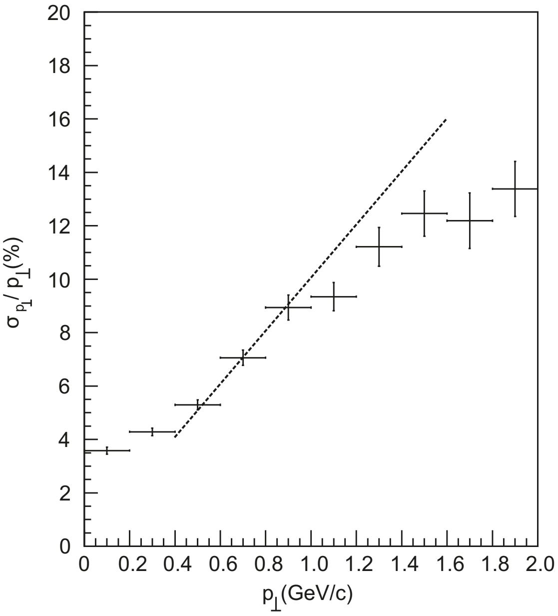


Figure 5.11: TPC momentum resolution for a simulated sample of muons as a function of the momentum perpendicular to the magnetic field. The dashed line indicates the targeted resolution. Reproduced from [174].

The momentum of the particle is estimated based on its curvature in the magnetic field. The momentum resolution of the TPCs as a function of the particle's momentum perpendicular to the beam axis is shown in Figure 5.11 for a sample of simulated muons. As could be expected, the momentum resolution is worse for higher-momentum particles due to their almost linear trajectory in the TPC.

5.4 Super-Kamiokande

The T2K experiment uses the Super-Kamiokande water Cherenkov detector shown in Figure 5.12 as its far detector. It is a cylindrical water Cherenkov detector, 39m in diameter and 42m in height, composed of two regions: an inner detector (ID) and an outer detector (OD) surrounding it. The OD is used to veto particles that are not produced inside the ID.

Super-K's overburden is about 1000 m of rock, corresponding to 2700 meters-water-equivalent, shielding it from most cosmic rays. The whole detector is filled with 50 kilotons of ultra-pure water (doped with gadolinium since 2020 [137, 178]), which is the interaction target for the neutrinos. The walls of the ID are lined with 11,146 inward-facing 20-inch PMTs, together, they provide a 40% photo-coverage of the ID. The OD inner walls are populated with 1,885 outward-facing 8-inch PMTs. A Tyvek and polyethylene layer provides optical shielding between the ID and OD regions.

As charged particles travel through water, they produce Cherenkov light if they are above threshold. The "fuziness" of the Cherenkov light is used to separate muons from shower-like particles (electrons and photons)

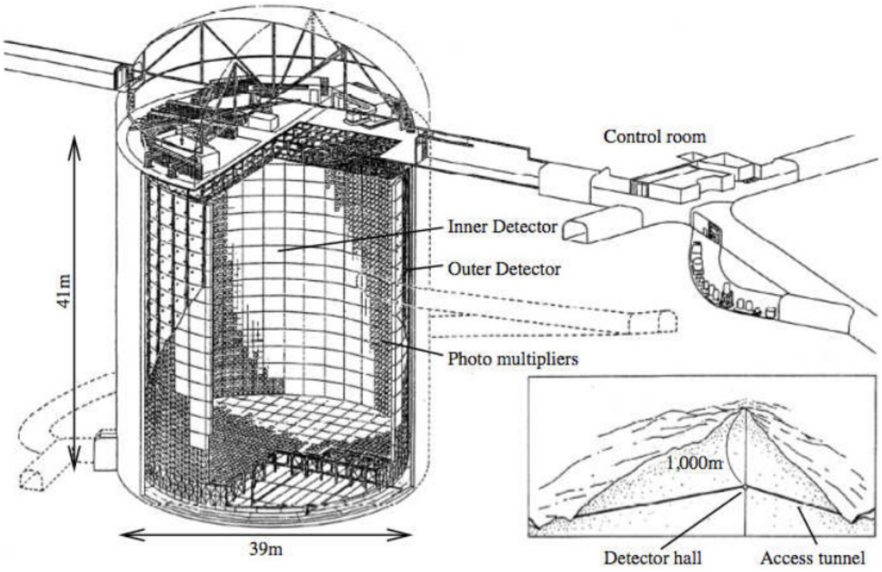
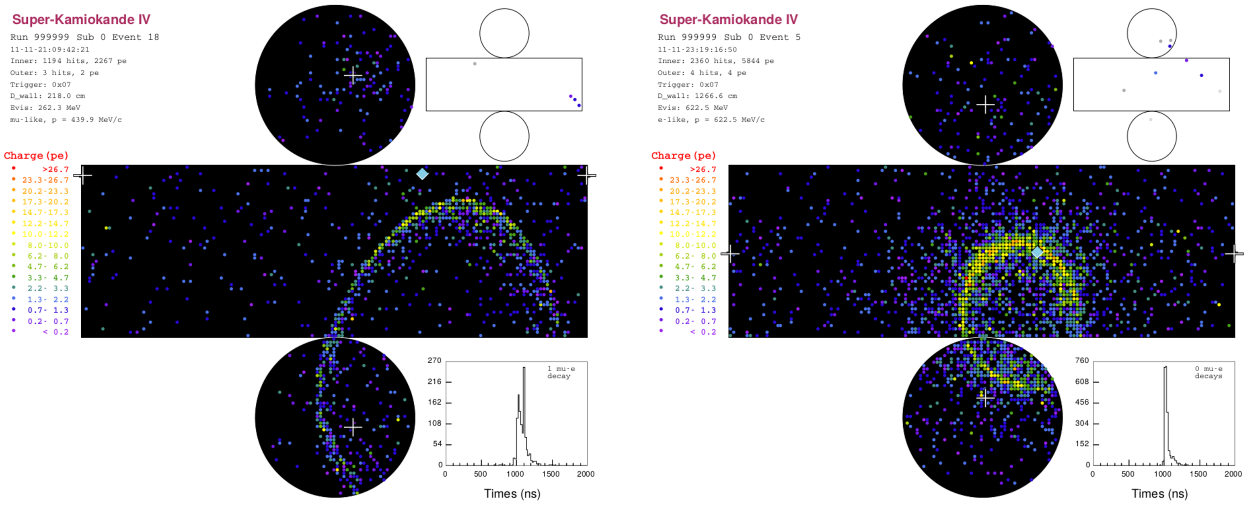


Figure 5.12: Presentation of the Super-Kamiokande detector. Adapted from [177]

.



```
Super-Kamiokande IV
Run 999999 Sub 0 Event 18
11-11-21:09:42:21
Inner: 1194 hits, 2267 pe
Outer: 3 hits, 2 pe
Trigger: 0x07
D_wall: 218.0 cm
Evis: 262.3 MeV
mu-like, p = 439.9 MeV/c
```

```
Super-Kamiokande IV
Run 999999 Sub 0 Event 5
11-11-23:19:16:50
Inner: 2360 hits, 5844 pe
Outer: 4 hits, 4 pe
Trigger: 0x07
D_wall: 1266.6 cm
Evis: 622.5 MeV
e-like, p = 622.5 MeV/c
```

(a) 439.9 MeV muon.

(b) 622.5 MeV electron

Figure 5.13: Example of Super-K simulated event displays for particles produced by neutrino interactions inside the ID which are identified as muons (a) and electrons (b). Figures by the T2K collaboration.

and the hit timings are used to reconstruct the particle's trajectory. The opening of the cone of light is used to infer the trajectory of the particle and the interaction vertex, whilst the total amount of light produced is used to measure its energy. Figure 5.13 shows examples of Cherenkov light rings identified as muon and electron.

By comparing the flavours of leptons produced by (anti-)neutrinos in the near and far detectors, the T2K experiment directly measures the θ_{13}, θ_{23} and δ_{CP} parameters of the PMNS matrix as well as the Δm^2_{32} parameter.

Chapter 6

Muon scattering on Carbon

6.1 Introduction

Since the 1970s, high-energy muons have been used to probe the structure of nuclei and nucleons. The European Muon Collaboration (EMC) notably discovered in 1984 that the quark distribution inside nucleons was affected by the nuclear environment, by looking at muon deep inelastic scattering (DIS) on deuterium and iron atoms [179].

Although high-energy muons have been used multiple times, there is barely any muon scattering data available at sub-GeV and GeV energies. The rare examples include the demonstration by the MICE collaboration of ionisation cooling of 140 MeV/c muons [117] and a study of the scattering of 172 MeV/c muons by Attwood *et. al.* [115] discussed in more detail in Section 6.5.1.3. This latter measurement, covering the range of -0.105 rad to 0.105 ($-6.0°$ to $6.0°$) was used to validate Geant4 models of muon scattering at low energies, as presented in [122].

The work presented in this thesis proposes to fill the experimental gap between $\sim$ 200 MeV/c and multi-GeV muon-carbon scattering data using the T2K ND280 detector. We measure the trajectory of about a million negatively charged muons before and after crossing a hydrocarbon target, namely the FGD1 sub-detector, and from there measure the muon scattering angle and transferred energy. Such a dataset can be used, as described in Section 2.2.1 to validate our modelling of neutrino-nucleus interactions. It can also be useful to tune the Geant4 models of GeV/c muon propagation through matter, helpful for instance in muon cooling, tomography and radiography [117, 180, 181].

Section 6.2 presents how we identify and select well-reconstructed muons which cross the TPC1, FGD1 and TPC2 sub-detectors. It then describes how the information from the two TPCs are used to propagate the trajectory of the muon into FGD1, accounting for energy lost due to ionisation and radiation, and estimate the angle by which the muon was deflected as it crossed the 33 cm-thick FGD1. In this analysis, we assume that at most one scatter took place in the FGD1, treating it as a thin target, essentially combining out the effect of multiple scattering (MS) into a single scatter. A future version of this analysis should treat the detector as a superposition of thin targets instead of a single one, to reduce biases from MS. Since muons are not expected to scatter more than once in FGD1 we deem the present approximation acceptable for this blue-sky analysis.

The analysis presented here uses two different selections, the "entire" sample, containing all muons crossing TPC1, FGD1, and TPC2, used to estimate the total muon scattering cross-section and the "scattered sample" containing only muons which are thought to have scattered inside FGD1. These samples are presented in Section 6.2.2 and 6.2.5 respectively.

In principle, one could study the scattering of muons of any origin (neutrino, cosmic ray, secondary interactions, etc.) with the same experimental method. In practice, requiring that muons cross the three sub-detectors is most easily fulfilled by muons produced in charged current ν_μ interactions upstream of TPC1. These events benefit from having been thoroughly studied and calibrated by the T2K collaboration. This analysis uses 7.78×10^{20} POT of data collected in runs 2, 3, 4 and 8 of T2K data taking. These runs correspond to stable ND280 operation in the FHC mode prior to its upgrade. Because of limited computational resources, the T2K collaboration only simulates a small amount of neutrinos interacting upstream of the ND280 detector (referred to as SAND MC or SAND events), focusing instead on the more interesting neutrino interactions taking place within the ND280 (referred to as ND280 MC). The analysis presented here does not distinguish muons produced inside ND280 from those produced outside it. We therefore expect a significant fraction (see Section 6.2.1 for more details) of the muons selected to have been produced in the sand upstream of ND280. The SAND MC simulation is known to be less accurate than the ND280 and we expect some data-MC discrepancies originating from this sample. We have decided to include SAND event instead of attempting to remove them from the sample to maximise the statistics.

This analysis uses the entire SAND MC simulation computed in FHC mode, which corresponds to 10.4×10^{20} POT. It uses a subset of the ND280 MC simulation corresponding to 23.5×10^{20} POT. For the entirety of the analysis, the ND280 MC is weighted by a factor $f = \mathrm{POT}_{\mathrm{SAND}}/\mathrm{POT}_{\mathrm{ND280}} = 0.44$. This weighting matches the ND280 and SAND simulated POT. Table 6.1 summarises the amount of data and MC used in this analysis. With some minor modifications, we could use the selection procedure presented here to study the interaction of muons in FGD2, which is a hybrid water-hydrocarbon target.

Sample	POT
ND280 MC	23.5×10^{20}
SAND MC	10.4×10^{20}
Data	7.78×10^{20}

Table 6.1: Summary of the data and MC samples used in the muon scattering analysis. Each ND280 MC event is weighted by 0.44 to match the total amount of SAND POT. This analysis uses the entirety of the SAND MC simulated by T2K.

Section 6.3 presents the analysis methods that were used to normalise the Monte-Carlo simulation to the data and the unfolding procedure that was used to factor out detector effects and retrieve true kinematic distributions. Section 6.4 goes on to describe the correction and systematic uncertainties associated with this measurement. Section 6.5 presents the first comparison between data and Geant4 simulation in the 200 MeV/c to 4.5 GeV/c momentum range and Section 6.6 compares the data and MC with predictions of muon scattering made with the experimental version of NEUT described in Section 2.2.2.

6.2 Event selection

The muon scattering analysis selects muon-like tracks that cross TPC1, FGD1 and TPC2. In practice, the selection tries to match together TPC1 and TPC2 tracks corresponding to the same muon if it has crossed FGD1. We define this as an “event”.

This analysis uses two selections, the “entire” selection containing all muons which travel through the three sub-detectors and the “scattered” selection, a subset of the first one, containing muons which have genuinely scattered in FGD1. In practice, the analysis relies exclusively on TPC1 and TPC2 momentum reconstruction and PID information to select events in both samples. As mentioned in Section 5.3.2.2, when charged particles cross the TPCs, they ionise the gas, liberating electrons which are detected by the Micromegas. The ND280 reconstruction software identifies clusters of hits in the TPCs and combines them to form tracks. The algorithm then stitches together tracks reconstructed in all sub-detectors to obtain the entire trajectory of a given particle.

Looking at generator level information, we consider that a pair of reconstructed TPC tracks is correctly selected (i.e. a signal event, s) if all of the following conditions are met:

- The outgoing particle (in TPC2) is the same particle (i.e. same Geant4 particle ID) or the daughter of the incoming particle (in TPC1)
- Both incoming and outgoing tracks are negative muons (Geant4 PDG code 13)
- Both particles cross FGD1

If one or more of these conditions are not met, then the event is categorised as background b. The selection procedure described in the following sections maximises the sample purity (i.e. fraction of true scatters in the sample) without significantly reducing its efficiency. This is done by maximising either of the two following figures of merit (FOMs):

$$\text{SSB} = \frac{s}{\sqrt{s+b}} \tag{6.1}$$

and

$$\text{PPE} = \text{purity} \times \text{purity} \times \text{efficiency} \tag{6.2}$$

depending on the variable of interest. In practice, those two FOMs are not equivalent in principle but give the same optimal point in most of the cases. In case they differ, we choose the FOM which corresponds to the tighter selection.

6.2.1 Entire sample selection

This section describes the selection procedure used to construct the “entire” sample containing muon-like events crossing TPC1, FGD1 and TPC2.

6.2.1.1 Event quality and multiplicity cut

For events to be accepted by the selection, they need to satisfy the global ND280 event quality requirement, which requires that they fall within the designated bunch time windows of the neutrino beam and that the detector is operational (e.g. the UA61 magnet is on). We also require that there is at least one reconstructed track in TPC1 and in TPC2, and that each track has at least 18 hits in the TPC. This requirement is ND280’s common definition of a “good” TPC track [174]. Tracks shorter than 18 hits are more likely to be cutting corners of the TPC and/or be poorly reconstructed.

6.2.1.2 Propagation quality

As will be described in Section 6.4, the particle’s momentum and direction measured by the TPCs are used to estimate the point of closest approach between the two TPC tracks. The propagation of a track forward or backwards from a given point at which the particle type, direction and momentum are known, is done by calculating the trajectory of the particle by taking small steps in the direction of travel, accounting at each step for the particle's curvature in the magnetic field and the energy lost to (or "gained from" when travelling backwards) the material according to the Bethe-Bloch formula for energy loss [146]. The propagation algorithm saves the muon momentum $\mathbf{p}_i$ and position $\{x_i(z_i), y_i(z_i), z_i\}$ at any point i along its trajectory. The step length is $d = 1$ cm for this analysis. This value corresponds to the highest accuracy achievable within the available computing resources. Section 6.2.3 discusses the precision of this propagation. We fit each set of points $x_i(z_i)$, $y_i(z_i)$ with a second-order polynomial to obtain a continuous form for the particle's trajectory as a function of the z coordinate:

$$\begin{cases} \tilde{x}(z) = a_x z^2 + b_x z + c_x \\ \tilde{y}(z) = a_y z^2 + b_y z + c_y \end{cases} \tag{6.3}$$

where the a, b and c coefficients are fitted using the Minuit2 minimiser [182, 183]. The coefficients are unique for each track and each direction of propagation. Because the ND280 magnetic field is in the x direction, the particle trajectories are mostly straight in the x direction, and therefore $a_x \approx 0$.

Given a TPC track i with start point $\mathbf{x}_0^i = (x_0, y_0, z_0)$, start momentum $\mathbf{p}_0^i$, end point $\mathbf{x}_f^i = (x_f, y_f, z_f)$ and end momentum $\mathbf{p}_f^i$, we measure the accuracy of the propagation by comparing $\mathbf{x}_f^i$ with the forward propagated (i.e. with starting position $\mathbf{x}_0^i$ and momentum $\mathbf{p}_0^i$) position of the particle in the $z = z_f$ plane : $\tilde{\mathbf{x}}^i(z_f|\mathbf{x}_0^i, \mathbf{p}_0^i) = (\tilde{x}(z_f), \tilde{y}(z_f), z_f)$. Similarly, we compare $\mathbf{x}_0$ with the backwards propagated position of the particle at $z = z_0$: $\tilde{\mathbf{x}}^i(z_0|\mathbf{x}_f^i, \mathbf{p}_f^i) = (\tilde{x}(z_0), \tilde{y}(z_0), z_0)$. In both MC and data we use the reconstructed start and end point of the tracks to validate the accuracy of the propagation tool. We define the mean propagation quality for a pair of TPC tracks $i \in 1, 2$ as:

$$\text{Mean quality} = \frac{\sqrt{\sum_i \left(\tilde{\mathbf{x}}_0^i - \mathbf{x}_0^i\right)^2 + \left(\tilde{\mathbf{x}}_f^i - \mathbf{x}_f^i\right)^2}}{\sum_i L_i} \tag{6.4}$$

where L_i is the total length of track i.

Importantly, because we are trying to select only muons, we set the particle mass and charge to be those of a negatively charged muon in the propagation for all TPC tracks, regardless of how they have been identified by the reconstruction algorithm. This means that by construction, the propagation quality is worse for particles which are not negative muons. It is expected to be especially poor for electrons and protons whose mass greatly differs from the muon mass and for positively charged particles, which will bend upwards instead of downwards in the magnetic field. Applying an upper cut on the value of the mean propagation quality removes tracks whose momentum has been poorly reconstructed and effectively acts as a first PID cut. Figure 6.1 shows the distribution of signal and background events as a function of the mean propagation quality in the MC, showing that a cut at 5% optimises the PPE FOM. Instead of applying this 5% cut on the mean propagation quality, we apply it to the individual track propagation quality defined as:

$$\text{Track quality} = \frac{\sqrt{(\tilde{\mathbf{x}}_0 - \mathbf{x}_0)^2 + (\tilde{\mathbf{x}}_f - \mathbf{x}_f)^2}}{L} \tag{6.5}$$

which allows us to remove single tracks instead of pairs of tracks, preserving a relatively high efficiency and reducing the risk of accepting a pair of tracks where one is well reconstructed and one is not, which would reduce the purity, in case this latter track is not a muon but for example a pion. The efficiency indicated in the plot is calculated as the fraction of true TPC muon track pairs correctly identified by the selection before the cut on the mean quality (i.e. after the data quality cut and requiring that there is a track in both TPCs). After applying the 5% cut on track propagation quality, the efficiency reduces from 74.1% to 73.8%, a small reduction in efficiency despite the large total number (8%) of TPC pairs removed from the sample by this cut. This is because a large fraction of muon tracks which were previously wrongly paired with poorly reconstructed tracks (corresponding to a background event) are now paired with the correct track (i.e. corresponding to a signal event) after the removal of poorly reconstructed tracks.

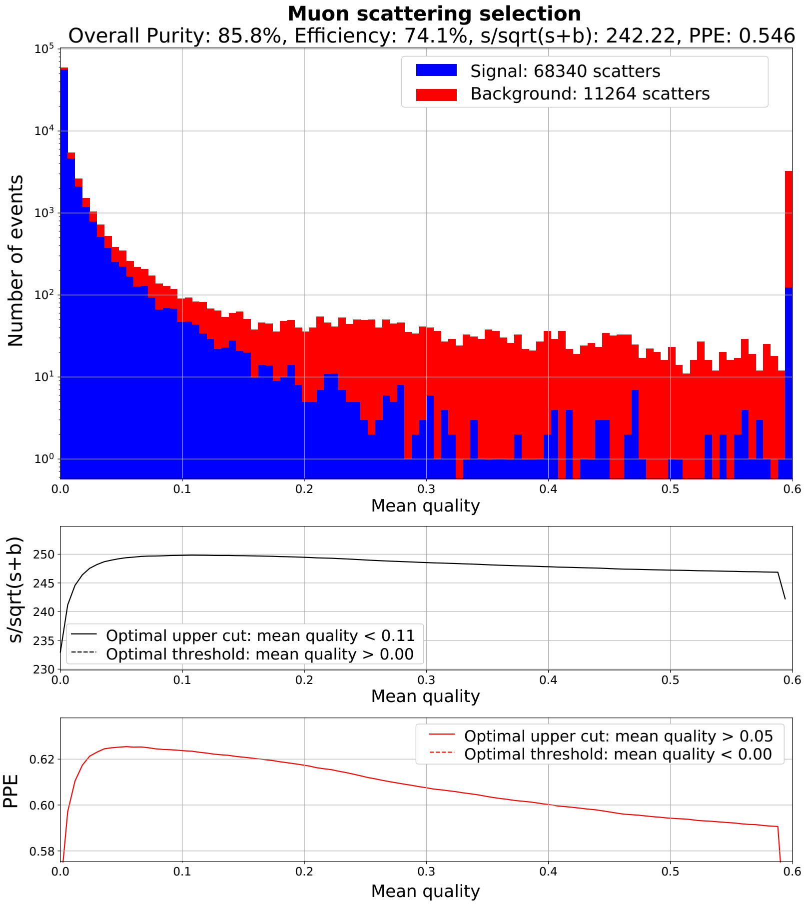


Figure 6.1: Mean propagation quality showing that poorly reconstructed events are mostly background events, as expected from the track propagation method described in the text. The PPE FOM peaks when an upper limit of 5% is applied. The purity and efficiency indicated in the title are prior to applying this cut.

	Pull_μ	Pull_e	Pull_p	Pull_π
Incoming track	[-8.80, 4.40]	[None, 1.00]	[-20, None]	[None, None]
Outgoing track	[-6.60, 4.00]	[None, 8.40]	[None, 11]	[None, None]

Table 6.2: Summary of the range of pulls accepted in the selection for the entire sample. “None” indicates that no upper or lower cut is applied.

6.2.1.3 Particle identification

The particle type is estimated based on the energy it deposits in the TPC. We define a pull as the difference between the observed energy deposition per unit length $dE^{\text{meas.}}/dx$ in the TPC and the expected energy deposited for a given particle i ($dE^{\text{exp.},\ i}/dx$) divided by the energy resolution for this particle type σ_i [184]:

$$\text{Pull}_i = \frac{\frac{dE^{\text{meas.}}}{dx} - \frac{dE^{\text{exp.},\ i}}{dx}}{\sigma_i}. \tag{6.6}$$

We set individual cuts on the value of Pull_μ, Pull_π, Pull_p and Pull_e to maximise the FOMs presented previously. Figure 6.2 shows the distribution of the Pull_μ variable, showing the optimal cut values. The cut values on the pull variables are presented in Table 6.2. We have tuned the value of the cut for the incoming and outgoing TPC segments separately to ensure the highest purity of the selection. We can see on Figure 6.2 that the sample is already very pure (with a purity greater than 93.0%) thanks to the previous selection cuts and that applying cuts on the value of the TPC pulls only marginally improves the purity of the selection. The distribution for the other pulls are presented in Appendix B.

6.2.1.4 Track matching and distance of closest approach

Once the selection has identified one (or more) muon-like, well-propagated TPC1 track(s) and one (or more) muon-like well-propagated TPC2 track(s), it needs to ensure that they correspond to the same muon.

To do so, we propagate the TPC1 track forward ($\tilde{\mathbf{x}}_\rightarrow = (\tilde{x}_\rightarrow, \tilde{y}_\rightarrow, z)$), from its end point, into FGD1 and we propagate the TPC2 track(s) backwards ($\tilde{\mathbf{x}}_\leftarrow = (\tilde{x}_\leftarrow, \tilde{y}_\leftarrow, z)$), from its start point, into FGD1, using the method presented in Section 6.2.1.2. For each TPC1 track, we measure the distance of closest approach (DCA) with each of the TPC2 tracks:

$$\text{DCA}(z_{\text{PCA}}) = \sqrt{\left(\tilde{x}_\rightarrow(z_{\text{PCA}}) - \tilde{x}_\leftarrow(z_{\text{PCA}})\right)^2 + \left(\tilde{y}_\rightarrow(z_{\text{PCA}}) - \tilde{y}_\leftarrow(z_{\text{PCA}})\right)^2} \tag{6.7}$$

where z_{PCA} is the z coordinate of the point of closest approach (PCA) between the two tracks. We identify the PCA as the point within FGD1 closest to both propagated tracks, as illustrated in Figure 6.3.

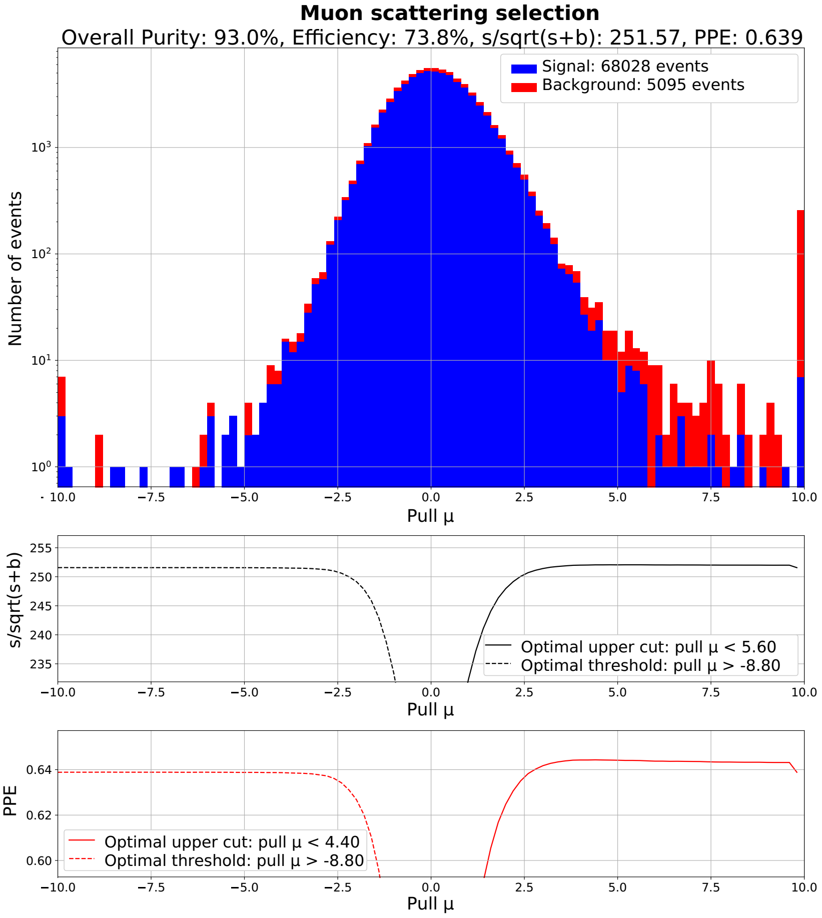


Figure 6.2: Distribution of the Pull$_\mu$ variable showing that events away from Pull$_\mu$ = 0 are mostly background events. The purity and efficiency indicated in the title are prior to applying cut on the pulls.

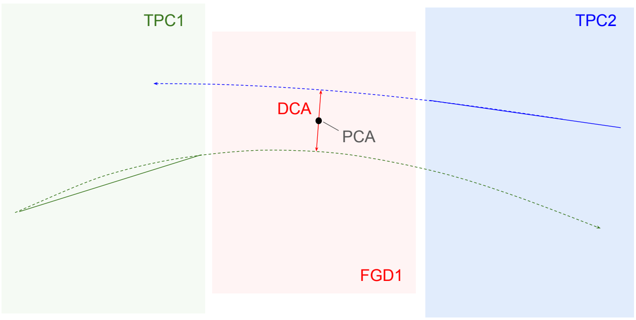


Figure 6.3: Diagram of the matching between two TPC tracks (solid line) showing their propagation (dashed line) into FGD1 and the point and distance of closest approach.

The selection algorithm uses a simple least-squares minimiser [182, 183] to find the value of z_{PCA} which minimises Equation 6.7. We set $z = 281.5$ mm (i.e. the middle of FGD1) as the prior fit value. We bound z_{PCA} to be within 115.620 mm and 447.375 mm (i.e. within FGD1). This bounding of z_{PCA} ensures that the reconstructed

PCA is within FGD1 instead of “anywhere” within the detector, which would otherwise typically happen for un-scattered muons for which the TPC1 and TPC2 tracks are collinear. Because of this, we cannot ensure that tracks cross FGD1 by applying a cut on the position of the PCA. Instead, we require that the DCA between the two tracks is less than 138 mm, which maximises the PPE FOM as shown in Figure 6.4. This DCA cut acts as both a fiducial volume cut and a “track matching quality”. In case there is more than one pair of tracks meeting the above requirement, the algorithm selects the pair(s) of tracks with the smallest total DCA.

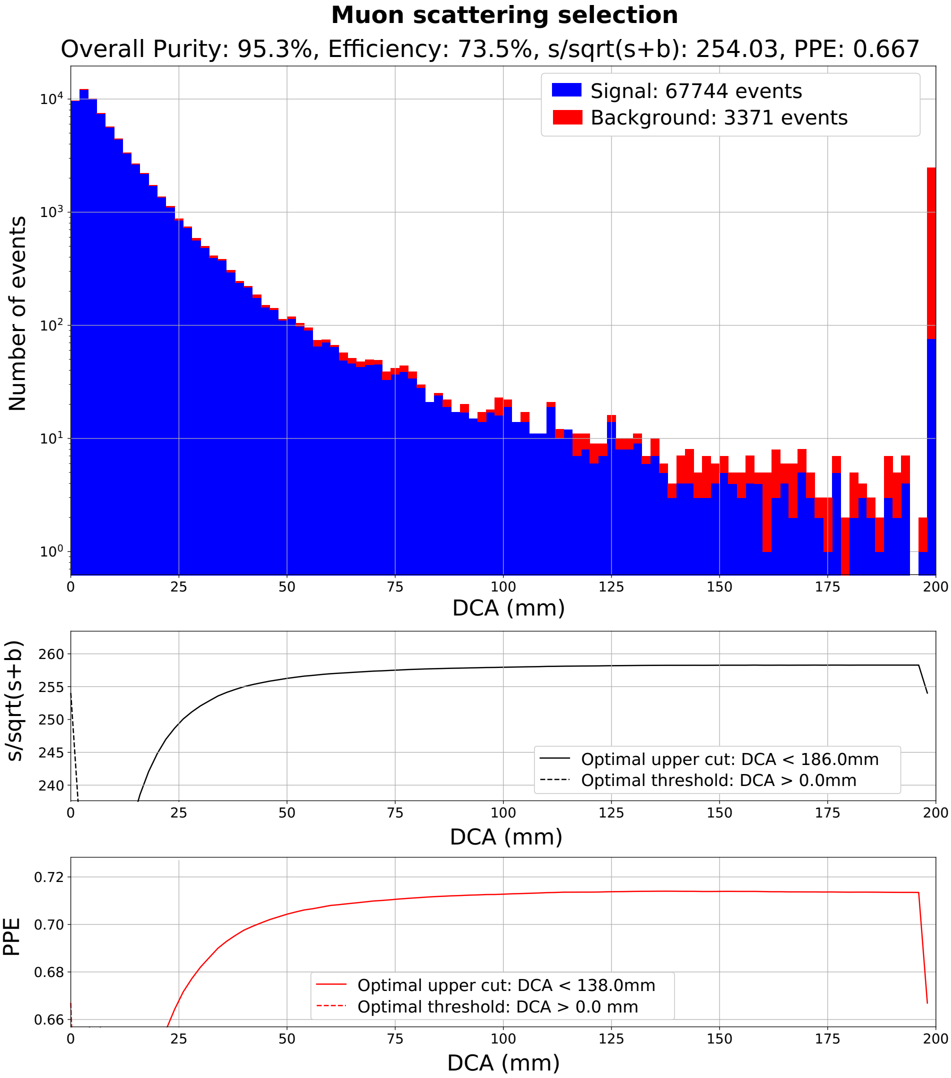


Figure 6.4: Distribution of the distance of closest approach showing that a cut at 138 mm optimises the PPE FOM. The purity and efficiency indicated in the title are prior to applying this cut.

6.2.2 Properties of the through-going sample

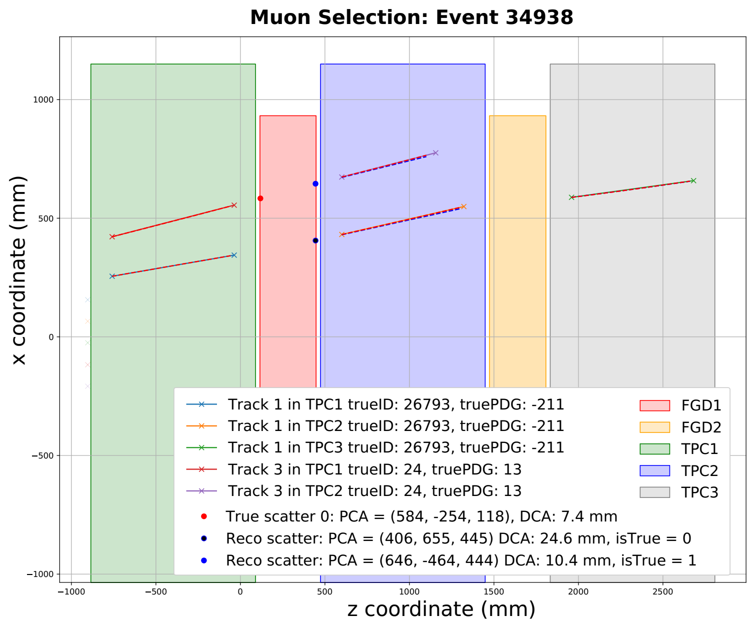
Sections 6.2.1.2 to 6.2.1.4 have described how we match pairs of TPC tracks which look like through-going muons. Figure 6.5 shows an example of an event where a true negative muon (PID = 13) and a true negative pion (PID = -211) crossed TPC1, FGD1 and TPC2. The reconstructed entry and exit points in the TPCs are marked by a cross and the propagated trajectory is shown as a dashed line. The downwards bending of the particles in the magnetic field is visible and well modelled by the propagation tool. We can infer from the trajectories that the muon has a higher momentum than the pion. The true and reconstructed PCAs are indicated in red and blue, respectively. There are no true PCA associated with the pion track since it is a background event. We can see that the true and reconstructed PCAs are respectively at the start and end of FGD1 for the through-going muon. This is because the two TPC segments are collinear (i.e. the muon does not scatter in FGD1), which causes the DCA fitter to reach its boundary value.

The entire selection accepts a total of 937,650 pairs of TPC1 and TPC2 tracks (later on referred to as candidate scatters) in 7.72×10^{20} POT of data and 814,245.6 candidate scatters in 1.0235×10^{21} POT of MC. The selection accepts 34.4% fewer candidate scatters per POT in the MC sample compared to the data (i.e. $795.55/10^{18}$ POT in MC compared to $1214.11/10^{18}$ POT in data). This difference is likely to be due to limitations in the SAND simulation. For example, an underestimation of the amount of material upstream of ND280 would greatly reduce the flux of SAND muons. Whilst this rate difference is not an issue for neutrino cross-section measurements for which through-going muons are an easy background to remove, it has to be addressed in this muon scattering cross-section analysis. This will be discussed in Section 6.3.1.

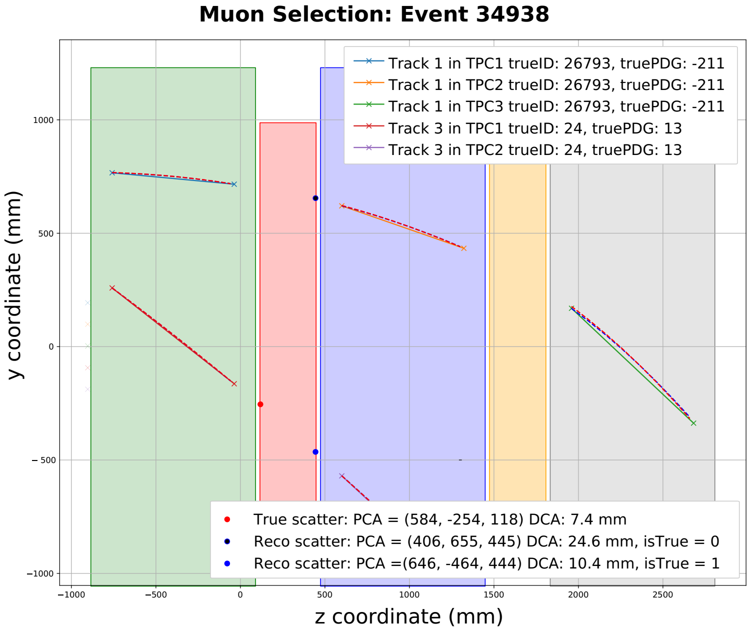
Figures 6.6a and 6.6b show the incoming kinematics for SAND and ND280 events respectively. About 68% of selected MC events belong to the SAND simulation, the remaining 32% are from the ND280 simulation. Most muons are incoming parallel to the z-axis. Low momentum muons typically enter TPC1 at higher angles due to having been bent by the magnetic field prior to reaching TPC1.

In Figure 6.6 we see that SAND and ND280 events have different kinematics, with particles produced in the sand typically incoming at a shallower angle and higher momentum than those produced inside ND280. This is because low-momentum muons produced from neutrinos interacting upstream of ND280 are more likely than higher-momentum ones to be absorbed or decay upstream of TPC1 or be bent away from TPC1. Muons produced in ND280 upstream of TPC1 (mostly in the PØD) and accepted by the selection tend to have a lower mean momentum since they do not have to cross as much material to reach TPC1. We note that there are only few events with momentum lower than about 250 MeV/c. This is because of the through-going requirement, which disfavours particles which significantly bend in the detector. We observe a relatively large number of high-momentum (> 2 GeV/c) events in the ND280 sample. These correspond to the high-energy tail of the neutrino spectrum and mis-reconstructed lower momentum particles.

We estimate the total purity of the entire sample and the contribution from each background using the MC simulation which we assume to be correct. Additional dedicated control studies outside the scope of this thesis would be necessary to assess the limitations of the MC. In this analysis, we split events into 13 categories presented in Table 6.3, one of them is the signal and the others are different types of backgrounds. The ND280



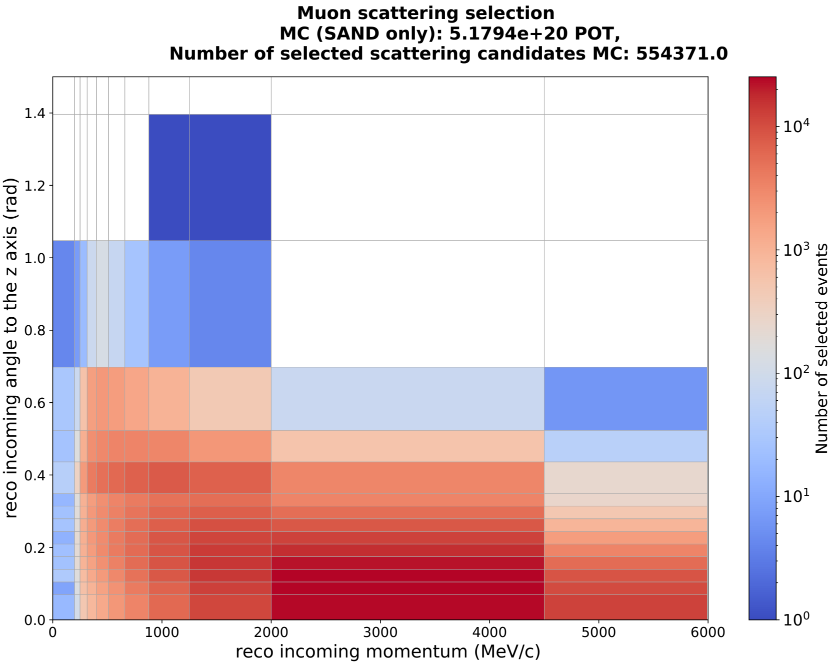
(a) xz projection



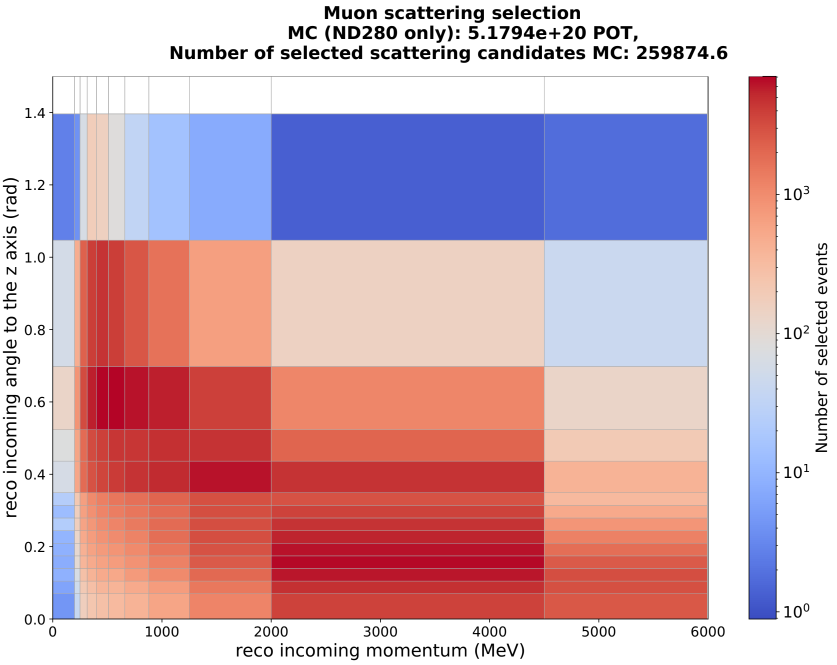
(b) yz projection

Figure 6.5: Visualisation of an event where a negative muon and a negative pion cross TPC1, FGD1 and TPC2, showing the reconstructed and true PCAs in FGD1. We observe that the selection process correctly matches the TPC tracks but misidentifies the pion as a muon.

software only stores true vertex information for neutrino interactions. It is therefore impossible to know *a priori* whether a given charged particle has elastically scattered or not. We use the term “affiliated” to qualify any pairs of TPC tracks corresponding to the same true particle or a true mother-daughter pair, as given by the



(a) SAND MC



(b) ND280 MC

Figure 6.6: Number of events selected per bin of incoming angle and incoming momentum in the through-going selection for SAND and ND280 MC simulations.

Geant4 truth. In Table 6.3, we refer to as "scatter" any events for which the two affiliated tracks also have the same PID.

The different background types can be split into three categories, as indicated on Figure 6.3:

- **reconstruction-induced backgrounds**, where the spatial reconstruction of the track failed (i.e. fake tracks, OOFV)

Catergory	Description
μ^- scatter	Tracks are affiliated, μ^-, truly cross FGD1 (signal)
OOFV	Tracks are affiliated, μ^-, do not cross FGD1
FakeTrack(s)	One or both of the tracks does not correspond to a true particle
π^- scatter	Tracks affiliated, both π^-, cross FGD1
μ^+ scatter	Tracks affiliated, both μ^+, cross FGD1
Other scatter	Tracks affiliated, cross FGD1, same PID$\neq\{\pi^-, \mu^\pm\}$
$\mu^- \to X$	Tracks affiliated, cross FGD1, incoming is μ^- outgoing is not μ^-
$\pi^- \to \mu^-$ decay	Track affiliated, cross FGD1, incoming one is π^-, outgoing is μ^-
$\mu^- \to e^-$ decay	Track affiliated, cross FGD1, incoming one is μ^-, outgoing is e^-
Other Affiliated	All other affiliated tracks crossing FGD1
Accidental μ^-	Tracks are not affiliated, both μ^-
Accidental non-μ^-	Tracks are not affiliated and one or both of them is not a μ^-
Other	Everything else

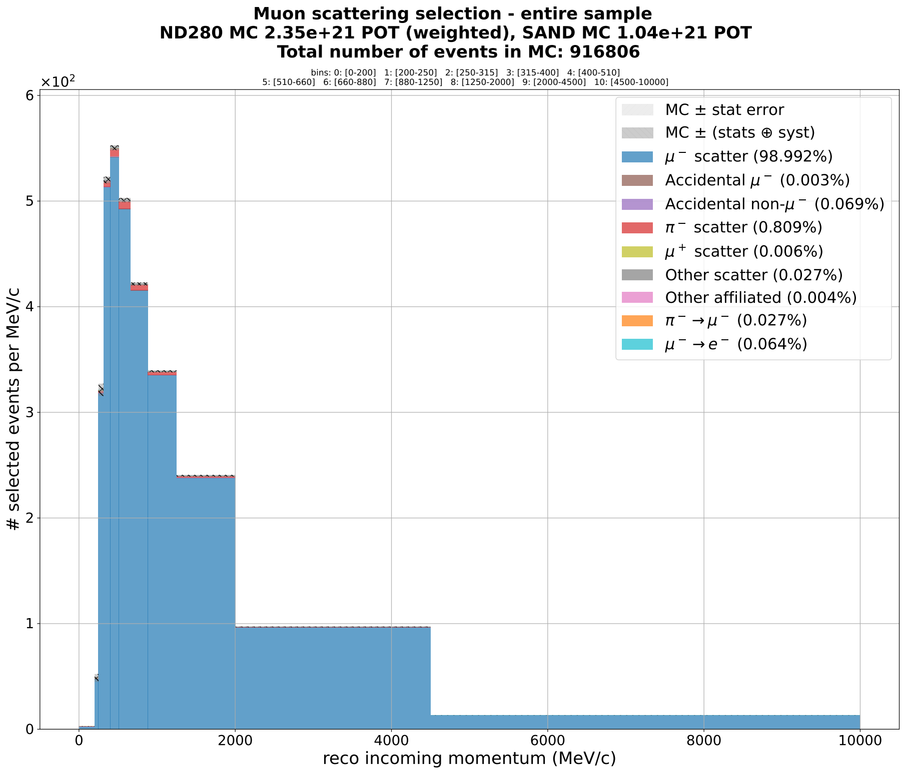
Table 6.3: Event categories, based on truth level information.

- **physical backgrounds**, where the incoming and outgoing tracks are affiliated but not muons
- **accidental backgrounds**, where the incoming and outgoing tracks are incorrectly matched.

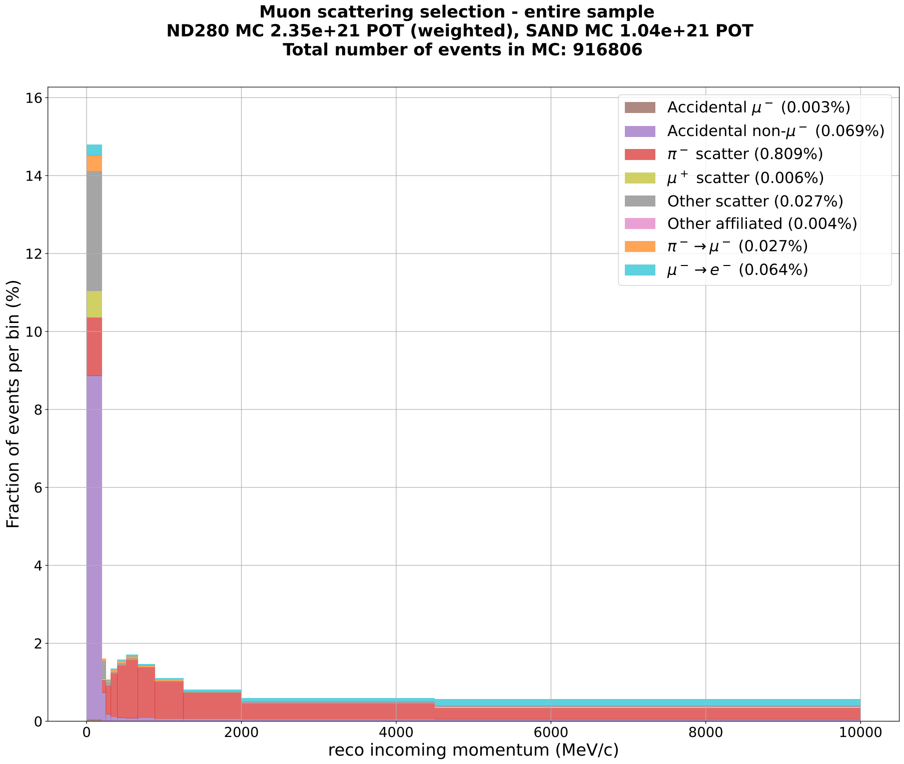
The overall purity of the through-going selection, also referred to as the entire sample, is 98.99%. Figure 6.7a shows the distribution of particle momenta as they exit TPC1 (later referred to as incoming momenta, implied into FGD1). Similarly, the outgoing momentum is defined as the initial particle momentum in TPC2. Figure 6.7b shows the contribution of each background. The selection is extremely pure with very high statistics.

We observe that the first momentum bin, ranging from 0 to 200 MeV, has a purity of 86.4%, much lower than the purities in the other bins of incoming momentum, which are all greater than 98%. Accidental backgrounds dominate in this first incoming momentum bin, whilst they are almost non-existent in higher momentum bins. This is because lower momentum particles curve more and are therefore more likely to cross paths with other tracks and be accepted by the DCA cut presented in Section 6.2.1.4. We decided against applying an additional momentum threshold cut, which would reduce the available phase-space without significant improvement to the selection's purity.

We see that the majority of background events in the selection correspond to genuine physical processes (92.4% of all backgrounds) and the rest of them are accidental backgrounds with almost no cases of reconstruction-induced background events. This highlights the performance of muon track reconstruction by the TPCs. The main background contamination is genuine negative pion scatters, corresponding to 79.9% of all backgrounds in this sample. This is expected since pions look almost identical to muons at these momenta. If we needed to further reduce the pion background, we could require that the outgoing track travels a long distance after leaving TPC2, e.g. reaching the SMRD, which pions, unlike muons, are unlikely to do. This additional selection step is not necessary for this analysis, seeing the low contamination level, but it could be a useful veto when trying to make a pion scattering measurement where muons are expected to be the largest background.



(a) Track momentum at the end of TPC1 for events selected in the through-going sample.



(b) Fraction of background in each incoming momentum bin.

Figure 6.7: Reconstructed incoming momentum for the entire sample: 916,806 events selected with less than 10GeV/c of momentum in 1.04×10^{21} POT of MC SAND and ND280 simulation. The values indicated in legend correspond to the total fraction of each event type in the sample.

We define the scattering angle as the angle between the instantaneous incoming and outgoing muon trajectories at the PCA, as given by the functional forms of the trajectories presented in Equation 6.3. Figure 6.8a shows the distribution of measured scattering angles for the entire sample. We note that the total number of events indicated in the title corresponds to the number of events with scattering angle less than 1.57 rad, not the total number of events selected in the sample. Events with scattering angle larger than 1.57 rad or incoming or outgoing momentum larger than 10 GeV/c are not included in the unfolding procedure presented in Section 6.3.2 nor in the comparisons of Sections 6.5 and 6.6. This cut reduces the range of cross-sections, preventing us from making an absolute cross-section measurement. Instead, this preliminary version of the analysis focusses on making comparisons between the data and MC predictions using the same selection cuts.

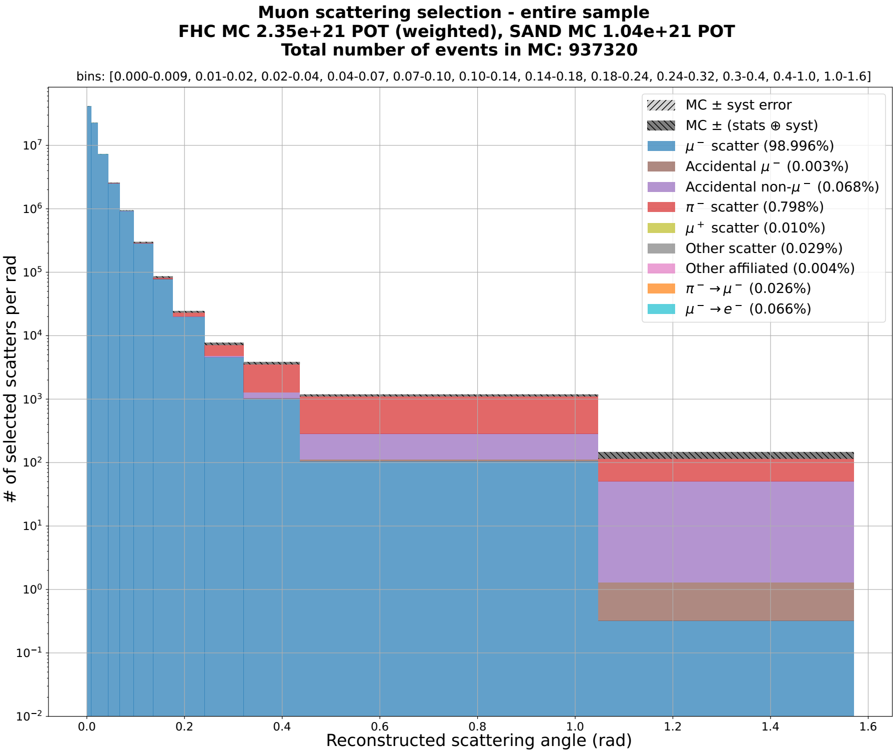
We observe that the majority of muons do not deviate as they travel through FGD1, with a scattering angle close to zero. Figure 6.8b presents the fraction of backgrounds in each scattering angle bin. We note that most events with scattering angles larger than 0.24 rad (13.8°) are background events. Comparing the MC with data in background-dominated bins will be useful to verify the background modelling, which is discussed in Section 6.5. Preliminary studies have estimated that the selection efficiency is on the order of 65%. Due to the large statistics available we have chosen a figure of merit for the selection that prioritises purity above efficiency.

6.2.3 Validation and correction of the propagation method

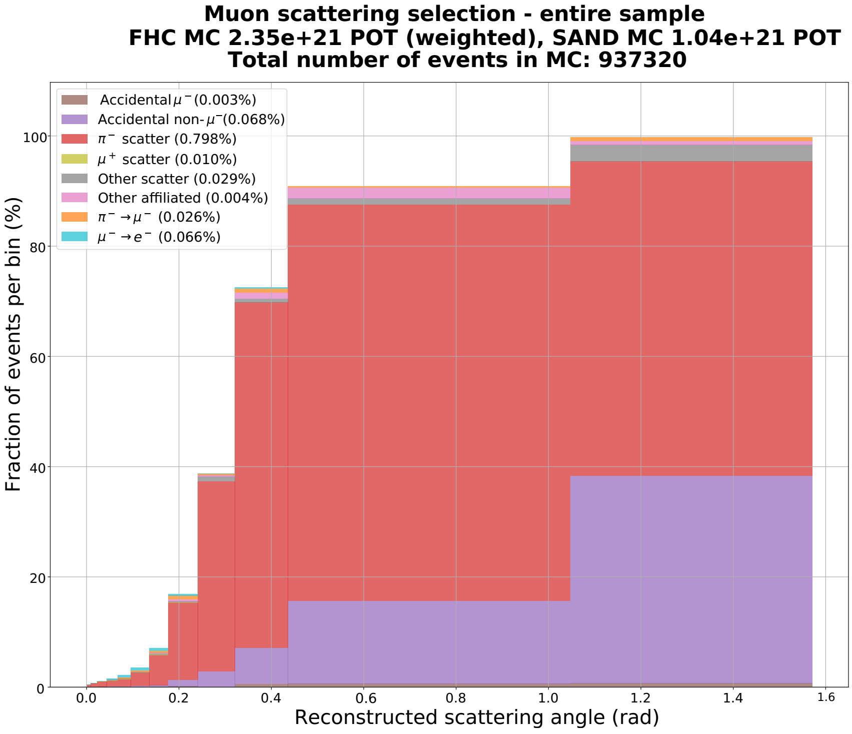
6.2.3.1 Correction to the momentum estimate

Section 6.2.1.2 presented how the TPC tracks are propagated into FGD1 and how we estimate the momentum of the muon at the PCA. To ensure that this estimate of the incoming and outgoing momenta is accurate, we look at a separate MC sample of muons truly produced in FGD1 and measured in TPC2. These muons originate mostly from charged current neutrino interactions, but can also originate from pions decaying inside FGD1.

We compare the generator-level starting momentum of the muon ($|\mathbf{p}_{\text{true}}|$, labelled "true" in this section and the next) with our estimate of the muon momentum at the true muon z starting point ($|\mathbf{p}_{\text{prop.}}|$, labelled "propagated" in this section and the next). This estimate is calculated using the propagation method presented in Section 6.2.1.2 and generator-level TPC2 momentum and direction as opposed to reconstructed ones. Figure 6.9a shows the fractional difference between the true and propagated muon momenta as a function of the true muon momenta. We can see that there is a bias in the propagation algorithm, causing it to overestimate the momentum of the muon at its starting point. This corresponds to an overestimation of the amount of energy lost in the material, since this is a backwards-propagated track. This implies that the incoming momentum, estimated from forward propagating the TPC1 track into FGD1, is likely to be underestimated. Any estimate of the energy transferred at the PCA based on these momenta is therefore likely to be biased.



(a) Distribution of the angle between the propagated TPC1 and TPC2 tracks at the PCA in FGD1 for the entire sample.



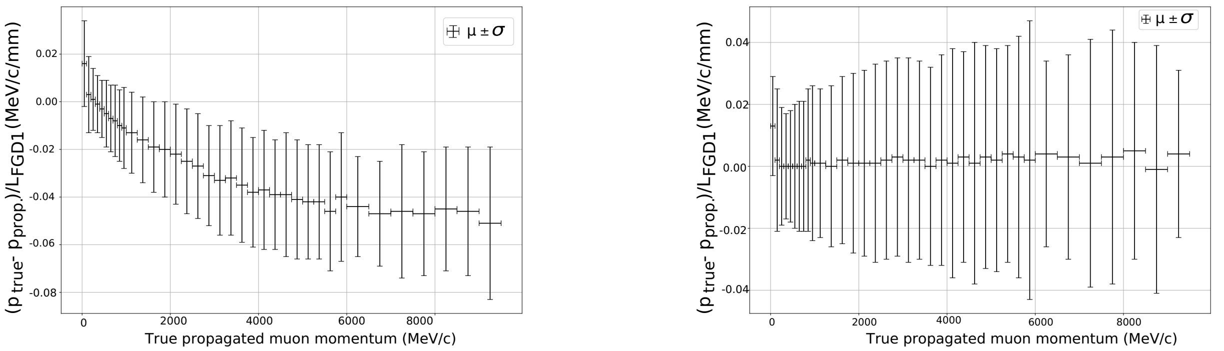
(b) Fraction of background in each scattering angle bin. The values indicated in legend correspond to the total fraction of background of each type in the sample, summing up to 1.004% of the total number of events (i.e. 9,410 background events in total out of which 7,480 are pion scattering events).

Figure 6.8: Scattering for the entire sample: 937,320 events selected with scattering angle smaller than 1.57 rad (90°) in 1.04×10^{21} POT of MC SAND and ND280 simulation.

We use this sample to derive a correction to the momentum estimated using the propagation tool. We can see from Figure 6.9a that this correction factor should depend on the particle's momentum. Assuming that it is a mis-modelling of dE/dx or a mismodelling of the track bending, we expect this correction to be proportional to the total length travelled in FGD1. We define an effective momentum correction factor per unit length travelled in FGD1:

$$C = \frac{|\mathbf{p}_{\text{true}}| - |\mathbf{p}_{\text{prop.}}|}{d_{\text{FGD1}}} \tag{6.8}$$

where d_{FGD1} is the distance travelled in FGD1, as measured using the propagation tool. We see on Figure 6.9a that the difference between the estimated and true momentum can be as large as 0.05 MeV/c/mm for high momentum muons. Considering that particles typically travel L = 330 mm in the FGD1 (i.e. its length along the beam axis) and if we assume that the PCA is in the middle of FGD1, the bias in the incoming muon momentum will be $CL/2 = 0.05 \times 165 = 8.25$ MeV/c and reciprocally, the outgoing muon momentum will be overestimated by the same 8.25 MeV/c leading to an overall bias in energy transfer of up to -16.5 MeV which is large compared to the expected total energy transfer of a few MeV for an elastic scatter. Section 6.2.6.3 describes in more detail how we estimate the energy transferred in a scatter.



(a) Before implementation of the correction.

(b) After implementation of the correction.

Figure 6.9: Difference between the true (i.e. generator-level) and propagated muon momentum divided by the distance travelled in FGD1, shown against the propagated momentum before (a) and after (b) the momentum correction is applied, showing the removal of momentum biases. The bars shown correspond to the 1σ spread of the distribution fitted with a Gaussian in each momentum bin which is significantly greater than the error on the mean.

As of now, it is not possible to know whether this bias is due to an issue with dE/dx estimation or with the estimate of the bending in the magnetic field. It is most likely a combination of both. We implement this correction to the propagated momenta, the result is shown in Figure 6.9b. We estimate that the systematic uncertainty on this correction is small and mostly due to the uncertainty on the FGD1 mass, which is about 1.7% [184]. We ignore this uncertainty in this preliminary version of the analysis, in future versions we will implement a dedicated systematic uncertainty parameter to ensure that the results are not biased.

6.2.3.2 Validation of the PCA estimation

Section 6.2.3.1 showed how a control sample of muons produced in FGD1 is used to correct for biases in the reconstructed muon momentum. We then use a sample of pion decays to evaluate the bias in our reconstruction of the position of the PCA. This sample is constructed (using generator-level information) by requiring that the TPC1 segment is a negative pion and the TPC2 segment is a negative muon, and that they are affiliated and cross FGD1. We compare the true starting point and initial momentum of the muon with the propagated initial momentum and PCA estimated using generator-level TPC information. In this case, we are using the pion mass to propagate the TPC1 segment forward into FGD1 instead of the muon mass. This sample is made from 2×10^{20} POT of ND280 simulation.

Figure 6.10 shows the relative difference between the true and propagated initial muon momentum, as a function of the distance between the true muon starting point and the propagated PCA between the muon and pion tracks. For pions which genuinely decayed in FGD1, a bias of up to 10 cm is observed in the position of the interaction point, despite using generator-level information as the propagation's start point. A maximal momentum bias of up to +8.4% is observed, which corresponds to a 197.7 MeV/c muon. Low-momentum particles are typically less accurately propagated than higher-momentum ones because the fixed stepping length is coarser relative to their bending radius in the magnetic field. In the future, momentum-dependent stepping lengths could be used to reduce this bias.

Muons produced in TPC1 (TPC2) are reconstructed with a momentum lower (higher) than their true value because the PCA is bound to be within FGD1, causing the method to overestimate (underestimate) the amount of energy lost. Figure 6.10 shows that this bias can be as large as 20% of the muon momentum.

To investigate whether this momentum bias is caused by mis-reconstruction of the PCA or by errors in the estimate of the energy lost along the particle's trajectory, Figure 6.11 shows the fractional difference between the true and propagated muon momentum at the true z starting point of the muon. There, we see that irrespective of their starting momentum, the momentum of muons produced in FGD1 from pion decays is within ±2% of its true value. We observe that muons produced in TPC1 have propagated momentum values ranging from +3.7% to −2.2% of the true initial muon momentum, slightly worse than for muons produced in FGD1, probably caused by a poorer modelling of the regions between FGD1 and the TPCs. Because we backwards propagate the muon using generator-level TPC2 information, we expect no biases in the momentum of muons truly produced in TPC2, and indeed, we do not see any.

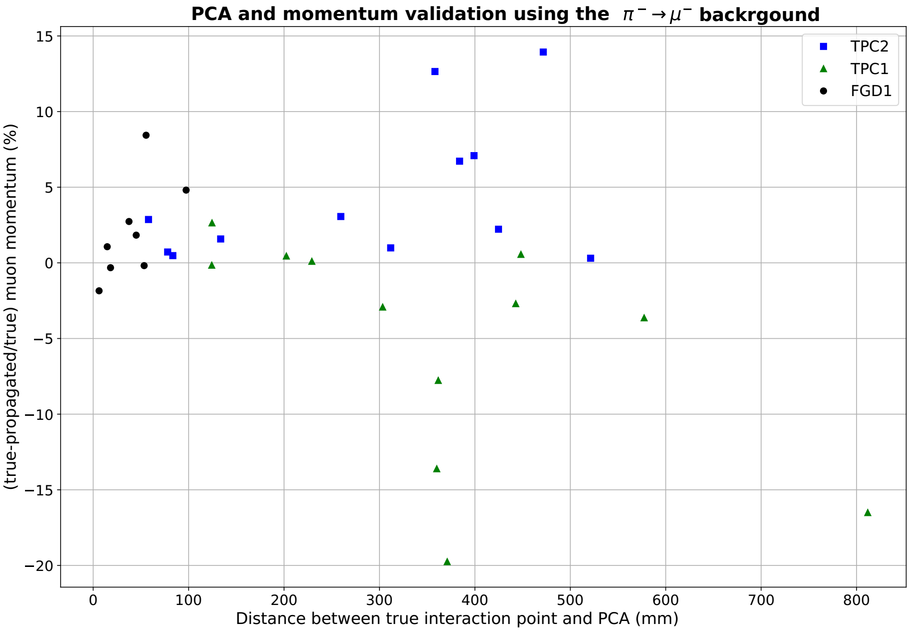


Figure 6.10: Fractional difference between the true and the propagated initial muon momenta at the PCA for muons produced from pion decays in FGD1 (circles), TPC1 (triangles) and TPC2 (squares).

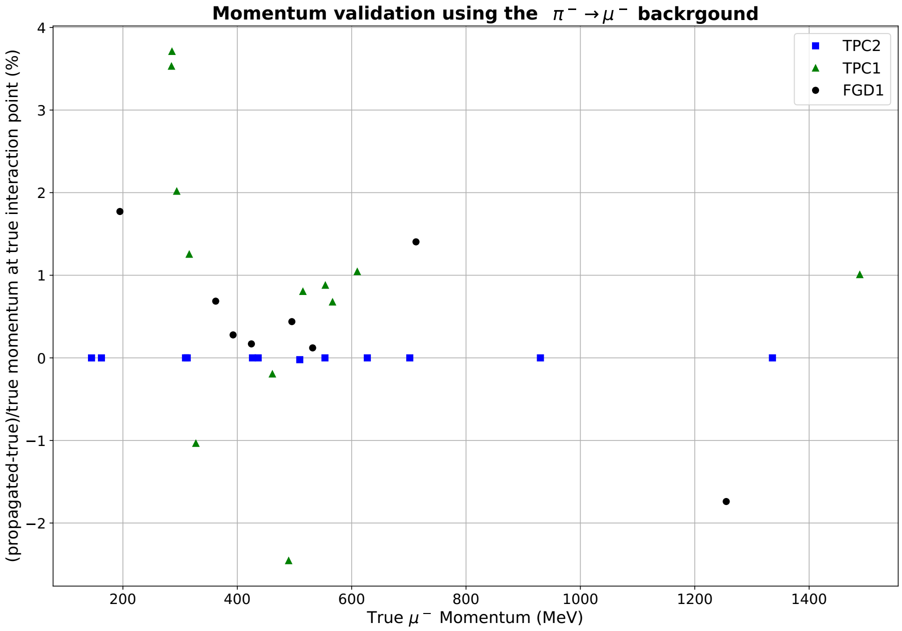


Figure 6.11: Fractional difference between the true and propagated momenta at the true z coordinate of the pion decay shown as a function of the true muon momentum.

Looking at muons originating from pion decays has shown that the propagation method is accurate to within 2% in momentum and that it reconstructs the PCA within about 10 cm of its true value. Because the data and

Monte-Carlo are processed with the same propagation method, we estimate that these biases cancel out when making data-MC comparison. Unfortunately, the final direction of the pion is not saved in the ND280 software, preventing us from estimating the bias on the scattering angle variable coming from the propagation method.

6.2.4 Scattered sample selection

From the entire sample, we identify a subset of events that are likely to be genuine scatters (later referred to as the “scattered sample”). This sample of scattered events can be compared with models of lepton scattering. We apply two additional cuts on top of the ones presented in Section 6.2.1: a fiducial volume cut and a scattering angle cut.

6.2.4.1 Fiducial volume cut

As discussed in Section 6.2.1.4, the estimation of the PCA is bound to be inside the FGD1 volume. Figure 6.12 shows the distribution of z_{PCA} for events in the entire sample, looking at about one percent of the data and MC simulation. Each bin corresponds to one XY module in FGD1.

We observe that about 36% of MC events and 48% of data events have a PCA reconstructed either in the middle of FGD1 (the prior value in the fit) or in the first or last layer of FGD1 (limited by the fit’s bounds). These two cases both correspond to cases where the muon is not scattered within FGD1. Either the fitter does not find a smaller DCA than at its prior value of the PCA (i.e. the muon goes straight through), or the minimal DCA value is outside the FGD1, causing the fitter to reach its boundary value, for example in cases where the muon scatters outside of FGD1 or in case of reconstruction issues. Whether these correspond to events where the muon has not scattered at all or events where the muon scattered outside of FGD1 does not matter for this analysis; neither of them should be included in the scattered sample. We remove them by requiring that the PCA is not reconstructed within 11.5 mm (about one layer) of the FGD1 edges nor within ±1 mm of the middle of FGD1. We note that a higher fraction of data events compared to MC are removed by this cut. We also observe that muons produced in the sand upstream of ND280 are typically more through-going than muons produced inside ND280 but still less than the data. This is described further in Section 6.5.2.

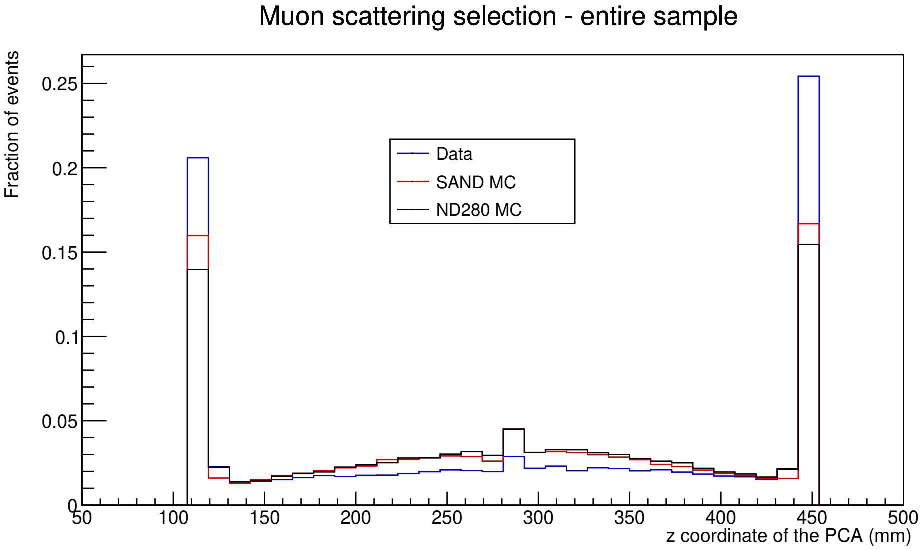


Figure 6.12: z coordinate of the PCA for events in the entire sample showing that about 36% of MC and 48% of data events have a PCA reconstructed in one of the two outer FGD1 layers or the central FGD1 layer. The fraction of events in the central bin is the same for the SAND and ND280 MC samples.

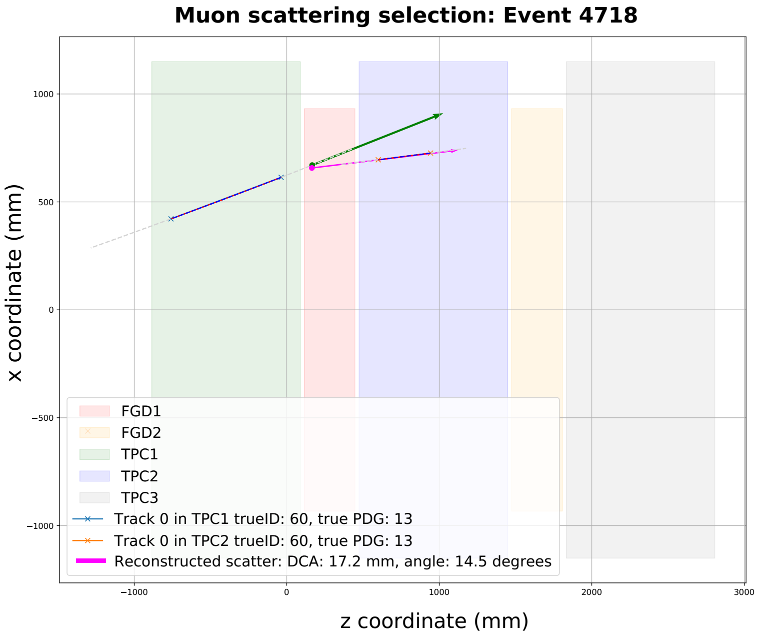
6.2.4.2 Scattering angle cut

To further reduce the contamination of unscattered muons and background events in the scattered sample, we apply a cut based on the value of the scattering angle θ. We require that $\theta > 1.3°$ to remove through-going events and $\theta < 0.24$ rad (13.8°) to remove the last four bins in Figure 6.8, which are background-dominated.

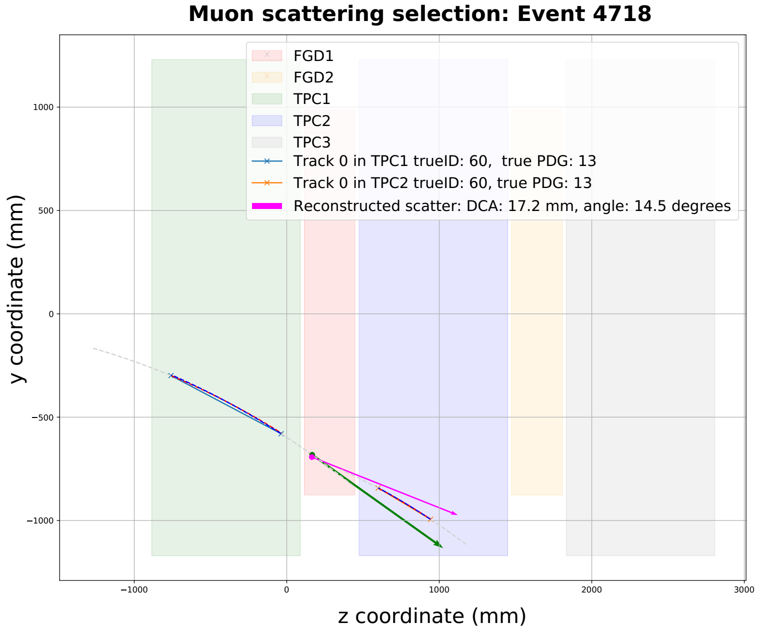
6.2.5 Properties of the scattered sample

As described above, the additional cuts applied to the scatter sample are designed to remove most unscattered events, they however restrict the sample in a way that is not unproduceable by other experiments. Future versions of the analysis will focus on moving away from these experiment-specific cuts and use more standard selection procedures (e.g. based on the kinematics and multiplicity of particles in the final state). Figure 6.13 shows an example of a true muon identified as having scattered. The DCA is 17.2 mm, the reconstructed scattering angle is 14.5°, and the incoming and outgoing tracks correspond to the same true muon.

When applied to the available 7.73×10^{20} POT of data, the scattered muon selection accepts 194,411 events (i.e. 20.7% of the entire sample). Applied to 1.04×10^{21} POT of MC, it accepts 195,566.8±3,324 events events (i.e. 24.0±0.4% of the entire sample); 58.2% of which are from the SAND simulation and 41.6% from the ND280 simulation. The data-MC rate discrepancy in the scattered sample is 25.0% (251.7 selected scatters per 10^{18} POT of data compared to 188.8±3.2 selected scatters per 10^{18} POT of MC), compared to 34.4% for the entire sample. This is discussed in greater detail in Section 6.5.2. The error estimation is presented in Section 6.4.



(a) xz projection

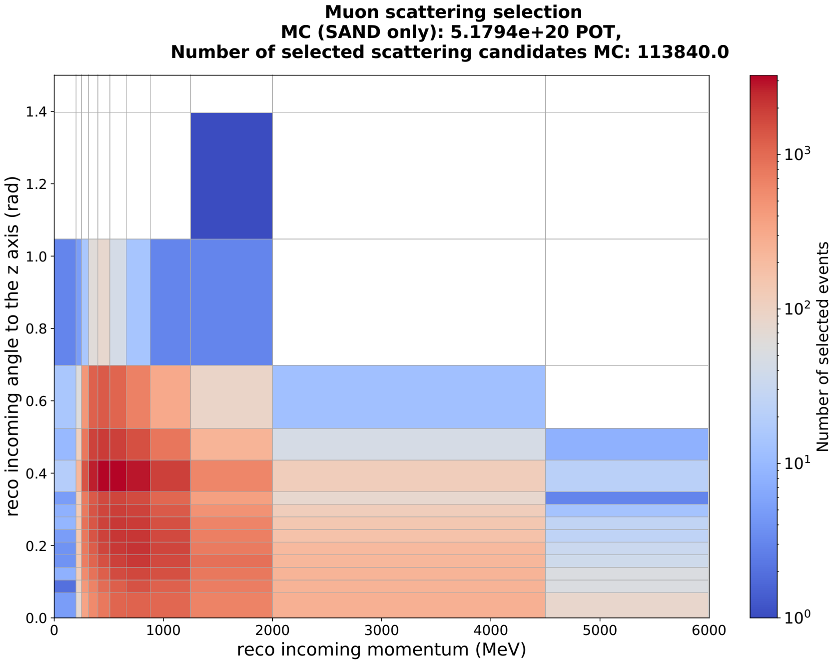


(b) yz projection

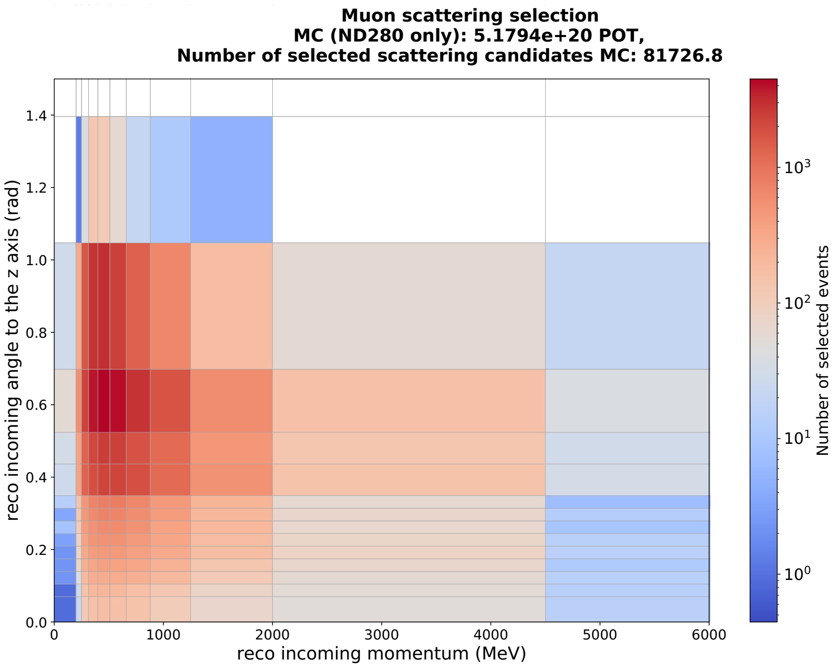
Figure 6.13: Visualisation of an event where a negative muon crosses TPC1, FGD1 and TPC2. The green and pink arrows indicate the instantaneous incoming and outgoing direction at the PCA. We reconstruct a scattering angle of 14.5°.

Figure 6.14 shows the initial momentum and angle for particles in the scattered sample.

We observe that the initial momentum distribution of SAND and ND280 MC events is more similar to each other than they were in the entire sample, reflecting the selection of genuine scatters, which typically occur



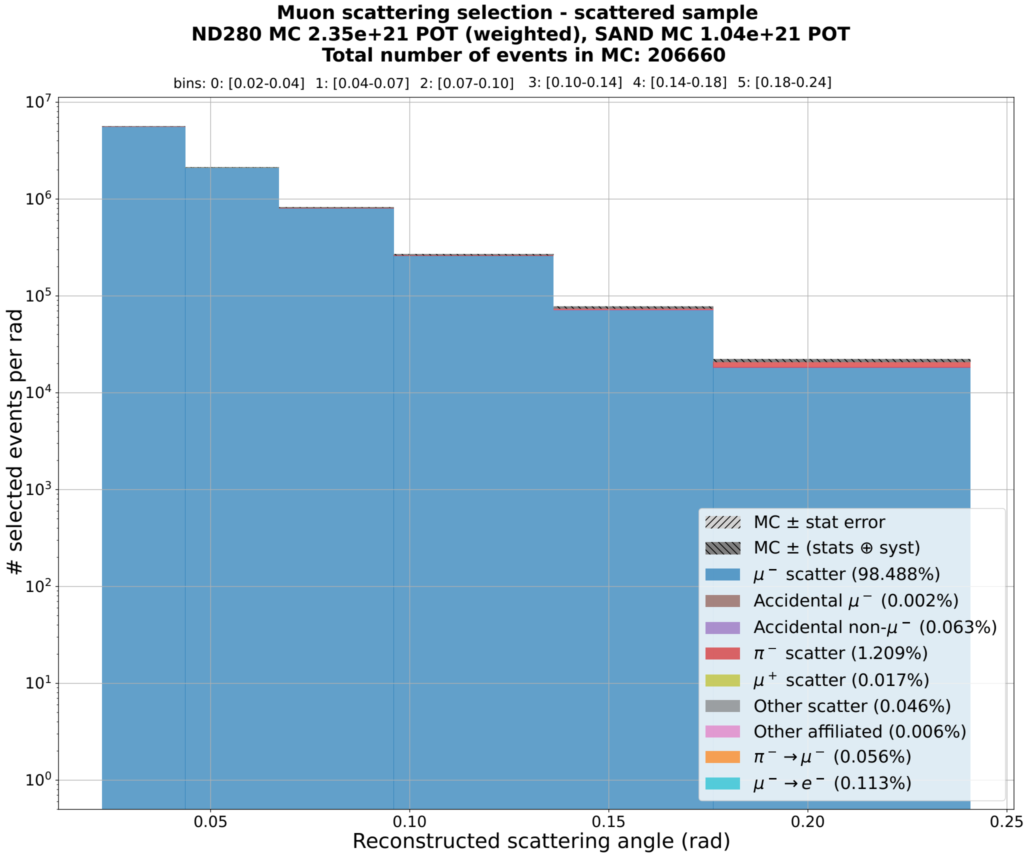
(a) SAND MC



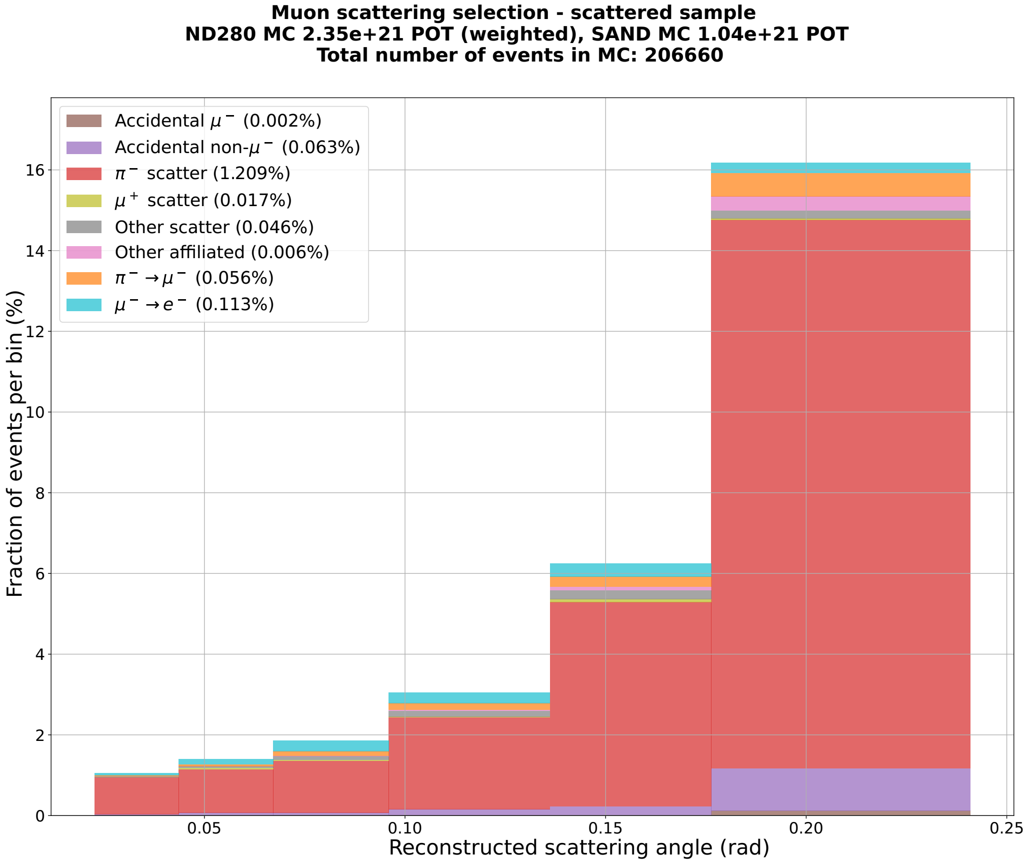
(b) ND280 MC

Figure 6.14: Number of events selected per bin of incoming angle and momentum in the scattered sample for SAND (top) and ND280 (bottom) MC simulations.

more at lower momenta. This preference for low-momentum particles explains why there are fractionally fewer SAND events in the scattered sample than in the entire sample.



(a) Distribution of reconstructed scattering angle in the scattered sample.



(b) Fraction of background in each incoming scattering angle bin for the scattered sample.

Figure 6.15: Reconstructed scattering angle for the scattered sample: 206,660 events selected out of 1.04×10^{21} POT of SAND and ND280 MC simulations.

Using the MC sample, we estimate that the purity of the scattered sample is 98.49%, slightly lower than that of the entire sample. Figure 6.15 shows the distribution of scattered angle for this sample. Once again, pion scattering is the largest contamination to the sample, making up 79.8% of all backgrounds, followed by muon decay to electron, corresponding to 7.4% of all backgrounds. Because the ND280 software does not store information about charged lepton scattering, it is impossible to measure the efficiency of the scattered sample selection.

6.2.6 Resolution and binning

We measure the detector resolution on a given observable by comparing (when available) its generator-level value (labelled true) with its reconstructed one (labelled reco). No generator-level information is available for the scattering angle and energy transferred variables. Instead, we use a proxy for the true value of these parameters calculated using the method described in Section 6.2.1.2 but based on the generator-level TPC observables (i.e. particle momentum and direction) instead of the reconstructed ones. In this section, we discuss the detector resolution on each of the variables of interest. We define the binning for variable x by requiring that at least 2σ (95.44%) of the events in true bin i ($x^i_{\text{min}} \leq x_{\text{true}} < x^i_{\text{max}}$) are reconstructed in the same bin ($x^i_{\text{min}} \leq x_{\text{reco}} < x^i_{\text{max}}$).

6.2.6.1 Incoming and outgoing momenta

The only variables used in this analysis for which the generator-level information is directly available in HighLAND are the particle's momentum inside TPC1 and TPC2. Figure 6.16 shows the distribution of reconstructed momentum for each bin of true incoming momentum in TPC1.

We observe that the differences between the true and reconstructed values follows a Gaussian distribution centred around zero, meaning that there are no biases in the momentum reconstruction. We have observed that TPC1 and TPC2 perform almost identically, in line with expectations [174]. We therefore use the same bins for incoming and outgoing momenta, shown in Figure 6.17.

6.2.6.2 Scattering angle

As described above, to estimate the detector resolution on the scattering angle, we compare the reconstructed value (θ_{reco}) with the one calculated using the same propagation method but based on generator-level TPC information (θ_{true}). Figure 6.18 shows the distribution of $\theta_{\text{reco}} - \theta_{\text{true}}$ in bins of θ_{true}.

We see on Figure 6.18 and Figure 6.19 that the true bins contain at least 2σ of the reconstructed values. We limit the range of the scattering angle bins to 1.57 rad (90°) as anything larger than this corresponds to particles

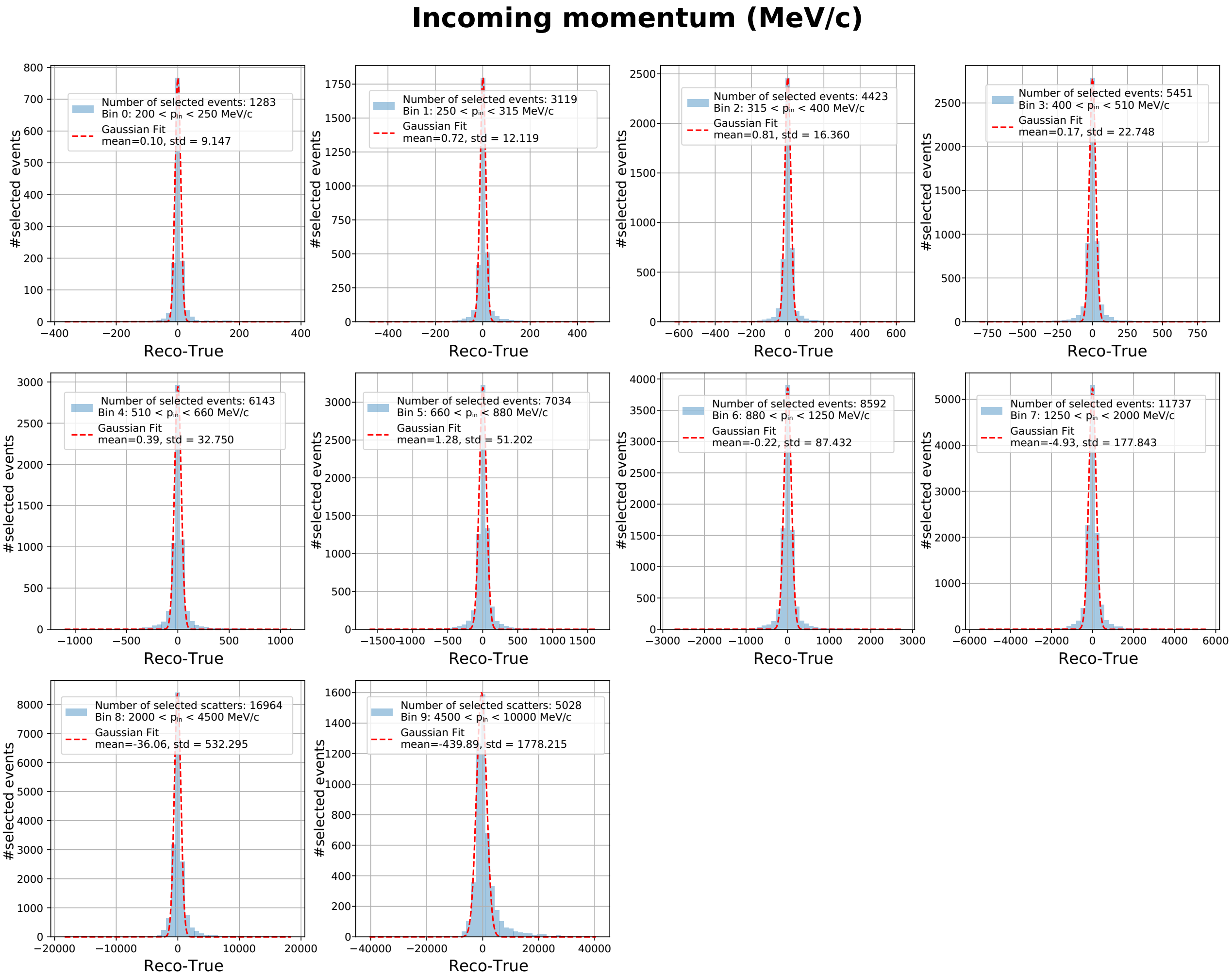


Figure 6.16: Difference between the reconstructed and true particle incoming momentum in bins of true momentum.

likely to cut corners of one (or both) of the TPCs and/or be poorly reconstructed.

6.2.6.3 Transferred energy

When looking at scattered particles, it is useful to estimate the amount of energy transferred in the scatter, as discussed in Chapter 2. In that chapter, we have also seen that the amount of energy transferred (q_0) in quasi-elastic and elastic scatters can be very small, on the order of tens of MeV or less. Knowing from Chapter 2 and Section 5.3.2.2 above that the ND280 TPCs momentum resolution is roughly 7.8%, we therefore have an uncertainty of about 50 MeV on the value of the incoming and outgoing momenta for an 650 MeV/c muon. Ignoring any uncertainty in the dE/dx that might arise from the Bethe-Bloch formula (other studies have shown that these errors are about 1% or less at these momenta [146]), there is still as $\sqrt{2} \times 50 = 72$ MeV/c uncertainty on the difference between the incoming and outgoing muon momenta, which is much larger than the expected energy transfer.

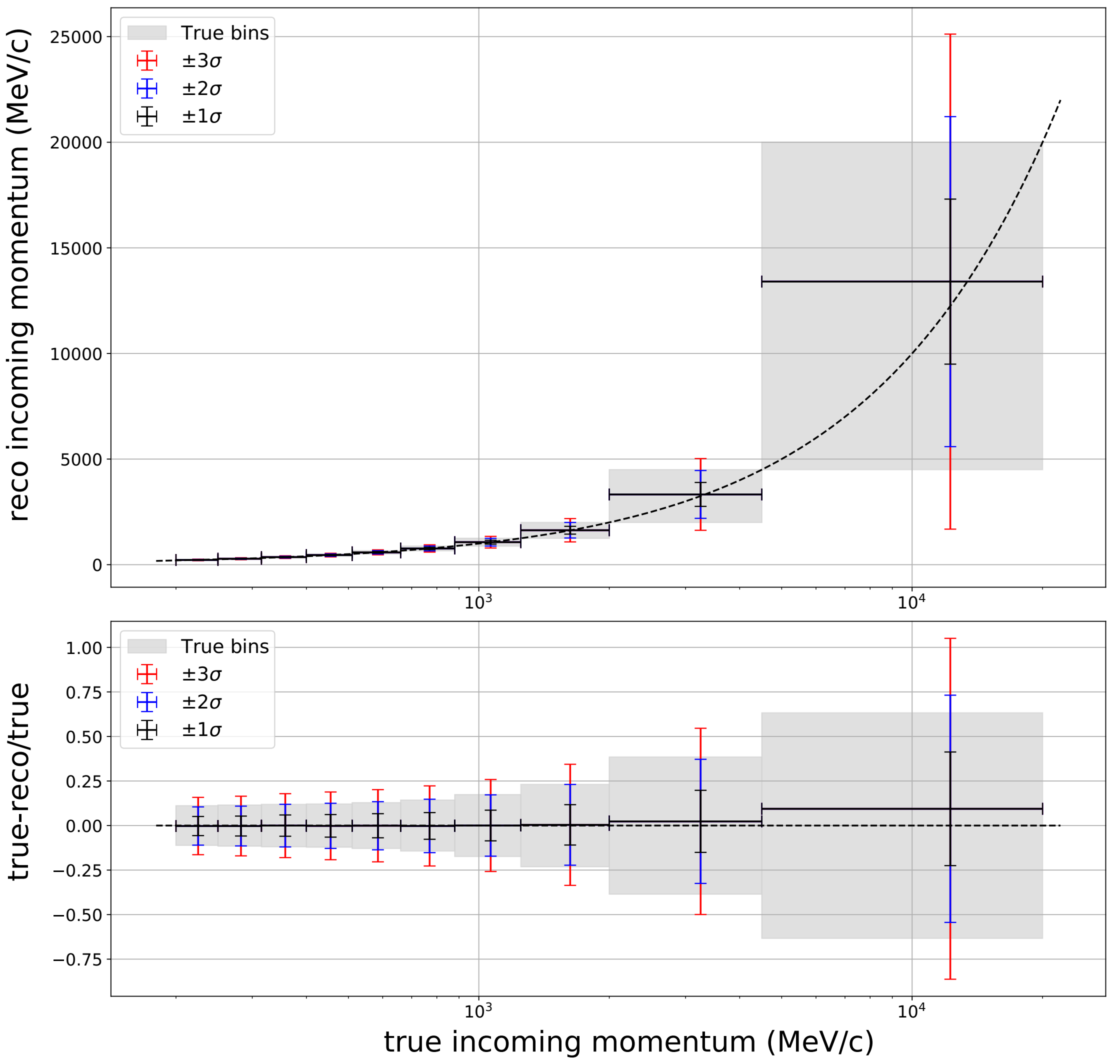


Figure 6.17: Bins of incoming momentum showing that 2σ of reconstructed momenta fall in the correct true bin. The dashed line corresponds to true = reco.

Instead, we estimate the amount of energy transferred using the binned distributions of incoming and outgoing momenta at the PCA ($p_{\text{in}}^{@\text{ PCA}}$ and $p_{\text{out}}^{@\text{ PCA}}$ respectively), as obtained from the propagation method. To illustrate, let us take an event for which $p_{\text{in}}^{@\text{ PCA}}$ is in bin i ($p_{\text{min}}^{i} \leq p_{\text{in}}^{@\text{ PCA}} < p_{\text{max}}^{i}$) and $p_{\text{out}}^{@\text{ PCA}}$ is in bin j ($p_{\text{min}}^{j} \leq p_{\text{out}}^{@\text{ PCA}} < p_{\text{max}}^{j}$). We estimate that the momentum transferred in this event is:

$$q_0 \approx E_i - E_j \tag{6.9}$$

where E_i is the mean muon energy in momentum bin i calculated as:

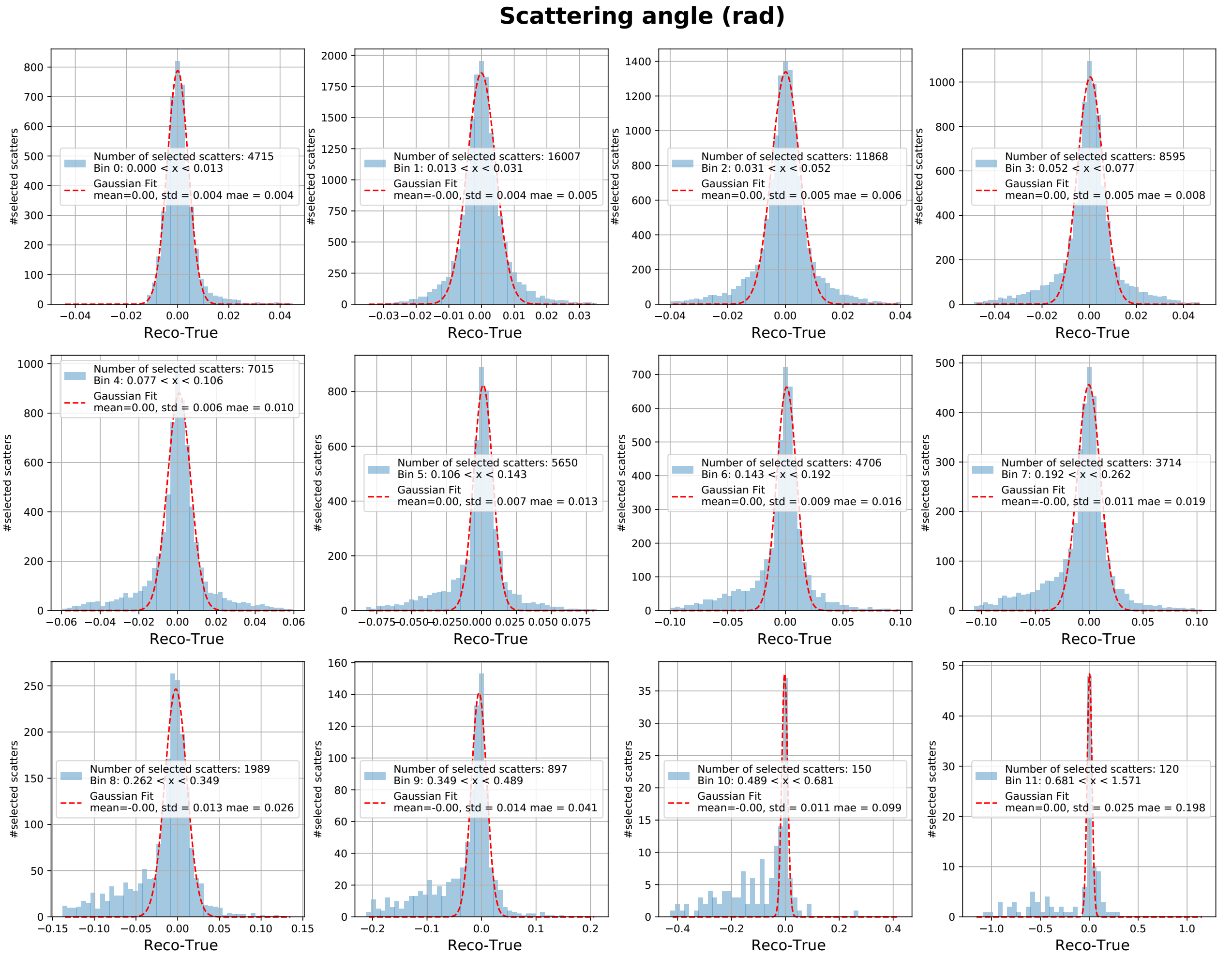


Figure 6.18: Difference between the reconstructed and true scattering angle in bins of true scattering angle. The distributions are fitted with a Gaussian.

$$E_i = \sqrt{\left(\frac{p^i_{\max} + p^i_{\min}}{2}\right)^2 c^2 + m_\mu^2 c^4} \tag{6.10}$$

.

According to this definition, all events for which the incoming and outgoing momenta lie in the same bin correspond to $q_0 = 0$ MeV. Figure 6.20 shows the distribution of reconstructed transferred energies for the entire sample (including unscattered particles) in panels of incoming momentum. Bins of negative q_0 mostly correspond to mis-reconstructions which cause the outgoing momentum to be larger than the incoming one, as will be further discussed in Section 6.3.2. We observe that for the large majority of events, no energy is transferred.

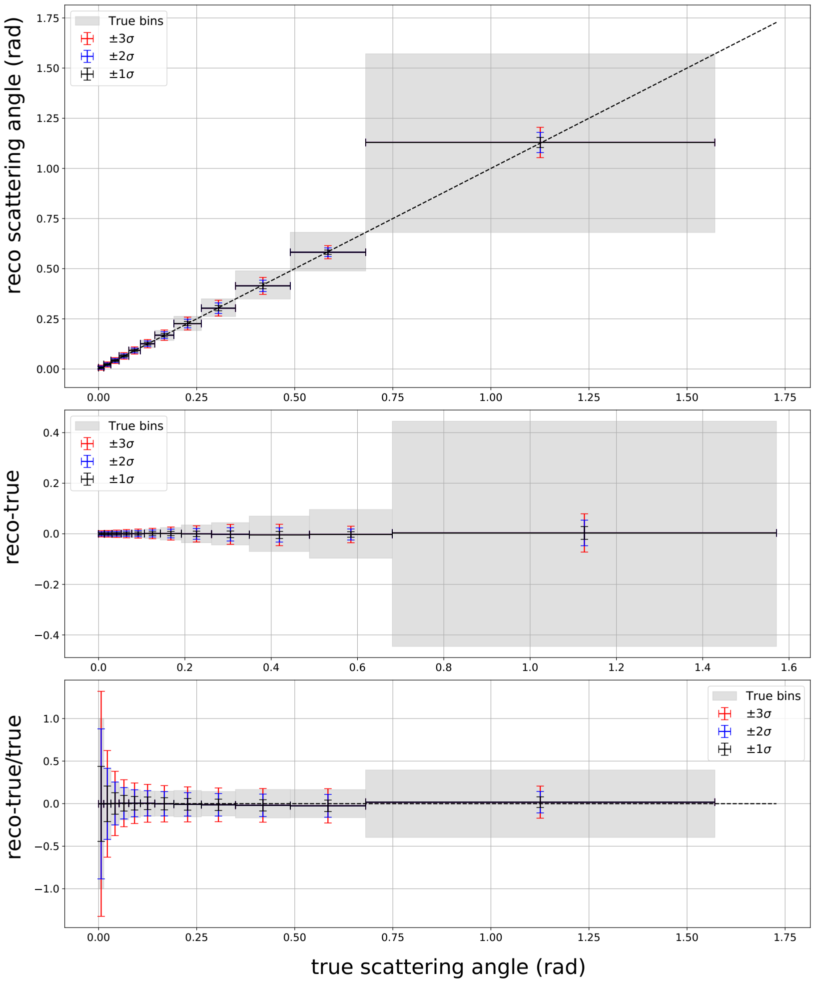


Figure 6.19: Detector resolution on the scattering angle variable. The dashed line corresponds to true = reco.

6.2.7 Selection summary

Tables 6.4 and 6.5 respectively present the purity p_i in each momentum and scattering angle bin for both samples. These values are slightly different from what is presented in the legend of plots above, as the normalisation procedure described in Section 6.3.1 is not applied in this latter case. We observe that the bin purity is roughly independent of momentum except for the first incoming momentum bin, which has a lower purity in both samples and for the last two bins in both incoming and outgoing momenta. In the scattered sample, the purity decreases with scattering angle with at 1% (16%) of events being backgrounds in the first (last) scattering angle bin.

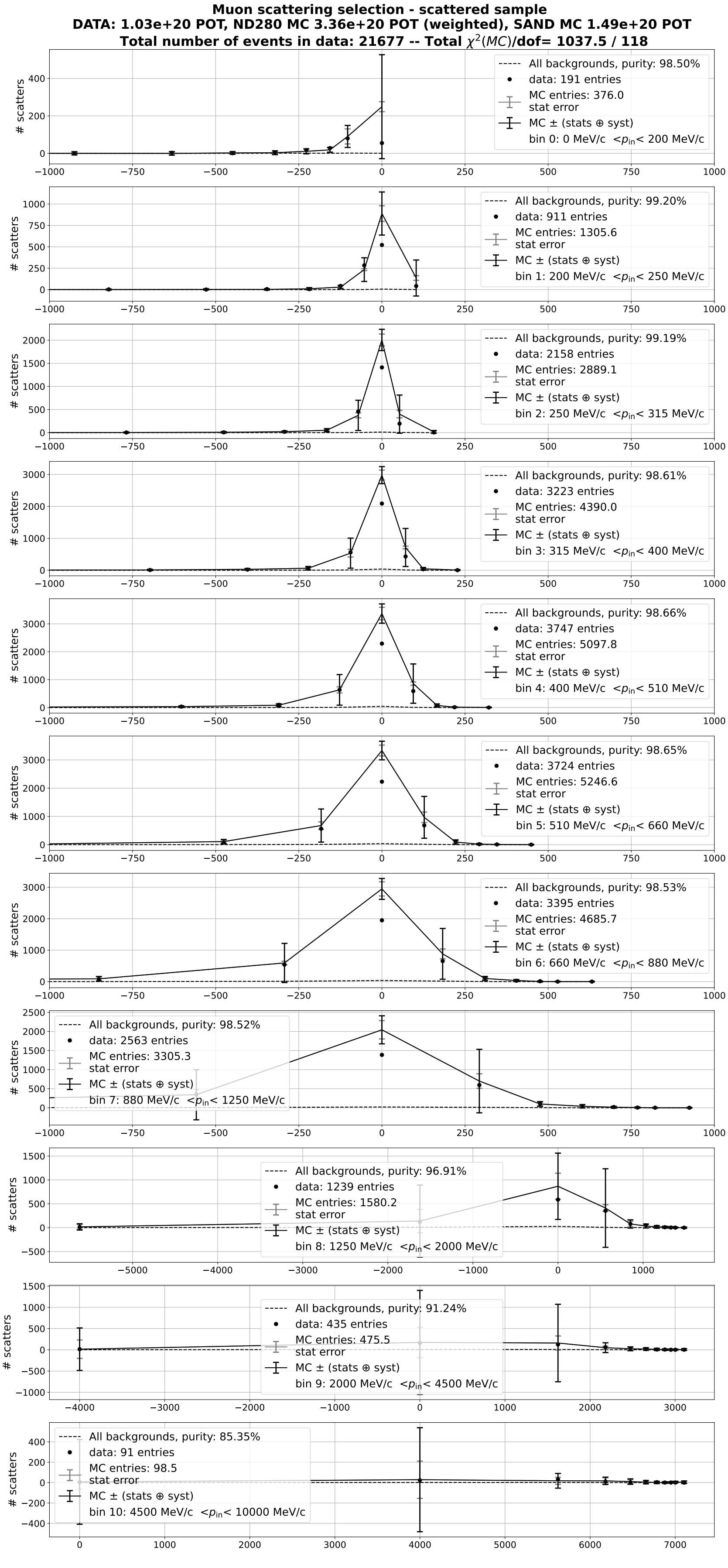


Figure 6.20: Reconstructed q_0 in bins of $p_{\rm in}^{@\ \rm PCA}$ for the entire sample.

Bin	Range (MeV/c)	Incoming momentum		Outgoing momentum	
		Entire sample	Scattered sample	Entire sample	Scattered sample
0	0–200	84.11% (540)	89.74% (212)	97.49% (14,343)	98.22% (10,085)
1	200–250	97.97% (2,272)	98.93% (1,500)	98.15% (20,147)	98.71% (14,185)
2	250–315	98.60% (20,232)	98.96% (14,414)	98.07% (32,872)	98.54% (22,354)
3	315–400	98.19% (42,993)	98.63% (29,489)	97.64% (45,544)	98.21% (29,328)
4	400–510	97.83% (59,402)	98.27% (37,678)	97.50% (55,716)	98.00% (32,165)
5	510–660	97.54% (73,746)	97.96% (39,765)	97.52% (67,144)	97.89% (31,390)
6	660–880	97.80% (89,759)	97.92% (35,603)	97.90% (84,129)	97.95% (28,722)
7	880–1250	98.26% (122,146)	97.93% (26,123)	98.37% (115,499)	97.81% (20,252)
8	1250–2000	98.70% (173,914)	96.53% (11,841)	98.80% (165,101)	96.18% (9,067)
9	2000–4500	99.11% (233,475)	89.47% (3,728)	99.24% (221,485)	92.32% (3,039)
10	4500–10000	99.25% (70,377)	85.99% (800)	99.43% (68,542)	91.58% (560)
Total	0-10000	**98.32% (827,956)**	**97.09% (229,453)**	**98.50% (825,592)**	**97.66% (196,156)**

Table 6.4: Bin purity p_i (number of events) for the reconstructed incoming and outgoing momenta for both entire and scattered samples using 1.04×10^{20} POT of MC not normalised to the data.

Bin	Scattering angle range (rad)	Entire sample	Scattered sample
0	0.000 - 0.009	99.29% (348,224)	-
1	0.009 - 0.023	98.81% (307,647)	-
2	0.023 - 0.044	98.42% (149,083)	98.47% (114,802)
3	0.044 - 0.067	98.09% (58,911)	98.20% (48,950)
4	0.067 - 0.096	97.58% (26,371)	97.87% (23,076)
5	0.096 - 0.136	96.34% (11,550)	96.72% (10,369)
6	0.136 - 0.176	93.38% (3,326)	94.06% (3,033)
7	0.176 - 0.241	84.13% (1,484)	84.99% (1,358)
8	0.241 - 0.321	62.40% (660)	-
9	0.321 - 0.436	27.22% (383)	-
10	0.436 - 1.047	8.15% (696)	-
11	1.047 - 1.570	0.00% (66)	-
Total	0.000 - 1.570	**98.18% (911,427)**	**97.86% (205,589)**

Table 6.5: Bin purity p_i (number of events) for the reconstructed scattering angle for both entire and scattered samples using 1.04×10^{20} POT of MC not normalised to the data.

6.3 Analysis methods

6.3.1 Normalisation to incoming muon flux

Unlike in neutrino scattering analyses, we know a lot about the initial state of the lepton in charged lepton scattering. In the present analysis, TPC1 gives a precise measurement of the incoming momentum and direction of the particle before it reaches FGD1. We use this information to scale (i.e. normalise) the Monte-Carlo to the observed data in 2D bins of incoming momentum and incoming angle with respect to the z axis in TPC1. Figure 6.21 shows the total number of events selected in the data (7.78×10^{20} POT) and MC (1.04×10^{21} POT) in bins of incoming momentum and angle with respect to the z axis.

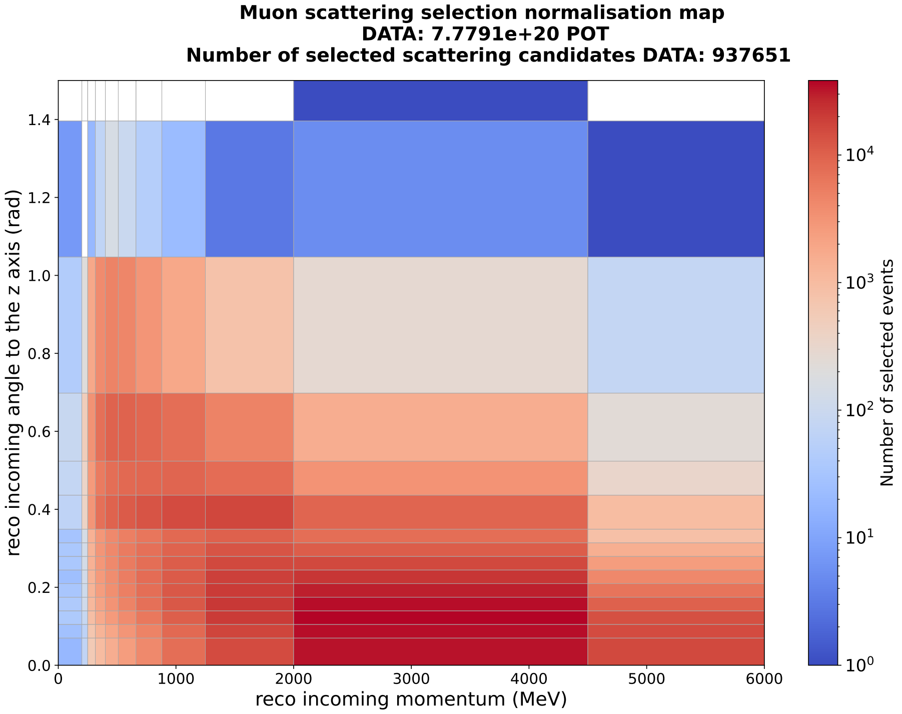
Figure 6.22 shows the one dimensional projections of Figure 6.21 onto the incoming momentum and incoming angle axes and Figure 6.23 shows the fraction of event per bin in each sample. On these figures we can see that there is an overall normalisation difference between data and MC, already discussed in Section 6.2.1. We also observe a noticeable difference in the shape of the incoming momentum distribution between the data and the Monte Carlo (MC) simulation. Specifically, the MC tends to peak at slightly lower momenta than the data. Within the MC, the SAND component peaks at higher momenta than the ND280 component and shows a significantly broader distribution. This shift in peak position might indicate an underestimation of the muon momentum in the SAND MC. In reality, muons produced in the sand are expected to have higher momenta than those produced in ND280. Given that the TPC's momentum resolution worsens at higher energies, this could explain the observed discrepancy. Interestingly there is almost no shape difference between the two distributions in term of incoming angle with respect to the z axis. This seems to indicate that the incoming muon angle and relative ND280 and SAND fractions are well modelled by the MC but that it tends to under-estimate the muon momentum.

We normalise the MC to the data by applying a weight w_i to all MC event in the (p_{in}, ϕ_{in}) 2D bin i. This weight corresponds to the ratio of number of events in bin i in the MC (n_i^{MC}) over the number of data events in this bin (n_i^{data}), appropriately scaled by the POT ratios:

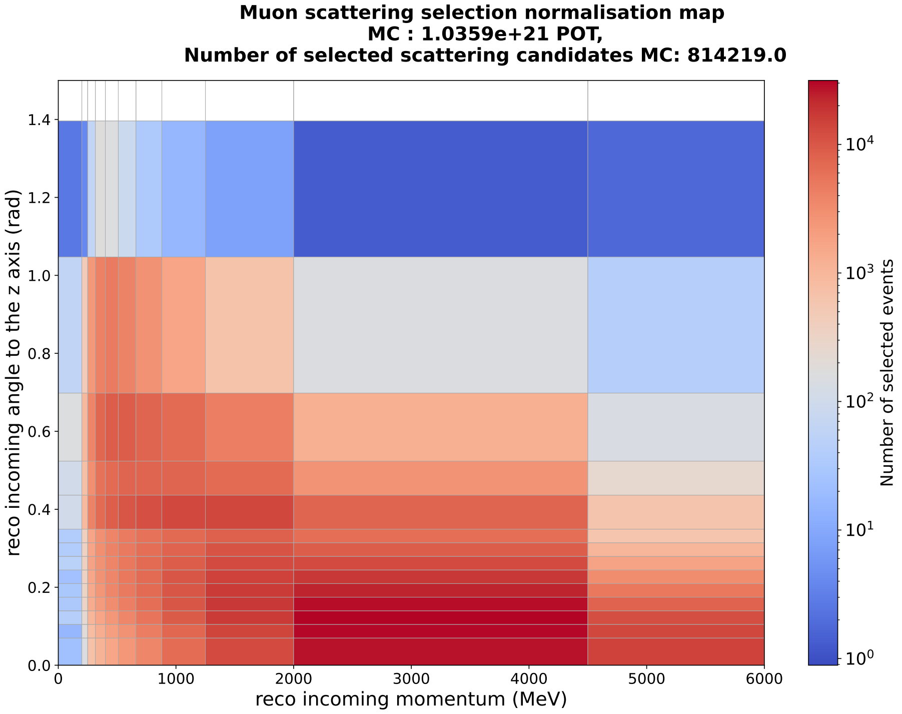
$$w_i = \frac{n_i^{\text{MC}}}{n_i^{\text{data}}} \times \frac{\text{POT}_{\text{ref}}^{\text{MC}}}{\text{POT}_{\text{measured}}^{\text{MC}}} \times \frac{\text{POT}_{\text{measured}}^{\text{data}}}{\text{POT}_{\text{ref}}^{\text{data}}} \tag{6.11}$$

where $\text{POT}_{\text{ref}}^{\text{data}}$ and $\text{POT}_{\text{ref}}^{\text{MC}}$ are the number of data and MC POT used to make the normalisation map and $\text{POT}_{\text{measured}}^{\text{data}}$ and $\text{POT}_{\text{measured}}^{\text{MC}}$ are the amount of data and MC POT used in a given data-MC comparison.

Figure 6.24 shows the normalisation weight w_i calculated using 7.78×10^{20} POT of data and 1.04×10^{21} POT of MC simulation. We observe that the normalisation weights are very uniform across the most populated region of phase space (250 MeV/c $< p_{\text{in}} <$ 4500 MeV/c, and $\phi_{\text{in}} <$ 1.07 rad, as shown from Figure 6.21) and roughly



(a) Data

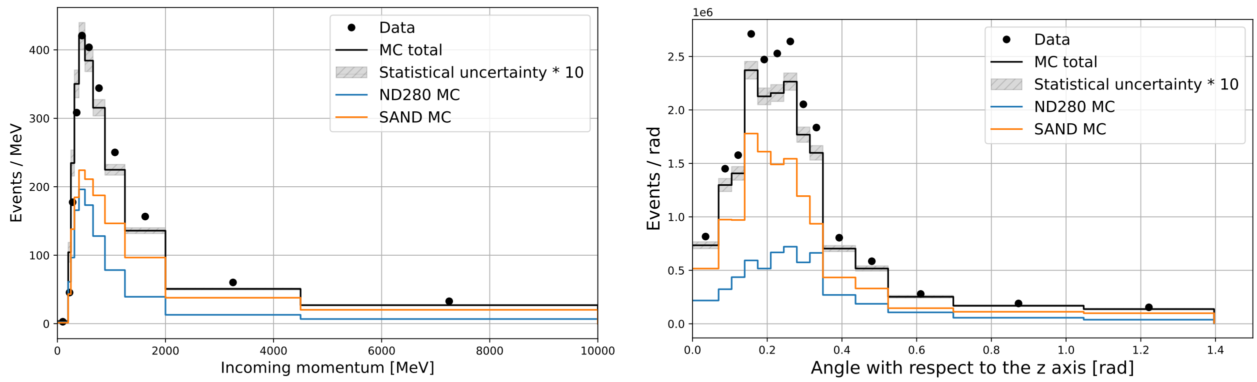


(b) MC

Figure 6.21: Number of events selected per bin of incoming angle and incoming momentum in the entire selection. The upper bins in momentum and angle are overflow bins.

equal to the ratio of total number of selected events.

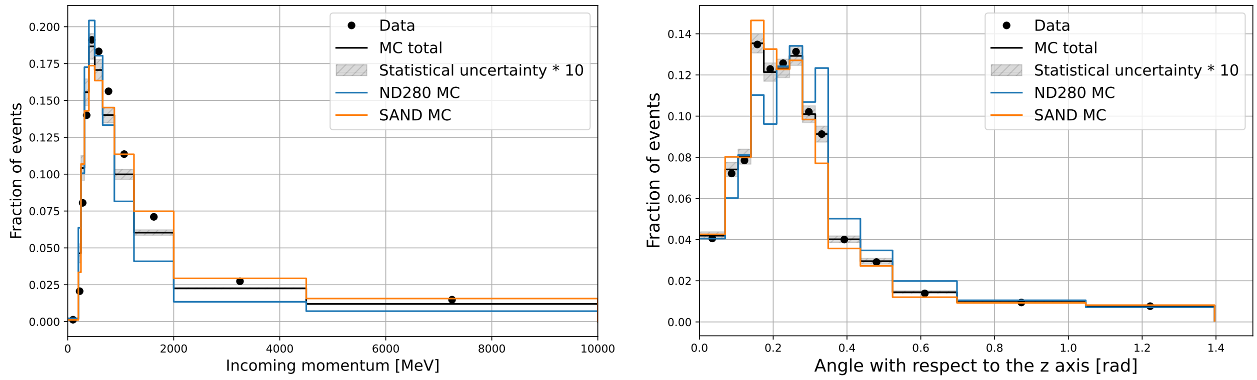
Applying the normalisation presented in Equation 6.11 removes any MC-data shape and rate differences caused by the mis-modelling of the incoming muon flux, allowing us to make an unbiased muon scattering comparison



(a) Incoming momentum

(b) Incoming angle with respect to the z axis

Figure 6.22: Distribution of incoming momentum (left) and angle with respect to the z axis (right) showing that more events are selected in data than in Monte-Carlo and that SAND (orange) and ND280 (blue) have slightly different incoming momentum flux and angular distributions, as discussed in the text.



(a) Incoming momentum

(b) Incoming angle with respect to the z axis

Figure 6.23: Fraction of events in each incoming momentum bin (left) and incoming angle bin (right). We observe that the shape of the MC distribution is in good agreement with those of the data in terms of angle to the z axis but not in terms of incoming momentum where the incoming momentum peak in MC is shifted to lower values compared to the data.

between the data and the Geant4 prediction.

6.3.2 Unfolding

To compare the data measured by ND280 with other datasets and theoretical model of lepton scattering, we need to remove background events from the sample and account for smearing of the observables caused by finite detector resolution and imperfect reconstruction. This procedure is called unfolding. The muon scattering analysis presented here uses a simple unfolding technique: given that n_i^{reco} events are measured in bin i for a reconstructed variable x_{reco}, we assume that the number of true muon scattering events n_j^{true} in bin j is given by:

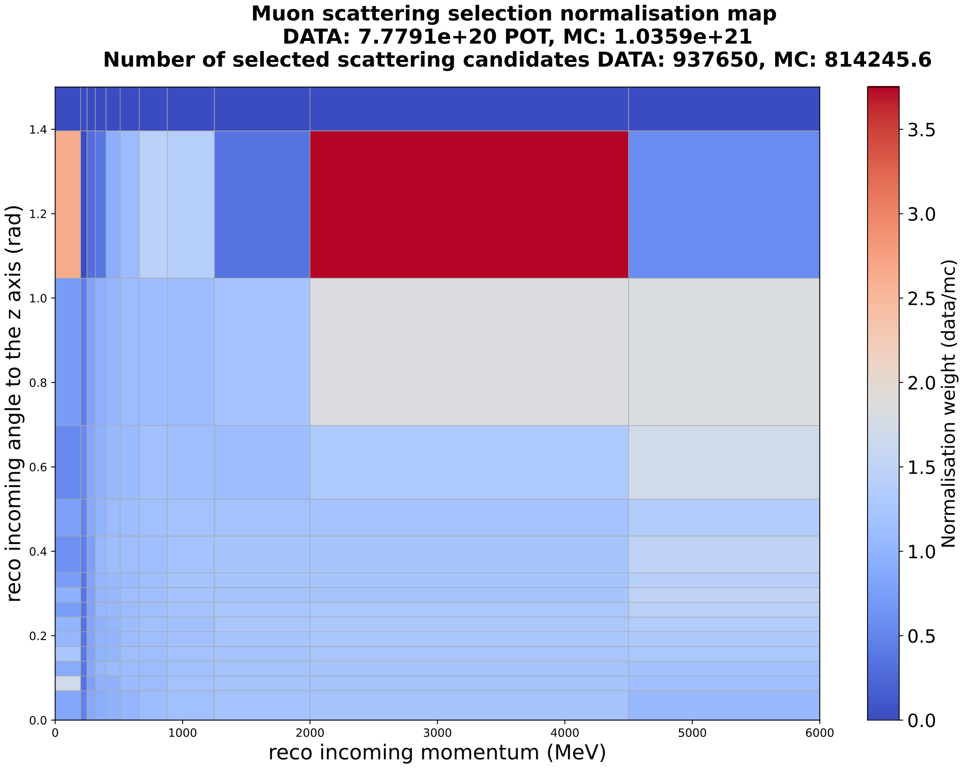


Figure 6.24: Normalisation weight w_i calculated for each (p_{in}, ϕ_{in}) bin.

$$n_j^{\text{true}} = \sum_i n_i^{\text{reco}} p_i U_{ij} \tag{6.12}$$

where i ranges between zero and N, the total number of bins, p_i is the purity of events in reconstructed bin i and U_{ij} is the unfolding matrix relating the distribution of events in bin i in reconstructed space with their true underlying distribution. The coefficient U_{ij} corresponds to the fraction of events in reconstructed bin i that lie in bin j in the true distribution. We use the same bins for both reconstructed and true distributions, such that the unfolding matrix U is an $N \times N$ square matrix where each row sums up to 1. We compute the unfolding matrix U and bin-by-bin purities, using the MC simulation where both true and reconstructed values of each observable are available. We compute one unfolding matrix for the entire sample and one for the scattered sample, since they are expected to behave differently and use the sample purities presented in Section 6.2.7. We apply the same unfolding to the MC and the data and unfold the variables of interest i.e. incoming and outgoing momenta, scattering angle and energy transferred.

The particle's momentum as it leaves TPC1 and enters TPC2 is the only variable for which we directly have generator-level information. As presented in Section 6.2.6, for the scattering angle and transferred energy we calculate a proxy of the generator level information using the same tools with true instead of reconstructed TPC observables.

6.3.2.1 Momentum

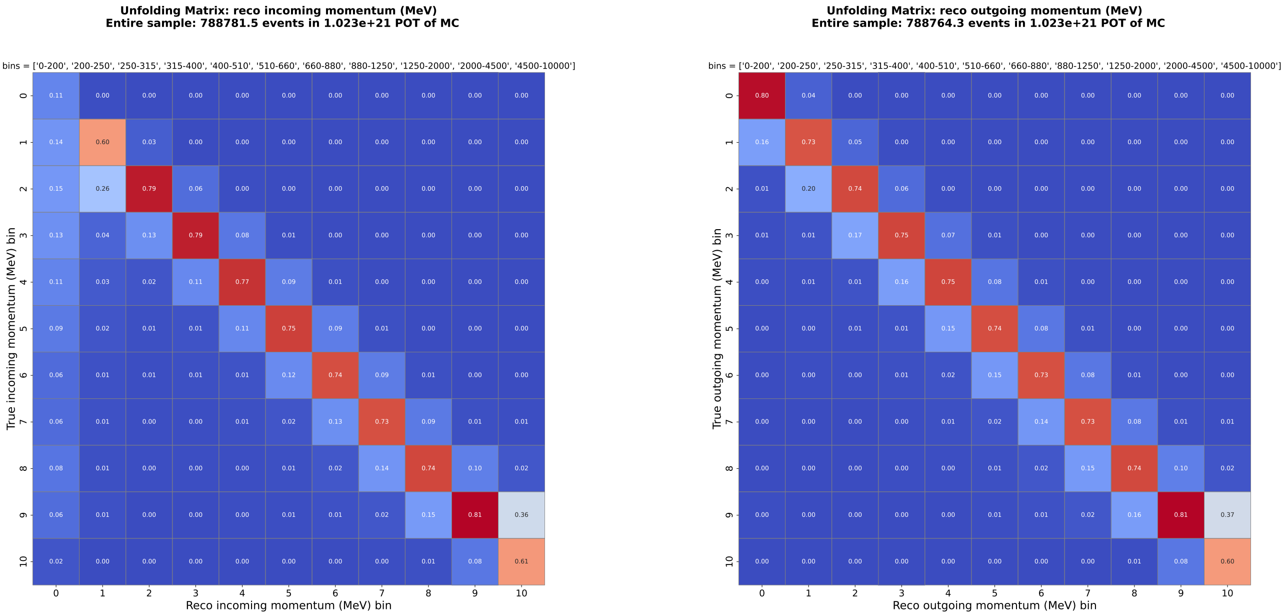
Figure 6.25 shows the unfolding matrix for the incoming and outgoing momenta in the entire and scattered sample. The unfolding matrices are mostly diagonal, with about 74% of events reconstructed in the correct bin. We note that in the entire sample about 12% (15%) of events are reconstructed with a higher incoming (outgoing) momentum and about 9% (8%) of events are reconstructed with a lower incoming (outgoing) momentum. In the scattered sample about 77% of events are reconstructed in the correct momentum bin. Interestingly, we note that the reconstruction is more momentum-dependent in the scattered sample than in the entire sample. We observe that lower (higher) true momenta are typically reconstructed with higher (lower) values in that sample. This is thought to be due to the fact that muons with a high momentum are less likely to scatter.

Finally, we observe in both samples that events reconstructed in the first incoming momentum bin (0-200 MeV/c) are mostly mis-reconstructed with only 11% (8%) of events correctly reconstructed in the entire (scattered) sample. This can easily be explained by the fact that low-momentum muons are unlikely to be accepted by either selection because of their large bending in the magnetic field and because they lose a lot of energy in FGD1. This means that incoming TPC tracks with reconstructed incoming momenta between 0 and 200 MeV/c are most likely mis-reconstructed higher momentum tracks, as we can see on the unfolding matrices. In a few cases, there are genuine low momentum particles most likely travelling upwards and/or at the top of TPC1 such that the downwards bending caused by the magnetic field is sufficiently compensated for the particle to reach TPC2.

6.3.2.2 Scattering angle

The unfolding matrix for the scattering angle variable gives some insight into the definition of a scattered event introduced in Section 6.2.4, which required that the reconstructed scattering angle was larger than 0.02 rad (1.3°). Figure 6.26 shows the unfolding matrix for the scattering angle variable for both samples.

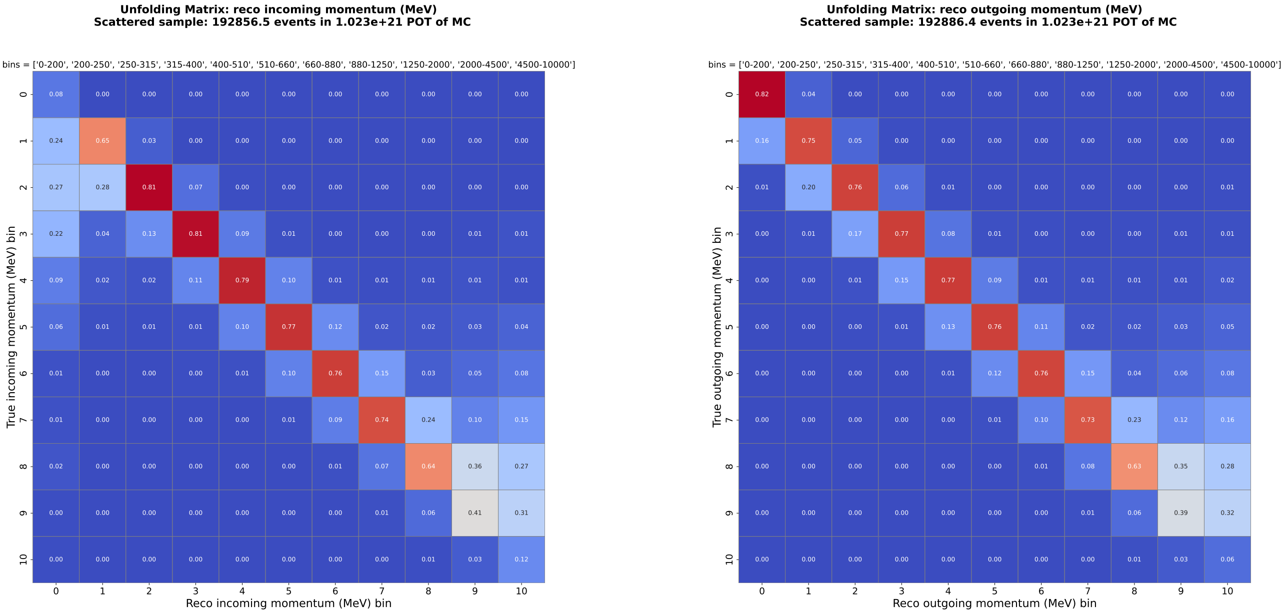
We see that the unfolding matrix is mostly diagonal. The reconstruction algorithm is more likely to overestimate (about 25% of events) than underestimate (about 5% of events) the scattering angle. We note that about 25% of events reconstructed in the second scattering angle bin (0.5° to 1.3°) have true scattering angles less than 0.5°. Such a high contamination of non-scattered events (i.e. compatible with $\theta = 0°$) is not acceptable, which leads to the cut at 1.3° presented in Section 6.2.4.2. We note that the additional PCA cut presented in Section 6.2.4.1 effectively reduces the smearing of the scattering angle and removes true through-going events from the sample. To illustrate this, note that for the entire sample, 5% of events in reconstructed bin 2 (1.3° - 2.5°) belong to true bin 0 (0° - 0.5°), whereas this contamination of non-scattered events in this bin is only 1% in the scattered sample.



Unfolding Matrix: reco incoming momentum (MeV)
Entire sample: 788781.5 events in 1.023e+21 POT of MC

Unfolding Matrix: reco outgoing momentum (MeV)
Entire sample: 788764.3 events in 1.023e+21 POT of MC

(a) Entire sample: incoming momentum (b) Entire sample: outgoing momentum



Unfolding Matrix: reco incoming momentum (MeV)
Scattered sample: 192856.5 events in 1.023e+21 POT of MC

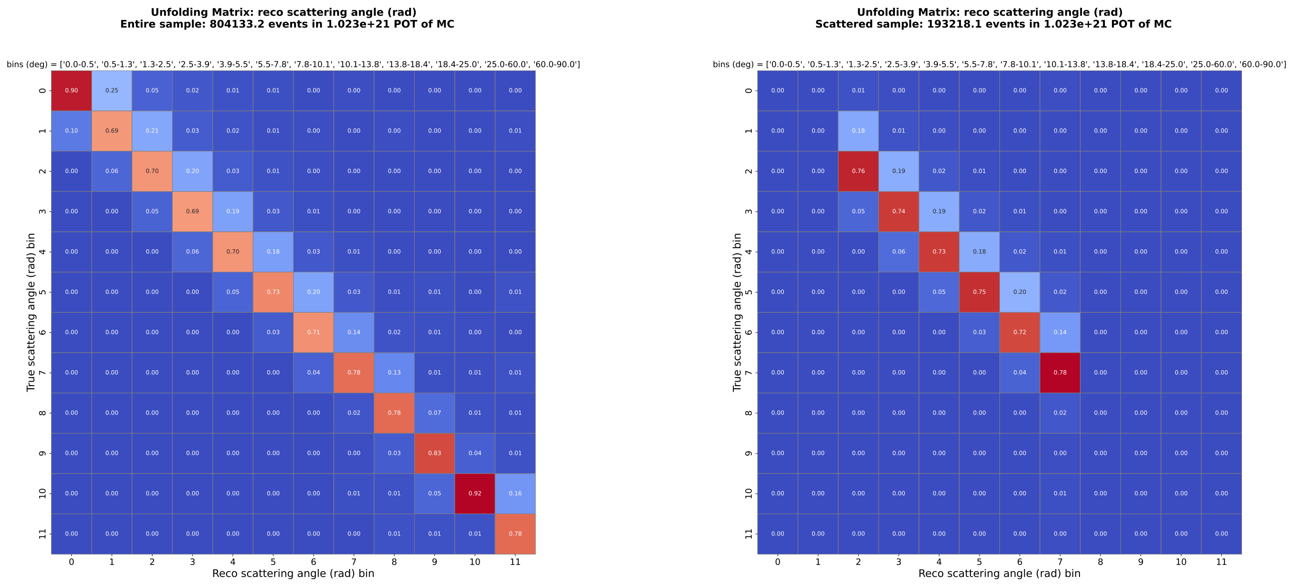
Unfolding Matrix: reco outgoing momentum (MeV)
Scattered sample: 192886.4 events in 1.023e+21 POT of MC

(c) Scattered sample: incoming momentum (d) Scattered sample: outgoing momentum

Figure 6.25: Unfolding matrices for the incoming and outgoing momenta in both selections.

6.3.2.3 Transferred energy

Section 6.2.6.3 discussed how we estimate the amount of energy transferred in the scatter by comparing the incoming and outgoing momenta at the PCA. Because the incoming and outgoing momenta are strongly correlated, they need to be treated together by the unfolding matrix. We define a new set of $M = N \times N$ bins where $N = 11$ is the number of incoming and outgoing momentum bins, such that the k^{th} bin where $k = i \times N + j$ holds all events whose incoming and outgoing momenta fall in bins i and j respectively. Figure 6.27 shows the purity of the scattered sample in bins of energy transferred. We note that the sample purity is generally



(a) Entire sample

(b) Scattered sample

Figure 6.26: Scattering angle unfolding matrices for the entire and scattered samples.

above 90%, and that bins corresponding to large differences between the incoming and outgoing momenta are background-dominated, as expected.

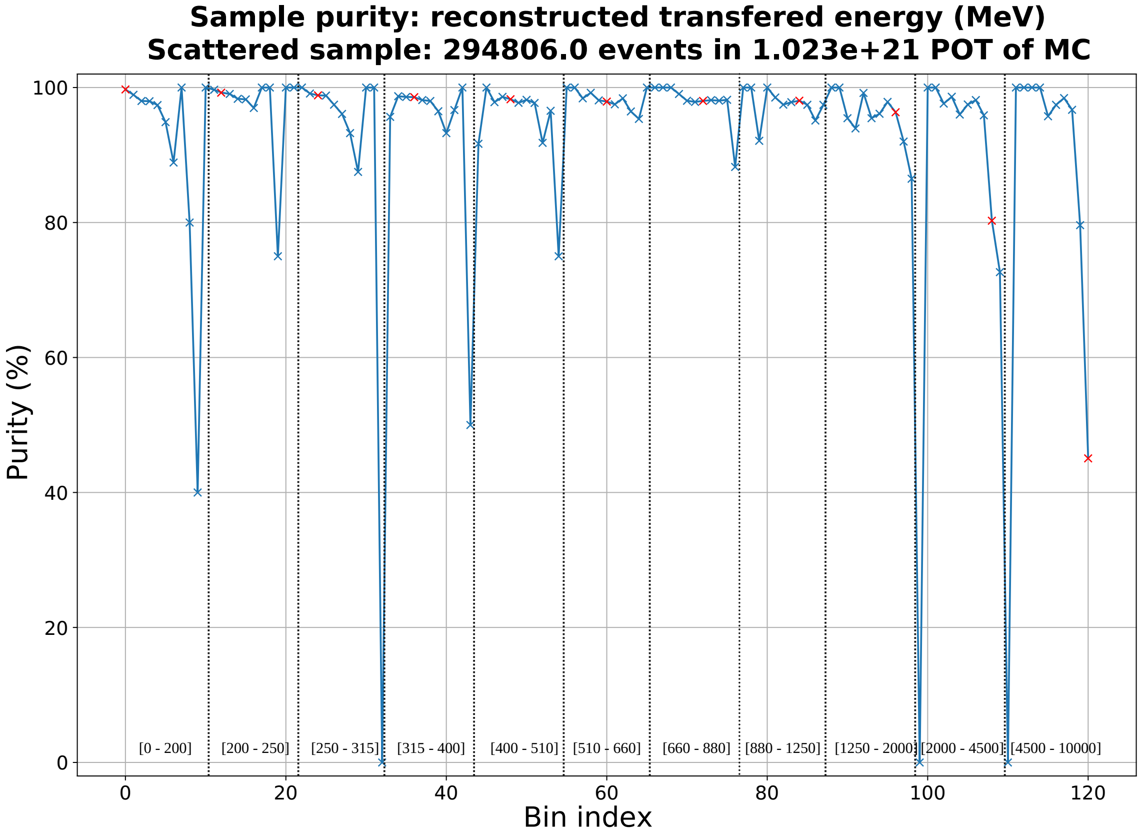


Figure 6.27: Sample purity in reconstructed bins of transferred energy. The bins of incoming momentum at the PCA are separated by dashed back lines, and the bin ranges (in units of MeV/c) are indicated. Bins of energy transferred for which the incoming and outgoing momenta fall in the same bin (i.e. $q_0 = 0$) are indicated in red.

The energy transferred unfolding matrix for the entire sample is shown in Figure 6.28 where the separation between bins of incoming energy is indicated with solid black lines. Bins corresponding to $q_0 = 0$ (i.e. incoming and outgoing momenta fall in the same bin) are highlighted in red. We observe that the unfolding matrix typically migrates events reconstructed with $q_0 \neq 0$ to bins of $q_0 = 0$ (i.e. within horizontal red lines), correcting for mis-reconstruction of the momenta.

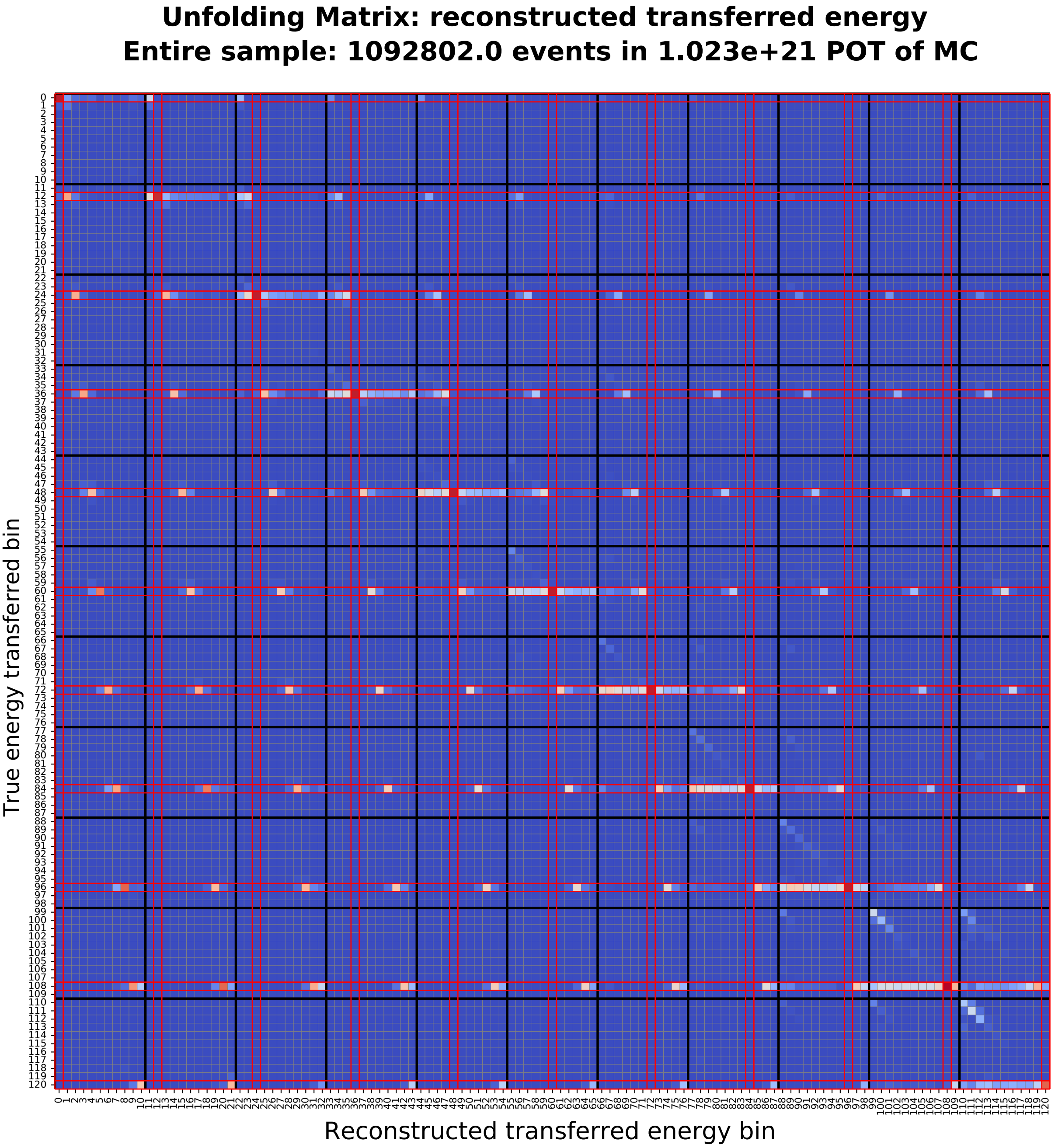


Figure 6.28: Unfolding matrix for the transferred energy, where the black lines correspond to bins of incoming momentum at the PCA and red lines to bins where the incoming and outgoing momenta at the PCA are in the same bin. The coefficients are between 0 (blue) and 1 (red).

6.4 Uncertainties

6.4.1 Systematic uncertainties

The muon selection and analysis presented above rely exclusively on TPC1 and TPC2 information to identify the particles and estimate the scattering kinematics. We therefore expect that TPC-related uncertainties will dominate the total systematic uncertainty on this measurement. Thanks to the normalisation procedure described in Section 6.3.1, this analysis is not affected by neutrino beam flux uncertainties. We, however, expect that uncertainties on the secondary interaction (SI) of pions will be a non-negligible source of error, because pion scattering is the dominant background in this analysis. Finally, because the value of the magnetic field is used in the propagation of the particle's trajectory, we expect that any uncertainty on the magnitude and direction of the magnetic field will also impact the reconstruction of tracks by the TPCs.

The magnitude of the uncertainty associated with each of these parameters as well as any discrepancies between the data and MC has been estimated using dedicated control samples. Examples of such studies can be found in Refs [185–187]. The data-MC discrepancies are lifted by applying a set of corrections to either the MC or data. This analysis uses all the default T2K corrections: dE/dx correction in data and MC, pile up corrections and general data quality correction accounting for periods of time where some detectors were malfunctioning. An additional set of corrections is applied only when estimating systematic uncertainties. These are corrections to the TPC momentum resolution, the TPC PID and FGD1 PID. All corrections are associated with an uncertainty. More information about the default HighLAND corrections can be found in [184].

The propagation of systematic uncertainties is done in HighLAND using the PSyChE (Propagation of Systematics and Characterisation of Events) package. In PSyChE, uncertainties are split into two types:

- **"Variation" systematics** which impact the scale (mean) or resolution of a variable (e.g. TPC momentum). This error is propagated by varying the given variable within a credible range and rerunning the entire selection. Variation systematics typically handle cases where an observable depends on another one (e.g. the reconstructed particle momentum depends on the strength of the magnetic field) and differences between MC and data (e.g. on the mean and resolution of a given observable). Depending on whether the variable is known or not, whether a correction needs to be applied with respect to the data and on whether the resolution on the observable is the same in data and MC, the smearing of the variable is done slightly differently to correctly account for uncertainties in the mean and resolution of the variable [184].

- **"Weight" systematics**, which can be divided into efficiency-like and normalisation-like systematics. In both cases, the uncertainty is propagated by applying a weight to each event without re-running the

> selection. Efficiency-like systematics address uncertainties in reconstruction and detection efficiency (e.g. matching or tracking efficiencies), whereas normalisation-like uncertainties affect the total event rate (e.g. FGD1 mass for neutrino cross-section measurements).

The muon scattering makes use of systematic dials already available in HighLAND and presented in Table 6.6. No uncertainty is currently applied to the momentum correction presented in Section 6.2.3.1. This source of uncertainty only affects the transferred energy parameter and is expected to be negligible compared to the overall low resolution we have on this variable.

Source	**Type**	**Description**
TPC momentum scale	Variable	Uncertainty estimated from a 2009 mapping of the magnetic field [186].
TPC Momentum Resolution	Variable	Uncertainty and correction obtained from tracks crossing multiple TPCs in data and MC samples [188]
TPC PID	Variable	The difference and smearing in the pull values are computed using a sample of events with known particle ID, in bins of momentum [189].
B field distortion	Variable	Corrections to the position of the TPC charge clusters obtained from **B** field maps [190]
TPC Charge ID efficiency	Weight	Capacity of TPCs to identify the particle charge, estimated by comparing charge ID by individual TPCs for a control sample of through-going muons [184].
TPC clustering efficiency	Weight	Corrects for differences between the data and MC in terms of number of TPC nodes for vertical and horizontal tracks individually based respectively on a CC sample of events starting in FGD1 and on a sample of cosmic muons [184, 191]
Pion secondary interaction	Weight	Uncertainty calculated based on a dedicated MC study [192, 193] of pion absorption and charge exchange validated against external pion scattering data [194].

Table 6.6: Summary of the source of systematic uncertainties used in the muon scattering analysis.

6.4.2 Toy throwing method

To estimate the total uncertainty on the number of events and estimate the bin-to-bin correlations, we generate a set of toy experiments and study the change in number of events in each bin caused by the variation of a given observable. For each toy experiment, we vary the parameters according to the systematic uncertainties presented in Table 6.6. The toy experiments need to be treated in the same way as the data. We compute four set of toy experiments, corresponding to the entire sample not unfolded, the entire sample unfolded, the scattered sample not unfolded and the scattered sample unfolded. We normalise each toy using the weights presented in Section 6.3.1, obtained from the nominal MC simulation. For each toy in an unfolded sample, we apply the nominal unfolding matrix presented in Section 6.3.2, essentially assuming that the unfolding matrix is perfectly accurate.

We have seen how the systematic uncertainties are accounted for using toy throws. To measure the total uncertainty on the number of events per bin, we also need to include statistical fluctuations in the number of events. Say that for a given toy t the number of events in bin i is n_i^t. We account for the statistical variations by drawing a new number of events in bin i for this toy from a Poisson distribution with expectation value n_t^i:

$$m_i^t = \text{Pois}(n_i^t) \tag{6.13}$$

.

The combined uncertainty on the number of events in bin i, accounting for resolution and shift in the scale is given by:

$$\sigma_i^{\text{tot}} = \sqrt{\left(\frac{\sigma_i}{\bar{m}_i}\right)^2 + \left(\frac{\bar{m}_i - n_i^{\text{nominal}}}{n_i^{\text{nominal}}}\right)^2} \tag{6.14}$$

where $\bar{m}_i$ is the mean number of events in bin i across all the toys, n_i^{nominal} is the nominal number of events in bin i and

$$\sigma_i = \sqrt{\frac{1}{N_{\text{toys}}} \sum_{t=0}^{N_{\text{toys}}} (m_i^t - \bar{m}_i)^2} \tag{6.15}$$

is the standard deviation of the number of events in bin i across all the toys. The total covariance matrix C is given by:

$$C_{ij} = \frac{1}{N_{\text{toys}}} \sum_{t=0}^{N_{\text{toys}}} (m_i^t - \bar{m}_i)(m_j^t - \bar{m}_j). \tag{6.16}$$

The correlation matrix R is given by:

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \tag{6.17}$$

Figure 6.29 shows how the number of events selected in the scattered sample varies over 500 toy throws. Each toy is normalised to 2.88×10^{20} POT of data, nominally corresponding to 71,132 scatters, and unfolded. All of the parameters presented in Table 6.6 are varied simultaneously, and the systematic variations were accounted for with the method presented above. The number of events in each bin roughly follows a Gaussian distribution centred at the nominal value, showing that the systematics are mostly smearing the distributions and barely shifting them. The bin-to-bin correlations in this unfolded sample are shown on Figure 6.30.

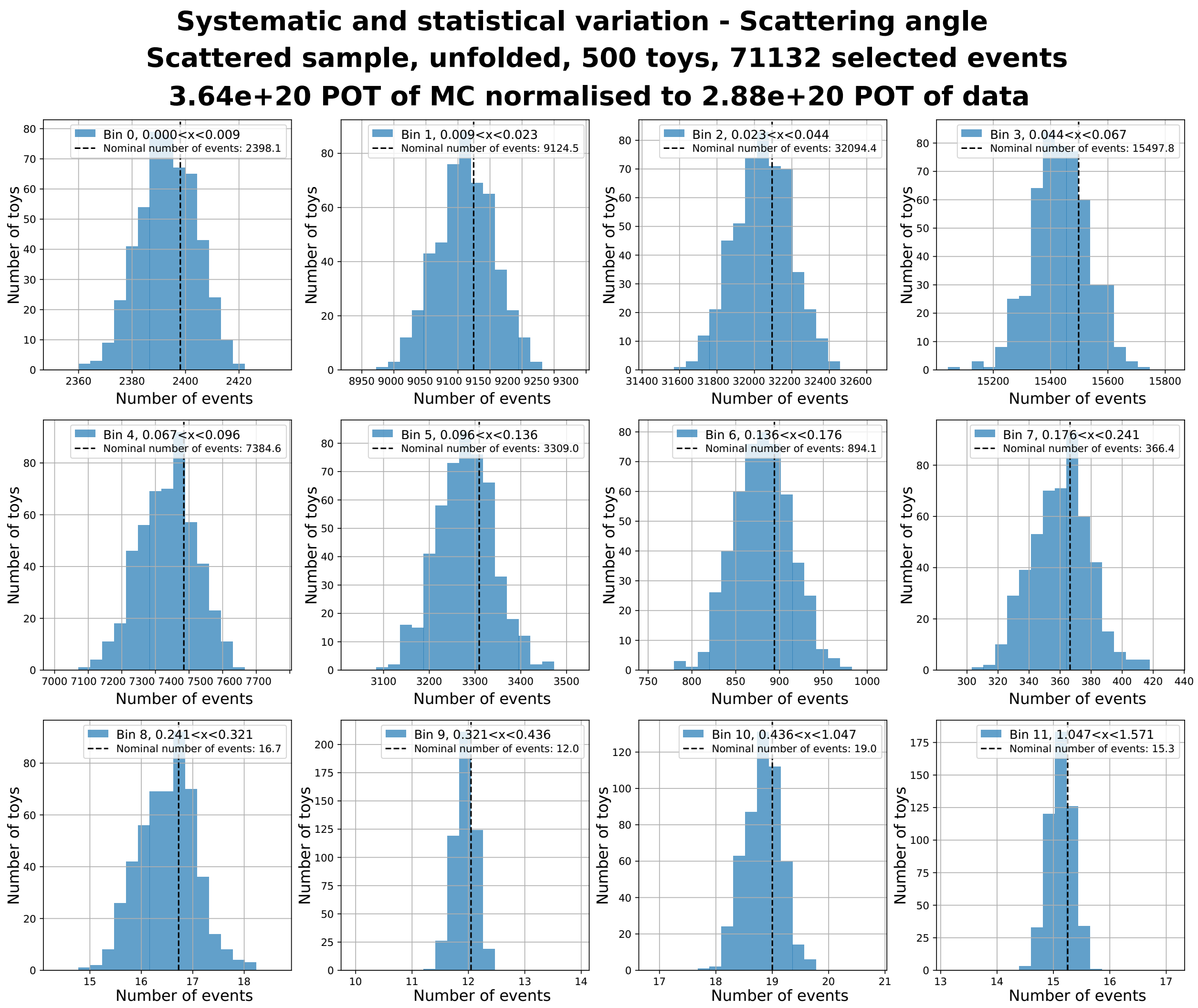


Figure 6.29: Distribution of the number of events selected in each scattering angle bin in the scattered and unfolded sample when varying the systematic parameters presented in Table 6.6 and accounting for systematic variations.

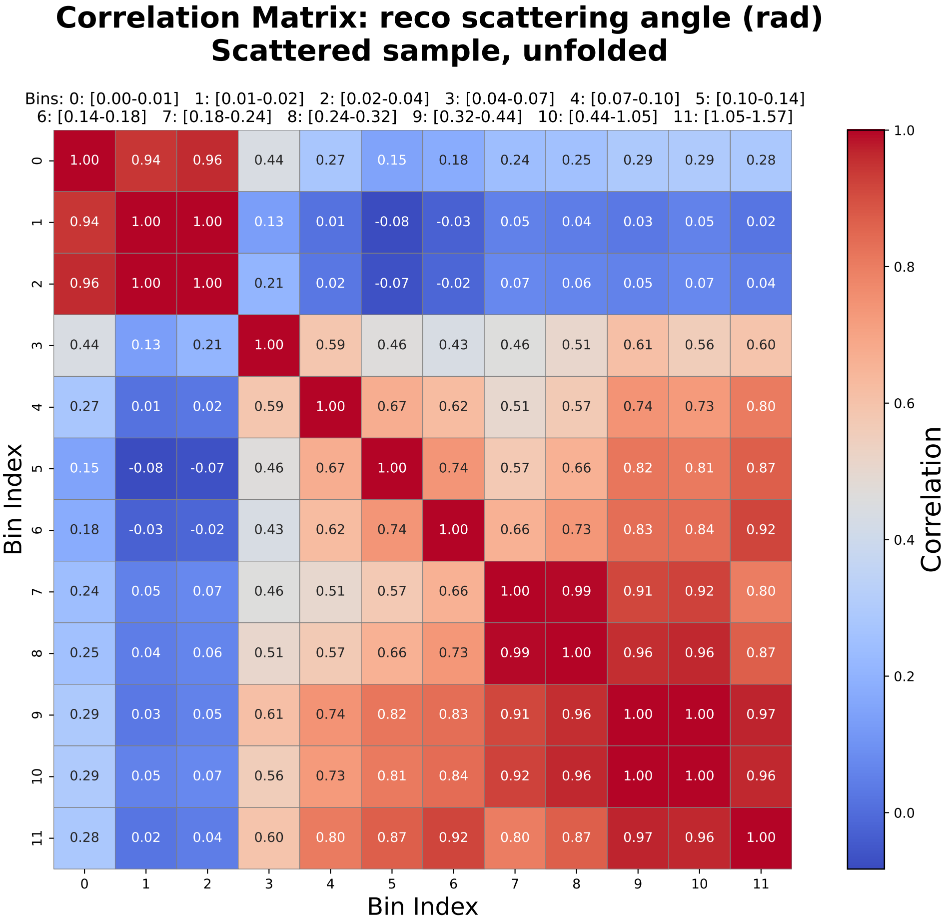
6.4.3 Relative uncertainties

6.4.3.1 Incoming and outgoing momenta

Figure 6.31 shows the fractional uncertainty on the incoming and outgoing momentum in the entire and the scattered sample, overlaid on the nominal number of selected events. In this section, we present the contribution from each systematic source using 100 toys. The total error used in the muon scattering measurement presented in Sections 6.5 and 6.6 were obtained with 500 toys. The uncertainty is slightly more precise when using a larger number of toys. Otherwise the two estimates do not significantly differ. The statistical uncertainty is given by:

$$\sigma_i^{\text{stat}} = \sqrt{|(\sigma_i)^2 - (\sigma_i^{\text{syst}})^2|} \tag{6.18}$$

where σ_i^{syst} is the systematic uncertainty, obtained from purely systematic toy throws (i..e not including the



	0	1	2	3	4	5	6	7	8	9	10	11
0	1.00	0.94	0.96	0.44	0.27	0.15	0.18	0.24	0.25	0.29	0.29	0.28
1	0.94	1.00	1.00	0.13	0.01	-0.08	-0.03	0.05	0.04	0.03	0.05	0.02
2	0.96	1.00	1.00	0.21	0.02	-0.07	-0.02	0.07	0.06	0.05	0.07	0.04
3	0.44	0.13	0.21	1.00	0.59	0.46	0.43	0.46	0.51	0.61	0.56	0.60
4	0.27	0.01	0.02	0.59	1.00	0.67	0.62	0.51	0.57	0.74	0.73	0.80
5	0.15	-0.08	-0.07	0.46	0.67	1.00	0.74	0.57	0.66	0.82	0.81	0.87
6	0.18	-0.03	-0.02	0.43	0.62	0.74	1.00	0.66	0.73	0.83	0.84	0.92
7	0.24	0.05	0.07	0.46	0.51	0.57	0.66	1.00	0.99	0.91	0.92	0.80
8	0.25	0.04	0.06	0.51	0.57	0.66	0.73	0.99	1.00	0.96	0.96	0.87
9	0.29	0.03	0.05	0.61	0.74	0.82	0.83	0.91	0.96	1.00	1.00	0.97
10	0.29	0.05	0.07	0.56	0.73	0.81	0.84	0.92	0.96	1.00	1.00	0.96
11	0.28	0.02	0.04	0.60	0.80	0.87	0.92	0.80	0.87	0.97	0.96	1.00

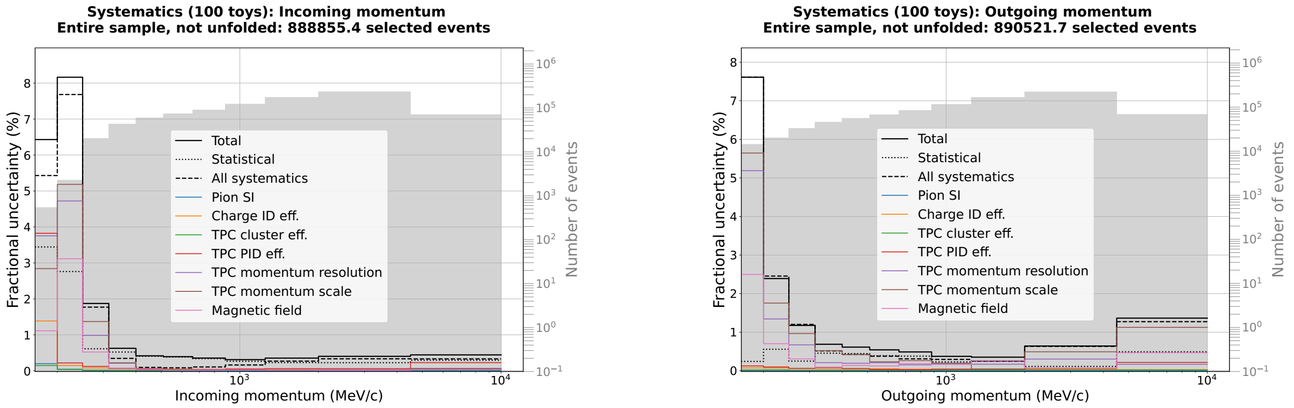
Figure 6.30: Correlation matrix for the scattering angle in the unfolded scattered sample showing the high correlations in the low and high scattering bins, highlighting the effects of the unfolding procedure. The edges of the bins are indicated above the matrix in units of radians.

statistical variation of Equation 6.13).

As could be expected, the dominant sources of uncertainty are related to the TPC momentum scale and resolution followed by the magnetic field distortions. We observe that the fractional uncertainty on the number of events is larger at low incoming and outgoing momenta. This is expected from previous discussions about the inaccuracies of the propagation method at lower momentum. In both samples, the error on the number of events selected with momenta below 250 MeV/c is $\sim 7\%$. At higher momenta the uncertainty is consistently below 2.5%. We note that in some rare cases, random systematic throws can partially compensate the systematic effects causing the combined error (solid black line) to be smaller than the systematic-only (dashed black line) error. This is for example the case in Figure 6.31b.

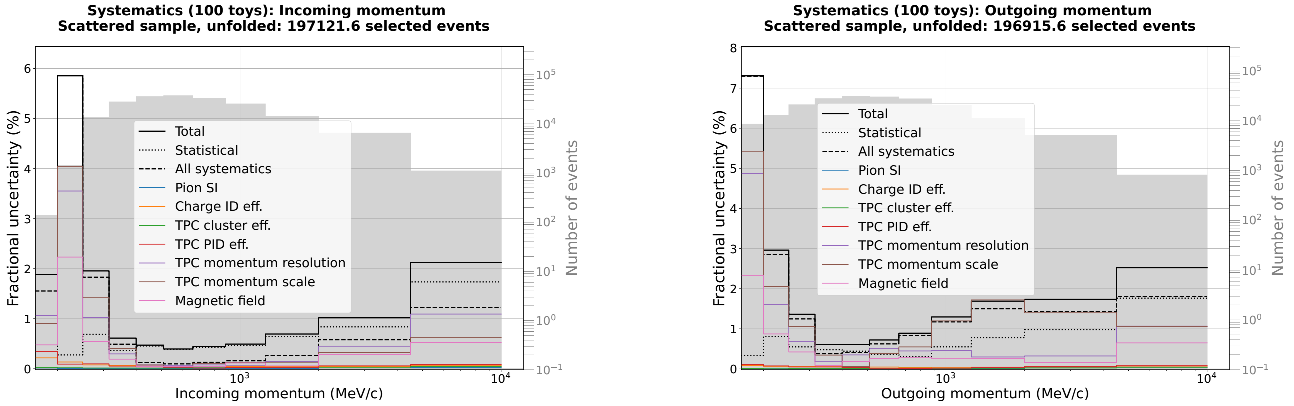
6.4.3.2 Scattering angle

The uncertainty on the number of events selected per bin of scattering angle in the entire and unfolded scattered sample is shown on Figure 6.32. On this figure we observe that the total uncertainty on the number of event



(a) Entire sample: incoming momentum

(b) Entire sample: outgoing momentum



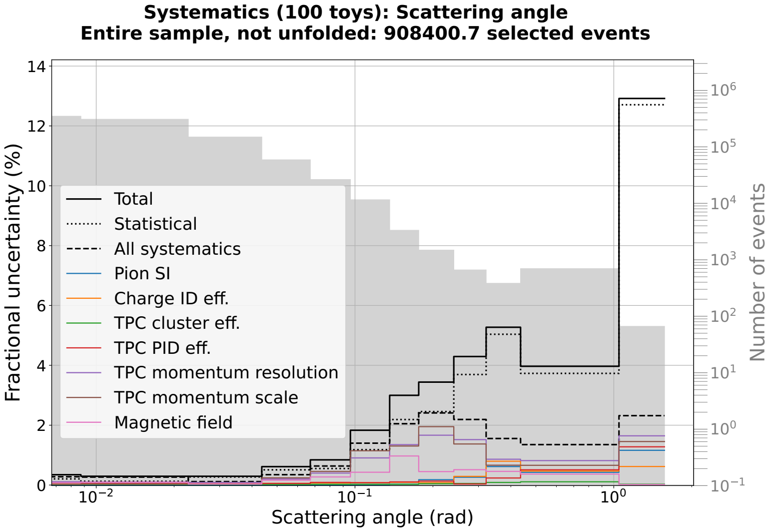
(c) Scattered sample unfolded: incoming momentum

(d) Scattered sample unfolded: outgoing momentum

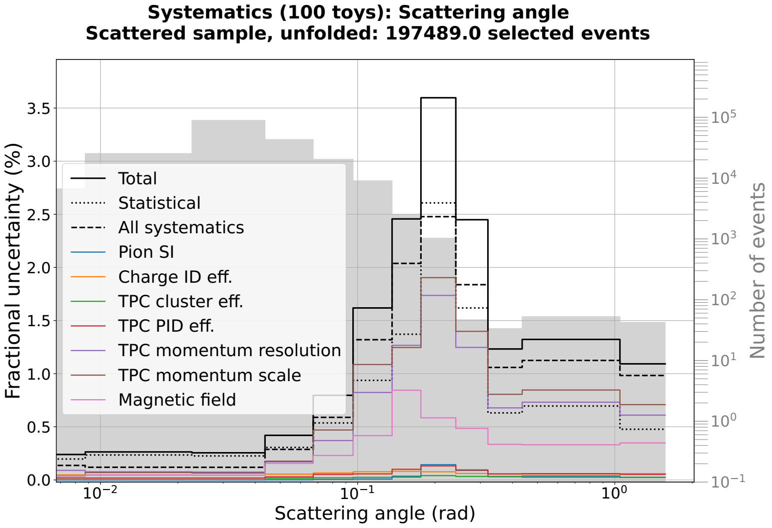
Figure 6.31: Uncertainty (solid line) on the number of events selected per bin incoming and outgoing momenta by both selections. The gray histogram shows the nominal number of events selected per bin.

per scattered angle bin is low for both samples at less than 0.5% for all scattering angles less than 0.067 rad (3.8°) which contains most of the data. The uncertainty on the number of events in the entire sample for bins of scattering angles larger than 0.24 (13.8°) is larger than 4% and statistically dominated. These bins are background dominated (see Section 6.2.2) and sparsely populated, which explains the large uncertainty. The maximal uncertainty in the entire sample is 13% corresponding to the last scattering angle bin. The total systematic error does not exceed 3% in either samples. In the scattered sample the maximum uncertainty on a bin is $\sim 3.5\%$ corresponding to the highest scattering angle bin included in the non-unfolded sample, i.e. bin 7 corresponding to $0.18 \leq \theta < 0.24$ rad.

For both samples the leading systematics are the TPC momentum scale and resolution. This is expected from the fact that the scattering angle is calculated based almost exclusively on the TPC's estimate of the particle's momentum, as are most of the variables used in the event selection. The next leading systematic is the magnetic field, which is another key ingredient in the propagation of the particle's trajectory. We note that the pion secondary interaction and TPC PID efficiency systematics becomes relevant (about 1% of the events) at high scattering angles in the entire sample. This is because pions, mostly likely to be mis-identified as muons,



(a) Entire sample, not unfolded



(b) Scattered sample, unfolded

Figure 6.32: Fractional uncertainty (solid line) on the scattering angle variable for the entire and the scattered sample split by contributions. The gray histogram shows the nominal number of events selected per bin.

make up the majority of the sample in these bins as seen in Figure 6.8b.

6.5 Data-MC comparison

Once we have selected muon-like particles crossing TPC1, FGD1 and TPC2, normalised the incoming flux to the data, identified those who scatter and factored out the detector smearing effects, we can compare the scattering kinematics between the data and MC simulation. We quantify the agreement between data and MC using a simple chi-squared test:

$$\chi^2 = (\mathbf{data} - \mathbf{MC})^{\mathrm{T}} C^{-1} (\mathbf{data} - \mathbf{MC}) \tag{6.19}$$

where **data** $= \{n_0^{\text{data}}, n_1^{\text{data}}, ..., n_f^{\text{data}}\}$ is the vector of number of data events in each bin for a given observable, similarly **MC** is the vector of number of MC events in each bin and C is the covariance matrix defined in Equation 6.16 which accounts for systematic and statistical uncertainties and bin to bin correlations. We compare χ^2 with the number of degrees of freedom (n_{dof}) equal to the number of non-empty bins. The value of χ^2/n_{dof} (reduced χ^2) indicates the agreement between the two distributions where $\chi^2/n_{\text{dof}} \approx 1$ indicates a good agreement, $\chi^2/n_{\text{dof}} >> 1$ indicates a poor agreement and $\chi^2/n_{\text{dof}} << 1$ typically indicates an overestimation of the uncertainties.

6.5.1 Through-going sample

6.5.1.1 Incoming and outgoing momentum

The simplest comparison that can be made between the MC and data is in terms of the incoming and outgoing momenta of the events in the entire sample. As explained in Section 6.3.1, the MC sample is normalised to the incoming momentum flux in data. Figure 6.33 shows the distribution of incoming momentum, which presents no difference between the MC and the data, confirming that the normalisation procedure behaves as expected.

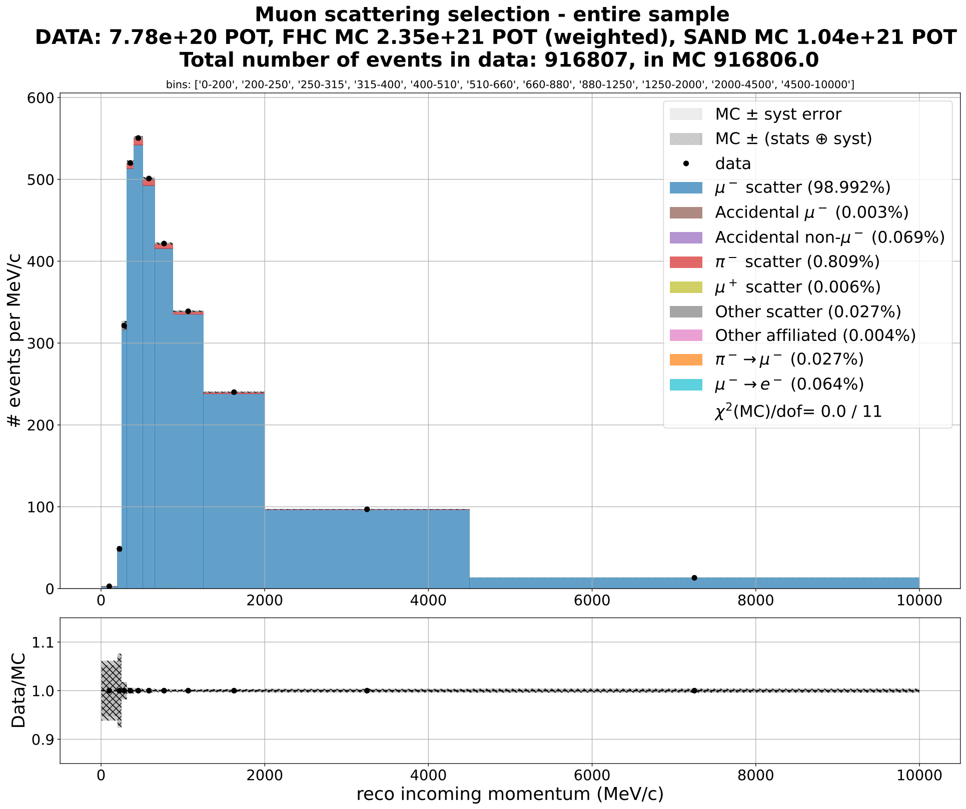


Figure 6.33: Number of events per MeV/c in bins of incoming momentum for data and MC showing that the normalisation procedure behaves as expected.

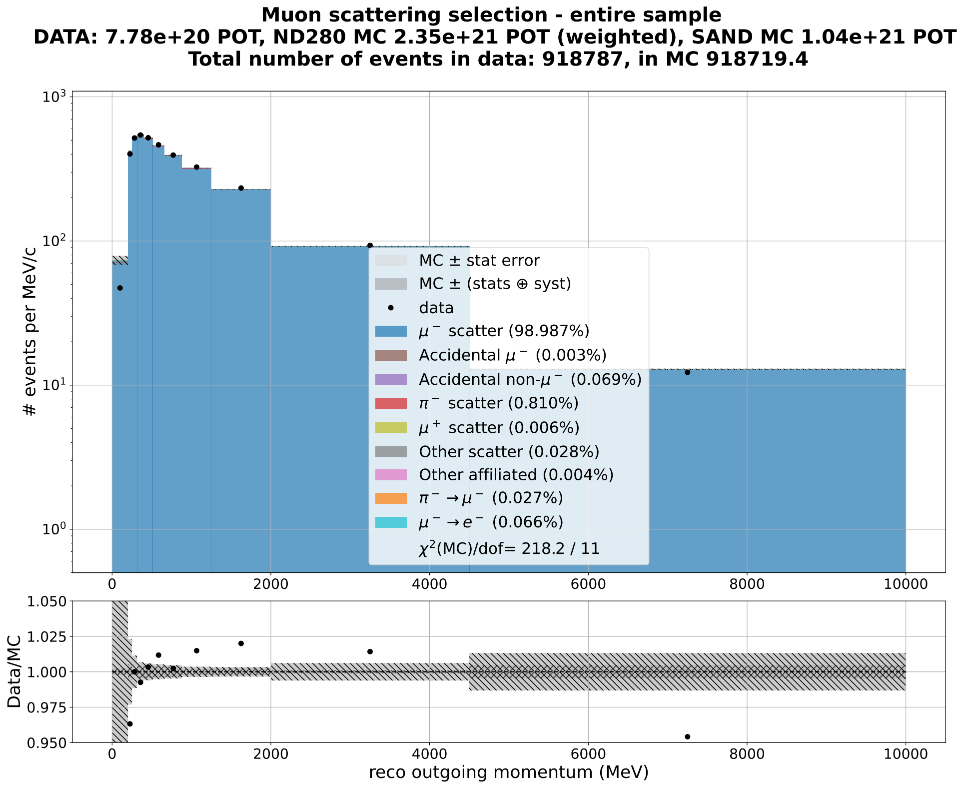
More interestingly, Figure 6.34 compares the distribution of outgoing momenta between the MC and data. We observe that despite the incoming momentum distributions being identical in data and MC, they are slightly different at the TPC2 entrance. While the overall agreement between MC and data is reasonable (χ^2/dof = 218.2/11), a localised discrepancy is visible around the momentum peak (approximately 400 MeV/c). We see that the MC underestimates the number of events before the peak (i.e. bins 0 to 2) and overestimates the number of events immediately after it (i.e. bins 5 to 9). This seems to indicate that the reconstructed peak in MC is shifted to slightly lower momentum compared to the data, which could be caused by small biases in the reconstruction algorithm or the modelling of scattering in MC, which is not entirely accounted for by systematics, since the observed discrepancies fall outside the systematics envelope.

The bin-to-bin correlation matrix is presented in Figure 6.34b. We note that there is a strong positive correlation between bins 2 and 4, which surround the peak. This smoothing effect can be explained by the finite reconstruction resolution and associated smearing. Additionally, we observe a slight anti-correlation between bins before the peak (bins 0 to 2) and bins after the peak (bins 3 to 8), which indicates the migration of events shifted out of the low-momentum bins into higher ones and vice versa. This effect, which could typically produce a shift in the peak position, is not sufficient to fully cover the observed shift, indicating that improvements should be made to the modelling of the detector response or reconstruction (for example in the SAND MC, which is known to be less accurate than the ND280 MC [195]) and/or that the Geant4 model of muon scattering is inaccurate.

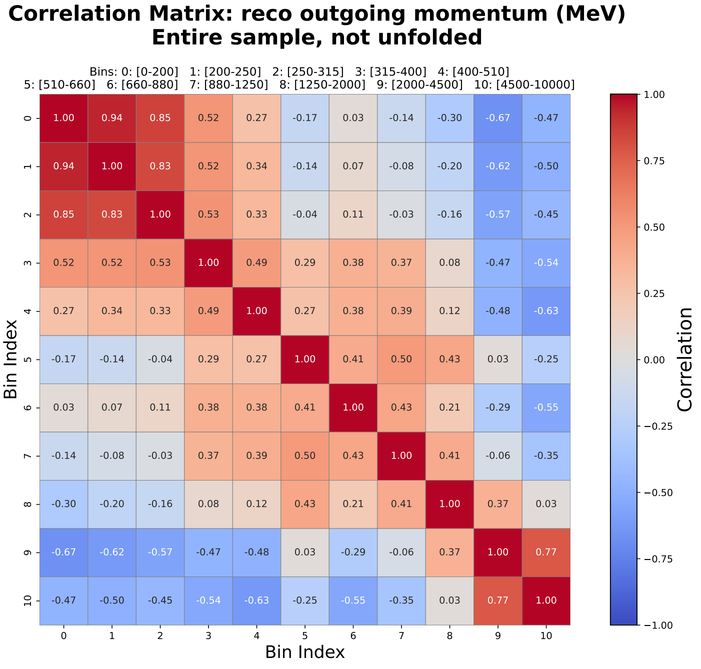
6.5.1.2 Scattering angle and total cross-section

The distribution of estimated scattering angle for all events in the entire selection is shown in Figure 6.35. We can see that the agreement between data and MC is generally poor with a reduced χ^2 of 110.5. Interestingly, the MC is consistent with the data in the four highest bins of scattering angle, which contain between 40% and 100% of background events. This indicates that the MC simulation is consistent with the data in background-dominated regions of the parameter space, an important sanity check.

We see more data-MC discrepancies in the low scattering angle bins. In the first bin $0 \leq \theta < 0.8^\circ$, we measure 5% less events in data compared to MC, significantly more than the sub-percent uncertainty on the number of events for this bin (see Figure 6.32a). In each of the following four bins, covering together the region $0.8^\circ \leq \theta < 3.85^\circ$, we measure about 4% more data than MC events, again significant with respect to the sub-percent uncertainty in these bins. The correlation matrix for this sample is given in Figure 6.36 where we can see that the first three bins, containing the vast majority of events are positively correlated with one another indicating that the shape discrepancy observed cannot be explained by systematic detector effects, unless there is a source of shape systematic that we have not accounted for. An example of such a source could be the ad-hoc energy loss correction introduced in Section 6.2.3.1, which was assumed to apply to both MC and data without further check. If this approximation is wrong, it could explain the MC-data discrepancy shown here, as the momentum estimate is directly used for calculating the scattering angle. This hypothesis is further strengthened by the



(a) Outgoing momentum

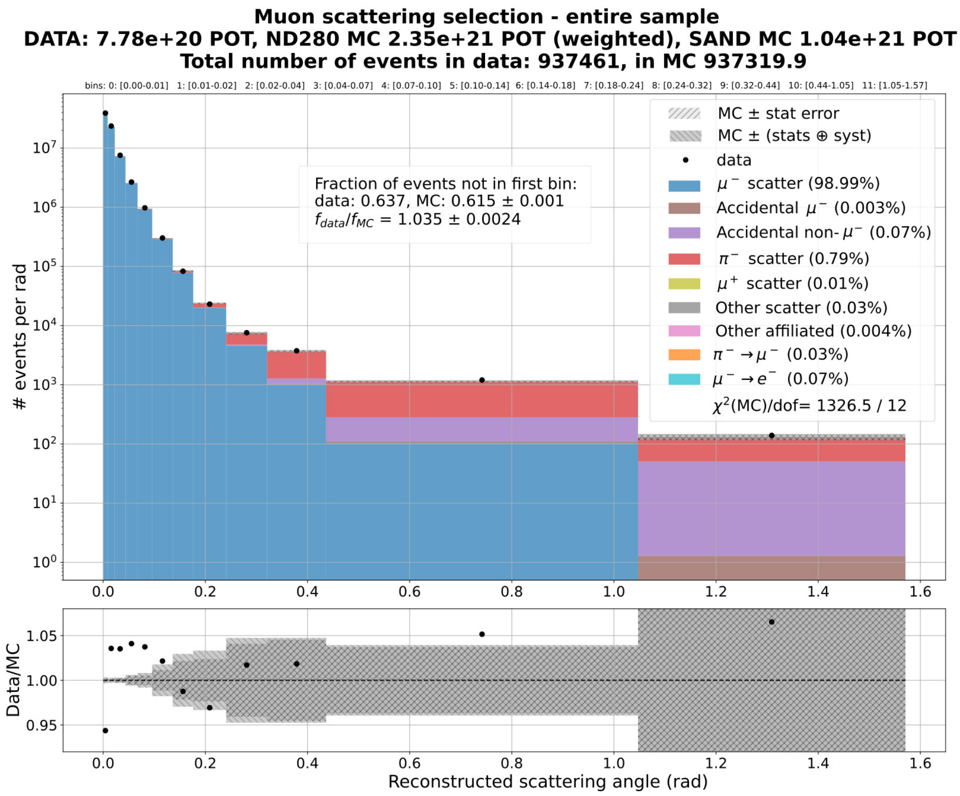


	0	1	2	3	4	5	6	7	8	9	10
0	1.00	0.94	0.85	0.52	0.27	-0.17	0.03	-0.14	-0.30	-0.67	-0.47
1	0.94	1.00	0.83	0.52	0.34	-0.14	0.07	-0.08	-0.20	-0.62	-0.50
2	0.85	0.83	1.00	0.53	0.33	-0.04	0.11	-0.03	-0.16	-0.57	-0.45
3	0.52	0.52	0.53	1.00	0.49	0.29	0.38	0.37	0.08	-0.47	-0.54
4	0.27	0.34	0.33	0.49	1.00	0.27	0.38	0.39	0.12	-0.48	-0.63
5	-0.17	-0.14	-0.04	0.29	0.27	1.00	0.41	0.50	0.43	0.03	-0.25
6	0.03	0.07	0.11	0.38	0.38	0.41	1.00	0.43	0.21	-0.29	-0.55
7	-0.14	-0.08	-0.03	0.37	0.39	0.50	0.43	1.00	0.41	-0.06	-0.35
8	-0.30	-0.20	-0.16	0.08	0.12	0.43	0.21	0.41	1.00	0.37	0.03
9	-0.67	-0.62	-0.57	-0.47	-0.48	0.03	-0.29	-0.06	0.37	1.00	0.77
10	-0.47	-0.50	-0.45	-0.54	-0.63	-0.25	-0.55	-0.35	0.03	0.77	1.00

(b) Correlation matrix

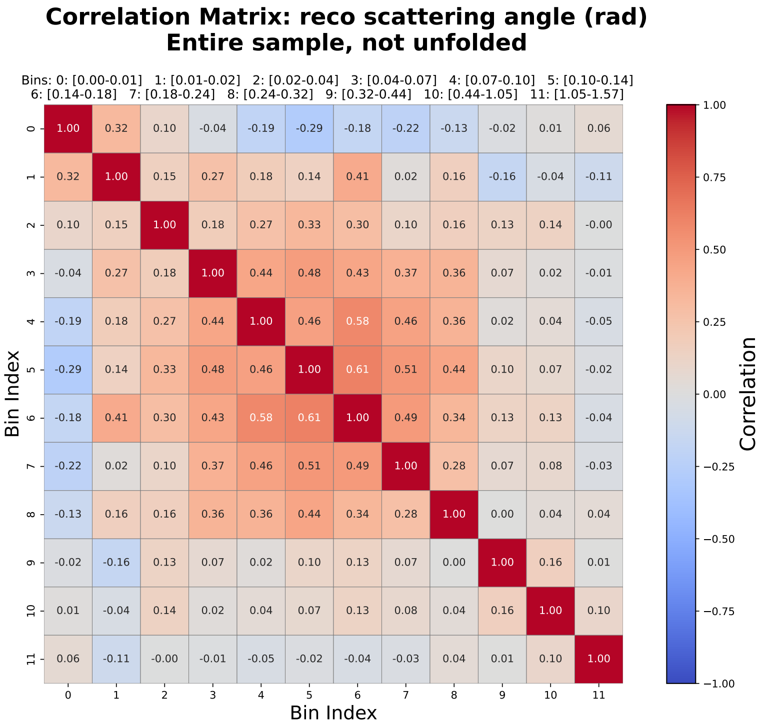
Figure 6.34: Data-MC comparison in bins of outgoing momentum in the entire sample presented alongside the corresponding correlation matrix.

shift in momentum presented in the previous section and motivates the implementation of a new systematic accounting for this effect.



Muon scattering selection - entire sample
DATA: 7.78e+20 POT, ND280 MC 2.35e+21 POT (weighted), SAND MC 1.04e+21 POT
Total number of events in data: 937461, in MC 937319.9

Figure 6.35: Comparison of scattering angles in the MC and data entire sample showing large discrepancies at low scattering angles.



Correlation Matrix: reco scattering angle (rad)
Entire sample, not unfolded

	0	1	2	3	4	5	6	7	8	9	10	11
0	1.00	0.32	0.10	-0.04	-0.19	-0.29	-0.18	-0.22	-0.13	-0.02	0.01	0.06
1	0.32	1.00	0.15	0.27	0.18	0.14	0.41	0.02	0.16	-0.16	-0.04	-0.11
2	0.10	0.15	1.00	0.18	0.27	0.33	0.30	0.10	0.16	0.13	0.14	-0.00
3	-0.04	0.27	0.18	1.00	0.44	0.48	0.43	0.37	0.36	0.07	0.02	-0.01
4	-0.19	0.18	0.27	0.44	1.00	0.46	0.58	0.46	0.36	0.02	0.04	-0.05
5	-0.29	0.14	0.33	0.48	0.46	1.00	0.61	0.51	0.44	0.10	0.07	-0.02
6	-0.18	0.41	0.30	0.43	0.58	0.61	1.00	0.49	0.34	0.13	0.13	-0.04
7	-0.22	0.02	0.10	0.37	0.46	0.51	0.49	1.00	0.28	0.07	0.08	-0.03
8	-0.13	0.16	0.16	0.36	0.36	0.44	0.34	0.28	1.00	0.00	0.04	0.04
9	-0.02	-0.16	0.13	0.07	0.02	0.10	0.13	0.07	0.00	1.00	0.16	0.01
10	0.01	-0.04	0.14	0.02	0.04	0.07	0.13	0.08	0.04	0.16	1.00	0.10
11	0.06	-0.11	-0.00	-0.01	-0.05	-0.02	-0.04	-0.03	0.04	0.01	0.10	1.00

Figure 6.36: Correlation matrix for the scattering angle observable in the entire sample.

We can compare the total muon scattering cross-section in data and MC by looking at the fraction f of events not in the first bin (i.e. scattered-like events). We can safely assume that the majority of events in the first bin (compatible with $\theta = 0$) are muons which have not scattered in FGD1, or whose multiple scatters destructively interfere, averaging to no direction change over FGD1. We make the approximation that all events in this first bin are unscattered muons. We measure the ratio of the fraction of events **not** in this first bin (i.e. the fraction of scattered events) in data and MC to be:

$$\frac{f_{\text{data}}}{f_{\text{MC}}} = 1.035 \pm 0.002 \tag{6.20}$$

The ratio between data and MC is incompatible with 1. Our measurement finds that the total muon scattering cross-section is under-estimated by $(3.5 \pm 0.2)\%$ in Geant4 compared to data. We do not expect this difference to arise from differences in momentum or angular reconstruction. This difference could be because Geant4 does not accurately model the scattering of low momentum muons, which is not surprising since it is primarily aimed at simulating higher energy scenarios where muons typically scatter less.

6.5.1.3 Comparison with 172 MeV/c μ-Carbon scattering

The muon scattering analysis presented here can be compared with previous experimental measurements of the scattering of 172 MeV/c muons on Carbon (corresponding to 1.53% radiation length) made by Attwood *et. al.* at TRIUMF in 2006 [115]. The two studies have a comparable amount of statistics (801,899 events for Attwood *et. al.* compared to 937,461 in the present analysis) and overlapping angular coverage, from -0.1 to 0.1 rad (from $-5.7°$ to $-5.7°$) for Attwood *et. al.* compared to -0.24 to 0.24 rad ($-13.8°$ to $13.8°$) in this analysis. In this section we compare the fraction of events which scatter at angles larger than 0.009 rad in the ND280 analysis and larger than 0.00895 rad in the study by Attwood *et. al.* For this comparison, we consider that these are scattered muons and compare the unfolded distribution of scattering angle presented in Figure 6.37 and Figure 6.38 for the ND280 and Attwood *et. al.* study respectively. The total number of data and MC events in each bin in the entire unfolded T2K sample is given in Table 6.7.

Because of the difference in muon momenta and total material budget one cannot make more than approximate comparisons between these two analyses. We account for the difference in scattering angle range by restricting our measurement to the $[-0.096$ rad, $+0.096$ rad$]$ range. We note that the fraction of scattered events f (i.e. events with $\theta > 0.009$ rad) in data and MC is slightly larger in this analysis than in ref [115] as presented in Table 6.8. This is expected from the rapid drop in muon total cross section at high momentum which is not fully compensated by the longer muon travel distance in Carbon (at least 33 cm in FGD1 compared to 2.50 mm in Attwood *et. al.*).

Interestingly, in the Attwood *et. al.* study, Geant4 over-estimates the total muon cross-section ($f_{\text{MC}}/f_{\text{data}} = 1.030 \pm 0.011$) whereas we see it under-estimates it in the present study ($f_{\text{MC}}/f_{\text{data}} = 0.971 \pm 0.003$). The

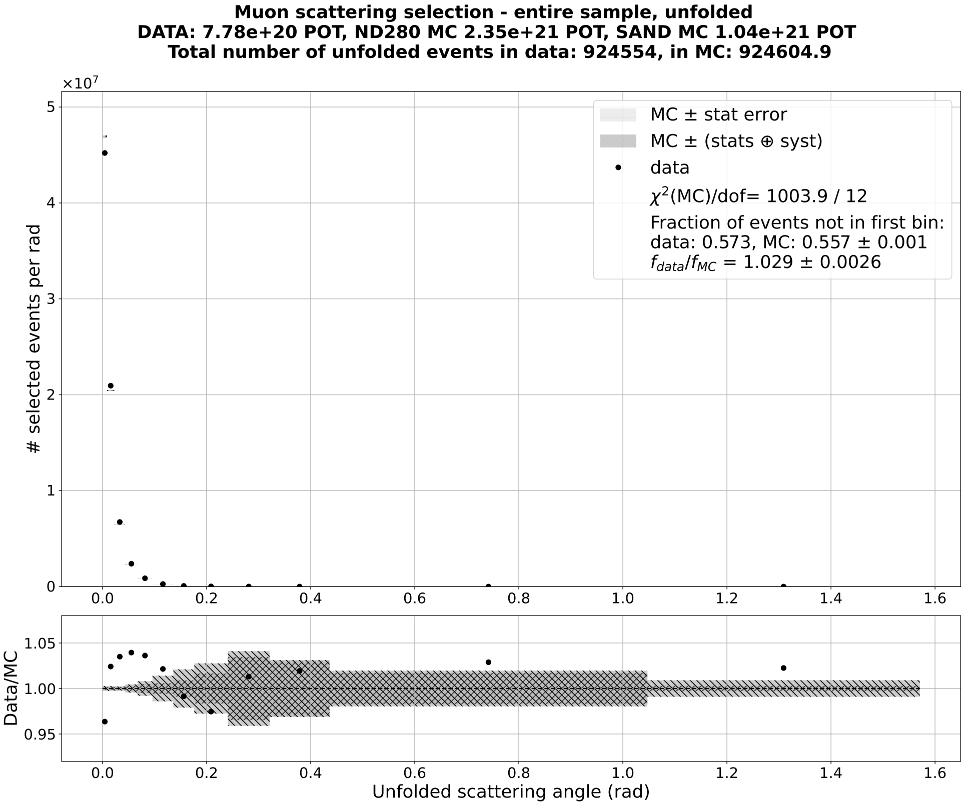


Figure 6.37: Unfolded distribution of scattering angles for the entire sample showing significant MC-data discrepancy at lower scattering angles.

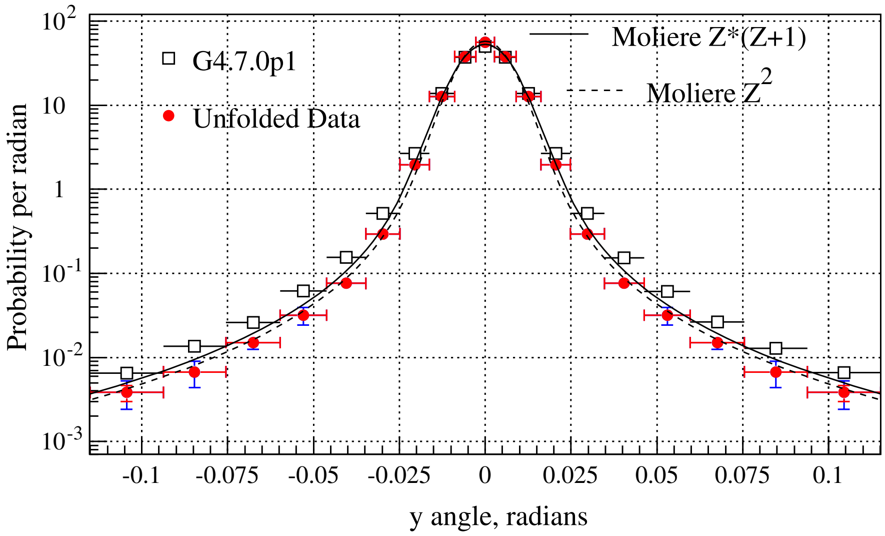


Figure 6.38: Probability for a 172 MeV/c muon to be deflected at a given angle when crossing 2.5mm of carbon, reproduced from [115].

MC-data discrepancies are of similar magnitudes in the two studies with values of $(+3.0 \pm 1.1)\%$ for Attwood *et. al.* and $(-2.9 \pm 0.3)\%$ in the present analysis. We expect that this discrepancy is momentum-dependent and that Geant4 models the interaction of high-momentum muons more accurately than those at lower momentum. The sample of through-going muons can be used to verify this hypothesis by further dividing it based on the

Bin	Bin range (rad)	Data Events	MC Events ± Uncertainty	(MC − Data)/Data (σ)
0	0.000 – 0.009	3.95×10^5	$4.09 \times 10^5 \pm 1.18 \times 10^3$	+3.6% (+12.0)
1	0.009 – 0.023	2.92×10^5	$2.86 \times 10^5 \pm 7.93 \times 10^2$	−2.1% (−8.3)
2	0.023 – 0.044	1.41×10^5	$1.36 \times 10^5 \pm 3.24 \times 10^2$	−3.5% (−15.0)
3	0.044 – 0.067	5.62×10^4	$5.40 \times 10^4 \pm 2.50 \times 10^2$	−3.9% (−8.7)
4	0.067 – 0.096	2.50×10^4	$2.41 \times 10^4 \pm 1.98 \times 10^2$	−3.6% (−4.3)
5	0.096 – 0.136	1.06×10^4	$1.04 \times 10^4 \pm 1.51 \times 10^2$	−2.0% (−1.6)
6	0.136 – 0.176	2.80×10^3	$2.83 \times 10^3 \pm 60.7$	+1.0% (+0.45)
7	0.176 – 0.241	1.19×10^3	$1.22 \times 10^3 \pm 34.7$	+2.6% (+0.89)
8	0.241 – 0.321	3.63×10^2	$3.58 \times 10^2 \pm 13.6$	−1.4% (−0.42)
9	0.321 – 0.436	1.59×10^2	$1.56 \times 10^2 \pm 5.19$	−2.1% (−0.63)
10	0.436 – 1.047	1.38×10^2	$1.34 \times 10^2 \pm 2.69$	−2.9% (−1.5)
11	1.047 – 1.570	6.79×10^1	$6.64 \times 10^1 \pm 0.617$	−2.2% (−2.4)

Table 6.7: Number of events in data and MC in the entire sample showing the fractional difference in each bin.

	$f_{\mathrm{data}}(\%)$	f_{MC} (%)	$f_{\mathrm{MC}}/f_{\mathrm{data}}$
ND280 analysis	56.6	55.0 ± 0.14	0.971 ± 0.003
Attwood *et. al.* [115]	61.4 ± 1.1	63.22 ± 0.02	1.030 ± 0.011

Table 6.8: Comparison of the fraction of events with $\theta > 0.009$ rad between this analysis and ref [115].

incoming momentum (essentially building pseudo mono-energetic muon samples) and measuring the MC-data discrepancy in total scattering cross section as a function of the muon momentum.

The study by Attwood *et. al.* uses Geant4 version 7.0p1 whereas the T2K study uses Geant4 version 9.4, which has been tuned using, amongst other results, the measurement by Attwood *et. al.*. In the light of the discrepancies observed between the two studies, one can hypothesise that the Geant4 tuning, and in particular the momentum scaling of the muon scattering cross-section, is inaccurate, motivating further measurements of sub-GeV muon scattering cross-section similar to the present work, ideally split into muon momentum slices.

6.5.2 Scattered sample

Comparing the data and MC scattered sample gives us further information on the modelling of scattering kinematics by Geant4.

6.5.2.1 Incoming and outgoing momentum

The distribution of incoming muon momentum in the scattered sample is shown alongside its correlation matrix in Figure 6.39. The first thing that we notice is that the scattering selection selects in total about 6.3% more events in MC compared to data (206,660 events in MC compared to 194,417 events in data).

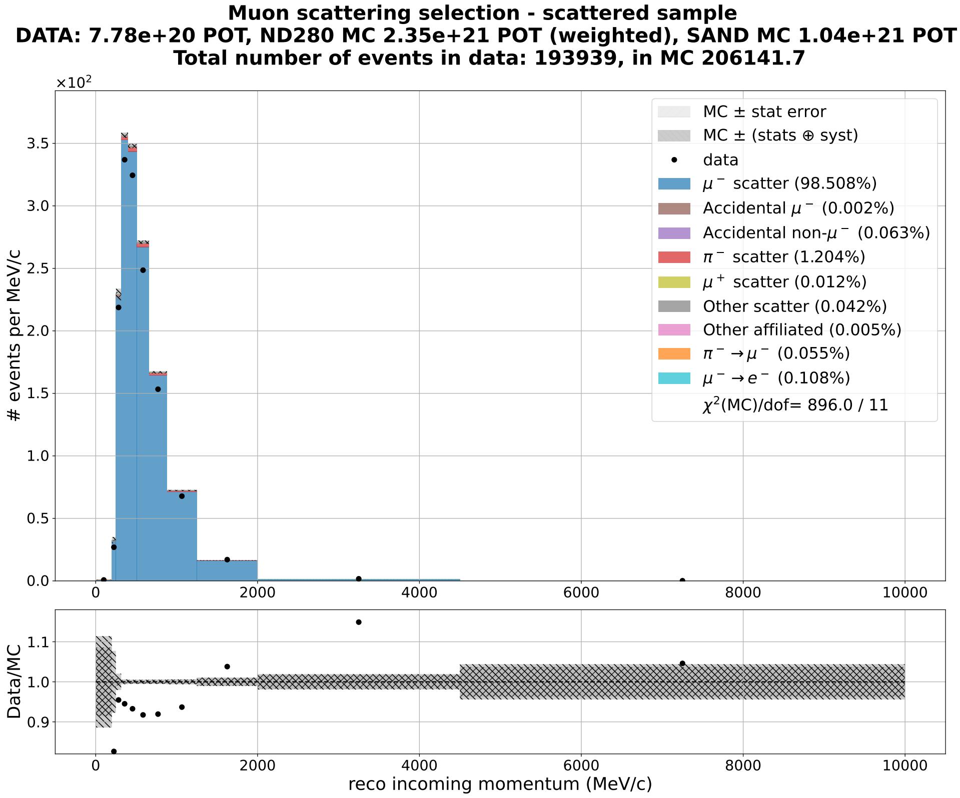
On top of this rate discrepancy, we observe a shape difference between the MC and data distributions. The peak in incoming momentum at about 400 MeV/c (bins 3 and 4) is narrower in MC compared to data. We read from the correlation matrix that bins 2 to 5 are mostly positively correlated with one another, except for bins 2 and 4, which have a small (-0.04) anti-correlation. The bin-to-bin correlations are not sufficient to explain this shape discrepancy. Remembering that the normalisation procedure described in Section 6.3.1 factors out any data-MC discrepancy in terms of total incoming momentum flux and incoming direction, we conclude that Geant4 significantly over-predicts the scattering cross-section for muons with momenta below 1 GeV/c. No known detector effects have been found to satisfactorily explain these rate and shape data-MC discrepancies.

Figure 6.40 shows the distribution of outgoing momenta and correlation matrix in MC and data in the scattered sample. We observe less shape discrepancy between the data and MC compared to the incoming momentum case but an overall slightly worse χ^2 agreement between the two distributions ($\chi^2(p_{\text{out}})/n_{\text{dof}} \approx 87$ and $\chi^2(p_{\text{out}})/n_{\text{dof}} \approx 81$). In this case, there is no significant difference in the peak width between the MC and data. We note a strong correlation between bins 0 to 2, which are before the peak, and that these bins are anti-correlated with bins after the peak (bins 3 to 8), which are, in turn, significantly correlated with each other. These correlations can easily cause a shift in the position of the peak as well as a broadening of it.

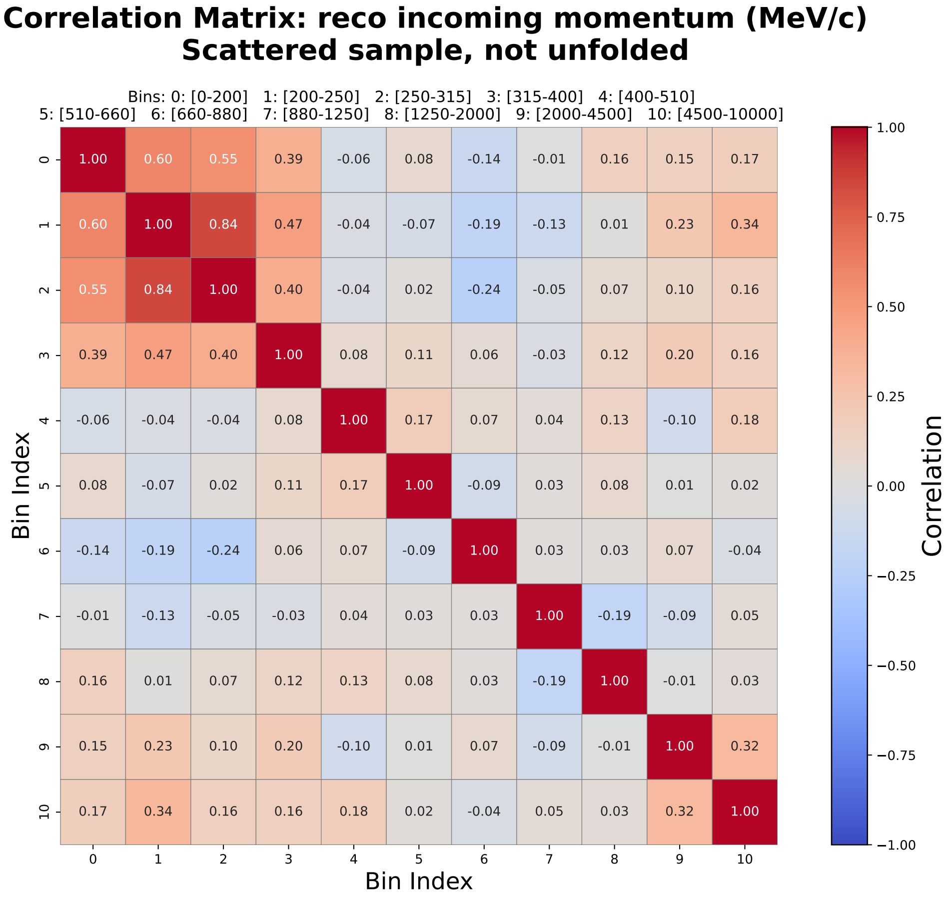
We observe once more that Geant4 over-predicts the scattering cross-section for GeV muons below 1 GeV/c. We note that the data-MC shape discrepancy observed in the incoming momentum distribution does not appear in the outgoing momentum distribution. This shape difference between the incoming and outgoing distribution is relatively small and can be explained by the detector smearing effects.

6.5.2.2 Scattering angle

We have observed significant rate discrepancies between Geant4 and data in terms of incoming and outgoing muon momentum. Figure 6.41 shows the distribution of scattering angles in the scattered sample and the associated correlation matrix. We note once more that there is both a shape and normalisation difference between the two distributions. In particular, the MC significantly over-predicts the rate of low-angle scattering, corresponding to the first and second bins covering [0.023 rad-0.044 rad] ([1.3° - 2.5°]) and [0.044 rad - 0.067 rad] ([2.5° - 3.85°]) respectively. There is good agreement between the data and the MC at higher scattering angles. It is interesting to compare the scattering angle distribution in the scattered sample with its equivalent in the entire sample shown in Figure 6.35. There we saw that Geant4 under-predicts (instead of over-predicts in the scattered sample) the number of events in the bins [0.023 rad - 0.044 rad], [0.044 rad - 0.067 rad] and [0.067 rad - 0.096 rad]. This shows that the overall data-MC rate difference in the scattered sample is mostly due to the PCA cut described in Section 6.2.4.1. Figure 6.12 showed that this cut removes about 30% more events from the data sample than from the MC one, which only is partially compensated by the larger total cross-section in data compared to MC reported in Equation 6.20. The correlation matrix shows that the shape discrepancy cannot be fully explained by the detector systematic effects.



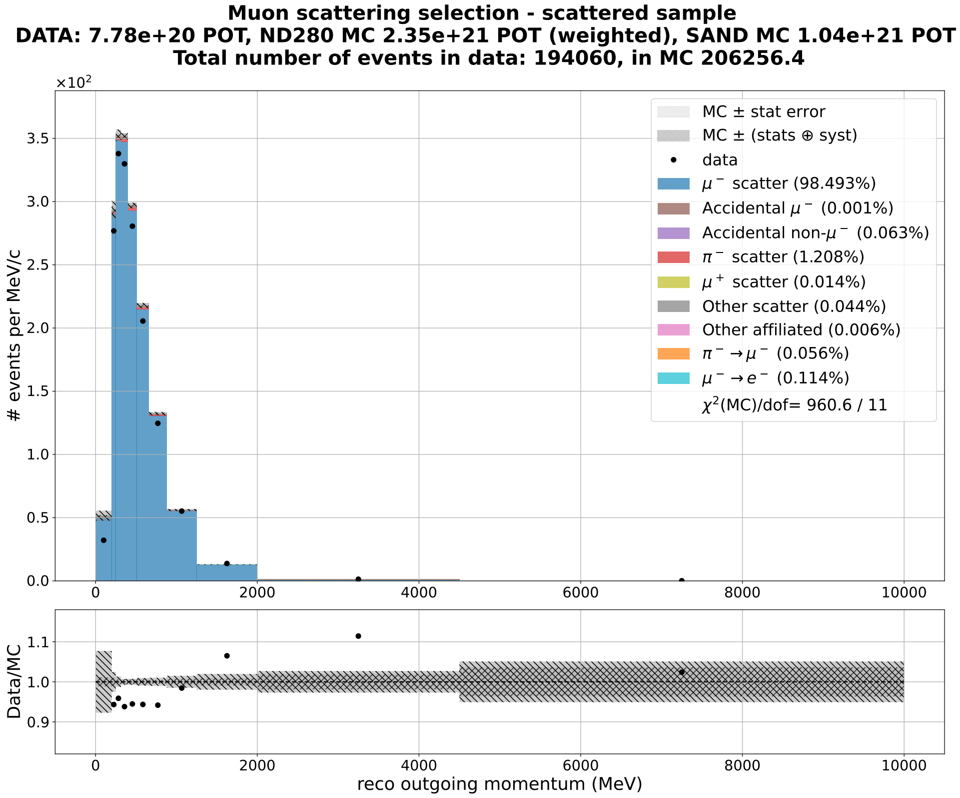
(a) Scattered sample: incoming momentum



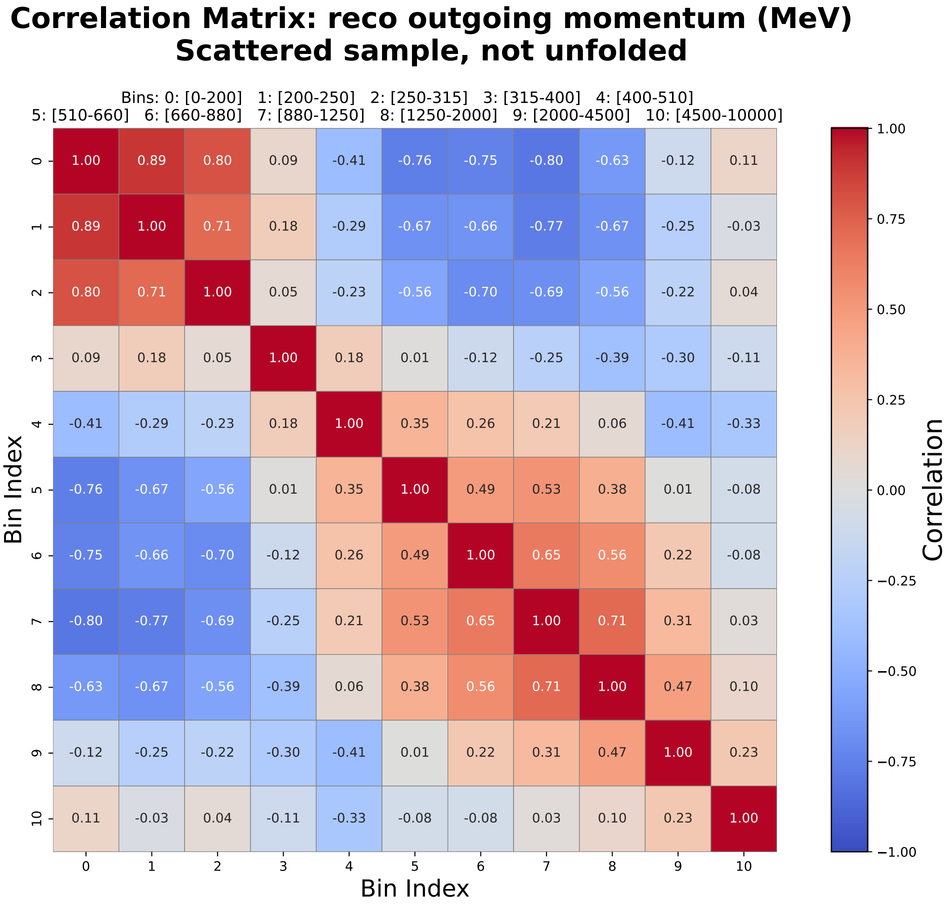
	0	1	2	3	4	5	6	7	8	9	10
0	1.00	0.60	0.55	0.39	-0.06	0.08	-0.14	-0.01	0.16	0.15	0.17
1	0.60	1.00	0.84	0.47	-0.04	-0.07	-0.19	-0.13	0.01	0.23	0.34
2	0.55	0.84	1.00	0.40	-0.04	0.02	-0.24	-0.05	0.07	0.10	0.16
3	0.39	0.47	0.40	1.00	0.08	0.11	0.06	-0.03	0.12	0.20	0.16
4	-0.06	-0.04	-0.04	0.08	1.00	0.17	0.07	0.04	0.13	-0.10	0.18
5	0.08	-0.07	0.02	0.11	0.17	1.00	-0.09	0.03	0.08	0.01	0.02
6	-0.14	-0.19	-0.24	0.06	0.07	-0.09	1.00	0.03	0.03	0.07	-0.04
7	-0.01	-0.13	-0.05	-0.03	0.04	0.03	0.03	1.00	-0.19	-0.09	0.05
8	0.16	0.01	0.07	0.12	0.13	0.08	0.03	-0.19	1.00	-0.01	0.03
9	0.15	0.23	0.10	0.20	-0.10	0.01	0.07	-0.09	-0.01	1.00	0.32
10	0.17	0.34	0.16	0.16	0.18	0.02	-0.04	0.05	0.03	0.32	1.00

(b) Scattered sample: incoming momentum correlation matrix

Figure 6.39: Comparison between the data and MC incoming momentum distributions in the scattered sample and associated correlation matrix.


Muon scattering selection - scattered sample
DATA: 7.78e+20 POT, ND280 MC 2.35e+21 POT (weighted), SAND MC 1.04e+21 POT
Total number of events in data: 194060, in MC 206256.4

(a) Scattered sample: outgoing momentum

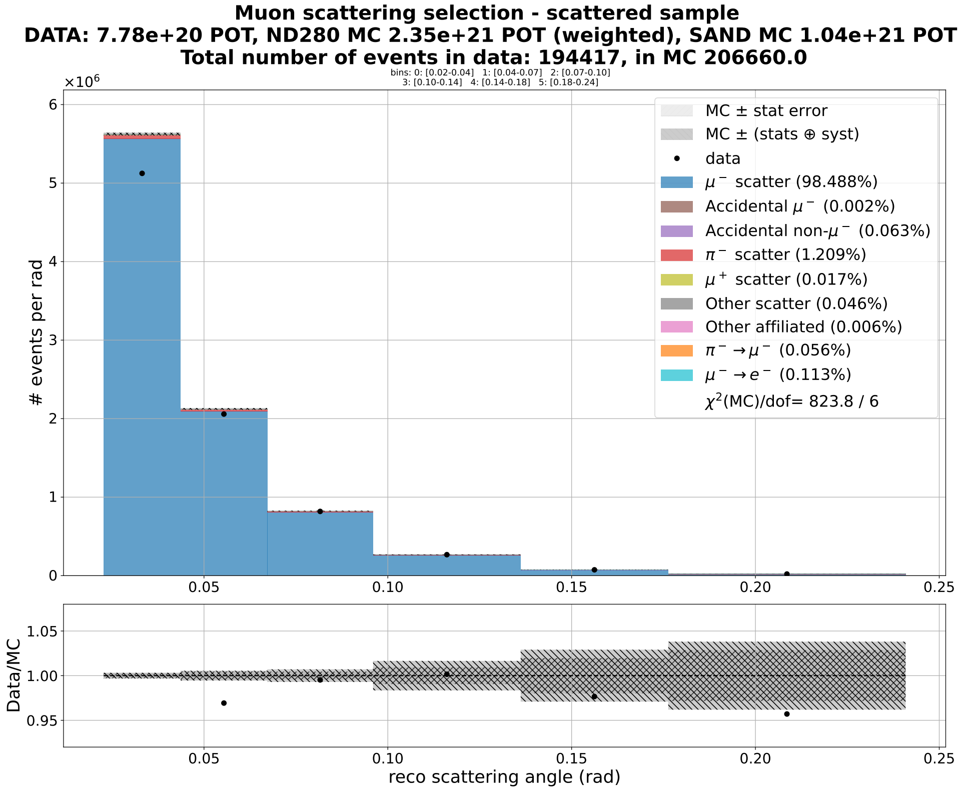

Correlation Matrix: reco outgoing momentum (MeV)
Scattered sample, not unfolded

Bins: 0: [0-200] 1: [200-250] 2: [250-315] 3: [315-400] 4: [400-510]
5: [510-660] 6: [660-880] 7: [880-1250] 8: [1250-2000] 9: [2000-4500] 10: [4500-10000]

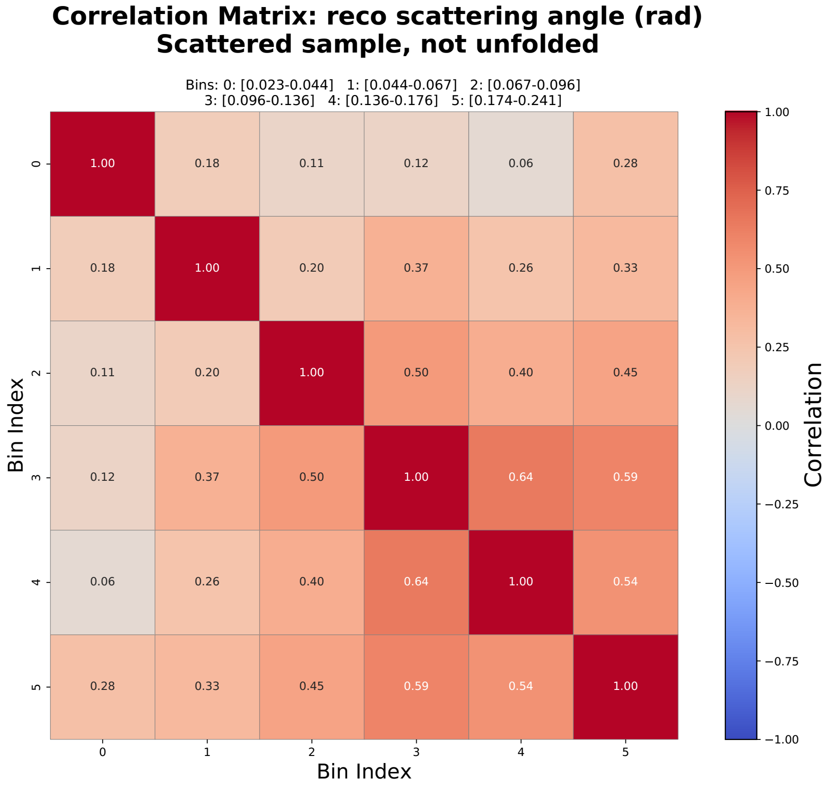
Bin Index	0	1	2	3	4	5	6	7	8	9	10
0	1.00	0.89	0.80	0.09	-0.41	-0.76	-0.75	-0.80	-0.63	-0.12	0.11
1	0.89	1.00	0.71	0.18	-0.29	-0.67	-0.66	-0.77	-0.67	-0.25	-0.03
2	0.80	0.71	1.00	0.05	-0.23	-0.56	-0.70	-0.69	-0.56	-0.22	0.04
3	0.09	0.18	0.05	1.00	0.18	0.01	-0.12	-0.25	-0.39	-0.30	-0.11
4	-0.41	-0.29	-0.23	0.18	1.00	0.35	0.26	0.21	0.06	-0.41	-0.33
5	-0.76	-0.67	-0.56	0.01	0.35	1.00	0.49	0.53	0.38	0.01	-0.08
6	-0.75	-0.66	-0.70	-0.12	0.26	0.49	1.00	0.65	0.56	0.22	-0.08
7	-0.80	-0.77	-0.69	-0.25	0.21	0.53	0.65	1.00	0.71	0.31	0.03
8	-0.63	-0.67	-0.56	-0.39	0.06	0.38	0.56	0.71	1.00	0.47	0.10
9	-0.12	-0.25	-0.22	-0.30	-0.41	0.01	0.22	0.31	0.47	1.00	0.23
10	0.11	-0.03	0.04	-0.11	-0.33	-0.08	-0.08	0.03	0.10	0.23	1.00

(b) Scattered sample: outgoing momentum correlation matrix

Figure 6.40: Comparison between the data and MC outgoing momentum distributions in the scattered sample and associated correlation matrix.



(a) Scattered sample: scattering angle



	0	1	2	3	4	5
0	1.00	0.18	0.11	0.12	0.06	0.28
1	0.18	1.00	0.20	0.37	0.26	0.33
2	0.11	0.20	1.00	0.50	0.40	0.45
3	0.12	0.37	0.50	1.00	0.64	0.59
4	0.06	0.26	0.40	0.64	1.00	0.54
5	0.28	0.33	0.45	0.59	0.54	1.00

(b) Scattered sample: scattering angle correlation matrix

Figure 6.41: Comparison between the data and MC scattering angle distributions in the scattered sample and associated correlation matrix.

6.5.2.3 Energy transferred

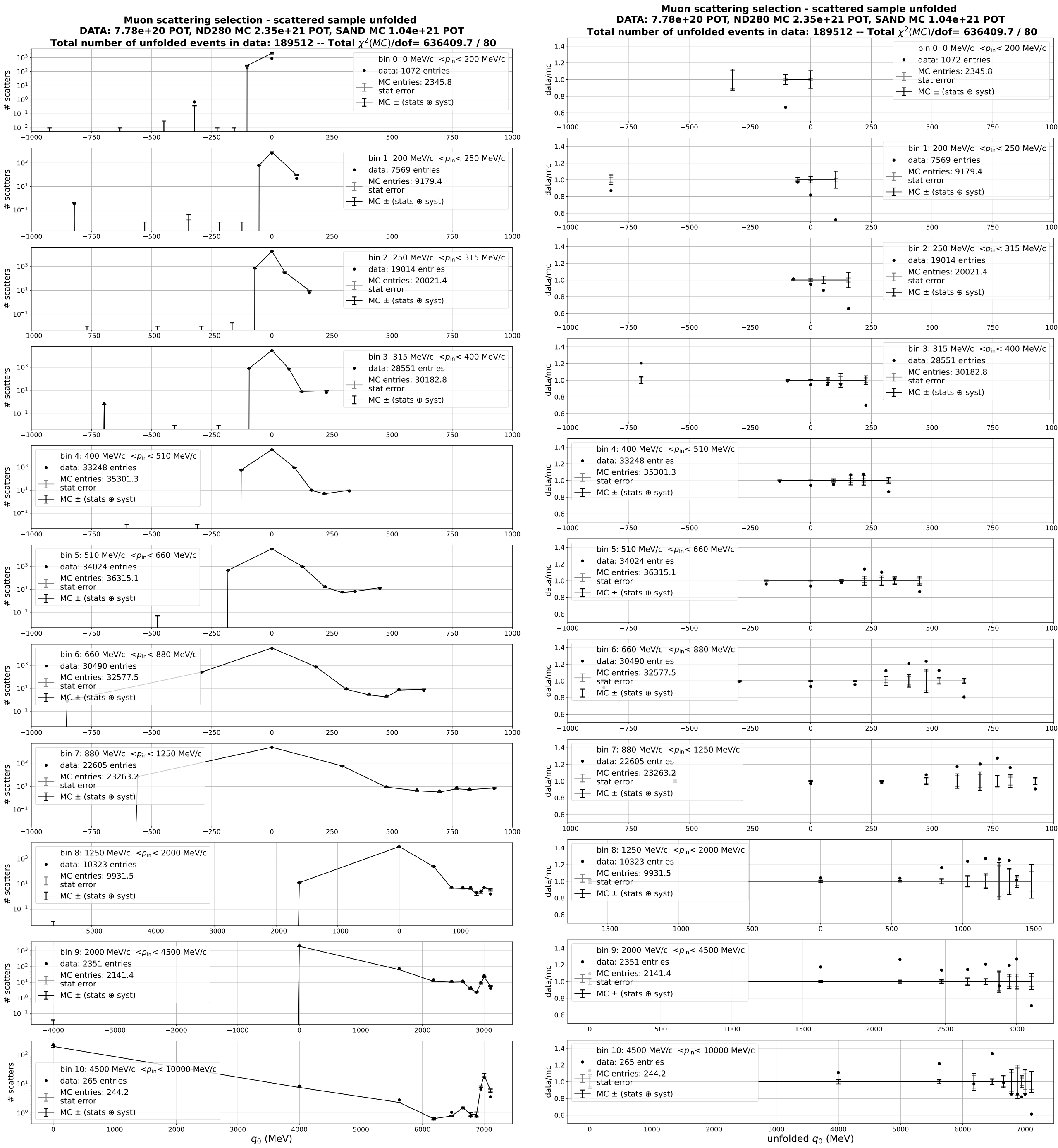
The amount of energy transferred is estimated from the incoming and outgoing momenta at the PCA as presented in Section 6.2.6.3. A comparison of the unfolded amount of energy transferred in the scatter is presented in Figure 6.42. The data and MC are presented in panels of incoming momentum at the PCA. We show the unfolded distribution to limit the contamination of mis-reconstructed events.

In Figure 6.42a we see that the majority of scatters do not transfer any energy (within our coarse detector resolution). These correspond to elastic and the lower tail of quasi-elastic scatters. We then note a dip region followed by an increase in the number of events at higher amount of energy transferred. This is first visible in the panel corresponding to bin 4 in incoming momentum ($400 < p_{in}^{@PCA} < 510$ MeV/c). There, we observe an increase in the number of events at $q_0 > 250$ MeV, which approximately corresponds to the turn-on point for resonant single pion production. Hints of this second peak is present in all subsequent momentum bins and gradually shift to higher values of q_0. The same behaviour is observed for SPP in electron scattering data [103], where we also note that the position of the q_0 peak depends on the scattering angle in electron scattering measurements. One could study this further by splitting up the selection based on the scattering angle of particles, thus obtaining a sample of approximately mono-energetic muons scattering at a given angle.

Looking at the ratio of events on Figure 6.42b we see that the MC mostly agrees with the data except in the dip region between the elastic (and quasi-elastic) peak and the possible pion production peak at higher q_0. This discrepancy between the data and MC in the dip region is about 10-20% which is significant compared to the uncertainty. This discrepancy in the dip region is observed in all bins with incoming momenta larger than 400 MeV/c. It can easily be explained by the fact that Geant4 does not explicitly model nucleon-nucleon correlations (either SRC or MEC) and instead treats nucleons as quasi-free inside the nuclear potential [196]. Geant4 uses a Bertini-style cascade where the low-energy photon (with $E_\gamma = q_0 < 10$ GeV) interacts directly with a single nucleon inside the nucleus, which is then propagated through the nucleus, accounting for interactions with other nucleons [119]. This method ignores all 2p2h contributions to the cross-section, leading to the observed under-prediction, whose magnitude is in line with what is expected from 2p2h contributions in this region.

The resonant pion production is modelled in Geant4 by the `G4PhotoNuclearProcess`, which simulates the excitation of a nucleon to a resonant state followed by its decay into pions and nucleons [118]. We see that the MC is above the data at higher values of q_0, which we have identified as the pion peak. This could indicate a mis-modelling of resonant pion production in Geant4. The precision of this measurement and the range of q_0 it covers would have to be increased to study this discrepancy in more depth.

Despite the unfolding, we observe a few events with a negative amount of energy transferred, corresponding to cases where the outgoing momentum is higher than the incoming one. These events are thought to correspond to issues in the propagation and estimation of the PCA procedures, which can be due to detector limitations



(a) Unfolded transferred energy (b) Unfolded transferred energy, data/mc ratio

Figure 6.42: Data-MC comparison of the amount of energy transferred for the scattered sample after unfolding. The MC prediction is shown as a solid line with error bars.

(e.g. incorrect dE/dx correction) or genuine physics effects (e.g. multiple scatters instead of a single one, biasing the reconstruction of the PCA). We note that the MC-data discrepancy in bins of negative energy transferred is not very large.

6.6 Data-MC-NEUT comparison

The modifications applied to the NEUT framework to simulate the quasi-elastic scattering of muons were presented in Section 2.2.2. To compare with the data and MC, we generate one million quasi-elastic muon-^{12}C scatterings in NEUT. The incoming flux from which the initial muon momentum is drawn is taken from the MC simulation. We ignore the contribution of Hydrogen to the scattering cross-section as it is a much smaller component of the scattering target mass than Carbon.

The muon scattering selection presented above is an inclusive measurement. We unfold the data and MC distributions to remove detector smearing. As discussed above, we believe that the large majority of events in the scattered sample have elastically scattered, with a small number of muons having quasi-elastically scattered, scattered off a correlated nucleon pair or produced pions. This means that the experimental NEUT model, which only includes quasi-elastic interactions, is not expected to model the data very well.

We estimate the agreement between the data and NEUT predictions using:

$$\chi^2 = (\mathbf{data} - \mathbf{NEUT})^{\mathrm{T}}(C + S)^{-1}(\mathbf{data} - \mathbf{NEUT}) \tag{6.21}$$

analogous to Equation 6.19 with $S = diag\left(\sqrt{n_i^{\mathrm{NEUT}}}\right)$ is a diagonal matrix accounting for the statistical error in each bin i in the NEUT sample.

6.6.1 Momentum

As described in Section 2.2.2, the NEUT framework samples the incoming momentum flux according to the total cross-section for a given process, where the cross-section for muon quasi-elastic scattering was presented in the same section, reproduced in Figure 6.43 for convenience. Figure 6.44 shows the distribution of incoming and outgoing momenta for the data, MC and NEUT scattered sample distributions. The total number of NEUT entries has been normalised to the total number of data in the scattered sample.

We observe that, as expected, the NEUT momentum distribution strongly differs from the data and Geant4 simulation. On the one hand, there is significantly less events with momenta lower than 660 MeV/c in NEUT compared to in the data. On the other hand, there are significantly more events with momenta higher than 880 MeV/c in the NEUT distribution compared to the data. This is easily explained by the momentum dependence of the different scattering processes: the lower momentum region is dominated by elastic scatterings, with the quasi-elastic cross-section rapidly dropping at energies below about 600 MeV, as shown on Figure 6.43. To the contrary, the quasi-elastic cross-section is high and almost flat at energies higher than about 800 MeV, causing NEUT to sample more from the high momentum region, producing the observed difference. The agreement

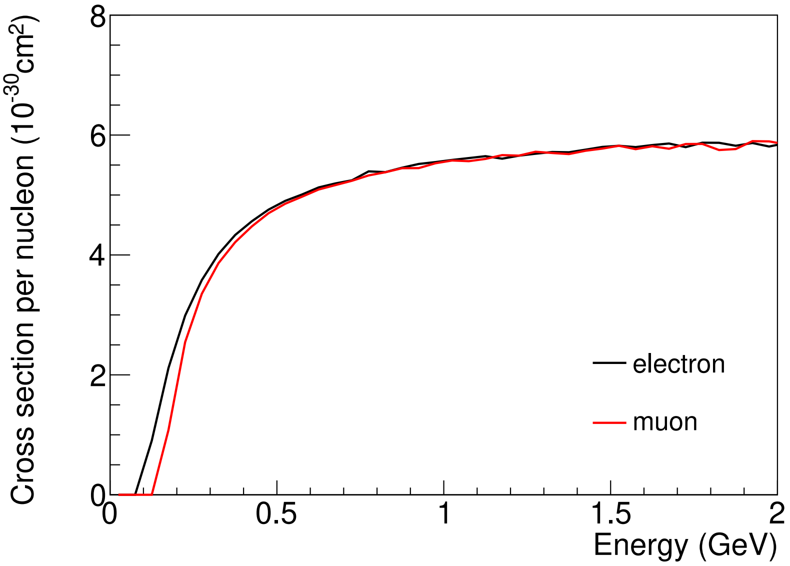


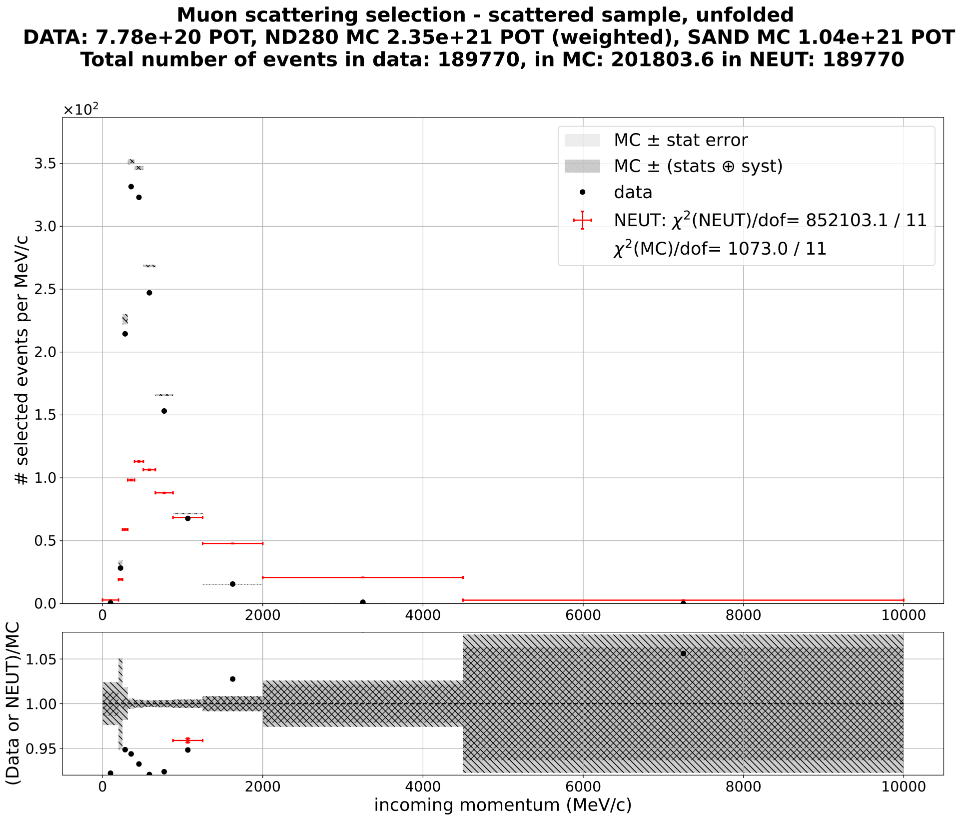
Figure 6.43: Total quasi-elastic cross-section as calculated by NEUT for electrons and muons.

between the data and both the MC and NEUT predictions is very poor with $\chi^2_{\mathrm{Geant4}} \sim 100$ and $\chi^2_{\mathrm{NEUT}} > 10^5$ respectively.

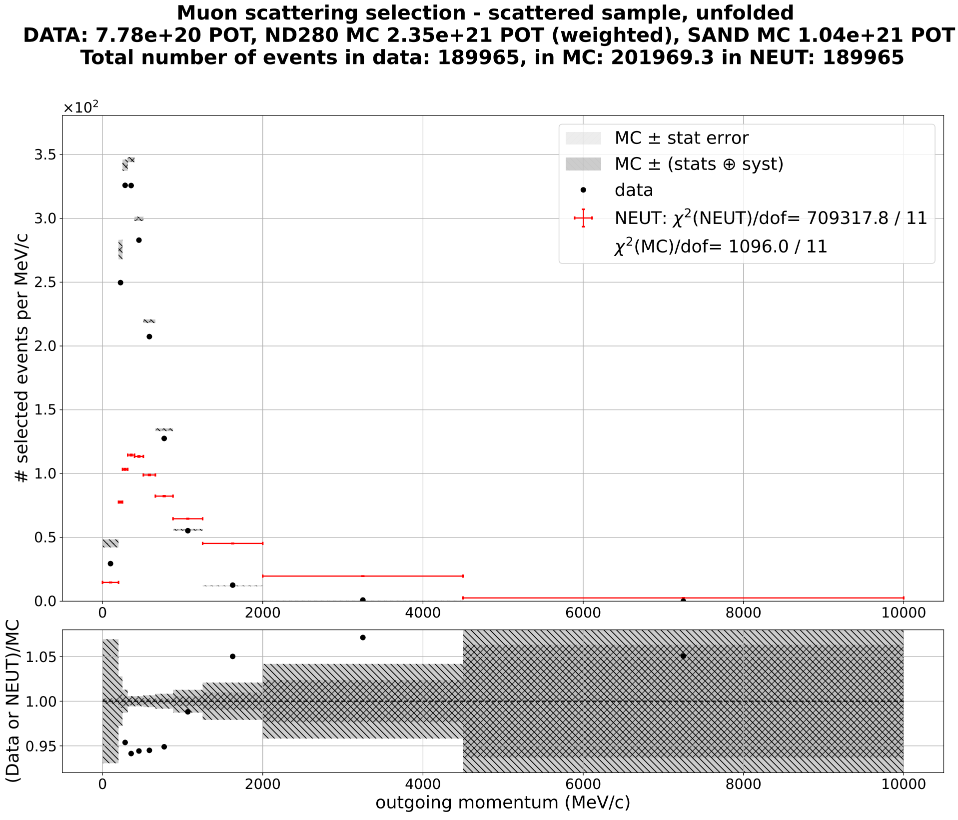
6.6.2 Scattering angle

As presented above, the scattering processes considered by NEUT and the data selections are different. Because more energy is transferred in quasi-elastic scatters, the muon is more likely to be deflected at a higher angle compared to the elastic case. We observe indeed on Figure 6.45 that the NEUT distribution of scattering angles is broader than the data and MC distributions with fewer events in the peak (i.e. between 0.02 rad and 0.07 rad) and more events with higher scattering angles.

Interestingly, we also note that the NEUT distribution peaks at a slightly lower scattering angle compared to the data. We believe that this is due to our definition of a scattered event, which imposes that the reconstructed scattering angle is greater than 0.02 rad. This means that all events in the first and second bin on Figure 6.45 are events with higher reconstructed angles unfolded back into bins of smaller true scattering angle. This seems to indicate that the scattered sample selection rejects some events which have genuinely scattered. Because this low-scattering angle region is expected to correspond mostly to elastic scatters, one needs to implement this scattering process in NEUT to make more precise comparisons.



(a) Scattered sample: unfolded incoming momentum



(b) Scattered sample: unfolded outgoing momentum

Figure 6.44: Comparison between the data, MC and NEUT of the incoming and outgoing muon momenta.

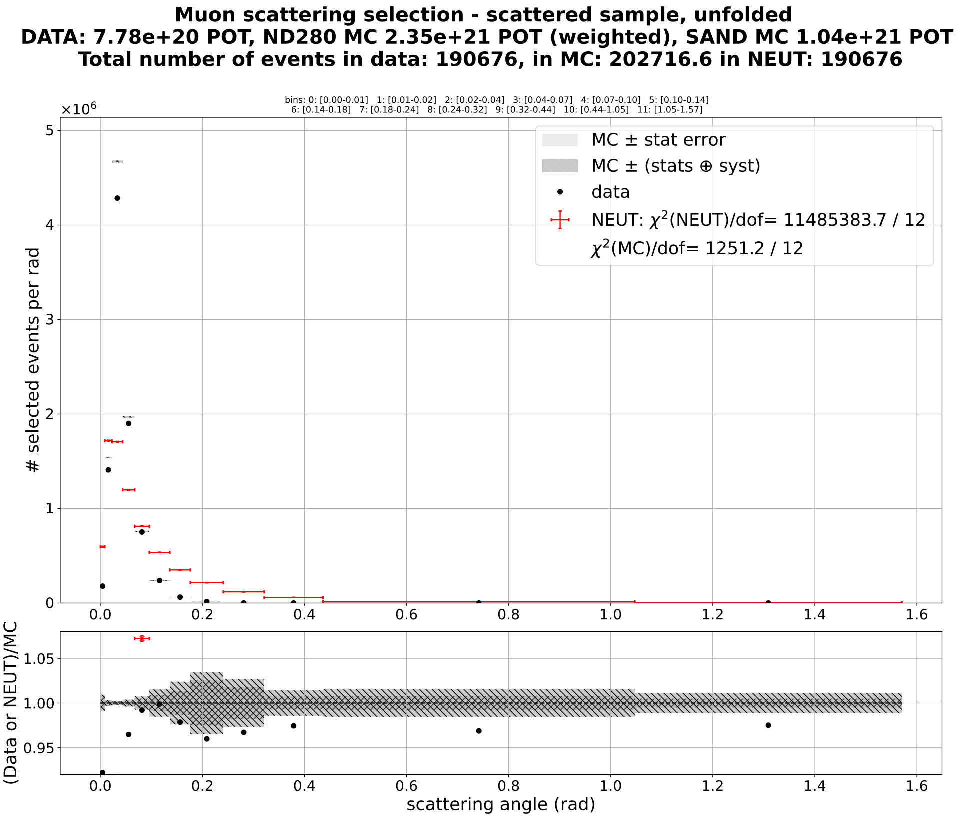


Figure 6.45: Unfolded distribution of scattering angles for the scattered sample.

6.6.3 Energy transferred

As mentioned earlier, the amount of energy transferred strongly depends on the scattering process considered, with elastic scatterings transferring typically less energy than quasi-elastic scatterings, for example. Figure 6.46 presents a comparison of the energy transferred in muon-carbon scatterings as modelled by NEUT and Geant4 (MC) with the data. Once again, we compare NEUT with the scattered sample after unfolding. We use the method presented in Section 6.2.6.3 to estimate the NEUT prediction of the transferred energy in the same way as is done for the data and MC samples.

Both the NEUT and Geant4 predictions are normalised such that they have the same number of events as the data in the first bin where $q_0 > 0$ MeV for each incoming momentum value. This bin, typically corresponding to $q_0 > 50$ MeV, is assumed to contain mostly quasi-elastic events. Elastic events are thought to be mostly contained in the $q_0 = 0$ bin. Events with $q_0 \leq 0$ are not included in this comparison.

We note that the NEUT prediction is above the data at relatively low amounts of energy transferred, i.e. in the second non-zero q_0 bin (i.e. first non-normalised bin) for each momentum panel below 1 GeV/c. We note that the discrepancy between the data and NEUT prediction in these bins reduces as the momentum increases, reaching approximately zero for muons with incoming momenta between 1.25 GeV/c and 2 GeV/c and crossing zero afterwards, with NEUT actually under-predicting the number of events in that q_0 bin for incoming momenta between 2 GeV/c and 4.5 GeV/c. Both of these effects can be explained by the fact that a significant fraction of events in the first $q_0 > 0$ MeV bin are actually elastic scatters, contrary to the assumption made above. The normalisation presented above therefore overestimates the number of quasi-elastic scatters, therefore broadening the quasi-elastic peak, something that is not seen in the data.

We note that NEUT under-predicts the number of events with a large amount of energy transferred, corresponding to resonant pion production, not simulated in this version of NEUT. Work is currently ongoing to adapt NEUT models of pion production (e.g. the DCC and HNVL models) to permit a comparison with the data in this higher q_0 region.

6.7 Summary and outlook

The muon scattering analysis presented in this Chapter is the first of its kind. We first described how we can select just under a million negative muons in the T2K dataset. These muons cross the TPC1, FGD1 and TPC2 sub-detectors with momenta between about 200 MeV/c and 4.5 GeV/c. Out of these, we have identified a subset of about 200,000 muons which we reconstruct as likely to have scattered in FGD1. We went on to estimate the amount of energy transferred in the scatter based on the momentum measured by TPC1 and TPC2. Both samples were found to have a very high purity ($> 98\%$), with the dominant background, pion scatters, representing about 80% of all backgrounds.

The muon flux was measured before crossing the hydro-carbon target, allowing us to normalise the MC to the incoming data in terms of incoming momentum and angle, removing the flux uncertainty. The total uncertainty on this measurement was found to be small, on the order of a few percent and dominated by TPC-related systematic uncertainties. The smearing of the observables caused by limited detector resolution was factored out using a simple unfolding procedure.

The data-MC comparison for the entire sample of muons showed some interesting discrepancies. In particular, Geant4 was observed to under-predict the total muon cross-section by $(2.9 \pm 0.3)\%$. A discrepancy of similar magnitude in the opposite direction was observed by Attwood *et. al.* in the scattering of 172 MeV/c muons [115].

The data-MC comparison in the scattered sample showed some rate discrepancies (6% more MC events than data events were selected) and some shape discrepancies in the incoming momentum, outgoing momentum and scattering angles distributions. These differences could not be explained by detector effects alone. The distribution of energy transferred in the scatter showed a clear elastic peak at q_0 and a probable single pion production peak. A good agreement between the simulation and the data was observed in the elastic scattering peak. Geant4 was observed to under-predict the number of events by about 10-20% in the dip region between these

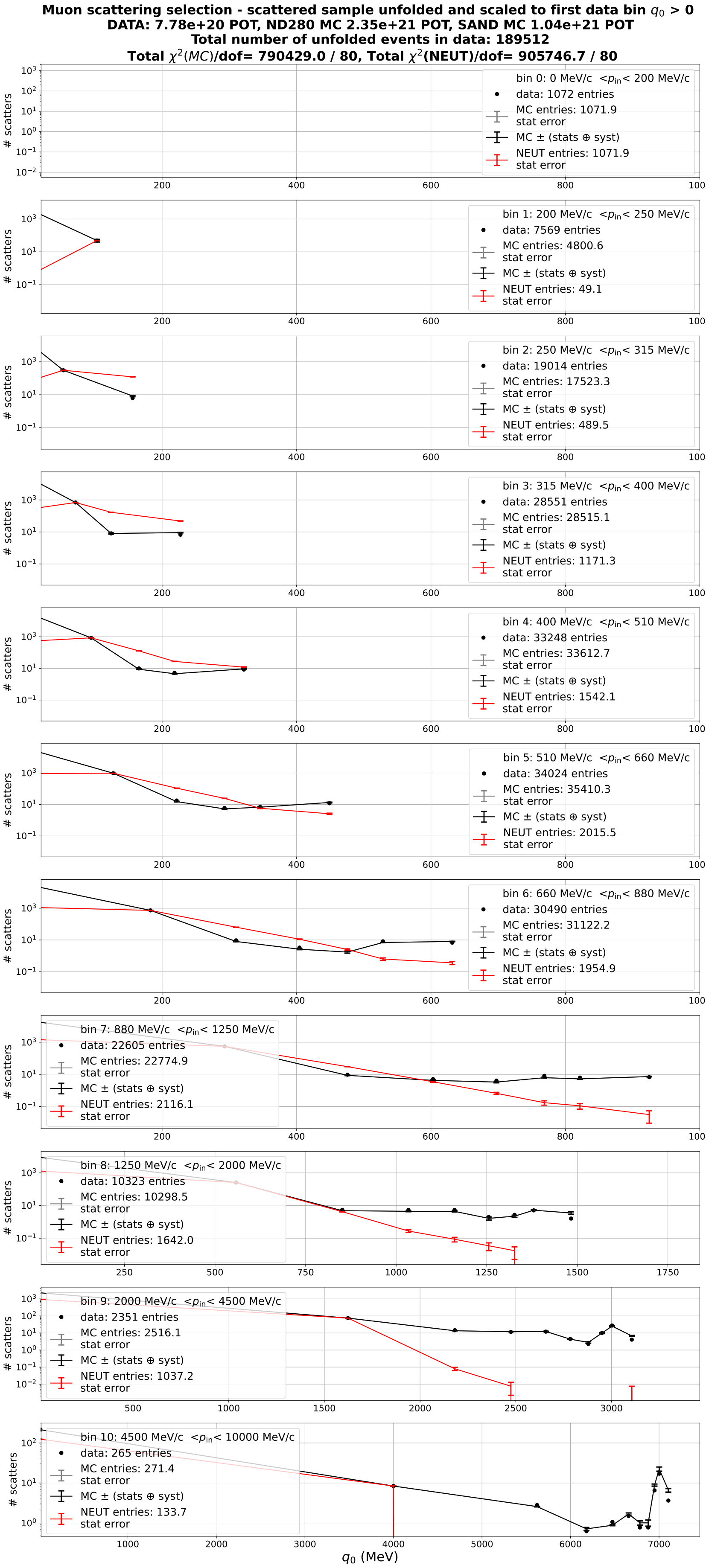


Figure 6.46: Comparison of the distribution of q_0 in muon scatters. The MC and NEUT predictions are normalised to the first non-zero q_0 bin.

two peaks, which typically corresponds to the muon interacting with a pair of correlated nucleons, something that is not modelled by Geant4. We observed that Geant4 over-estimates the number of events in the single pion production peak by about 15%.

Finally, we presented a comparison of the data with predictions of muon quasi-elastic scattering made using an experimental version of the NEUT neutrino event generation framework. The discrepancies observed with the data could all be explained by the fact that the data sample also contains elastic, 2p2h and SPP scatterings instead of only quasi-elastic scatters.

The analysis presented here opens the door to a new range of studies, presented in Table 6.9, that could be made using the already available T2K data prior and post ND280 upgrade, which would directly contribute to the improvement of our neutrino interaction models, reduce the associated systematic uncertainty and help future generation experiments to reach their Physics goals.

Measurement	**Use**	**Comments**
Combined μ^--^{16}C and μ^--^{16}O	Initial nuclear state modelling	Using TPC2, FGD2 and TPC3
μ^+ scattering	$\sigma_{\nu_\mu}/\sigma_{\bar{\nu}_\mu}$ (Radiative corrections)	Similar to [75], measurement of the Coulomb field distortion)
$\pi^\pm$ scattering	Secondary interactions	Easily adaptable from the current analysis, large muon contamination expected. ND280 Upgrade is a good candidate
$\mu^\pm + A \to \mu^\pm + A + p$	Neutrino modelling	Quasi-elastic channel, allowing direct comparison with NEUT
$\mu^\pm + A \to \mu^\pm + A + \pi$	Neutrino modelling	Modelling of SPP and FSI
$\mu^\pm + A \to \mu^\pm + A + \gamma$	Real photon production (radiative corrections)	Photons are difficult to detect at ND280

Table 6.9: Summary of some of the measurements that could be made using the T2K charged particle data and how they could benefit the overall systematic reduction in neutrino oscillation experiments. Most of these measurements will be performed on Oxygen by the WCTE collaboration.

Chapter 7

Conclusions and Outlook

As the next generation of long-baseline neutrino experiments starts taking data, the need to reduce the systematic uncertainty on the oscillation measurements is becoming a priority in the field of neutrino physics. This thesis presented two studies directly working towards the reduction of these uncertainties.

The first one, presented in Chapter 4, developed the method for charged particle identification and momentum measurement for the WCTE experiment. This study measured the PID capability of novel Aerogel Cherenkov Threshold detectors in the sub-GeV range. It also calibrated the energy of photons produced in the tagged photon setup, measuring its energy resolution.

The characterisation of the beam presented in this chapter was essential to optimise WCTE's data taking and allow it to maximise its physics output using the limited amount of beam time available. It also represents the first characterisation of the CERN T09 beam since its renovation, helpful to any other user of the beam in the sub-GeV momentum range. The methods and tools developed for this analysis are used in WCTE's 2024 and 2025 operations and acts as truth-level information on particles' nature and momentum. This information is necessary for all of WCTE's analyses, which include electron, muon, pion and photon scattering studies, in addition to the validation of existing and development of new particle identification tools for water Cherenkov detectors and the study of detector response and energy reconstruction. All of these studies directly contribute to reducing the systematic error on detector response, secondary interactions and/or neutrino scattering models, for current and next-generation neutrino oscillation experiments. The work presented in this thesis has already been continued by WCTE collaborators using the 2024 and 2025 data runs. In particular, improvements in the selection purity were achieved using non-linear cut lines and after removing the reflector film, which was identified by this analysis as a scintillator. Additionally, a significant amount of background was reduced thanks to tighter timing cuts and a better understanding of the backgrounds, thanks to dedicated studies built upon the outcome of the work presented here. WCTE completed its taking data at the beginning of June and is currently being disassembled. Some of its components (mPMTs, calibration systems) are directly shipped to Japan and will become part of the IWCD and HK detectors. Over the next few months and years, the WCTE collaboration will analyse its data and publish its findings in time for the start of HK's data-taking.

The second analysis presented in this thesis is a measurement of the inclusive muon scattering cross-section on carbon. This measurement is the first of its kind in the 200 MeV/c to 4.5 GeV/c momentum range. It uses two ND280 TPCs to characterise the flux of muons before and after crossing a 33 cm hydrocarbon target, the FGD1 sub-detector. A total of just under one million muons were selected with a purity higher than 98%. From this sample, 200,000 events were identified as having scattered in the target. A comparison of the data with the Geant4 prediction revealed significant discrepancies, showing the inaccuracies of the Geant4 simulation of low-momentum muon scattering. In particular, the Bertini model in Geant4 was observed to under-estimate the total muon cross section by $(2.9\pm0.03)\%$. This discrepancy is of a similar magnitude and opposite direction to the one observed by Attwood *et. al.* in the scattering of 172 MeV/c muons [115]. The uncertainty on the ND280 measurement is a third of the uncertainty on the measurement by Attwood *et. al.*

A more detailed comparison between the data and MC in the scattered sample showed some significant shape differences in the momentum and scattering angle distributions that could not be explained by detector effects. The distribution of energy transferred showed a 10-20% discrepancy in the dip region between the elastic and single pion production peaks, corresponding to an underestimation by Geant4 of the number of interactions with nucleons in correlated pairs, a known limitation of Geant4 in this energy range [119]. The number of events in the range of energy transfer corresponding to single-pion production was larger in Geant4 than in the data by about 15%. Finally, a comparison between the data and an experimental version of NEUT showed significant but explicable differences due to the inclusive nature of the measurement.

This novel dataset opens the way to testing neutrino event generators with muons, similarly to what has been done with electrons. It can be used to directly verify the muon radiative corrections in this energy range, an important contributor to the uncertainty on the electron-to-muon neutrino cross-section ratio, which limits HK's resolution on δ_{CP} [55] and delays its discovery of CP violation in the lepton sector, if it exists. This dataset can also be used to identify and address limitations of the available simulations of low-momentum muon scattering in preparation for a future muon collider.

The next steps for this work consist in further developing the NEUT framework to include elastic scattering, 2p2h, and single pion production in muon scattering events, thus allowing a direct validation of the models and tools used in neutrino event generators. Alternatively, one could adapt the inclusive selection described here, requiring, for example, that there is one proton track and no pion track, to produce a sample of quasi-elastic muon scatterings that can be directly compared with the NEUT predictions. The upgraded ND280 detector is very well suited to perform such a measurement, thanks notably to the high nucleon and pion detection capability of the super Fine Grained Detector (sFGD), and the new top and bottom high-angle TPCs, which permit a measurement of the scattering of cosmic muons in sFGD. Furthermore, the tools developed in this analysis can be used to study the scattering of other charged particles, for example pions, whose secondary interactions are a significant source of uncertainty in neutrino oscillation measurements. This measurement is complementary to the measurement of the muon scattering cross-section in oxygen by WCTE.

Bibliography

[1] S Navas and others (Particle Data Group). “Review of Particle Physics”. *Physical Review D* 110.3 (2024), p. 30001. DOI: `10.1103/PhysRevD.110.030001` (page 1).

[2] K Abe and others. “Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target”. *Eur. Phys. J. C* 83.9 (2023), p. 782. DOI: `10.1140/epjc/s10052-023-11819-x` (pages 1, 8, 9, 12, 13, 18, 84, 85, 88, 90, 91).

[3] A D Sakharov. “Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe”. *Sov. Phys. Usp.* 34.5 (1991), p. 392. DOI: `10.1070/PU1991v034n05ABEH002497` (page 1).

[4] Hyper-Kamiokande Proto-Collaboration et al. *Hyper-Kamiokande Design Report.* 2018. URL: `https://arxiv.org/abs/1805.04163` (pages 1, 6, 7, 10, 23, 27).

[5] R Acciarri and others (DUNE Collaboration). “Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 1: The LBNF and DUNE Projects”. *arXiv preprint* (2016). URL: `https://arxiv.org/abs/1601.05471` (page 1).

[6] T. Adam et al. “JUNO Conceptual Design Report” (Aug. 2015). URL: `http://arxiv.org/abs/1508.07166` (pages 1, 6, 7).

[7] Collaboration WCTE. *Water Cherenkov Test Experiment (WCTE) Annual Report.* Tech. rep. Geneva: CERN, 2023. URL: `https://cds.cern.ch/record/2857041` (pages 1, 24).

[8] K Abe et al. “The T2K experiment”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 659.1 (2011), pp. 106–135. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2011.06.067`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900211011910` (pages 1, 84–86, 90).

[9] Henri Becquerel. “Sur les radiations émises par phosphorescence”. *Comptes Rendus de l’Académie des Sciences* 122 (1896), pp. 420–421. URL: `http://www.bibnum.education.fr/physique/radioactivite/sur-les-radiations-invisibles-emises-par-les-corps-phosphorescents` (page 2).

[10] L Meitner and O Hahn. “Über die Absorption der β-Strahlen”. *Physikalische Zeitschrift* 12 (1911), pp. 177–188 (page 2).

[11] J Danysz. "Über die Absorption der β-Strahlen". *Physikalische Zeitschrift* 14 (1913), pp. 417–424 (page 2).

[12] J Chadwick. "The intensity distribution in the magnetic spectrum of β particles from radium B and radium C". *Verhandlungen der Deutschen Physikalischen Gesellschaft* 16 (1914), pp. 383–391 (page 2).

[13] Glenn F Knoll. *Radiation Detection and Measurement.* 4th ed. Wiley, 2010 (page 2).

[14] Wolfgang Pauli. *Letter to the Physical Institute of the Federal Institute of Technology, Zürich.* 1930 (page 2).

[15] C L Cowan et al. "Detection of the Free Neutrino: A Confirmation". *Science* 124.3212 (1956), pp. 103–104. DOI: `10.1126/science.124.3212.103` (page 2).

[16] Frederick Reines and Clyde L. Cowan. "Free antineutrino absorption cross section. I. Measurement of the free antineutrino absorption cross section by protons". *Physical Review* 113.1 (1959). ISSN: 0031899X. DOI: `10.1103/PhysRev.113.273` (page 2).

[17] M Goldhaber, L Grodzins, and A W Sunyar. "Helicity of Neutrinos". *Physical Review* 109.3 (1958), pp. 1015–1017. DOI: `10.1103/PhysRev.109.1015` (page 2).

[18] C G Sudarshan and R E Marshak. "Chirality Invariance and the Universal Fermi Interaction". *Physical Review* 109.6 (1958), pp. 1860–1862. DOI: `10.1103/PhysRev.109.1860` (page 2).

[19] R P Feynman and M Gell-Mann. "Theory of the Fermi Interaction". *Physical Review* 109.1 (1958), pp. 193–198. DOI: `10.1103/PhysRev.109.193` (page 2).

[20] C S Wu et al. "Experimental Test of Parity Conservation in Beta Decay". *Physical Review* 105.4 (1957), pp. 1413–1415. DOI: `10.1103/PhysRev.105.1413` (page 2).

[21] C D Anderson and S H Neddermeyer. "Cloud Chamber Observations of Cosmic Rays at 4,300 Meters Elevation and Near Sea-Level". *Physical Review* 51 (1937), pp. 884–886. DOI: `10.1103/PhysRev.51.884` (page 3).

[22] I Hinchliffe and C H Llewellyn Smith. "Lepton number violation and muon decay". *Nuclear Physics B* 88 (1975), pp. 423–460. DOI: `10.1016/0550-3213(75)90207-4` (page 3).

[23] D Bryman et al. "Search for the Decay μˆ+ eˆ+ γ". *Physical Review Letters* 50 (1983), p. 1546. DOI: `10.1103/PhysRevLett.50.1546` (page 3).

[24] G. Danby et al. "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos". *Physical Review Letters* 9.1 (1962). ISSN: 00319007. DOI: `10.1103/PhysRevLett.9.36` (page 3).

[25] K Kodama et al. (DONUT Collaboration). "Observation of tau neutrino interactions". *Physics Letters B* 504.3 (2001), pp. 218–224. DOI: `10.1016/S0370-2693(01)00307-0` (page 3).

[26] SLD ALEPH DELPHI and L3 OPAL Collaborations. “Precision electroweak measurements on the Z resonance”. *Physics Reports* 427.5–6 (2006), pp. 257–454. DOI: `10.1016/j.physrep.2005.12.006` (page 3).

[27] R Davis. “Attempt to Detect the Antineutrinos from a Nuclear Reactor by the 37Cl + $\nu \rightarrow$ e + 37Ar Reaction”. *Physical Review* 97.3 (1955), pp. 766–769. DOI: `10.1103/PhysRev.97.766` (page 3).

[28] R Davis, D Harmer, and K Hoffman. “Search for Neutrinos from the Sun”. *Phys. Rev. Lett* 20.1205 (1968). DOI: `https://doi.org/10.1103/PhysRevLett.20.1205` (page 3).

[29] Q R Ahmad et al. “Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory” (2002). DOI: `10.1103/PhysRevLett.89.011301` (pages 3, 84).

[30] B Pontecorvo. “Inverse beta processes and nonconservation of lepton charge”. *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247 (pages 3, 4).

[31] B Pontecorvo. “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”. *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725 (pages 3, 4).

[32] B. Aharmim et al. “Measurement of the cosmic ray and neutrino-induced muon flux at the Sudbury neutrino observatory”. *Physical Review D* 80.1 (July 2009). ISSN: 15502368. DOI: `10.1103/PhysRevD.80.012001` (page 3).

[33] Y Fukuda et al. *Evidence for Oscillation of Atmospheric Neutrinos.* Tech. rep. 1998, p. 21 (pages 3, 84).

[34] Z Maki, M Nakagawa, and S Sakata. “Remarks on the Unified Model of Elementary Particles”. *Progress of Theoretical Physics* 28.5 (May 1962), pp. 870–880. ISSN: 0033-068X. DOI: `10.1143/PTP.28.870`. URL: `https://doi.org/10.1143/PTP.28.870` (page 4).

[35] R L Workman et al. (Particle Data Group). “Neutrino Masses, Mixing, and Oscillations”. *Prog. Theor. Exp. Phys.* 2022.8 (2022), p. 083C01. URL: `https://pdg.lbl.gov/2022/reviews/rpp2022-rev-neutrino-mixing.pdf` (page 4).

[36] V Barger et al. “Matter effects on three-neutrino oscillations”. *Phys. Rev. D* 22 (1980), p. 2718. DOI: `https://doi.org/10.1103/PhysRevD.22.2718` (page 5).

[37] L Wolfenstein. “Neutrino oscillations in matter”. *Phys. Rev. D* 17 (1978), p. 2369. DOI: `https://doi.org/10.1103/PhysRevD.17.2369` (page 5).

[38] A De Gouvêa, J Jenkins, and B Kayser. “Neutrino Mass Hierarchy, Vacuum Oscillations, and Vanishing |Ue3 |”. *Phys. Rev. D* 71.113009 (2005). DOI: `https://doi.org/10.1103/PhysRevD.71.113009` (page 5).

[39] Charles Naseby. “Understanding the Impact of an Expanded Neutral Current Pion Production Model on Long-Baseline Oscillation Analyses at T2K”. PhD thesis. Imperial College London, 2022. URL: `https://www.imperial.ac.uk/a-z-research/high-energy-physics/group-information/theses/` (page 5).

[40] M. A. Acero et al. “An Improved Measurement of Neutrino Oscillation Parameters by the NOvA Experiment” (Aug. 2021). DOI: `10.1103/PhysRevD.106.032004`. URL: `http://arxiv.org/abs/2108.08219http://dx.doi.org/10.1103/PhysRevD.106.032004` (page 5).

[41] I Esteban et al. “NuFit-6.0: Updated global analysis of three-flavor neutrino oscillations” (Oct. 2024). DOI: `10.1007/JHEP12(2024)216`. URL: `http://arxiv.org/abs/2410.05380http://dx.doi.org/10.1007/JHEP12(2024)216` (pages 5, 6).

[42] H Nunokawa, S J Parke, and J W F Valle. “CP Violation and Neutrino Oscillations”. *Prog. Part. Nucl. Phys.* 60 (2008), p. 338. URL: `https://arxiv.org/abs/0710.0554` (page 6).

[43] R N Mohapatra and A Y Smirnov. “Neutrino Mass and New Physics”. *Ann. Rev. Nucl. Part. Sci.* 56 (2006), p. 569. URL: `https://arxiv.org/abs/hep-ph/0603118` (page 6).

[44] G Altarelli and F Feruglio. “Discrete Flavor Symmetries and Models of Neutrino Mixing”. *Rev. Mod. Phys.* 82 (2010), p. 2701. URL: `https://arxiv.org/abs/1002.0211` (page 6).

[45] S F King. “Models of Neutrino Mass, Mixing and CP Violation”. *J. Phys. G* 42 (2015), p. 123001. URL: `https://arxiv.org/abs/1510.02091` (page 6).

[46] S T Petcov. “Discrete Flavour Symmetries, Neutrino Mixing and Leptonic CP Violation”. *Eur. Phys. J. C* 78 (2018). URL: `https://doi.org/10.1140/epjc/s10052-018-XXXX-X` (page 6).

[47] Angel Abusleme et al. “TAO Conceptual Design Report”. *ArXiv Preprint* (2020). DOI: `https://doi.org/10.48550/arXiv.2005.08745` (page 6).

[48] JUNO Collaboration et al. “Potential to identify neutrino mass ordering with reactor antineutrinos at JUNO”. *Chinese Phys. C* 43.3 (2024). DOI: `10.1088/1674-1137/ad7f3e`. URL: `http://arxiv.org/abs/2405.18008http://dx.doi.org/10.1088/1674-1137/ad7f3e` (page 6).

[49] T Zhu. “Long-baseline neutrino oscillation sensitivities with Hyper-Kamiokande and impact of Intermediate Water Cherenkov Detector Hyper-K Long-baseline sensitivities and impact of IWCD”. *Proceedings of Science.* 2022. URL: `https://inspirehep.net/files/a26c195f9bfdbb880b33820d934ccc83` (page 6).

[50] DUNE Collaboration et al. “Low exposure long-baseline neutrino oscillation sensitivity of the DUNE experiment” (Sept. 2021). URL: `http://arxiv.org/abs/2109.01304` (page 6).

[51] Hyper-Kamiokande Collaboration. “Sensitivity of the Hyper-Kamiokande experiment to neutrino oscillation parameters using acceleration neutrinos” (May 2025). URL: `http://arxiv.org/abs/2505.15019` (pages 6, 8).

[52] DUNE Collaboration et al. “Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 2: The Physics Program for DUNE at LBNF” (2015). URL: `http://arxiv.org/abs/1512.06148` (page 6).

[53] DUNE Collaboration. “Prospects for Beyond the Standard Model Physics with DUNE Atmospheric Neutrinos”. *Eur. Phys. J. C* 81 (2021), p. 322. DOI: `10.1140/epjc/s10052-021-09007-w`. URL: `https://arxiv.org/abs/2008.12769` (page 7).

[54] L. Alvarez-Ruso et al. “NuSTEC White Paper: Status and Challenges of Neutrino-Nucleus Scattering” (June 2017). DOI: `10.1016/j.ppnp.2018.01.006`. URL: `http://arxiv.org/abs/1706.03621http://dx.doi.org/10.1016/j.ppnp.2018.01.006` (pages 8, 9, 15–17).

[55] T. Dieminger et al. “Uncertainties on the $\nu\mu/\nu$e, $\nu\mu/\nu$e and νe$/\nu$e cross-section ratio from the modelling of nuclear effects and their impact on neutrino oscillation experiments” (2023). URL: `http://arxiv.org/abs/2301.08065` (pages 9, 19, 163).

[56] A Bodek. “Muon internal bremsstrahlung: a conventional explanation for the excess electron-neutrino events in MiniBoone” (2007). DOI: `https://doi.org/10.48550/arXiv.0709.4004`. URL: `http://arxiv.org/abs/0709.4004` (page 9).

[57] O Lalakulich, U Mosel, and K Gallmeister. “Neutrino energy reconstruction in quasielastic-like scattering in the MiniBooNE and T2K experiments”. *Phys. Rev. C* 86.5 (2012), p. 54606. DOI: `10.1103/PhysRevC.86.054606`. URL: `https://link.aps.org/doi/10.1103/PhysRevC.86.054606` (page 9).

[58] Y Hayato and L Pickering. “The NEUT Neutrino Interaction Simulation”. *Phys. J. Spec. Top.* 230 (2021), pp. 4469–4481. DOI: `10.1140/epjs/s11734-021-00287-7`. URL: `http://arxiv.org/abs/2106.15809http://dx.doi.org/10.1140/epjs/s11734-021-00287-7` (pages 10, 11, 17, 18, 20).

[59] D Freedman. “Coherent effects of a weak neutral current”. *Physical Review D* 9.5 (1974), pp. 1389–1392. DOI: `https://doi.org/10.1103/PhysRevD.9.1389` (page 10).

[60] D Akimov et al. “Observation of coherent elastic neutrino-nucleus scattering”. *Sceince* 357.6356 (2017), pp. 1123–1126. DOI: `10.1126/science.aao0990`. URL: `https://arxiv.org/abs/astro-ph/9801320v1` (page 10).

[61] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino cross sections across energy scales”. *Reviews of Modern Physics* 84.3 (Sept. 2012), pp. 1307–1341. ISSN: 00346861. DOI: `10.1103/RevModPhys.84.1307` (page 11).

[62] L Alvarez-Ruso et al. “Recent highlights from GENIE v3”. *European Physical Journal: Special Topics* 230.24 (2021), pp. 4449–4467. ISSN: 19516401. DOI: `10.1140/epjs/s11734-021-00295-7` (pages 11, 19).

[63] O. Benhar et al. “Spectral function of finite nuclei and scattering of GeV electrons”. *Nuclear Physics A* 579.3-4 (Oct. 1994), pp. 493–517. ISSN: 0375-9474. DOI: `10.1016/0375-9474(94)90920-2` (pages 11, 12).

[64] Ulrich Mosel. *Neutrino Interactions with Nucleons and Nuclei: Importance for Long-Baseline Experiments.* Oct. 2016. DOI: `10.1146/annurev-nucl-102115-044720` (page 11).

[65] M Born. “Quantenmechanik der Stoßvorgänge”. *Zeitschrift für Physik* 38 (1926), pp. 803–827. DOI: `10.1007/BF01397184` (page 12).

[66] L D Landau and E M Lifshitz. *Quantum Mechanics: Non-relativistic Theory.* Pergamon Press, 1965 (page 12).

[67] J A Caballero et al. “Analysis of factorization in (e,e’p) reactions: A survey of the relativistic plane wave impulse approximation”. *Nuclear Physics A* 632 (1998), pp. 323–362. DOI: `https://doi.org/10.1016/S0375-9474(97)00817-8` (page 12).

[68] O Benhar and D Meloni. “Total neutrino and antineutrino nuclear cross sections around 1 GeV”. *Nuclear Physics A* 789.1-4 (2007), pp. 379–402. ISSN: 0375-9474. DOI: `10.1016/J.NUCLPHYSA.2007.02.015` (page 12).

[69] G. Shen et al. “Inclusive neutrino scattering off the deuteron from threshold to GeV energies”. *Physical Review C - Nuclear Physics* 86.3 (Sept. 2012). ISSN: 1089490X. DOI: `10.1103/PhysRevC.86.035503` (page 12).

[70] Particle Data Group. “Review of Particle Physics”. *Journal of Physics G: Nuclear and ParticlePhysics* 37.075021 (2010). DOI: `10.1088/0954-3899/37/7A/075021` (page 13).

[71] C. H. Llewellyn Smith. “Neutrino reactions at accelerator energies”. *Physics Reports* 3.5 (June 1972), pp. 261–379. ISSN: 0370-1573. DOI: `10.1016/0370-1573(72)90010-5` (page 13).

[72] R. Bradford et al. “A New Parameterization of the Nucleon Elastic Form Factors”. *Nuclear Physics B - Proceedings Supplements* 159.1 (Sept. 2006), pp. 127–132. ISSN: 09205632. DOI: `10.1016/j.nuclphysbps.2006.08.028` (page 13).

[73] D H Wilkinson. *ANALYSIS OF NEUTRON ß-DECAY.* Tech. rep. 1982, pp. 474–504 (page 13).

[74] A. Bodek et. al. “Extraction of the axial nucleon form factor from neutrino experiments on deuterium”. *Journal of Physics: Conference Series* 110.8 (2008). DOI: `10.1088/1742-6596/110/8/082004` (page 13).

[75] S.J. Barish et al. “Study of neutrino interactions in hydrogen and deuterium: Description of the experiment and study of the reaction nu +d →mu +p +p”. *Physical Review D* 16.11 (1977). DOI: `10.1103/PhysRevD.16.3103` (pages 13, 18, 161).

[76] L A Ahrens et al. “A study of the axial-vector form factors and second-class currents in antineutrino quasielastic scattering”. *Physics Letters B* 202.2 (1988). DOI: `https://doi.org/10.1016/0370-2693(88)90026-3` (page 13).

[77] S. L Adler. “Tests of the Conserved Vector Current and Partially Conserved Axial-Vector Current Hypotheses in High-Energy Neutrino Reactions*”. *NUM B ER* 135 (1964), p. 24. DOI: `10.1103/PhysRev.135.B963` (page 13).

[78] T W Donnelly. “Superscaling of inclusive electron scattering from nuclei”. *Physical Review C* 60.065502 (1999). DOI: `10.1103/PhysRevC.60.065502` (page 14).

[79] A Klustová. “Measuring the Muon Neutrino Magnetic Moment in the NOvA Near Detector”. PhD thesis. Fermi National Accelerator Laboratory, 2025. URL: `https://lss.fnal.gov/archive/thesis/2000/fermilab-thesis-2025-02.pdf` (page 14).

[80] A. A. Aguilar-Arevalo et al. “First measurement of the muon neutrino charged current quasielastic double differential cross section”. *Physical Review D - Particles, Fields, Gravitation and Cosmology* 81.9 (May 2010). ISSN: 15507998. DOI: `10.1103/PhysRevD.81.092005` (page 14).

[81] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. “The nucleon axial mass and the MiniBooNE quasielastic neutrino-nucleus scattering problem”. *Physics Letters B* 707.1 (2012), pp. 72–75. ISSN: 03702693. DOI: `10.1016/j.physletb.2011.11.061` (page 14).

[82] M. Martini et al. “Unified approach for nucleon knock-out and coherent and incoherent pion production in neutrino interactions with nuclei”. *Physical Review C - Nuclear Physics* 80.6 (Dec. 2009). ISSN: 1089490X. DOI: `10.1103/PhysRevC.80.065501` (page 14).

[83] R. Gran et al. “Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 GeV”. *Phys. Rev. D* 88.113007 (July 2013). DOI: `10.1103/PhysRevD.88.113007`. URL: `http://arxiv.org/abs/1307.8105http://dx.doi.org/10.1103/PhysRevD.88.113007` (page 14).

[84] G. D. Megias et al. “Inclusive electron scattering within the SuSAv2 meson-exchange current approach”. *Physical Review D* 94.1 (2016). ISSN: 24700029. DOI: `10.1103/PhysRevD.94.013012` (page 14).

[85] I. Ruiz Simo et al. “Relativistic model of 2p-2h meson exchange currents in (anti)neutrino scattering”. *Journal of Physics G* 44.6 (2017). ISSN: 13616471. DOI: `10.1088/1361-6471/aa6a06` (page 14).

[86] R. Subedi et al. “Probing cold dense nuclear matter”. *Science* 320.5882 (2008), pp. 1476–1478. ISSN: 00368075. DOI: `10.1126/science.1156675` (page 14).

[87] M Sehgal and D Rein. “Neutrino-Excitation of Baryon Resonances and Single Pion Production”. *Annals of Physics* 133 (1981), pp. 79–153. DOI: `https://doi.org/10.1016/0003-4916(81)90242-6` (page 15).

[88] K M Graczyk, J Zmuda, and J T. Sobczyk. “Electroweak form factors of the Δ (1232) resonance”. *Physical Review D - Particles, Fields, Gravitation and Cosmology* 90.9 (2014). ISSN: 15502368. DOI: `10.1103/PhysRevD.90.093001` (page 15).

[89] Krzysztof M. Graczyk and Jan T. Sobczyk. *Erratum: Form factors in the quark resonance model (Physical Review D - Particles, Fields, Gravitation and Cosmology (2008) 77 (053001)).* Apr. 2009. DOI: `10.1103/PhysRevD.79.079903` (page 15).

[90] C Berger and L. M. Sehgal. “Lepton mass effects in single pion production by neutrinos”. *Physical Review D* 76.11 (2007). ISSN: 15507998. DOI: `10.1103/PhysRevD.76.113004` (page 15).

[91] Krzysztof M. Graczyk and Jan T. Sobczyk. “Lepton mass effects in weak charged current single pion production”. *Physical Review D - Particles, Fields, Gravitation and Cosmology* 77.5 (Mar. 2008). ISSN: 15507998. DOI: `10.1103/PhysRevD.77.053003` (page 15).

[92] S. X. Nakamura, H. Kamano, and T. Sato. “Dynamical coupled-channels model for neutrino-induced meson productions in resonance region”. *Physical Review D - Particles, Fields, Gravitation and Cosmology* 92.7 (Oct. 2015). ISSN: 15502368. DOI: `10.1103/PhysRevD.92.074024` (page 16).

[93] T. Sato. “Neutrino–nucleon reactions in resonance region”. *European Physical Journal: Special Topics* 230.24 (2021), pp. 4409–4418. ISSN: 19516401. DOI: `10.1140/epjs/s11734-021-00284-w` (page 16).

[94] E. Hernández, J. Nieves, and M. J.Vicente Vacas. “Single π production in neutrino-nucleus scattering”. *Physical Review D* 87.11 (2013). ISSN: 15507998. DOI: `10.1103/PhysRevD.87.113009` (page 16).

[95] M. Kabirnezhad. “Single pion production in electron-nucleon interactions”. *Physical Review D* 102.5 (2020). ISSN: 24700029. DOI: `10.1103/PhysRevD.102.053009` (page 16).

[96] C Berger and L. M. Sehgal. “Partially conserved axial vector current and coherent pion production by low energy neutrinos”. *Physical Review D* 79.5 (2009). ISSN: 15507998. DOI: `10.1103/PhysRevD.79.053003` (page 16).

[97] L Aliaga et al. “Summary of the NuSTEC Workshop on Neutrino-Nucleus Pion Production in the Resonance Region”. *FERMILAB-CONF-20-602-ND-SCD* (2020). DOI: `https://doi.org/10.48550/arXiv.2011.07166` (page 16).

[98] A Bodek and U K Yang. “Higher twist, ξ scaling, and effective LO PDFs for lepton scattering in the few GeV region”. *J. Phys. G* 29 (2003), pp. 1899–1905. DOI: `10.1088/0954-3899/29/8/369` (page 16).

[99] A Bodek, I Park, and U K Yang. “Improved low Q2 model for neutrino and electron nucleon cross sections in few GeV region”. *Nuclear Physics B - Proceedings Supplements* 139.1-3 SPEC. ISS. (2005), pp. 113–118. ISSN: 09205632. DOI: `10.1016/j.nuclphysbps.2004.11.208` (page 16).

[100] T Sjöstrand. “High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4”. *Computer Physics Communications H* 82 (1994), pp. 74–89 (page 16).

[101] A Bodek and T Cai. "Removal energies and final state interaction in lepton nucleus scattering". *European Physical Journal C* 79.4 (2019). ISSN: 14346052. DOI: `10.1140/epjc/s10052-019-6750-3` (page 17).

[102] A. M. Ankowski, O Benhar, and M Sakuda. "Improving the accuracy of neutrino energy reconstruction in charged-current quasielastic scattering off nuclear targets". *Physical Review D* 91.3 (2015). ISSN: 15502368. DOI: `10.1103/PhysRevD.91.033005` (page 17).

[103] S Abe. "Implementation and investigation of electron-nucleus scattering in NEUT neutrino event generator". *ArXiv Preprint* (2024). URL: `https://arxiv.org/abs/2412.07466` (pages 17–21, 153).

[104] Tomasz Golan, Cezary Juszczak, and Jan T. Sobczyk. "Effects of final-state interactions in neutrino-nucleus interactions". *Physical Review C - Nuclear Physics* 86.1 (July 2012). ISSN: 1089490X. DOI: `10.1103/PhysRevC.86.015505` (page 17).

[105] E. Oset and L. L. Salcedo. "Delta self-energy in nuclear matter". *Nuclear Physics A* 468.3-4 (July 1987), pp. 631–652. ISSN: 0375-9474. DOI: `10.1016/0375-9474(87)90185-0` (page 18).

[106] L. L. Salcedo et al. "Computer simulation of inclusive pion nuclear reactions". *Nuclear Physics A* 484.3-4 (July 1988), pp. 557–592. ISSN: 0375-9474. DOI: `10.1016/0375-9474(88)90310-7` (page 18).

[107] E. S.Pinzon Guerra et al. "Using world $\pi\pm$ -nucleus scattering data to constrain an intranuclear cascade model". *Physical Review D* 99.5 (2019). ISSN: 24700029. DOI: `10.1103/PhysRevD.99.052007` (page 18).

[108] Oleksandr Tomalak et al. "QED radiative corrections for accelerator neutrinos". *Nature Communications* 13.1 (Sept. 2022), p. 5286. ISSN: 2041-1723. DOI: `10.1038/s41467-022-32974-x`. URL: `https://doi.org/10.1038/s41467-022-32974-x` (pages 18, 24).

[109] A De Rújula, R Petronzio, and A Savoy-Navarro. "Radiative corrections to high-energy neutrino scattering". *Nuclear Physics B* 154 (1979), pp. 394–426. DOI: `10.1016/0550-3213(79)90039-7` (page 18).

[110] M Day and K McFarland. "Differences in quasielastic cross sections of muon and electron neutrinos". *Physical Review D* 86.5 (2012). ISSN: 15507998. DOI: `10.1103/PhysRevD.86.053003` (page 18).

[111] K. Abe et al. "Search for Electron Antineutrino Appearance in a Long-Baseline Muon Antineutrino Beam". *Physical Review Letters* 124.16 (2020). ISSN: 10797114. DOI: `10.1103/PhysRevLett.124.161802` (page 18).

[112] O. Lalakulich and U. Mosel. "GiBUU and Shallow Inelastic Scattering" (2013). URL: `http://arxiv.org/abs/1303.6677` (page 19).

[113] Jakub Zmuda et al. "NuWro monte carlo generator of neutrino interactions-first electron scattering results". *Acta Physica Polonica B*. Vol. 46. 11. Jagellonian University, Nov. 2015, pp. 2329–2334. DOI: `10.5506/APhysPolB.46.2329` (page 19).

[114] M Khachatryan et al. "Electron-beam energy reconstruction for neutrino oscillation measurements". eng. *Nature* 599.7886 (2021), pp. 565–570. ISSN: 0028-0836. DOI: `https://doi.org/10.1038/s41586-021-04046-5` (pages 20, 24).

[115] D Attwood et al. "The scattering of muons in low-Z materials". *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 251.1 (2006), pp. 41–55. ISSN: 0168-583X. DOI: `https://doi.org/10.1016/j.nimb.2006.05.006`. URL: `https://www.sciencedirect.com/science/article/pii/S0168583X06006756` (pages 20, 24, 96, 146–148, 159, 163).

[116] C Accettura et al. "Towards a Muon Collider" (2023). URL: `http://arxiv.org/abs/2303.08533` (page 20).

[117] M. Bogomilov et al. "Demonstration of cooling by the Muon Ionization Cooling Experiment". *Nature* 578.7793 (2020), pp. 53–59. ISSN: 14764687. DOI: `10.1038/s41586-020-1958-9` (pages 20, 96).

[118] J. Allison et al. "Recent developments in GEANT4". *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 2016), pp. 186–225. ISSN: 01689002. DOI: `10.1016/j.nima.2016.06.125` (pages 20, 21, 153).

[119] A Heikkinen, N Stepanov, and J. P. Wellisch. "Bertini intra-nuclear cascade implementation in Geant4". *CHEP* (2003). DOI: `https://doi.org/10.48550/arXiv.nucl-th/0306008`. arXiv: `0306008 [nucl-th]` (pages 21, 153, 163).

[120] Afroditi Papadopoulou. "Lepton-Nucleus Scattering Measurements for Neutrino Interactions and Oscillations". PhD thesis. 2016. DOI: `https://doi.org/10.48550/arXiv.2301.0311` (page 21).

[121] J. Apostolakis et al. "The performance of the geant4 standard em package for LHC and other applications". *Journal of Physics: Conference Series*. Vol. 119. 3. Institute of Physics Publishing, July 2008. DOI: `10.1088/1742-6596/119/3/032004` (page 22).

[122] V. N. Ivanchenko et al. "Geant4 models for simulation of multiple scattering". *Journal of Physics: Conference Series*. Vol. 219. 1 PART 3. Institute of Physics Publishing, 2010. DOI: `10.1088/1742-6596/219/3/032045` (pages 22, 96).

[123] W Y Ma et al. "Current status of final-state interaction models and their impact on neutrino-nucleus interactions". *Journal of Physics: Conference Series* 888.1 (Sept. 2017), p. 12171. DOI: `10.1088/1742-6596/888/1/012171`. URL: `https://dx.doi.org/10.1088/1742-6596/888/1/012171` (page 24).

[124] K Abe et al. "Calibration of the Super-Kamiokande detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 737 (Feb. 2014), pp. 253–272. ISSN: 0168-9002. DOI: `10.1016/j.nima.2013.11.081`. URL: `http://dx.doi.org/10.1016/j.nima.2013.11.081` (pages 24, 33).

[125] Andrew Missert. “Improving the T2K Oscillation Analysis With fiTQun: A New Maximum-Likelihood Event Reconstruction for Super-Kamiokande”. *Journal of Physics: Conference Series* 888 (Apr. 2017), p. 12066. DOI: `10.1088/1742-6596/888/1/012066` (page 24).

[126] Nicholas Prouse, Patrick de Perio, and Wojciech Fedorko. “Machine Learning Techniques to Enhance Event Reconstruction in Water Cherenkov Detectors”. *Physical Sciences Forum* 8.1 (2023). ISSN: 2673-9984. DOI: `10.3390/psf2023008063`. URL: `https://www.mdpi.com/2673-9984/8/1/63` (page 24).

[127] M Anghinolfi et al. “Quasi-elastic and inelastic inclusive electron scattering from an oxygen jet target”. *Nuclear Physics A* 602.3 (1996), pp. 405–422. ISSN: 0375-9474. DOI: `https://doi.org/10.1016/0375-9474(96)00093-0`. URL: `https://www.sciencedirect.com/science/article/pii/0375947496000930` (page 24).

[128] Li Qing-Run, Chen Sheng-Zhong, and Zhao En-Guang. “Low-energy pion scattering from nuclei and the α-particle model”. *Nuclear Physics A* 384.3 (1982), pp. 466–474. ISSN: 0375-9474. DOI: `https://doi.org/10.1016/0375-9474(82)90346-3`. URL: `https://www.sciencedirect.com/science/article/pii/0375947482903463` (page 24).

[129] E Richard et al. “Measurements of the atmospheric neutrino flux by Super-Kamiokande: Energy spectra, geomagnetic effects, and solar modulation”. *Phys. Rev. D* 94.5 (Sept. 2016), p. 52001. DOI: `10.1103/PhysRevD.94.052001`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.94.052001` (page 24).

[130] T Mueller and S Mine. *SK pi0 systematic errors for the nu_e analysis with T2K 6.4e20 POT (RUN1-4) data.* T2K Technical Note. 2013 (page 24).

[131] K Okumura. *π^0 background study using hybrid π^0 sample at Super-K.* T2K Technical Note. 2011 (page 24).

[132] E Montbarbon et al. “The CERN East Area Renovation”. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 461 (2019), pp. 98–101. ISSN: 0168-583X. DOI: `https://doi.org/10.1016/j.nimb.2019.08.028`. URL: `https://www.sciencedirect.com/science/article/pii/S0168583X1930583X` (page 25).

[133] Luigi Lavitola. “multi-PMT electronics system for Hyper-Kamiokande”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1054 (2023), p. 168461. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2023.168461`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900223004515` (pages 27, 28).

[134] S Aiello et al. “The KM3NeT multi-PMT optical module”. *Journal of Instrumentation* 17.07 (July 2022), P07038. DOI: `10.1088/1748-0221/17/07/P07038`. URL: `https://dx.doi.org/10.1088/1748-0221/17/07/P07038` (page 27).

[135] S Adrián-Martínez et al. “Deep sea tests of a prototype of the KM3NeT digital optical module”. *The European Physical Journal C* 74.9 (2014), p. 3056. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3056-3`. URL: `https://doi.org/10.1140/epjc/s10052-014-3056-3` (page 27).

[136] Gola Mohit et. al. “Assembly, testing, and installation of mPMT photosensor for the Water Cherenkov Test Experiment” (2025). DOI: `https://doi.org/10.48550/arXiv.2504.07216` (pages 28, 29).

[137] K Abe et al. “Second gadolinium loading to Super-Kamiokande”. *Nuclear Instruments and Methods in Physics Research Section A* 1065 (2024), p. 169480. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2024.169480`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900224004066` (pages 30, 34, 94).

[138] B A Moffat et al. “Optical calibration hardware for the Sudbury Neutrino Observatory”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 554.1–3 (Dec. 2005), pp. 255–265. ISSN: 0168-9002. DOI: `10.1016/j.nima.2005.08.029`. URL: `http://dx.doi.org/10.1016/j.nima.2005.08.029` (page 32).

[139] M R Anderson et al. “Optical calibration of the SNO+ detector in the water phase with deployed sources”. *Journal of Instrumentation* 16.10 (Oct. 2021), P10021. ISSN: 1748-0221. DOI: `10.1088/1748-0221/16/10/p10021`. URL: `http://dx.doi.org/10.1088/1748-0221/16/10/P10021` (page 32).

[140] T Akaishi and others (the EMPHATIC Collaboration). “EMPHATIC: A proposed experiment to measure hadron scattering and production cross sections for improved neutrino flux predictions”. *arXiv* (2019) (page 35).

[141] P A P Nghiem et al. “Core-halo issues for a very high intensity beam”. *Applied Physics Letters* 104.7 (Apr. 2014). ISSN: ISSN 0003-6951. DOI: `10.1063/1.4866400`. URL: `https://www.osti.gov/biblio/22283109` (page 40).

[142] D R Bellamy. “Photomultiplier Tube Spectrum Fitting in Particle Physics”. *Nuclear Instruments and Methods* A223.3 (1984), pp. 461–465. DOI: `10.1016/0168-9002(84)90757-3` (page 44).

[143] A. A. L. Craplet and J. Kvita. *Sub-GeV particle identification and beam flux measurement at CERN East Hall T9 beamline.* Tech. rep. WCTE Technical Note, 2025 (pages 45, 46, 48, 57, 59, 61, 76).

[144] J A Appel et al. “Performance of a lead-glass electromagnetic shower detector at fermilab”. *Nuclear Instruments and Methods* 127.4 (1975), pp. 495–505. ISSN: 0029-554X. DOI: `https://doi.org/10.1016/0029-554X(75)90653-9`. URL: `https://www.sciencedirect.com/science/article/pii/0029554X75906539` (page 50).

[145] E Armengaud et al. “Development of ^100Mo-containing scintillating bolometers for a high-sensitivity neutrinoless double-beta decay search”. *The European Physical Journal C* 77.11 (2017), p. 785. DOI: `10.1140/epjc/s10052-017-5343-2`. URL: `https://doi.org/10.1140/epjc/s10052-017-5343-2` (page 59).

[146] R L Workman and Others. “Passage of Particles Through Matter”. *Prog. Theor. Exp. Phys.* 2022.8 (2022), p. 083C01. DOI: `10.1093/ptep/ptac097`. URL: `https://pdg.lbl.gov/2022/reviews/rpp2022-rev-passage-particles-matter.pdf` (pages 61, 100, 121).

[147] Geant4 Collaboration. *Geant4 User’s Guide for Application Developers: Defining Materials.* 2023. URL: `https://geant4.web.cern.ch/documentation/manuals` (page 61).

[148] M. H. Ahn et al. “Indications of Neutrino Oscillation in a 250 km Long-Baseline Experiment”. *Physical Review Letters* 90.4 (2003). ISSN: 10797114. DOI: `10.1103/PhysRevLett.90.041801` (page 84).

[149] P. Adamson et al. “Measurement of neutrino oscillations with the MINOS detectors in the NuMI meam”. *Physical Review Letters* 101.13 (Sept. 2008). ISSN: 10797114. DOI: `10.1103/PhysRevLett.101.131802` (page 84).

[150] K Abe and others. “Observation of Electron Neutrino Appearance in a Muon Neutrino Beam”. *Phys. Rev. Lett.* 112.6 (Feb. 2014), p. 61802. DOI: `10.1103/PhysRevLett.112.061802`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.112.061802` (pages 84, 90).

[151] Y. Abe et al. “Indication of reactor ν e disappearance in the double chooz experiment”. *Physical Review Letters* 108.13 (Mar. 2012). ISSN: 00319007. DOI: `10.1103/PhysRevLett.108.131801` (page 84).

[152] J. K. Ahn et al. “Observation of reactor electron antineutrinos disappearance in the RENO experiment”. *Physical Review Letters* 108.19 (May 2012). ISSN: 00319007. DOI: `10.1103/PhysRevLett.108.191802` (page 84).

[153] K Abe and others. “Precise Measurement of the Neutrino Mixing Parameter θ_23 from Muon Neutrino Disappearance in an Off-Axis Beam”. *Phys. Rev. Lett.* 112 (2014), p. 181801. DOI: `10.1103/PhysRevLett.112.181801` (page 84).

[154] K Abe et al. *T2K ND280 Upgrade – Technical Design Report.* 2020. URL: `https://arxiv.org/abs/1901.03750` (pages 84, 85).

[155] Yasuhiro Takeda et al. “Exchange and observation systems for charge stripper foils at the J-PARC 3GeV-RCS”. *Nuclear Instruments and Methods in Physics Research Section A* 590.1-3 (June 2008), pp. 213–220. ISSN: 0168-9002. DOI: `10.1016/J.NIMA.2008.02.091` (page 86).

[156] K. Abe et al. “T2K neutrino flux prediction”. *Physical Review D* 87.1 (2013). ISSN: 15507998. DOI: `10.1103/PhysRevD.87.012001` (pages 86, 87, 89).

[157] S. Bhadra et al. “Optical transition radiation monitor for the T2K experiment”. *Nuclear Instruments and Methods in Physics Research Section A* 703 (2013), pp. 45–58. ISSN: 0168-9002. DOI: `10.1016/J.NIMA.2012.11.044` (page 86).

[158] K Suzuki et al. “Measurement of the muon beam direction and muon flux for the T2K neutrino experiment”. *Progress of Theoretical and Experimental Physics* 2015.5 (2015), p. 053C01. DOI: `10.1093/ptep/ptv054`. URL: `https://doi.org/10.1093/ptep/ptv054` (page 87).

[159] N Abgrall et al. *Status report to the proposal P330 Report from the NA61/SHINE experiment at the CERN SPS*. Tech. rep. 2010. URL: `http://na61.web.cern.ch` (page 87).

[160] N. Abgrall et al. “Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c”. *Physical Review C* 84.3 (2011). ISSN: 1089490X. DOI: `10.1103/PhysRevC.84.034604` (page 87).

[161] Stanford Linear et al. *FLUKA: A Multi-Particle Transport Code*. Tech. rep. 2005. URL: `https://www.slac.stanford.edu/pubs/slacreports/reports16/slac-r-773.pdf` (page 87).

[162] E889 COllaboration. *Long Baseline Neutrino Oscillation Experiment: Design Report*. Tech. rep. Brookhaven National Laboratory, Apr. 1995. URL: `https://digital.library.unt.edu/ark:/67531/metadc692670/` (page 88).

[163] Sacha E Kopp. “Accelerator Neutrino Beams”. *Physics Reports* 439.3 (Sept. 2007), pp. 101–159. DOI: `https://doi.org/10.1016/j.physrep.2006.11.004`. URL: `https://arxiv.org/pdf/physics/0609129` (page 88).

[164] LBNE Collaboration et al. *The Long-Baseline Neutrino Experiment: Exploring Fundamental Symmetries of the Universe*. 2014. URL: `https://arxiv.org/abs/1307.7335` (page 90).

[165] K Abe et al. *Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —*. 2011. URL: `https://arxiv.org/abs/1109.3262` (page 90).

[166] A Astbury et al. *CERN-SPSC-78-06*. Tech. rep. CERN, 1978. URL: `http://cds.cern.ch/record/319371` (page 90).

[167] J Altegoer and The NOMAD Collaboration. “The NOMAD experiment at the CERN SPS”. *Nuclear Instruments and Methods in Physics Research Section A* 404 (1998), pp. 96–128. DOI: `10.1016/S0168-9002(97)01079-6` (page 90).

[168] S. Aoki et al. “The T2K Side Muon Range Detector (SMRD)”. *Nuclear Instruments and Methods in Physics Research, Section A* 698 (2013), pp. 135–146. ISSN: 01689002. DOI: `10.1016/j.nima.2012.10.001` (page 91).

[169] D. Allan et al. “The electromagnetic calorimeter for the T2K near detector ND280”. *Journal of Instrumentation* 8.10 (Oct. 2013). ISSN: 17480221. DOI: `10.1088/1748-0221/8/10/P10019` (page 91).

[170] S. Assylbekov et al. “The T2K ND280 off-axis pi-zero detector”. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 686 (2012). ISSN: 01689002. DOI: `10.1016/j.nima.2012.05.028` (page 91).

[171] Still B. *π 0 and the T2K Experiment.* Tech. rep. 2010. DOI: `https://doi.org/10.48550/arXiv.1006.1034` (page 91).

[172] K. Abe et al. "Measurement of the electron neutrino charged-current interaction rate on water with the T2K ND280 π0 detector". *Physical Review D* 91.11 (2015). ISSN: 15502368. DOI: `10.1103/PhysRevD.91.112010` (page 91).

[173] P. A. Amaudruz et al. "The T2K fine-grained detectors". *Nuclear Instruments and Methods in Physics Research, Section A* 696 (2012). ISSN: 01689002. DOI: `10.1016/j.nima.2012.08.020` (page 92).

[174] N Abgrall et al. "Time projection chambers for the T2K near detectors". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 637.1 (2011), pp. 25–46. DOI: `10.1016/j.nima.2011.02.036` (pages 92–94, 99, 120).

[175] B. Jamieson. "Experience with the Time Projection Chambers for the T2K Near Detectors". *Physics Procedia.* Vol. 37. 2012. DOI: `10.1016/j.phpro.2012.02.402` (page 93).

[176] I. Giomataris et al. "Micromegas in a bulk". *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 560.2 (2006). ISSN: 01689002. DOI: `10.1016/j.nima.2005.12.222` (page 93).

[177] S. Fukuda et al. "The Super-Kamiokande detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.2-3 (Apr. 2003), pp. 418–462. ISSN: 0168-9002. DOI: `10.1016/S0168-9002(03)00425-X` (page 95).

[178] K. Abe et al. "First gadolinium loading to Super-Kamiokande". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1027 (Mar. 2022), p. 166248. ISSN: 0168-9002. DOI: `10.1016/J.NIMA.2021.166248` (page 94).

[179] J. J. Aubert et al. "The ratio of the nucleon structure functions F2N for iron and deuterium". *Physics Letters B* 123.3-4 (Mar. 1983), pp. 275–278. ISSN: 0370-2693. DOI: `10.1016/0370-2693(83)90437-9` (page 96).

[180] Nagamine Kanetada. "Advanced muon radiography with compact accelerator system". *Proc. Jpn. Acad. Ser. B* 80.4 (2004), pp. 179–182 (page 96).

[181] S. Pesente et al. "First results on material identification and imaging with a large-volume muon tomography prototype". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 604.3 (June 2009), pp. 738–746. ISSN: 0168-9002. DOI: `10.1016/J.NIMA.2009.03.017` (page 96).

[182] L Moneta, O Zapata, and H Dembinsky. "New Developments in Minuit2". *CHEP 2023.* 2023. URL: `https://indico.jlab.org/event/459/contributions/11597/attachments/9469/13862/Minuit2_CHEP2023.pdf` (pages 100, 103).

[183] F James. *MINUIT – Function Minimization and Error Analysis*. Tech. rep. CERN-D506. CERN, 1994. URL: `https://root.cern.ch/download/minuit.pdf` (pages 100, 103).

[184] P Bartet et al. *$\nu\mu$ CC event selections in the ND280 tracker using Run 2+3+4 data*. 2019 (pages 102, 112, 135, 136).

[185] A Cervera-Villanueva, C Giganti, and F Sanchez. *PID systematics of the Time Projection Chambers for the ν_μ analysis*. 2009 (page 135).

[186] E Frank, A Marchionni, and M Messina. *B-field calibration and systematic errors*. 2009 (pages 135, 136).

[187] L Escudero et al. *Study of TPC momentum resolution from tracks that cross multiple TPCs and the associated systematic error on the ν_μ CC and CCQE analyses*. 2010 (page 135).

[188] A Cervera and L Escudero. *Study of momentum resolution and scale using tracks that cross multiple TPCs*. 2014 (page 136).

[189] S Bordoni et al. *The TPC Particle IDentification algorithm with production 6B*. Jan. 2015 (page 136).

[190] C Bojechko et al. *Measurement and Correction of Magnetic Field Distortions in the Time Projection Chambers*. 2012 (page 136).

[191] A Fiorentini. *ND280 TPC Cluster Efficiency*. 2015 (page 136).

[192] C Bojechko et al. *CCQE-like and CC-non-QE-like $\nu\mu$ event selections in the ND280 tracker using Run 1+2 data*. 2013 (page 136).

[193] C Bojechko et al. *CC-multiple-pion $\nu\mu$ event selections in the ND280 tracker using Run 1+2+3+4 data*. 2016 (page 136).

[194] I Navon et al. "True Absorption and Scattering of 125-MeV Pions on Nuclei". *Phys. Rev. Lett.* 42.22 (May 1979), pp. 1465–1469. DOI: `10.1103/PhysRevLett.42.1465`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.42.1465` (page 136).

[195] J Lagoda. "Sand muon studies in ND280". 2015 (page 143).

[196] J Allison et al. "Geant4 developments and applications". *IEEE Transactions on Nuclear Science* 53.1 (2006), pp. 270–278. DOI: `10.1109/TNS.2006.869826` (page 153).

Appendix A

Momentum error propagation

The mean momentum for particle population x is:

$$\mathrm{p}_x^{\text{initial}} = \mathrm{p}_x^{\text{at TS1}} + \delta_{\mathrm{p}_x} = \frac{v_x m_x}{\sqrt{1-\left(\frac{v_x}{c}\right)^2}} + \delta_{\mathrm{p}_x}. \tag{A.1}$$

where v_x is the mean particle velocity:

$$v_x = \frac{L}{\tau_x + \tau_e - \frac{L}{c}}. \tag{A.2}$$

The error on the particle velocity is:

$$\sigma_{v_x} = \sqrt{\left(\frac{\partial v_x}{\partial \tau_e}\right)^2 \sigma_e^2 + \left(\frac{\partial v_x}{\partial \tau_x}\right)^2 \sigma_x^2 + \left(\frac{\partial v_x}{\partial L}\right)^2 \sigma_L^2} \tag{A.3}$$

where:

$$\frac{\partial v_x}{\partial \tau_e} = \frac{\partial v_x}{\partial \tau_x} = \frac{-L}{(\tau_x + \tau_e - \frac{L}{c})^2} \text{ and } \frac{\partial v_x}{\partial L} = \frac{\tau_x + \tau_e}{(\tau_x + \tau_e - \frac{L}{c})^2}, \tag{A.4}$$

substituting Equation A.4 into Equation A.3 to give the total error on the particle's velocity:

$$\sigma_{v_x} = \sqrt{\frac{L^2\sigma_e^2}{\left(\tau_x + \tau_e - \frac{L}{c}\right)^4} + \frac{L^2\sigma_x^2}{\left(\tau_x + \tau_e - \frac{L}{c}\right)^4} + \frac{\left(\tau_x + \tau_e\right)^2\sigma_L^2}{\left(\tau_x + \tau_e - \frac{L}{c}\right)^4}}. \tag{A.5}$$

The total error on the initial particle's momentum is:

$$\sigma_{p_x^{\text{initial}}} = \sqrt{\left(\frac{\partial \mathrm{p}_x^{\text{initial}}}{\partial v_x}\right)^2 \sigma_{v_x}^2 + \left(\frac{\partial \mathrm{p}_x^{\text{initial}}}{\partial \delta_{\mathrm{p}_x}}\right)^2 \sigma_{\delta_{\mathrm{p}_x}}^2} \tag{A.6}$$

where:

$$\frac{\partial \mathrm{p}_x^{\text{initial}}}{\partial \delta_{\mathrm{p}_x}} = 1 \text{ and } \frac{\partial \mathrm{p}_x^{\text{initial}}}{\partial v_x} = \frac{m_x}{\left(1 - \left(\frac{v_x}{c}\right)^2\right)^{3/2}} \tag{A.7}$$

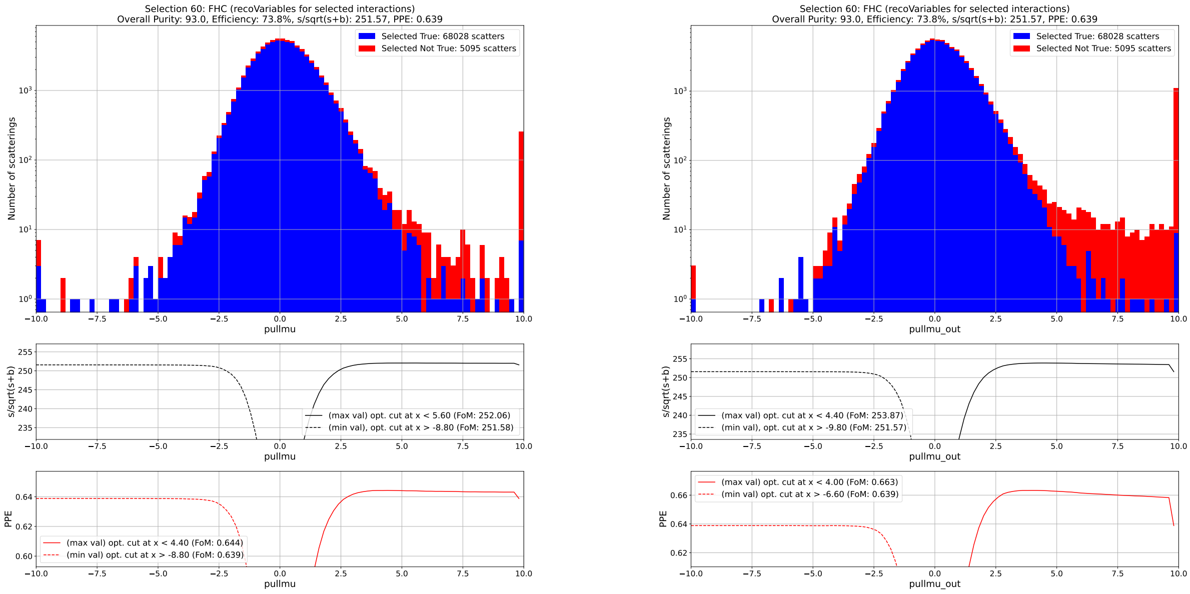
giving:

$$\sigma_{p_x^{\text{initial}}} = \sqrt{\frac{m_x^2 \sigma_{v_x}^2}{\left(1 - \left(\frac{v_x}{c}\right)^2\right)^3} + \sigma_{\delta_{\mathrm{p}_x}}^2} \tag{A.8}$$

where σ_{v_x} is given in Equation A.5 and the values of σ_L, $\sigma_{\delta_{\mathrm{p}_x}}$, σ_e and σ_x are indicated in Table 4.5.

Appendix B

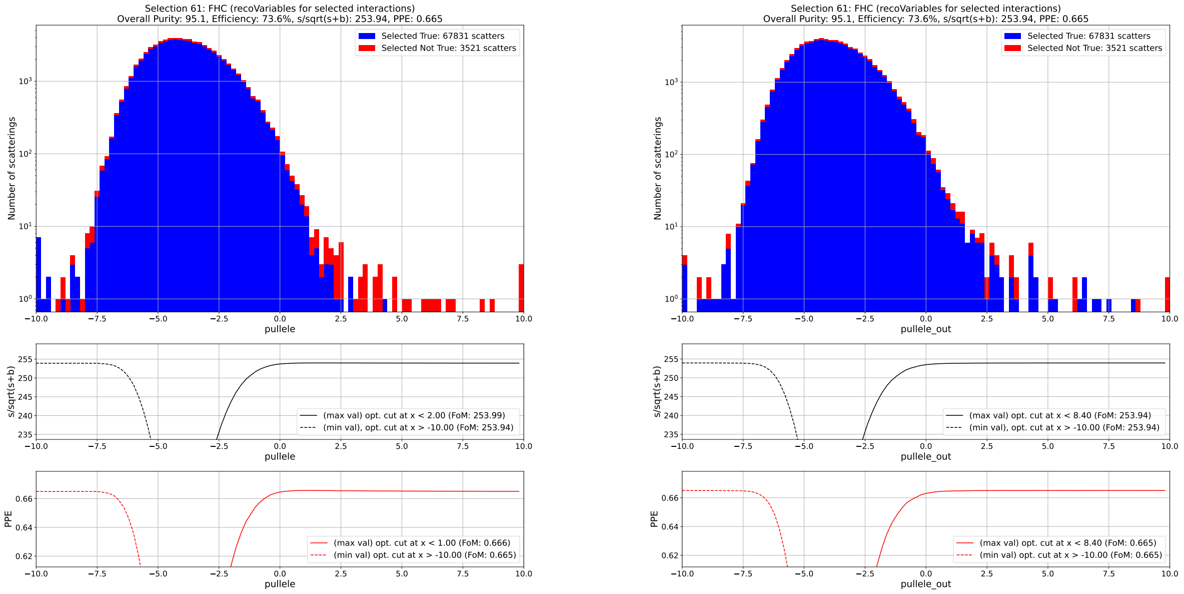
TPC pulls



(a) Incoming TPC1 track

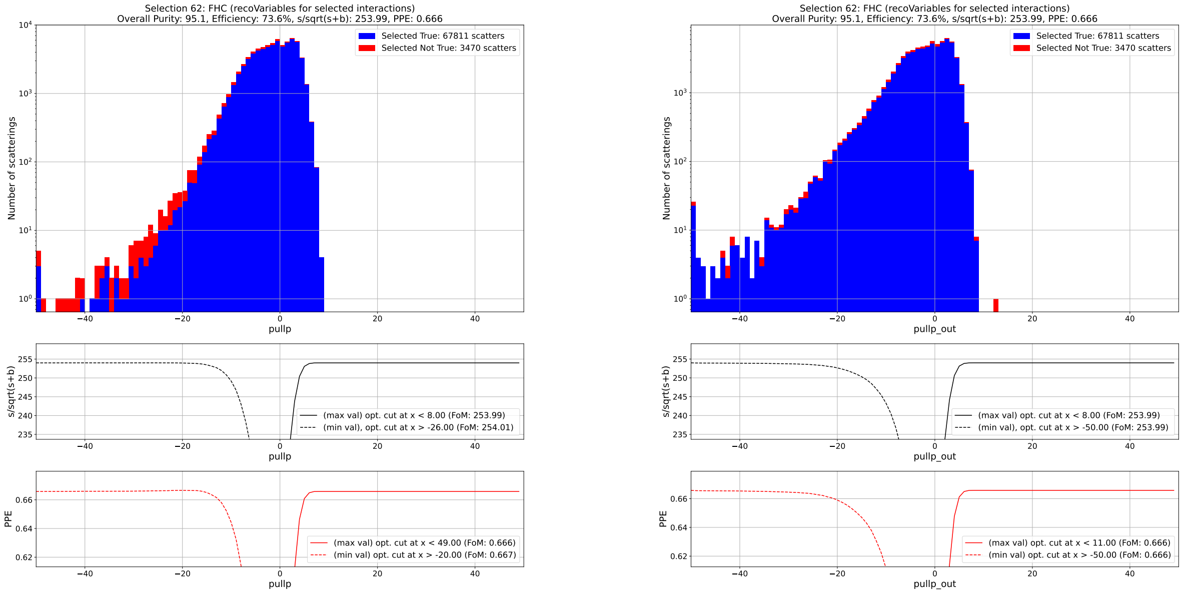
(b) Outgoing TPC2 track

Figure B.1: Distribution of the muon pull of the incoming and outgoing TPC tracks. The data quality and propagation cut are applied.



(a) Distribution of the electron pull for the incoming TPC track. The data quality, propagation and pullmu cuts are applied.

(b) Distribution of the electron pull for the outgoing TPC track. The data quality, propagation and pullmu cuts are applied.



(c) Distribution of the proton pull for the incoming TPC track. The data quality, propagation, pullmu and pullele cuts are applied.

(d) Distribution of the proton pull for the incoming TPC track. The data quality, propagation, pullmu and pullele cuts are applied.

Figure B.2: TPC electron and proton pull distributions.