Imperial College London,
Department of Physics

Measurement of lead-target neutrino interactions using the whole T2K near detector

Luke Pickering

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy, 7th June 2017

Abstract

This thesis presents a selection of neutrino interactions occurring within the Electromagnetic Calorimeters (ECals) of the T2K off-axis near detector, ND280. ND280 is situated 2.5° off axis with respect to the J-PARC neutrino beam and is used to constrain the unoscillated neutrino flux for T2K oscillation analyses. The motivation for an ECal-target sample is three-fold: to investigate the neutrino flux across a wider range of off-axis angles than is possible with the main near-detector target; to measure the rate of muon production with a larger angular acceptance than the standard near detector event samples; and to investigate neutrino interactions on lead, a heavy nuclear target on which there is little published data.

To enable this analysis, the ECal vertex reconstruction was integrated with the rest of the ND280 reconstruction software. For the first time, the accurate particle identification (PID) and momentum reconstruction capabilities of the ND280 Tracker are used for particles produced in neutrino interactions in the ECals. PID and kinematic reconstruction is not possible using ECal information alone.

Using this newly facilitated reconstruction of ECal-originating final-state particles, distributions of the reconstructed kinematics for a sample of charged-current inclusive muon-neutrino interaction candidates are compared between data and simulation. This highlights that a deficiency in the ND280 simulation of entering background particles constitutes a more significant problem than previously believed.

The data–simulation agreement seen is reasonable: the shape-only test statistic for the muon-like sample was evaluated to be $\chi^2 = 245$ for 104 analysis bins. This motivates the further use of the ND280 ECals for constraint of the T2K flux prediction as well as heavy-target, neutrino scattering measurements.

Acknowledgements

I want to start by highlighting the support and love of my Mum, Carolyn, and my Grandfather, Terry. The two people who most-championed my academic success are unfortunately no longer with us, but their spirit is evident in everything I do. I miss you both terribly. This is for you.

To my Grandmother, Josephine, thank you for all the nurturing support that you gave me growing up. The time spent playing Manic Miner after school especially stands out.

To my Dad, who is the most generous, funny, and caring man that I know. It was you who instilled in me the love of maths and programming from a young age, which has led me to this point. I remember you once saying to me

> *The thing about computer programmers is, they get paid to do what they'd do at home anyway.*

Over the past four years, that comment turned out to have been prescient.

I'd also like to mention my youngest brother Mark, my middle brother, James, his wonderful new family, Paige and Imogen, as well as Jo and Gemma, and Scott and Stephen. You have all been a source of laughter and joy, creating a relaxing environment for the scant down-time during the trying 'writing up' period.

The work presented here was carefully balanced on top of a stack of software used to simulate and analyse particle data from ND280. This has been built by a great many people over the last decade; my work would not have been possible without all of their contributions. I'd especially like to thank Dom and Sasha for their readily offered help on all things *recon*.

I have learned so much over the last four years from so many people, I'd like to name and shame a few, but many other have also been positive influences. Asher, you never failed to have useful advice on analysis techniques, statistics, or physics and your mentorship helped me reach this point. Studying transverse kinematics with Xianguo and Stephen granted me a number opportunities that I otherwise wouldn't have been able to take. I owe you both a great deal of thanks. I had a lot of fun and learned an indispensable amount about how to better approach data analysis while working with team NUISANCE: Patrick, Callum, and Clarence. I hope that we can try our best to do good science together in the future.

Phill arrived at Imperial at exactly the right time for me to bug him with the every terror-inducing plot that I made on the path to submitting. You always gladly offered advice of where to hunt for an answer, and as a result, I was able to finish this thesis satisfied with the body of work presented within. Thanks again Phill.

To Yoshi and Morgan, thank you both for offering me salient presentation and grammatical pointers, and making sure that I trundled in the right direction throughout my doctorate. I wouldn't be where I am (wherever that is) without your supervision these past four years. However, `vim` and `FORTRAN` are both rubbish... sorry.

I owe so much to Dom. This analysis would never have reached the stage that it has without the many hours we spent discussing how best to use the ECal data. Thank you so much.

I'd also like to briefly thank the administrative and sysadmin staff in the HEP group: Paula, Ray, and Simon. Much of my limited productivity was due to your time spent making day-to-day life on Level 5 run so smoothly.

Thanks to all the T2KUK students and postdocs who have helped make the last four years such an amazing experience. Especially Trevor, Kirsty, Stephen, Raj, and David from Oxford; Dom, Pip, Wing, and Clarence from Imperial; and Dave and Martti. From amazing holidays spent skiing down mountains and snorkelling off pacific beaches to countless delicious dinners, drinks, karaoke sessions, and mid-afternoon coffee breaks. Many, many thanks guys.

Finally, extra special thanks to Kirsty, who has helped make my doctorate the most formative and fulfilling time of my life. The love and support that you've shown me has been so important to my sanity. Thank you for everything that you do. You're my favorite.

Dedication

For Carolyn, Terry, Roland, and Jo

Copyright declaration

Statement of originality

This thesis and the analysis presented in it are my own. Figures and results reproduced from other sources have been cited throughout.

The first and second chapters briefly introduce the neutrino, its interactions with matter, and the T2K experiment; as a result, they contain reference to the work of a great many authors who have directly and indirectly made this analysis possible. The third and fifth chapters describe the use of reconstruction and uncertainty assessment tools that were developed by members of the T2K collaboration. A previous ECal-target selection was developed by Dominic Brailsford, for which the ECal vertex reconstruction algorithms were written. The integration of that algorithm with the rest of the ND280 reconstruction is original work, motivated by the research presented herein. The fourth chapter describes the event selection, which was built on top of the HighLAND2 analysis framework, developed by members of the T2K collaboration. However, the tailored ECal-target event selection and diagnostic tools were developed by myself. The results and conclusions presented in the final two chapters were my own. NuWro ReWeight, described in Appendix: B, was developed by myself, Patrick Stowell, and Jan Sobczyk.

Luke Pickering

'Yes, but quantum mechanics blows that out of the water. According to the Heisenberg Uncertainty Principle it is impossible to look at a sub-atomic particle and know both where it is and where it's going at the same time.'
Samantha Carter, SG-1

'Nothing's certain, 'cause of quantum.'
Lu-Tze, History Monk

Contents

Abstract **2**

Acknowledgements **3**

Copyright declaration and Statement of originality **6**

1 Introduction **13**

1.1 Neutrinos and flavour oscillation 13

1.2 Accelerator-based neutrino oscillation experiments 15

1.2.1 Neutrino beams 18

1.3 Neutrino interactions with matter 19

1.3.1 Interactions with free nucleons 21

1.3.2 Interactions with nuclei 24

1.3.3 Interaction simulations 28

2 The T2K near detector and ECal-target motivations **30**

2.1 The J-PARC neutrino beam 31

2.2 The magnetised off-axis near detector: ND280 35

2.3 The ND280 Tracker 38

2.4 The ND280 Barrel ECal . . . 42
2.5 The ND280 software . . . 44
2.6 Motivation for an ECal-target sample . . . 47
2.6.1 Previous ND280 ECal measurements . . . 49

3 ND280 Reconstruction **57**
3.1 ECal reconstruction . . . 59
3.1.1 Standard reconstruction . . . 59
3.1.2 Vertex track-finding . . . 61
3.1.3 Vertex candidate reconstruction . . . 65
3.1.4 Reconstruction performance . . . 66
3.2 ND280 global track reconstruction . . . 81
3.2.1 TPC tracking efficiency . . . 81
3.2.2 Inter-detector object matching . . . 84
3.3 Global vertex candidate reconstruction . . . 90
3.4 Future work . . . 96

4 ND280 ECal interactions selection **97**
4.1 The fiducial volume . . . 98
4.1.1 The previously-used fiducial volume definition . . . 98
4.1.2 A data-motivated fiducial volume . . . 99
4.2 The event selection . . . 108
4.2.1 Golden events . . . 109
4.2.2 Silver events . . . 133

4.2.3 Entering background selections . . . 136
4.2.4 Selection summary . . . 137
4.3 Future Work . . . 140

5 Sources of error **144**
5.1 Neutrino beam . . . 145
5.2 Neutrino interactions . . . 150
5.3 ND280 detector . . . 158
5.3.1 Tracker reconstruction . . . 158
5.3.2 ECal tracking and Tracker matching . . . 159
5.3.3 ECal mass . . . 164
5.3.4 Propagated uncertainties . . . 167

6 ECal-as-Target analysis results **170**
6.1 The data sample . . . 170
6.2 Selected event rates . . . 170
6.2.1 Data–simulation comparison . . . 176

7 Conclusion **187**
7.1 Future improvements . . . 188
7.2 Future uses for an ECal-target sample . . . 191

A ECal-isolated vertexing data–simulation comparisons **194**
A.1 ECal vertex track comparisons . . . 194
A.2 ECal vertex candidate comparisons . . . 200

B NuWro ReWeight **206**

B.1 Event Reweighting 206

B.2 NuWro Reweight 208

B.3 Free Nucleon Model Tuning 211

C A lightyear of lead **214**

List of Figures **225**

List of Tables **227**

Bibliography **240**

Chapter 1

Introduction

1.1 Neutrinos and flavour oscillation

The neutrino is an electrically neutral particle that interacts predominantly through the weak force. Proposed in 1930 by Wolfgang Pauli to explain the continuous beta particle energy spectra seen in nuclear decays, it was first observed by Cowan and Reines in 1956 [1]. Three 'flavours' of neutrinos have been experimentally observed, corresponding to the three generations of matter; the electron-, muon-, and tau-neutrino. Measurements of the decay width of the Z weak vector boson at the LEP electron–positron collider constrain these three to be the only light, weakly interacting neutrinos [2]. For many years, the neutrino was believed to be massless—indeed, the neutrino is taken as massless in the Standard Model of Particle Physics, but there is now an overwhelming body of evidence that neutrinos have a small but non-zero mass [3, 4, 5, 6, 7]. This evidence comes from observations of flavour change in neutrinos produced in nuclear reactions in the Sun, cosmic ray interactions in the atmosphere, and in man-made accelerator-based neutrino beams and nuclear reactors. Neutrinos produced in the atmosphere and from man-made sources have been shown to change flavour as a function of their production flavour, energy, and the distance travelled between production and observation—this process is known as *neutrino oscillation*. While neutrino flavour oscillation has been firmly established, some of the parameters governing the oscillation probabilities as a

function of neutrino energy and distance travelled are still only loosely constrained.

Neutrino oscillations occur because the three flavour eigenstates, ν_e, ν_μ, ν_τ, are not the same as the energy, or propagation, eigenstates, ν_1, ν_2, ν_3. The relationship between the two bases is encapsulated in the Pontecorvo-Maki-Nakagawa-Sakata matrix, $\mathbf{M}_{\text{PMNS}}$, and can be expressed as

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix}}_{\mathbf{M}_{\text{PMNS}}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \tag{1.1}$$

This 3×3 unitary matrix can be expressed in terms of three real angles, θ_{12}, θ_{13}, θ_{23}, and one phase, δ_{CP}. This can be parameterised as

$$\mathbf{M}_{\text{PMNS}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & C_{23} & S_{23} \\ 0 & -S_{23} & C_{23} \end{bmatrix} \begin{bmatrix} C_{13} & 0 & S_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -S_{13}e^{+i\delta_{\text{CP}}} & 0 & C_{13} \end{bmatrix} \begin{bmatrix} C_{12} & S_{12} & 0 \\ -S_{12} & C_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1.2}$$

where $C_{ij} = \cos\theta_{ij}$ and $S_{ij} = \sin\theta_{ij}$. The probability for a muon neutrino, of energy E_ν, travelling in a vacuum to be observed as an electron neutrino after a distance L is

$$\begin{aligned} P(\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_e) = \; & -4\sum_{i<j}^{3} \text{Re}\left[U_{\mu i}\cdot U_{ei}^{\star}\cdot U_{\mu j}^{\star}\cdot U_{ej}\right] \sin^2\left(\Delta m_{ij}^2 \tfrac{L}{4E_\nu}\right) \\ & \mp 2\sum_{i<j}^{3} \text{Im}\left[U_{\mu i}\cdot U_{ei}^{\star}\cdot U_{\mu j}^{\star}\cdot U_{ej}\right] \sin\left(\Delta m_{ij}^2 \tfrac{L}{2E_\nu}\right), \end{aligned} \tag{1.3}$$

where $\Delta m_{ij}^2 = m_i^2 - m_j^2$ is the mass squared difference of the neutrino mass states. The probability for electron anti-neutrino appearance, $P(\bar{\nu}_\mu \to \bar{\nu}_e)$, differs by a sign in the second term. For the explicit derivation of the oscillation probabilities, the reader is directed to Ref. [8]. From Eq. 1.3, it can be seen why the term 'oscillation' is used: the probability for a neutrino

produced in one flavour state to be later detected as another is sinusoidally dependent on L/E_ν.

The oscillation parameter currently of most interest is δ_{CP}, which can generate differences in how neutrinos and anti-neutrinos oscillate. For some values of δ_{CP}, the neutrino sector can generate sizeable CP violation—potentially resulting in matter–anti-matter asymmetry. Models exist that are capable of using this source of CP violation to explain the observed deficit of anti-matter in the visible universe. For an overview of these 'Leptogenesis' models, the reader is directed to § 4 of Ref. [9]. The anti-matter deficit is one of the greater unsolved problems in our best pictures of the physical laws that govern the universe. The search for CP violation in the neutrino sector is predominantly performed by long-baseline oscillation experiments exposed to man-made muon neutrino beams.

1.2 Accelerator-based neutrino oscillation experiments

Man-made neutrino beams are used to study flavour oscillations generated by PMNS mass-flavour mixing. Experiments generally use neutrino beams with peak energies of a few GeV and investigate oscillation over distances of hundreds of kilometres. The neutrino beams are usually first sampled near to the production site with so-called 'near detectors'. Near detectors are used to characterise a neutrino beam before significant oscillation has taken place; the properties of neutrino oscillations can then be inferred from measurements made after oscillations have taken place. The first accelerator-based, long-baseline neutrino oscillation experiment was the Main Injector Neutrino Oscillation Search (MINOS), which started taking data in 2005 and made world-leading measurements of two neutrino oscillation parameters: one mass squared splitting, $|\Delta m^2_{32}|$, and one mixing angle $\sin^2(2\theta_{23})$ [10]. These measurements have since been surpassed by the two second generation accelerator-based experiments, T2K [11] and NOνA [12], which are both taking data at the time of writing. The confidence intervals for the current measurements of these two parameters, which are measured by investigating muon neutrino disappearance, are reproduced in Figure 1.1. A summary of the relevant current and future long-baseline neutrino oscillation experiments is shown in Table. 1.1.

Experiment	Years active	Oscillation parameter sensitivities
Super-K [3]	1996–	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{23})$
MINOS [13]	2005–2012	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(2\theta_{23})$, $[\sin^2(\theta_{13}), \delta_{CP}]$
MINOS+ [14]	2013–2015	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(2\theta_{23})$, $[\sin^2(\theta_{13}), \delta_{CP}]$
IceCube [15, 16]	2011–	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{23})$, $[\text{sign}(\Delta m^2_{32})]$
T2K [11]	2010–	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{13})$, $\sin^2(\theta_{23})$, $[\delta_{CP}, \text{sign}(\Delta m^2_{32})]$
NOνA [12]	2013–	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{13})$, $\sin^2(\theta_{23})$, $[\delta_{CP}, \text{sign}(\Delta m^2_{32})]$
T2K-II [17]	Exp. 2020–	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{13})$, $\sin^2(\theta_{23})$, δ_{CP}, $[\text{sign}(\Delta m^2_{32})]$
Hyper-K [18]	Exp. 2026	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{13})$, $\sin^2(\theta_{23})$, δ_{CP}, $[\text{sign}(\Delta m^2_{32})]$
DUNE [19]	Exp. 2028	$\lvert\Delta m^2_{32}\rvert$, $\sin^2(\theta_{13})$, $\sin^2(\theta_{23})$, δ_{CP}, $\text{sign}(\Delta m^2_{32})$

Experiment	Energy/GeV	Baseline/km	Neutrino source
Super-K	0.2–10	15–13,000	Atmospheric neutrinos
MINOS	3 (peak)	735	On-axis NuMI beam
MINOS+	7 (peak)	735	On-axis NuMI beam
IceCube	10–100	10–12,700	Atmospheric neutrinos
T2K	0.6 (peak)	295	Off-axis J-PARC (200–750 kW)
NOνA	2 (peak)	810	Off-axis NuMI beam
T2K-II	0.6 (peak)	295	Off-axis J-PARC (750–1300 kW)
Hyper-K	0.6 (peak)	295	Off-axis J-PARC (1.3 MW)
DUNE	2 (peak)	1300	DUNE beam

Table 1.1: Summary of the long-baseline neutrino oscillation experiments mentioned in this chapter. The years that each experiment ran, or are expected to run; the neutrino oscillation parameters that each experiment aims to constrain, where bracketed parameters are those for which an experiment has weaker sensitivity; and approximate neutrino energies and oscillation baselines are presented.

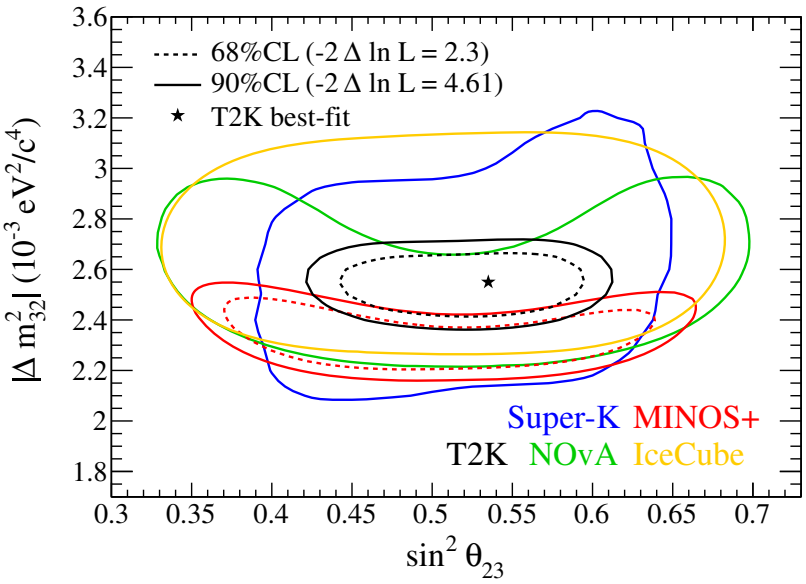


Figure 1.1: The best measurements of $|\Delta m^2_{32}|$ and $\sin^2(\theta_{23})$ and the associated two-dimensional 90% confidence contours from a number of long-baseline oscillation experiments investigating muon neutrino disappearance. Figure from Ref. [20]

This thesis describes an investigation performed using the T2K near detector. The T2K experiment investigates muon neutrino flavour change between J-PARC [21], on the east coast of Japan, and the Kamioka observatory 295 km to the west [11]. The unoscillated T2K neutrino beam is first sampled 280 m from the production site by the near detector, which is named ND280[1]. The neutrino flux is later observed by the 'far detector', Super-Kamiokande [22], after the neutrinos have travelled along an experimental baseline of 295 km. This baseline corresponds to the shortest distance at which maximal oscillation occurs for initially muon neutrinos with approximately 0.6 GeV of energy—the T2K beam peak energy. In 2014, the T2K experiment reached its design goal by making the first observation of electron neutrino appearance [7]. Since then, the updated T2K collaboration physics goals include contributing to the measurement of δ_{CP}, which will require precise measurement of the oscillation probabilities for both neutrinos and anti-neutrinos [23]. To achieve this, an accurate prediction of the neutrino flux is needed and the details of neutrino interactions with matter must be well known.

At the time of writing, long-baseline oscillation experiments are just starting to become sensitive to the value of δ_{CP}, but any evidence for δ_{CP} taking CP violation-generating values is still just

[1] Attempts were made to arrive at a more inventive near detector name, but failed.

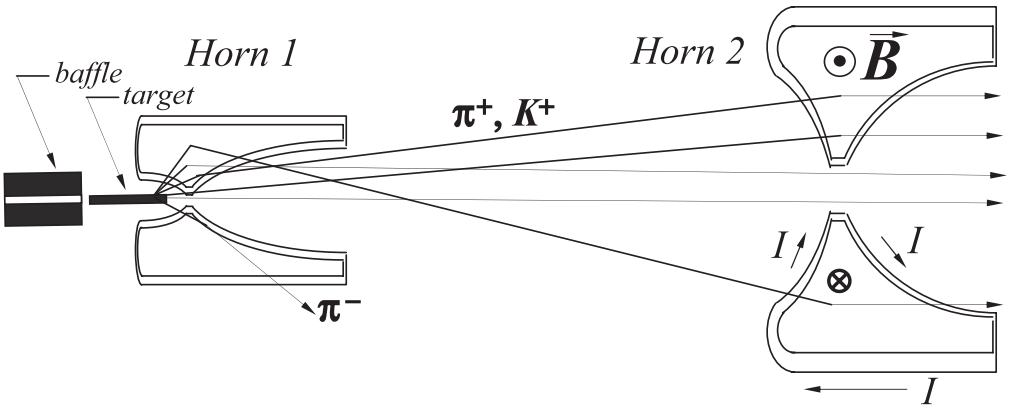


Figure 1.2: The Neutrinos at the Main Injector (NuMI) proton beam target and two-horn focussing system. The NuMI beam was used for MINOS [14] and currently produces the neutrino beam for the NOνA [12] and MINERνA [26] experiments. Electromagnetic horns are used to focus and charge select a beam of secondary hadrons, which are produced when protons from an accelerator beam impact on a target. If the current, I, is reversed, negatively (positively) charged particles can be focussed (defocussed). Figure from Ref. [13]

a tantalising hint [20]. The next generation of long baseline neutrino experiments, DUNE [19] and Hyper-K [18], project good sensitivity to δ_{CP} if systematic uncertainties can be suitably reduced.

1.2.1 Neutrino beams

Accelerator-based beams of muon neutrinos are created by focussing short-lived mesons, produced through hadronisation in a proton beam target, and allowing them to decay in flight to neutrinos and other particles. While there are designs for muon storage-ring 'neutrino factories' [24], there is no agreed-upon timescale for the construction of such sources. The focussing of the mesons is performed by one or more electromagnetic neutrino 'horns', which have a wide acceptance for incoming particles and are capable of sign-selecting and focussing the mesons into an approximately collinear meson beam [25]. A cartoon of the MINOS target and horn configuration is reproduced in Figure 1.2. The polarity of a focussing horn can be chosen to select (negative) positive mesons, which results in a (anti-matter-) matter-dominated neutrino beam. The beam comprises mostly muon neutrinos, which are produced through the decay of charged pions and kaons, $\pi^{\pm}/\text{K}^{\pm} \rightarrow \mu^{\pm} + \overset{(-)}{\nu}_{\mu}$ (equivalent decays to electrons are strongly helicity suppressed).

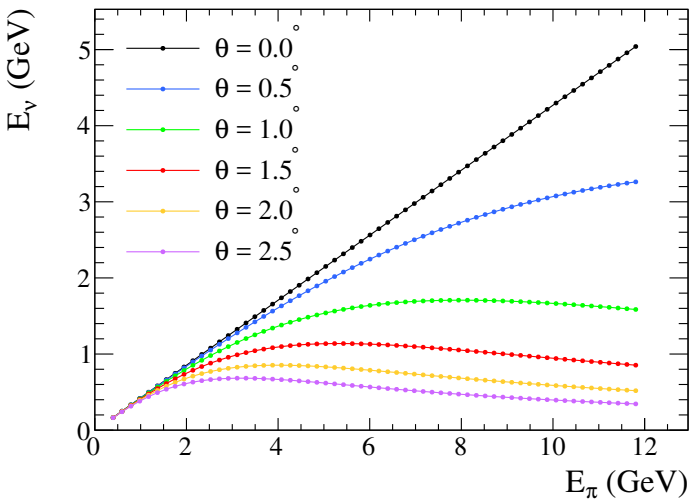


Figure 1.3: The neutrino energy spectrum predicted as a function of parent pion parent and the angle between the momentum of the neutrino and the pion—the off-axis angle. The flux seen at greater off-axis angles has a finer spectral width in neutrino energy. Figure from Ref. [27].

Off-axis beams

Both T2K and NOνA have far detectors that are situated away from the central axis of the neutrino beam. Because of the kinematics of the boosted meson decay in flight, this results in a more fine-band neutrino energy spectrum. The neutrino energy spectrum as a function of decay parent pion energy, for a range of off-axis angles, is shown in Figure 1.3. By placing the detectors off axis, the number of events from high-energy neutrinos—which have a different oscillation probability at the far detector—can be strongly suppressed.

1.3 Neutrino interactions with matter

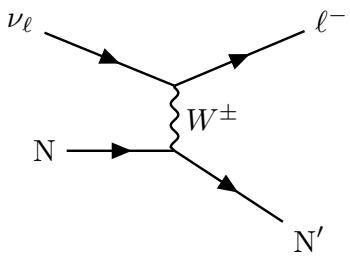
To study the properties of flavour oscillation, a significant number of neutrinos must be observed after they have travelled the baseline distance. Neutrinos are detected by the measurement of charged final-state particles produced through weak interactions with matter. However, the quantity of interest when studying oscillations is the number of neutrinos of each flavour that passed through the detector, not the number that were observed. The true number of neutrinos must be inferred through neutrino interaction and detector response models. At the

time of writing, the properties of neutrino interactions with matter are not well-known. In this precision era of neutrino oscillation measurement, these uncertainties are fast-becoming the limiting factor on future sensitivity [28].

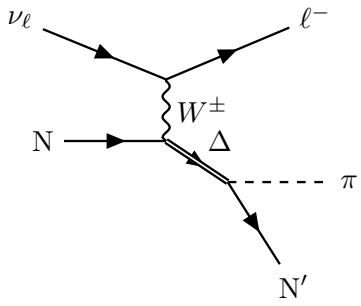
There are a number of competing factors which constrain the practical energy regime in which neutrino oscillations can be studied. Neutrinos are most-often observed through charged current interactions which produce a flavour-equivalent charged lepton; this production of mass has an associated energy threshold that places a lower limit on the useful neutrino energy[2]. The oscillation probability (Eq. 1.3) is periodic in the ratio L/E_ν of the experimental baseline, L, to the neutrino energy, E_ν. To attain the highest probability of flavour change at the far detector, this ratio must be tuned so that $\sin\left(\Delta m^2_{ij}\frac{L}{2E_\nu}\right) \simeq 0$—the most common choice is at the so-called 'first oscillation maxima', where $\Delta m^2_{ij}\frac{L}{2E_\nu} \simeq \pi$. The neutrino interaction rate should also be considered. On the one hand, the inclusive charged-current neutrino–matter cross section rises approximately proportionally to neutrino energy. But on the other hand, higher neutrino energies necessitate longer experimental baselines to maintain the L/E_ν ratio, and beam dispersion causes an attenuation in neutrino flux that is proportional to $1/L^2$. Therefore, to achieve the largest oscillation signal, the shortest baseline and lowest energy above threshold is the most desirable. A further complicating consideration arises when an experiment aims to make use of 'matter effects' [29, 30] to determine $\text{sign}(\Delta m^2_{32})$; here, higher neutrino energies render the difference between oscillations in matter and in a vacuum more apparent. However, the higher energy neutrino beam necessitates a longer experimental baseline to maintain the L/E_ν ratio.

For all intents and purposes, neutrinos interact with matter purely through the weak nuclear force. At very low four-momentum transfer the vector bosons that mediate the weak force scatter coherently with large-scale structures—e.g., entire nuclei. At high four-momentum transfer—above approximately 5–10 GeV^2—the vector bosons scatter off the quarks that make up nucleons. Interactions in the high four-momentum transfer regime are comparatively straightforward to model [31]. Current and planned neutrino oscillation experiments have neutrino beams in the $E_\nu \sim \mathcal{O}\,(1\text{–}5\,\text{GeV})$ energy region. This energy range constitutes a transition regime

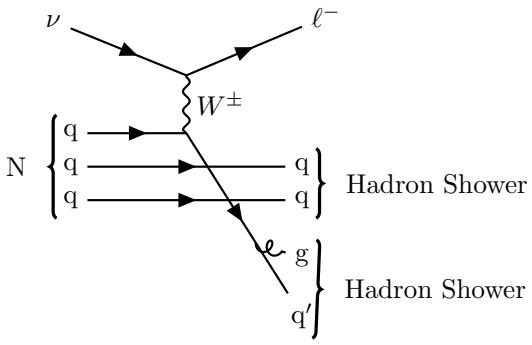
[2] For this reason, tau neutrino appearance cannot be effectively investigated at current oscillation experiments.



(a) Quasi-elastic scattering (CCQE)



(b) Single pion production (SPP)



(c) Deep inelastic scattering (DIS)

Figure 1.4: Example Feynman diagrams for three charged-current neutrino scattering channels.

where neutrinos scatter from both bound nucleons within a nucleus and nucleon-constituent quarks. Building accurate models in this non-perturbative regime is a difficult task [31]. Current interaction models are pieced together from constituent, non-interfering free-nucleon interaction channels, to which additions and corrections are made to account for the effects of scattering off bound nucleons.

1.3.1 Interactions with free nucleons

Interactions with free nucleons are broadly split into three channels, quasi-elastic scattering (QE), Single Pion Production (SPP), and Deep Inelastic Scattering (DIS). Example diagrams for the three processes are shown in Figure 1.4. The predicted total charged-current cross sections as a function of neutrino energy for each interaction type, as predicted by the NuWro [32] neutrino interaction simulation, are presented in Figure 1.5.

Charged-current quasi-elastic scattering (CCQE) from free nucleons is often modelled in the Lewellyn Smith formalism [33]. This channel has historically been regarded as the 'golden' interaction channel because it is possible to calculate the initial neutrino energy using just the

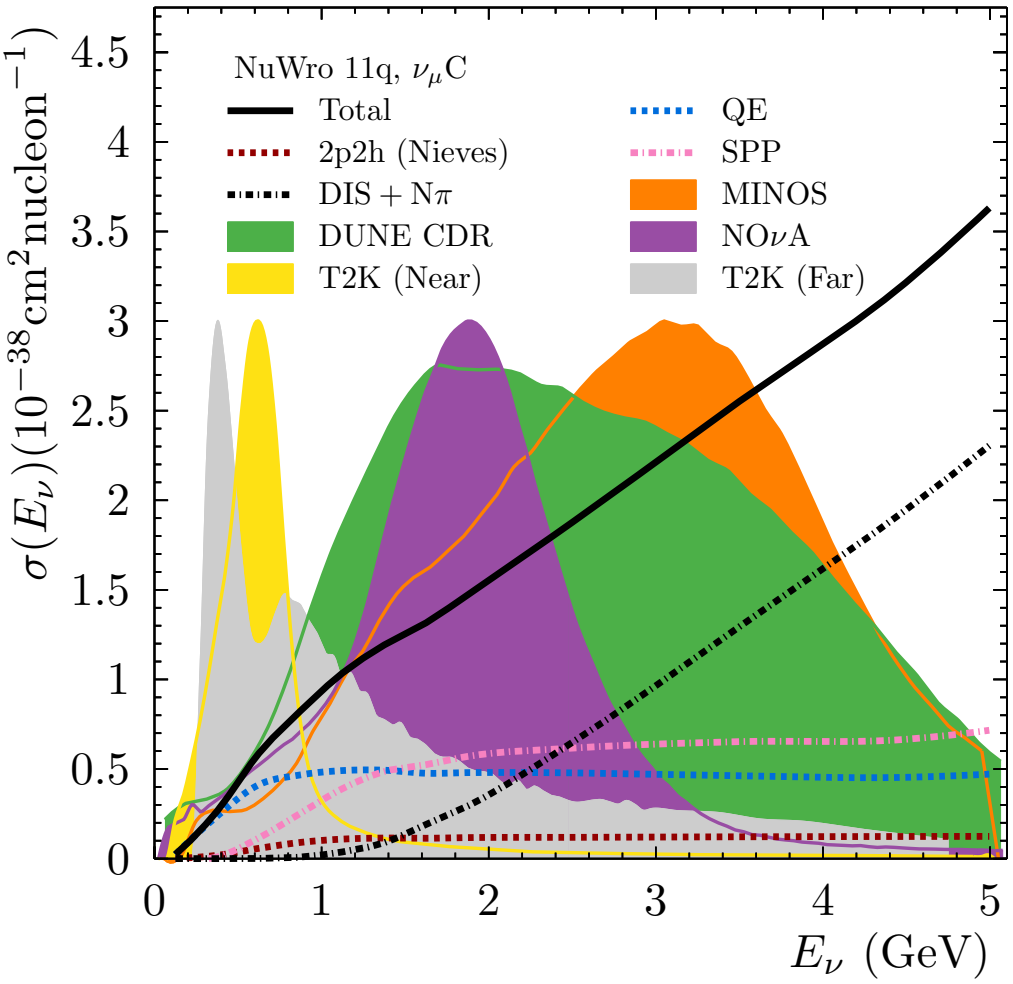


Figure 1.5: The charged-current cross section components (curves), as predicted by NuWro [32], overlaid on some experimental muon neutrino flux profiles (solid areas). Each flux is the predicted unoscillated flux sampled by a detector near to the production site, except for the T2K (Far) flux, which is the T2K (Near) flux convoluted with the muon neutrino disappearance probability for a 295 km baseline. The muon neutrino total cross-section and individual channel cross-sections vary significantly over the spectral width of current accelerator-based neutrino beams. Furthermore, the neutrino flux near to the beam production site and after oscillation differ in shape. To infer the properties of neutrino oscillation from observed neutrino interaction rates, scattering cross-sections must be well-understood.

observed final-state charged lepton kinematics as

$$E_{\text{rec}}^{\text{QE}} = \frac{2M_{\text{N}}E_\ell - M_\ell^2 + M_{\text{N}'}^2 - M_{\text{N}}^2}{2\left(M_{\text{N}} - E_\ell + |\vec{p}_\ell|\cos\left(\theta_\ell\right)\right)}.$$

Here, M_{N} is the mass of the initial state nucleon, which is assumed at rest[3]; $M_{\text{N}'}$ is the mass of the final-state nucleon; and E_ℓ, $\vec{p}_\ell$, and θ_ℓ are the energy, three momentum, and polar angle of the final-state charged lepton respectively. As the neutrino energy governs the oscillation probability, precision oscillation experiments must be able to reconstruct the neutrino energy spectrum in an unbiased way. However, as neutrino sources most often have a wide neutrino energy profile, it is not possible to use CCQE interactions in isolation. As highlighted in Figure 1.5, the flux profiles from each current and planned experiment span energies where the interaction rate of non-QE processes is significant.

Single pion production mostly occurs via neutrino-induced baryon resonance excitation and subsequent decay (an example diagram is depicted in Figure 1.4b). Non-resonant SPP channels may also be accounted for; these processes may occur through diagrams such as depicted in Figure 1.6 [34]. SPP models are significantly more complex than CCQE models. Historically, the model of Rein and Sehgal has been used [35], but more recent models such as the HNV model [34] and the GiBUU model [36] have shown better predictive power [37] and are capable of simultaneously describing pion production in electron scattering processes. Neutrino energy reconstruction for SPP interactions requires the measurement of both the final-state leptonic and hadronic systems—this renders accurate energy reconstruction more difficult. As can be seen in Figure 1.5, the deep inelastic scattering channel forms a significant contribution to the total cross-section for neutrino energies greater than about 1.5 GeV. The inclusive DIS cross section as a function of four momentum transfer is calculated from nucleon structure functions, usually determined from empirical parton distribution functions [38]; the final-state hadronic system is often simulated by fragmentation models, such as PYTHIA [39]. Neutrino energy reconstruction for DIS interactions requires accurate sampling of the full final-state hadronic system.

[3]For nuclear targets, this should be replaced with the effective mass of the target nucleon, which includes the nucleon binding energy

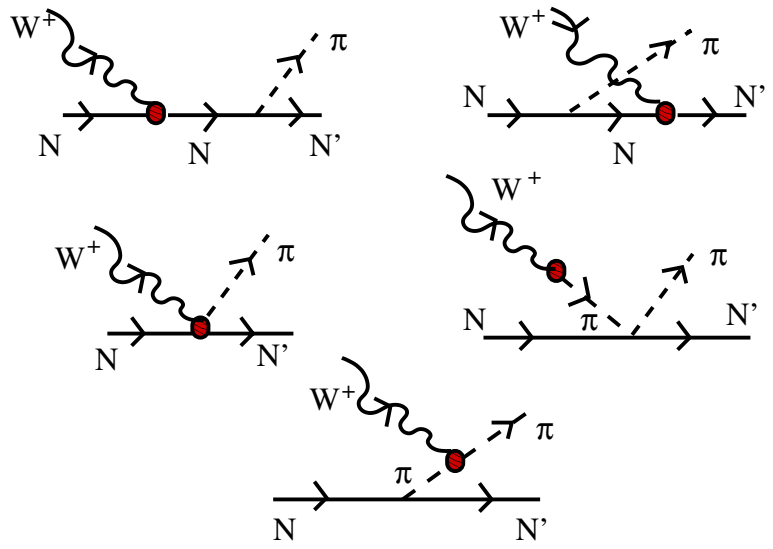
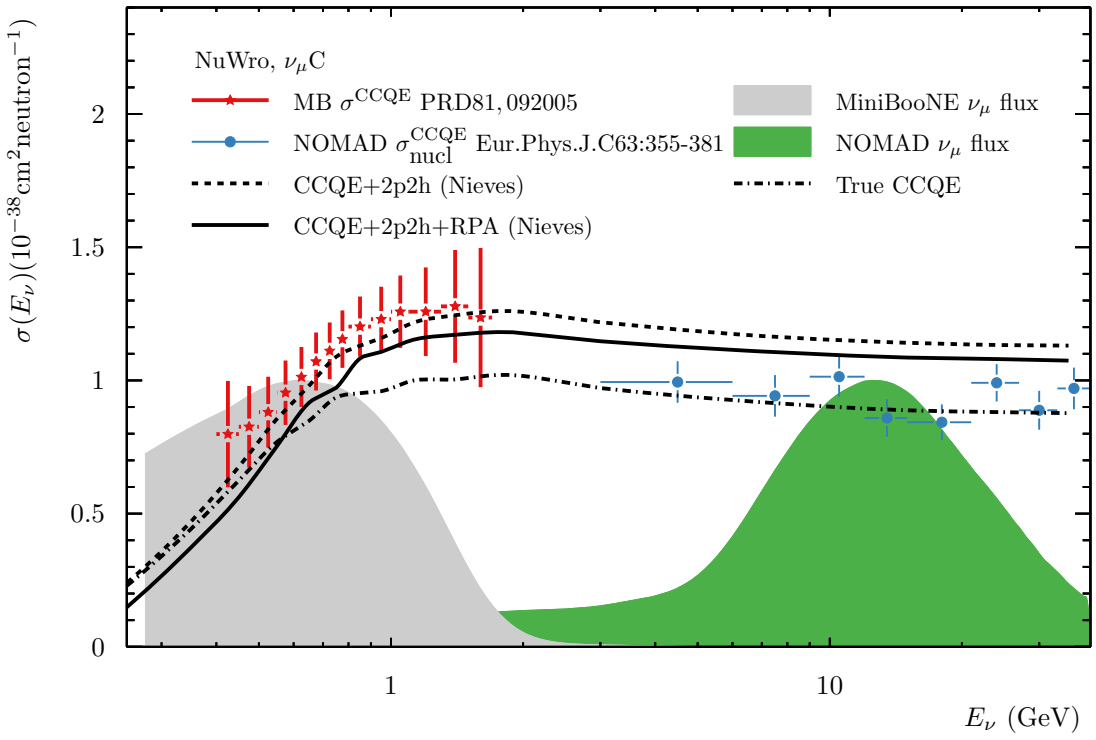


Figure 1.6: Example neutrino-induced single pion production diagrams, used in the HNV model, that do no include any intermediate baryon resonance. Figure from Ref. [34]

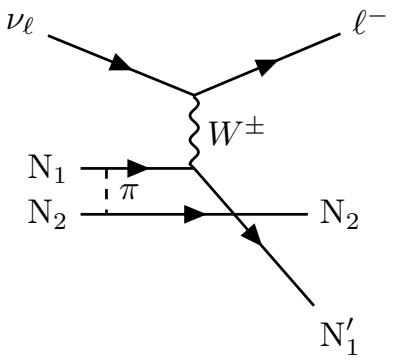
Free parameters in the neutrino–nucleon interaction models are often determined through comparison to historic hydrogen and deuterium bubble chamber scattering data, which are largely free of the effects of interactions with nucleons bound within nuclei, which will be introduced in the next section. An example tune of the NuWro CCQE and SPP models to data from the ANL and BNL bubble chamber experiments is presented in Appendix B.

1.3.2 Interactions with nuclei

All current and planned neutrino scattering experiments use nuclear targets. For a number of years, it was assumed that the free-nucleon interaction models could be simply extended to nuclear targets by boosting into the momentum distribution of a simple nuclear model and accounting for the effect of nucleon binding energies. A discrepancy between measurements of the true CCQE cross section at lower energies by MiniBooNE [40], and at higher energies by NOMAD [41]—shown in Figure 1.7a—led to the conclusion that significant components were missing from this simple picture. The MiniBooNE result was best-described with a quasi-elastic axial mass—an effective parameter in the dipole nucleon form factor used for CCQE [33]—of $M_{\text{A}} \simeq 1.3$ GeV. The NOMAD data were consistent with global fits to bubble chamber data, which preferred $M_{\text{A}} \simeq 1.03$ GeV [31]. This discrepancy resulted in the inclusion of a number



(a)



(b) CC 2p2h

Figure 1.7: (a) The NuWro prediction for the flux-unfolded, carbon-target, muon neutrino CCQE cross section over the $E_\nu \sim \mathcal{O}\,(1\text{–}40\,\mathrm{GeV})$. The 'corrected' CCQE data from the MiniBooNE [40] and NOMAD [41] experiments are overlaid. (b) A cartoon of a 2p2h interaction, where a neutrino scatters off two bound nucleons. For an in-depth description of a 2p2h interaction model, the reader is directed to Ref. [42].

of updated model components, some of which can be consistently combined and some that cannot. The unifying feature of the latest models is the inclusion of multi-nucleon effects. These come in the form of extra interaction channels, such as multi-nucleon currents, often termed '2p2h' or 'two particle two hole' interactions (with diagrams similar to Figure 1.7b) [42, 43]; more careful consideration of the nuclear model, such as the Benhar spectral function [44]; and mean-field corrections to the mediator boson propagators—calculated using the 'Random Phase Approximation' (RPA) [45, §3 A].

In neutrino–nucleus interactions, hadrons produced in the primary interaction will propagate out of the nucleus. During this process, re-interactions of the hadrons can take place that modify the observable hadronic system—referred to here as final-state interactions (FSI). These processes are often modelled by a semi-classical cascade in which hadrons are 'stepped' through the nuclear medium[4] and allowed to interact with other nucleons [47]. Such interactions can hinder the investigation of neutrino–nucleus interactions by obsfucating the properties of the primary interaction.

A sample of muon neutrino quasi-elastic interaction candidates, which is selected by requiring one final-state muon and no pions, may contain significant contamination from 2p2h events, where the primary interaction has a pion-less final state, and other events in which charged pions were absorbed by FSI processes. This renders obtaining a high-purity QE event sample difficult, even in the limit of a perfect detector. An example of the bias when using $E^{\mathrm{QE}}_{\mathrm{REC}}$ to reconstruct the neutrino energy for true QE events, 2p2h events, and other events with no final-state, charged pions is shown in Figure 1.8. This highlights the need for well-understood interaction physics when performing precision neutrino oscillation studies. While the T2K and NOνA collaborations have different approaches to investigating oscillations, both use some form of 'mapping' [5] between observed event properties and true neutrino energy. Even if the energy reconstruction does not presume the $E^{\mathrm{QE}}_{\mathrm{REC}}$ formula, any such mapping will contain model assumptions and these assumptions must be well motivated so as not to introduce bias into

[4] The use of a 'free first step' or formation zone has been shown to improve predictions [46].

[5] Whether this takes the form of a full neutrino flux and interaction model which is constrained by near detector data [48], or an un-smearing matrix used to unfold from near detector data to a 'true' flux shape [49], this is where assumptions about neutrino–nucleus interactions are evident.

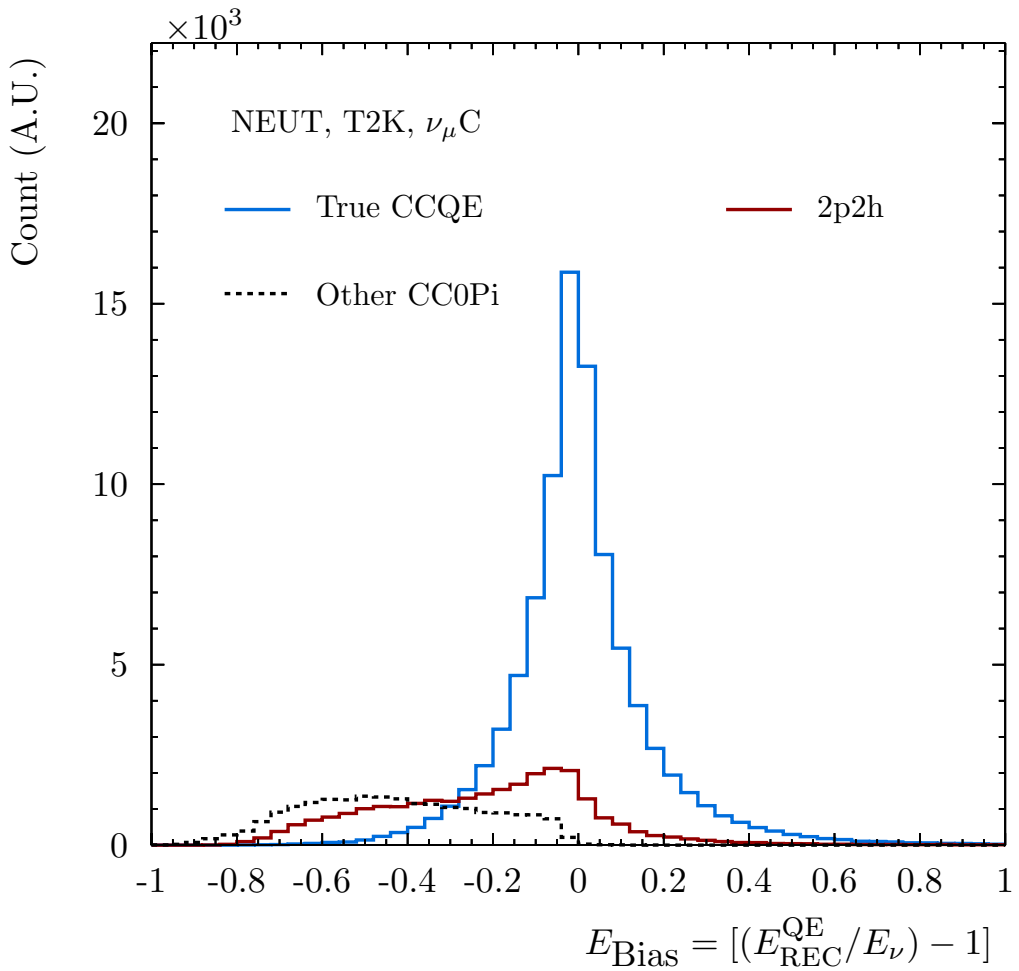


Figure 1.8: The quasi-elastic reconstructed energy bias for true CCQE interactions and other non-QE channels with no final-state pions. $E^{\text{QE}}_{\text{REC}}$ is an unbiased estimator for true CCQE interactions, with a width caused by the unknown event-by-event Fermi motion of the target nucleon. For non-QE interactions, $E^{\text{QE}}_{\text{REC}}$ systematically underestimates the true energy.

the inferred values of the oscillation parameters. While the use of a near detector does reduce the reliance on cross section models, like-for-like comparisons are not possible. Oscillation causes the far detector to be exposed to a different flux shape than the near detector, and by design, this difference is most prominent at the peak energy. The T2K near- and far-detector muon neutrino flux shapes are shown in Figure 1.5. Furthermore, when studying electron neutrino appearance the properties of electron neutrino interactions with matter must also be well-understood—it is not currently clear that they are [50]. As a result of these complex and coupled model components, the only feasible way to interpret real neutrino scattering data is through the use of 'event generators'.

1.3.3 Interaction simulations

Event generators use Monte Carlo techniques to produce fully-simulated neutrino interaction events from theoretical models. Their use allows interpretation of neutrino scattering data in the context of full interaction models, after corrections for detector response have been made. The validity of any such interpretations is related to the degree to which the model can predict the data in the absence of fine-tuning. Event generators provide an indispensable tool in the development of particle physics analyses. However, over-use of the simulation can 'bake in' bias, or cause interesting discrepancies to be discarded when investigating real data—an example of this is candidly presented [51], specifically in Figure 20 and 21, where the interaction model used in the simulation is shown to affect the extracted cross section.

The inclusive muon neutrino charged-current cross section prediction from two interaction simulations, NEUT v5.3.5 [52] and NuWro version 11q, for a hydrocarbon and a lead target[6] are shown in Figure 1.9. The two predictions are similar for lower energies where CCQE scattering dominates, but differences in the treatment of SPP interactions and the transition region between SPP and DIS interactions are evident for energies corresponding to the tail of the T2K near detector flux.

The next chapter describes the T2K near detector and motivates the selection of a predominantly lead-target muon neutrino interaction candidate sample.

[6] These are the relevant nuclei found in the ND280 electromagnetic calorimeters, which will be the main interaction target for the analysis presented in this thesis.

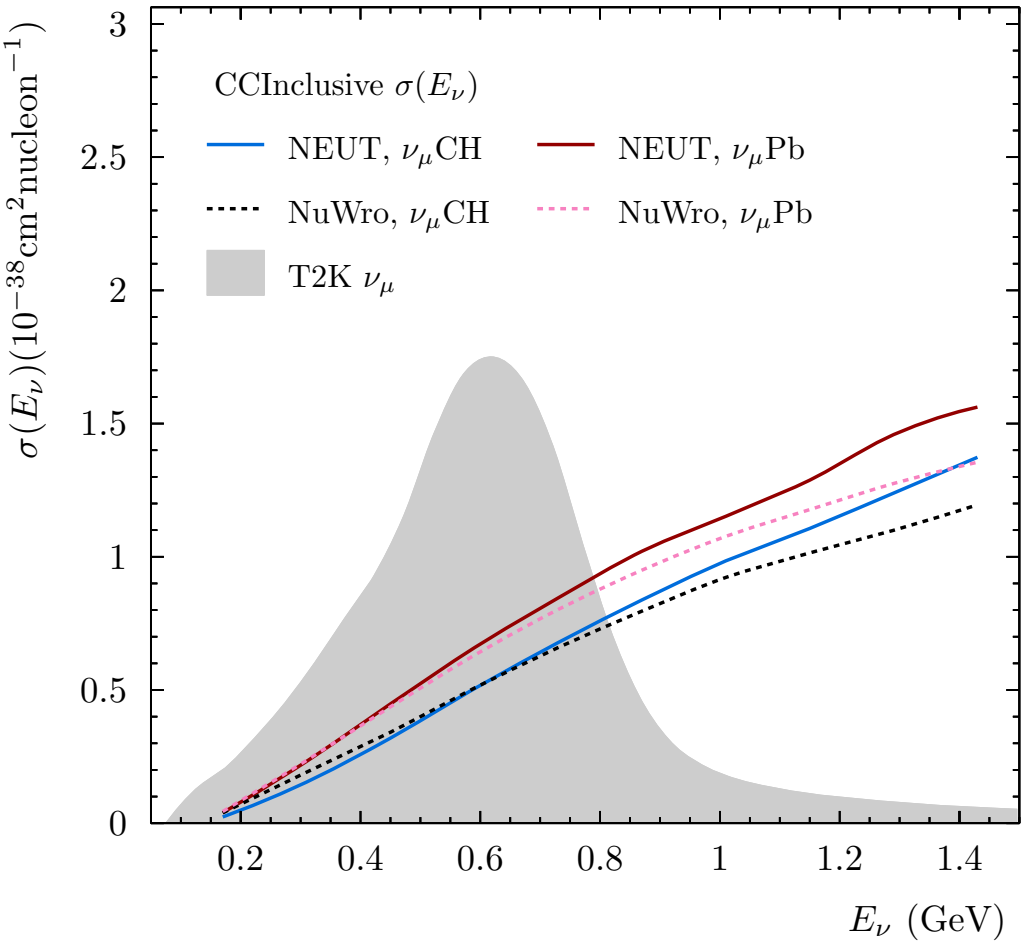


Figure 1.9: The flux-unfolded, charged-current muon neutrino inclusive cross section (curves) for lead and hydrocarbon targets, as predicted by NEUT [52] and NuWro. The T2K muon neutrino near detector flux prediction (solid area) is also shown [53].

Chapter 2

The T2K near detector and ECal-target motivations

The T2K experiment investigates muon neutrino flavour change over a 295 km oscillation baseline [11]. The experimental apparatus consists of a number of different components, each of which are important to the highly-significant oscillation measurements that have been published [54, 55, 48, 7, 56, 57, 20]. This thesis focuses exclusively on an analysis of neutrino scattering occurring within the off-axis near detector, ND280. The other components of the experiment have been discussed in great detail elsewhere. The on-axis near detector, INGRID [58], is used to sample the neutrino beam direction, which ensures that the flux that the other detectors are exposed to remains consistent. A number of high statistics cross-section measurements have also been performed using INGRID [59, 51, 60]. The off-axis far detector, Super-Kamiokande, was first used in the discovery of muon neutrino disappearance as a function of distance travelled by observing muon neutrinos arriving from cosmic ray interactions in the atmosphere. In this seminal discovery, the flux of muon neutrinos seen to be arriving from above was significantly greater than that of those arriving from below—where the neutrinos had traversed the diameter of the planet before being observed [3].

The next section briefly introduces the J-PARC proton driver and the production of the neutrino beam. The rest of the chapter discusses the T2K off-axis near detector and motivation for the

analysis presented from § 4 onwards.

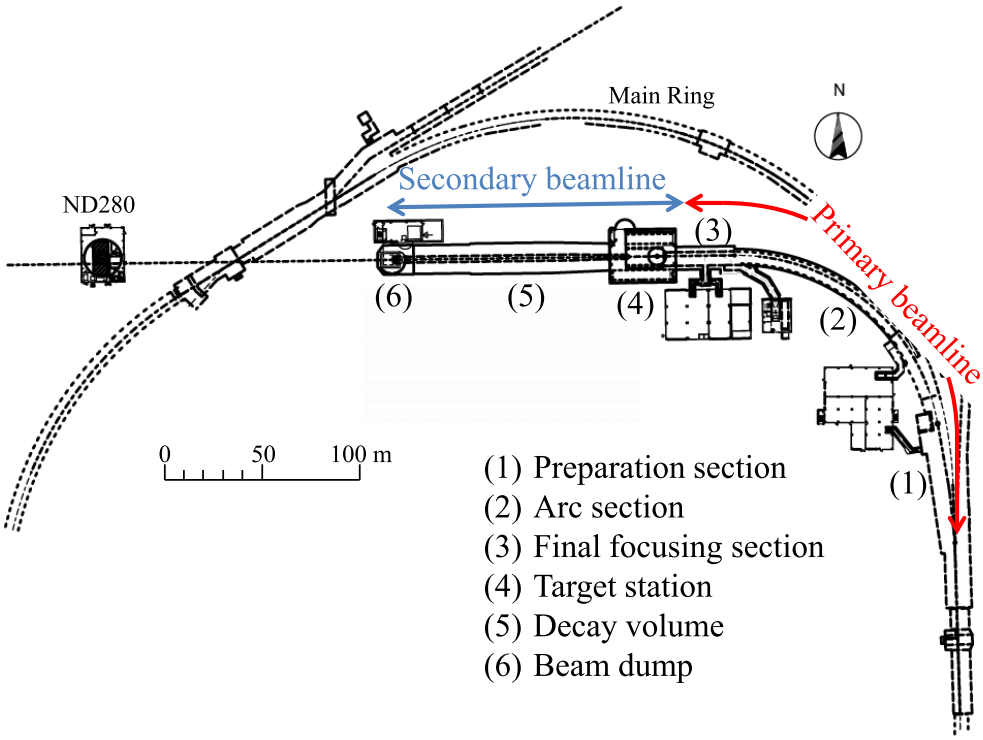
2.1 The J-PARC neutrino beam

The J-PARC neutrino beam is produced by impacting 30 GeV protons onto a graphite target. A short description is given here, and the interested reader can find more details of the neutrino beam production in Ref. [53].

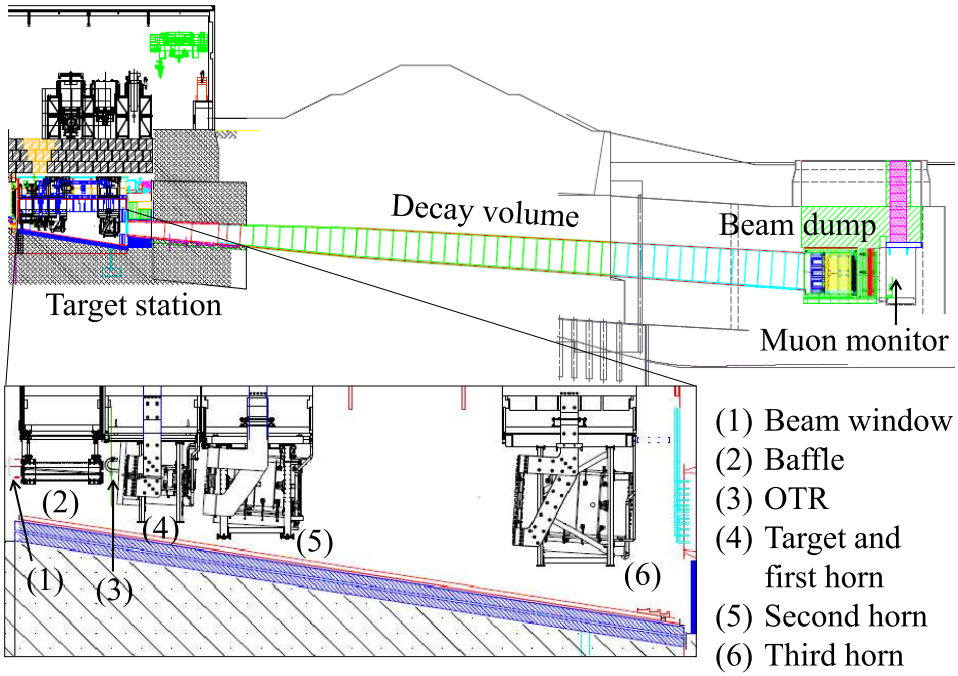
Protons are accelerated up to 30 GeV by the Main Ring (MR)—the third and final stage in the J-PARC proton accelerator facility [21]—and directed down the neutrino beamline, which is depicted in Figure 2.1a. The hadrons produced in proton interactions in the target are charge-selected and focussed by three neutrino horns [61]. The secondary beam of particles is then left to decay in-flight while travelling down a 96 m tunnel; many decay to neutrinos and other particles. Surviving hadrons are stopped in the beam dump and through-going muons are sampled by the muon monitor, which is used to provide pulse-by-pulse monitoring of the properties of the secondary beam [62]. A diagram showing the target station, decay volume, and beam dump is reproduced in Figure 2.1b.

Protons are extracted from the MR into the neutrino beam line in 'spills', which consist of eight narrow bunches of protons. Each bunch is approximately 5 ns wide and separated from the next bunch by about 580 ns [62]. This structure is reflected in the resulting neutrino beam. Currently, the MR repetition rate (which corresponds to the spill frequency) is 0.4 Hz [62]. Planned upgrades to the MR will allow for an increased repetition rate, which will facilitate the J-PARC neutrino beam in surpassing the design power of 750 kW and reaching approximately 1 MW [17].

The number of Protons Per Spill (PPS) and total number of protons received, or Protons On Target (POT), for the first four T2K data-taking periods are presented in Figure 2.2a. The peak beam intensity is approximately 1.2×10^{14} P_{PS}, which for the proton energy and repetition rate of the MR corresponds to a power of about 235 kW. The average beam power has continued to rise since Run 4 because of an increased number of protons in each bunch. For the most recent



(a)

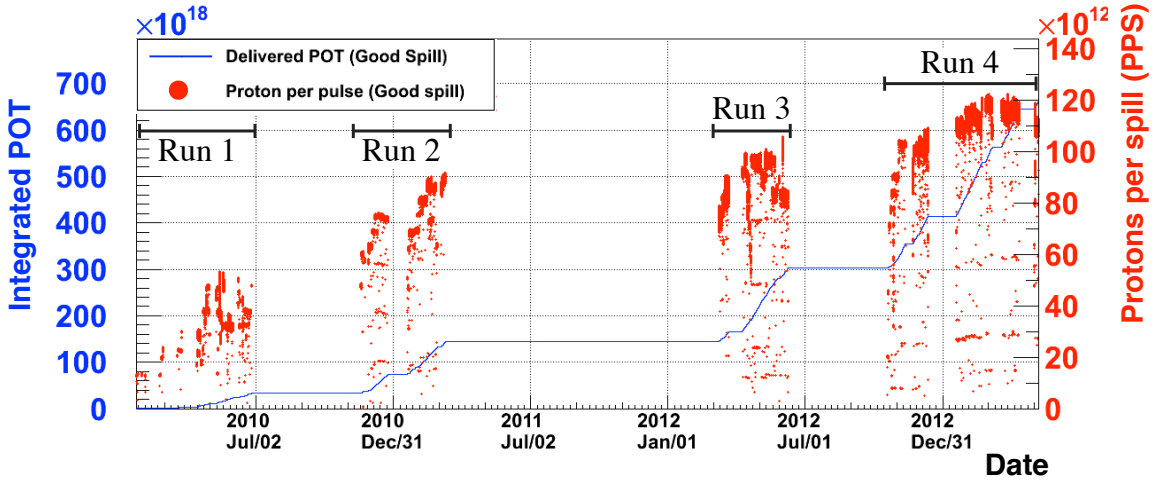


(b)

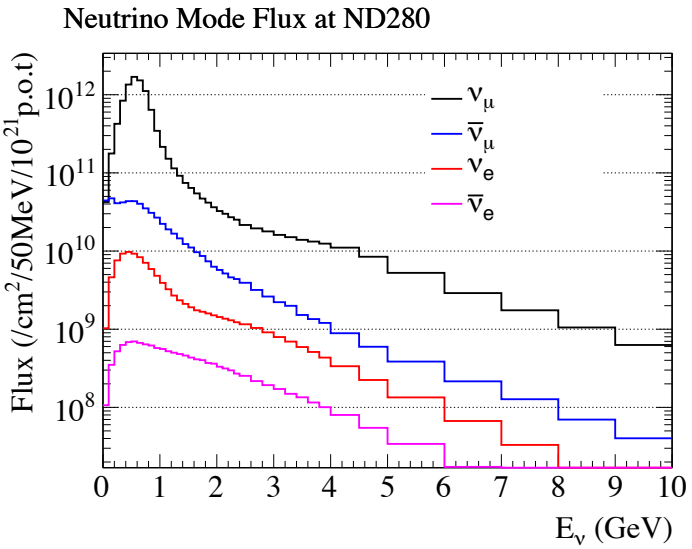
Figure 2.1: Schematics of the primary and secondary beamlines that are involved in the production of the J-PARC neutrino beam. The lower figure shows the secondary beamline, in which a beam of secondary particles produced from hadronic interactions in the target are focussed and allowed to decay in-flight to produce the neutrino beam. Just before the protons arrive at the target, the Baffle collimates, and the Optical Transition Radiation (OTR) monitor measures the two dimensional profile of the beam. The OTR is described in detail in Ref. [63]. Figures from Ref. [53]

run (Run 8), the beam power peaked at about 450 kW. More-recent near detector data are excluded from analysis as they were either not processed at the time of writing, or were taken in a predominantly anti-neutrino beam (*c.f.* § 1.2.1).

The predicted muon neutrino flux at the near detector in neutrino mode is shown in Figure 2.2b. The neutrino flux variation as a result of the angular span of the near detector is presented in § 2.6 and the effect of uncertainties related to the flux prediction on the analysis will be discussed in § 5.1.



(a)



(b)

Figure 2.2: (a) The protons per spill and the total number of protons on target up to the end of T2K Run 4. The beam intensity (PPS) has steadily increased through the life of the experiment. Data from the latter part of Run 3 and from the whole of Run 4 are used in the analysis presented here. (b) The neutrino mode flux predictions at ND280, sub-divided by neutrino species. Figure from Ref. [64]

2.2 The magnetised off-axis near detector: ND280

ND280 is a magnetised, off-axis, general purpose neutrino detector—a diagram of ND280 is shown in Figure 2.3. Multiple sub-detectors are housed within a large electromagnet which produces a 0.2 T field in the centre of the detector during nominal operation. The innermost region, the Tracker, contains three gaseous argon Time Projection Chambers [65] (TPCs), interleaved with two finely segmented plastic scintillator tracking detectors (Fine-Grain Detectors [66]—FGDs), and capped at the downstream end with a sampling electromagnetic calorimeter (DsECal). The two FGDs are used as the primary neutrino interaction target. The most upstream FGD, FGD1, consists entirely of layers of plastic scintillator bars, which provides good vertex position reconstruction and tracking of charged particles near to an interaction vertex. Alternate layers are oriented at 90 degrees to each other to facilitate three-dimensional reconstruction. The more downstream FGD, FGD2, has some water-filled layers which act as a passive target for beam neutrino interactions. Measurements from both FGDs are combined to investigate differences in neutrino interactions on carbon and water targets, as water is the target material at the far detector. The TPCs provide high-precision position tracking, momentum reconstruction, and energy loss measurements for charged particles, which affords accurate Particle IDentification (PID).

The Tracker is surrounded by another electromagnetic calorimeter, the Barrel ECal [67], which consists of six separate modules. Both the Barrel and Downstream ECals are built from alternating layers of active plastic scintillator bars and lead absorber sheets—an illustration of how a charged particle activates bars in the two orientations, or 'view's, of a module is shown in Figure 2.4a. The lead layers increase the number of radiation lengths across each module to span the necessary range of particle energies that can be fully contained and sampled. The ECal provides good discriminating power between electrons and muons which enter from the Tracker—electrons will shower leaving cone-like energy deposits whereas muons will not, leaving track-like signatures [67]. The ECal sub-detector almost completely surrounds the Tracker in an attempt to measure all charged outgoing particles from interactions occurring within.

The most-upstream end of the detector houses the Pi-zero Detector (PØD) a detector designed

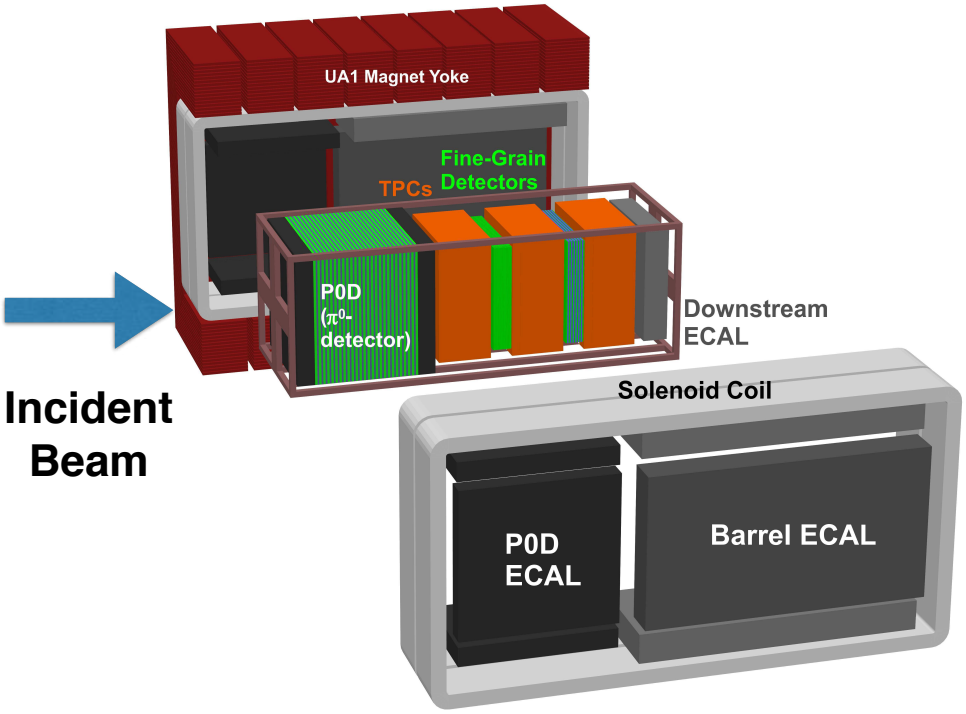
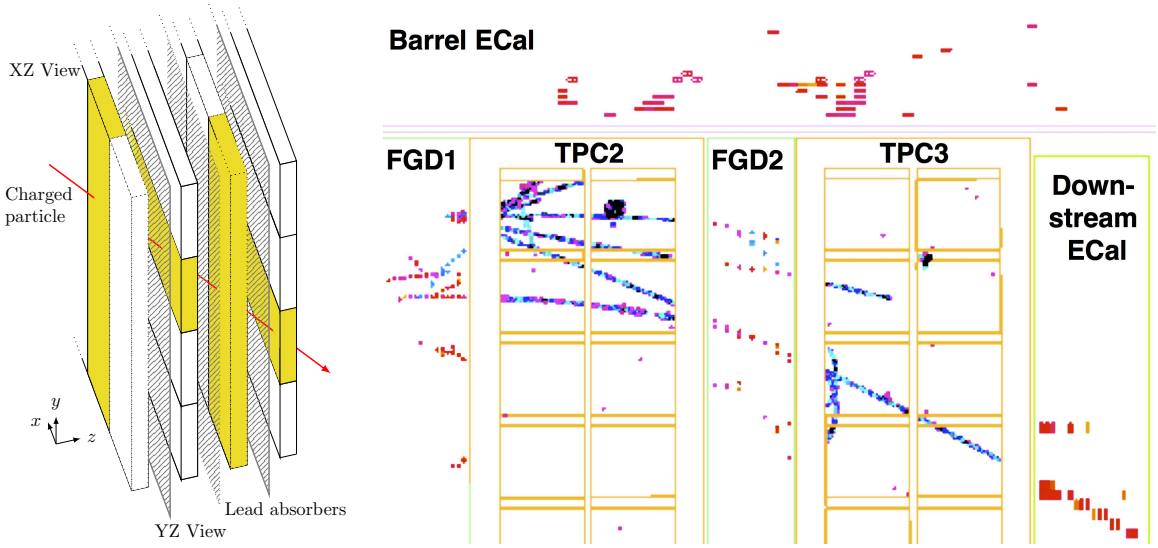


Figure 2.3: A diagram depicting the various sub-detectors that make up the T2K off-axis near detector.

to study neutrino-induced neutral pion production [68]. Decays of neutral pions to photons lead to electromagnetic showers that can be mistaken for electron neutrino interactions at Super-K. A good understanding of the neutrino-induced neutral pion production rate is important to maximise the significance of the electron neutrino appearance measurement. The PØD is constructed from alternating scintillator layers and water bags, where the bags can be drained of water and filled with air allowing water-target measurements through the statistical subtraction of 'water' and 'air' data. Lastly, in-amongst the iron clams of the magnet return yoke sits the Side Muon Range Detector (SMRD) [69]. The majority of entering and exiting charged particles will leave some trace in the SMRD.

An example of the event topology used to sample the un-oscillated neutrino beam for input into the T2K oscillation analysis can be seen in Figure 2.4b. An interaction candidate can be seen in FGD1 which produces a number of charged particles which pass through the rest of the Tracker and leave deposits in the Barrel and Downstream ECals.

ND280 is a general purpose detector that is used to measure a variety of neutrino interaction channels. At the time of writing, published measurements include: carbon-target muon neutrino



(a) ECal layer configuration (b)

Figure 2.4: (a) Exploded view of the ECal scintillator and absorber configuration. Alternate layers of scintillator bars are oriented at 90° to facilitate 3D reconstruction. (b) ND280 event display showing a candidate interaction in FGD1, the charged final state particles can be seen travelling through the tracker with one leaving a deposit in the DsECal. The neutrino beam travels left to right.

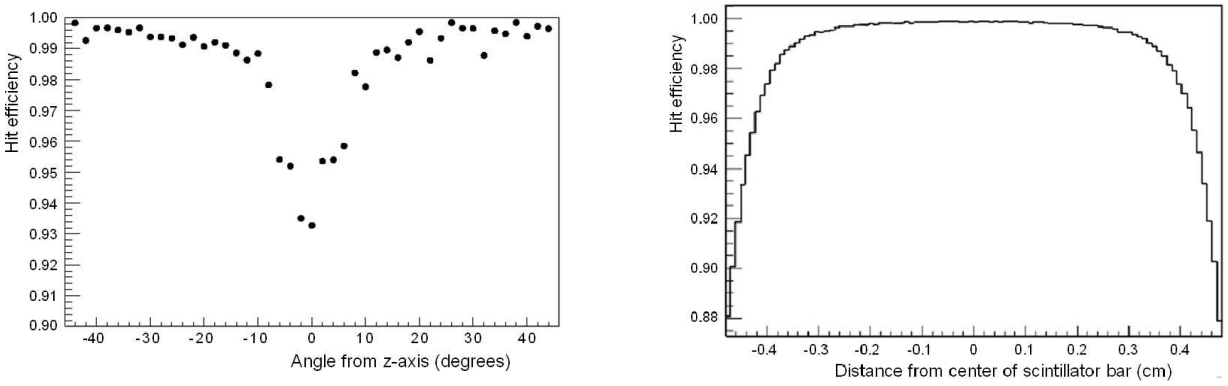
inclusive charged-current cross section [70], carbon-target electron neutrino inclusive charged-current cross section [71], and carbon-target muon neutrino true CCQE cross section [72], as well as many others [73, 74, 71, 50]. Perhaps the two most interesting measurements are the CC0π [75] and CC1π [76] measurements on carbon and water targets respectively. Both have signals defined by the topologies of particles that leave the nucleus (*i.e.* after FSIs), rather than signal definitions that require corrections that account for FSIs. For example, the extraction of a 'true' CCQE cross section from an event sample with no observed pions involves using the interaction model to determine the contamination from other interaction channels with similar observable final state topologies. This inherently introduces unnecessary model assumptions into the published data and adds little extra information to final-state topology-based results. These two analyses have spearheaded a focus on model-independent results within the neutrino scattering community, which is beginning to become the standard practice.

2.3 The ND280 Tracker

The ND280 Tracker is used as the main detector element for the majority of analyses performed using ND280. The FGDs present a significant target mass of $\sim 1.1 \times 10^3$ kg that results in a reasonable interaction rate. The scintillator bars have a 10×10 mm cross section and are arranged in thirty alternately-oriented layers which allows accurate three-dimensional vertex reconstruction and precise tracking of final state particles leaving an interaction. Every layer is oriented perpendicular to the beam axis, which gives the best reconstruction efficiency for forward-going final-state particles. Activated bars are read out from one end via wavelength-shifting (WLS) fibres coupled to Multi-Pixel Photon Counters (MPPCs), silicon photo-detectors developed for T2K that are capable of functioning effectively within a magnetic field [77]. Each bar is co-extruded with a titanium oxide-based reflective coating to increase the scintillation light read out by the MPPCs [66].

For charged particles that are produced and stop within the same module, the FGD reconstruction algorithms can estimate momentum for each particle hypothesis by the length of a reconstructed track and also provide energy loss-based PID. Because of the orientation of the FGD scintillator layers, tracks that are approximately perpendicular to the beam axis do not cross many layers and will be poorly reconstructed. As the FGDs are not used as the primary target by the analysis presented here, it is only important that their hit efficiency is high enough to allow charged particles to be followed through a module. The FGD hit efficiencies for cosmic muon tracks as a function of the angle between the reconstructed track and the long bar axis, and trajectory distance from the centre of the bar is shown in Figure 2.5. The efficiency is suppressed for particles that travel approximately along a bar or those that only pass through the very edge of a bar, but for tracks passing through an FGD module from a TPC or a Barrel ECal module this efficiency is more than adequate.

The ND280 TPCs play a significant role in the analysis presented here, providing precise momentum measurement and PID for final-state charged particles leaving a Barrel ECal module. Each TPC is filled with a mostly-argon gas mixture which is readily ionised by through-going charged particles. The liberated electrons are drifted by a 25 kV potential difference to readout

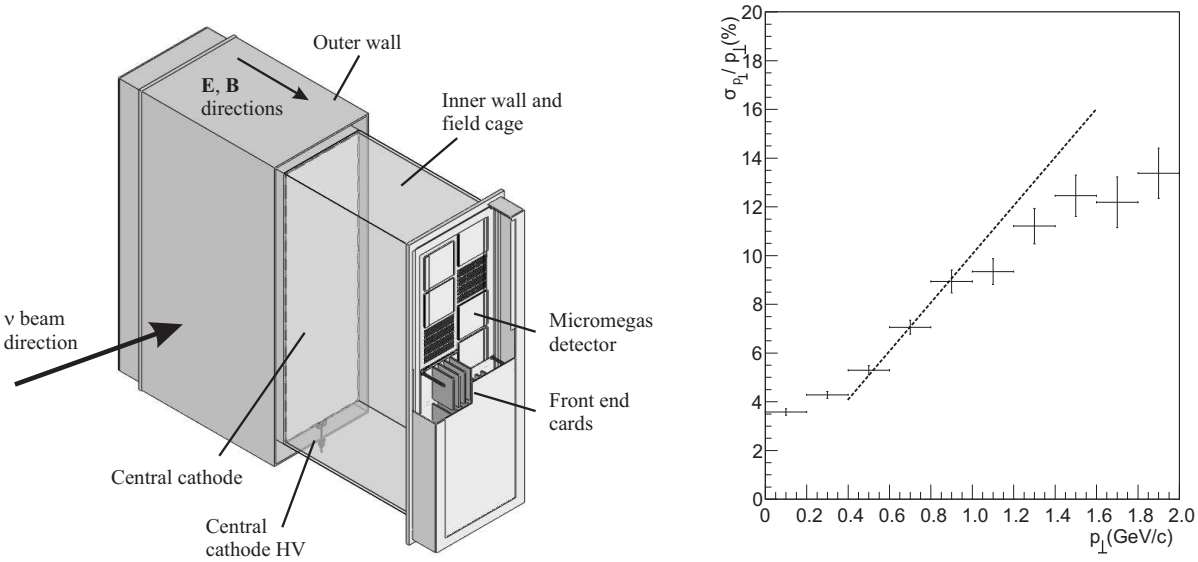


(a) (b)

Figure 2.5: The FGD hit efficiency as a function of track angle relative to the long axis of a bar, (a), and distance from the centre of a bar, (b). Figures from Ref. [66]

planes on the anodes at either side of each module. A schematic of one ND280 TPC module is shown in Figure 2.6a. The functional unit of the readout system is a MICROMEGAS module [78], a 36×34 cm planar unit of 48×36 readout pads. Clusters of hits are formed from neighbouring pads, which are clustered by row or column depending on the local track angle of the ionisation signature; *i.e.* for a particle travelling perfectly horizontally, along a row of 48 pads, vertical clustering would be used and would result in 48 distinct position measurements. As the drift potential is uniform across the volume, the drift velocity of the ionisation deposit is approximately constant and can be used to determine the relative positions of clusters in the direction of the electric field, x. Without any other timing information it is not possible for the absolute x position to be determined because the drift time for the earliest part of deposit is not known. The neighbouring scintillator detectors are used to provide a reference time measurement for each track, which is used to precisely determine the position in the drift dimension using the mean drift velocity. This technique affords a spatial resolution for tracks of within $\mathcal{O}(0.1\text{ mm})$, which was one of the design requirements of the system [65].

The momentum and electric charge of a charged particle can be reconstructed by fitting the observed track trajectory, taking into account the presence of the electric and magnetic fields. The track curvature allows the momentum component transverse to the magnetic field to be determined, which along with the reconstructed track direction yields the estimated particle momentum. The fractional momentum resolution as a function of the transverse momentum is shown in Figure 2.6b; it can be seen that the momentum resolution meets the design goal [65].



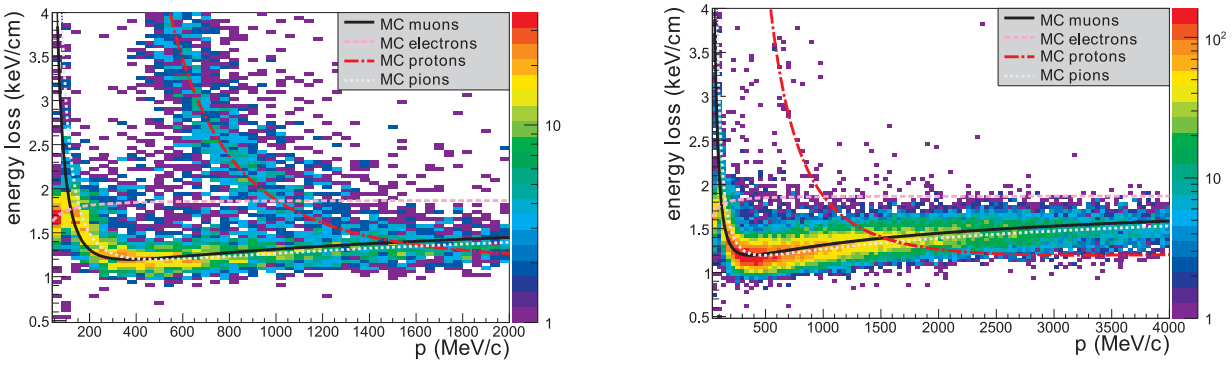
(a) (b)

Figure 2.6: (a) A diagram depicting a TPC module. The potential difference between the central cathode and anodes on either x face drifts electrons released by ionisation of the TPC gas towards the MICROMEGAS detector modules. (b) The TPC reconstructed transverse momentum resolution as a function of transverse momentum. The dotted line shows the design target resolution. Figures from Ref. [65].

The TPC particle identification uses measurements of the ionisation energy loss that a charged particle experiences when propagating through the TPC gas. The expected energy loss is described by the Bethe formula,

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \frac{nz^2}{\beta^2} \left(\frac{e^2}{4\pi\varepsilon_0} \right)^2 \left[\ln\left(\frac{2m_e c^2 \beta^2}{I\left(1-\beta^2\right)} \right) - \beta^2 \right],$$

where n, z, and I are the electron number density, proton number, and the mean excitation energy of the material respectively. The relevant properties of the propagating particle are encapsulated within the relativistic beta factor that includes the mass and momentum of the particle. For a deposit with a given reconstructed momentum, an expected energy loss can be calculated for each PID hypothesis from the known particle masses. A PID likelihood can be formed between the observed and expected energy losses for each hypothesis from the observed momentum-by-curvature. The expected and observed energy loss profiles for relevant positive and negative charged particles are shown in Figure 2.7. The energy loss is calculated per cluster by summing up the total measured charge per activated pad and dividing by the local track



(a) (b)

Figure 2.7: The measured energy loss and momentum for positive ((a)) and negative ((b)) charged particles passing through the TPC. Overlaid are the expected energy losses for muons, electrons, protons and charged pions. Figures from Ref. [65]

length through the cluster pads. The estimated energy loss per distance travelled for the whole deposit is then calculated as a truncated mean which includes the 70% of the clusters with the lowest calculated energy loss. This removes a long tail that would otherwise reduce the accuracy of the estimator. The energy loss resolution of the TPCs was determined to be better than the design requirement of 10% [65]. However, because of the similarity in masses of muons and equivalently charged pions, discrimination between the two purely through this technique is very poor. For the majority of interesting event topologies that include a muon and a pion, they will be of opposite charge, which reduces the impact of this issue.

At the time of writing, there are ongoing efforts to select neutrino interactions occurring on argon nuclei within the TPC drift gas; in the current data set there are expected to be $\mathcal{O}(1000)$ such events [79]. Because of the low density and high precision tracking within the TPC, it is expected that the hadronic final state of argon-target interactions can be investigated with unprecedented precision. As liquid argon is used as an active target for a number of current [80, 81] and future [19] neutrino experiments, such a measurement may provide a valuable constraint on the nuclear effects involved in argon-target neutrino interactions.

2.4 The ND280 Barrel ECal

The Barrel ECal modules are used as the primary interaction target in this analysis. Each module is constructed from 32 layers built from bars of plastic scintillator and 1.75 mm-thick lead absorber sheets. The bars are oriented so that alternate layers lie with the long bar axis parallel or perpendicular to the neutrino beam axis. This construction was used to meet the design requirement of optimally sampling charged particles that exit the Tracker; it also means that high-angle particles produced in interactions occurring within an ECal module are well-oriented for efficient reconstruction. The four top and bottom modules, and the two side modules have the same dimensions and each module contributes a target mass of 5500 kg and 8200 kg respectively. Each scintilator bar has a cross section of 10×40 mm, which results in reasonable tracking and vertex reconstruction resolution. The assessment and discussion of the ECal tracking and vertexing performance is the focus of the next chapter. The ECal bars are read out by WLS fibres attached to MPPCs, similarly to the FGD bars. Bars that are oriented parallel to the neutrino beam axis are read out by an MPPC at both ends (double-ended) because the length of these bars (3.84 m) is comparable to the attenuation length of the WLS fibres ($\sim$ 4 m) (Figure 9)[67]. For single-ended bars, the far end of the WLS fibre is mirrored to increase the scintillation light directed toward the MPPC.

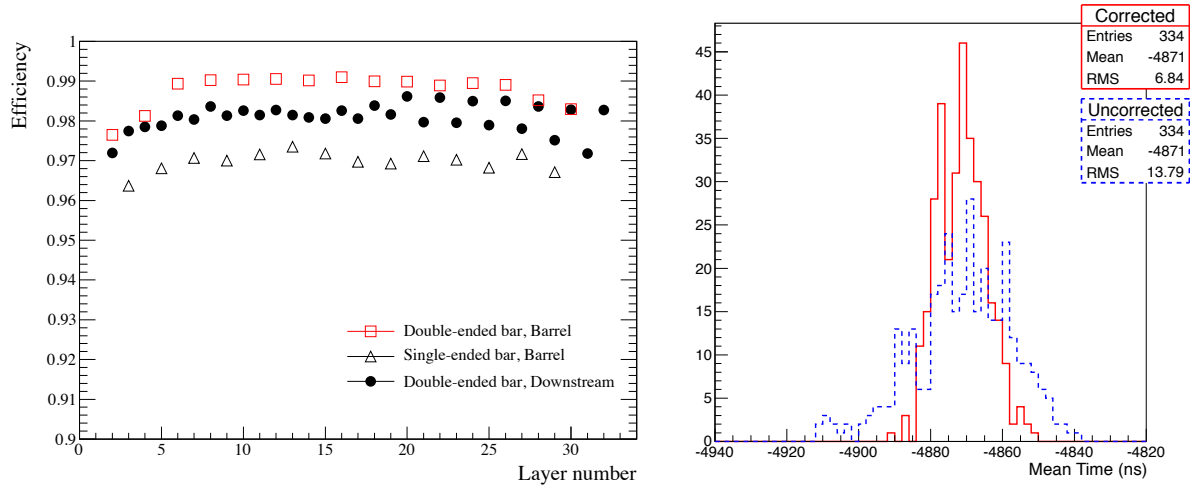
When a pixel of an MPPC is hit by a photon, a charge pulse is generated. The charge from each MPPC is collected onto a capacitor, from which the integrated charge is read every 580 ns. The capacitor readout times are chosen so that each beam bunch arrives close to the centre of an integration period.

The ECal response is calibrated in two stages. The online calibration uses the dark noise spectra from each channel to tune the over-voltage of the corresponding MPPC so that a single photon hit generates a consistent charge response. This calibration affects the recorded data and is performed approximately once per week when taking data as the MPPC response is sensitive to ambient temperature changes [82]. The offline calibration uses samples of cosmic muons to ensure that the hit timing and charge response to through-going particles is consistent across the detector. The attenuation properties of each WLS fibre are determined from the

relationship between the recorded charge and the reconstructed distance along a bar that a muon passed.

The ECal MPPC channels are read by a hierarchy of data acquisition (DAQ) subsystems. Blocks of up to 16 MPPCs are read by a single 'TriP-T' chip, a microprocessor originally designed for the DØ experiment at Fermilab. Four of these chips are housed on each of the 334 'TriP-T Front end Boards' (TFBs), with twelve 'Readout Merger Modules' (RMMs) combining the output of up to 48 TFBs each. The RMM output, along with the DAQ output from the other ND280 sub-detectors, is built into an 'event', which corresponds to the recorded detector response for a single detector trigger, such as a beam spill.

The two most important properties of the ECal for this analysis are the hit efficiency and the hit timing resolution, which are used in the hit clustering algorithms (described in the next chapter). The hit efficiency as a function of layer number, determined from a sample of through-going cosmic muons, is shown in Figure 2.8a. The efficiency is reasonably high and consistent throughout the bulk of each ECal module, the efficiency for the double-ended bars is slightly higher; this is expected because the MPPC at either end may record a hit. The hit time for the same sample of muons, averaged over the readout channels connected to each of the 334 TFBs, is presented in Figure 2.8b. The spread in the calibrated mean hit time is more than sufficient for efficient hit clustering.



(a) (b)

Figure 2.8: (a) The ECal hit efficiency as a function of layer number for a sample of through-going cosmic muons. The efficiency is reasonably high and flat across each type of bar. (b) The mean hit time from each ECal readout subsystem for a sample of cosmic muons. After calibration, a timing resolution of about seven nanoseconds was achieved, which is sufficient for effective hit clustering. Figures from Ref. [67].

2.5 The ND280 software

The software used in the simulation and analysis of data from ND280 was designed to be modular. The long-established `ROOT` framework [83] is used throughout for data persistence, analysis, and visualisation.

When processing real data from the detector, the raw binary event output of the DAQ systems is first converted to 'digits' in a `ROOT`-based event format. These digits contain information about the time and 'charge' of an activated detector element[1]. Sub-detector specific and time-dependent calibrations are applied to the digits, which are used to create the detector hits that are processed by the reconstruction algorithms.

When simulating neutrino beam interactions, a number of stages are performed before the simulated detector response can be processed in an equivalent way to the real data. The first stage is to simulate a realistic flux of neutrinos. The proton interactions and secondary re-interactions in the T2K target are simulated by FLUKA [84] and GCALOR [85] respectively.

[1]For the different detection elements, the 'charge' of a hit has different meanings. For example, the 'charge' of an ECal hit is related to the number of photons counted by each MPPC. It is the job of the reconstruction and analysis algorithms to translate this into physically useful information.

Both models are tuned to experimental hadronisation data, most notably, the FLUKA model is tuned to data from the NA61/SHINE [86] experiment at CERN [53]. The products of these interactions are propagated out of the target, through the electro-magnetic horns, and into the decay volume by a GEANT3-based model [87]. Neutrinos produced in decays of these particles are recorded as a 'vector' of simulated neutrinos. This vector contains information on the production species, position, direction, and energie of each simulated neutrino. These neutrinos are then traced through the ND280 detector geometry and allowed to interact with material in the detector. Interaction probabilities are determined by weighting positions along the trace according to the total neutrino cross-section for the relevant material mixture and neutrino species. These cross-sections are calculated by a neutrino interaction simulation; in the ND280 software, this is most often NEUT, but can also be NuWro or GENIE [88]. In this way, a vector of neutrino properties, interaction positions, and target-nuclei is built up. For each neutrino that interacts, the event generator simulates a primary interaction and any FSIs that may take place. For a given neutrino energy, interacting with a given target nuclei, the primary interaction is chosen proportionally from the relative cross-sections of each interaction channel (*c.f.* Figure 1.5 and § 1.3.1). The output of the event generator is the list of particles that escaped the target nucleus in the simulation. In addition to neutrinos which interact within ND280, a significant number of interactions occur in the earth between the beam dump and the near detector. These so-called 'sand' interactions[2] produce a flux of particles—mostly muons—which enter the detector from outside. For some analyses, this constitutes a significant source of background and as a result must be simulated. Unfortunately, while these are simulated as before, but with a different geometry, the further processing of the final state particles is kept separate from those produced in interactions occurring within ND280. This means that it is not currently possible to determine the effect of 'sand muons' within the same simulated beam spill as signal interactions, they can only be used to characterise entering backgrounds. These first two simulation stages are not specific to the ND280 software.

The particles produced by the event generators are then propagated through the ND280 detector geometry by a Geant4-based model [89], which incorporates a number of secondary interaction

[2]So called because the earth surrounding the ND280 complex is mostly sand.

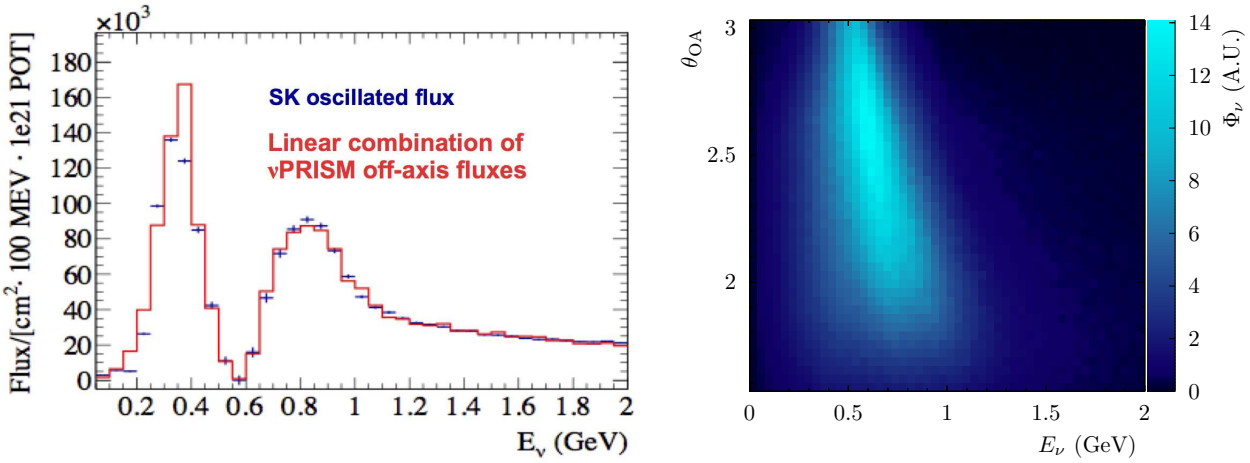
processes. Energy loss and electro-magnetic showering models are used to determine where energy is deposited in the detector. Energy deposits that occur within active detection elements are used to simulate the detector response. In addition to the simulated digits induced by beam neutrinos, each sub-detector has a noise simulation, which can inject extra charge into the simulated readout electronics according to data-motivated dark noise distributions. The output of the electronics simulation is qualitatively similar to the real detector output, albeit with extra 'simulation truth' information associated with each simulated digit. The processing from this stage through calibration, reconstruction, and analysis is performed by the same algorithms that process the real data. The reconstruction algorithms make extensive use of the RECPACK [90] toolkit.

Versions of the ND280 software are tied to 'Productions'. A Production encompasses all the information, including a specific version of the software, needed to simulate ND280 data and process the real data (a 'Real Data Production', RDP). At the time of writing, Production 6 is used for the main analyses and Production 7 is near-finalised. Production 7 will include a significant number of simulation, calibration, and reconstruction improvements over Production 6. The additions to the ECal reconstruction software, which will be the focus of the next chapter, motivated the use of the candidate Production 7 software and calibrations for this analysis. As no new features are planned before the release of Production 7, this constitutes the most up-to-date ND280 calibrations and reconstruction algorithms, which will form the basis for future analysis. Unfortunately, it was not feasible to run the simulation from the event generation stage. The analysis presented herein uses Production 7 software from the electronics simulation stage onwards, which notably incorporates a much improved dark noise simulation for the ND280 ECal. This means that NEUT v5.3.2 and Geant4 9.4 were used, instead of Geant4 10.0 and the most recent version of NEUT, neither of these changes significantly impact the interpretations of the data.

2.6 Motivation for an ECal-target sample

This thesis presents a selection of beam neutrino interactions occurring within the ND280 Barrel ECal. The main motivation for such a sample is three-fold: to provide a high-statistics neutrino flux constraint across a wider range of off-axis angle than is possible in the FGDs; to investigate the rate of muon production outside of the angular acceptance of the FGD, which may provide insight into nuclear effects; and to measure lead-target neutrino interactions, a heavy nuclear target on which there is very little published data. The geometry of the ND280 Tracker limits the polar angular acceptance for charged particles produced in neutrino interactions occurring within the FGD to less than about 50 degrees (Analysis II in Ref. [75]). The polar angle herein is defined with respect to the axis of symmetry of the Tracker, which is in line with the neutrino beam axis to within 1.7 degrees [20]. A muon neutrino charged-current interaction sample that uses the Barrel ECal as a target may provide a complementary investigation of 'high-angle' interactions, for which $50^\circ \lesssim \theta_\mu \lesssim 130^\circ$.

One of the proposed near-detector designs for future oscillation experiments in the J-PARC neutrino beam, νPRISM, aims to use measurements taken at different off-axis angles to mitigate the systematic error induced by the use of incorrect neutrino interaction models [91]. An arbitrary effective neutrino flux can be constructed after the data have been recorded by using a linear combination of separate measurements taken at different off-axis angles. An analysis of νPRISM data would be able to mock up a measurement using the predicted oscillated flux at the far detector and compare observed event rates without a large reliance on the neutrino interaction model. An example of such an effective flux is shown in Figure 2.9a. The correct combination of separate measurements needed to form an effective flux relies on an accurate understanding of the true neutrino flux. The off-axis flux prediction across the whole of the face of ND280 is shown in Figure 2.9b. The Barrel ECal spans a significantly larger range of off-axis angle than the FGD alone—currently the only target used when constraining the flux prediction for oscillation studies. The observed muon neutrino interaction rate per target mass in the Barrel ECal is shown in Figure 2.10: the measured rate can be seen to vary as a function of the off-axis angle. A high statistics measurement over a range of off-axis angles may provide



(a) (b)

Figure 2.9: (a) A comparison of the oscillated J-PARC muon neutrino spectrum predicted at Super-Kamiokande and a linear combination of simulated measurements taken at different off-axis angles. The νPRISM experiment will be able to modify the effective flux shape during analysis. (b) The neutrino energy spectrum as a function of the off-axis angles that the near detector is exposed to, calculated from the full J-PARC neutrino beam-line simulation.

a constraint of the off-axis flux that could improve the flux model for future analyses, including those made by νPRISM.

The off-axis near detector at T2K is planned to be upgraded for an extended T2K data-taking period, termed T2K-II [17]. One proposed design re-orientates the FGD and TPC sub-detectors with a focus on enabling high-angle measurements from hydrocarbon-target interactions. Measurements of high-angle final-state particles from FGD-target interactions are not possible with the current ND280 geometry (Figure 2.3) as high-angle particles will not travel through the Time Projection Chamber (TPC). However, high-angle measurements, which utilise the TPCs for precise momentum measurement and particle identification, are already possible using the Barrel ECal as an active target. Such measurements would provide an interesting constraint on high-angle muon kinematics from interactions with a combined lead-hydrocarbon target.

The study of neutrino interactions is interesting for its own sake. The sum total of published lead-target neutrino-scattering data is shown in Figure 2.11[3]. Both results appear to show that the simulation is able to predict the data to within the quoted uncertainties. However, the CHORUS data are expressed relative to some unknown neutrino flux constant, C, and the

[3] At the time of writing, a very interesting new analysis that includes a number of kinematic projections of a new MINERνA lead-target sample has just been submitted for publication. An eprint can be found in Ref. [92]

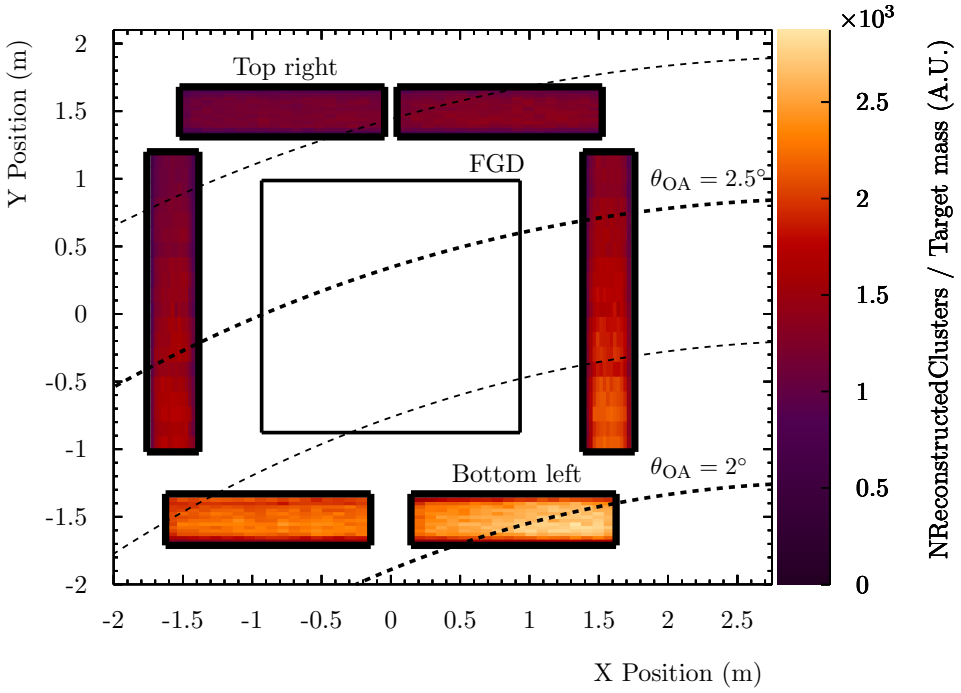
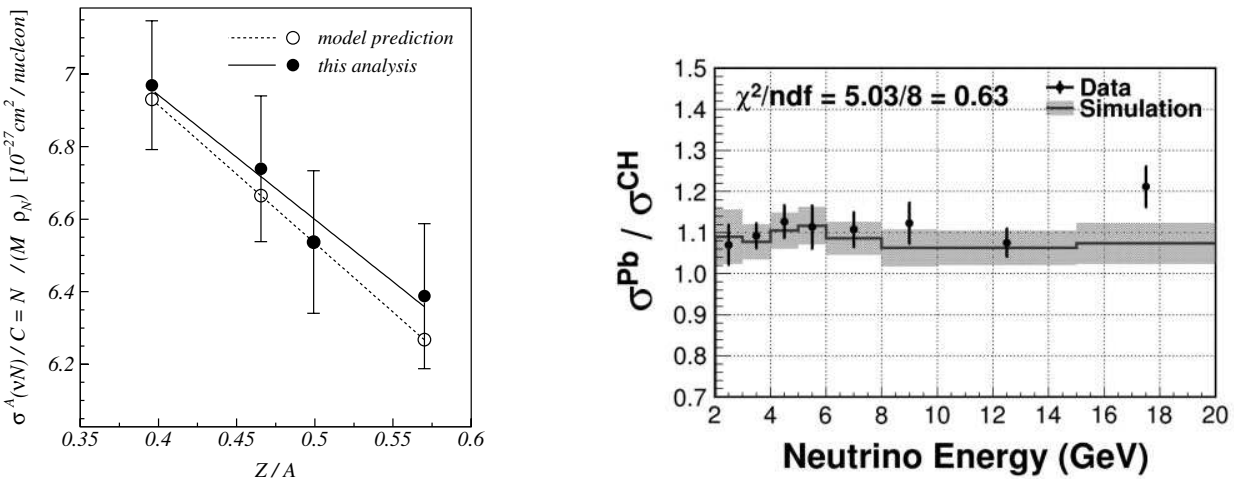


Figure 2.10: The number of deposits observed in each module of the ND280 Barrel ECal, normalised to the detector mass contained within each bin. Contours of constant off-axis angle are overlaid. Significantly more deposits are seen in the bottom left module, which is exposed to the smallest off-axis angle.

MINERνA data are presented as a ratio with carbon-target measurements, which may contain distinct model biases. Furthermore, the MINERνA data are presented as a function of reconstructed neutrino energy, and as briefly discussed in § 1.3.3, such a mapping is detrimentally model-dependent. It would be more interesting to investigate lead-target neutrino scattering data in distributions of interaction model-independent observables. The ND280 detector contains a number of nuclear targets that can be used, in a similar way to the CHORUS result, to investigate the scaling of nuclear-target effects with nucleon number. A robust lead-target analysis is needed if ND280 is to be used to its full physics potential.

2.6.1 Previous ND280 ECal measurements

Interactions occurring within the ND280 ECal have been assessed twice before. The first analysis, performed in 2013, attempted to fit for the flux-averaged, charged-current inclusive lead-target cross section by selecting a sample of neutrino interaction candidates reconstructed to start within the ND280 ECal [95]. The fit varied the overall cross section normalisation of



(a) CHORUS (b) MINERνA

Figure 2.11: Recent lead-target measurements from the CHORUS [93] and MINERνA [94] experiments. CHORUS measured the total charged-current cross section relative to some unknown constant, C, for lead ($Z/A \simeq 0.39$) and other targets. The MINERνA analysis presents the ratio of extracted lead-target and carbon-target cross sections per nucleon. Both measurements are relative to a carbon-target sample.

a lead, a carbon, and a background component, forming a test statistic between the predicted and observed event rate in each ECal module. An FGD charged-current inclusive selection was used as a separate constraint on the normalisation of the carbon cross section. Unfortunately, the fit favoured an unphysical negative cross section for the lead-target component of the event sample (§ 5.2.4)[95]; this has been attributed to a significant, unresolved data–simulation discrepancy. One significant source of disagreement arose because no magnetic field was simulated in the flux return yoke of the ND280 magnet [96]. In the actual detector, more negatively charged particles travelling through the flux return yoke were directed into (out of) the bottom (top) ECal modules. This mis-modelling has since been rectified (§ 6)[96].

The second analysis for neutrino interactions in the ECAL, performed in 2015, involved the implementation of new vertex reconstruction algorithms (§ 5)[96], which will be discussed in §3. Further undetermined disagreement between simulation and data produced another unsatisfactory result.

The analysis selected interaction candidates using a fiducial volume, all deposits that were reconstructed as starting outside of this volume were discarded. A diagram of the fiducial volumes used to select one-track and multi-track interaction candidates is shown in Figure 2.12.

The active volume of a Barrel ECal module is defined as the bounding box which contains all of the scintillator bars and lead absorber sheets in a module. Selected interaction candidates were then sorted into 'signal' and 'reverse', or background-enhanced, samples for each ECal module by a number of other cuts which are detailed in § 7 of Ref. [96]. A similar fit to that implemented in the 2013 analysis was used to extract the overall normalisations of three templates: carbon-target, lead-target, and 'background'. The pre- and post-fit selected event rates are shown in Figure 2.13. An evaluation of the relevant shape-only systematic uncertainties was encoded into a covariance matrix that was used to calculate a test statistic (Eqn. 8.1)[96]—the corresponding correlation matrix is shown in Figure 2.14a. The selected and reverse samples can be seen to be strongly anti-correlated, which is expected from the shape-only treatment. Any event property variations caused by the systematic assessment that resulted in candidates migrating across a cut threshold will have moved events from the reverse sample to the corresponding selected sample bin, or *vice versa*. The shape of the background template and the significant number of carbon-target and lead-target events selected into the background-enhanced sample meant that none of the templates could be varied while not incurring a significant penalty due to the anti-correlations evident in the covariance matrix. As highlighted in Figure 2.14b, the reverse sample contained 'background' events originating from a wide variety of sources: interactions in the magnet, interactions outside of ND280 ('sand events'), interactions in the Tracker or in ECal modules other than the one that the candidate was reconstructed in. Each of these sources may have been mis-modelled in understandable ways but by controlling the sum total with a single normalisation parameter, most of the power to resolve any simulation–data differences was lost. Furthermore, the events that contribute to the background template in the ECal and FGD samples were likely to be very different, while the systematic uncertainty treatment aimed to account for this, the uniform treatment of the 'background' normalisation parameter appears highly susceptible to modelling deficiencies in the outer detector.

While it was not determined what exactly led to the poor fit results, data–simulation discrepancies that differ from ECal module to module were apparent in some of the distributions used in the selection. For example, the number of detector hits within a reconstructed deposit (§ 8)[96], reproduced in Figure 2.15a, shows a data excess of deposits with a low number of

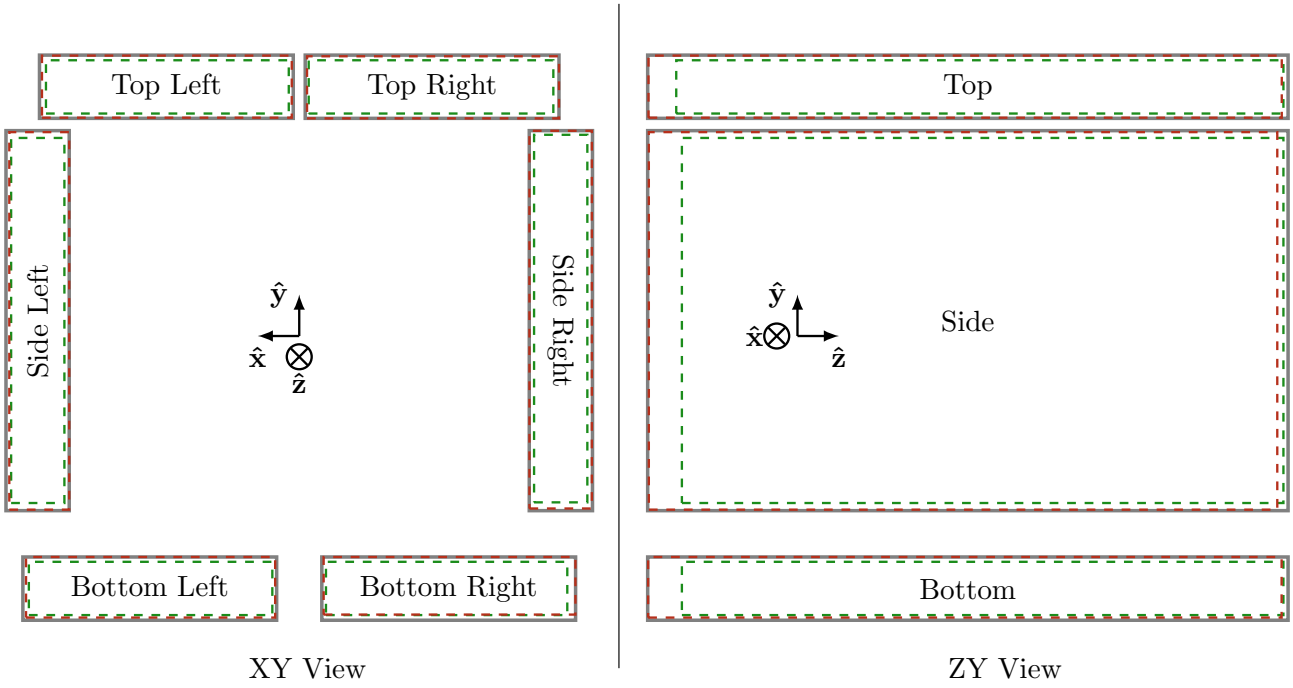
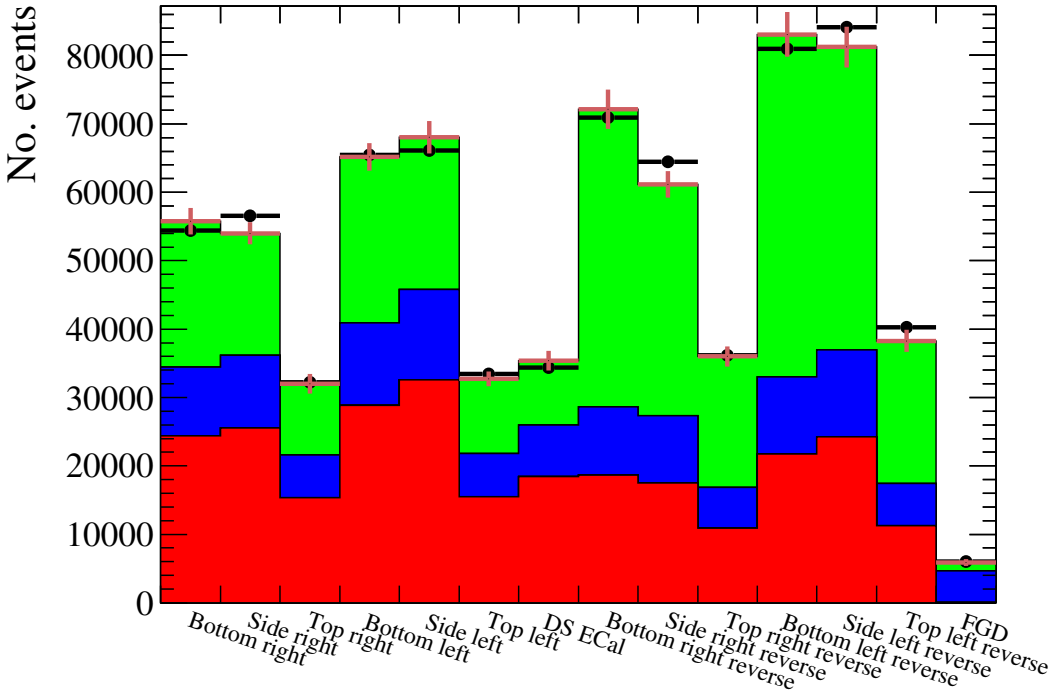


Figure 2.12: The ECal fiducial volume definitions used in the 2015 analysis (§ 7.6.1.1)[96]. The solid grey regions show two-dimensional projections of the active Barrel ECal volumes and the dashed green (red) lines show the fiducial volume used to select one-track (multi-track) vertex candidates.

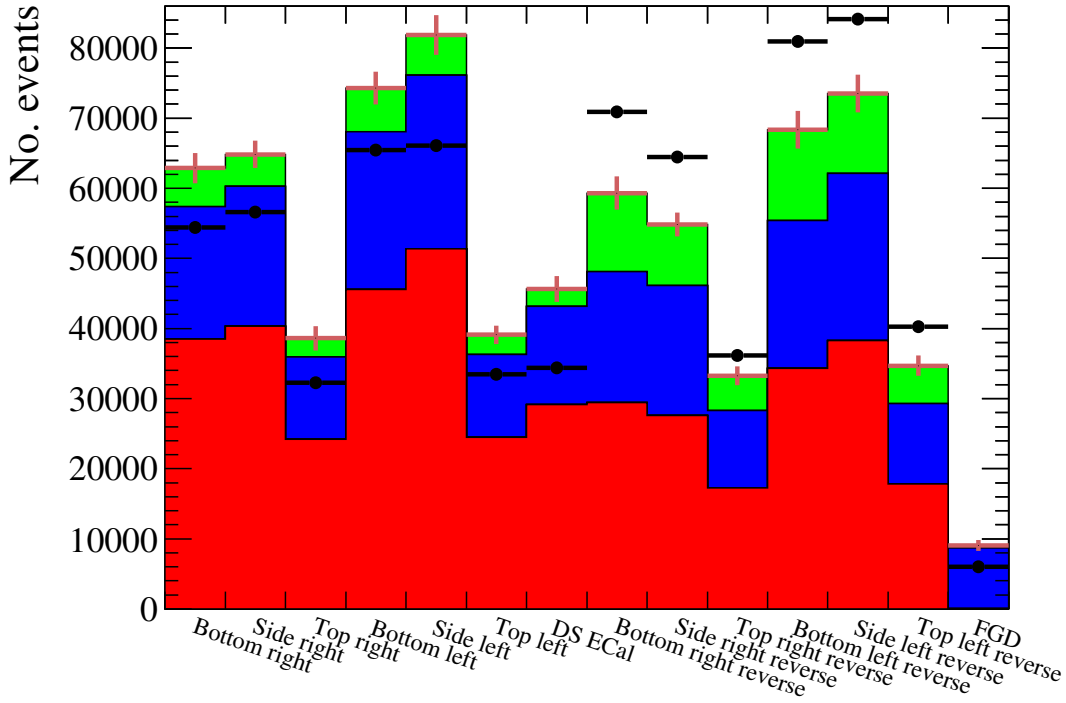
hits. A second peak at $\simeq$ 40 hits—which corresponds to through-going particles which enter the detector approximately perpendicular to the layer orientation and only activate a few bars per layer—is more prominent in the simulation than the data. The cut thresholds used in § 7 of Ref. [96] to select charged-current inclusive interaction candidates were tuned to efficiently retain a high-purity sample of true charged-current ECal-target interactions in the simulated data set. This resulted in cuts placed through regions of data–simulation disagreement, which can be seen in the two-dimensional cut on the cluster charge[4] and number of hits in a cluster—shown in Figure 2.15b—that was tuned to select a region of low background to signal ratio. The selection of events contributing to this region of disagreement reduces the validity of the systematic uncertainty estimation.

In the analysis presented herein, the strategy was reassessed in an attempt to constrain and mitigate the many potential sources of mis-modelling that plagued these previous analyses. The data analysed in the 2015 analysis were taken at the end of Run 3 (*c.f.* Figure 2.2a) between 2012-05-27 and 2012-06-09; in this analysis, those data are used to validate and cross-check the

[4]It can be seen in Figure A.4 that similar distributions of track charge are still poorly simulated.

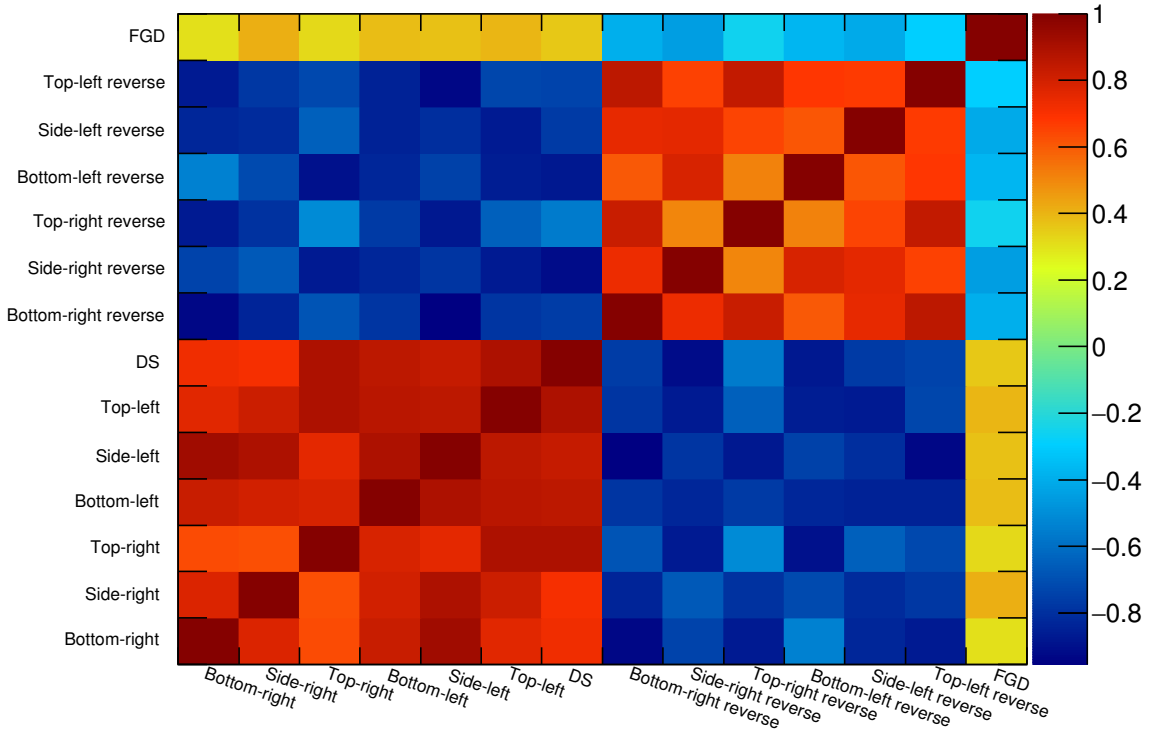


(a) Pre-fit

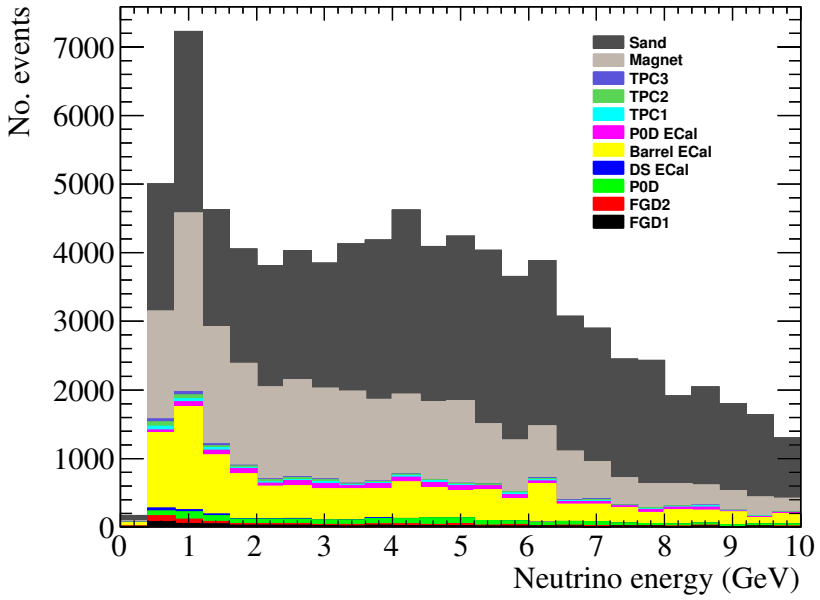


(b) Post-fit

Figure 2.13: The pre- (a) and post-fit (b) selected event rates per ECal module from the 2015 analysis [96, Figure 8.41, Figure 8.42]. The black markers show the observed event rate in the data and the red crosses show the total event rate and associated systematic uncertainty in each sample. The pre-fit appears reasonable by eye, but when the covariance matrix is included in the test statistic evaluation, the normalisation of the lead (red), carbon (blue), and background (green) contributions cannot be satisfactorily fit to the observed data.

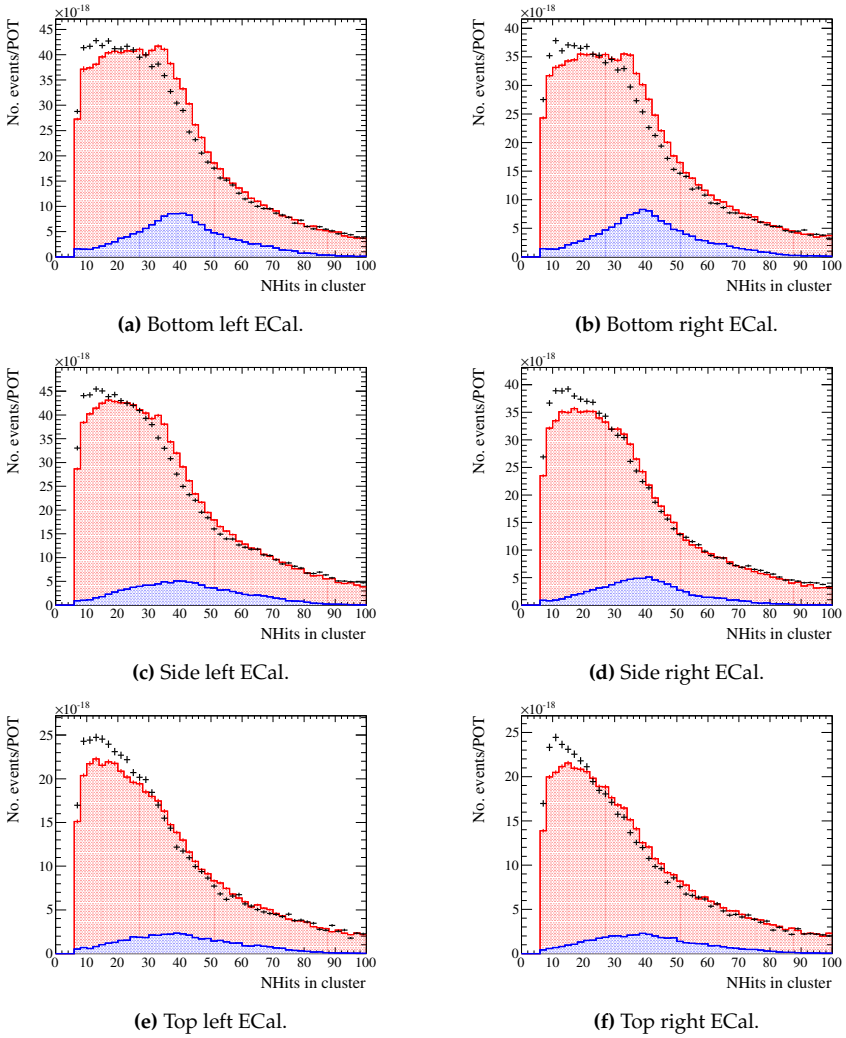


(a)



(b)

Figure 2.14: (a) The pre-fit correlation matrix which encodes the effects of systematic uncertainties on the ECal selection [96, Figure 8.44]. The uncertainty assessment found that the selected and reverse samples in a given ECal module are strongly anti-correlated, while selected samples from different ECal modules are strongly correlated with each other, similarly for the reverse sample. (b) The background template for the reverse bottom left Barrel ECal subdivided by true neutrino interaction position [96, Figure 8.48]. 'Sand' refers to interactions occurring outside ND280 and 'Magnet' refers to interactions occurring in the iron flux return yokes of the magnet.



(a) Bottom left ECal.

(b) Bottom right ECal.

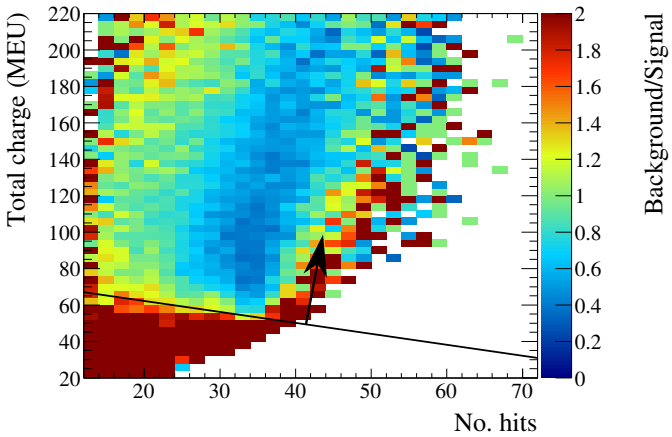
(c) Side left ECal.

(d) Side right ECal.

(e) Top left ECal.

(f) Top right ECal.

(a)



(b)

Figure 2.15: (a) The distribution of the number of detector hits in each reconstructed deposit shows significant data–simulation disagreement [96, Figure. 8.]. The shape of the disagreement differs between ECal modules. The observed distribution is shown by the black markers and the simulated distribution is subdivided into interactions occurring within ND280 (red) and interactions occurring outside of ND280 (blue). Only events which passed the fiducial volume cut, shown in Figure 2.12, were included. (b) The tuning of the two-dimensional track charge and hit multiplicity cut to maximise the true signal selection resulted in a threshold that passed through a region of poorly simulated phase space—as seen in (a). Events above the black line passed this cut.

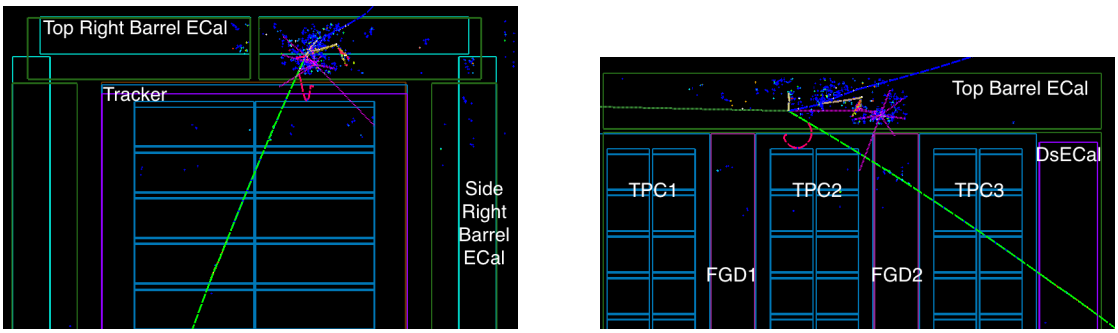
reconstruction and selection. To disambiguate this data set from those introduced in § 6, they are referred to as the 'Run3Cb' data set.

Chapter 3

ND280 Reconstruction

When a neutrino interaction occurs inside ND280, charged final state particles may deposit energy in any of the various sub-detectors that make up the detector. Figure 3.1 shows a simulated interaction in one of the Barrel ECal modules. Any interesting properties of such charged particles need to be reconstructed from the deposits that are recorded by the data acquisition system. During the first stage of reconstruction, sub-detector 'isolated' reconstruction is performed on deposits within each detector module—these may be fully contained deposits or segments of a longer track. These detector-isolated results are then combined by the ND280-wide reconstruction algorithm, or *global reconstruction*, to better use the measurement strengths of each sub-detector.

The developments described in this chapter concern improvements made to the ECal-isolated vertex reconstruction and global track matching and vertexing algorithms. These improvements were made to facilitate a more robust ECal-target interaction candidate selection that could make use of ND280-wide reconstruction information. The key result of the developments made is that a charged particle produced in a neutrino interaction within an ECal module can now be followed through the rest of the detector. This allows precise momentum measurements and PID estimates to be associated to the reconstructed final state tracks of ECal-target interaction candidates. The ECal-isolated vertex reconstruction validation (§ 3.1.4), the TREx high-angle reconstruction validation (§ 3.2.1), the inclusion of the ECal vertex reconstruction into global



(a) Back view (b) Right view

Figure 3.1: The simulated energy deposit from an ECal interaction occurring in the Top Right Barrel ECal. The neutrino (dotted green) can be seen arriving from the left and interaction within the Barrel ECal active volume. A number of final state particles can be seen leaving energy deposits (coloured crosses) in the detector. A final state muon (green dashed) travels through the Tracker and the Downstream ECal.

reconstruction (§ 3.3), and the validations of the ECal–tracker matching and global track parameter reconstruction (§ 3.2.2 and § 3.3) are all original work. The ECal-isolated vertexing was developed for the 2015 ECal-target analysis and the global reconstruction algorithms were developed as the final stage of the standard ND280 reconstruction software, which was originally used for oscillation studies.

The ECal-isolated vertexing algorithms developed by Brailsford [96] were ECal-only reconstruction algorithms—*i.e.* the reconstructed interaction candidates were not able to be used by the ND280-wide reconstruction[1]. This limited the information available for analysis.

For measurements of interactions occurring within the fiducial volume and kinematic acceptance of the ND280 Tracker, reconstruction and selection efficiency effects due to the position within a detector, azimuthal angle, and occasionally polar angle may be integrated over. This is justifiable when the efficiency is flat and well-understood in these projections, the neutrino interaction model is well constrained, and the neutrino flux can be treated as homogenous across the fiducial volume; in the Barrel ECal, however, none of these are defensible assumptions. To be confident in the robustness of the sample presented here, the efficiency for each stage of the reconstruction should be understood in each relevant part of phase space. This would be of critical importance were a neutrino-scattering cross section measurement to be performed using

[1] the ECal vertex reconstruction algorithm is not the 'mainstream' ECal reconstruction algorithm, which is based on track–shower discrimination and is fully integrated with the rest of ND280 reconstruction

the ECal as an active target. Throughout, reconstruction efficiencies are presented in multiple kinematic projections with care taken to reduce effects from integrating over other dimensions of phase space.

3.1 ECal reconstruction

The ECal-isolated vertexing algorithms make use of the Hough Transform [97] to find straight, track-like features in amongst clusters of detector hits. A brief overview of the ECal-isolated vertexing algorithm is given in this section, followed by validations of the reconstruction performance. The reconstruction algorithm is summarised in Figure 3.2. The ECal-isolated vertexing algorithms and any reconstruction thresholds determined in Ref. [96] have been kept as-is.

3.1.1 Standard reconstruction

The original ECal reconstruction algorithms were designed to maximise the discrimination power between showering and track-like particles—specifically for the reconstruction of neutral pions and to provide robust muon–electron separation [67]. The algorithms were developed in the expectation that they would be used to reconstruct deposits left by a single, entering particle. The reconstruction starts by forming two-dimensional clusters of calibrated detector hits, firstly by a simple nearest neighbour search and then by conditionally combining the nearest neighbour clusters based on similarity in charge-weighted position, time, and cluster shape [96, §4.4]. Hits can only be added to a cluster if they occurred within 50 ns of at least one clustered hit. Any isolated hits that appear to be related to the combined clusters in hit time or position can also be absorbed. The clusters from each two-dimensional view are then matched between the views to form three-dimensional clusters of hits. This matching calculates a matching likelihood between each cluster from one view and every cluster in the other view. The most likely match above some minimum likelihood threshold is made and the matched clusters are ignored for future iterations; the process is repeated until no more matches can be made. The likelihood is built from distributions of ‘correct’ and ‘incorrect’ matches in three

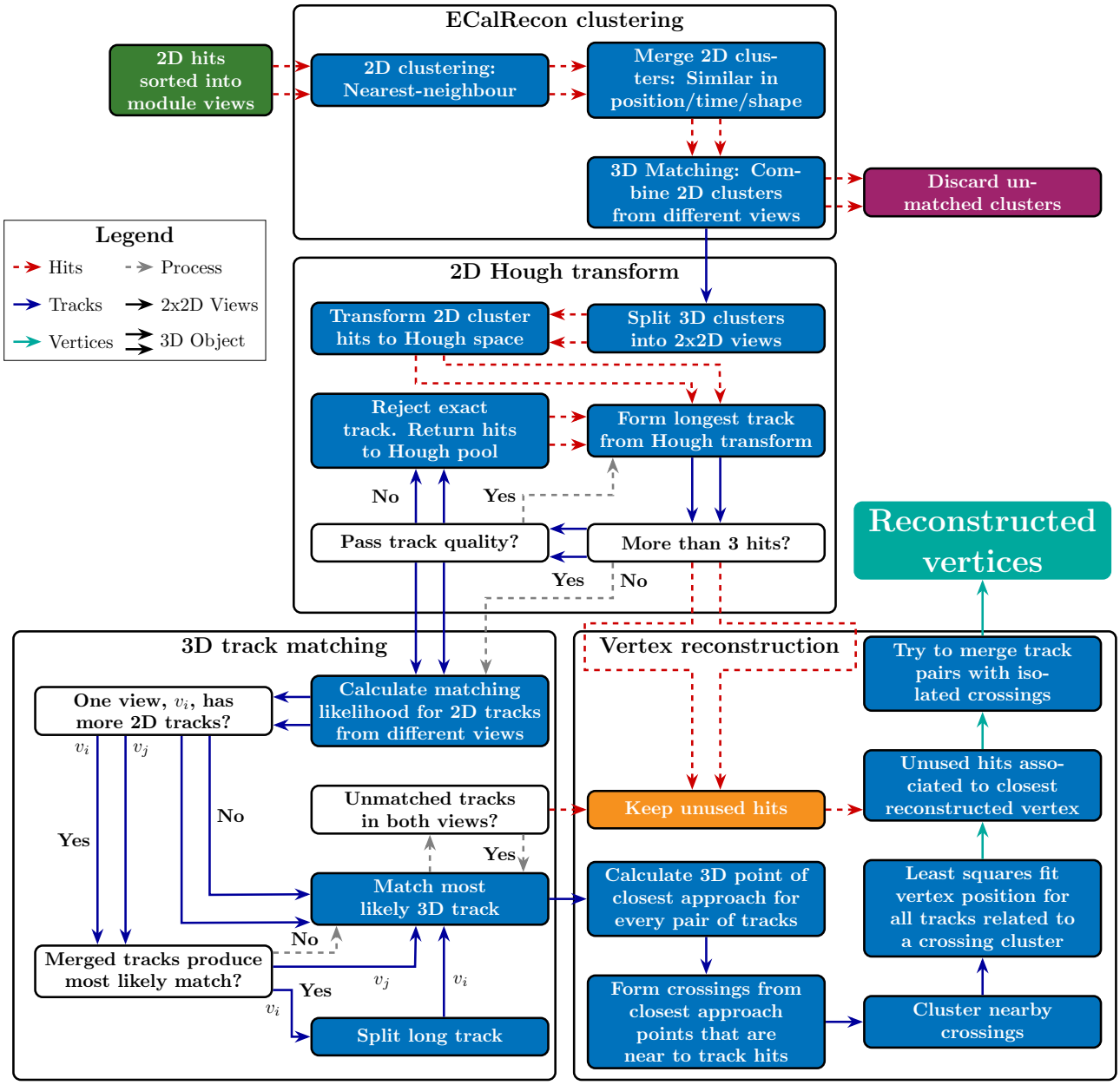


Figure 3.2: A summary flow chart that depicts how ECal scintillator hits are reconstructed as ECal-isolated vertexing algorithm used to reconstructed neutrino interactions that occurred within an ECal module.

observables; the distributions were determined from an analysis of simulated neutrino beam data. The three observables are: the ratio of total hit charge in each view and the separation of the 'top' and 'bottom' of each two-dimensional cluster in the coordinate shared between the two views—the *running coordinate*, which corresponds to the layer number of the hit bars. The expectation is that if two clusters describe two views of the same physical process, each will have a similar extent in the shared coordinate and each will have clustered approximately half of the visible energy. The results of the matching are three-dimensional clusters of distinct detector hits—each hit can only be used by a single cluster.

The standard algorithms further process these three-dimensional clusters to reconstruct the total electromagnetic energy deposited from the visible charge distribution and calculate PID discriminators. These discriminators can be used to separate shower-like deposits (left by electrons, photons, and neutral pions) from track-like deposits left by minimally-ionising particles (muons and high-momentum charged pions) and highly-ionising particles (protons). The results of these reconstruction stages, while important for the other uses of ECal information in ND280 analysis, are not used herein and so will not be discussed in more detail here. Interested readers can find more information in Ref. [67].

3.1.2 Vertex track-finding

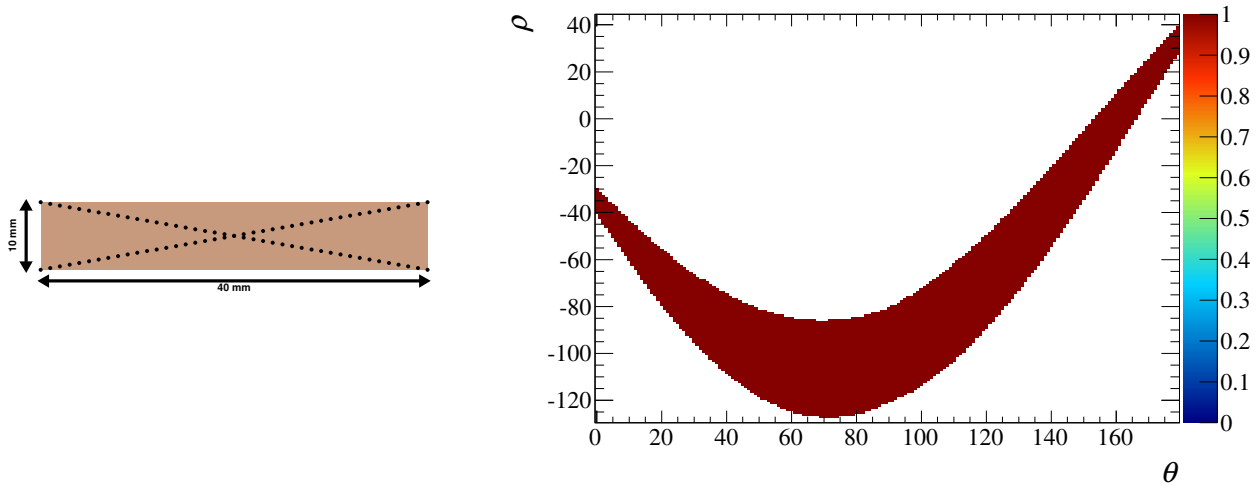
The ECal-isolated vertexing post-processes the three-dimensional clusters produced by the standard ECal reconstruction. Each cluster has each of the two views processed by the two-dimensional Hough transform, which transforms the distribution of hit positions to a distribution of straight lines that pass through each hit. In this distribution, track candidates that contain many of the hits from the reconstructed cluster appear as peaks.

The Hough lines are parameterised by an angle, θ, and a displacement of closest approach from the origin, ρ, with respect to some two-dimensional cartesian coordinate system. As each cluster is transformed entirely independently of the other clusters, the origin is taken to be the charge-weighted mean hit position in each view—this choice has no impact other than keeping the range of ρ reasonably small. The relationship between cartesian points on the line, (x, y),

and the track parameters, (ρ, θ), is given by $\rho = x\cos\theta - y\sin\theta$. This parameterisation is used to avoid a problem for tracks perpendicular to the $\hat{x}$ axis which result in a singularity for the standard cartesian line equation, $y = mx + c$. For each point in position space through which a track could have passed, the track parameters of all straight lines which pass through the point are added to the Hough-space distribution. For a cluster of detector hits, any charged tracks that might be reconstructed must have only passed through bars that were 'hit'[2] and not bars that were not. The ECal scintillator bars are far from pixel-like detection elements—candidate tracks could have passed through them in many ways that would have caused them to register a hit. It was noted in [96, §5.2.1] that the most reduced set of points through which any infinite Hough line through a bar must pass is given by placing the points on the diagonals of the bar cross section as shown in Figure 3.3a. An example Hough transform of a single ECal bar is shown in Figure 3.3b. For each hit in a cluster, the transformed contributions to the track parameter space are summed together to form the whole cluster (ρ, θ) distribution. An example Hough transform of a simple cluster is shown in Figure 3.4. The parameters of tracks that pass through hits left by the simulated muon appear as a peak in the Hough parameter space—this can be seen at $\rho \simeq 20\,\text{mm}$ and $\theta \simeq 170°$. In this way, analysis of the Hough parameter space highlights the track parameters of all straight trajectories that may have contributed to a cluster.

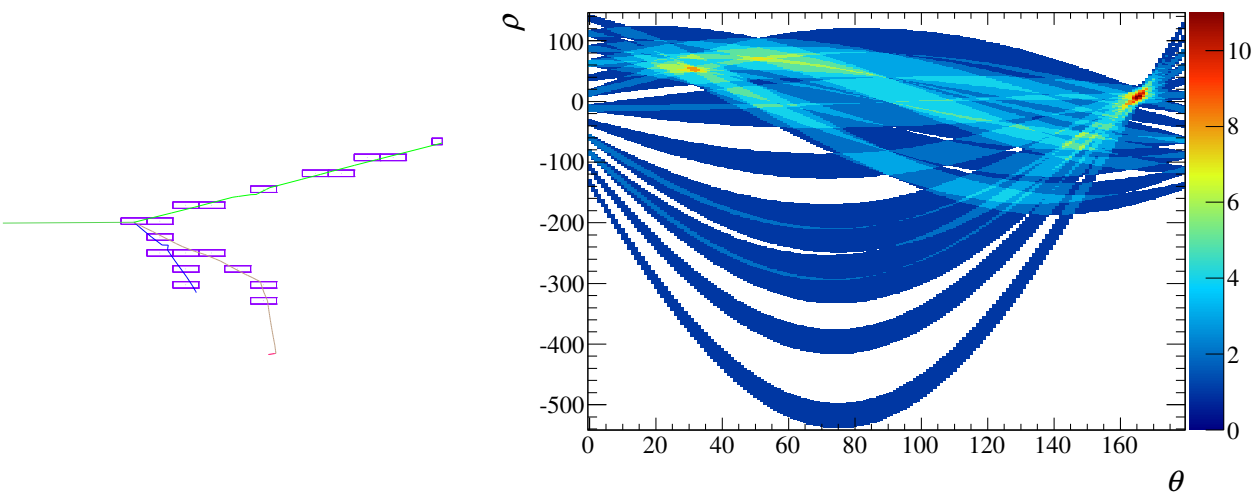
The two-dimensional track-finding proceeds iteratively by choosing the highest point in the Hough parameter space and forming a two-dimensional track from all contributing hits. This 'proto-track' is then checked for quality; if it passes, the contributions from the used hits are removed from the hough parameter space distribution and the process continues until no new tracks can be formed. A two-dimensional track candidate must use at least three hits—fewer than this provides too little information to have any degree of confidence that both views of the deposit were left by the same particle [96]. While the Hough transform picks out the infinite lines that pass through a representative set of points, the transform is not influenced by any missed hits that might be expected if a charged particle propagated along the candidate trajectory. To enforce the construction of contiguous candidate tracks, proffered Hough tracks

[2] or it may have passed through in-active material



(a) (b)

Figure 3.3: (a) Each ECal scintillator bar is represented as a cross of points to be Hough transformed. (b) The Hough transform of a single, activated bar in Hough track-parameter space. Figures from Ref. [96].



(a) (b)

Figure 3.4: (a) A two-dimensional event display of a simulated neutrino (green, enters left) interaction in the ECal. The simulated charged particles, and the scintillator bars that they deposit energy in, can be seen to the right of the interaction vertex (b) The Hough transform of the hit cluster in (a). The highest points in the two-dimensional parameter space describe infinite, straight 'tracks' which pass through the greatest numbers of hits in the cluster. Figures from Ref. [96].

are rejected if a single ECal layer is missed along a track—*i.e.* if contributing hits occur in only two out of three consecutive layers. Tracks are also rejected if more than one scintillator bar within a layer is missed—where the candidate track includes hits from the same layer, which are separated by two or more consecutive bars without hits. Hits from rejected tracks are not allowed to appear in exactly the same combination again but are returned to the track-finding pool to be used in the formation of other track candidates.

Once a set of two-dimensional tracks has been extracted from each view of the three-dimensional cluster, they are matched together to form three-dimensional tracks. The matching is performed in the same way and uses the same discriminating observables as the matching in the standard ECal reconstruction—the cluster charge ratio and difference in position in the shared coordinate. The probability distributions used to form the likelihoods were re-calculated using the results of the Hough track-finding. The matching uses different distributions depending on the number of reconstructed Hough tracks—one, two, or three or more tracks in each view [96, §5.2.5]. An extra complication arises for the matching of vertex tracks in that what appears as one long straight track in one view may be clearly two or more tracks in the other view. This can be resolved by checking whether pair-wise combinations of Hough tracks from the view with more tracks form more likely matches with long tracks in the other view. If the combined track results in a greater likelihood than any other matches, the crossing point of the combined tracks in the shared coordinate is used to split the long track into two sub-tracks. The four tracks are then included in the list of tracks for further matching. This process is described in detail in § 5.2.5 of Ref. [96]. Unlike the standard ECal reconstruction matching procedure, no minimum likelihood threshold is used, and so the matching continues until all the two-dimensional tracks are used from at least one of the views. This results in some very poor-quality matches that need to be removed before further processing. The three-dimensional tracks produced by the track-finding and matching processes then need to be combined into reconstructed vertex candidates.

3.1.3 Vertex candidate reconstruction

In the 2015 analysis, the three-dimensional Hough tracks were considered the final result of the reconstruction; the reconstruction of vertex candidates was part of the 'interaction candidate selection'. For this analysis, the vertex candidate reconstruction was ported back to the ECal-isolated reconstruction so that the vertex candidates could be used by ND280 global reconstruction. The steps and thresholds used to reconstruct ECal vertex candidates was motivated in §7.4 of Ref. [96] and have been left unchanged.

First, the three-dimensional Hough tracks are checked for quality and poor-quality tracks are rejected. Poor-quality tracks are those that only include hits from a single ECal layer in one of the views or those in which the two-dimensional tracks in the two views do not overlap in the shared coordinate. Surviving tracks are then clustered to form vertex candidates. A point of closest approach is calculated for each pair-wise combination of tracks. These points are compared with the positions of the detector hits that make up the two contributing tracks. If some of the hits are within 200 mm, then the crossing point is kept for further processing [96, §7.4.2, Table 7.2]. The crossing points are then clustered together if the distance between any two hit positions in a potential cluster is less than 140 mm [96, §7.4.2, Table 7.2]; the tracks associated with each of the clustered crossings are then used to form 'proto-vertex' candidates. The reconstructed position of each candidate is then taken as the point which minimises the sum squared distance of closest approach to each of the contributing tracks. It should be noted that there is no requirement that each track is only included in a single reconstructed vertex.

The final stage of the ECal-isolated vertex reconstruction attempts to merge three-dimensional tracks. Charged particles that travel a significant distance within a module will often curve due to the magnetic field inside the detector. However, as the Hough transform finds straight tracks in clusters of hits, long curving trajectories may be reconstructed as more than one three-dimensional track. The track merging tests all track pairs which have an isolated crossing point—*i.e.* the calculated track crossing point is more than 140 mm away from any other crossing point. The combination of track merging and the ability for tracks to be included in multiple vertex candidates allows the ECal-isolated vertexing to fully reconstruct neutrino

interactions which produce a multi-track deposit near the interaction point and a long curving deposit left by a propagating minimally ionising particle, such as a muon or a charged pion. This is a key topology of interest when searching for interactions occurring within an ECal module. Such deposits would most-likely have been mis-reconstructed as a single track or a large, distended 'shower' by the standard ECal reconstruction.

Each of the tracks associated with each of the reconstructed vertex candidates are then passed to global track reconstruction to be matched with reconstructed objects in other sub-detectors. While being matched by global reconstruction, the individual tracks are temporarily dissociated from their vertex candidates and treated as separate, independent tracks that are all matched with the equivalent priority. In the next section, the performance of the ECal-isolated vertexing is assessed and comparisons between data and simulation are presented and discussed.

3.1.4 Reconstruction performance

To fulfil the requirements of this analysis, the ECal-isolated vertex reconstruction needs to be able to accurately reconstruct neutrino interaction vertices and charged particles that deposit energy in an ECal module. The important reconstructed track properties are the module-exit direction and position, and the associated uncertainties. It is also important that the reconstructed interaction vertex position does not exhibit any significant bias.

The Inward Muon particle gun sample

It is often useful to be able to simulate particles with arbitrary distributions of starting position within the detector and initial kinematics; the General Particle Source (GPS) of the Geant4 framework, often referred to as a 'particle gun', performs this function. This technique can be used to investigate the reconstruction of important particle topologies in isolation, and also allows the simulation of topologies that are not predicted by any interaction model.

A sample of fake charged-current interaction vertices was used to assess the reconstruction performance for muons originating in a Barrel ECal module passing through the Tracker—

the main topology of interest for this analysis. The fake vertices were created by firing one or two particles from the same point within the detector geometry. In each ECal module, three samples of fake vertices were simulated: muon, muon + proton, and muon + charged pion. The muon initial-state kinematics were restricted so that the generated muons were fired approximately towards the Tracker. The protons (pions) were fired isotropically with kinetic energies uniformly in the range $[50, 250]$ MeV ($[0, 400]$ MeV). The protons and pions were included to account for the effect of multi-particle final states on the muon reconstruction performance. The fake interaction vertices were distributed uniformly over the active volume of each Barrel ECal module to preclude including regions of significantly different than average performance.

Track reconstruction

The muon tracking efficiency of the ECal vertex reconstruction is shown in Figures 3.5–3.7 as a function of the muon kinematics, p_μ, θ_μ, and ϕ_μ, which are all defined according to the ND280 coordinate system, *c.f.* § 2.6. These efficiencies were calculated with the Inward Muon particle gun sample. The fake vertex positions were selected to occur within the fiducial volume of each barrel module[3]. This selection was imposed to mitigate any phase space-sensitive edge effects where the muon left the ECal module without passing through enough scintillator bars to be reconstructed. The tracking efficiency can be seen to be very high across the momentum phase space sampled (Figure 3.5) and plateaus at $\sim 90\%$ for muons with momenta greater than $400\,\mathrm{MeV}/c$. For the polar and azimuthal projections, the momenta of included muons was restricted to $p_\mu > 400\,\mathrm{MeV}/c$; this is to mitigate effects from the shape of the low momentum efficiency. This restriction is used for Figure 3.6 and Figure 3.7—it is not carried through to the interaction candidate selection. In these projections, the tracking efficiency is near unity except in regions where poorer reconstruction is to be expected. Tracks that travel approximately in the plane of a layer will cross enough bars in both views for the reconstruction to perform well; this corresponds to tracks with $\cos(\theta) > 0.9$ or $\phi \simeq 0, 180, -180$ ($\phi \simeq 90, -90$) for tracks in the top or bottom (side) modules. In the most important regions of track parameter space,

[3]The fiducial volume is described in section 4.1.2

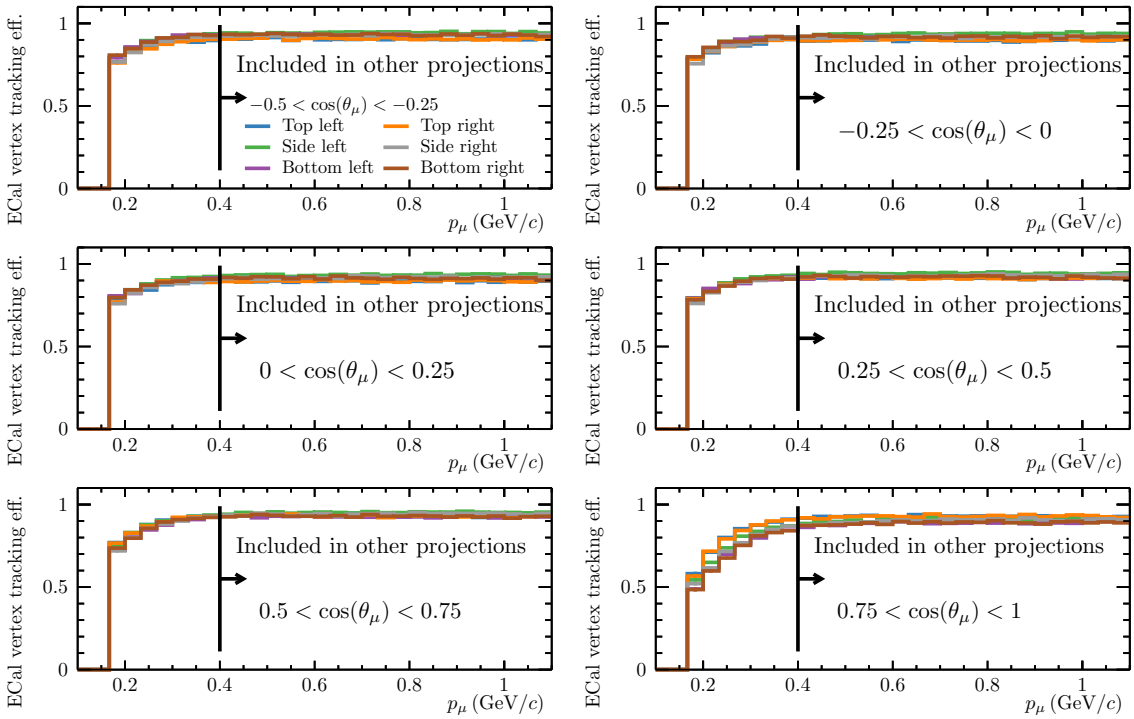


Figure 3.5: The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle momentum. The region $p_\mu < 400\,\mathrm{MeV}/c$ is excluded when projecting the efficiency onto the angular phase spaces. For higher momentum muons, the tracking efficiency is approximately flat.

the ECal tracking efficiency is more than sufficient for this analysis.

The reconstructed module-exit position, direction and associated uncertainties are the quantities used to construct a matching likelihood between ECal-isolated tracks and tracks from other sub-detectors. The global matching algorithm is described in detail in section 3.2.2. In each ECal module, one of the three coordinate axes is treated as the 'running coordinate' for the reconstructed track parameters. The running coordinate is used to parameterise the reconstructed direction: if x is taken as the running coordinate, then the direction vector is given by $(1, {}^{dy}/_{dx}, {}^{dz}/_{dx})$. The shared coordinate between neighbouring scintillator layers is the most natural choice for running coordinate. The reconstructed state vector for an ECal track in a side barrel module, which has the x axis as the running coordinate, is described by three position coordinates, two direction 'slopes', and an inverse momentum[4] estimate, $\{x, y, z, {}^{dy}/_{dx}, {}^{dz}/_{dx}, {}^{1}/_{p}\}$. The inverse momentum estimate is used during global reconstruction to take account of the helical trajectory followed by a charged particle propagating in an electric and magnetic field.

[4] the inverse momentum is proportional to the curvature in a magnetic field

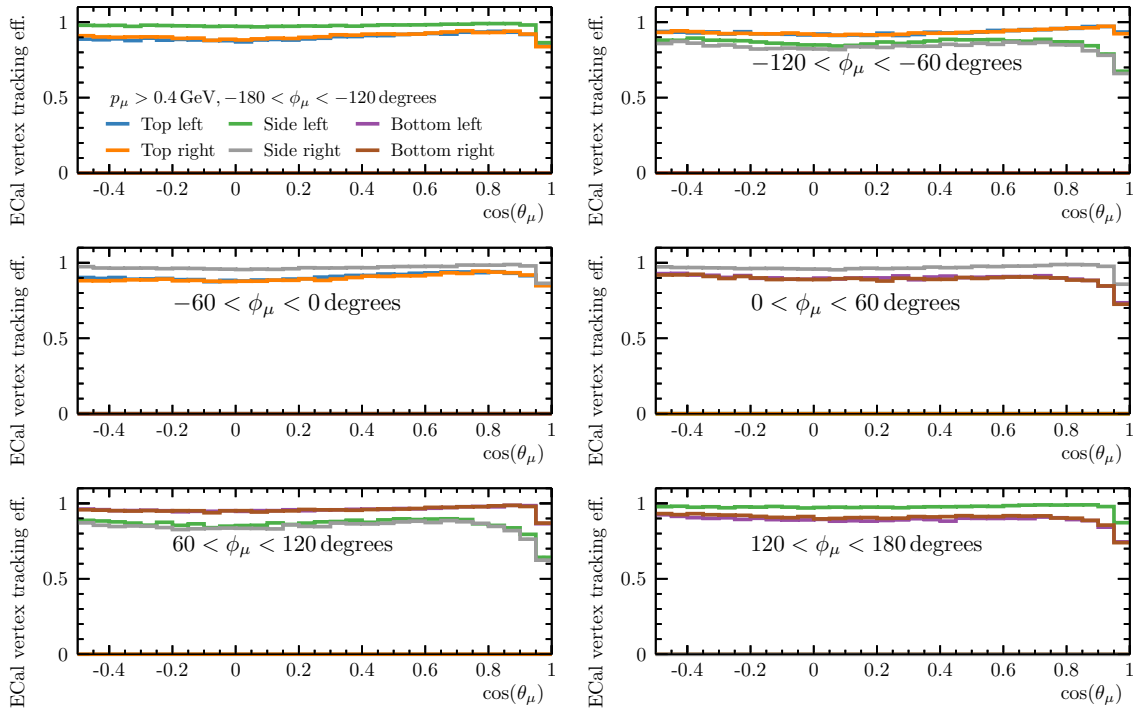


Figure 3.6: The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle polar angle cosine. The efficiency is flat for muons that do not travel approximately in the plane of a scintillator layer. The simulated muons shown in this projection are limited to $p_\mu > 400\,\text{MeV}/c$.

For tracks in the tracker, where precise momentum measurements are facilitated by the TPCs[5], this momentum estimate is used and updated by the reconstruction. For ECal-isolated reconstruction the inverse momentum is largely ignored and seeded with a value of $1\,\text{GeV}^{-1}\,c$. This estimate is only used when fitting long, merged Hough tracks where the trajectory curvature in the magnetic field is important. For tracks that are eventually merged with tracker components by global reconstruction, the whole track is re-fit, and this re-fit is seeded by the momentum state determined from the tracker component.

The important metric when assessing the suitability of the reconstructed properties for global matching are the track parameter 'pulls'

$$\varphi = \frac{X_{\text{reconstructed}} - X_{\text{true}}}{\sigma_{X,\text{reconstructed}}}, \tag{3.1}$$

where $\sigma_{X,\text{reconstructed}}$ is the reconstructed uncertainty in track parameter X. The Inward Muon

[5]TPC position resolution is $\sim 1\,\text{mm}$ [65, Figure 24–26] and momentum resolution is $\sim 0.1 \times p_\text{T}$ [65, Figure 27]

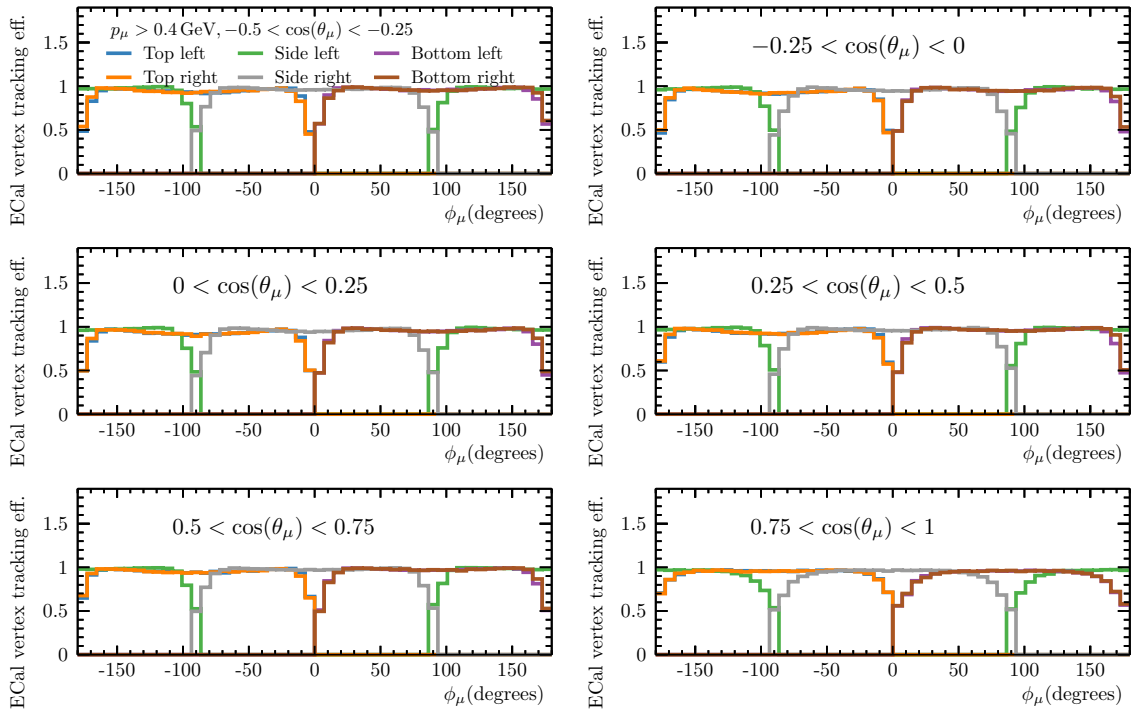


Figure 3.7: The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle azimuthal angle. The range of simulated muon angles is restricted for each module so that they propagate toward the tracker. The efficiency is flat for muons that do not travel approximately in the plane of a scintillator layer. The simulated muons shown in this projection are limited to $p_\mu > 400\,\text{MeV}/c$.

particle gun sample was used to assess the ECal-isolated vertex track reconstruction in terms of parameter pulls. To provide the most consistent comparison between reconstructed and true properties, the reconstructed state of the track at the end closest to the tracker is propagated along the running coordinate to a matching plane with the closest corresponding true state. In this way, the model used for the propagation of reconstructed states, *i.e.* the RECPACK model, is used to compare the reconstructed and true track parameters at an equivalent position. The Geant4 model used to propagate the true particles through the ND280 geometry is not available in the reconstruction framework and so could not be used for a fair comparison.

The pulls for the non-running position coordinates of the track-end closest to the tracker for the top and bottom—where $\hat{y}$ is the running coordinate—and the side barrel modules are shown in Figure 3.8. Pull distributions that follow the normal distribution highlight that the estimated parameter is unbiased and that the quoted uncertainties are neither over-, nor under-, estimated. It is clear from Figure 3.8 that the track end position reconstruction is unbiased. However, the shape of the distribution about $\varphi = 0$ shows that the uncertainty estimation provided by the reconstruction is not ideal. The shape of the peak signifies an over-conservative estimation of the uncertainty for some tracks, while the slow fall-off shows that for other tracks the uncertainty is significantly under-estimated.

The pulls for the relevant slopes for the top, bottom, and side barrel modules are shown in Figure 3.9. The direction pulls have similar features to the position pulls, however the plateau-like tail is even more prominent here, suggesting that significant under-estimation of the direction uncertainty is a systematic problem.

This situation is far from ideal and is partly a consequence of the generalised approach to reconstructed state determination offered by RECPACK. When fitting for the position, direction, and associated uncertainties given a set of hits, the ECal bars are modelled as a gaussian, rather than a top-hat function. This over-generalisation contributes to the situation shown in Figure 3.8 and Figure 3.9. A more detector-aware approach to uncertainty estimation may well provide a more accurate uncertainty estimation—the reconstruction should take account of the geometry of the ECal bars and layers. However, the discriminator thresholds used by global

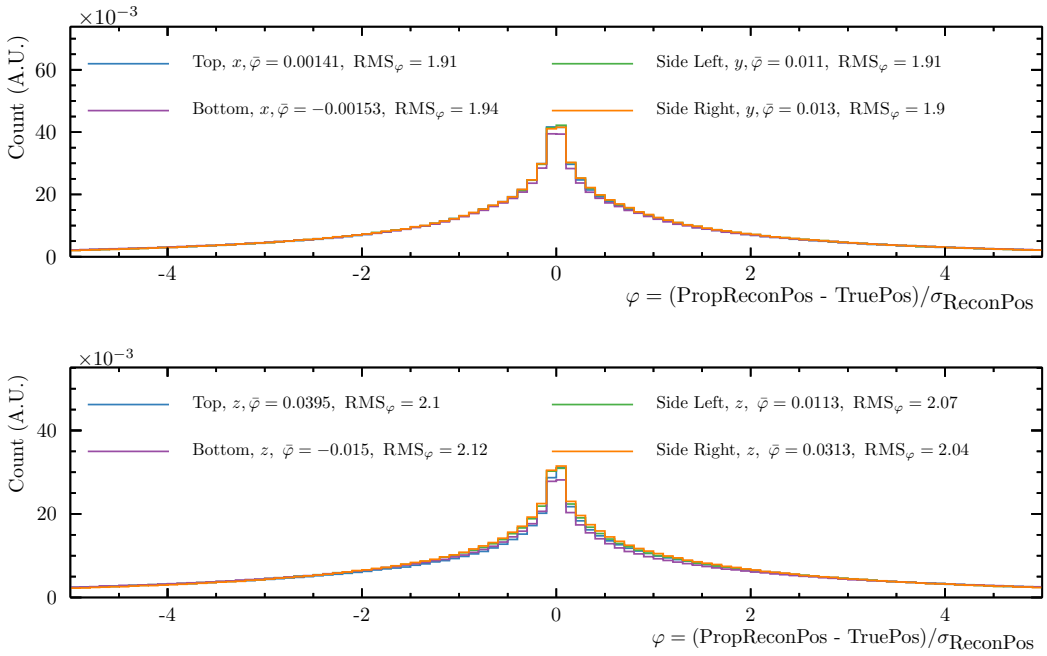


Figure 3.8: The distribution of track position pulls for the non-running coordinates for the top, bottom, and side Barrel ECal modules. The sharp peak and long tail highlight that for some tracks the uncertainty estimate is too conservative and for others too small, respectively.

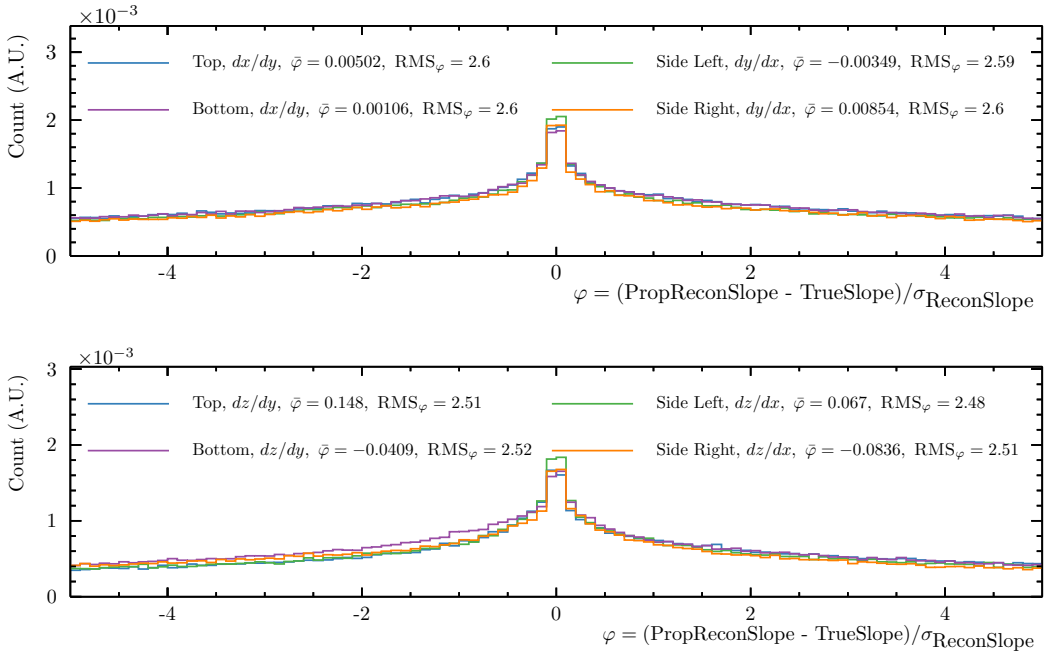


Figure 3.9: The distribution of relevant track slope pulls for the top, bottom, and side Barrel ECal modules. The sharp peak and long tail highlight that for some tracks the uncertainty estimate is too conservative and for others too small, respectively.

matching were tuned with such reconstruction in place and a robust re-tuning of the global reconstruction was outside the scope of this analysis. Reconstruction improvements should be a significant consideration for future work. This analysis used the track-state reconstruction as described above. The ECal-Tracker matching efficiency is assessed in § 3.2.2. The discrepancy between the apparent matching rate in data and simulation is assessed in § 5.3.2.

The next subsection focusses on the important components of the ECal-isolated vertex reconstruction.

Vertex reconstruction

In this analysis the fiducial volume of the ECal modules is re-assessed and designed to be conservative (*c.f.* § 4.1.2). Unless there is significant bias in the reconstructed positions of true ECal-target interactions, the net reconstructed vertex migration will be a negligible consideration. As the uncertainty on the vertex position is not used in this analysis, the only important quantity is the absolute reconstructed position bias. This is best shown by the distribution of position residuals,

$$r = X_{i,\text{recon}} - X_{i,\text{true}}, \tag{3.2}$$

for $i \in x, y, z$. The distribution of residuals for the top, bottom, and side Barrel modules is shown in Figure 3.10. These distributions were determined using the Run 3-equivalent simulation data set; only reconstructed vertex candidates where the true vertex occurred within the same Barrel ECal module were used.

It can be seen from Figure 3.10 that the reconstructed vertex position is not biased in the x or y directions, and any bias evident in the z direction is smaller than the 40 mm width of a single ECal bar. The spread of the reconstructed position residual is consistent in each module and prong multiplicity. Single track vertices will always be placed in the centre of a bar as the reconstructed position is placed at the upstream end of the track. For the running coordinate, this results in a residual distribution that is almost discretised; the majority of neutrino interactions occur within the lead absorber sheets in-between layers of scintillator. As

a result, the residual for these coordinates is not peaked about $r = 0$.

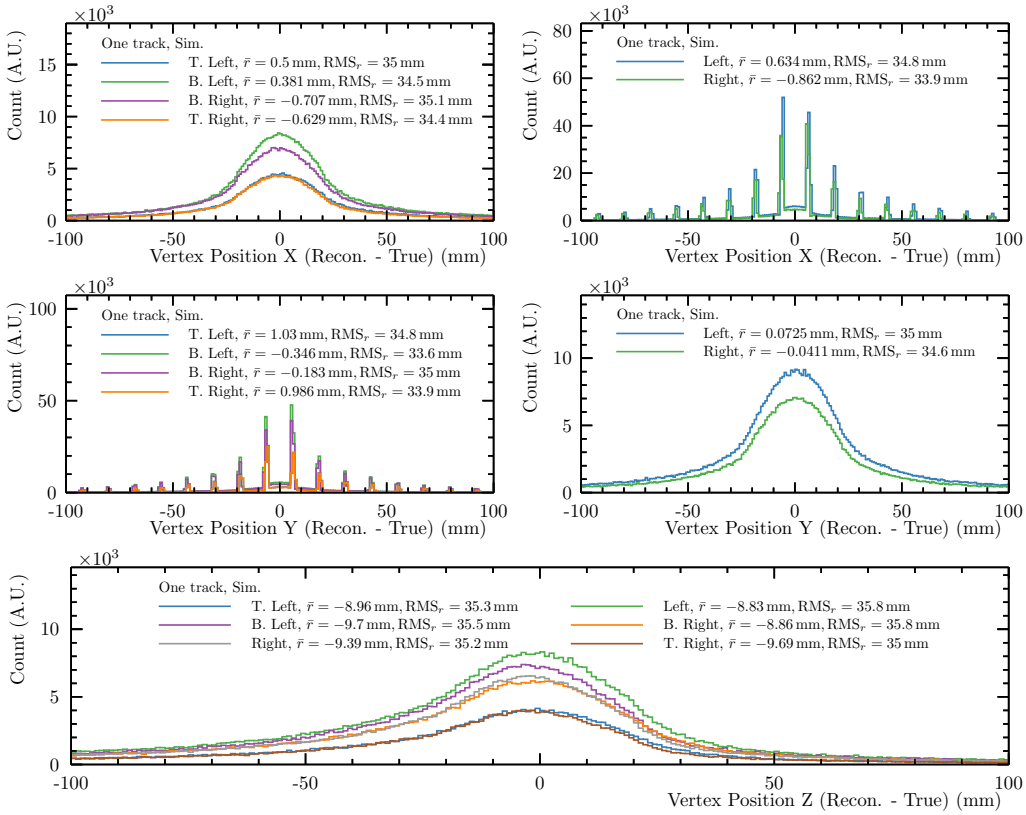
While the number of reconstructed tracks will not be a key component of the selection presented here, future analyses may want to try and create more granular, topology-based samples in which well-understood individual track reconstruction would be more important. The distribution of the true track topologies that contribute to a vertex candidate with a given number of reconstructed tracks is shown Figure 3.11. When combined with the distribution of reconstructed tracks per vertex shown in Figure 3.15, it is clear that most reconstructed vertex candidates underestimate the number of true tracks. This is to be expected from the ECal geometry and limitations in the reconstruction. The inability for one hit to be shared by multiple tracks and the minimum number of six hits per reconstructed track limits the reconstruction efficiency for short tracks from a vertex. The reconstruction necessarily prioritises creating long tracks. This is not a deficiency for this analysis where the most important reconstructed properties are the vertex position and the correct matching of an ECal track with a Tracker track. However, improved short track reconstruction efficiency may open up new uses for an ECal-target sample.

The next subsection contains comparisons between simulation and real data for a variety of reconstructed properties of reconstructed ECal-isolated vertex tracks and candidates.

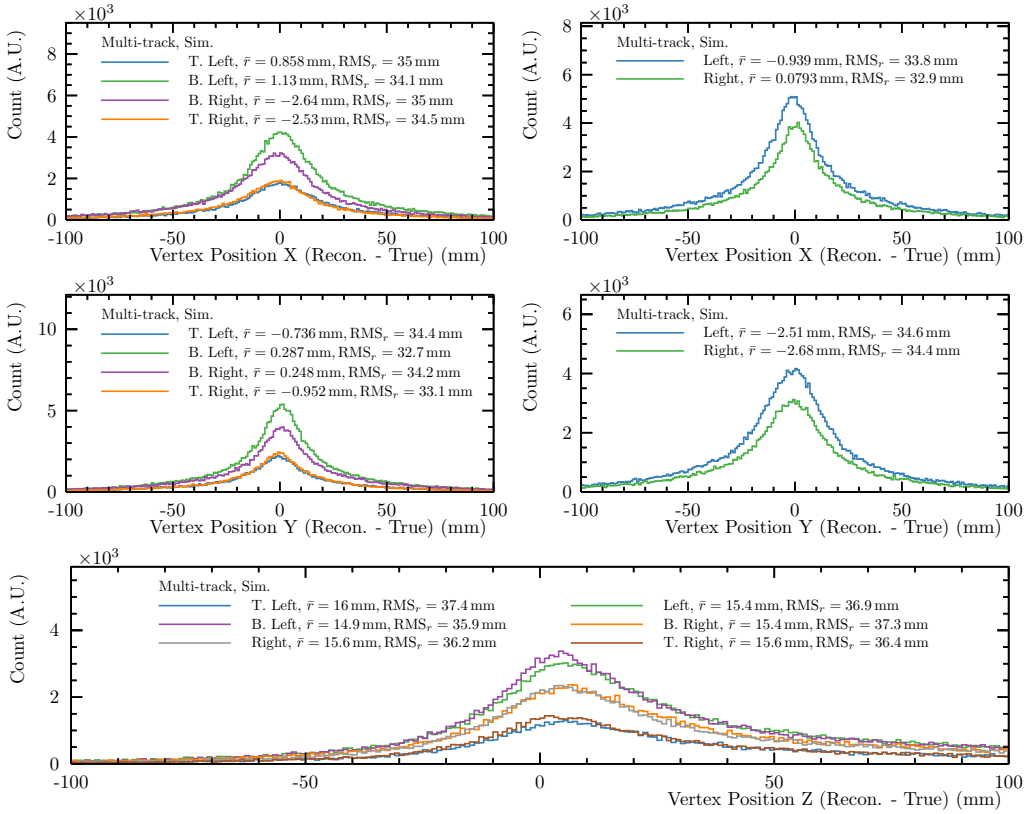
Comparison between simulation and data

The following distributions are presented as a low-level check that the ECal vertex reconstruction behaves similarly between data and simulation. All of these distributions are 'physics sensitive' and may not be expected to display exact agreement if the simulation is imperfect. The most important distributions to see good correspondence in are the track directions, shown in Figure 3.12 and Figure 3.13. The agreement here is satisfactory; there are no significant discrepancies that might indicate a systematic difference in raw reconstruction performance between data and simulation.

The reconstructed ND280 track length is used as an observable by some of the samples presented



(a) One track vertex candidate position residuals



(b) Multi-track vertex candidate position residuals

Figure 3.10: The distributions of the residuals for reconstructed vertex candidates with vertices occurring within the Barrel ECal active regions. The distributions are separated into position coordinate, reconstructed prong multiplicity (one (a), and more than one (b)), and module running coordinate (top/bottom and side). The spread of reconstructed vertex position is approximately 35 mm and no significant biases are evident for any specific dimension, module or track multiplicity.

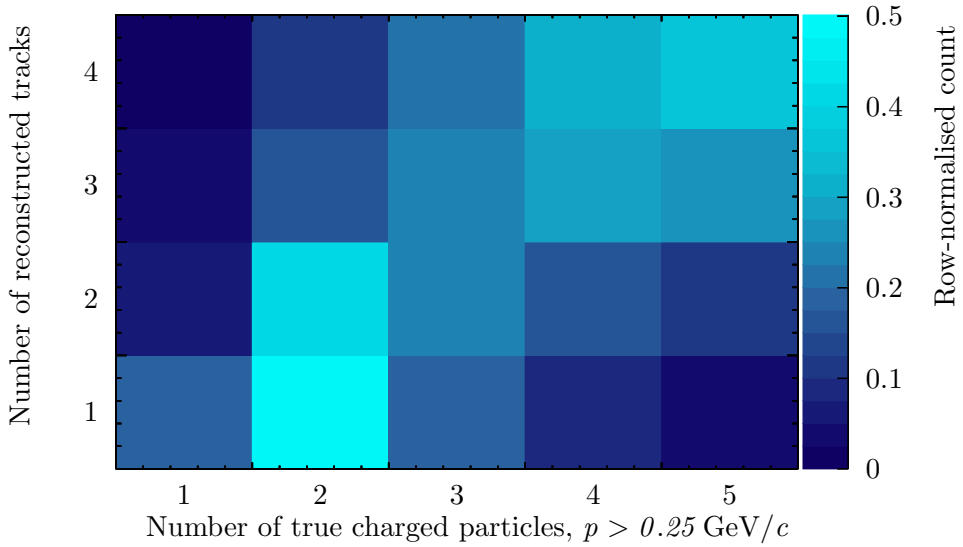


Figure 3.11: The truth content matrix for the number of reconstructed tracks. Each row is normalised to show the fractional content of true topologies for each reconstructed topology. True topologies are defined by a number of produced charged particles with greater than 250 MeV/c of momentum to increase the tracking probability. The quality cuts, described in § 3.1.3, are the only cuts made on the reconstructed tracks.

in the next chapter; the track length for ECal-isolated tracks is shown in Figure 3.14.

Other distributions are presented in Appendix: A.1 and Appendix: A.2 for interest and posterity.

A general feature of these distributions is that the normalisation is generally over-predicted in the bottom modules, approximately correctly predicted in the side modules and under-predicted in the top modules. One plausible explanation would be a remaining mis-modelling in the magnetic field simulation inside of the ND280 flux return—more entering particles may be directed out of the top module and in to the bottom module by the outer magnetic field when compared to the data. It was noted in § 6 of Ref. [96] that the magnetic field was not simulated in the flux return of the ND280 magnet; this was cited as a major contributor to the unsatisfactory result of the 2013 ECal-target analysis. An *ad-hoc* correction, implemented into the ND280 simulation appeared to reduce the discrepancy, but the field model was very simple. The implementation and assessment are presented in § 6 of Ref. [96]. It is also possible that the discrepancy is caused, in part, by a deficiency in the neutrino flux model. These discrepancies, and the repeated failure to produce robust ECal-target neutrino interaction cross-section results, are the main motivation for the analysis presented here.

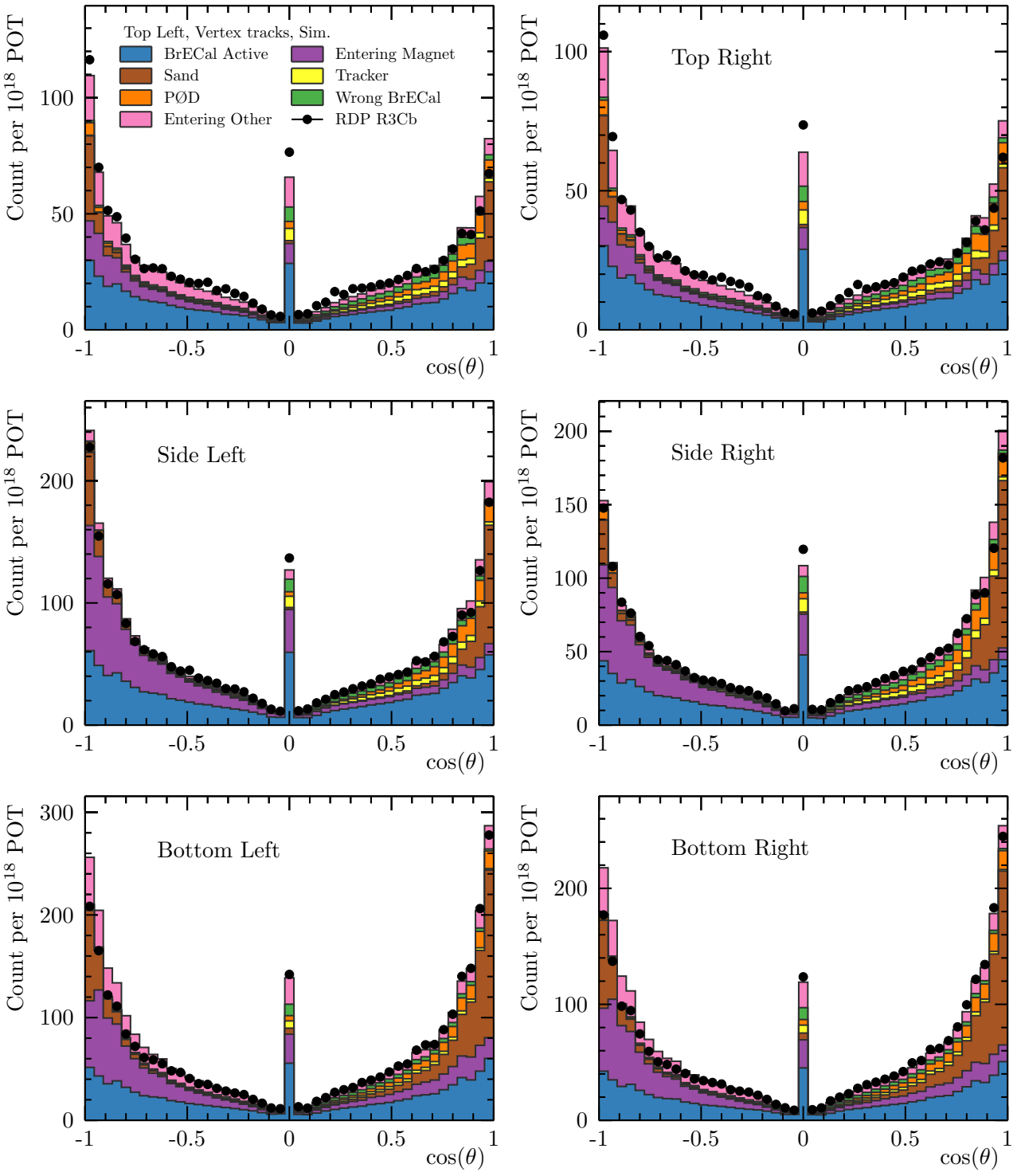


Figure 3.12: The POT-normalised, polar track angle for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. The distribution drawn from the simulation is subdivided by the true interaction topology. The data overlay is calculated from the Real Data Processing (RDP) Run3Cb neutrino beam data sample.

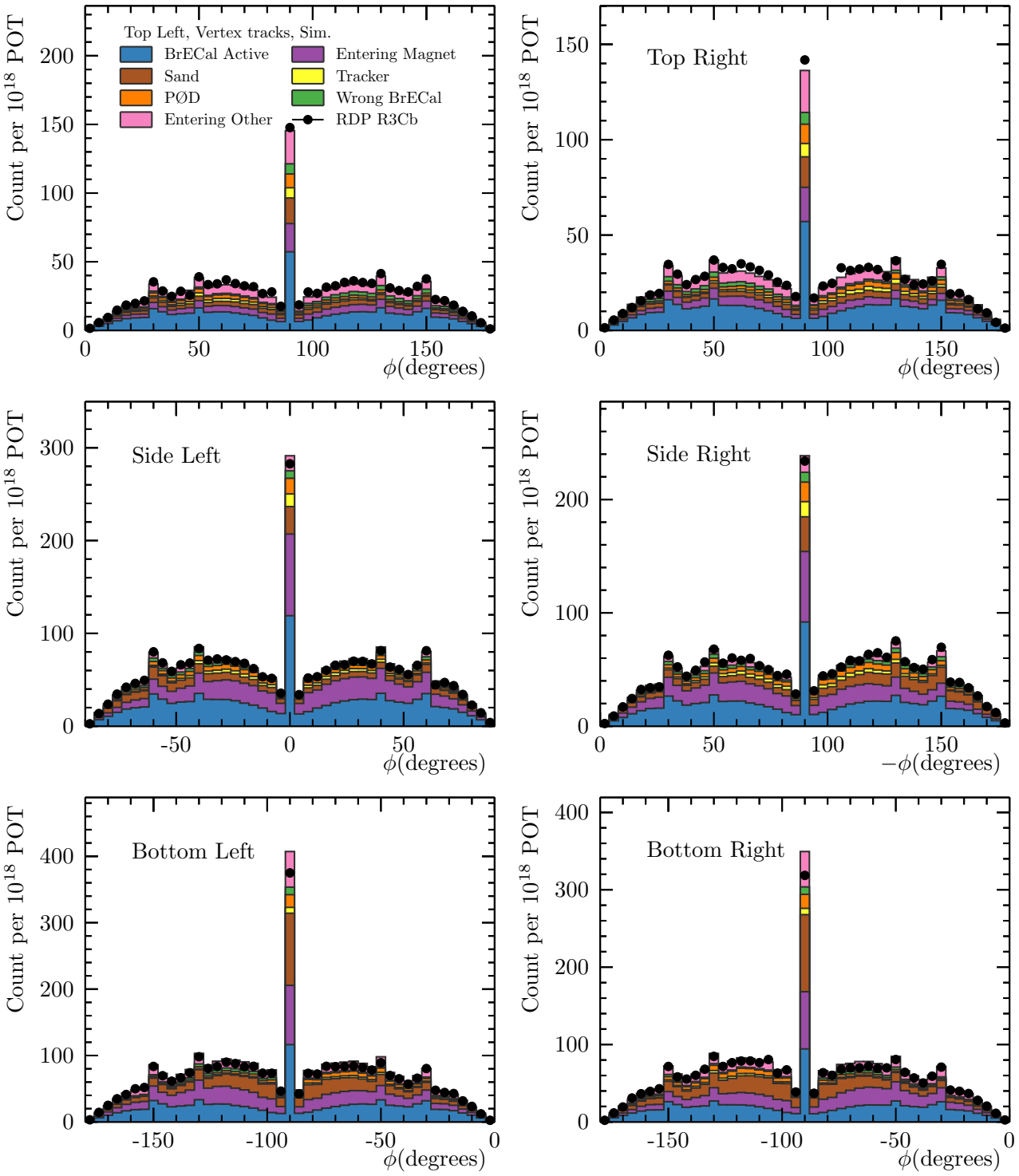


Figure 3.13: The POT-normalised, azimuthal track angle for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. The distribution drawn from the simulation is subdivided by the true interaction position.

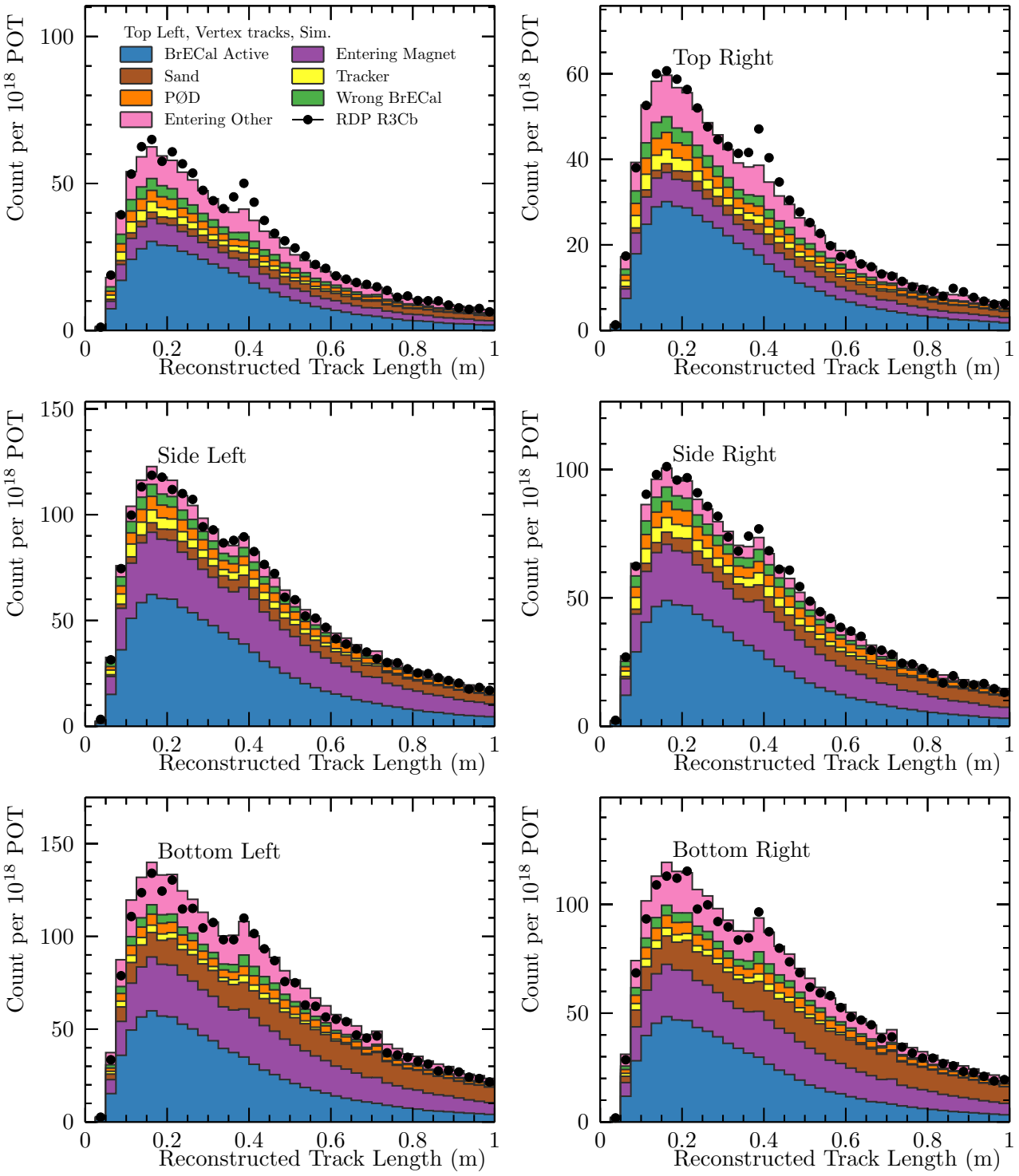


Figure 3.14: The POT-normalised, track length for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. The distribution drawn from the simulation is subdivided by the true interaction position.

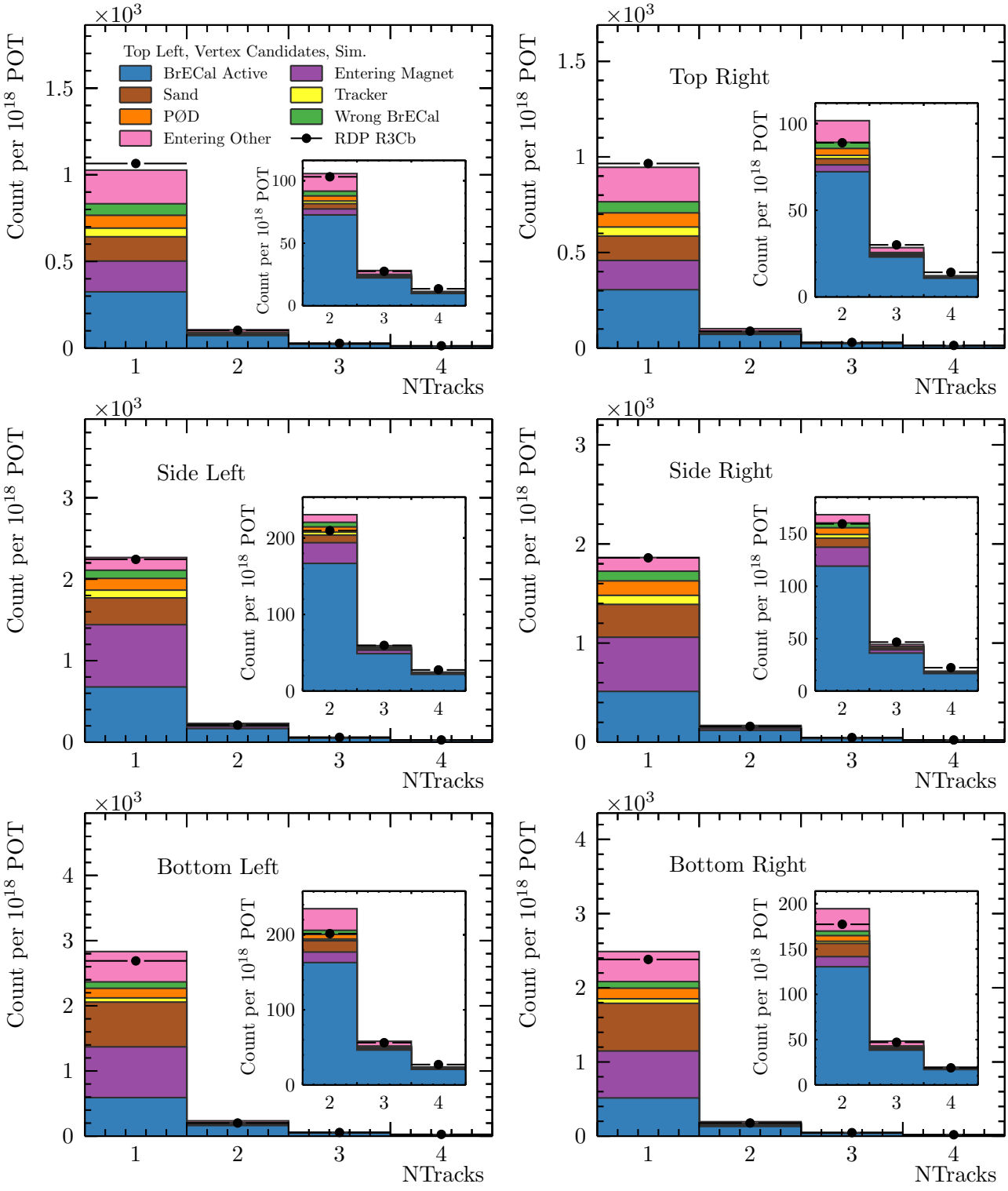


Figure 3.15: The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. The distribution drawn from the simulation is subdivided by the true interaction position. The inlay figures focus on the distribution of multiplicity for vertex candidates with more than one associated track.

3.2 ND280 global track reconstruction

Global track reconstruction takes the objects reconstructed by the individual sub-detector reconstruction algorithms and attempts to match them together to form ND280-wide reconstructed particle hypotheses. At the time of writing the reconstruction software used for TPC-isolated reconstruction had been newly re-written; this was to allow for the reconstruction of neutrino interactions occurring on the gas in the TPCs themselves. This new software, named TREx (TPC Reconstruction Extension), is explained in detail in Ref. [79]. While TREx underwent extensive validation during its development and it was written to be an improvement over the previous software, tpcRecon, it has not been used in a published T2K analysis. The improved performance of TREx motivates its use in this analysis. The next section presents a study validating the use of TREx in the context of ECal-target interactions with TPC-entering muons.

3.2.1 TPC tracking efficiency

The tracking efficiency of the TPC is expected to be very good. However, it has only been extensively assessed for neutrino interactions occurring within the PØD and the FGD. This analysis necessitates effective tracking over a different range of incident angles than is spanned by particles that enter from the PØD or one of the FGDs.

A comparison of the muon tracking efficiency for tpcRecon and TREx as a function of the true momentum, polar angle, and azimuthal angle at the true entry point to the TPC is shown in Figure 3.16, Figure 3.17, and Figure 3.18, respectively. The efficiency is projected onto the true muon kinematics at the entry point to the TPC, rather than the true kinematics at the production point in the ECal. This avoids integrating over any effects of the ECal starting position, such as any energy loss and re-scattering that may occur before the particle enters the TPC. The position on the TPC detector face is still averaged over.

The efficiencies are only shown for the left-hand Barrel ECal modules—the symmetry of the detector means that the results for the right-hand modules are qualitatively the same. The

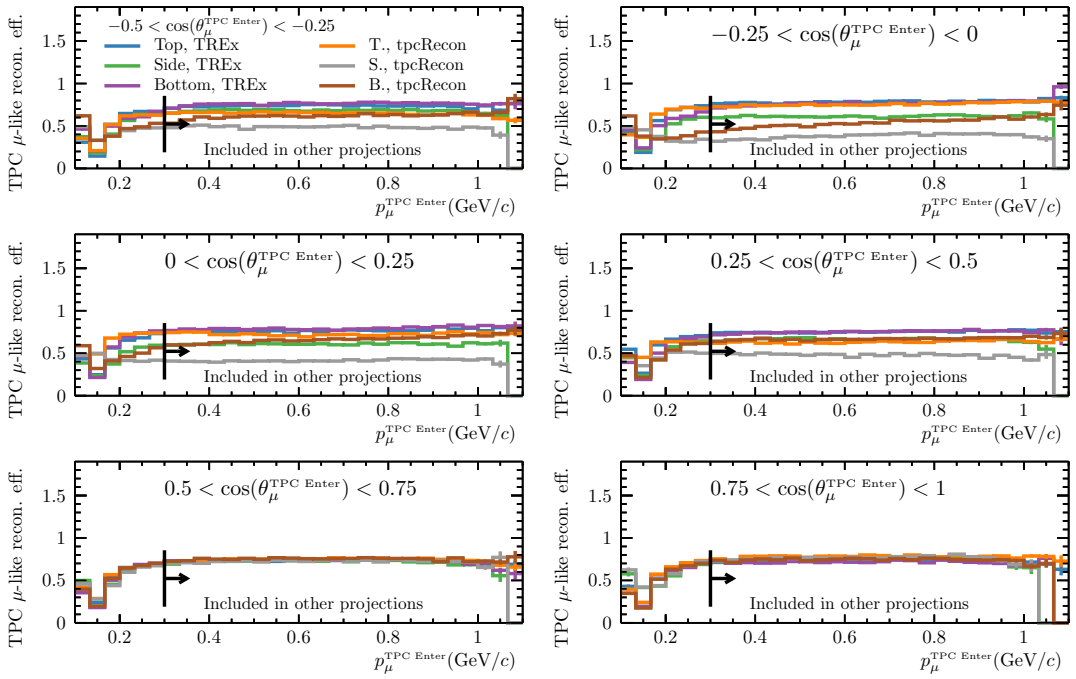


Figure 3.16: Muon tracking efficiencies as a function of true momentum at the entry point to the TPC subdetector. In every bin, TREx matches or out-performs tpcRecon when comparing between muons that came from the same ECal module. The region $p_\mu^{\text{TPC Enter}} < 300\,\text{MeV}/c$ is excluded when projecting the efficiency onto the angular phase spaces.

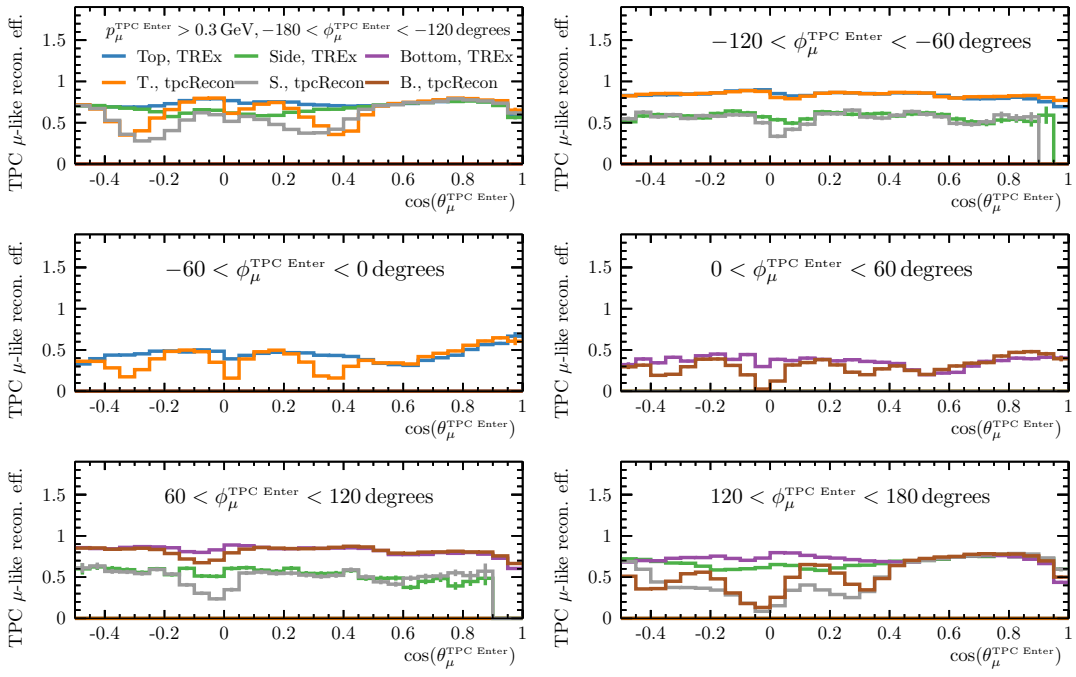


Figure 3.17: Muon tracking efficiencies as a function of true polar angle at the entry point to the TPC subdetector. In every bin, TREx matches or out-performs tpcRecon when comparing between muons that came from the same ECal module and displays a significantly more consistent efficiency across the phase space.

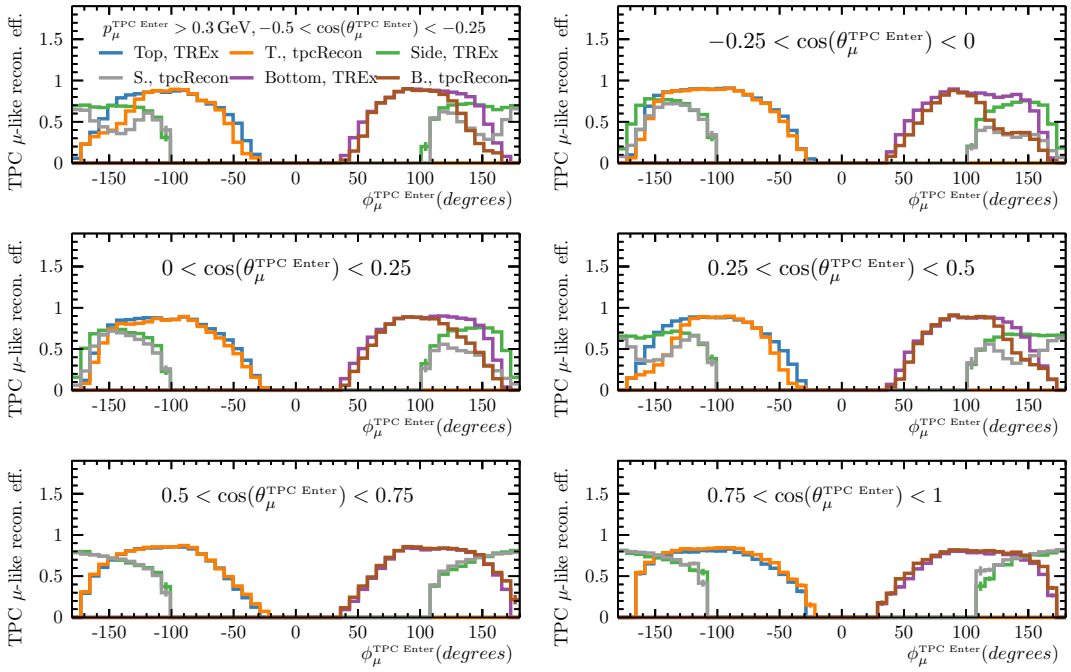


Figure 3.18: Muon tracking efficiencies as a function of true azimuthal angle at the entry point to the TPC sub-detector. In every bin, TREx matches or out-performs tpcRecon when comparing between muons that came from the same ECal module and displays a more smoothly varying efficiency across the relevant phase space for each module.

tracking efficiency is taken as the probability of reconstructing a high-quality[6], muon-like track given that a true muon passed through an active region of the TPC sub-detector. A recurrent difficulty when presenting efficiency-like metrics is disentangling effects that are most evident in projections that are not presented. The denominator of this efficiency requires that the true muon passed through the active region of the TPC, but some number of these will not have passed through a large enough region to have left a track long enough to pass the quality cut. Such effects are mitigated where possible in the design of the metrics, but are often summed over. The efficiency as a function of momentum is approximately flat for muons that enter the TPC with more than 300 MeV/c of momentum—this is used as a restriction when presenting the two angular projections. It is clear from Figure 3.17 and Figure 3.18 that TREx offers more consistent tracking performance across the regions of phase space important for this analysis. As a result of this study, it was decided that TREx should be used as the TPC reconstruction software for this analysis.

For completeness, the TREx tracking efficiency as a function of the azimuthal angle is shown

[6]High-quality here refers to a TPC track with at least 16 MICROMEGAS pad clusters (*c.f* 2.3)

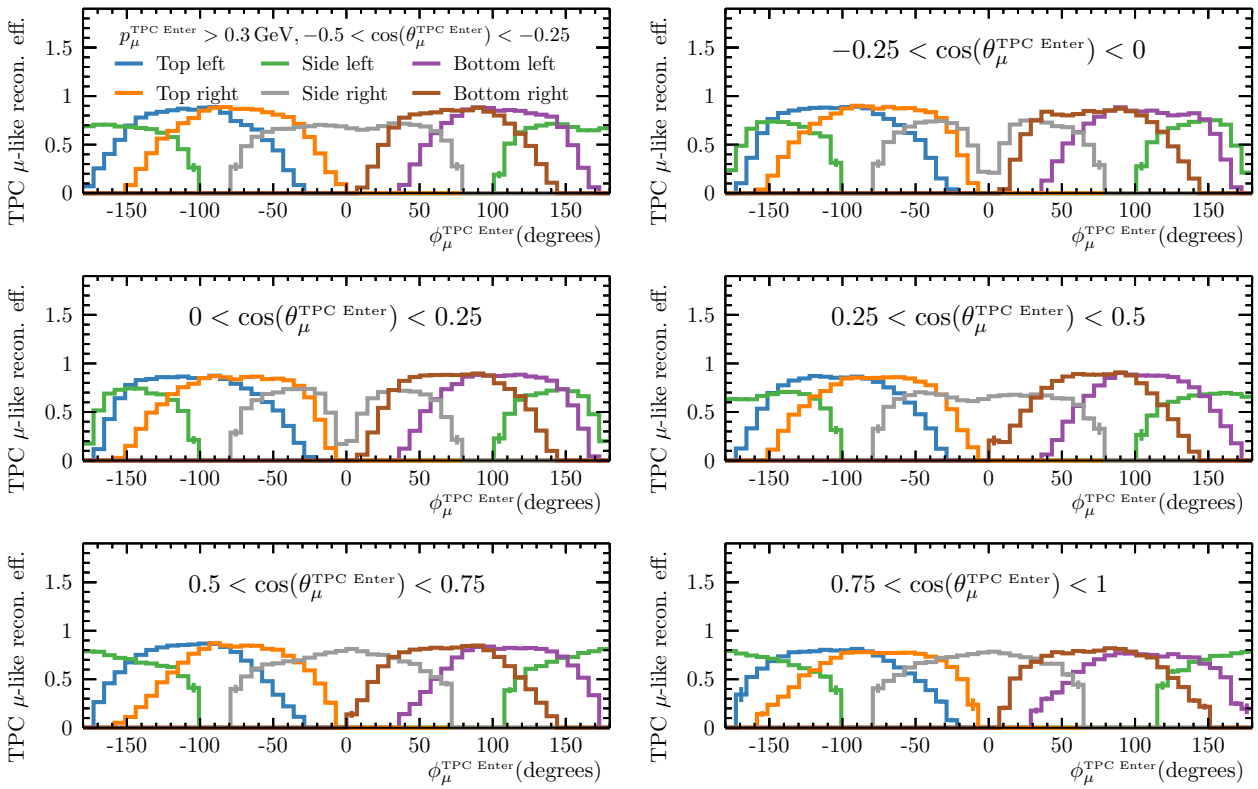


Figure 3.19: TREx muon tracking efficiency as a function of the true azimuthal angle at the entry point to the TPC subdetector for muons entering from each of the Barrel ECal modules. The overall tracking efficiency is consistent except for $\cos(\theta_\mu^{\text{TPC Enter}}) \simeq 0, \phi_\mu^{\text{TPC Enter}} \simeq \{-180, 0\}$ where the muon is travelling approximately (anti-)parallel to the magnetic field.

for all Barrel ECal modules in Figure 3.19. The sum of contributions from each barrel module leaves few regions of the global azimuthal angle outside of the reconstruction acceptance. For tracks travelling nearly (anti-)parallel to the magnetic field, $(-)x$, the tracking becomes very poor—elsewhere the efficiency has a wide peak at around 75%.

The rest of this section introduces the method of inter-detector matching and presents a study assessing the ECal–Tracker matching efficiency. The interfacing of the ECal-isolated vertexing results with global vertex reconstruction is then described and the performance of the particle kinematic reconstruction at the vertex is assessed.

3.2.2 Inter-detector object matching

The final stage of the pattern matching at ND280 attempts to reconstruct single particle tracks that cross sub-detector boundaries by matching and merging reconstructed objects across mul-

tiple sub-detectors. This process uses a Kalman filter to perform track-following; a brief description of the use of a Kalman filter for such pattern recognition is given here, for more details see § 2.4.2 of Ref. [97].

A Kalman filter is an algorithm which processes time-ordered measurements of a dynamical system. While accounting for measurement noise or other stochastic effects in the input measurements, the Kalman filter builds up a best estimate of the parameters of some dynamical model for the system being measured. The Kalman filter used at ND280 models a charged particle propagating under the effects of an electric and magnetic field. It filters reconstruction nodes—position, positional uncertainty, and time measurements—and produces a discriminator for whether each filtered node was likely reconstructed from deposits that were left by the particle being followed. Accepted nodes are used to update the best guess of the current dynamical properties of the particle, which affects the discriminator for later nodes to be accepted.

Global reconstruction uses the Kalman filter to estimate the position and momentum of a particle in the Tracker, where such measurements can be quite precise. To match these objects to reconstructed deposits in the outer detectors, the reconstructed properties of Tracker tracks are propagated through the inactive volumes of the detector to matching surfaces on the faces of the outer detectors. A χ^2-like discriminator, formed between the propagated Tracker object position and direction and the position and direction of the outer detector object, is used to decide whether the two objects should be matched (Figure 3.20). The discriminator is calculated as:

$$\chi^2 = (X_i^{\text{prop}} - X_i^{\text{outer}}) V_{ij}^{-1} (X_j^{\text{prop}} - X_j^{\text{outer}}), \tag{3.3}$$

where $i, j \in \{x, y, z, {}^{dx}/_{dy}, {}^{dz}/_{dy}\}$ enumerate the relevant matching state parameters—in this case for a state with running coordinate y—and V_{ij} is the corresponding calculated covariance matrix element. Tracks that are successfully matched are re-fit with the Kalman filter to incorporate and propagate the new information gained from the match. The new, matched track is returned to the pool to match again. Global matching processes potential matches between specific sub-detector pairs until no further matches are possible. It then proceeds onto the next sub-detector

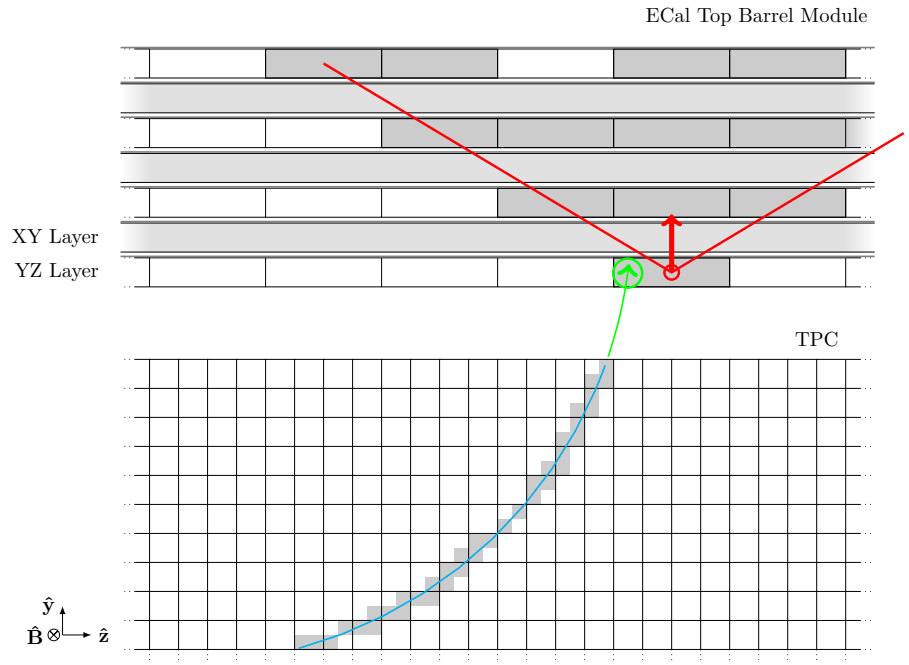


Figure 3.20: The reconstructed position and momentum information from the TPC-exiting end of a track is extrapolated to the active region of an ECal to be used in inter-detector object matching.

pair topology. The first step is to attempt to match Tracker tracks[7] to PØD deposits. This is followed by matching PØD–Tracker deposits to Downstream ECal deposits, then Barrel ECal objects, and finally PØDECal and SMRD objects. In this way, ND280-wide deposits can be effectively reconstructed, giving priority to matches between the most fine-grained detectors before making matches to deposits in the outer detectors.

The most important matching topology to the analysis presented here is between ECal and Tracker objects.

Comparison between data and simulation

The discriminator distribution for all attempted Tracker–Barrel ECal object matches for the Run3Cb RDP sample and a reduced sample of the Run3-equivalent simulation data set is shown in Figure 3.21. Global track matching applies a threshold of $\chi^2 < 70$ when matching objects across sub-detector boundaries. This threshold was determined using single muon particle guns

[7] TPC and FGD objects are passed to global reconstruction pre-matched by a preprocessing algorithm—trackerRecon.

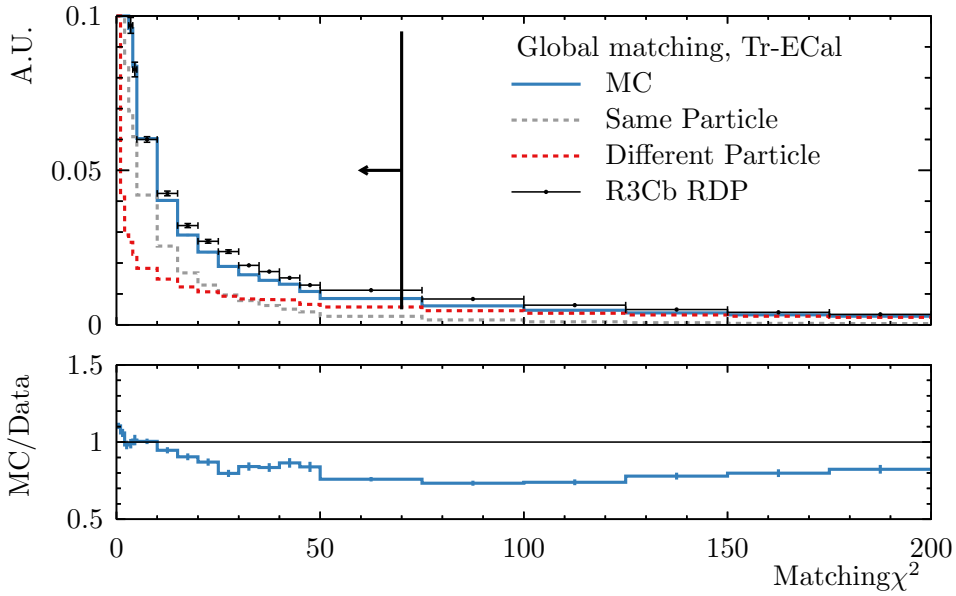


Figure 3.21: The χ^2 distribution for attempted Tracker–Barrel ECal object matches. Global track matching imposes a threshold of $\chi^2 < 70$ in an attempt to maximise the efficiency × purity of matched tracks.

and using the standard ECal reconstruction. It was tuned to maximise the efficiency × purity of the matched deposits. The threshold is placed in the tail of the distribution, this is in part due to the poor uncertainty estimation of ECal objects discussed in 3.1.4. The distributions shown here were produced using the ECal vertexing reconstruction, but the matching efficiency and purity are still adequate. The correspondence between data and simulation is satisfactory in the region where the majority of matches are made and becomes less good in the tail of the distribution. Exact agreement should not be expected as the χ^2 distribution is highly sensitive to any mis-alignment of the sub-detectors in the geometry used to interpret the data. This geometry model was originally constructed using design plans and surveys and is iteratively updated using data from the detector—it is known not to be a perfect representation of the real detector. The distribution of the hadronic deposits about a vertex candidate will be sensitive to the accuracy of the neutrino interaction model, any significant mis-modelling in the shape of hadronic deposits about a vertex may also be reflected in the matching discriminator.

This analysis requires ECal–tracker object matching that is efficient and consistent across the polar angular phase space. The matching efficiency is assessed in the next section.

ECal–Tracker matching efficiency

The tracking efficiencies of the ECal and the TPC for the topology of interest have been shown in 3.1.4 and 3.2.1, respectively. Any inefficiencies as a result of the tracking are ignored in the following distributions by requiring a truth-matched, reconstructed track in both the relevant ECal module and the TPC. The matching efficiency is then the probability of finding the two reconstructed tracks in the same global object; this is presented as a function of the three relevant particle kinematics in Figures 3.22–3.24. The efficiency as a function of the polar angle highlights an oversight in the global reconstruction. For muons that enter the TPC from the top modules at a forward, but high, angle ($0 < \cos(\theta_{\mu}^{\text{TPC Enter}}) \lesssim 0.25$) there is a significant drop in efficiency which is not evident for even slightly backwards muons. Conversely, for muons entering from the bottom with a backward, but high, angle ($-0.3 \lesssim \cos(\theta_{\mu}^{\text{TPC Enter}}) < 0$) the inefficiency shape observed for the top modules appears reflected about $\cos(\theta_{\mu}^{\text{TPC Enter}}) = 0$. One reason for the significantly less consistent reconstruction observed when using tpcRecon is poor performance for tracks that curve back on themselves due to their orientation in the magnetic field. An explicit check for curving back tracks is used by the global matching algorithms to stop such tracks from being matched to objects in the outer detectors. It is likely that when using the much-improved reconstruction that TREx offers, this check can be disabled. Unfortunately, a re-assessment of the global matching performance with this check disabled was outside the scope of the study presented here. The findings of these studies, that present the clear motivation for using TREx and highlight a significant oversight in the global matching algorithm for high angle tracks, will be addressed in the next versions of the ND280 analysis software.

Once all global track matches have been made, each global track is re-fit under three particle hypotheses, muon, electron, and charged pion. These fits are performed under the separate hypotheses that the particle started at each end of the track—forward and 'flipped' kinematic fits—resulting in six sets of reconstructed kinematics for each global track. This re-fit includes an estimation of the energy lost by a particle as it travelled through active and in-active regions of the detector; sub-detectors that do not provide reconstructed momentum estimates are not

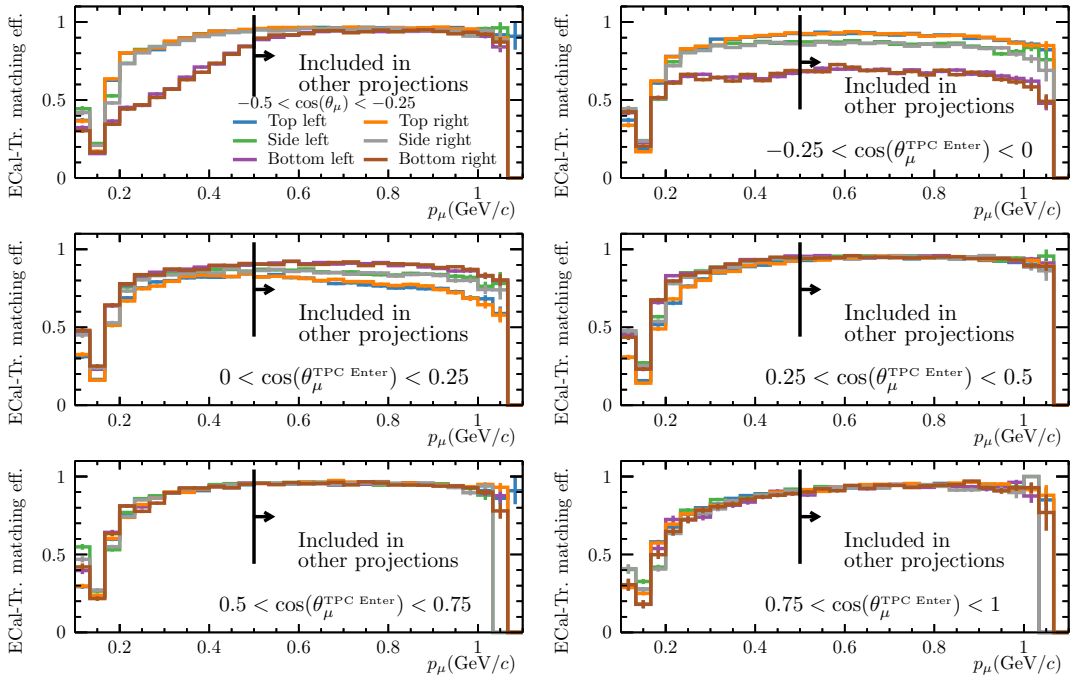


Figure 3.22: The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of momentum at the entry point to the TPC. The region $p_\mu^{\text{TPC Enter}} < 500\,\text{MeV}/c$ is excluded when projecting the efficiency onto angular distributions. For higher momentum muons, the matching efficiency is approximately flat.

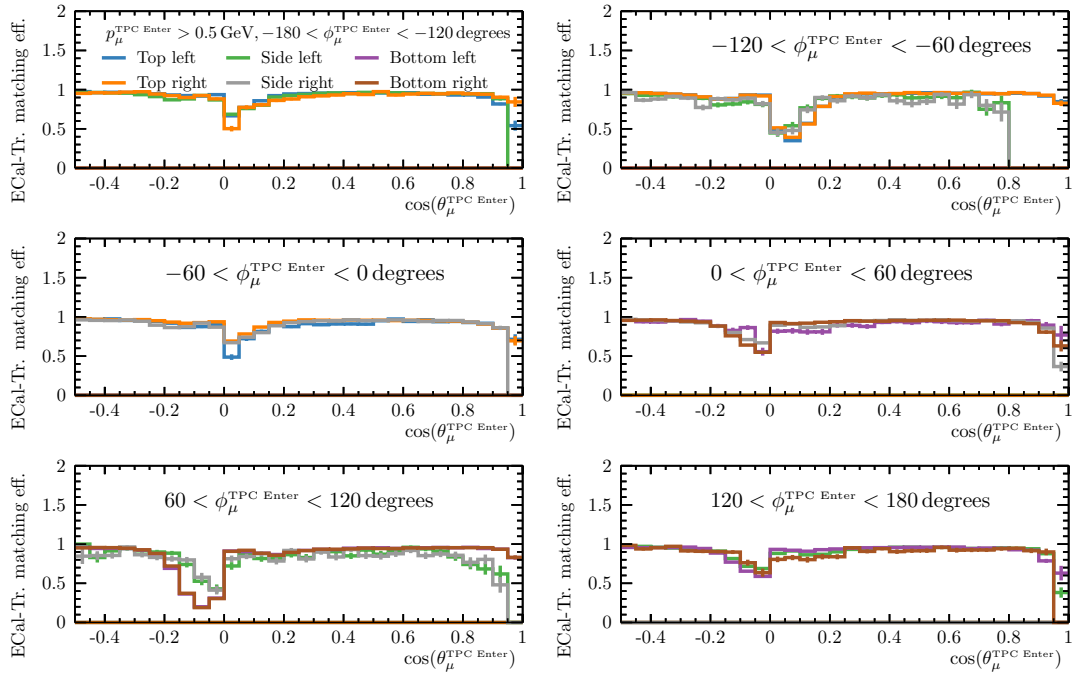


Figure 3.23: The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of polar cosine angle at the entry point to the TPC. Only tracks with $p_\mu^{\text{TPC Enter}} > 500\,\text{MeV}/c$ are included. The matching efficiency is flat over a large range of the polar angle projection. The sharp drop around $\cos(\theta_\mu^{\text{TPC Enter}}) = 0$ is caused by an explicit check that removes curving back tracks from global matching.

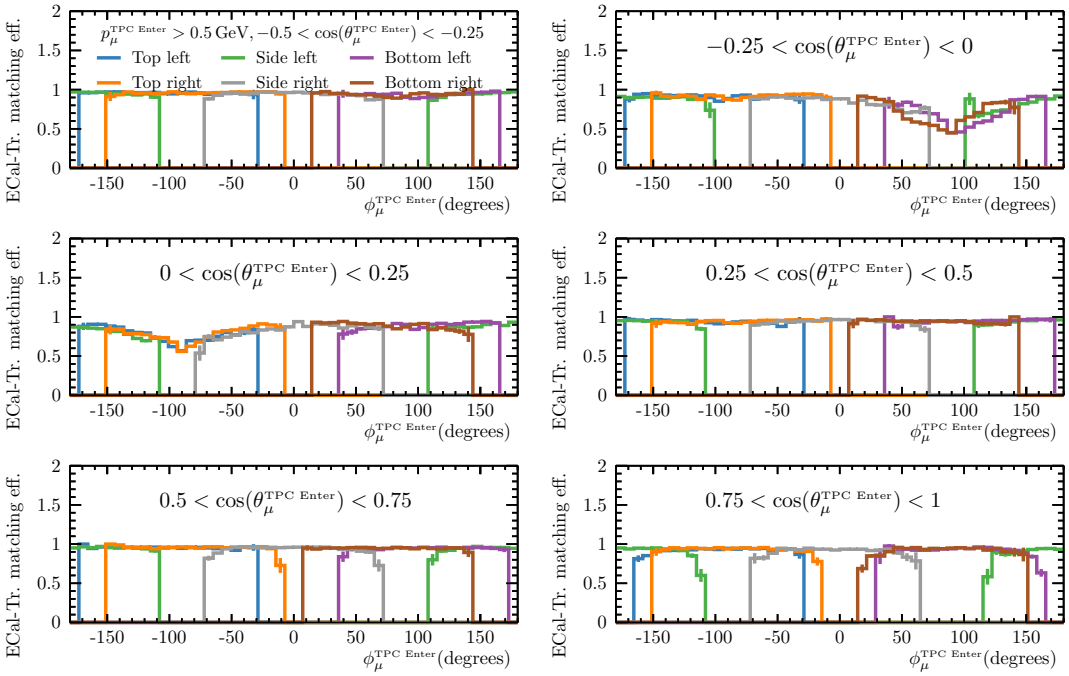


Figure 3.24: The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of azimuthal angle at the entry point to the TPC. Only tracks with $p_\mu^{\text{TPC Enter}} > 500\,\text{MeV}/c$ are included. The curving back track check can be seen to manifest in module-specific regions of the azimuthal projection.

used to inform the kinematic fits, so the energy loss must be estimated from a model. In this way, the reconstructed initial momentum of a particle produced in an interaction within an ECal module can be estimated by an associated deposit left in the TPC—the energy lost before the particle was measured by the TPC is compensated for by an energy loss model. The accuracy of this process is assessed in the next section. The most likely of the six hypotheses is then passed to global vertexing which performs the ND280 vertex candidate reconstruction.

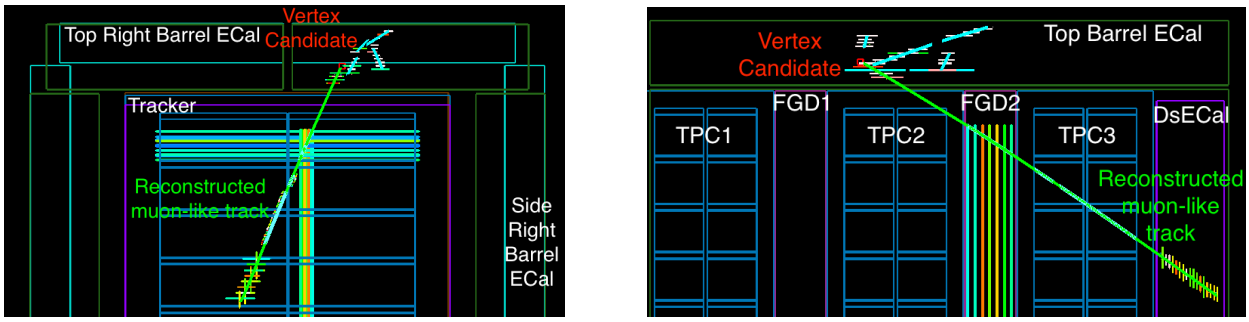
3.3 Global vertex candidate reconstruction

For this analysis, only vertices reconstructed as occurring within the Barrel ECals need to be carefully considered—any tracks that start upstream of a vertex candidate may be used in a veto, but the precise positioning of interaction vertices occurring in other detectors is not necessary. The reconstructed vertex candidates output by the ECal-isolated vertexing are therefore used for the global vertex candidate reconstruction.

The reconstruction proceeds by creating 'proto-vertices' from each of the separate vertex candidates output by ECal-isolated vertexing. Each ECal track is then replaced with the global track that now contains it[8]. At this stage, multiple vertices may be associated to the same track. For example, if a muon passed from the top left Barrel module to the bottom right, through one of the TPC modules, both of the ECal modules would have reconstructed the contained deposits as 'vertex candidates' with a single associated track; both of these would have been replaced by the global object that includes the TPC deposit. Vertex candidates are then grouped based on shared global tracks. The most upstream vertex candidate in each group becomes the primary candidate and any others become secondary candidates stored as child objects of the primary vertex. The shared track is removed from the secondary candidates and any that have no unique global track associations are discarded. In this way, the example presented above would be reconstructed as a single vertex candidate with a single global track association which spanned two Barrel ECal modules and the Tracker.

The position of the primary vertex candidate may have to be readjusted as a result of the candidate merging process. For example, consider a muon which originates from within FGD1 and leaves the detector, depositing energy in side Barrel ECal and the SMRD. It will initially have its 'ECal vertex' position placed in the Barrel ECal module, even though this is now in the centre of a long global track. The choice of vertex position is intertwined with the ability to reconstruct, and the purity of reconstructed, backward-going particles. If the vertex is always placed at the most upstream end of any tracks associated with the vertex, no particles will appear to be backward-going. Truly high-angle and backward final states will be removed from a selection by any entering background cut. But if the vertex is always placed at the vertex candidate position which contains the most associated tracks, then truly forward going particles that undergo a re-scatter within the ECal fiducial volume may be reconstructed as vertex candidates in the ECal with one backward and one forward final state—in truth both tracks were left by the same particle. This would significantly contaminate selections of events with true backwards final states, which are very rare. Until recently, ND280 reconstruction has always assumed that all particles were forward-going. This is an acceptable first assumption as

[8]The global track which contains an ECal vertexing track may just be the same track if no matches to it were made.



(a) Back view (b) Right view

Figure 3.25: An event display depicting the results of the reconstruction on the simulated ECal interaction shown in Figure 3.1. The vertex candidate (red box) position and associated muon-like Tracker track can be seen. The reconstruction of such a deposit was not possible before the global reconstruction improvements that are described in this chapter.

the vast majority of neutrino interactions result in very forward final state particles. The result of running the reconstruction on the simulated energy deposit shown in Figure 3.1 is depicted in Figure 3.25. The vertex position is well reconstructed and the Tracker track, which has a muon-like PID, was correctly associated to the global vertex.

For the current analysis, the decision was made not to attempt 'sense'-, or time-, based reconstruction—calibrating the inter-detector timing to the precision required is a difficult problem (*c.f.* § 4.3). Including the track sense in the reconstruction would necessitate significant systematic assessment. For primary vertex candidates with a single associated global track, the vertex position is placed at the upstream end of the track. For primary vertices with multiple associated global tracks the position remains where it was determined to be by ECal-isolated vertexing. This method provides an attractive compromise between allowing for some correctly reconstructed backward objects, when the ECal vertex has multiple tracks, without introducing difficult-to-assess systematic uncertainties associated with using the track timing. These criteria are clearly sensitive to the neutrino interaction simulation and if the model is not representative, will result in discrepancy.

An unfortunate oversight in the global vertex reconstruction becomes apparent when two tracks from the same ECal vertex candidate are merged into the same global track. This may result in a global track end that is quite far away from the ECal vertex that it was associated to—it may even extend into the neighbouring detector. For some deposits it may be correct to merge

the two ECal tracks, in which case the original ECal vertex reconstruction was at fault. For other deposits, the second ECal track may have been left by a different particle, travelling approximately back-to-back with the first. Reconstructing such a deposit as a single track will cause a bias in the global momentum estimate due to extra energy loss compensation. It was deemed too complex to fix this oversight given how infrequently it causes problems; instead, reconstructed vertices that exhibit this pathology are cut by the selection by comparing the position of the start of the global track with the reconstructed vertex position.

In the next section, the performance of the global momentum and track-angle estimation is assessed for tracks associated to a vertex candidate in the ECal.

Reconstruction performance

The reconstructed momentum and polar angle residual distributions are shown in Figure 3.26 and Figure 3.27. The distributions were calculated using the Inward Muon particle gun sample. Included events were restricted to those with well-reconstructed, muon-like TPC components that were associated to a vertex within the ECal active region from which the particle guns were fired. This topology is representative of the true topology of interest for this analysis. The default reconstructed momentum estimate assumes that the track end closest to the associated reconstructed vertex is the track start—the 'flipped' global track hypothesis is used when the 'end' of the track is closest to the vertex position. While the biases evident in Figure 3.26 should be acknowledged when interpreting the data, there is no reason to believe that the cause should manifest differently for real data events. The momentum residuals between the reconstructed TPC momentum compared to the true momentum at the TPC entry point are shown in Figure 3.28. These exhibit a small skew for muons emanating from the side modules, but the bias is significantly less prominent than in Figure 3.26. It is therefore likely that the global momentum bias arises mostly from the energy loss correction. This might be because of a bias in the reconstructed track end position that systematically moves the reconstructed position away from the TPC—this would result in an over-correction. However, Figure 3.10a shows no such bias. The detector model used to calculate the energy loss correction (RECPACK) is more

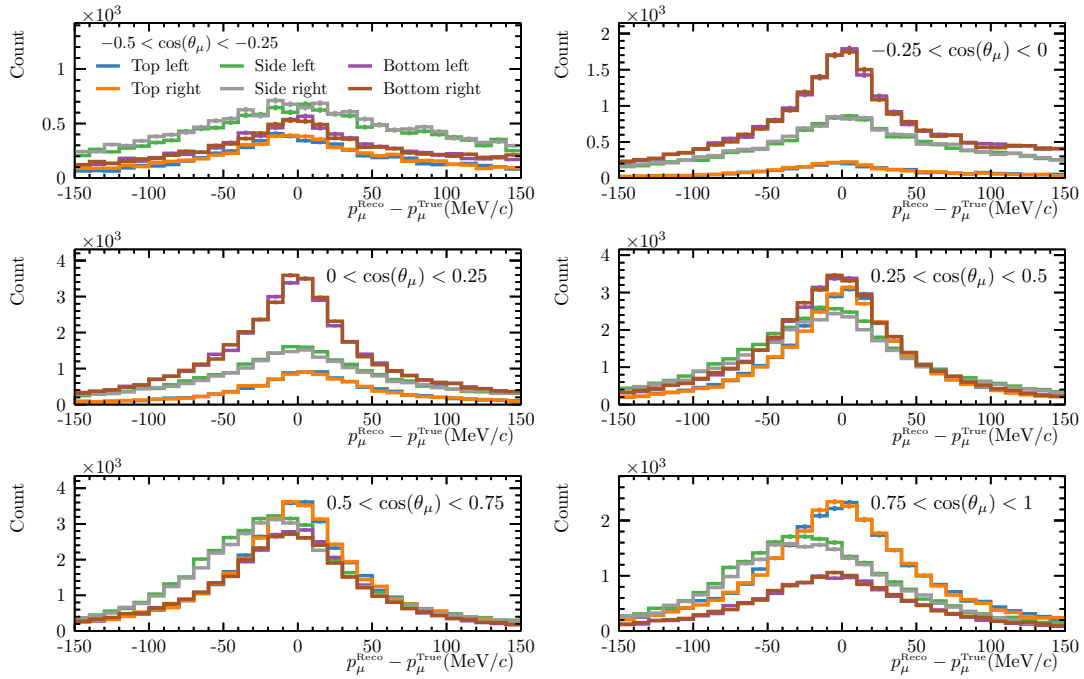


Figure 3.26: The momentum residuals for globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. The distribution is further separated into bins of true production polar angle. A significant bias can be seen for forward going tracks originating in the side ECal modules.

simple than the one used by Geant4 to simulate the true particles produced by the interaction generator. This is to increase the computational efficiency of the reconstruction software. It is possible that an over-simplification in the inter-detector spaces between the TPC and the side Barrel ECal modules contributes to the bias. The bias is not observed to the same degree for the top and bottom modules.

While the work presented in this chapter addresses the necessary reconstruction issues to begin investigating muons produced in interactions within the Barrel ECal, there is scope for improvement.

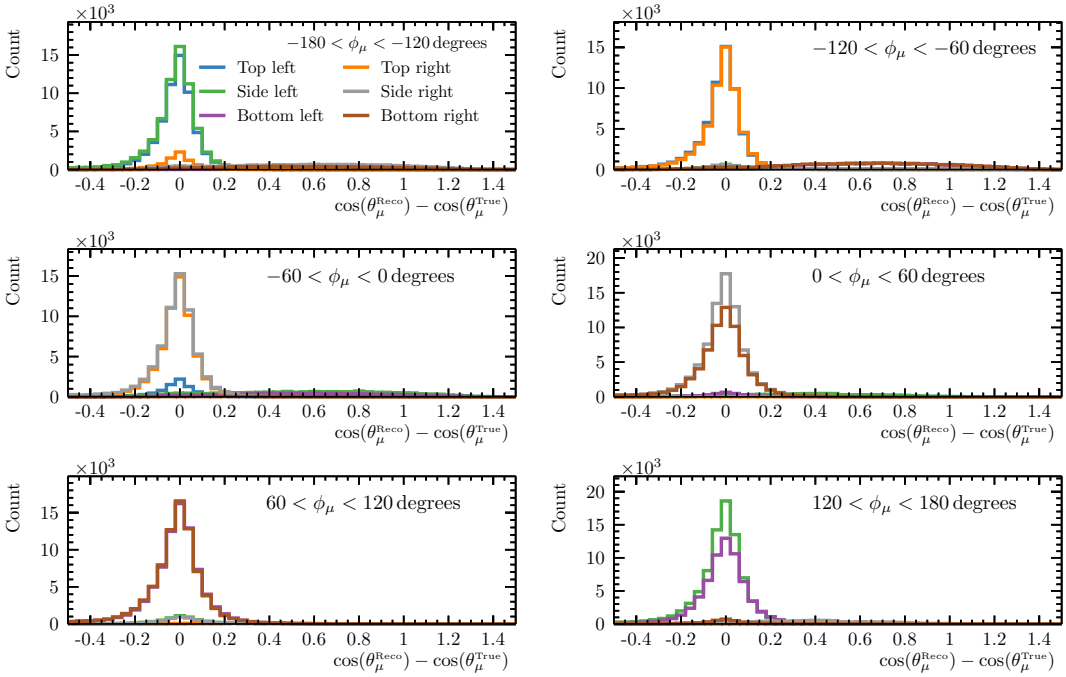


Figure 3.27: The residuals of the polar angle cosine for globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. The distributions are largely unbiased, but have a long tail extending into a region where the reconstruction imparts a significant forward bias onto angle estimate for specific regions of ϕ_μ phase space.

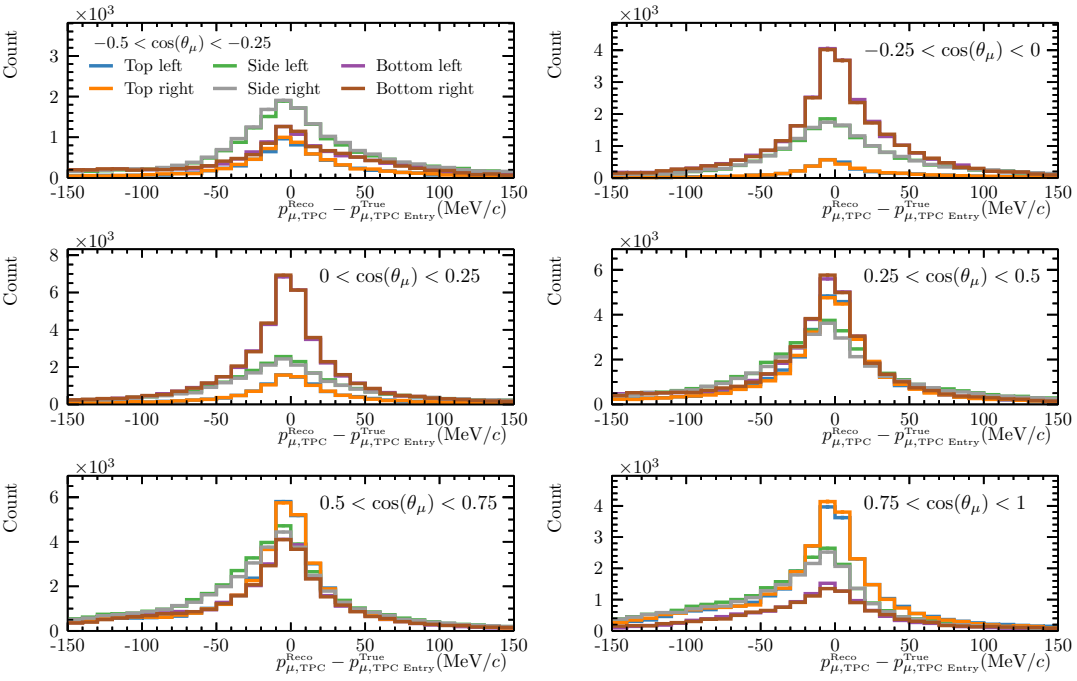


Figure 3.28: The momentum residuals for the TPC component of globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. A small bias may be expected as the true momentum used is that at the TPC detector entry point, which will not be quite the same position as the nearest reconstructed node within the TPC active volume. The biases for muons entering from the side Barrel ECal modules are less evident than in Figure 3.26, which suggests that the effect is not solely due to the TPC reconstruction.

3.4 Future work

Using the Barrel ECal as a target presents a unique opportunity within ND280 to perform a precise, high-statistics study of high angle and backwards particles. This is because of the scintillator layer orientation relative to the beam axis, the high target mass, and the proximity to the tracker. Conversely, when attempting a similar analysis that uses the FGD as a target, there is a much more limited phase space for high-angle final states to leave well-reconstructed deposits within the TPC—most will travel straight into a Barrel ECal module. The reconstruction improvements presented here have made such an ECal-target analysis possible, but there is still scope for improvement. Future work should focus on developing a new procedure for motivating ECal-isolated track slope uncertainties (*c.f.* § 3.1.4), which will allow a re-tune of the global matching thresholds. Further assessment of TREx, specifically concerning the reconstruction and global matching of backward curving tracks, would also be necessary to be confident in the selection (*c.f.* § 3.2.1). A detailed investigation into the inter-detector timing resolution and the use of the track sense to choose the reconstructed vertex candidate point may significantly reduce the amount of forward-going background being reconstructed as ECal vertex candidates with backward-going final states.

Future analyses may also want to use more ECal information. Improvements to the short track reconstruction efficiency would be facilitated by allowing Hough tracks to share hits. Investigation into tuning the ECal-isolated particle identification algorithms for Hough tracks may allow selections of events with and without charged pions—a distinction important to model builders (*c.f.* § 1.3.2).

Chapter 4

ND280 ECal interactions selection

This chapter describes the samples and selection methods used to build ECal interaction candidate samples and investigates the selection performance on simulated neutrino beam data. Real data are also used where possible to motivate or cross-check any choices made. The Real Data (RDP) and simulated beam data (Sim.) used in this chapter are the Run3Cb and Run3-equivalent data sets respectively. The Run3Cb real data set was chosen as it had already been analysed for ECal interactions in the 2015 analysis [96]. While this analysis does not aim to be fully blind, it is best to use statistically independent development samples and final analysis samples where possible.

The 2015 ECal-target selection, described in [96], exclusively used ECal information; the developments described in § 3 allow for a more robust interaction candidate selection and the event samples presented in more illuminating projections. In this chapter, four event samples will be introduced: two that select interaction candidates that occur within the Barrel ECal, and two that aim to characterise the sources of entering background. The next section presents the choice of a new, data-motivated fiducial volume for the Barrel ECal.

4.1 The fiducial volume

4.1.1 The previously-used fiducial volume definition

In the 2015 ECal-target analysis, a fiducial volume was defined to select well-reconstructed vertices from neutrino interactions that occur within the ECal active volume—as opposed to mis-reconstructed, entering background. The fiducial volume definition was tuned to maximise $\phi_{\text{sel.}} = \epsilon \times \eta^2$, where ϵ is the selection efficiency and η the selection purity [96, § 7.6.1.1]. This optimisation resulted in a very permissive fiducial volume—approximately a single scintillator bar is excluded in each dimension except the upstream face, where a larger active veto region is used. The fiducial volume that results from this definition is depicted in Figure 2.12. This choice retains a high event selection efficiency by design, but renders any efficiency corrections susceptible to mis-modelling, both in the modelling of neutrino interactions occurring within the ECal and in the distributions of, and detector response to, entering background particles. Such a permissive fiducial volume would result in true interactions that occurred near the edges of an ECal module being poorly reconstructed because of missed tracks, yet still being selected. A number of the selection cuts used in the 2015 analysis were dependent on the number of reconstructed tracks [96, § 7.6]. It is expected that the track reconstruction efficiency falls off sharply in the few layers closest to the edge of each module. The rate at which tracks are missed, and hence the frequency that the reconstructed vertex position is biased, is a function of the ECal detector response[1], the neutrino interaction and hadronic re-interaction models, and true interaction position within each module. In the 2015 analysis, every phase space dimension except the ECal module in which the event candidate occurred was summed over.

Integrating over regions of sharply varying efficiency imposes model assumptions into any inferences made. Event selections can contain cuts which are tuned in the context of a given simulation that may include inappropriate models. The results of these selections on simulated data may then be used to determine efficiency-like properties of the selection, which are often used to inform 'corrections' of the observed data. If the efficiency correction is not flat in the

[1] The total ECal vertex track charge is known to be poorly simulated: *c.f.* Figure A.4 and Figure A.5

relevant regions of phase space it will alter the shape of the data—the full shape of an efficiency correction is often unseen because many relevant dimensions of phase space are never explicitly investigated. The alteration to the shape of the data will be correct if, and only if, the simulation is a good representation of nature. The introduction of such irrevocable model dependence severely impacts the utility and longevity of published data. Even when explicit corrections are not made, any inferences about the data in the context of a poorly-understood, non-flat selection efficiency will be biased. If an efficiency correction can be shown to be approximately flat in all integrated-over dimensions of phase space, then any bias becomes effectively a normalisation bias— simpler to control than a shape uncertainty only apparent in some dimension of phase space that the data are rarely projected onto. Regions of flat selection efficiency should be used wherever possible.

A fiducial volume that can be shown to exhibit an approximately flat reconstruction efficiency is highly desirable as it better-justifies summing over the true interaction position within each Barrel ECal module.

4.1.2 A data-motivated fiducial volume

A strong motivation for this selection was to provide a set of cuts which select true Barrel ECal interactions with a high degree of confidence. This allows discrepancies between the simulated prediction and the observed data to be investigated in isolation from other effects of poor modelling in the outer detectors. An appropriate fiducial volume choice is a significant step towards such a sample. While the neutrino flux normalisation and spectral shape do vary across the Barrel modules because of the off-axis effect, the variations across a single module should be small. Lead-target interactions occurring near the centre of an active region should therefore produce a similar distribution of final state particles as those occurring near the edge of the same module. The detector response should also be uniform across a module as each pair of alternately oriented layers is constructed similarly through the bulk of the detector[2]. The ratio of the number of reconstructed single-track events to multi-track vertex candidates

[2]discounting un-cooperative electronics.

Module	$\lvert\Delta x\rvert$ Constraint /mm	$\lvert\Delta y\rvert$ Constraint /mm	$\lvert\Delta z\rvert$ Constraint /mm
Top/Bottom	650	125	1450
Side	650	900	1450

Table 4.1: The restrictions placed on the integrated-over dimensions in the distributions of $\Delta\{x, y, z\}$ that are used to motivate the choice of fiducial volume.

can then be expected to be flat in regions of consistent reconstruction efficiency. It should be expected that interactions that occur near the edges of a Barrel module will result in a higher rate of missed charged particles where particles have left the active ECal region without leaving a reconstructable deposit. Entering background particles are also most likely to be reconstructed as a single-track vertex near the outer edge of a module. Therefore looking at the ratio

$$R^{\text{P}} = \frac{\text{N single-track reconstructed vertices}}{\text{N multi-track reconstructed vertices}},$$

across each ECal module should highlight regions of consistent reconstruction efficiency and minimal entering background. Furthermore, this ratio does not use any truth information from the simulation and so can be used to check the correspondence of track acceptance between real data and simulation.

The rates of one-track and multi-track reconstructed vertex candidates as a function of the z distance from the centre of each module, Δz, are shown in Figure 4.1 and Figure 4.2, respectively. The Δx distribution and the Δy distribution for multi-track vertex candidates can be found in Appendix: A.2. To mitigate edge effects from dimensions that are not presented, the other two dimensions are restricted; the restrictions are described in Table. 4.1.

The correspondence between the shape of the rates for the data and simulation is good throughout the bulk of the detector. Significant discrepancy can be seen near the outer faces of each module—a good motivation for not tuning the fiducial volume on simulation alone. It is interesting to note that the interaction rate does noticeably vary across the side modules as a function of y position. The variation is also apparent in the data and can be seen in Figure 4.3.

The regions of poor track-reconstruction efficiency can be judged from the distribution of multi-track vertex candidates. In the z dimension, the rate of multi-prong vertex candidates can be

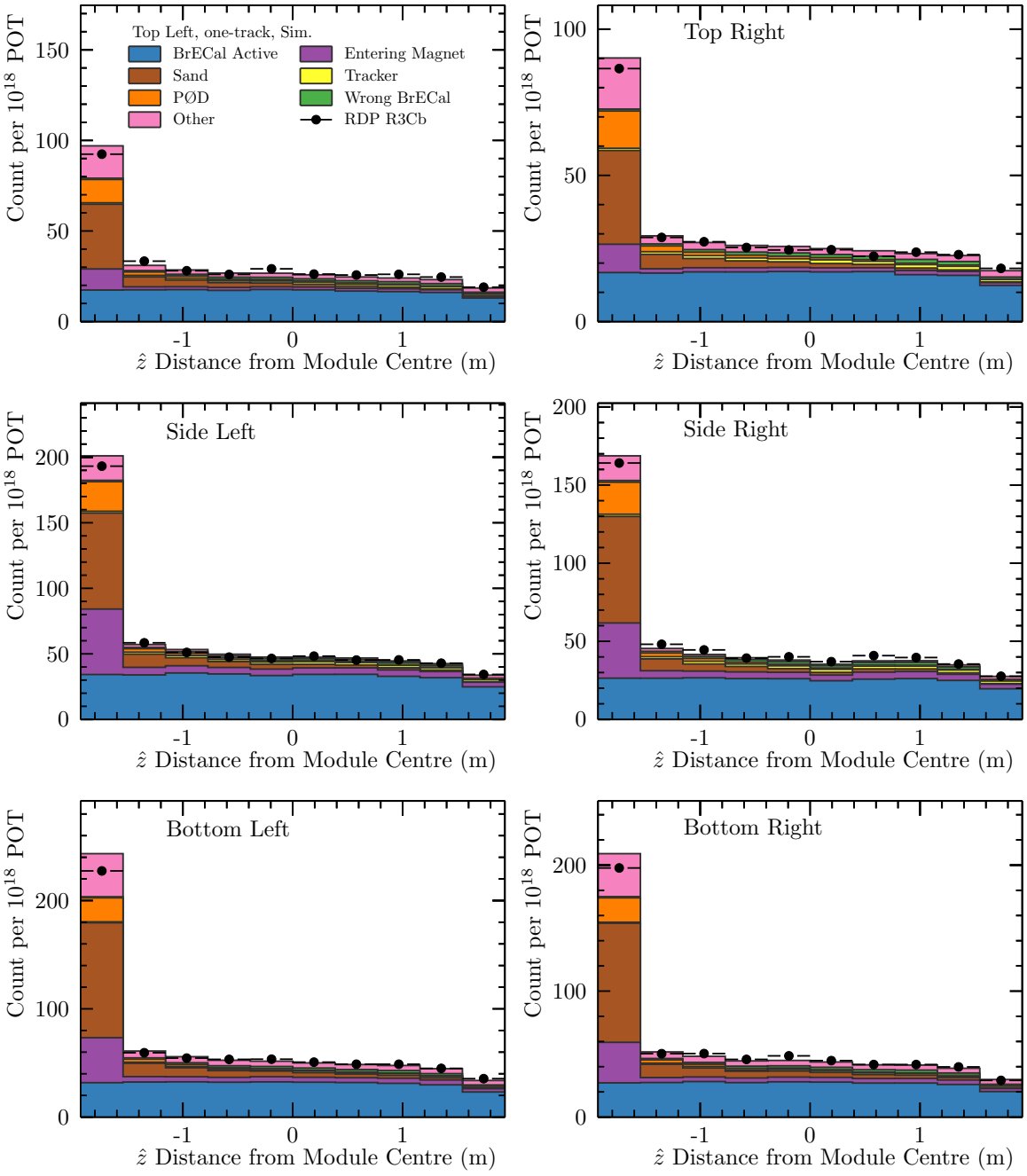


Figure 4.1: The rate of reconstructed one-track ECal-isolated vertex candidates as a function of the z distance from the centre of the module. The simulation predicts that a large proportion of the one prong vertices that are reconstructed near the upstream face of each module can be attributed to entering sand muons.

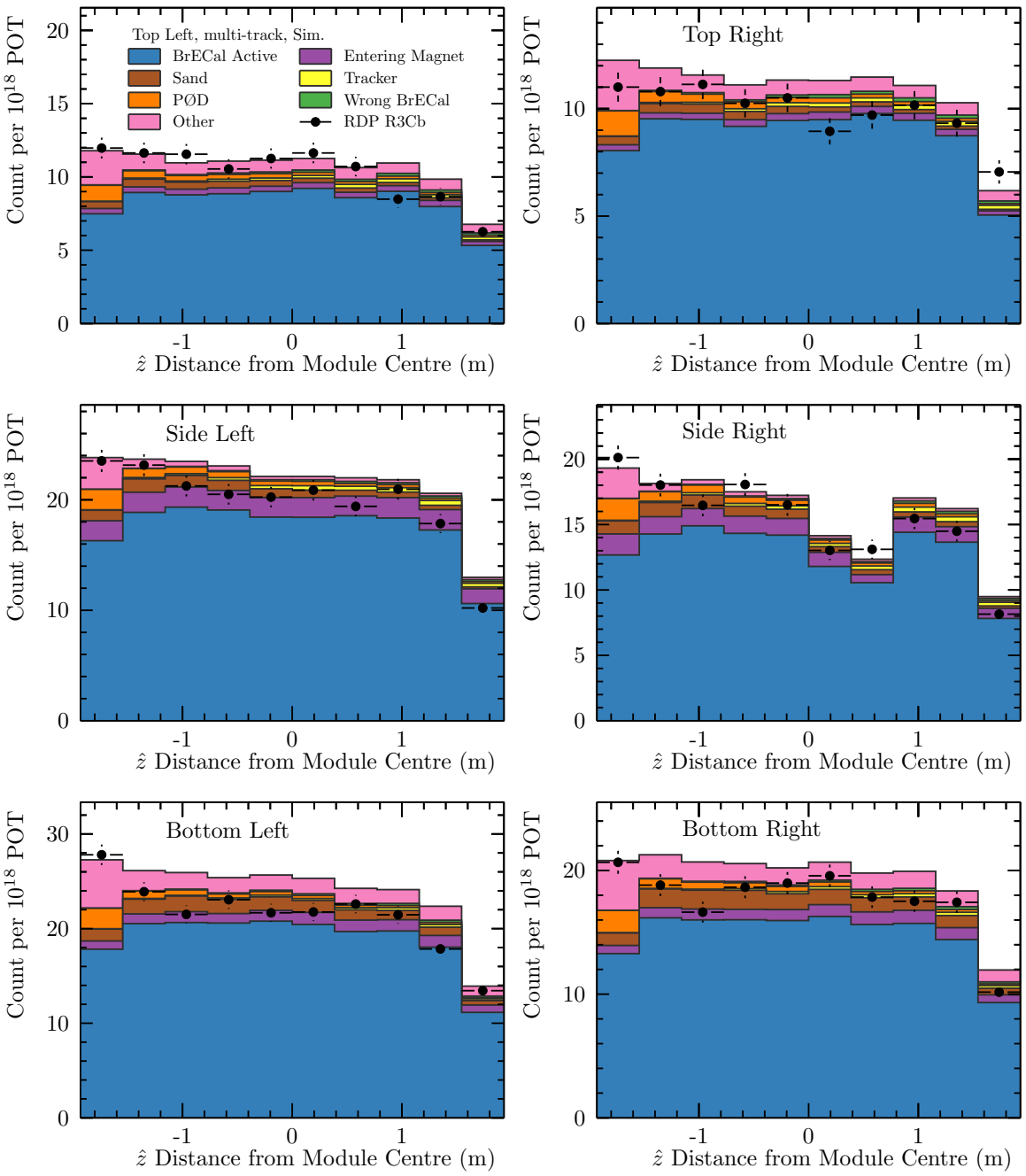


Figure 4.2: The rate of reconstructed multi-track ECal-isolated vertex candidates as a function of the z distance from the centre of the module. A possible localised data deficit may be apparent at $\Delta z \sim -1\,\mathrm{m}$, it is difficult to decide with the data-statistics available in the control sample. The feature observed in the side right module can be attributed to the damage sustained to the ECal electronics in the Great East Japan Earthquake.

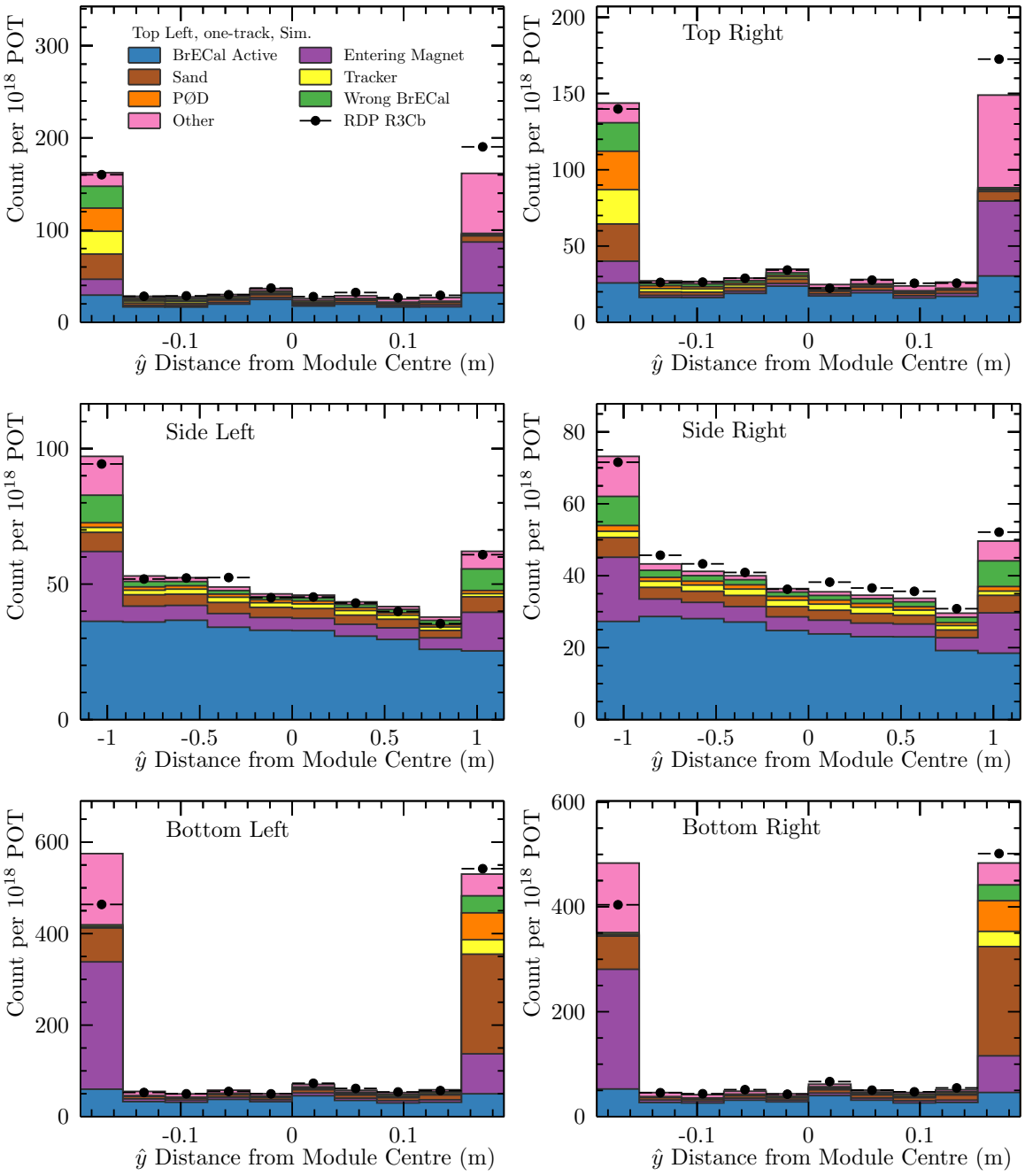


Figure 4.3: The rate of reconstructed one-track ECal-isolated vertex candidates as a function of the y distance from the centre of the module. The simulation prediction is subdivided by true interaction topology. The interaction rate can be seen to vary over the extent of the side modules due to the large span of angular distance from the neutrino beam axis within a single module. The 'other' template is predominantly aluminium-target interactions on the support frame of each ECal module and the ND280 magnet coils.

Module	$\mathcal{M}_{\text{FV}}$ /kg	$\mathcal{M}_{\text{Act}}/\mathcal{M}_{\text{FV}}$	$\mathcal{M}^{x}_{\text{FV}}/\mathcal{M}_{\text{FV}}$	$\mathcal{M}^{y}_{\text{FV}}/\mathcal{M}_{\text{FV}}$	$\mathcal{M}^{z}_{\text{FV}}/\mathcal{M}_{\text{FV}}$
Top/Bottom	1558	3.55(1)	1.52(5)	1.72(6)	1.34(8)
Side	2804	2.95(5)	1.72(6)	1.26(9)	1.34(8)

Table 4.2: The fiducial masses of the two Barrel ECal module types. The associated mass ratios show the relative increase in mass inside the 'selected' volume if the whole fiducial volume restriction is ignored ($\mathcal{M}_{\text{Act}}/\mathcal{M}_{\text{FV}}$), or one dimension at a time is ignored—*e.g.* x, $\mathcal{M}^{x}_{\text{FV}}/\mathcal{M}_{\text{FV}}$.

seen to fall off at the downstream end of each module; here there are not enough bars downstream of the interaction to adequately reconstruct the tracks. This can be seen in Figure 4.2. It is clear that the majority of entering background is reconstructed near the outer faces of each module. Specifically, the vast majority of the entering background from the magnet flux return and sand muons are reconstructed on the largest outer face and the upstream face of each module, as seen in Figure 4.3 and Figure 4.1. Selections of such vertex candidates can be used to constrain distributions of entering background that may be poorly modelled in the simulation.

The distributions of R^{P}_{x}, R^{P}_{y}, and R^{P}_{z} for data and simulation, are shown in Figure 4.4, Figure 4.5, and Figure 4.6 respectively. The choice of fiducial volume was motivated to remove the sharp change in R^{P} at the edges of each module. The regions marked as included show the chosen fiducial volume. The choice was made by-eye to preclude the introduction of model-dependence from tuning to the simulation. A possibly interesting feature can be seen in Figure 4.6 where an apparent localised increase in the ratio is observed at about 1.2 m upstream of each module centre—but only for data. This appears to be mostly due to a relative deficit in the number of multi-track vertex candidates, rather than an excess of one-track candidates. This can be seen from the comparison of Figure 4.1 and Figure 4.2.

The two-dimensional projections of the chosen fiducial volume, relative to the active regions of each Barrel ECal module, can be seen in Figure 4.7. The fiducial masses of the modules and the relative decrease in mass imparted by each dimension of the fiducial volume cut, and the total fiducial volume cut, is shown in Table. 4.2. Even the restrictive fiducial volume presented here retains a total fiducial mass of 11, 840 kg; as a result, this analysis is unlikely to be statistically limited. The next section introduces the vertex candidate selection methods.

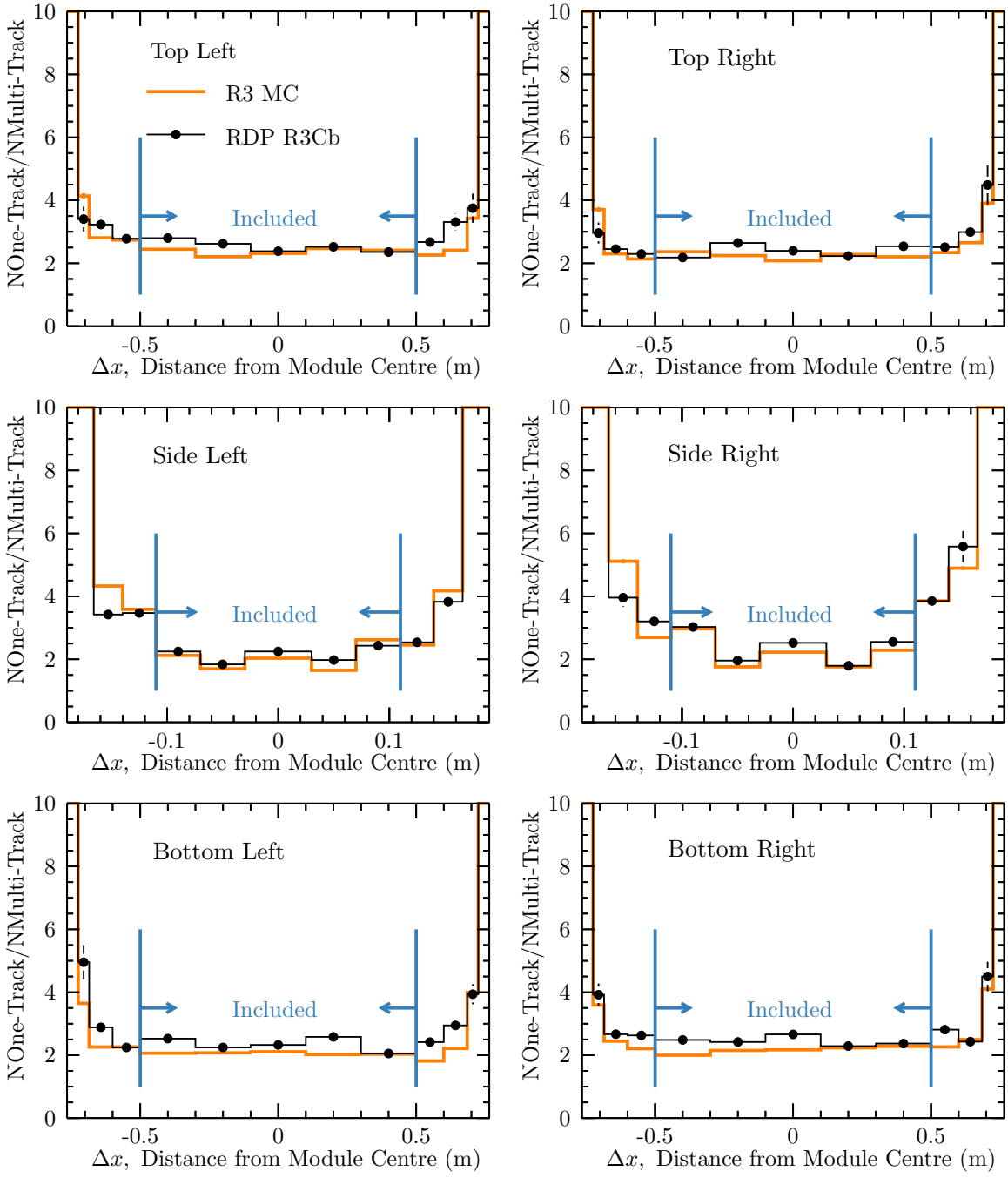


Figure 4.4: The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function x distance from the module centre, Δx. The regions marked as included constitute the fiducial volume for this analysis.

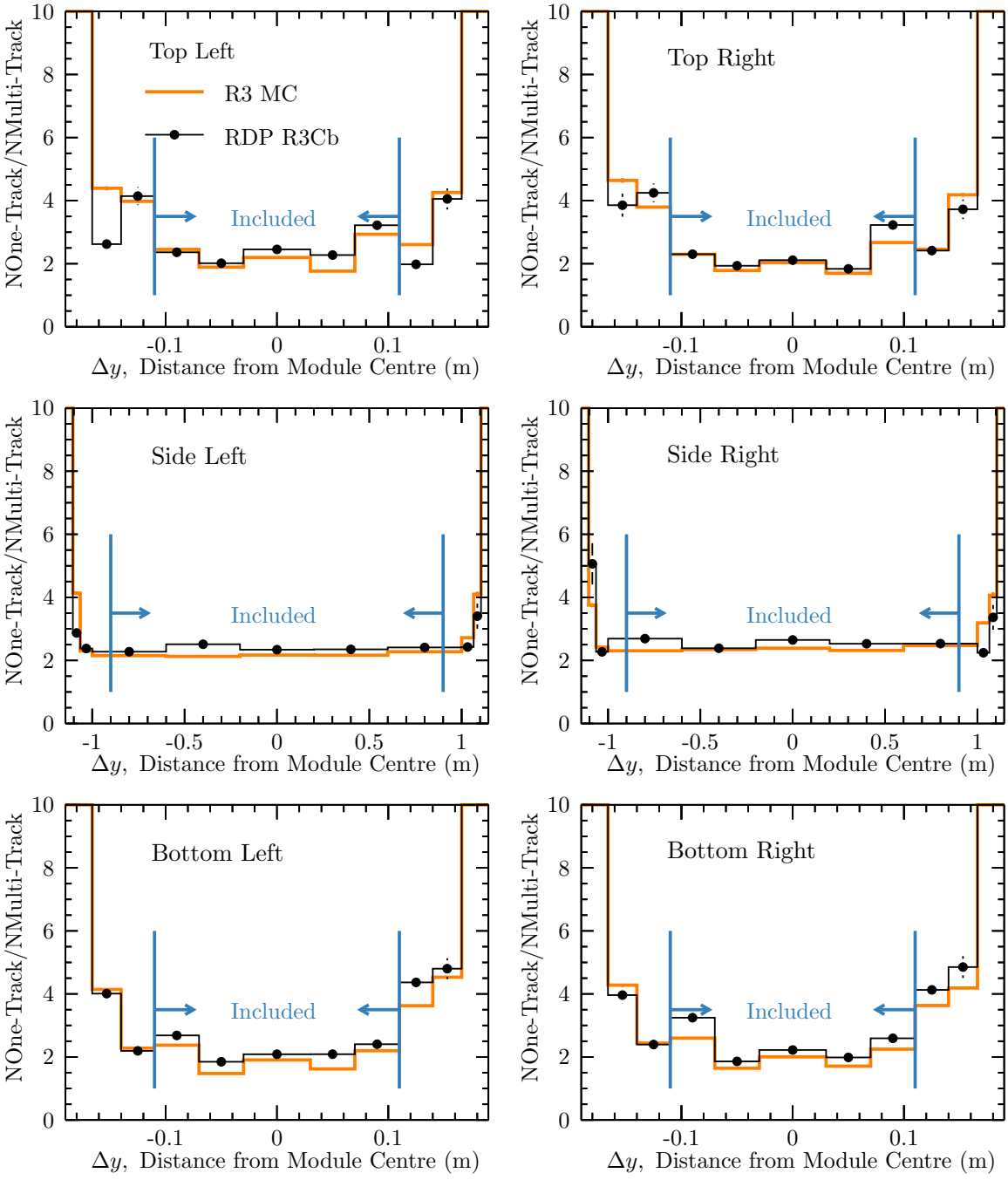


Figure 4.5: The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function y distance from the module centre, Δy. The regions marked as included constitute the fiducial volume for this analysis.

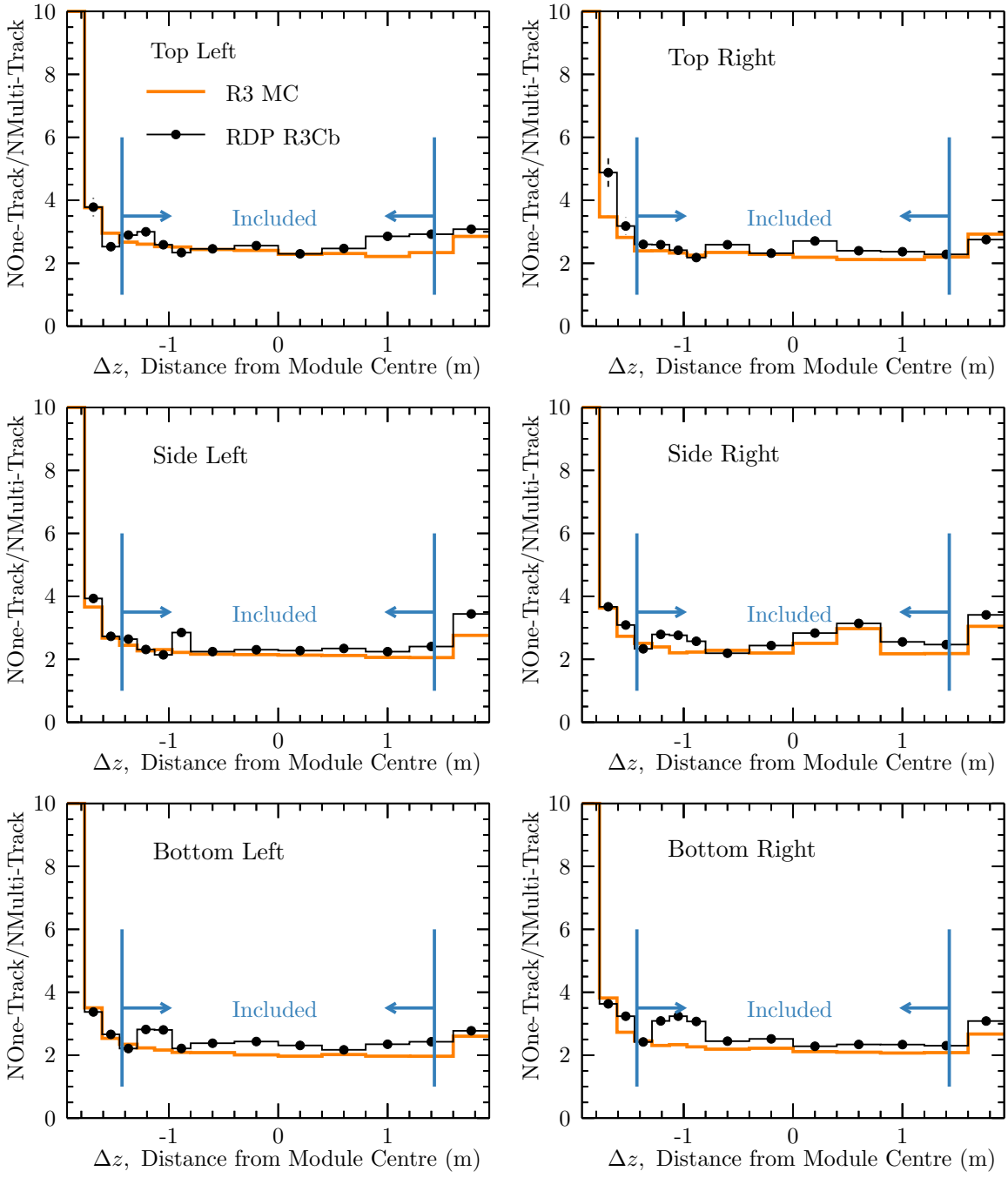


Figure 4.6: The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function z distance from the module centre, Δz. The regions marked as included constitute the fiducial volume for this analysis.

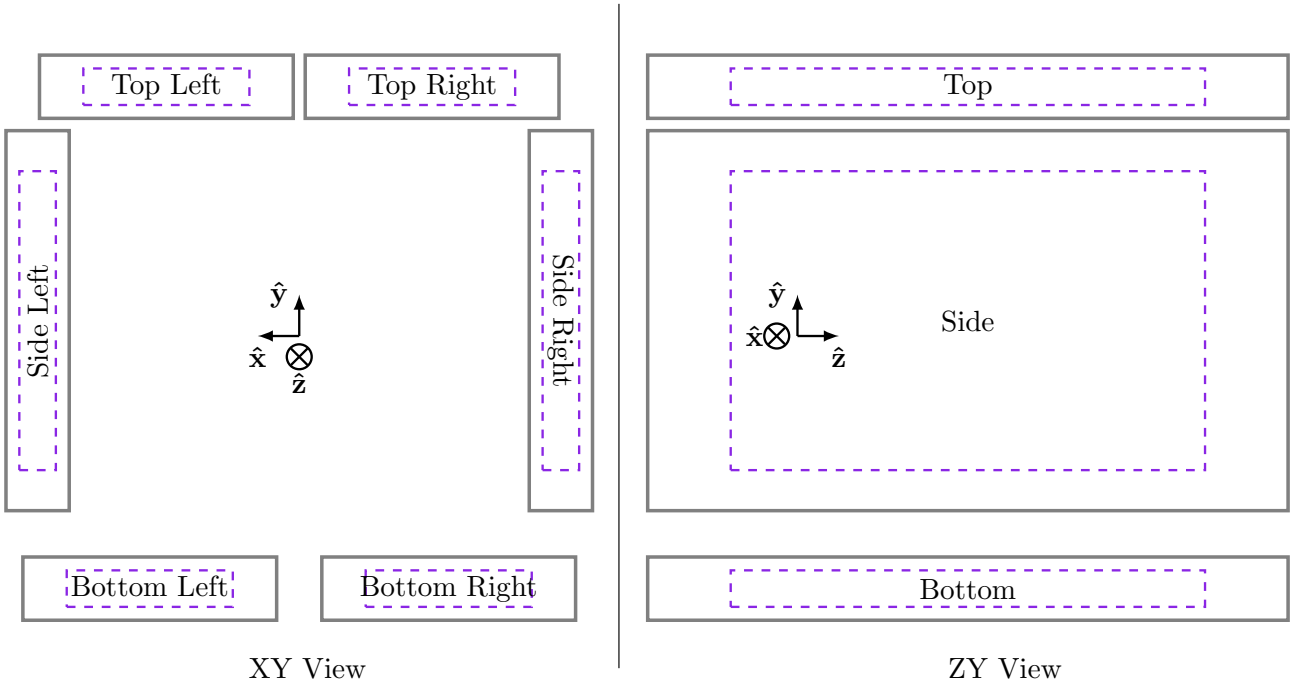


Figure 4.7: The fiducial volume definition used in this analysis. The solid grey regions show the ECal active volumes, and the dashed purple lines show the fiducial volume used to select interaction candidates.

4.2 The event selection

The aim of this analysis is twofold: to provide a robust measurement of interactions occurring within the ECal, and to attempt to isolate poorly modelled regions that have plagued previous ND280 ECal-target analyses. The fiducial volume is used to select high-quality ECal interaction candidates in regions of approximately uniform ECal reconstruction efficiency and entering background contamination. However, to highlight regions of poor data–MC correspondence, reconstructed deposits that fall outside the fiducial volume should not be simply discarded.

Four mutually exclusive event selections, designed to separate events based on the degree of confidence in reconstructed interaction position and available kinematic information, are described in the rest of this section. A cartoon of the example event topologies for the two ECal interaction candidate samples, the 'golden' and 'silver' event samples, is shown in Figure 4.8. The golden and silver samples differ by the existence of Tracker information associated to the ECal vertex candidate.

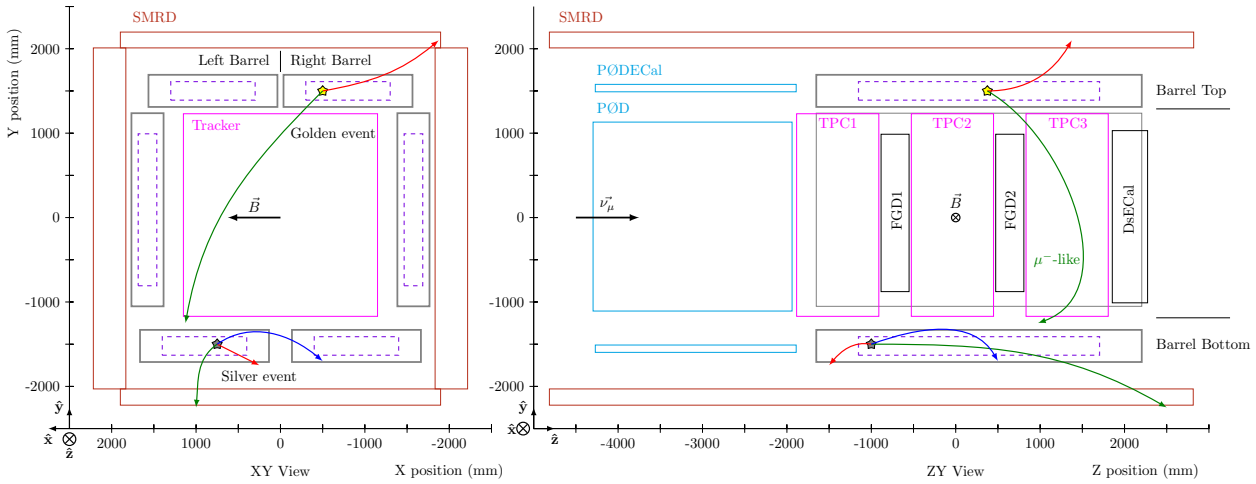


Figure 4.8: Examples of the two interaction candidate event topologies. Golden event candidates are those that have a reconstructed vertex position within the fiducial volume of the Barrel ECal (dashed lines) and have an associated muon-like Tracker track. Silver event candidates must also be reconstructed within the fiducial volume, but explicitly do not have any associated Tracker tracks. No extra restrictions are placed on the topologies of other associated tracks.

4.2.1 Golden events

Some of the most basic observables that can be used to constrain neutrino charged-current interaction models are the kinematic of the final state charged lepton. The integration of the ECal-isolated vertexing with ND280 global reconstruction allows the association of tracks with momentum and PID measurements from the tracker to ECal interaction candidates.

The golden event selection is designed to sample muon neutrino interactions, occurring within the ECal active volume, which produce a final state muon that was measured by the tracker. In each Barrel ECal module, some restricted region of initial azimuthal angle will result in a final state muon propagating towards the tracker.

Each pair of modules offers advantages and disadvantages for this selection. In the bottom modules, the magnetic field will bend forward-going negative muons away from the Tracker; but backward particles would be bent towards the TPC. The bottom modules are closer to the beam axis and thus see a higher interaction rate. The side modules have the largest fiducial mass, but a global momentum bias apparent for muons starting in the side modules was presented in 3.3. The right side module sustained damage to two TFBs in The Great East

Japan Earthquake, which renders two sections of it inactive. The top modules see the lowest interaction rate as they are furthest from the beam axis, but there is very little inactive space between the module edges and the tracker region compared to the other four modules, which have readout and control electronics between them and the Tracker.

The next section discusses the geometric and electro-magnetic detector acceptance for muons produced within the Barrel ECal to pass through an active region of a TPC. This complements the discussions on reconstruction efficiency presented throughout 3.

TPC acceptance

The Inward Muon particle gun sample (*c.f.* § 3.1.4) was used to assess the TPC detector acceptance for muons produced in ECal interactions. The detector acceptance should be approximately a binary function, a muon propagating from a given position with a given three momentum either will, or will not, pass through the TPC-active volume[3]. For each point in the muon initial-state phase space, the detector geometry and electro-magnetic environment define an exact trajectory (ignoring stochastic effects). When projecting the acceptance into a lower dimension, variations in any integrated-over dimension of phase space can lead to a non-binary acceptance distribution. The acceptance in some projection is defined as

$$\text{Acceptance}(X_i) = \frac{N^{\text{TPC}}(X_i)}{N(X_i)}, \tag{4.1}$$

where $X_i, i \in \{x, y, z, p, \phi, \theta\}$ is a relevant initial state property, $N^{\text{TPC}}(X_i)$ is the number in the sample with a true muon passing through the TPC active region, and $N(X_i)$ is the total number in the sample with a given initial state property, X_i. For simplicity, the position within a module will be always integrated over. This is a reasonable simplification as the fiducial volume precludes most edge effects from the reconstruction. The muon TPC-active acceptance fractions, projected onto initial momentum, polar angle, and azimuthal angle, are shown in Figure 4.9, Figure 4.10, and Figure 4.11 respectively.

[3] Stochastic effects may also result in a probabilistic acceptance.

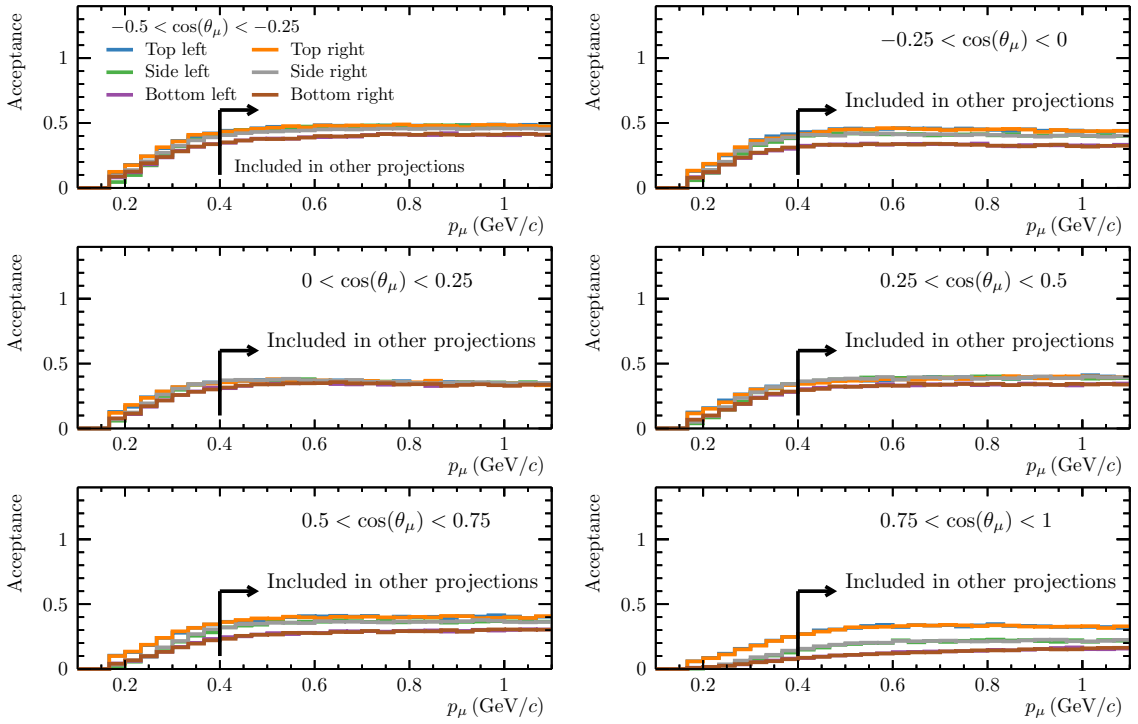


Figure 4.9: The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon momentum. The muon momentum restriction shown in this figure is only applied to the other projections of this sample, it is not used as a selection cut for the interaction candidate selections.

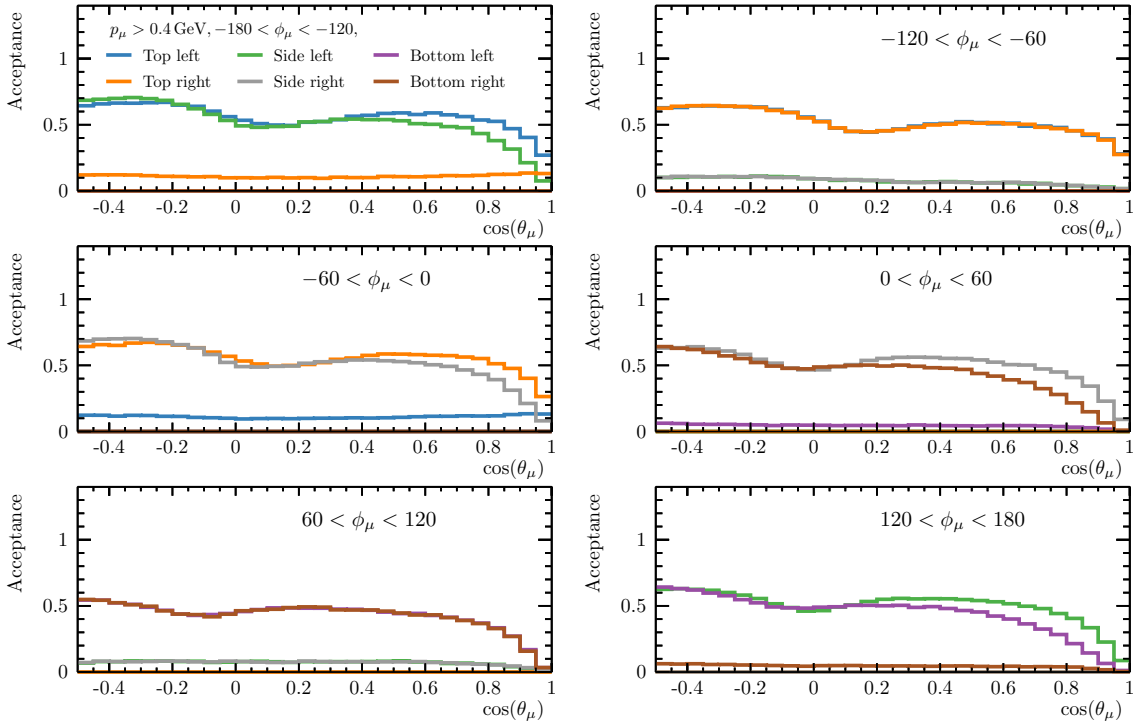


Figure 4.10: The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon cosine polar angle. The momentum of muons included is restricted to $p_\mu > 0.4\,\text{GeV}/c$ to avoid averaging over a varying momentum acceptance.

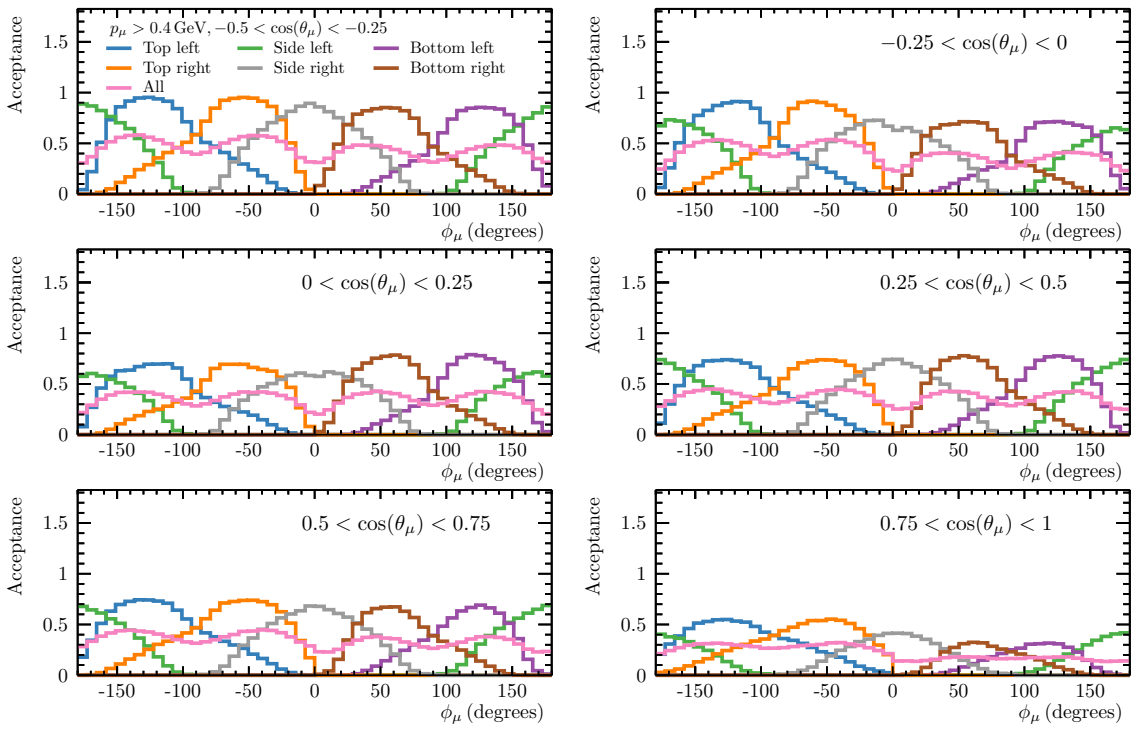


Figure 4.11: The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon azimuthal angle. The momentum of muons included is restricted to $p_\mu > 0.4\,\text{GeV}/c$ to avoid averaging over a varying momentum acceptance.

Due to the geometry of the ECal modules, forward-going, energetic muons will not curve into the TPC and so have an expectedly low acceptance, otherwise the acceptance is good. For future analyses that try to de-convolve the detector effects, the flat acceptance for $p_\mu > 0.4\,\text{GeV}/c$, and the slowly varying acceptance for $\cos(\theta_\mu) \lesssim 0.75$ show promise. To provide robust data for the constraint of interaction models, no attempt should be made to measure, or 'correct for', regions of phase space outside of the detector acceptance. For example, understanding the rate at which true muons with $\cos(\theta_\mu)_{\text{True}} > 0.75$ are reconstructed as $\cos(\theta_\mu)_{\text{Recon}} < 0.75$ and correcting for it, is more robust than trusting the simulation to 'fill in' for the majority of the events in the forward region. As all data–simulation comparisons presented herein are of reconstructed observables, consideration of such subtleties is not necessary.

Selection cuts

The selection steps for all four samples are summarised in Figure 4.12. While the selection is presented as a flow, all cuts are fully independent and the ordering of the cuts has no impact

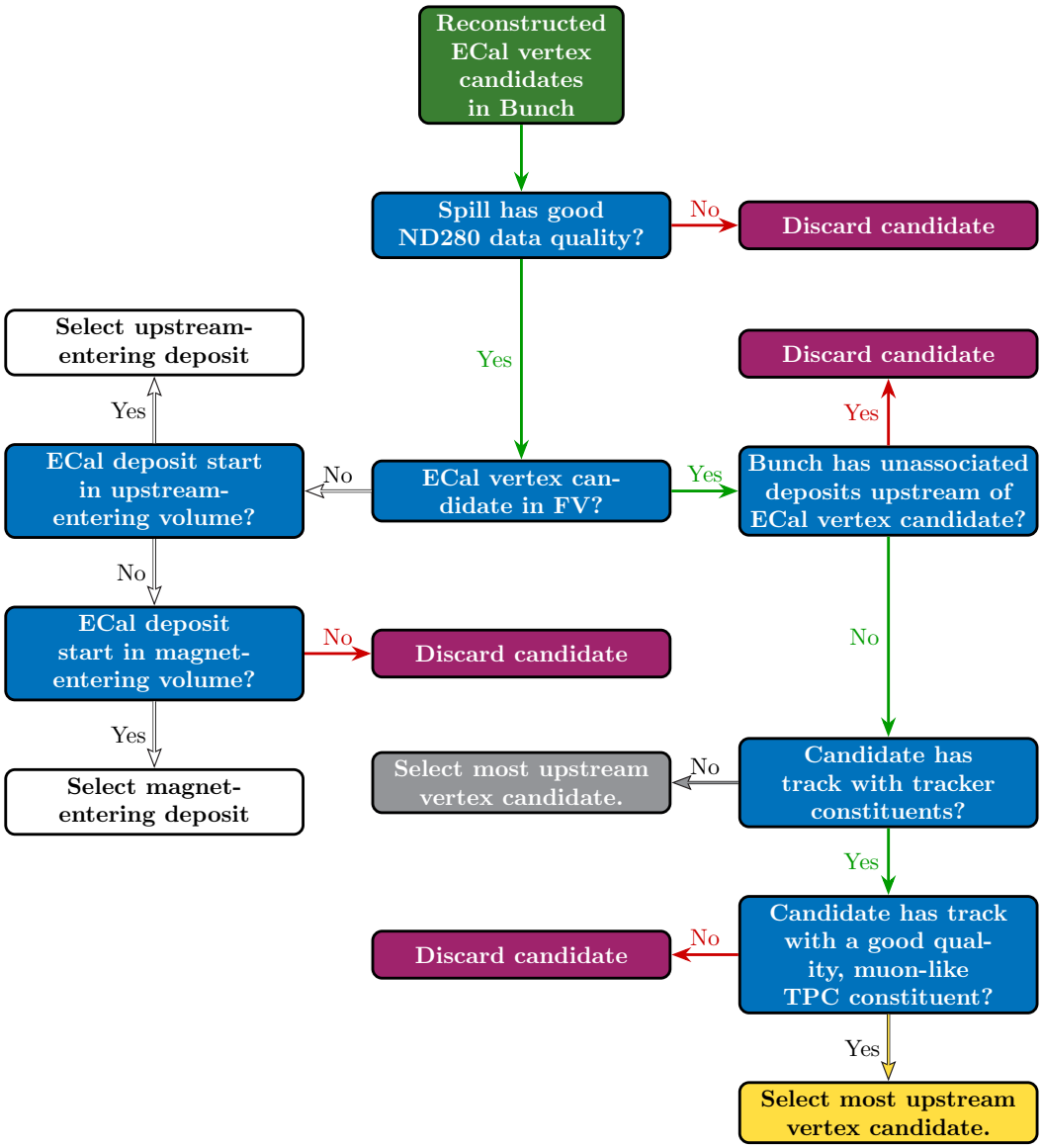


Figure 4.12: A flow schematic describing the selection steps leading to the four analysis samples: magnet- and upstream-entering backgrounds and silver and golden ECal interaction candidates.

on how candidates are selected or discarded from each sample.

The first stage of the selection sorts the vertex candidates into bunches, which correspond to the bunch structure of the J-PARC proton beam pulses (*c.f.* § 2.1). The timing window for each bunch is defined relative to a beam spill trigger sent by the J-PARC neutrino beamline. Candidates that did not fall within the time span of a bunch were discarded; this occurred for less than 1% of candidates. For a variety of reasons, the quality of the data recorded by ND280 may be inadequate for analysis because of the operational state of the detector. This is decided after the data have been recorded and the 'data quality' flags are checked on a spill-by-spill basis. Vertex candidates reconstructed within spills that were flagged as having bad ND280 data quality were discarded.

There are six upstream veto flags, each of which tags deposits that occur upstream of the vertex candidate and have no association to it. The veto cuts were kept separated by detector system to allow for a granular assessment of any discrepancy in the rate that they fire. A veto was used for deposits starting in the PØD, the PØDECal, the Tracker, the Barrel ECal, the SMRD, or any deposits reconstructed as starting outside of an active detector[4]. The Tracker track quality cut, and muon particle identification threshold are standard ND280 analysis cuts and are discussed and motivated in § 4.6 of Ref. [98]. The second fiducial volume cut, placed on the Tracker track start position was chosen to remove poorly-reconstructed vertex tracks. If the interaction did take place within the ECal, but the reconstruction matched two ECal tracks left by different particles, the global momentum would be biased due to incorrect energy loss compensation. The other likely possibility is that global track matching made the correct decision and the reconstructed ECal deposit was left by a curving, or re-scattering entering charged particle. In this case, the cut will increase the purity of the selected events. The frequency with which such matches occur is sensitive to the hadronic final state distributions, as well as the re-scattering probability for entering particles and may be imperfectly modelled.

[4]This can happen due to reconstruction failure or a vertex reconstructed from two or more entering tracks may be positioned outside an active region

Selection performance

The effect of the six positional upstream veto cuts are presented in Figures 4.13 to 4.18. These are so-called '$N-1$' plots, which show the effect of each cut as if it was the last cut in a sequence . The only event candidates that enter each plot are those that pass all cuts other than the one on the dimension presented. In this way, the effect of each individual cut on the purity and efficiency of the selection is presented. Each veto removes a region of lower purity than the selected sample, with the possible exception of the SMRD veto. The choice to include data in these figures was to allow for any selection-breaking discrepancies to be highlighted before looking at the analysis data sample[5]. While on average the veto rate for data is slightly higher, the agreement is certainly good enough to proceed. Each positional veto will remain in use when analysing the full data set.

To show the effect of the fiducial volume cut, the cut in each coordinate axis is separated so that when investigating the effect of the cut in x, the cut for y and z are applied. The '$N-1$' plots for the x, y, and z axes of the fiducial volume cut are shown in Figure 4.19, Figure 4.20, and Figure 4.21 respectively.

The shape observed for some modules for some of the individual coordinate axes can be seen to be due to the requirement of a Tracker track. The '$N-2$' plot for the x fiducial volume restriction, where the presence of a Tracker track is no longer predicated, can be seen in Figure 4.22.

Each of the above '$N-1$' plots show the purity of the final selection; the selection efficiency and purity are summarised in Figure 4.23. The effect of each selection cut is displayed by presenting the efficiency and purity with that cut *ignored*. The efficiency is defined, with respect to the acceptance discussed in the previous chapter, as

$$\text{Efficiency}(\mathrm{m}_i) = \frac{N^{\text{Selected, Signal}}(\mathrm{m}_i)}{N^{\text{TPC}}(\mathrm{m}_i)}.$$

The Barrel ECal module is iterated by m_i and $N^{\text{Selected, Signal}}(\mathbf{x})$ is the number of selected true

[5] Unfortunately, it didn't perform this task particularly well *c.f.* § 6.2

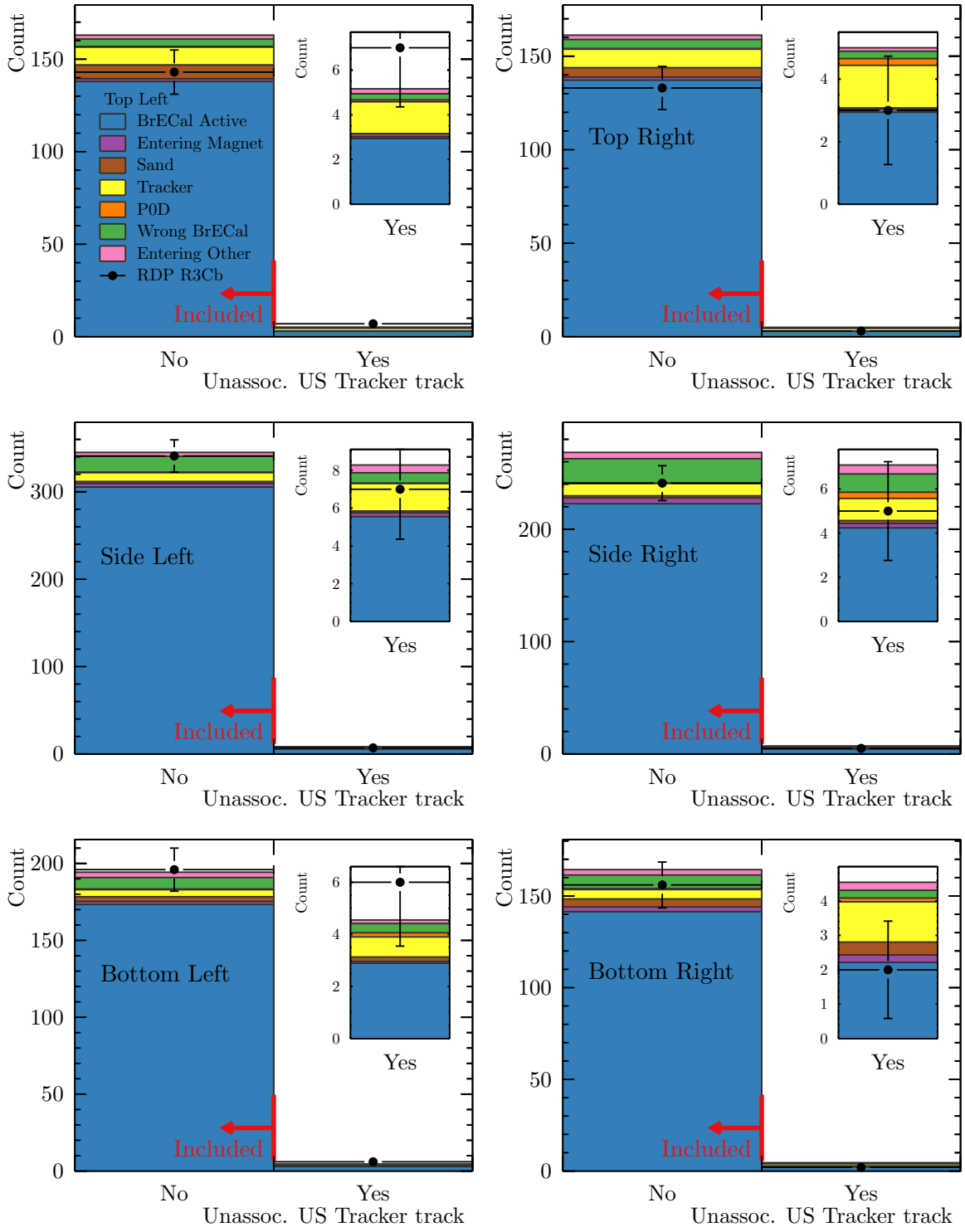


Figure 4.13: The effect of the tracker upstream veto on the golden candidate sample. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

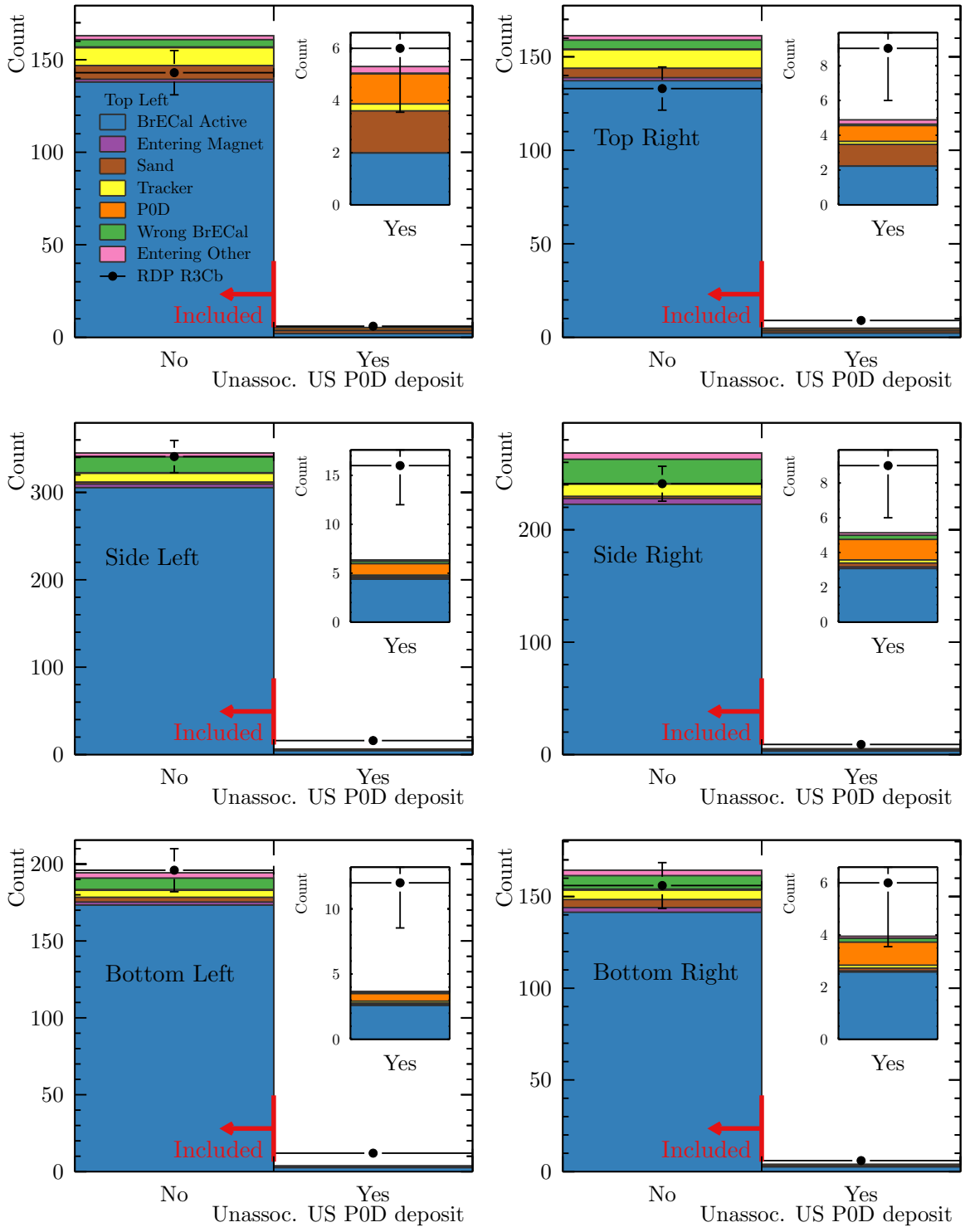


Figure 4.14: The effect of the upstream PØD veto on the golden candidate sample. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

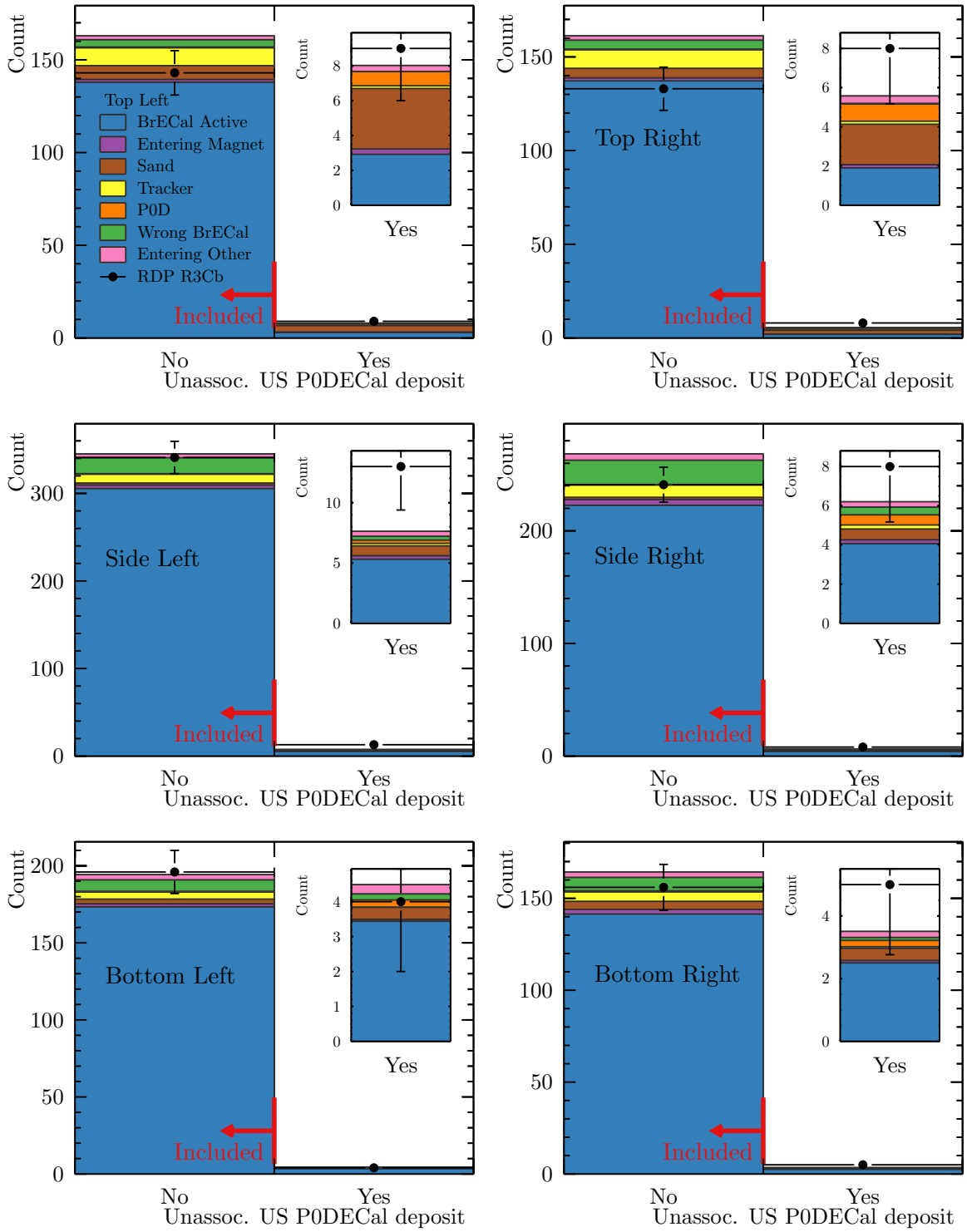


Figure 4.15: The effect of the upstream PØDECal veto on the golden candidate sample. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

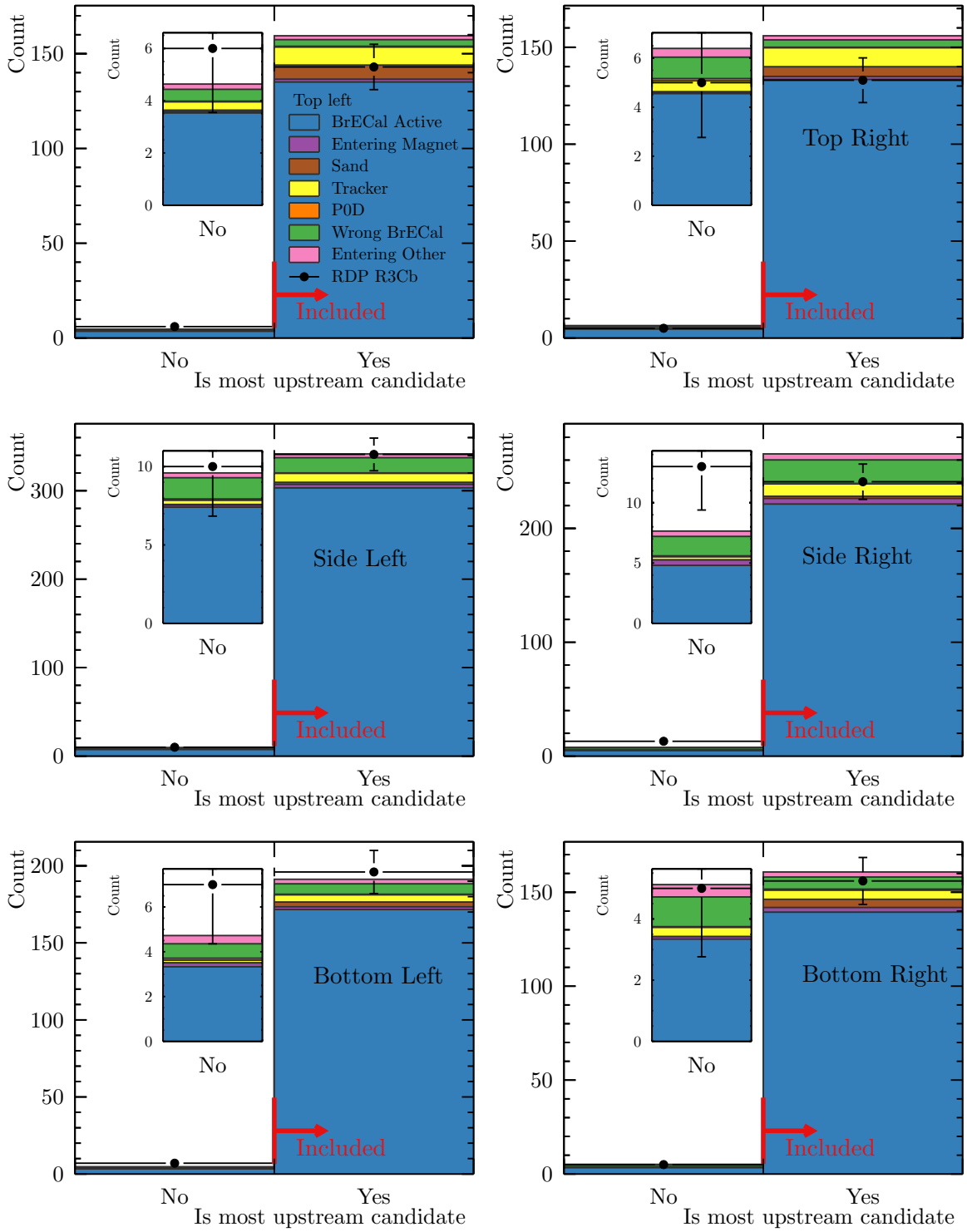


Figure 4.16: The effect of the upstream Barrel ECal veto on the golden candidate sample. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

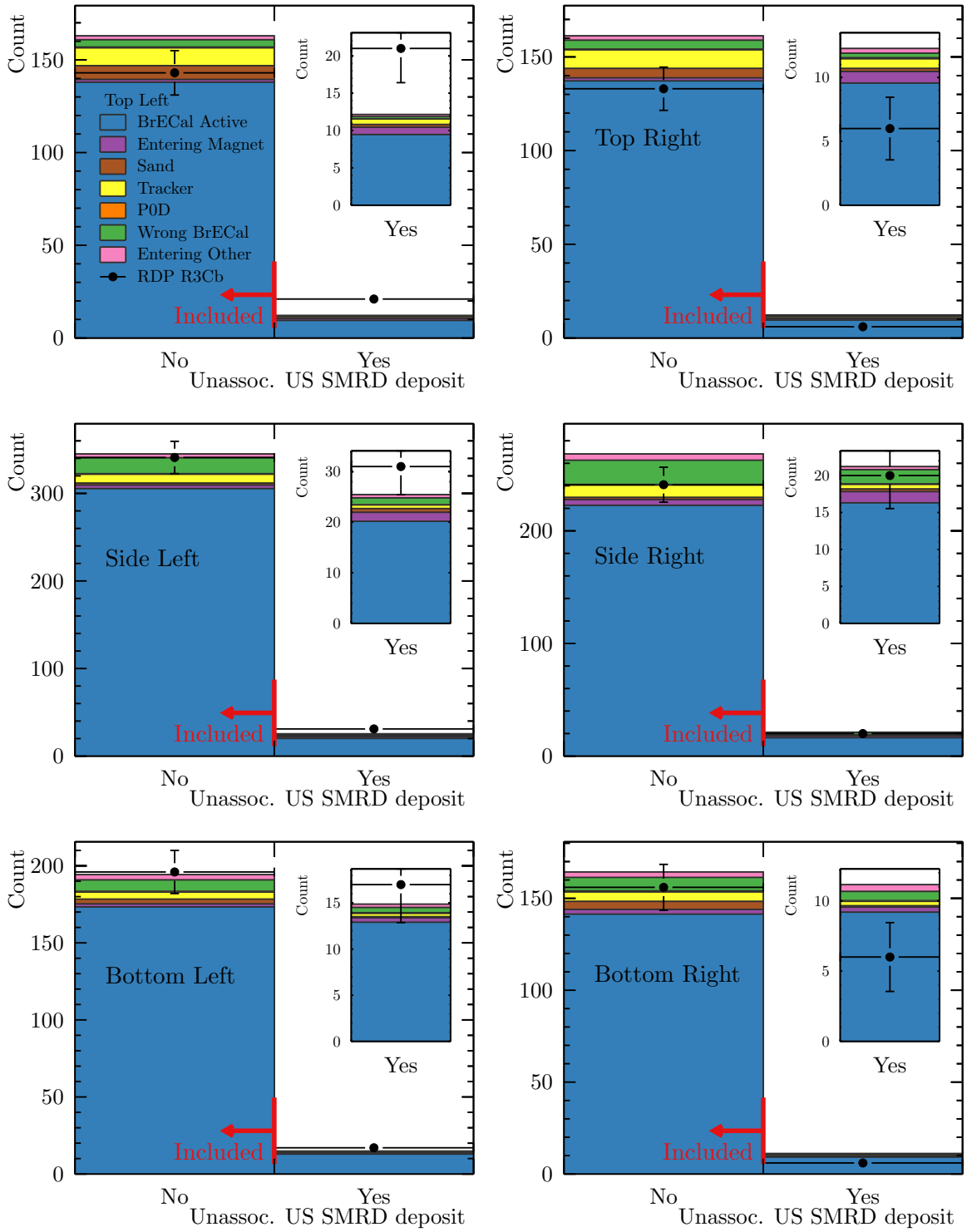


Figure 4.17: The effect of the upstream SMRD veto on the golden candidate sample. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

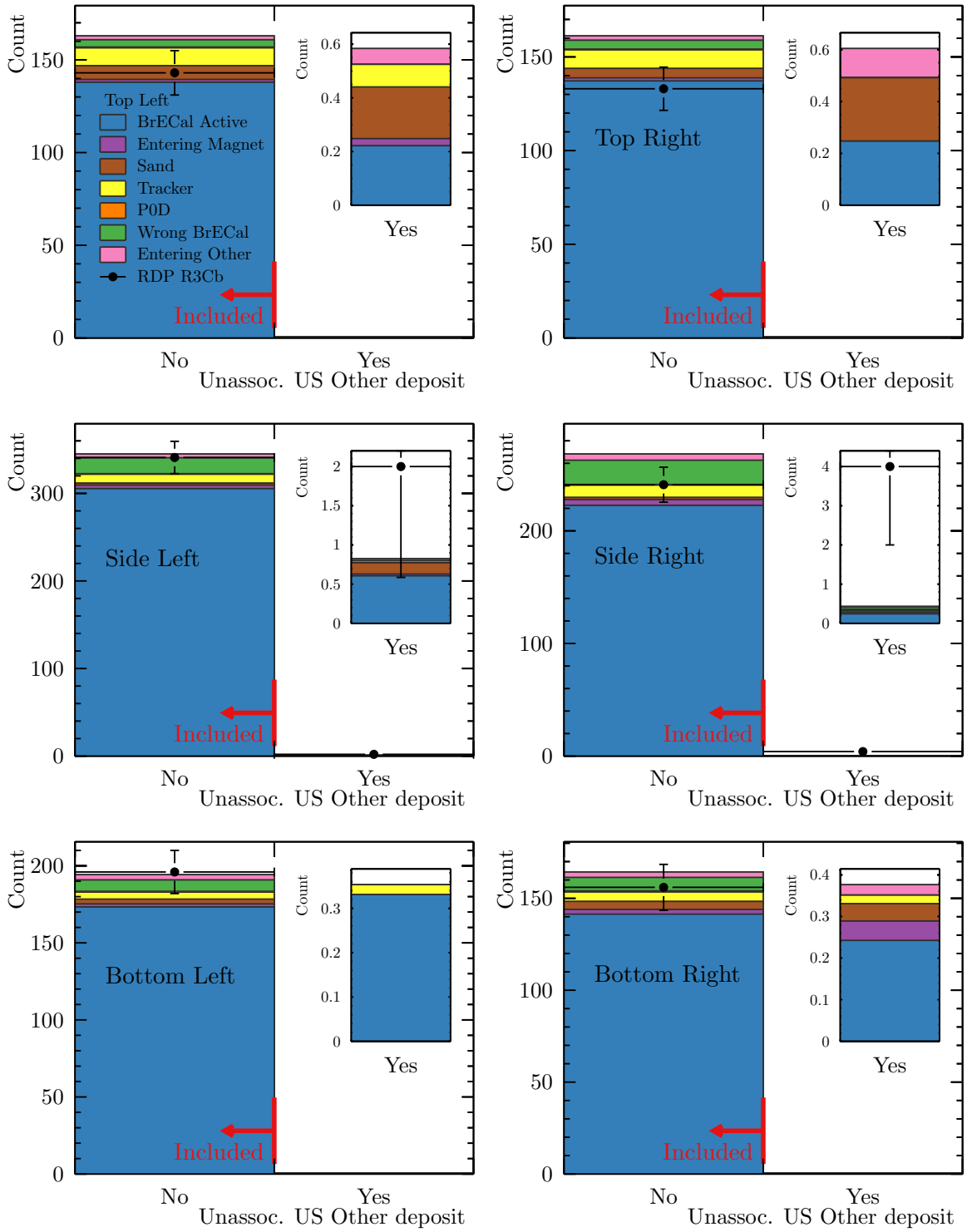


Figure 4.18: The effect of the upstream 'other' veto on the golden candidate sample—deposits reconstructed to start outside the any active detector region. All other selection cuts have been applied. The inlay figure in each pane shows a magnification of the excluded region.

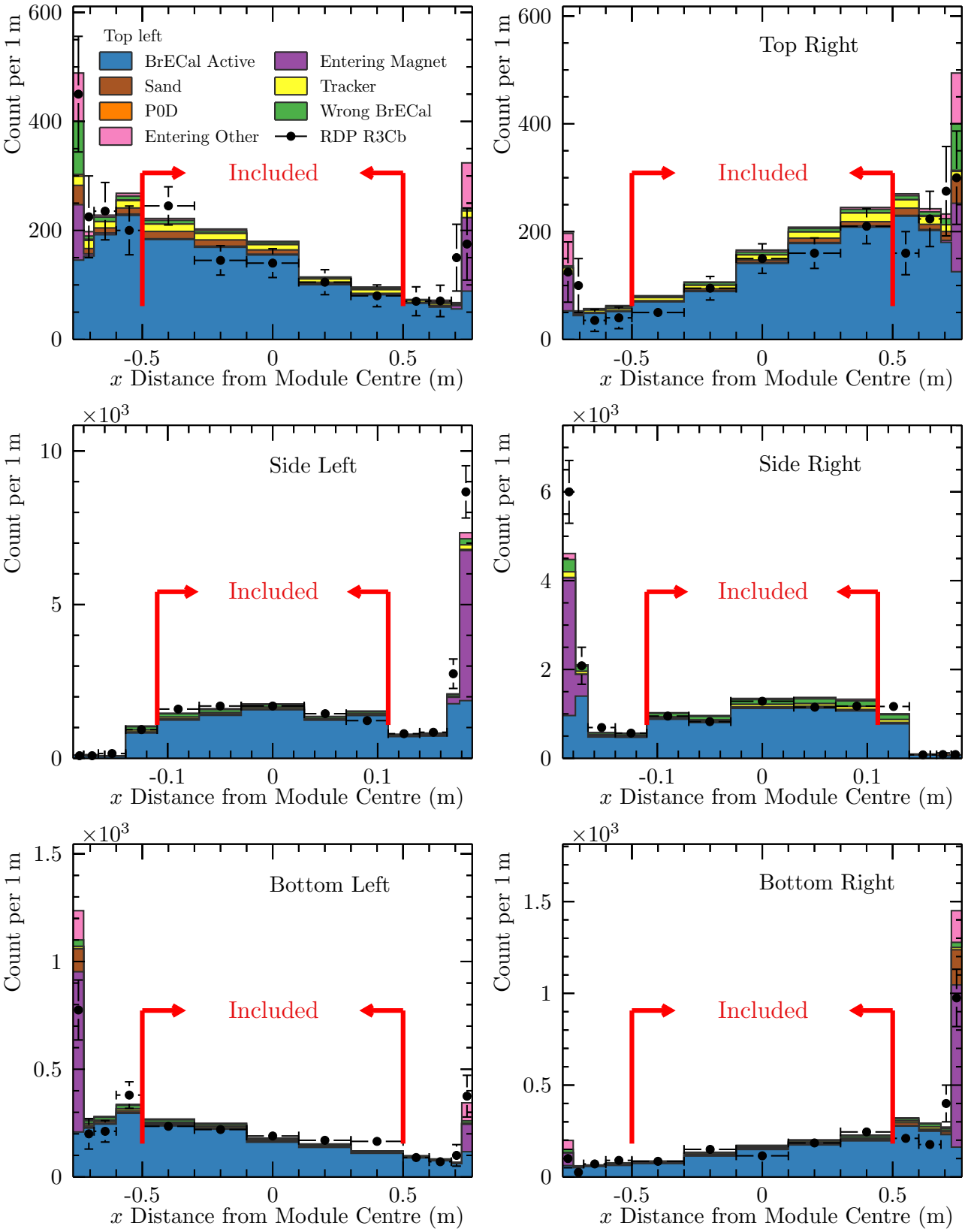


Figure 4.19: The effect of the fiducial volume cut in the x direction on the golden candidate sample. All other selection cuts, including the fiducial volume cut in the y and z directions, have been applied.

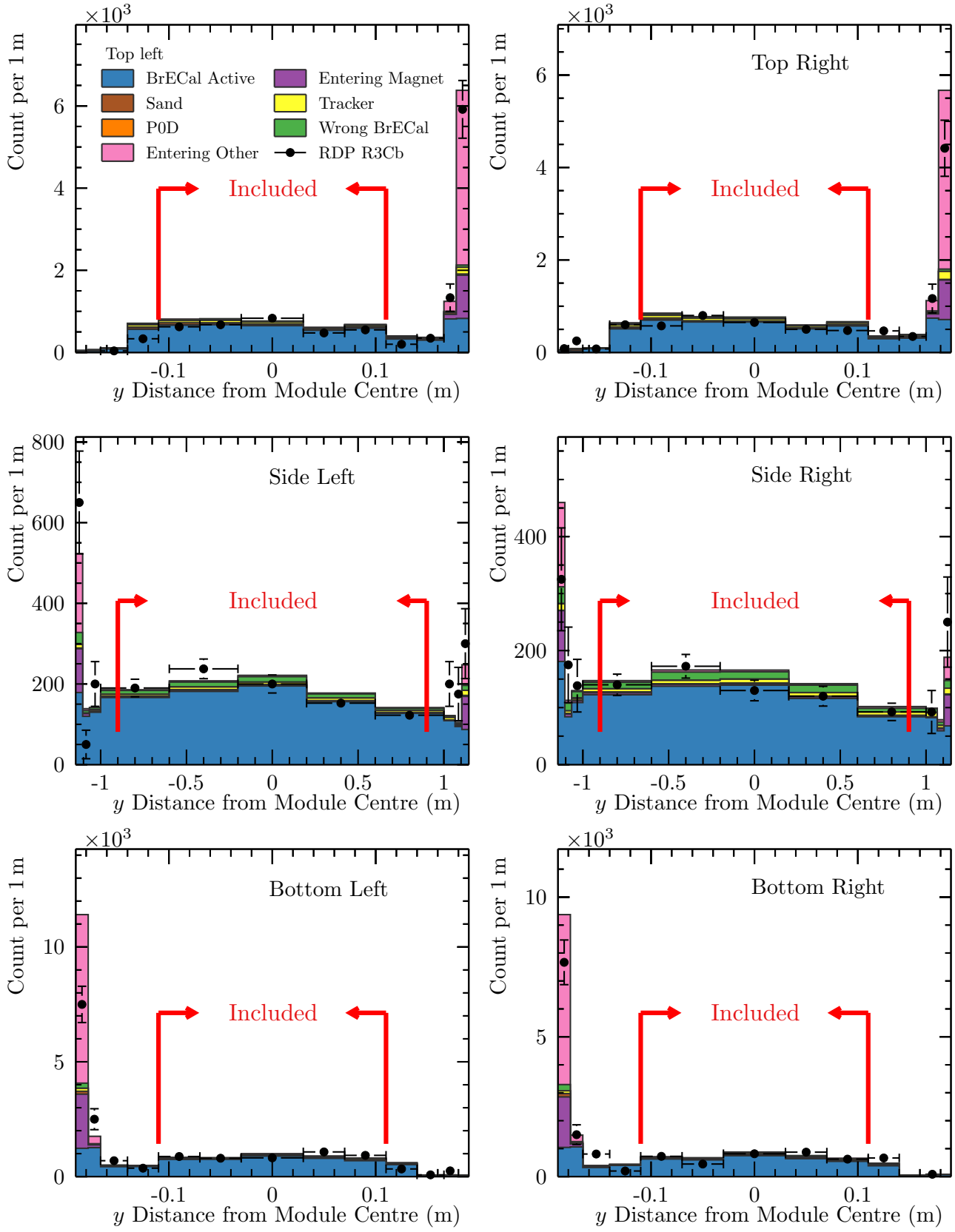


Figure 4.20: The effect of the fiducial volume cut in the y direction on the golden candidate sample. All other selection cuts, including the fiducial volume cut in the x and z directions, have been applied.

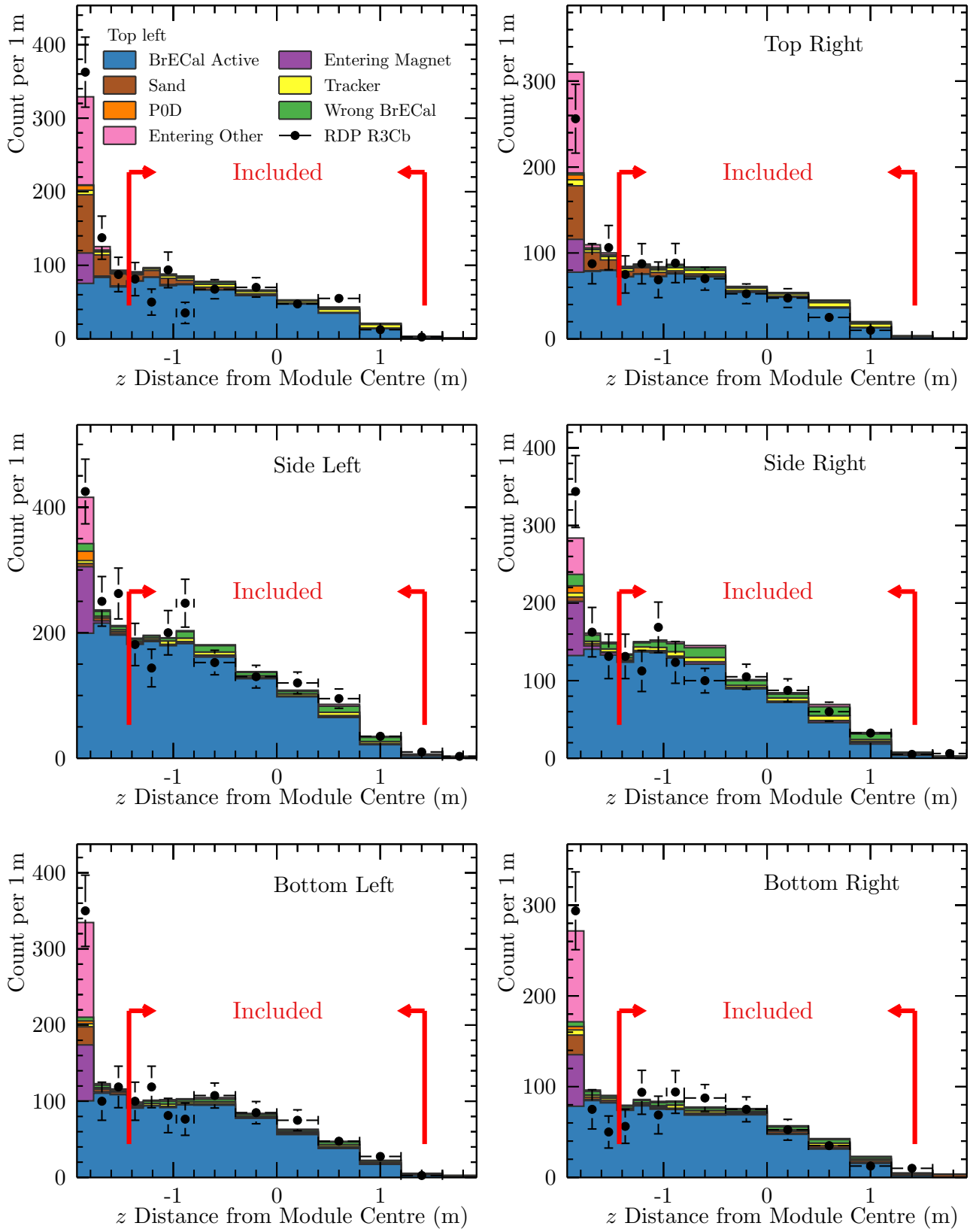


Figure 4.21: The effect of the fiducial volume cut in the z direction on the golden candidate sample. All other selection cuts, including the fiducial volume cut in the x and y directions, have been applied.

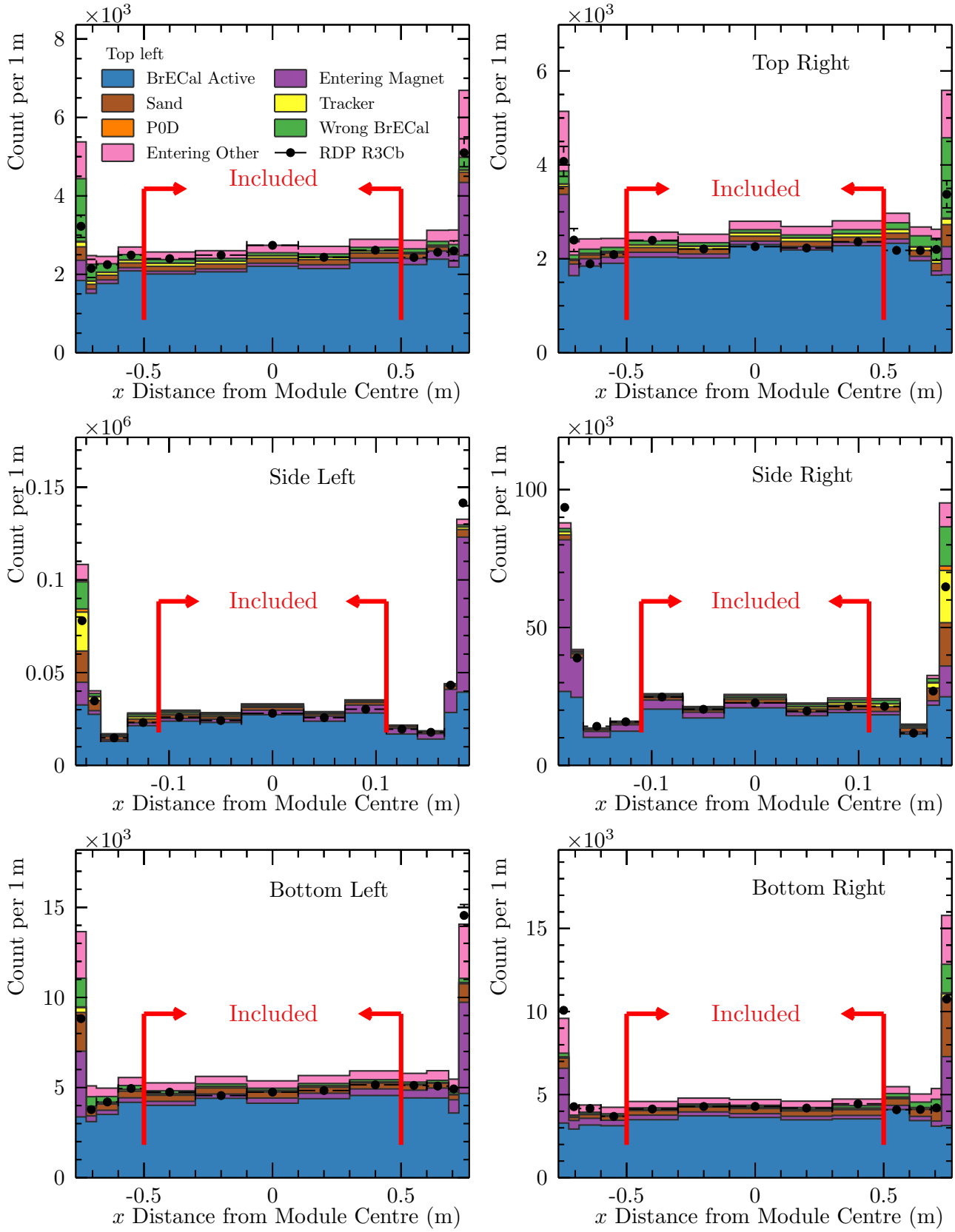


Figure 4.22: The effect of the fiducial volume cut in the x direction on the golden candidate sample. All other selection cuts have been applied except for the requirement of a Tracker track association. By comparison to Figure 4.19, it is clear that the Tracker track requirement imposes the shape on the included region.

interactions that occurred in, and were reconstructed in, the active volume of ECal module, m_i. $N^{\mathrm{TPC}}(\mathrm{m}_i)$ is defined as in 4.1. The purity is defined as

$$\mathrm{Purity}(\mathrm{m}_i) = \frac{N^{\mathrm{Selected,\ Signal}}(\mathrm{m}_i)}{N^{\mathrm{Selected}}(\mathrm{m}_i)},$$

where $N^{\mathrm{Selected}}(\mathrm{m}_i)$ is the number of selected interactions that were reconstructed in m_i. As could have been qualitatively determined from the '$N-1$' plots above, the effect on the selection from the veto cuts and the Tracker track fiducial volume cut are marginal. The majority of the selection power comes from the fiducial volume cut along the running coordinate in each ECal module, which cuts out entering background from the magnet flux return. These cuts also have the largest impact on efficiency as they discard the largest mass fraction from each module. A 'projected' efficiency increase is calculated for each dimension of the fiducial volume cut. The average selection efficiency in that module is multiplied by the relevant mass ratio in Table. 4.2. This effectively extrapolates the full selection efficiency over the fiducial volume into the selection volume, which has been increased by the relaxation of one dimension of the fiducial volume cut. If the calculated efficiency increase is less than the projected efficiency increase, then the fiducial volume cut in that dimension removes a region of lower average selection efficiency than the average selection efficiency in the fiducial volume. For each dimension, in each module, the calculated efficiency gain is less than or equal to the projected gain.

While the overall selection efficiency is low, it is known that for large regions of muon kinematic phase space the tracking and matching efficiencies are good and reasonably flat (*c.f* § 3.1.2, § 3.2.1, and § 3.2.2). The large target mass of the Barrel ECal modules will result in a high interaction rate, therefore a flat, well understood efficiency is more important than a high efficiency. Furthermore, the simulation-predicted purity of the selected sample is about 80%, and consistent across the modules. This bodes well for the future use of such a sample to constrain both the flux and interaction model by utilising the degeneracy-breaking off-axis angle variation across the detector.

The next section presents the content of the golden candidate sample for the simulation and control sample data.

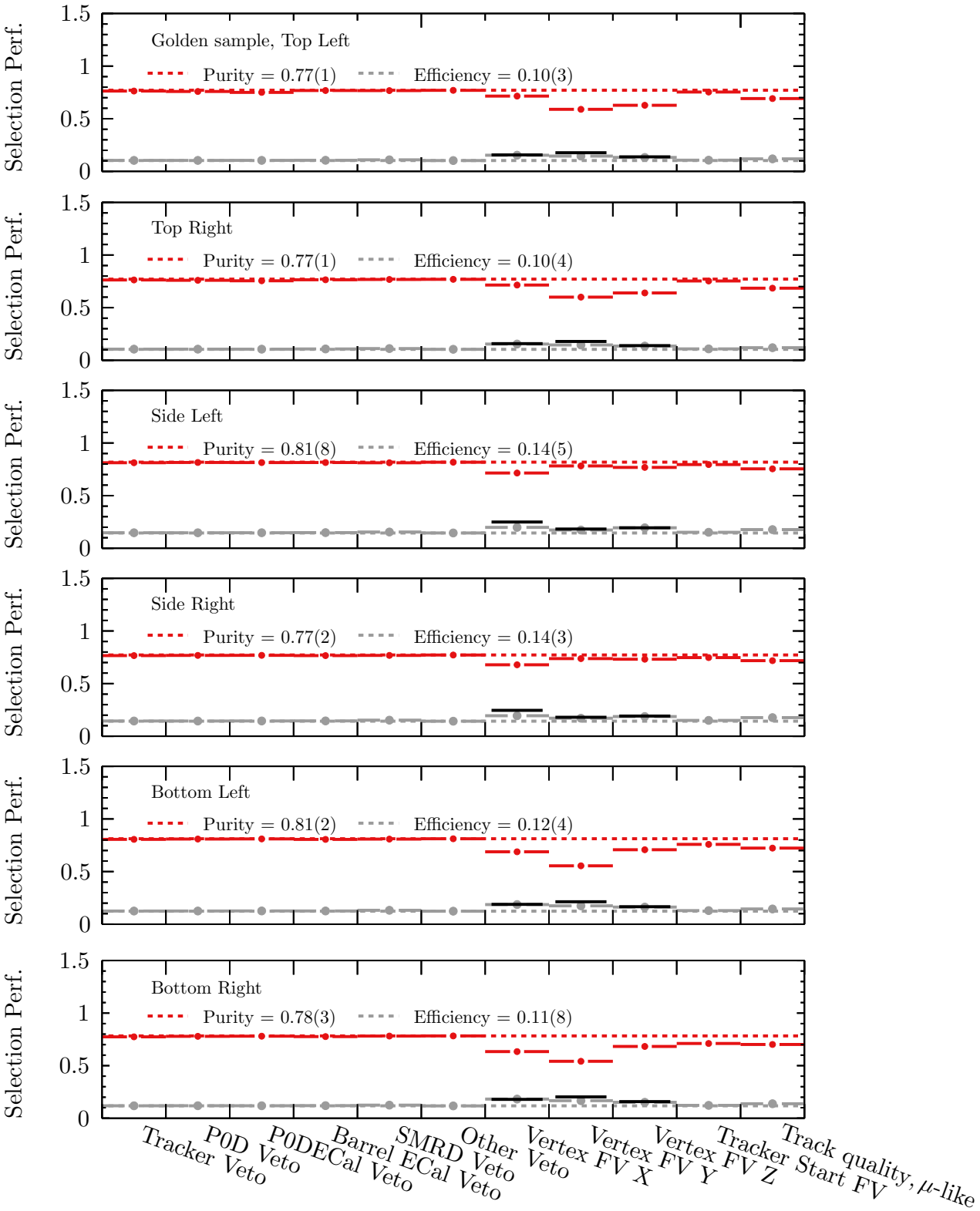


Figure 4.23: The effect of each cut on the efficiency and purity of the selection. The difference between the selection performance with each cut individually ignored and the full selection performance (dashed line) highlights the effect of a single cut with all other cuts accounted for. The black lines for each dimension of the fiducial volume cut show the naive expectation for the efficiency gain when ignoring one dimension of the fiducial volume restriction. This naive gain is calculated by extrapolating the selection efficiency over the fiducial mass to the selection mass available when ignoring one dimension. The mass ratios can be found in Table. 4.2

Sample content

The double differential event rate for the golden sample is shown in Figure 4.24. The binning was chosen to be more granular about the peak rate at $p_\mu \simeq 0.6$ GeV/c, but with momentum bins no smaller than 100 MeV/c to reduce the sensitivity to event migration between non-neighbouring bins because of the reconstruction smearing. This distribution is summed over all of the Barrel ECal modules; when comparing to the full data-set, data–simulation discrepancies may be more visible in distributions of candidates from individual modules. The normalisation of the simulated prediction, for a high purity sample of ECal-target events, is satisfactory. Without including systematic uncertainty estimates for the simulation, it is hard to quantitatively compare the shape of the simulation and the data. This is the focus of the next two chapters. The large statistical uncertainty also renders insightful, qualitative statements difficult to make. However, there are no particularly worrying regions of the two-dimensional phase space.

The reconstruction smearing of global track parameters, in an approximately isotropic muon sample, has already been presented in § 3.3. It is interesting to investigate the effect of smearing for the full neutrino beam simulation sample, in a relevant binning. The one-dimensional truth content matrices for the momentum and polar angle binning used in Figure 4.24 are shown in Figure 4.25, and Figure 4.26 respectively. In the truth content matrix, each row is normalised so that the sum of the values along the row is one. Each row shows the fraction of candidates in one reconstructed phase space bin that were generated in each true bin. Perfect reconstruction would result in a diagonal distribution, with smearing causing contributions to the near-diagonal regions. Only interactions that were correctly reconstructed as occurring within a Barrel ECal module are included. The binning for the 'true' and 'reconstructed' variables need not be the same when constructing such a matrix; here, a more granular true axis would not offer any more insight. These distributions cannot depict two-dimensional migrations, however, they fulfill the purpose of gaining an intuitive understanding of the kinematic smearing in the relevant binning. For momentum $p_\mu \gtrsim 0.4\,\text{GeV}/c$, the matrices are diagonally dominated—as may have been expected from Figure 3.26. It is clear from Figure 3.27 that the angular reconstruction performs

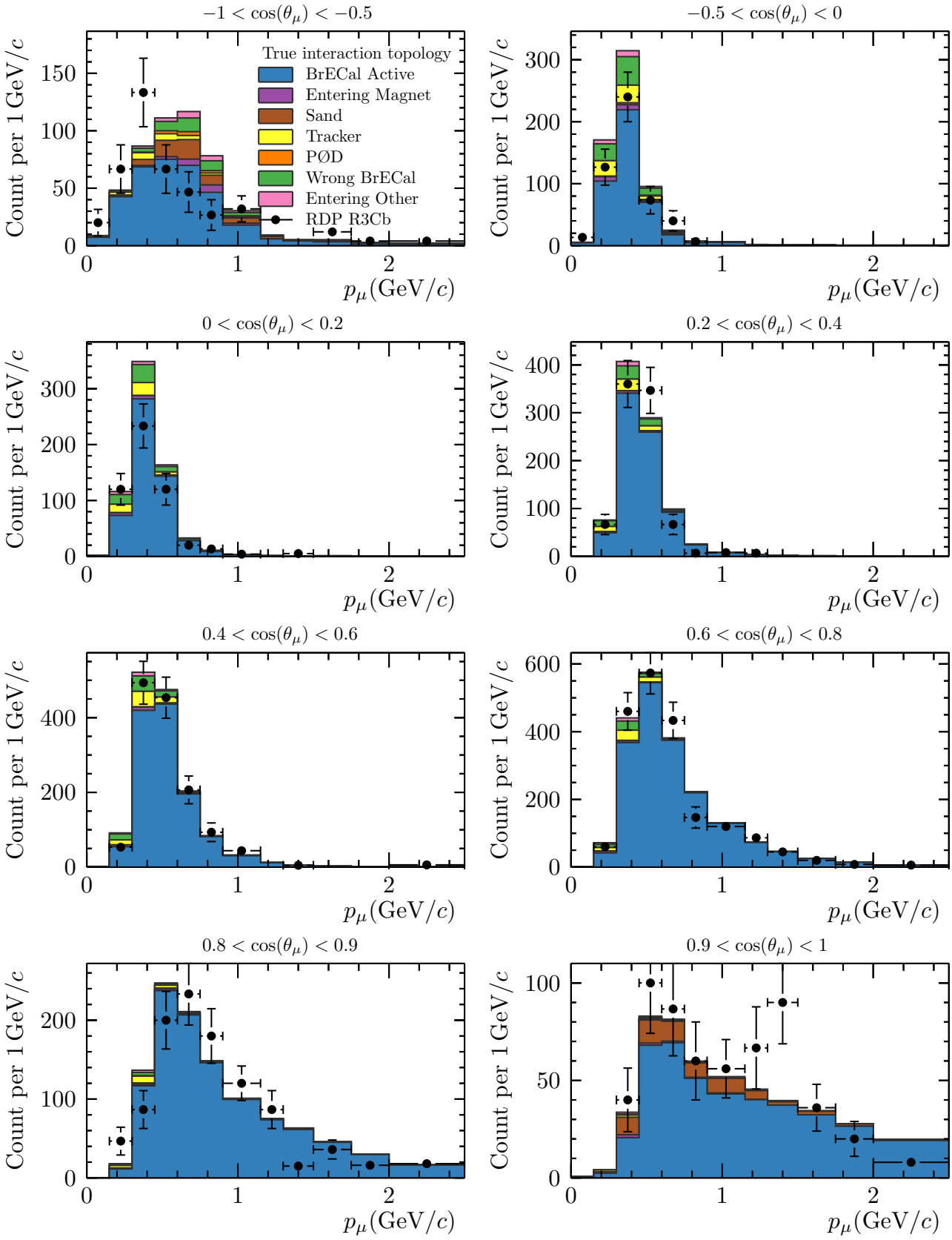


Figure 4.24: The contents of the golden interaction candidate sample projected onto muon momentum and polar angle. The simulation prediction is subdivided by true interaction topology.

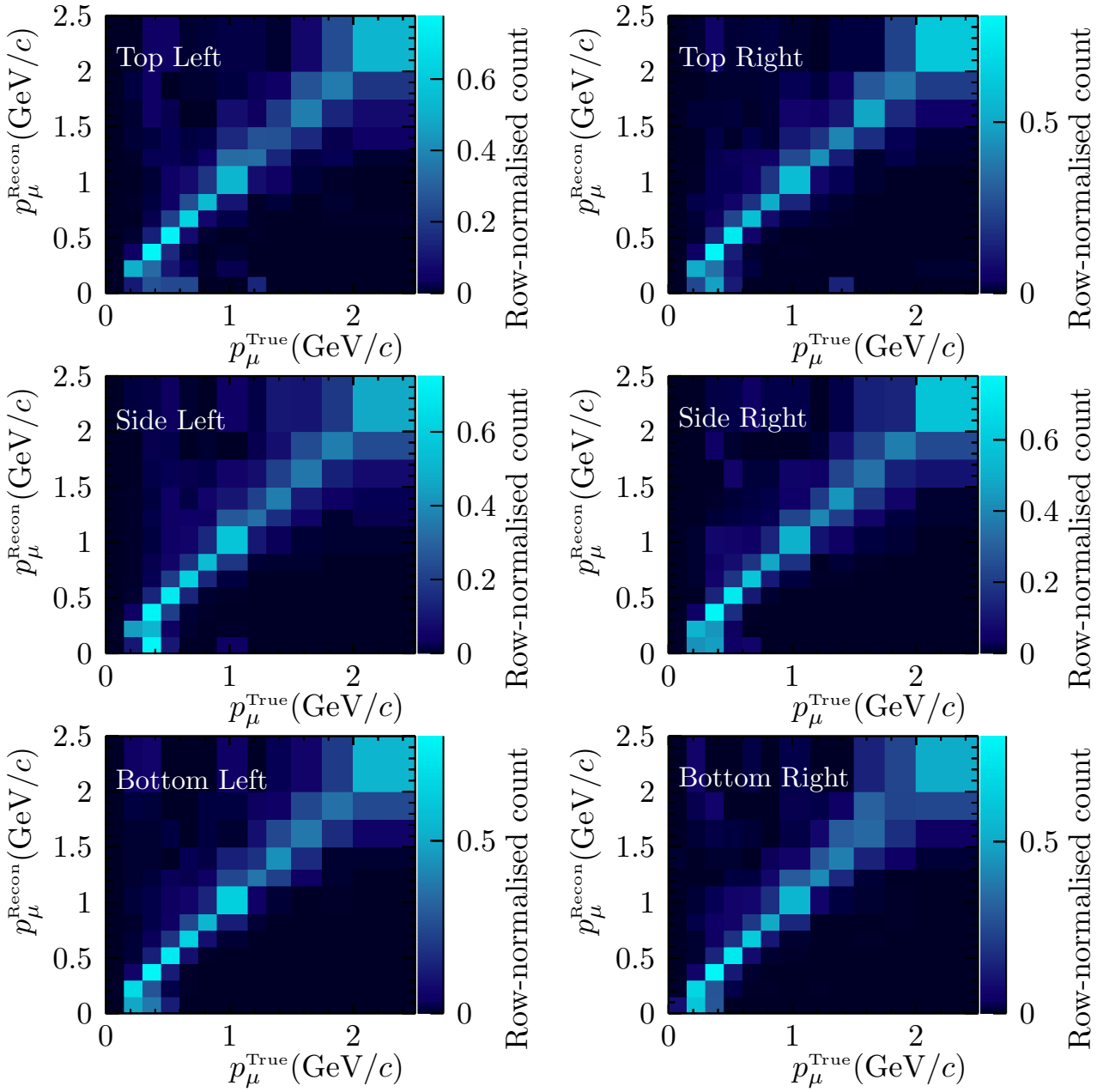


Figure 4.25: The one-dimensional muon momentum truth content matrix for true ECal interactions in the golden candidate sample. Each row depicts the fractional contributions to a single reconstructed bin from each true bin.

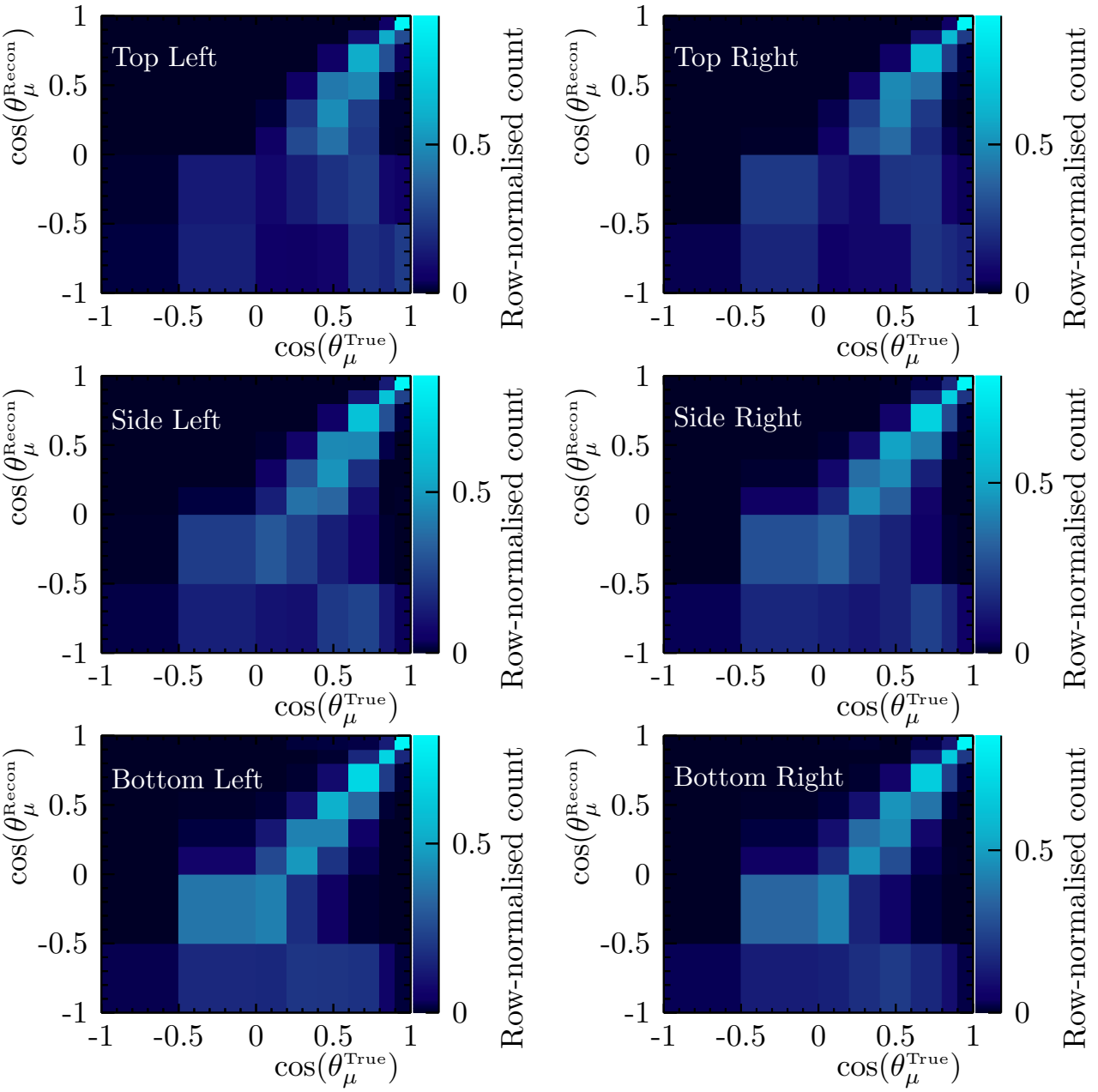


Figure 4.26: The one-dimensional muon polar angle truth content matrix for true ECal interactions in the golden candidate sample. Each row depicts the fractional contributions to a single reconstructed bin from each true bin. The backward bins contain very few true backward-events and are mostly filled with mis-reconstructed muons. Discrepancy between simulation and data in these bins may highlight poor modelling of the processes which cause such backward migration.

well, but Figure 4.26 shows that the backward-reconstructed bins are mostly contributed to by forward-going muons. The interaction model, phase space acceptance, and initial neutrino momentum coincide to produce very few true events in the backward-most bin. Instead, this bin contains true, forward events which re-scatter, or curve, and result in mis-reconstructed vertex candidates downstream of the true interaction. Any significant data–simulation discrepancy in the backward-most bin may come from a number of different effects that would be difficult to separate in the current analysis. However, such a discrepancy may hint at a deficiency in the muon re-scattering model. While it may be unfortunate that the backward bins are heavily contaminated by mis-reconstructed interactions, it does not impact on the robustness of any data–simulation comparisons unless the reconstruction responds significantly differently for real data. Validations of the ECal reconstruction presented in § 3.1.4 and Appendix: A suggest that this is not the case.

Model dependence can enter into an analysis when regions of phase space that are connected by the model are integrated over. For example, for a given neutrino energy and a given muon production polar angle the interaction model predicts a muon initial momentum probability distribution. If the selection is projected onto just muon momentum, and corrections or interpretations based upon the simulation are applied that were determined assuming an incorrect correspondence between the production momentum and polar angle, the interpretations are accordingly biased to the simulation model. An equivalent consideration was discussed when motivating the fiducial volume choice in § 4.1.1. The relevant observables in this selection are the muon production momentum and polar angle, and these are always compared in two-dimensional projections. Incorrect interpretations can still be made if the reconstruction behaves differently between data and simulation, or if the geometry or electromagnetic field simulations are poor, or if the model used to propagate the simulated particles through the geometry is deficient. If correlations in the interaction model do not change slowly and smoothly over the $(p_\mu, \cos(\theta_\mu))$ projections of phase space, then the finite bin widths in either dimension of the projection could cause bias. The interaction model should not polarise the muon production in the azimuthal angle and the reconstruction efficiency has been assessed and shown to vary smoothly over this production angle. By projecting onto the phase spaces relevant for

the interaction model and assessing the efficiency in important , integrated-over dimensions of phase space, the selection aims to be as robust as possible and provide a new, granular way to examine interactions in the outer detectors at ND280.

The selected golden events can be loosely termed the 'signal' sample for this analysis. However, because of the requirement that a charged track is reconstructed as passing through the TPC, the full phase space acceptance is limited. To be more sensitive to data–simulation discrepancy, more samples need to be selected. The next section discusses the silver event sample and presents a comparison of the content between simulation and control sample data.

4.2.2 Silver events

A second sample of ECal interaction candidates was selected in each module—the silver candidate events. The selection follows as in § 4.2.1, except for an inversion of the associated Tracker track requirement.

The content of the silver sample for the nominal simulation is shown in Figure 4.27. As the distribution has no tracker component by design, the events are binned in the reconstructed length and angle of the longest associated track. The binning was chosen to be significantly wider than the track end position residual spread (§ 3.1.4), but there is no constraint that the track is ECal-isolated; it just may not have a Tracker component. The correspondence between data and simulation is reasonable, notably, the predicted normalisation is underestimated in the two forward-most bins. These bins will contain more deposits left by energetic neutrinos interacting with the detector and it may be reasonable that errors in the high-energy neutrino flux model manifest as discrepancies in the forward-most bin. The selection performance as a function of cuts ignored is shown in Figure 4.28. To ensure exclusivity of the golden and silver signal definitions, the denominator of the efficiency used in Figure 4.28 enforces that a true muon from the selected interaction did not pass through the TPC active region.

In addition to studying interaction candidates reconstructed within the fiducial volume of the Barrel ECals, data–simulation discrepancy may be apparent in distributions of properties re-

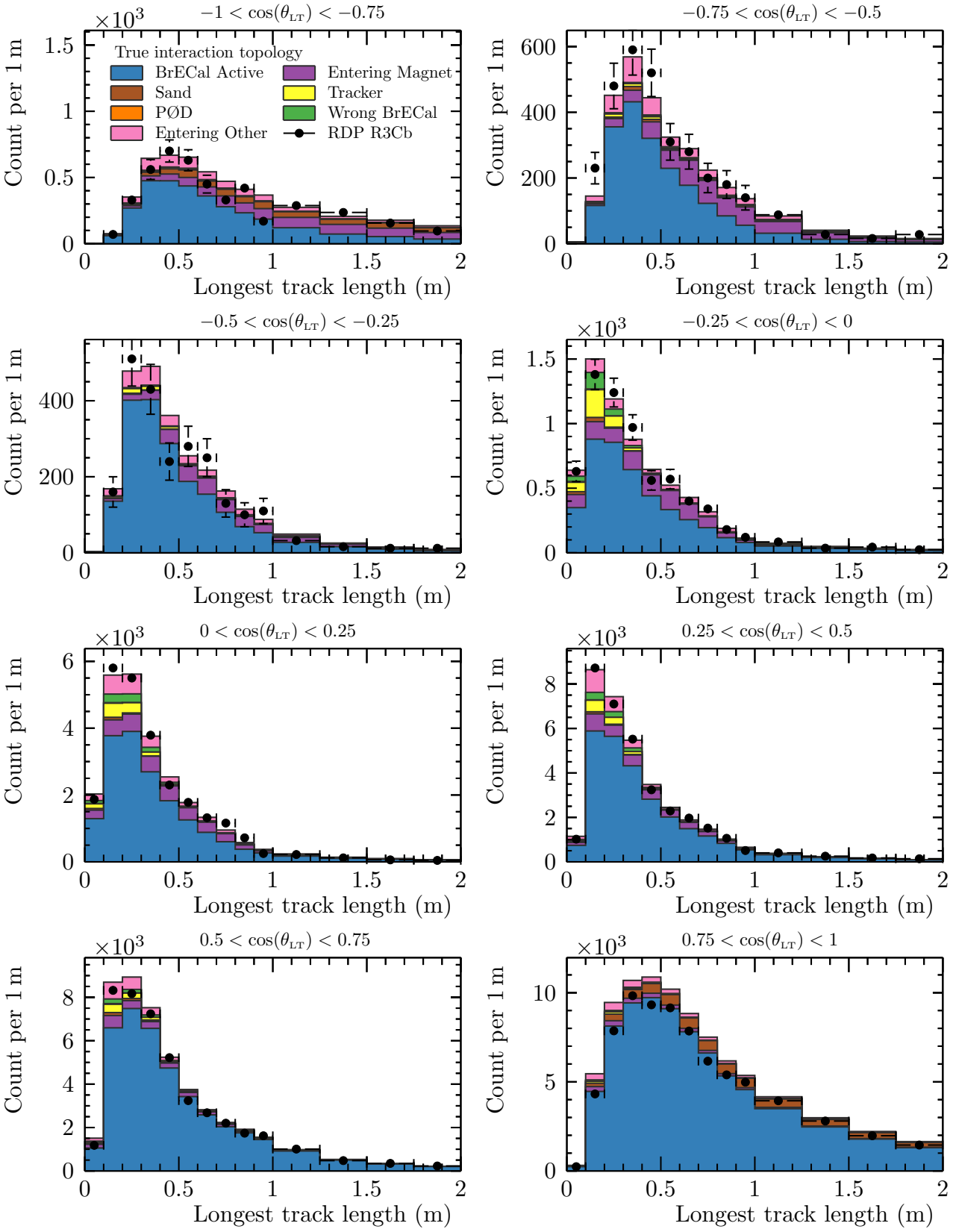


Figure 4.27: The contents of the silver interaction candidate sample projected onto longest track length and polar angle. The simulation prediction is subdivided by true interaction topology.

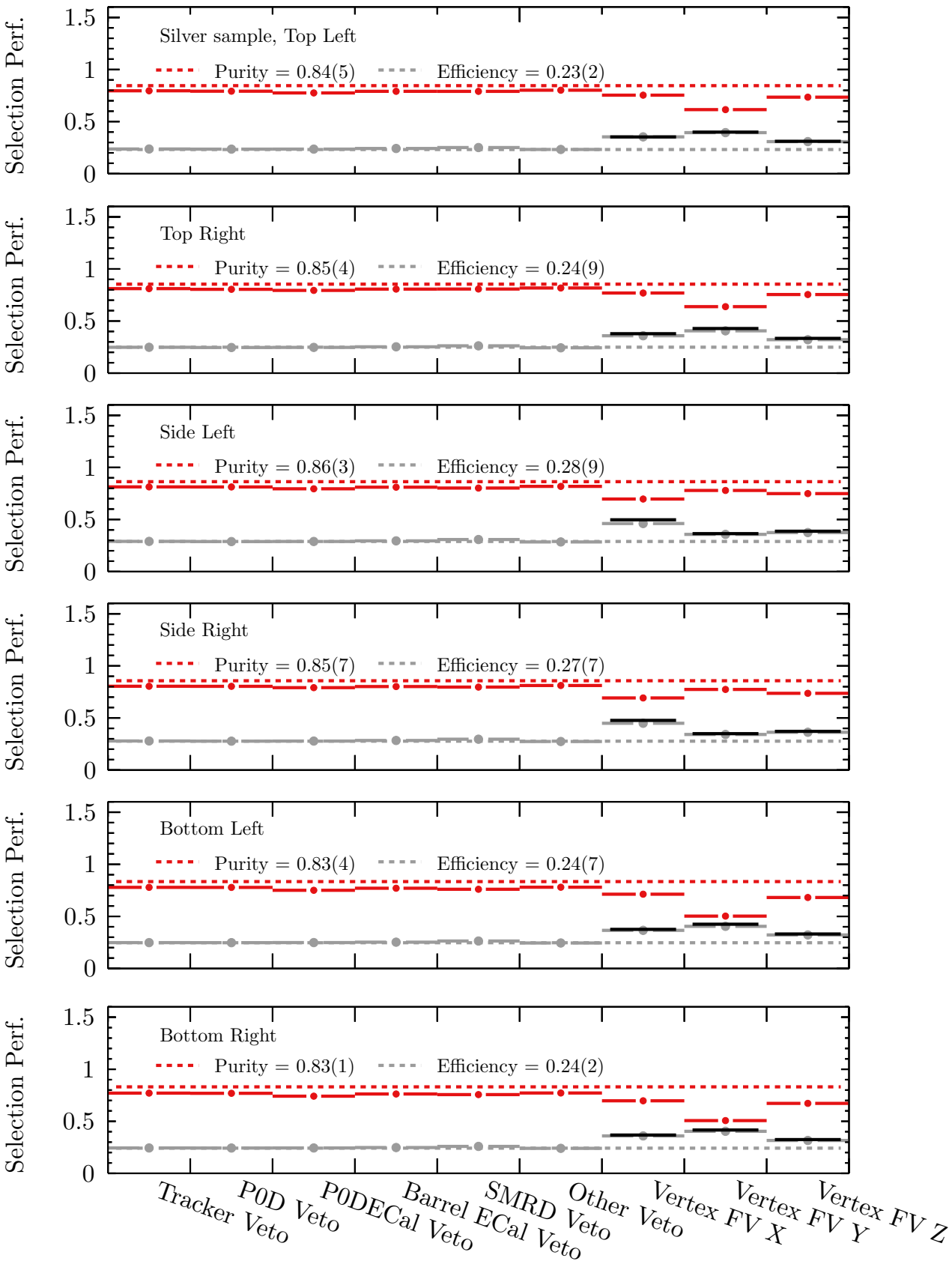


Figure 4.28: The effect of each cut on the efficiency and purity of the selection. The difference between the selection performance with each cut individually ignored and the full selection performance (dashed line) highlights the effect of a single cut with all other cuts accounted for. The black lines for each dimension of the fiducial volume cut show the naive expectation for the efficiency gain when ignoring one dimension of the fiducial volume restriction. This naive gain is calculated by extrapolating the selection efficiency over the fiducial mass to the selection mass available when ignoring one dimension. The mass ratios can be found in Table. 4.2

constructed from deposits that appear to have been left by particles entering the Barrel ECal modules.

4.2.3 Entering background selections

In the previous analysis, deposits that started outside the fiducial volume were simply discarded. A 'reverse' sample, containing events which passed the fiducial volume cut, but then failed any other cut, was used to constrain all classes of background simultaneously [96, § 8.2]. Such an analysis strategy places too much confidence on the ability of the simulation to predict the combined effect of all sources of background accurately. In an attempt to remedy this, two entering background samples were selected in this analysis. The first aims to provide a representative sample of deposits left by charged particles entering from interactions within the magnet flux return. The second aims to characterise background entering through the the upstream face of the detector, which should largely consist of sand muons. Both of these samples are selected purely by reconstructed vertex position. A limitation of the selection presented in Ref. [96] is the high signal content in the background-constraining selections, about 40%. The effective constraint of either entering backgrounds or signal with such a sample is highly model-dependent. As a result, the selection volumes are kept thin—a single bar or layer thickness—to limit the number of true ECal interactions contained within the sample. The selection volumes are depicted in Figure 4.29.

The nominal simulation content for the magnet- and upstream-entering samples is shown in Figure 4.30 and Figure 4.31 respectively. The significant amount of aluminium-target interactions occur on the Barrel ECal support frame and the aluminium coils of the ND280 magnet. The entering background samples have no limitation on whether the selected track contains a tracker segment. As the selected track may span the whole of ND280, the track length provides a proxy observable for particle kinetic energy. As a result, the samples provide a constraint on both the rate and approximate kinematics of entering background particles.

A general feature of the correspondence between data and simulation is that angular bins which contain significant sand contribution are often over-estimated by the simulation. This may be

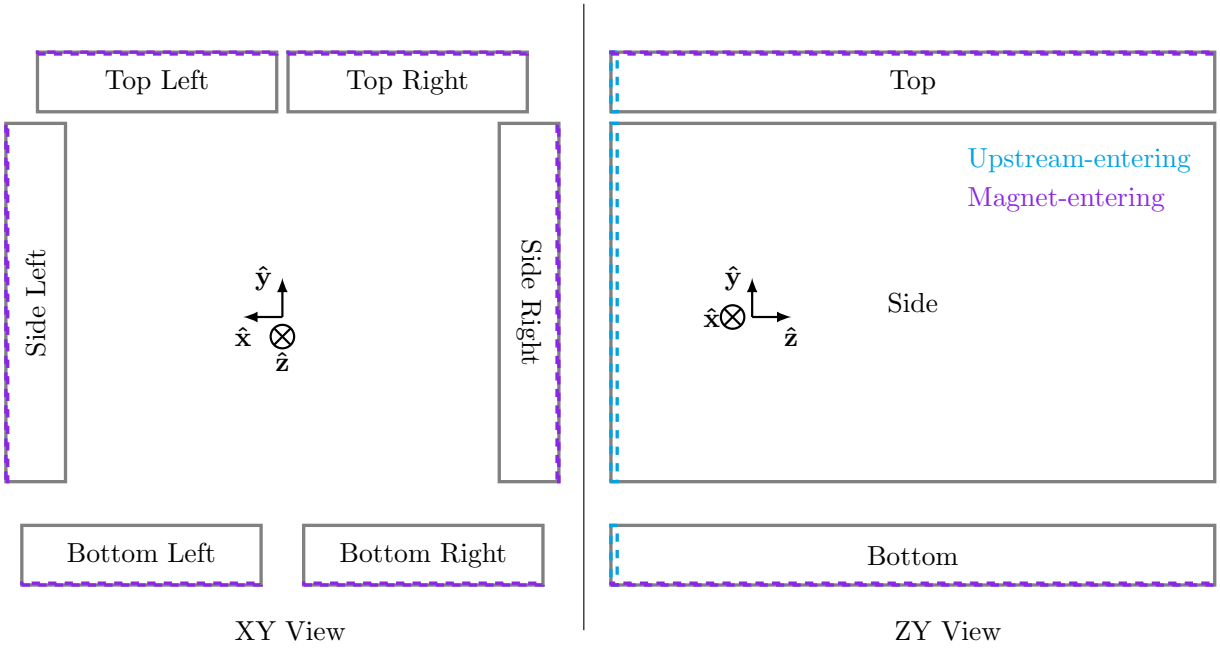


Figure 4.29: The definitions of the two entering background samples in each module which aim to select tracks entering from the magnet and the upstream ECal face.

because of a mis-modelling in the sand muon simulation, either as a result of a deficiency in the neutrino flux simulation or the simulation of the magnetic field in the return yoke.

4.2.4 Selection summary

The selection presented here differs from the selection in the 2015 analysis in a number of ways:

- The Barrel ECal fiducial volume was re-evaluated to conserve regions of consistent track reconstruction efficiency and entering background rate.
- The effect of the selection cuts was compared with a statistically independent data sample. While no analysis choices were made after comparing to the data, the correspondence gives confidence that the selection is behaving comparably on the simulated and observed data.
- The distributions of entering background were granularly sampled with two selections. This is in contrast to the 2015 analysis, which discarded deposits that started outside the fiducial volume and instead attempted to constrain the varied sources of background with the interaction candidate sample.

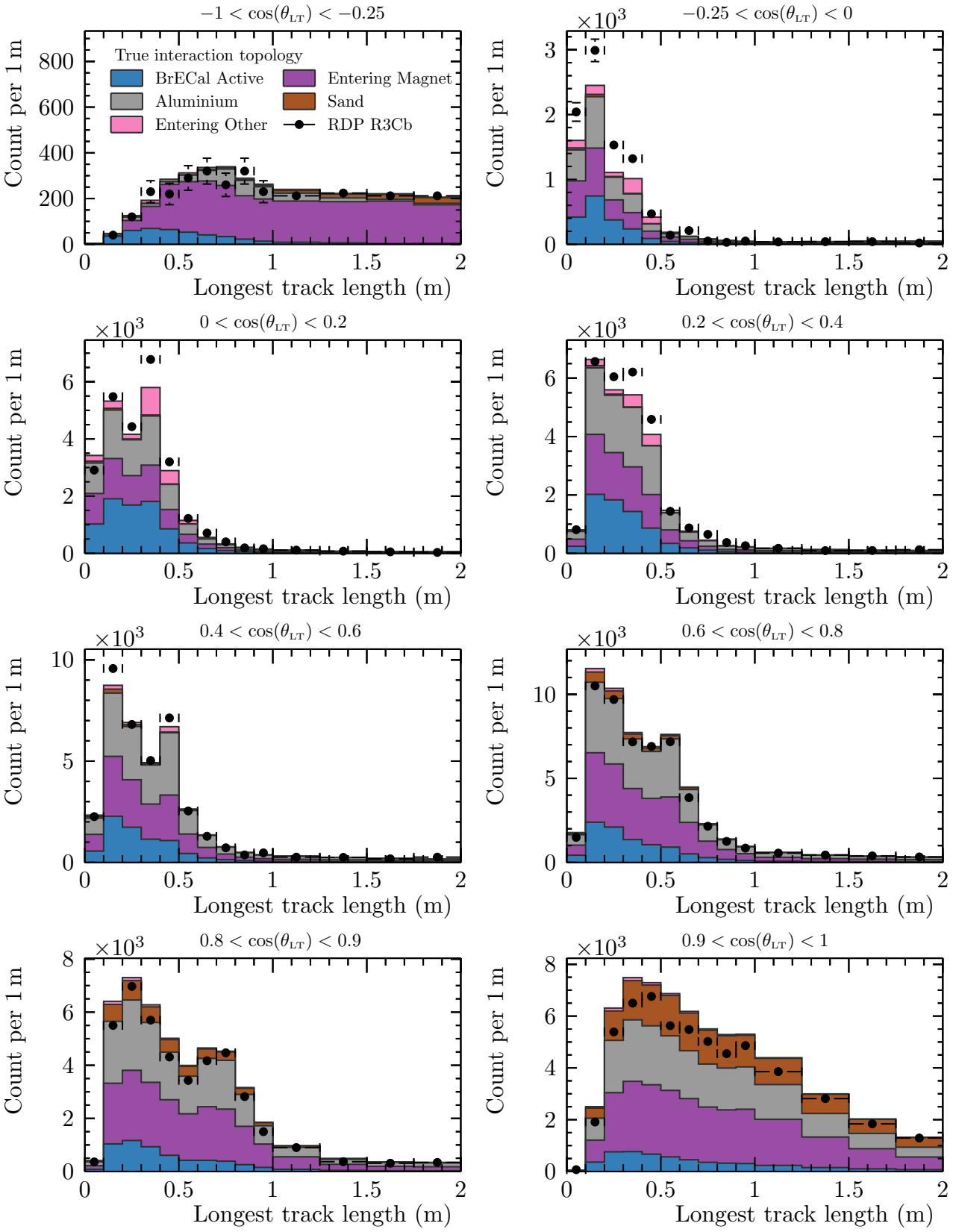


Figure 4.30: The contents of the magnet entering background sample projected onto longest track length and polar angle. The simulation prediction is subdivided by true interaction topology.

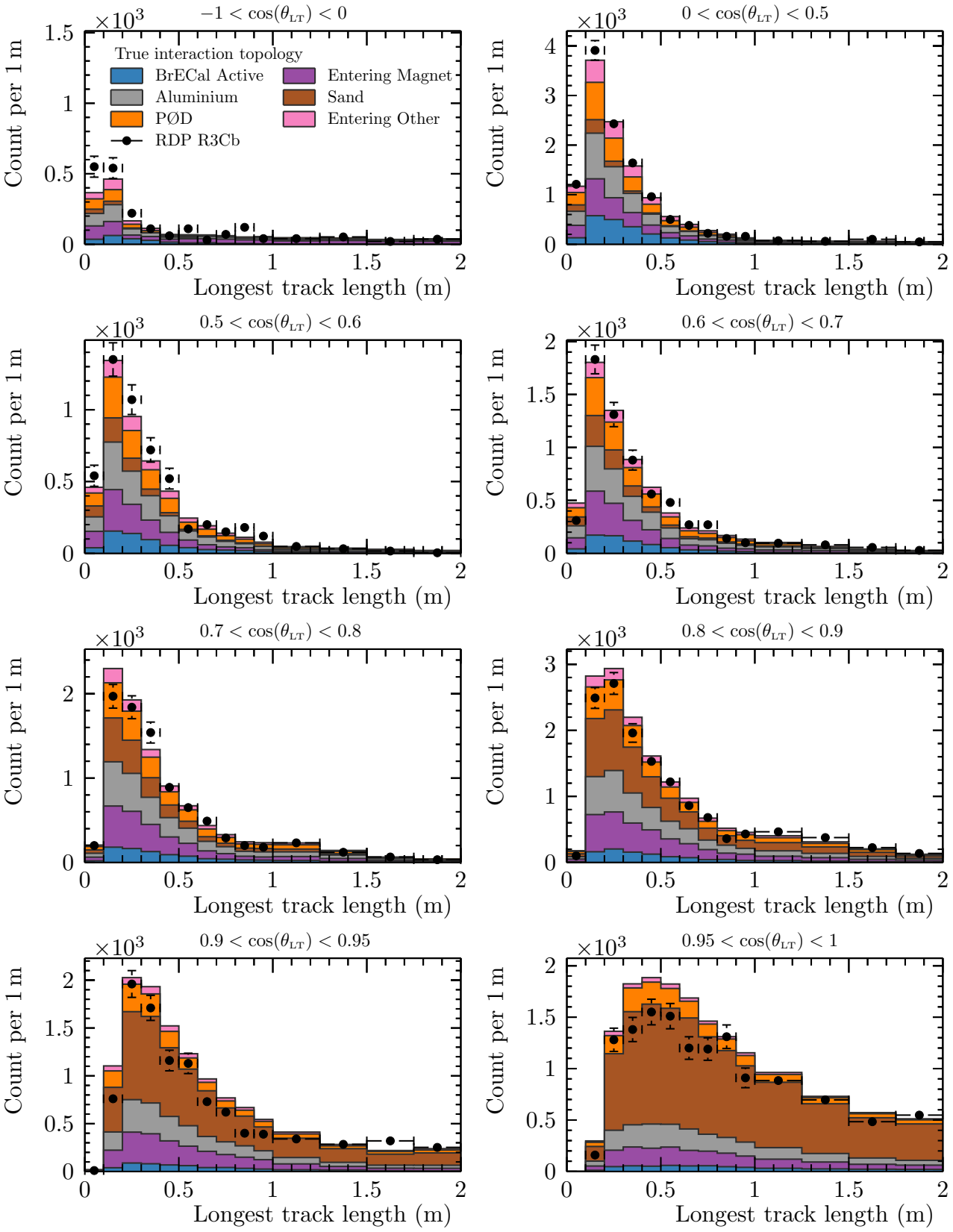


Figure 4.31: The contents of the upstream entering background sample projected onto longest track length and polar angle. The simulation prediction is subdivided by true interaction topology.

- The fiducial volume performs the majority of the background rejection.
- No attempt was made to separate charged-current and neutral-current in the silver ECal interaction sample. In the 2015 analysis, this separation motivated complex multi-dimensional cuts that were finely tuned to the background rejection capabilities of the simulated ECal detector response.
- The integration of the ECal-isolated vertex reconstruction with ND280 global reconstruction allowed for information from the whole detector to be used in the selection, most notably the use of the detector-wide upstream veto and the PID capabilities of the Tracker.

4.3 Future Work

The dominant background topology in the high-angle and backward bins of the golden selection is interactions which were reconstructed to occur in the wrong ECal module. While it is possible that some of these are due to global matching failures, it is most likely that a re-scattering caused the vertex candidate position to be incorrectly reconstructed. The use of track timing may allow some of these events to be reconsidered as signal originating from a different module. It would be important when assessing the systematic uncertainty associated with using the track timing to be able to make such a choice at the 'selection' level—even if the reconstruction was improved as discussed in 3.4.

The reconstructed time of flight for the selected golden event candidates is presented in Figure 4.32. For the simulation prediction, the vast majority of the mis-reconstructed background in the backward bins could be re-assessed if the time of flight was used to re-position the vertex candidate. The simulation is clearly able to distinguish track sense based upon timing information—more study would be required to properly assess the uncertainties associated with making such choices for data events.

As shown in Figure 4.26, the majority of the simulated signal events reconstructed to be backward-going are in-truth, forward-going. A Michel electron tag might be able to iden-

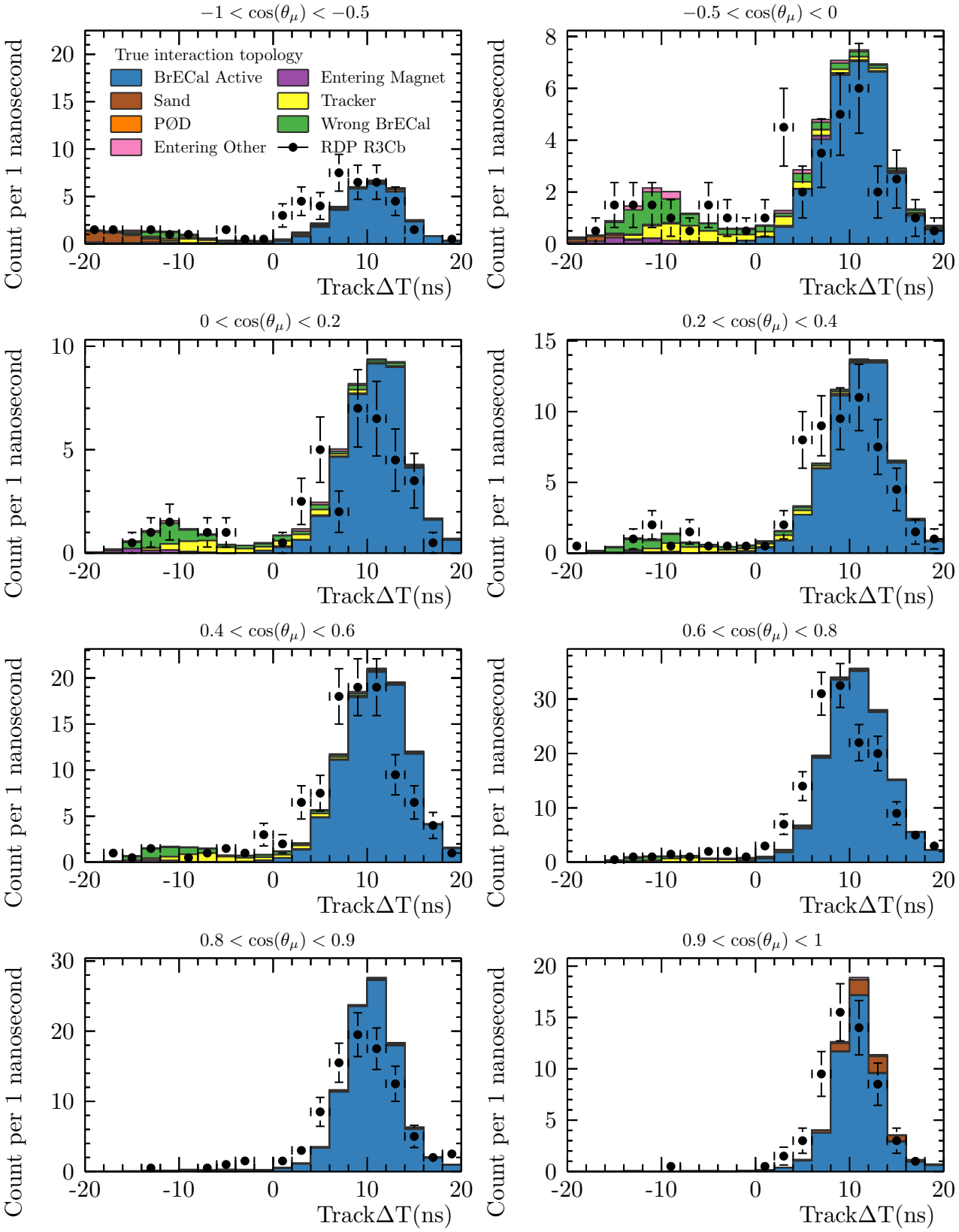


Figure 4.32: The time of flight for the muon-like global track of selected golden event candidates. Tracks reconstructed with the wrong sense in the high-angle and backward bins are clearly separable by timing information.

tify the track end for particles that stop in an active detector. The active stopping volume for selected muons from Barrel ECal interactions is shown in Figure 4.33. The requirement that muons stop in an active detector will clearly shape the selection efficiency in the true muon momentum and polar angle phase space—this would severely limit the confidence with which the data could be used to constrain flux or neutrino interaction models. The addition of a Michel electron tag may be able to increase the selection purity for backward-going tracks, but its effect on the selection would need be to be studied in great detail.

The selection presented above provides a robust sample in which to granularly assess regions of data–simulation agreement and disagreement. It is also a basis for further studies, such as those mentioned in this section.

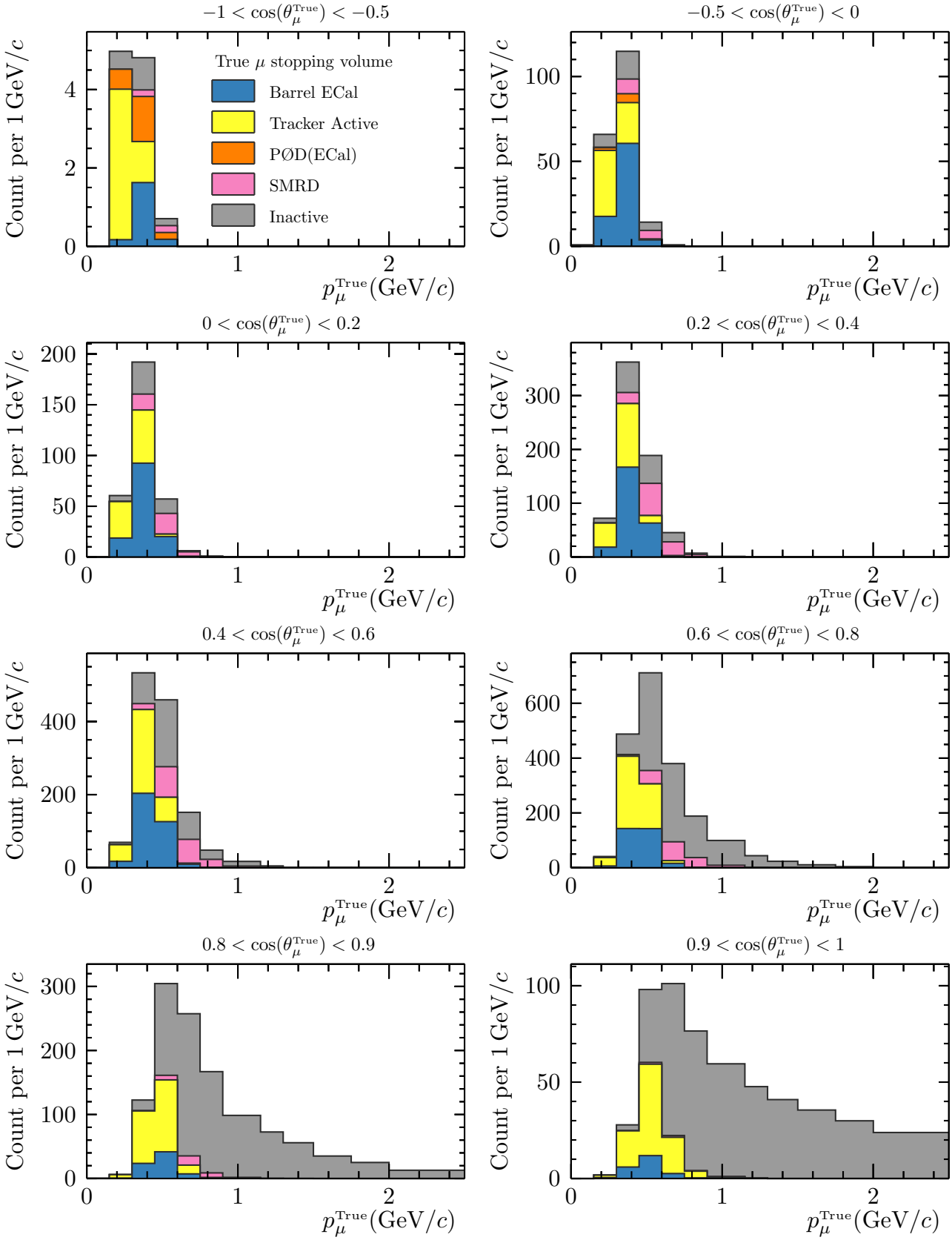


Figure 4.33: The true stopping volume for muons produced in Barrel ECal interactions. Muons that stop outside of an instrumented detector volume are shown as 'Inactive'. It would not be possible to use Michel electron detection to perform muon PID for these particles. High momentum, forward muons frequently escape the detector. High angle and backward muons stop in a variety of sub-detector systems.

Chapter 5

Sources of error

This chapter discusses the most important sources of error that are relevant to this analysis. The effect of each error needs to be propagated to an uncertainty on the predicted event rate in each bin of the golden interaction candidate sample. For a robust uncertainty assessment, covariances between pairs of analysis bins also need to be understood. As no part of the simulation model is to be fitted to the data, the uncertainty for the full prediction has been assessed and presented.

All the uncertainties described in this chapter are propagated to the selection in the same way. For some ensemble of toy parameter values, thrown within uncertainties, a covariance matrix of the binned event rate was constructed. For the majority of the nuisance parameters discussed in this chapter, the response to some parameter variation can be fully encapsulated by an event weight; but a number of the detector parameters induce reconstructed kinematic variations that may cause momentum bin migration. Covariance matrices determined from uncorrelated sources of uncertainty are then summed together to give the full uncertainty band on the predicted event rate.

The software and inputs used to determine the response of a given event to the variation of parameters controlling the neutrino beam, neutrino interactions, and the tracker reconstruction were not developed by the author. However, the study into the ECal-tracker matching and tracking uncertainty, in § 5.3.2, is original work.

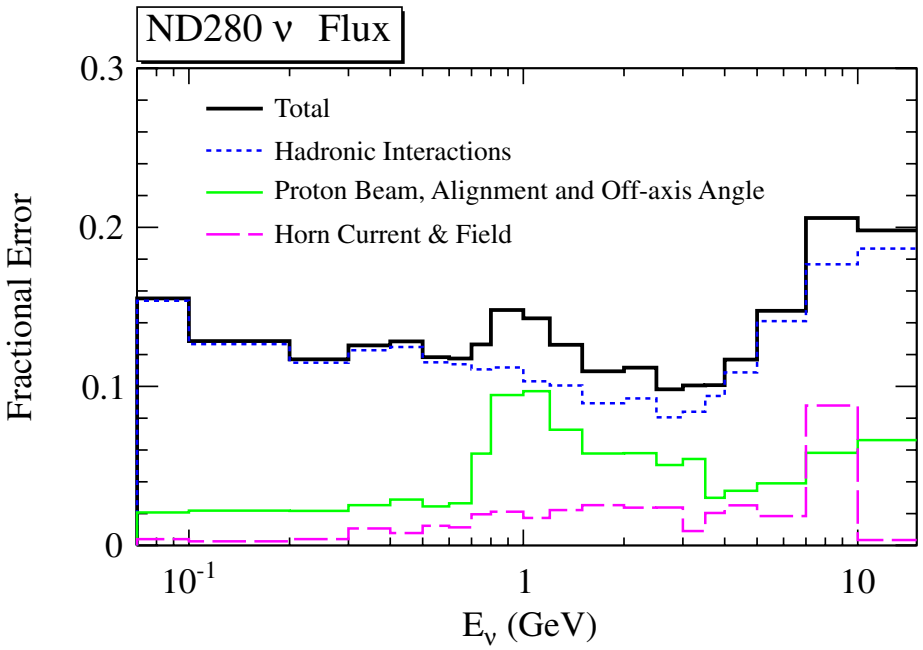


Figure 5.1: The fractional uncertainties in the muon neutrino flux prediction at ND280 from the three main sources of error. For neutrino energies near to the peak energy (0.6 GeV), the total flux error is influenced by uncertainties in hadronic interactions in the target and by uncertainties in the properties of the proton beam. Figure from Ref. [53]. These uncertainties are constrained by comparison to external hadron-beam data from the NA61/SHINE experiment [86] (*c.f.* § 2.5).

5.1 Neutrino beam

The neutrino beam spectral shape and normalisation uncertainties are ultimately determined by an assessment of the effect of variations of fundamental physics parameters in the proton beam simulation, hadronisation in the beam target, and the magnetic focussing horns. The fractional error from each of these sources for the muon neutrino flux prediction at ND280 is reproduced in Figure 5.1. In the peak of the flux distribution (*c.f.* Figure 2.2b), uncertainties in both the hadronic interactions in the target and properties of the proton beam contribute significantly to the total flux uncertainty.

The determination of the flux prediction response to variations of the fundamental parameters is too computationally demanding to run at 'analysis' time. Instead, uncertainties on these parameters are propagated to an uncertainty on the neutrino beam energy spectrum, which is provided as a covariance matrix binned in neutrino energy and species. For a detailed description of the J-PARC neutrino beam simulation and hadronisation tuning, the reader is

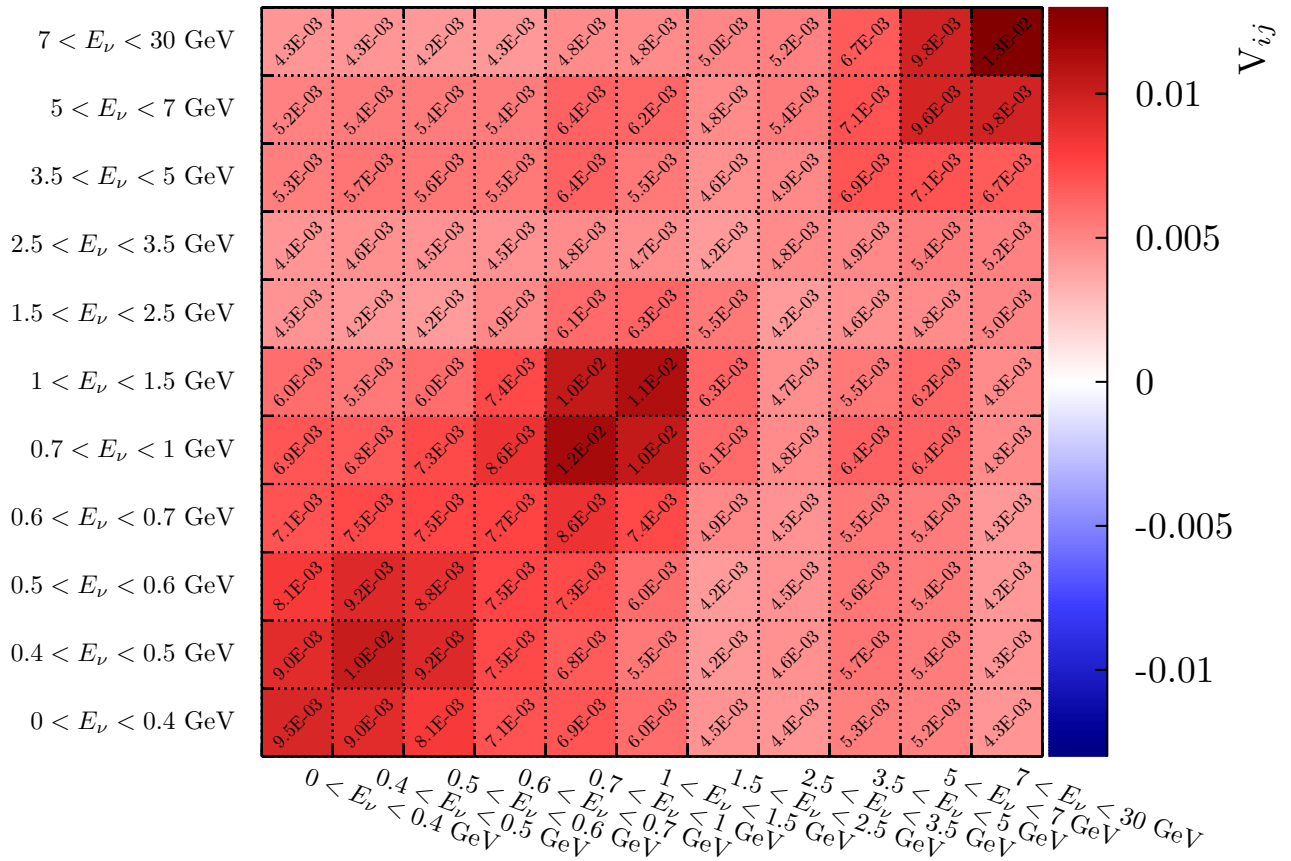


Figure 5.2: Covariance matrix describing the correlated uncertainties in the muon neutrino flux at the ND280 when running in 'neutrino mode'. The neutrino flux uncertainties can be seen to be positively correlated throughout.

directed to Ref. [53]. The covariance matrix for the muon neutrino beam component is shown in Figure 5.2—the dominant component for this analysis (*c.f.* Table. 5.1). Random vectors of correlated neutrino energy bin weights can be thrown from the matrix and used to weight selected interaction candidates by true energy of the neutrino that produced them.

The number of selected events per analysis sample in a Run3Cb-equivalent simulated data set is presented in Table. 5.1. The total non-muon neutrino contamination is below five percent for each sample—only the golden sample actively selects muon neutrino events.

Unfortunately, no uncertainties were available for the neutrino flux prediction used in the sand muon simulation. Instead, an *ad-hoc*, uncorrelated, fractional uncertainty of ten percent was used. This was motivated by the fractional uncertainty determined from non-sand events, which is between six and ten percent.

Neutrino species	Golden sample	Silver sample	Magnet sample	Upstream sample
Muon neutrino	1214	22368	28479	6796
Muon anti-neutrino	21.0	583	692	187
Electron neutrino	2.8	473	303	100
Electron anti-neutrino	< 0.1	31.9	17.4	7.3
Total	1238	23457	29493	7091

Table 5.1: The number of selected Barrel ECal interactions in a sample of Run3Cb POT-equivalent simulated neutrino beam data, separated by true neutrino species. Non-integer counts are because of flux tuning and POT scaling. The interactions may have occurred, in truth, outside of a Barrel module.

The effect of the flux tuning, and the uncertainty band generated using 1000 toy throws from the four neutrino species flux uncertainty matrix, is shown in Figure 5.3. The visible width of each component of the uncertainty band is such that two sources that contribute similarly to the total bin variance will appear with the same width. For each source of uncertainty, σ_i, the visible width, w_i, is determined as

$$w_i = \sqrt{\sum_j \sigma_j^2} \times \frac{\sigma_i^2}{\sum_j \sigma_j^2}. \tag{5.1}$$

For the majority of analysis bins, the neutrino flux uncertainties are significant. The analysis bin correlation matrix for the flux uncertainty is presented in Figure 5.4. As expected from Figure 5.2, the effect of the flux uncertainty is strongly correlated across the analysis bins.

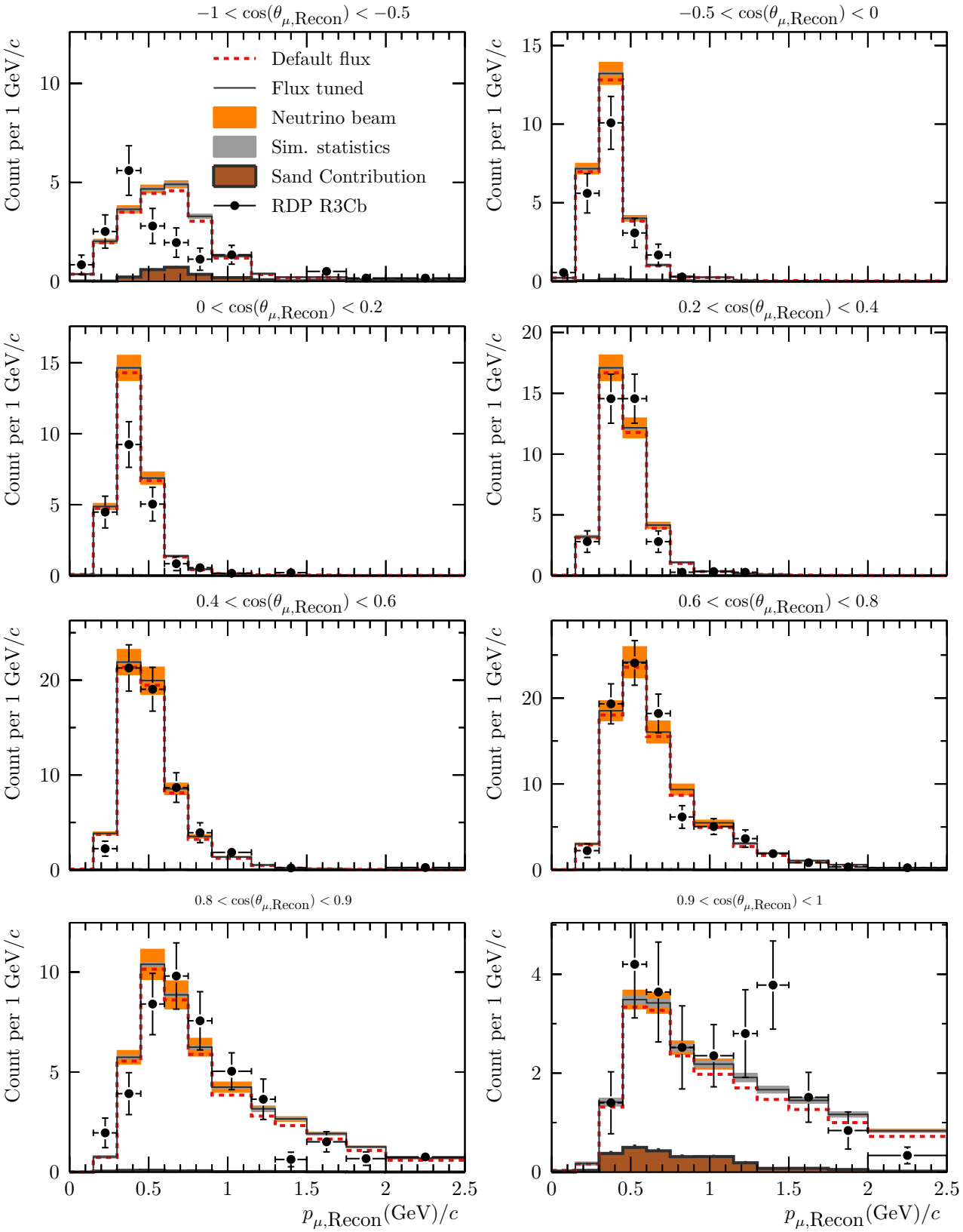


Figure 5.3: The effect of the neutrino beam uncertainties propagated to the golden interaction sample. The result of the flux tuning compared to the default neutrino flux is also shown. An *ad-hoc* 10% uncertainty is included on the sand muon component.

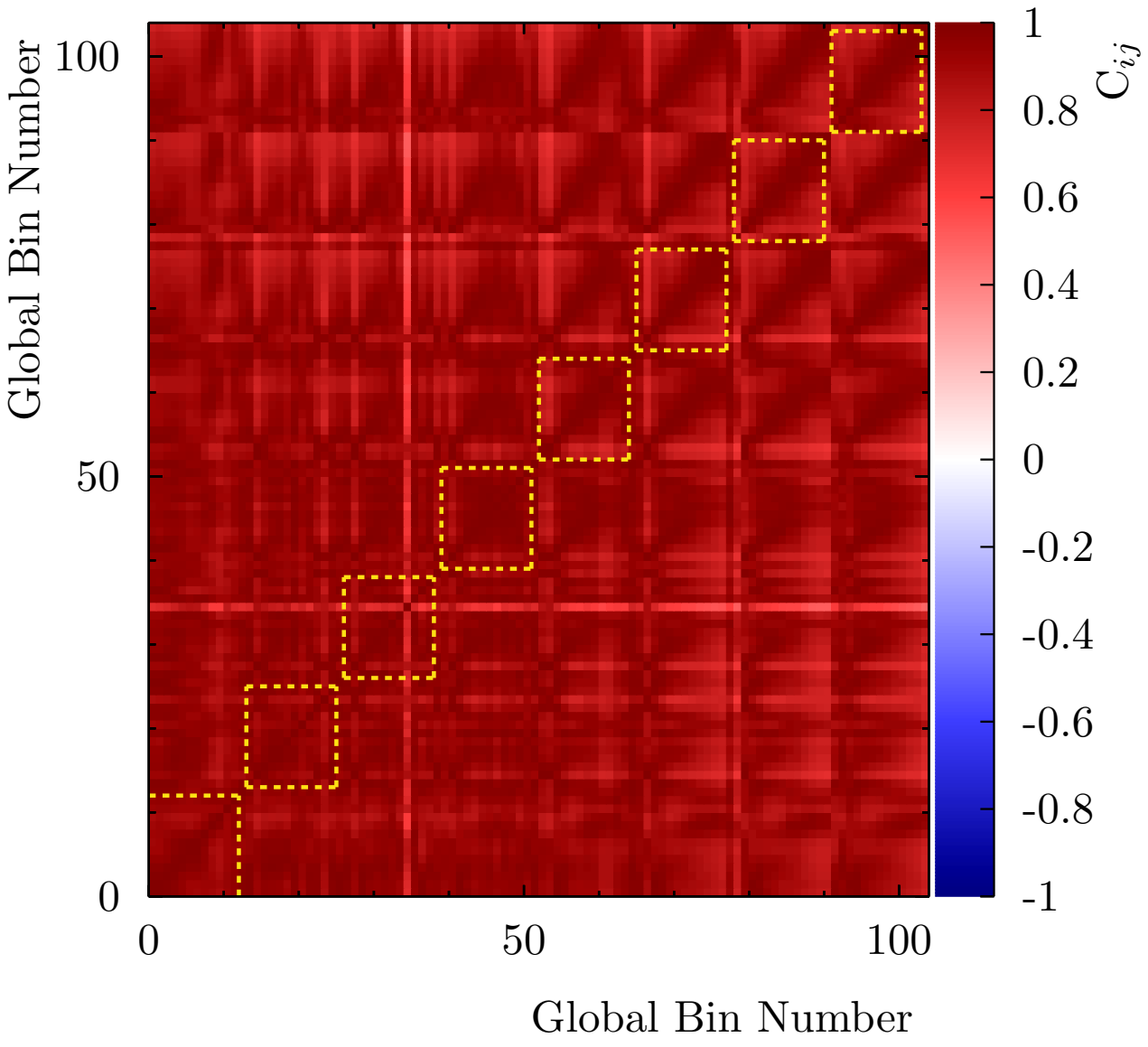


Figure 5.4: The correlation matrix determined from the neutrino beam uncertainties in golden sample analysis bin number. The dotted regions correspond to the polar angle bins shown in Figure 5.3. The bottom left (top right) region corresponds to the backward-most (forward-most) reconstructed muons. Within each region the binned momentum increases from 0 GeV/c to 2.5 GeV/c. Each angular bin has a corresponding momentum overflow bin, these are the gaps to the above-right of each dotted region. The propagated neutrino flux uncertainties can be seen to be strongly, positively correlated—the effect on the muon kinematic distribution is largely a normalisation error.

5.2 Neutrino interactions

As discussed in § 1.3.2, at the time of writing, the details of neutrino interactions on nuclear targets are a limiting factor in interpreting neutrino scattering data. The current models are built out of multiple constituent models, some of which have possibly overlapping components that may double-count nuclear effects[1]. There is a significant global effort—which includes neutrino interaction theorists, simulation developers, and experimentalists—to develop and constrain consistent, complete interaction models. Unfortunately, because of the non-perturbative energy-momentum transfer regime in which accelerator-based neutrino physics is performed, this is a non-trivial task [31]. Estimating well-motivated parameter uncertainties from models that we know to be poor or incomplete is difficult. An effective, but model-dependent, method of estimating such uncertainties is to fit state of the art model predictions to large, high-quality neutrino–nucleus interaction samples. Unfortunately, there is very little published lead-target neutrino scattering data (*c.f.* § 2.6). The currently available models have significant trouble consistently explaining the available carbon-target data [99]—the unconstrained extrapolation of the findings of carbon-target analyses to a nucleus as large as lead is questionable at best. Ideally, a high-purity sample, such as presented in § 4, would be used to begin to constrain the model predictions for lead-target interactions. Unfortunately, there was not enough time left to robustly perform such investigations after having developed and assessed the selections. This is left as a so-called 'treat in store'.

The response to a variation from each of the parameters discussed in this section is determined by cross section re-weighting. This technique is described in more detail in Appendix: B, alongside a description of NuWro reweight and fits of the neutrino-nucleon interaction parameters to historic bubble-chamber data.

[1] The current default NEUT model has been humorously referred to as a 'Franken-model'. CCQE interactions on carbon use Benhar's nuclear spectral function [44] to model the nucleus and calculate the final state kinematics—this model includes multiple nucleon ejection through short-range nucleon–nucleon correlations. More multi-nucleon effects are included through the use of Nieves 2p2h, which is calculated using a Local Fermi Gas (LFG) nuclear spectral function [42]. However, single- and multi-pion production, and deep inelastic scattering interactions are simulated with a simple Relativistic Fermi Gas (RFG) nuclear spectral function. All interactions on nuclei other than carbon, oxygen, or iron use an RFG nuclear spectral and the Lewellyn-Smith formalism for true CCQE interactions [33]. Finally, the simulation of the hadronic cascade assumes an LFG-like momentum-density distribution for spectator nucleons.

The interaction model used here is motivated by a NEUT parameter tuning performed in support of the T2K oscillation analysis. NEUT predictions were compared to nuclear target data from the MINERvA and MiniBooNE experiments and a nominal model was selected [99]. While I was not involved in that analysis, I am a core developer of the open-source interaction generator fitting and comparison framework that grew out of the work, NUISANCE. No explicit discussion of NUISANCE is presented here, however, the interested reader is directed to [100]. The chosen interaction model consistently uses an RFG nuclear spectral function, with target-dependent Fermi-surface momenta (p_{F}) and first nucleon removal energies (E_{b}) informed by electron–nucleus scattering data, *e.g.* [101]. True CCQE interactions are modelled in the Lewellyn-Smith formalism [33]. A suppression of CCQE interactions at low energy-momentum transfer is included through a relativistic 'Random Phase Approximation' (RPA) treatment, which accounts for nuclear-medium corrections to the W propagator [45]. Other multi-nucleon effects are included through the Nieves 2p2h model [42]. Single pion production is modelled with the full Rein-Sehgal resonant pion production model, which also includes 'background' single pion production from non-resonant, half-integer isospin processes [35]. The energy-momentum transfer for deep inelastic scattering interactions is calculated with the Bodek-Yang model [38] and the hadronisation is handled by PYTHIA [39]. The transition from resonant pion production to deep-inelastic scattering is modelled by a smooth function that governs the fraction of events generated with each of the single pion production and DIS models. Hadronic re-interactions are modelled by a full, semi-classical cascade simulation, which uses the Oset [102] hadronic interaction processes and cross-section parameterisation. The NEUT cascade parameters are tuned to thin-target, hadron beam scattering data (Table A.1 from Ref. [103] shows the pion scattering data used).

The model described above and used herein differs from the model used during the simulation of the neutrino beam data that has been presented thus far. That model uses the Benhar nuclear spectral function for CCQE interactions on carbon, oxygen, and iron; and takes the quasi-elastic axial mass to be $M_{\mathrm{QE}}^{\mathrm{A}} = 1.21\,\mathrm{GeV}$. That model is re-weighted to the chosen model, which uses an RFG nuclear spectral function and an updated, nominal quasi-elastic axial mass of $M_{\mathrm{QE}}^{\mathrm{A}} = 1.03\,\mathrm{GeV}$. The effect of these tuning weights on the predicted event rate is shown in

Figure 5.5.

The relevant physics parameters that can be varied in the neutrino interaction model and nuclear model are presented in Table. 5.2; parameter nominal values, prior uncertainties and any relevant applicability limits are also shown. For the neutrino–nucleon interaction parameters, a brief description of the CCQE and single pion production parameters is given in Appendix: B. The 2p2h normalisation parameters control the overall probability of 2p2h interactions—*i.e.* the nominal value of 1 corresponds to using the Nieves model predicted cross section for 2p2h interactions; a variation to $2\text{p2h}^{\text{C}}_{\text{NORM}} = 1.2$, results in 1.2 times as many 2p2h interactions on carbon, which increases the total event rate. The $2\text{p2h}^{\text{Other}}_{\text{NORM}}$ parameter controls the 2p2h normalisation for all nuclear targets heavier than oxygen. The effect of the uncertainties on hadronic cascade processes on the golden sample event rate will be small—a few more or less charged pions may escape the nucleus, travel through the TPC, and be mis-identified as muons. The uncertainty on the hadronic cascade is also taken into account. For a detailed description of the implementation and motivation of these uncertainties, the reader is directed to [103]. All of the parameters, with the exception of the 2p2h normalisation parameters, are thrown according to a Gaussian probability distribution with a mean at the parameter nominal and a width of the parameter uncertainty. Parameter limits are observed by truncating the Gaussian distribution.

The predicted true target nuclei composition of the four samples is summarised in Table. 5.3. The contamination from targets other than lead and carbon is about six percent and sixteen percent for the golden and silver samples respectively. Such backgrounds would need to be handled carefully if these data were to be published as a lead- and carbon-target scattering cross section.

The chosen uncertainties for the neutrino–nucleon interaction parameters is known to be conservative. As shown in Appendix: B, the single pion cross-section response to variations of $M^{\text{A}}_{\text{RES}}$ and $C^{\text{A}}_{5}(0)$ is strongly correlated—the use of uncorrelated parameter throws in Figure 5.5 results in a conservative assessment of the single pion production uncertainty.

The simulation ‘pass-through’ information required to calculate responses to interaction pa-

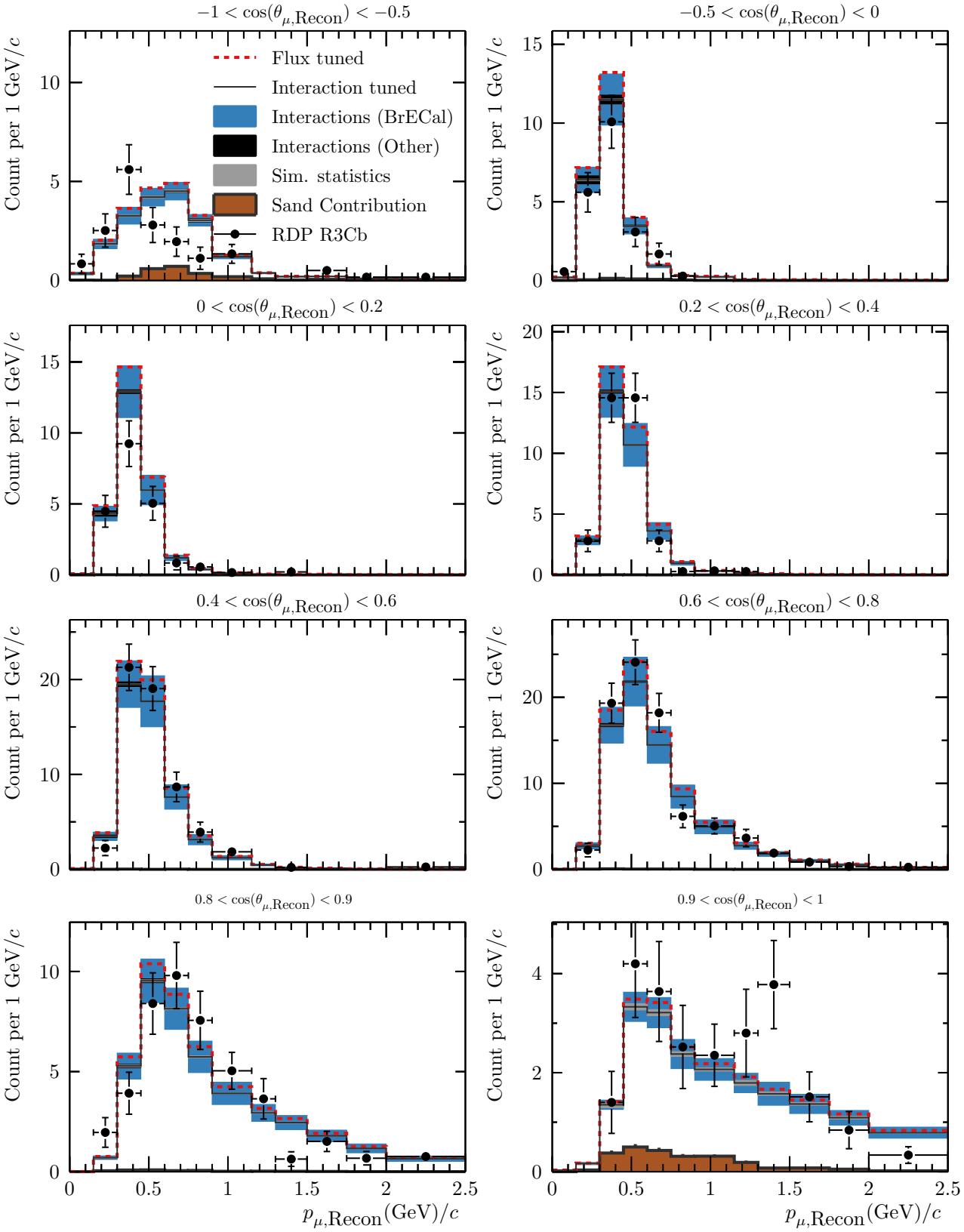


Figure 5.5: The effect of the neutrino interaction, nuclear model, and hadronic re-interaction uncertainties propagated to the golden interaction sample. The difference between the tuned interaction model and the original simulation model is also shown. The effect of the nuclear model and hadronic re-interaction uncertainties are negligible. The uncertainty band is subdivided into interaction uncertainties resultant from neutrino interactions that truly occurred within a barrel ECal module and interactions that occurred elsewhere.

Parameter class	Name	Nominal value	Prior uncertainty	Limits
Nuclear model	$p_{\mathrm{F}}^{\mathrm{C}}$ /MeV	217	30	[200,275]
	$E_{\mathrm{b}}^{\mathrm{C}}$ /MeV	25	9	[12,42]
	$p_{\mathrm{F}}^{\mathrm{O}}$ /MeV	217	30	[200,275]
	$E_{\mathrm{b}}^{\mathrm{O}}$ /MeV	25	9	[12,42]
	$p_{\mathrm{F}}^{\mathrm{Al}}$ /MeV	233	30	[205,275]
	$E_{\mathrm{b}}^{\mathrm{Al}}$ /MeV	28	9	[20,56]
	$p_{\mathrm{F}}^{\mathrm{Fe}}$ /MeV	250	30	[205,275]
	$E_{\mathrm{b}}^{\mathrm{Fe}}$ /MeV	33	9	[20,56]
	$p_{\mathrm{F}}^{\mathrm{Pb}}$ /MeV	245	30	[205,275]
	$E_{\mathrm{b}}^{\mathrm{Pb}}$ /MeV	44	9	[20,56]
Neutrino–nucleon interactions	$M_{\mathrm{QE}}^{\mathrm{A}}$/GeV	1	0.2	N/A
	$\mathrm{2p2h}_{\mathrm{NORM}}^{\mathrm{C}}$	1	1	[0,2]
	$\mathrm{2p2h}_{\mathrm{NORM}}^{\mathrm{Other}}$	1	1	[0,2]
	$M_{\mathrm{RES}}^{\mathrm{A}}$/GeV	0.95	0.2	N/A
	$C_5^{\mathrm{A}}(0)$	1.01	0.2	N/A

Table 5.2: The nominal values, widths, and applicability limits for the most relevant neutrino interaction parameters. Parameters are thrown with Gaussian PDFs, centred on the nominal value, with a width of the prior uncertainty; where applicable, the distributions are truncated at the limits. The two 2p2h normalisation parameters are thrown with a uniform distribution between the limits. The nominal values for all parameters, with the exception of $M_{\mathrm{QE}}^{\mathrm{A}}$, were set to the default values in NEUT v5.3.2 [52]. The nominal value of $M_{\mathrm{QE}}^{\mathrm{A}}$ was informed by the results of Ref. [99]. The uncertainties on the nuclear model parameters were set at the default values from the T2K event reweighting tools. Those on the neutrino–nucleon interaction parameters were set to conservative values, informed by the default uncertainties in NEUT reweight, but somewhat inflated.

Target nuclei	Golden sample	Silver sample	Magnet sample	Upstream sample
Hydrogen	10.8	306	69.2	44.8
Carbon	365	6483	1757	905
Oxygen	9.9	106	54.4	21.1
Aluminium	22.3	1308	11484	2557
Silicon	1.5	21.0	10.5	3.6
Chlorine	0.1	2.1	1.9	61.5
Titanium	2.8	47.9	11.5	4.6
Iron	16.3	1558	12226	2025
Cobalt	0.1	9.1	3.4	14.3
Copper	0.6	3.9	4.7	81.9
Lead	807	13603	3862	1357

Table 5.3: The number of selected Barrel ECal interactions, separated by target nuclei, in a sample of Run 3Cb POT-equivalent simulated neutrino beam data. Non-integer counts are due to flux tuning and POT scaling. Interaction targets that constitute less than ten expected interactions in all samples have been omitted. The interactions may have occurred, in truth, outside of a Barrel module.

rameter variations is not available for the sand muon sample. As a result no interaction uncertainties are evaluated for the sand component of the golden sample. The effect of the neutrino interaction uncertainties on the flux of sand muons entering ND280 is likely obscured by uncertainties in the neutrino flux simulation and muon propagation through the earth surrounding the ND280 pit. Instead, a sample which includes regions of significant sand muon content, such as described in § 4.2.3, should be used to provide a simultaneous constraint on the entering sand muon flux.

The effect of the neutrino interaction, nuclear model, and hadronic re-interaction uncertainties are shown in Figure 5.5. The freedom in the nuclear model and pion re-interaction parameters contribute negligibly to the total uncertainty. The associated correlation matrix is presented in Figure 5.6. The propagated uncertainties are highly correlated across the analysis bins, as should be expected from the cross-section response to variations in $M^{\mathrm{A}}_{\mathrm{QE}}$, $M^{\mathrm{A}}_{\mathrm{RES}}$, and $C^{\mathrm{A}}_{5}(0)$, which are discussed in Appendix: B.

In this section, the model uncertainties for interactions occurring both within a barrel ECal module and for entering background events have been presented. If this sample were to be used to extract the true interaction rate, the uncertainties on the 'signal' model would be largely irrelevant as the true rate would be extracted. These uncertainties are presented in this analysis to highlight the range of distributions that the model is capable of describing[2]. The large, blue uncertainty bands in Figure 5.5 do not themselves constitute a problem for using this sample to constrain true lead-target neutrino interaction physics.

[2]One dimensional projections of correlated uncertainties are almost *always* misleading. For better results, combine with a covariance matrix and evaluate a test statistic that takes account of the correlations.. In Figure 5.5, the χ^2 per analysis bin, as in Eq. 3.3, is 3.64, which corresponds to quite poor agreement.

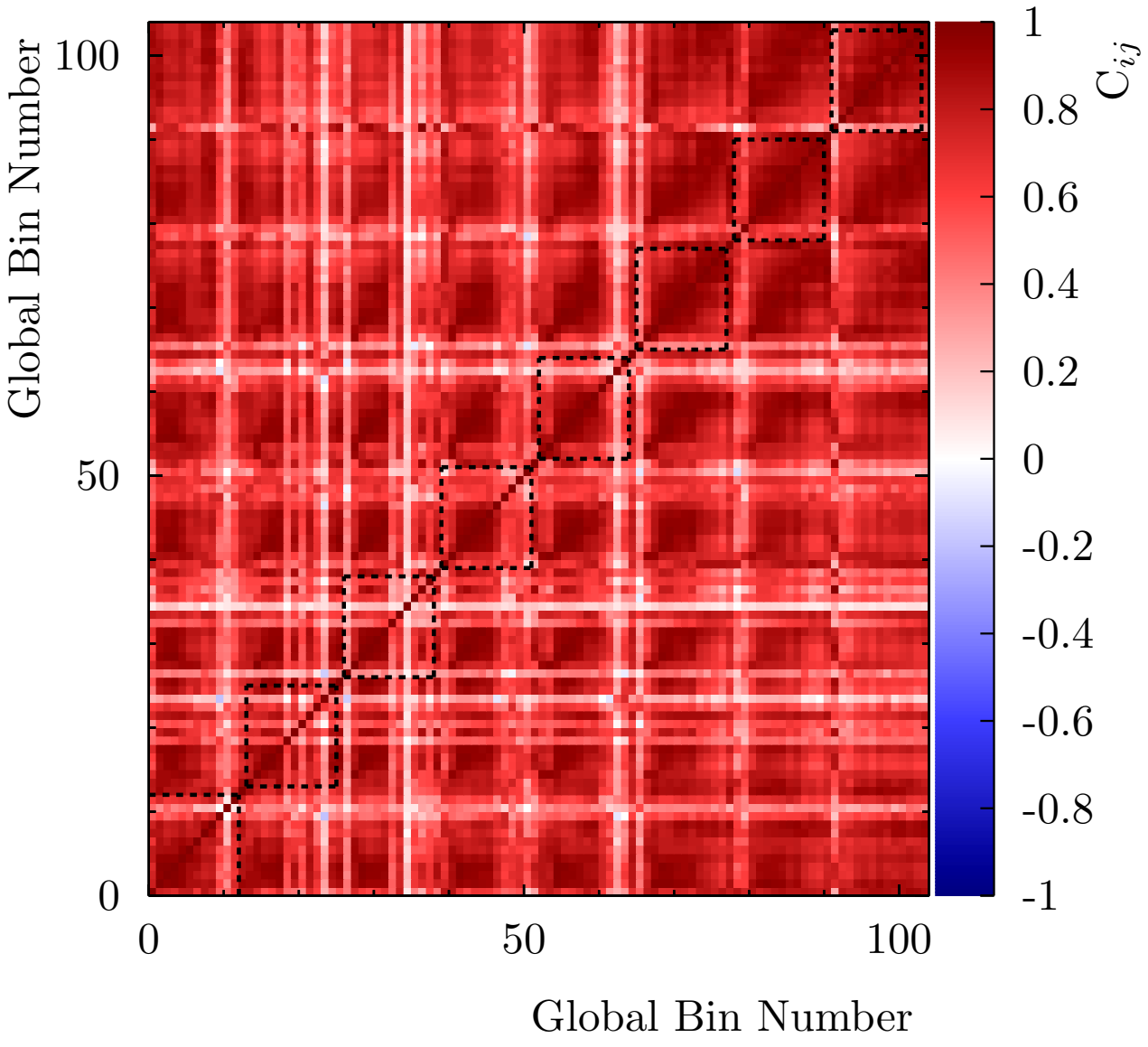


Figure 5.6: The correlation matrix determined from the combined effect of all interaction uncertainties in golden sample analysis bin number. The errors can be seen to be strongly correlated between analysis bins, this should be taken into account when assessing the goodness of fit of the nominal model to the observed data.

5.3 ND280 detector

A number of uncertainties related to the performance of the ND280 detector were assessed. The next section briefly describes the inclusion of Tracker uncertainties, which were not developed by the author. The following section describes the assessment of the ECal–tracker matching systematic which was developed for this analysis.

5.3.1 Tracker reconstruction

The performance of the TPC momentum reconstruction and particle identification is a source of error. The Tracker uncertainties used in this analysis are capable of inducing variations in the reconstructed momentum of a global track and the TPC PID discriminators used to decide if a track is muon-like. As a result of these variations, selected event candidates can move between analysis bins and candidates that were originally selected may be cut because the PID hypothesis has changed. The variations arise from uncertainties in the simulation of the ND280 magnetic field within the Tracker, uncertainties in the reconstructed momentum scale determined from a given TPC deposit, and uncertainties in the particle ionisation energy deposit and charge response of the TPC. Variations of the reconstructed momentum within the momentum resolution, which for TPC measurements is a function of the true momentum component transverse to the magnetic field, are also accounted for. The Tracker uncertainties that were deemed appropriate for inclusion in this analysis were TPC momentum resolution, TPC momentum scale, magnetic field distortions, and TPC PID variations. The investigations into how best to parameterise and constrain these uncertainties were performed during the development of the T2K oscillation analysis near detector samples and are described in detail in § 5.2 of Ref. [98].

The distribution of induced momentum variations over the ensemble of toy throws is shown in Figure 5.7. The binning is chosen to be wide enough to avoid significant bin-to-bin migrations. However, such migrations are not intrinsically a problem as long as they are well modelled—they are accounted for in the uncertainty covariance matrix.

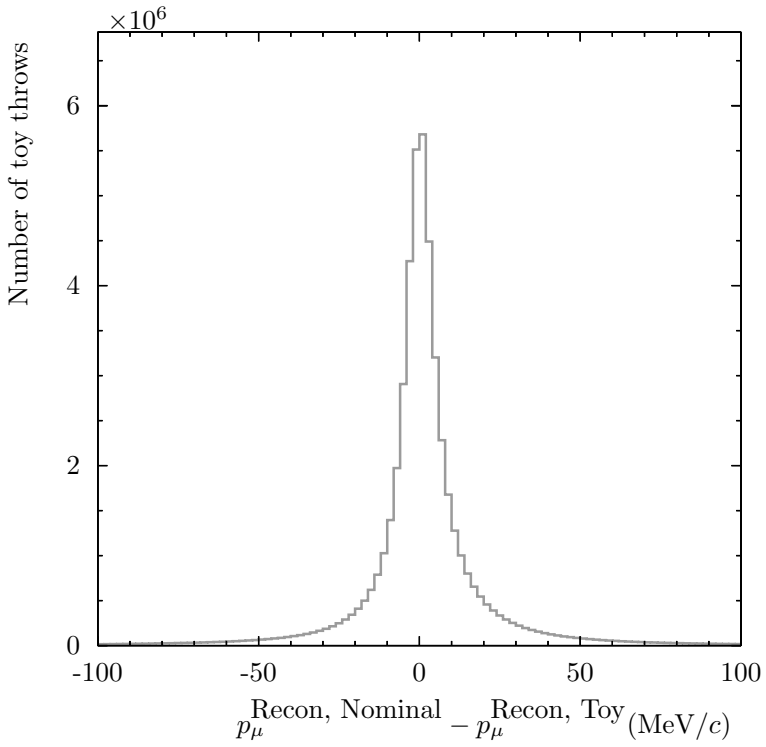


Figure 5.7: The reconstructed momentum variations, induced by the TPC momentum scale, momentum resolution, and magnetic field distortion systematic assessment, for 1000 toy throws of the simulated golden event candidate sample.

5.3.2 ECal tracking and Tracker matching

The selection of a golden event requires the existence of a Tracker constituent track. The ND280 global matching algorithms have been discussed and assessed in § 3.2.2. If the matching or ECal tracking efficiencies differ between data and simulation, then a raw comparison between the number of selected events will results in incorrect inferences[3]. While it is not possible to assess the true matching efficiency for data, an efficiency-like quantity can be constructed which motivates correction and uncertainty propagation.

To assess any difference in the matching and ECal tracking efficiencies between data and simulation, a control sample of Tracker deposits that appear to be entering or leaving a barrel ECal module was selected. The Inward Muon particle gun sample (*c.f.* § 3.1.4) was used to determine the volumes that contain the starting positions of Tracker segments that were left by vertex-like deposits occurring within the Barrel ECal. The closest reconstructed Tracker

[3] As discussed in § 4.1.2 and § 4.2.1

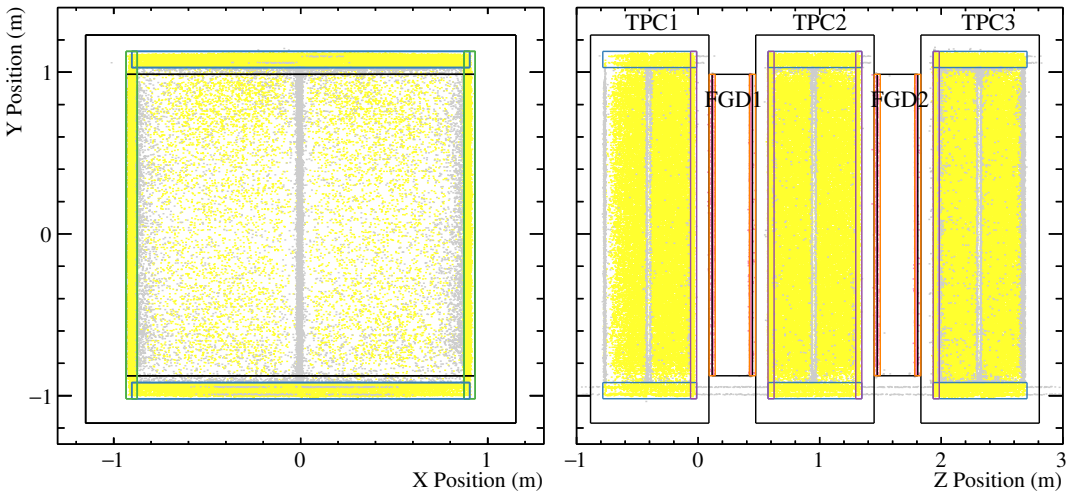


Figure 5.8: The reconstructed starting and ending positions of tracker segments of global tracks. Twenty volumes were used to select track segments which appear to be entering the tracker. Yellow markers demarcate positions that were reconstructed as occurring within one of the TPC-active face volumes, magenta markers show those that were reconstructed as entering or exiting an FGD module, and grey markers highlight tracker segments that started elsewhere. The unselected positions (grey) appear to cluster around the TPC central cathode (XY view) and micromegas gaps (YZ view); this suggests that they are mostly segments of broken, through-going tracks that failed to be matched together by the global track reconstruction.

position to a sample of fake ECal vertices is shown in Figure 5.8. Volumes at the outer faces of each TPC, with the exception of the upstream-most face of TPC1 and downstream-most face of TPC3, were used to select deposits that appear to be entering, or exiting, each TPC module. Upstream and downstream volumes were also used for each FGD, but because of the geometry of the detector it is rare for an FGD segment to be the first Tracker segment in a global track that enters or exits a Barrel ECal module.

The selected Tracker segments may be segments of a longer reconstructed track or the whole of an isolated Tracker track. The only selection criteria is that the start or end of the Tracker segment of the full track is reconstructed as occurring in one of the volumes shown in Figure 5.8. For each of these candidates, the reconstructed track direction at the selected position is used to define a 'search' direction. The search direction corresponds to the reconstructed direction that the particle came from. Practically, if the dot product of the reconstructed direction with the vector between the selected segment end and the other end of the segment is positive, then the search direction is the negative of the reconstructed particle direction, otherwise the reconstructed direction is used. If a single global track has a Tracker segment which both

starts in one selection volume and ends in another, then both ends will be entered into the control sample separately. The search direction is used to define a straight line extrapolation from the selected segment end. If this line enters a Barrel ECal module, then the Tracker segment is said to be pointing at the Barrel ECal and is selected. The simple straight line extrapolation was used to preclude the inclusion of biases from the RECPACK propagation used in global matching. In this way, a control sample for the global ECal–Tracker matching is formed without a reliance on the same models used to decide when tracks should be matched during the reconstruction. A straight line extrapolation becomes a very poor estimate of the trajectory of a charged particle in the magnetic field after a short distance, therefore only Tracker segments which start or end near to the edge of the active region of a TPC module were selected.

The true matching efficiency should be a function of the true particle momentum and polar angle at the entry point to the TPC. The contents for the control sample selected from the Run 3-equivalent simulated data set and the Run 3Cb real data set is shown in Figure 5.9. The sample is projected onto reconstructed Tracker momentum and search direction. As the global momentum is a function of the matching efficiency, if a track is matched to an outer detector object, the reconstructed global momentum will contain an energy loss compensation, using the reconstructed Tracker momentum is the only consistent approach. The simulated TPC momentum estimate was shown to be well behaved in in § 3.3.

The probability for a tracker segment that is pointing towards a Barrel ECal module to be associated to a segment within the same ECal module is designated ξ, and is shown in Figure 5.10. The agreement for regions of higher momentum is generally good, although the data statistics are lower, while the predicted efficiency for lower momentum, backward search directions[4] is over-estimated. A one-dimensional assessment in either relevant parameter would have obscured such subtleties. The agreement of the POT-scaled rate of selected candidates in the control sample, shown in Figure 5.9, is generally reasonable, but there are regions of statistically significant discrepancy in the ξ distribution. This suggests that the matching efficiency is poorly estimated for some regions of this phase space.

[4] often from forward-going tracks

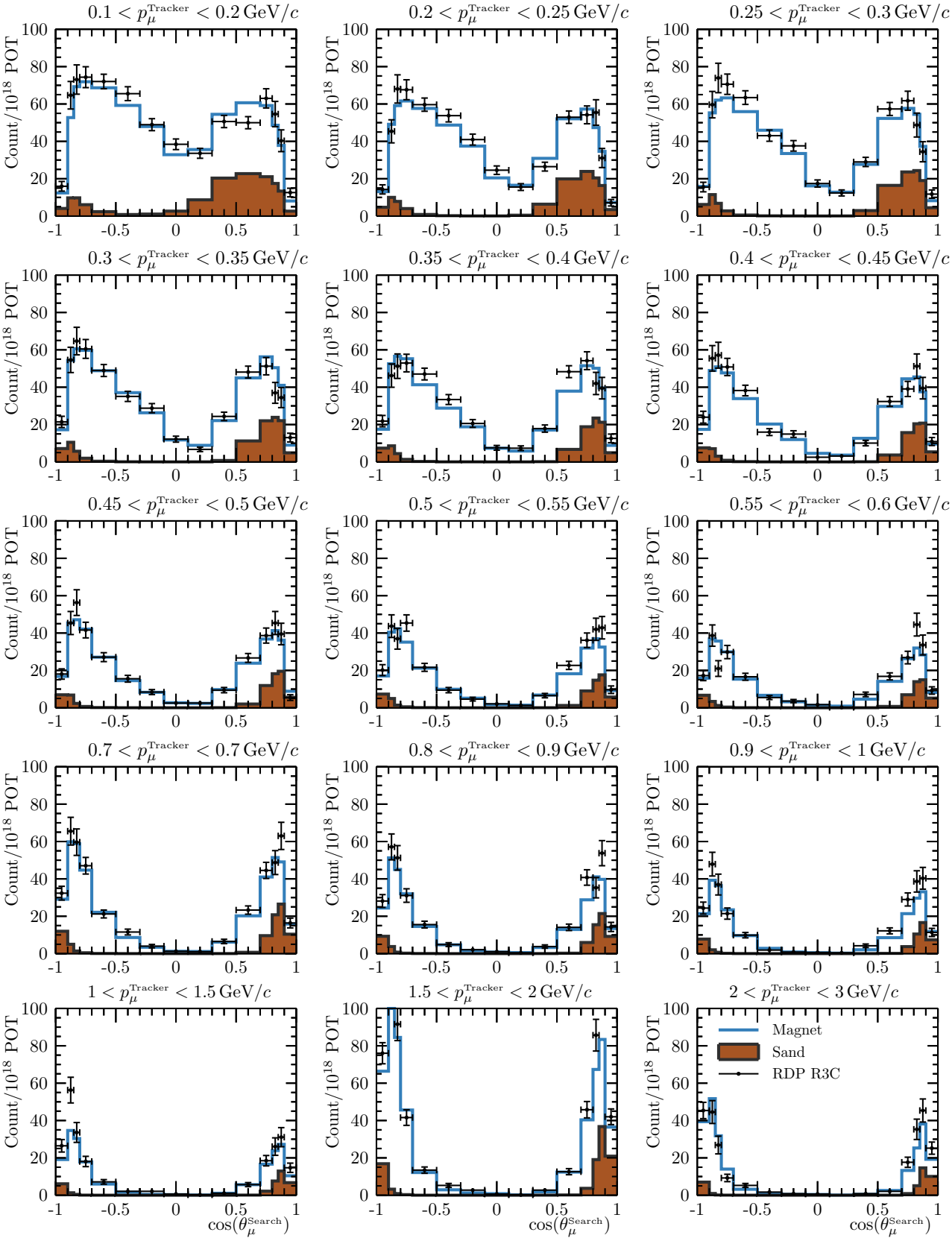


Figure 5.9: The ECal–Tracker matching control sample projected onto the Tracker segment 'search' direction polar angle and the reconstructed TPC momentum. The agreement between the simulation and data is good across these dimensions of phase space.

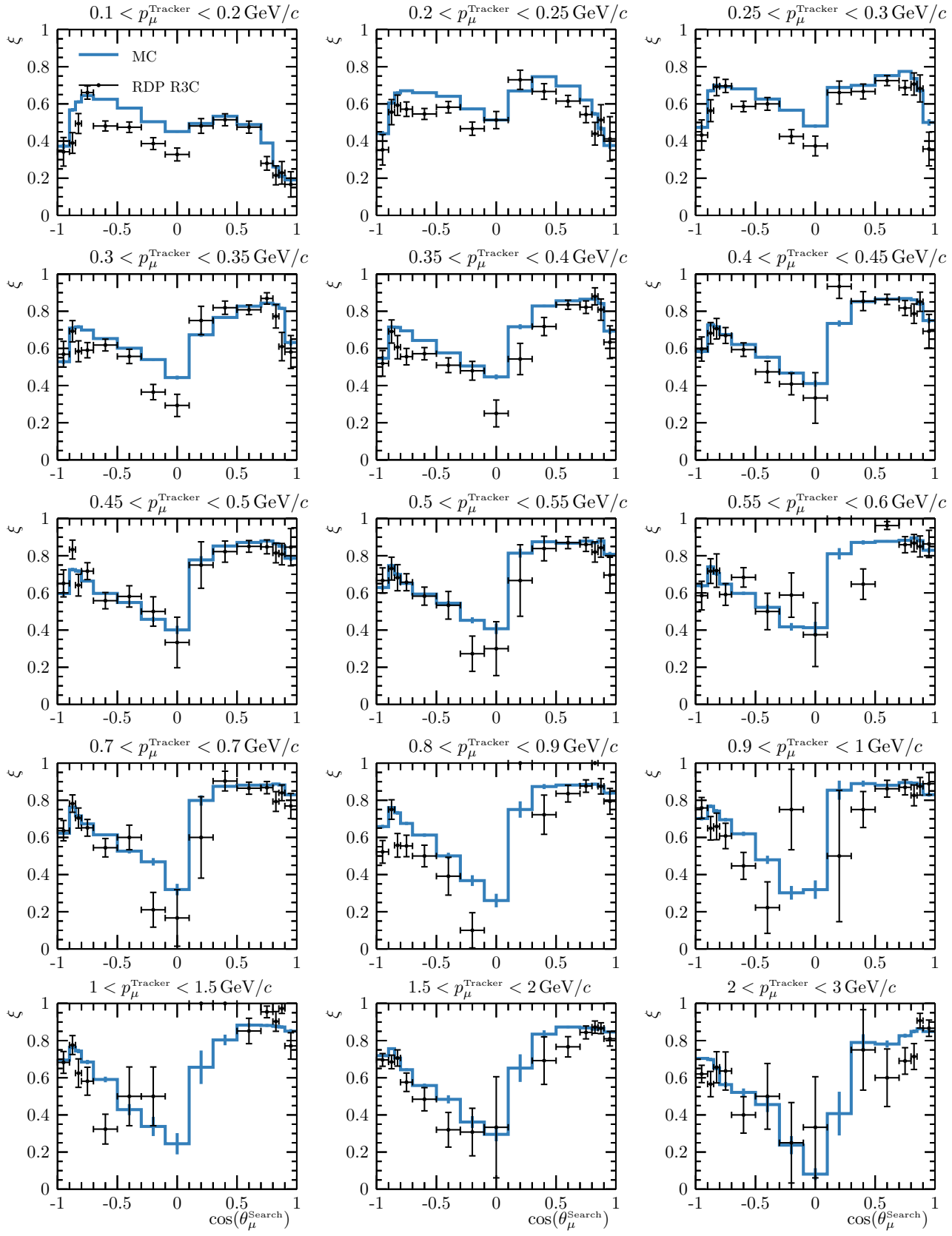


Figure 5.10: The distribution of the ECal–Tracker matching and ECal tracking efficiency-like metric. This characterises the probability that a tracker segment which appears to be entering the active region of a Barrel ECal module is associated to a reconstructed Barrel ECal track segment in the correct module. The difference between ξ in data and the simulation is used to apply an efficiency-like correction to the predicted event rate.

The ξ metric encapsulates the systematic uncertainty assessment. The application of the findings were separated into a correction and an uncertainty. For a given track with both ECal and Tracker segments, the Tracker segment end closest to the reconstructed vertex is found. Only the closest tracker segment end is used as the correction aims to account for tracks that were matched and as a result selected in the simulation, but that may not have been in data or *vice versa*[5]. The reconstructed search direction and Tracker momentum of the selected segment end were then used to find the appropriate bin in the $\xi\left(p_{\mu}^{\text{TRACKER}}, \cos\left(\theta_{\mu}^{\text{SEARCH}}\right)\right)$ distribution (*c.f.* Figure 5.10). A correction weight was calculated from the ratio of the central value for the data distribution with the central value for the simulation, in the relevant ξ bin. This weight attempts to correct for regions of poorly-estimated matching efficiency. The uncertainty on this correction was then propagated by an ensemble of toys which throw the correction weight, calculated for each simulated track, to random values distributed within the combined statistical uncertainties of the data and simulation.

In this assessment, neutrino beam data were used so that the relevant regions of the matching parameter phase space were adequately populated. It would be of interest to further assess the matching using more neutrino beam data, as well as cosmic muon data, to investigate whether the findings remained correct with a larger, more varied data sample.

5.3.3 ECal mass

Manufacturing processes come with associated tolerances, and each scintillator bar and each lead absorber sheet used in the construction of the Barrel ECals will vary within those tolerances. This manifests as an uncertainty on the target mass, to which the neutrino interaction rate is proportional, for a given neutrino flux. The uncertainties in the dimensions of each ECal component are presented in Ref. [67].

The uncertainties on the dimensions of the ECal bars and lead absorbing sheets are given in Table. 5.4. The dimensions of the scintillator bars and absorber sheets are the only non-neglible

[5]Were time of flight, or Michel tags to be introduced into the selection then the matching efficiency of the other Tracker segment end would need to be accounted for when assessing systematic affects in the the time of flight or Michel tagging efficiency.

Component	Property	Top/Bottom module	Side module
Layer construction	Number ∥ layers	15	
	Number ⊥ layers	16	
Lead absorber	Number /layer	2	4
	Width /mm	765^{+4}_{-0} (764)	2330^{+4}_{-0} (2288)
	Height /mm	$1.75^{+0.1}_{-0.1}$	
	Length /mm	3858^{+4}_{-0} (3856)	964.5^{+4}_{-0} (964)
Scintillator bar	Number ⊥ layers	96	
	Length ⊥ /mm	1520 ± 0.1	2280 ± 0.1
	Number ∥ layers	38	57
	Length ∥ /mm	3840 ± 0.1	
	Width /mm	$40^{+0.0}_{-0.4}$	
	Height /mm	$10^{+0.0}_{-0.4}$	

Table 5.4: The dimensions and manufacturing tolerances of the Barrel ECal scintillator bars and lead absorber sheets. ∥ and ⊥ refer to the layer orientation, whether the length of the bars are aligned parallel or perpendicular to the detector $\vec{z}$ axis. The GEANT4 geometry uses the nominal values except for where specified by bracketed values. The differences are negligible except the width of the lead absorber sheets which is simulated as 1.8% too short.

Components	Module	Design /kg	Simulation geometry /kg
Scintillator bars	Top/Bottom	1878	1904
	Side	2817	2856
Lead absorbers	Top/Bottom	3581	3627
	Side	5454	5432

Table 5.5: The component masses for the two Barrel ECal module orientations, as calculated from the detector simulation geometry and the design specifications given in Ref. [67].

contribution to the total active-mass uncertainty. The geometry used in the detector simulation differs slightly from the survey presented in Ref. [67]. The only notable difference is the width of the lead sheets used in the side modules, which is simulated as 42 mm shorter than was detailed in Ref. [67].

The quality assurance that the scintillator bars underwent ensures an upper-limit to the dimensions of bar cross-section, the lead sheets were manufactured with minimum dimensions and associated single-sided uncertainties. It is unclear from Ref. [67] how best to simulate the variations when throwing toy ECals—should the quoted uncertainties be used as a width for a Gaussian distribution, or did the quality assurance result in a uniform distribution within the quoted tolerance?

The component mass extracted from the geometry and the nominal mass from Ref. [67] are compared in Table. 5.5. The differences in the absorber mass can be explained by the lack of any antimony contamination in the simulated geometry and the different side-module absorber layer widths. The difference in bar mass can be attributed to a discrepancy in the bar density when determined from the geometry compared to when calculated from the plastic composition, fibre channel radius, and fibre composition taken from Ref. [67]. Unfortunately, due to a bookkeeping error, the T2K technical note that described the ECal geometry implementation was not available for further investigation. However, the most significant effects of these uncertainties will be apparent at the edges of each active region. The conservative fiducial volume and unbiased vertex position reconstruction help to mitigate edge effects. Within the bulk of

each layer, the ECal bars will have been packed as closely together as possible, which should reduce the effects of any uncertainties except the layer height uncertainties.

The uncertainty on the layer height was propagated to the analysis bins by 1000 uniform toy layer throws within the uncertainties quoted in Table. 5.4. The ratio of the thrown width to nominal width was used as an event weight for interactions on the absorber layers and plastic scintillator bars separately. The throws were handled artificially conservatively so that for each toy throw, every layer component grew or shrunk to the thrown size; in reality some layers would be slightly too thick and some too thin in a single ECal module. The effect of the mass uncertainty is less prominent than the simulation statistical uncertainties in every bin.

5.3.4 Propagated uncertainties

The combined effect of the ND280 detector uncertainties is shown in Figure 5.11. The significant detector uncertainty, apparent in the backward-most momentum bins, arises mostly from the TPC reconstruction uncertainties. The correlation matrix for the combined detector uncertainties is shown in Figure 5.12. Some analysis bins can be seen to be anti-correlated, which is expected as momentum variations induce bin migration between nearby momentum bins. The forward regions of correlated uncertainty are caused by the assessment of the ECal mass uncertainty. In the distributions presented in this section, it appears that the matching correction generally pulls the prediction closer to the data. The interaction model tuning, shown in Figure 5.5, shows a similar overall reduction in the predicted event rate. in the next chapter the fully tuned prediction is compared to the high statistics data sample.

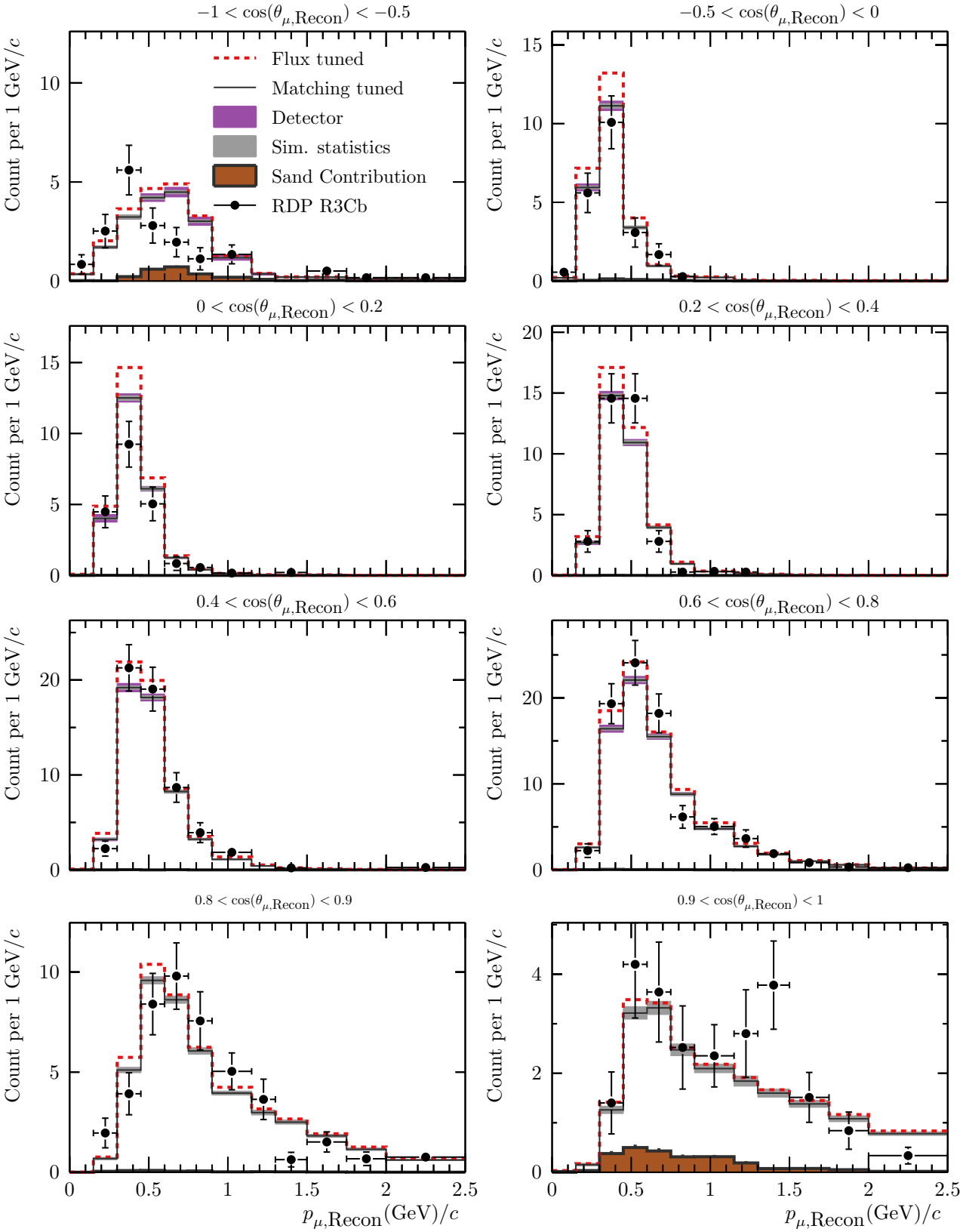


Figure 5.11: The combined effect of the ND280 detector uncertainties propagated to the golden interaction sample. The effect of the ECal–Tracker matching efficiency correction is also shown.

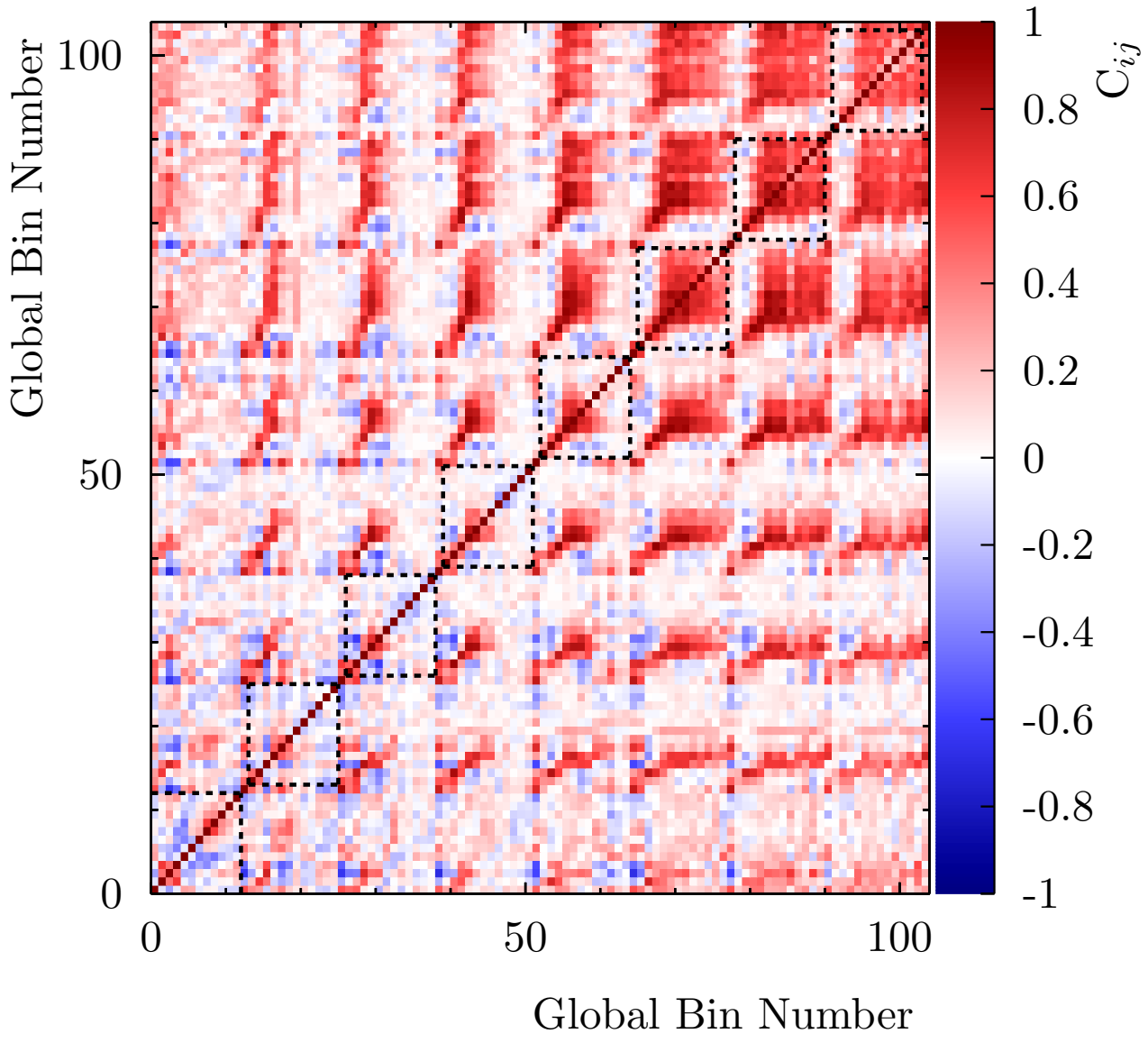


Figure 5.12: The correlation matrix calculated from the propagation of the ND280 detector uncertainties in golden sample analysis bin number. The dotted regions correspond to the polar angle bins shown in other figures. The visible anti-correlations occur because of bin migrations induced by the TPC momentum measurement uncertainties.

Chapter 6

ECal-as-Target analysis results

6.1 The data sample

In the previous two chapters, the four selections and the uncertainty treatment have been presented in the context of a reduced data set. The full comparisons use a larger data set that comprises the rest of T2K Run 3C and the whole of T2K Run 4. At the time of writing, this constitutes the majority of the fully-processed neutrino-mode data. Run 1 and 2 are excluded to keep the detector configuration consistent after the damage sustained to the right side ECal at the end of Run 2—both of these run periods only contain a comparatively small amount of integrated POT.

The Run 4 data set is subdivided into 'Air' and 'Water', which correspond to the contents of the PØD water bags—this distinction makes no direct difference to the analysis presented here, however, the intensity of the J-PARC proton beam increased between the two run periods.

6.2 Selected event rates

The reconstruction-level comparison of the golden sample events for the full data set is shown in Figure 6.1. The overall agreement of the analysis data set with the predicted distribution

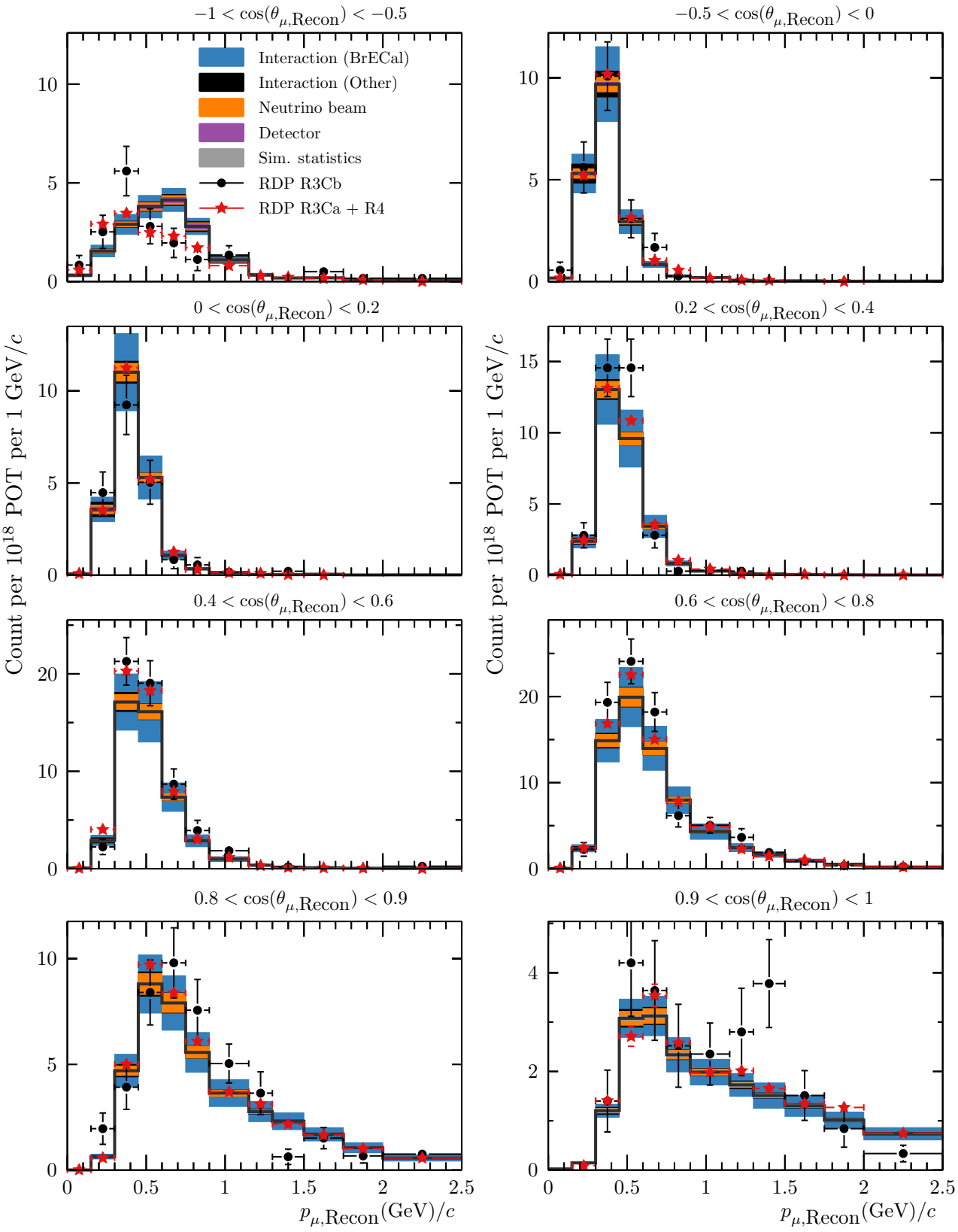


Figure 6.1: The golden sample reconstructed muon kinematic distribution for the simulation and the two sets of real data used in this analysis. Note that the interaction uncertainties on true barrel ECal events (blue) would not be included when extracting the event rate but are presented here to cover the distributions that the model is capable of describing.

Name	Dates	POT /10^{18}	Sample content/10^{18} POT			
			Golden	Silver	Upstream	Magnet
Run 3Ca	12/[04/08–05/25]	111	49.5(7)	951(3)	421(2)	1304(3)
Run 3Cb	12/[05/27–06/09]	23.8	51(1)	967(6)	420(4)	1304(7)
Run 4 Water	12/10/19–13/02/06	162	48.3(5)	903(2)	418(2)	1289(2)
Run 4 Air	13/[02/06–05/08]	176	45.6(5)	883(2)	416(2)	1287(2)

Table 6.1: The number of selected candidates per 10^{18} POT analysed for each sample. A deficit of events is statistically evident for the two Run 4 data sets compared to the two Run 3 data sets.

is acceptable—while the higher statistics data appear to be within the predicted uncertainty band for the majority of bins, many of the error sources result in strongly correlated event rate uncertainties.

It may not be clear from Figure 6.1, but the full analysis data set has a lower normalisation per POT than the Run 3Cb data. The number of selected event candidates within each sample, subdivided by run period, is given in Table. 6.1. The fractional discrepancy for the golden and silver selected candidates in the Run 4 Air data set is about 10%, which is significantly larger than the statistical uncertainties. This indicates some change in conditions between the run periods that the ECal interaction selections emphasise—note that the entering background selections seem less affected.

The effect of the silver selection cuts on this run-dependent data discrepancy is shown in Figure 6.2. The silver sample was chosen as it exhibits smaller statistical uncertainties than the golden sample. The discrepancy can not be resolved by the removal of any single selection cut. However, if all of the veto cuts are ignored, the selected event rates per POT agree between the data sets to within statistical uncertainties. The apparently unifying effect of ignoring the fiducial volume cuts is most likely due to a dilution of the source of the discrepancy, as opposed to a resolution. The most plausible explanation for the observed discrepancy is an increase of coincidence, or 'pile up', veto triggers. This occurs when an unrelated energy deposit, from a separate neutrino interaction, results in a veto flag that causes an otherwise well-reconstructed ECal interaction to be discarded. The number of deposits reconstructed to start within the

Barrel ECal fiducial volume, as a function of proton beam intensity, is presented in Figure 6.3. The fraction of such event candidates which are associated to a veto flag can be seen to be linearly related to the beam intensity, which has risen between and during T2K run periods. It is likely that this increase is mostly because of coincident veto triggers, as opposed to correct triggers which reject poorly reconstructed background. In the simulated data set, the veto cuts reject more background than signal which motivated their use in § 4.2.1. This can also be seen in Figure 6.3.

Unfortunately, each run-equivalent simulated data set is generated using the mean run proton beam intensity. For Run 3, the simulated beam intensity is one in which the real J-PARC beam effectively never ran. When the rate of multiple interactions within the active detector in a single beam bunch is low, this is not a significant worry. However, it is clear from Figure 6.3 that even at the Run 3 proton beam intensity, this is already not the case. Furthermore, as the sand muon simulation is generated completely separately from the rest of the neutrino beam simulation, it is not currently possible to analyse the effect of sand muons entering the detector within the same beam bunch as an interaction within ND280. For analyses that must use harsh veto cuts to identify rare or complex processes, this constitutes a significant deficiency in the simulation. The veto fraction can be seen to be poorly predicted within the simulated data set and it is possible that the separation of ND280 interactions and sand muons is a significant contributor to this mis-modelling. This deficiency will continue to become more apparent as the J-PARC proton beam intensity increases.

As can be seen in Figure 6.2, the Barrel ECal veto is involved in generating this discrepancy. This veto is practically the same as the requirement that a selected interaction candidate is the most upstream reconstructed deposit in the ECal subsystem—a cut that was used in the 2015 analysis. It is possible that the lack of simultaneous sand simulation contributed to the unsatisfactory results of the analysis presented in Ref. [96].

Under the assumption that the discrepancy between data-taking periods is mostly as a result of coincidence veto flags, the shape of the data should be unaffected; the net effect is an increase in well-reconstructed interaction candidates being vetoed by unrelated physical processes. The

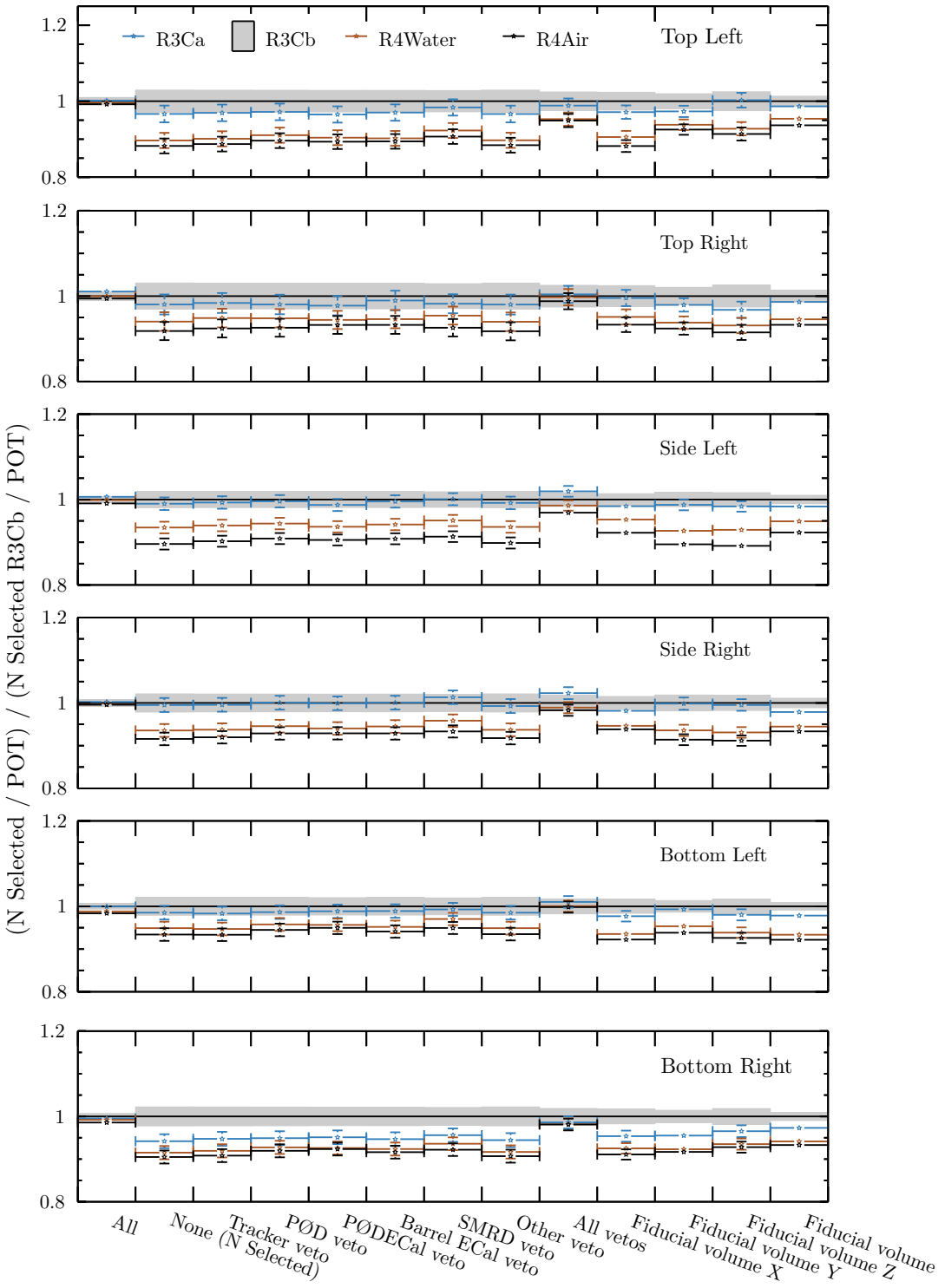


Figure 6.2: The effect of the silver selection cuts, and combinations thereof, on the total rate discrepancy between the different data taking periods. Each column shows the total number of selected events in the sample when removing the corresponding selection cut(s). The two left-most columns show the total number of ECal vertex candidates before any selection and the number of selected candidates given the full silver selection. The uncertainties on each data point are the statistical uncertainties for the relevant sample. The removal of all veto cuts resolves the discrepancy to within statistical uncertainties.

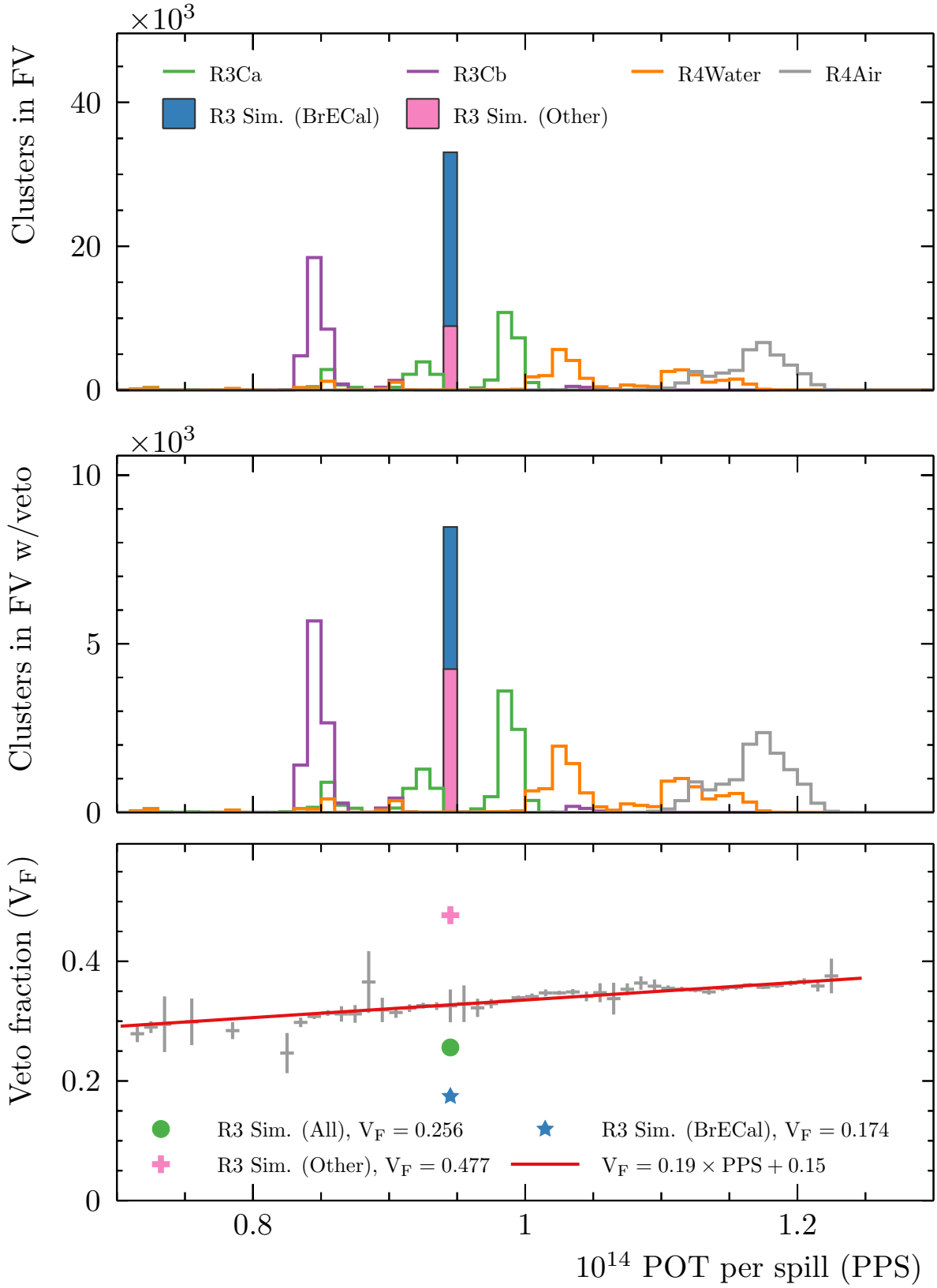


Figure 6.3: The veto fraction as a function of proton beam intensity. The number of vertex candidates reconstructed as starting in the fiducial volume and the subset for which an upstream veto was triggered are shown in the top two panes. A straight line fit to the ratio of these distributions extracted a slope of 0.188(7) and an intersect of 0.147(7). The simulated neutrino beam data are generated with a fixed, run-average proton beam intensity and predicts the observed veto fraction poorly.

form of V_f(PPS) presented in Figure 6.3 could be used to reweight the observed events to reproduce a constant veto fraction. However, as the veto fraction is apparently not well modelled in the simulation, the relative contribution of correct vetos and coincidence vetos to the total fraction is difficult to infer with any confidence. Therefore, modifying the data to emulate the coincidence rate determined from the simulation seems counter-productive. If a rate comparison was needed, it would be better to run the simulation with a range of proton beam intensities and reweight the simulated intensity distribution to match the real distribution. Then the rate uncertainty resultant from coincidence vetos would just effect the prediction and not be shared between a poorly-justified data modification and an imperfect simulation. Modifying the neutrino vector generation to simulate a range of beam intensities was outside the scope of this thesis, but this work highlights that the deficiency should be addressed.

As the overall normalisation cannot be relied upon, the following section presents shape-only comparisons of the final results for the interaction sample. These comparisons take into account the off-axis variations by including events from each of the Barrel ECal modules. The shape of the charged-current inclusive muon momentum spectrum is sensitive to the contributions from different interaction channels, so little information is sacrificed by the shape-only treatment.

6.2.1 Data–simulation comparison

The shape-only comparisons for the golden, silver, magnet-entering, and upstream-entering candidate samples are shown in Figure 6.4, Figure 6.7, Figure 6.8, and Figure 6.9 respectively. The correlation matrix corresponding to the the total, shape-only, propagated event rate uncertainty is shown in Figure 6.5. The fractional rate uncertainty from the most prominent error sources can be seen in Figure 6.6.

For the golden sample, the shape-only correspondence between the data and the prediction in the angular bins other than the backward-most bin is surprisingly good. While the reconstruction is less trustworthy in the backward-most bin, the apparent shape difference is intriguing. As discussed in §4.2.1, the true content of this angular bin is mostly mis-reconstructed, forward-going events. Processes that cause migration to this bin are not easily distinguishable with the

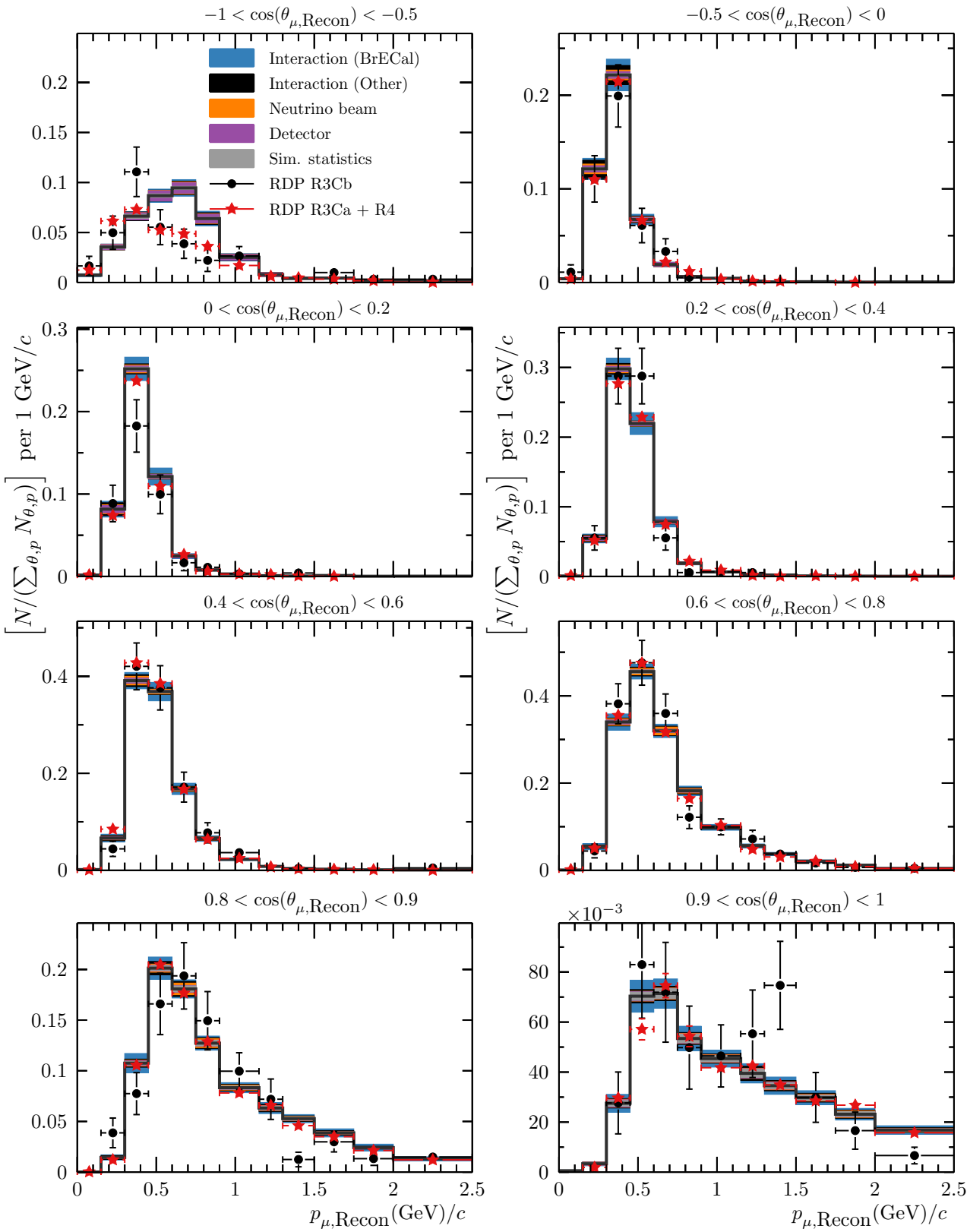


Figure 6.4: The shape-only distribution of reconstructed muon kinematics for all selected events in the golden sample. The whole two dimensional distribution is normalised so that the sum over all angular and momentum bins is unity—this includes a momentum overflow bin for each angular bin. Within each angular bin, the distribution is then scaled down by the bin-widths to give a consistent one dimensional histogram.

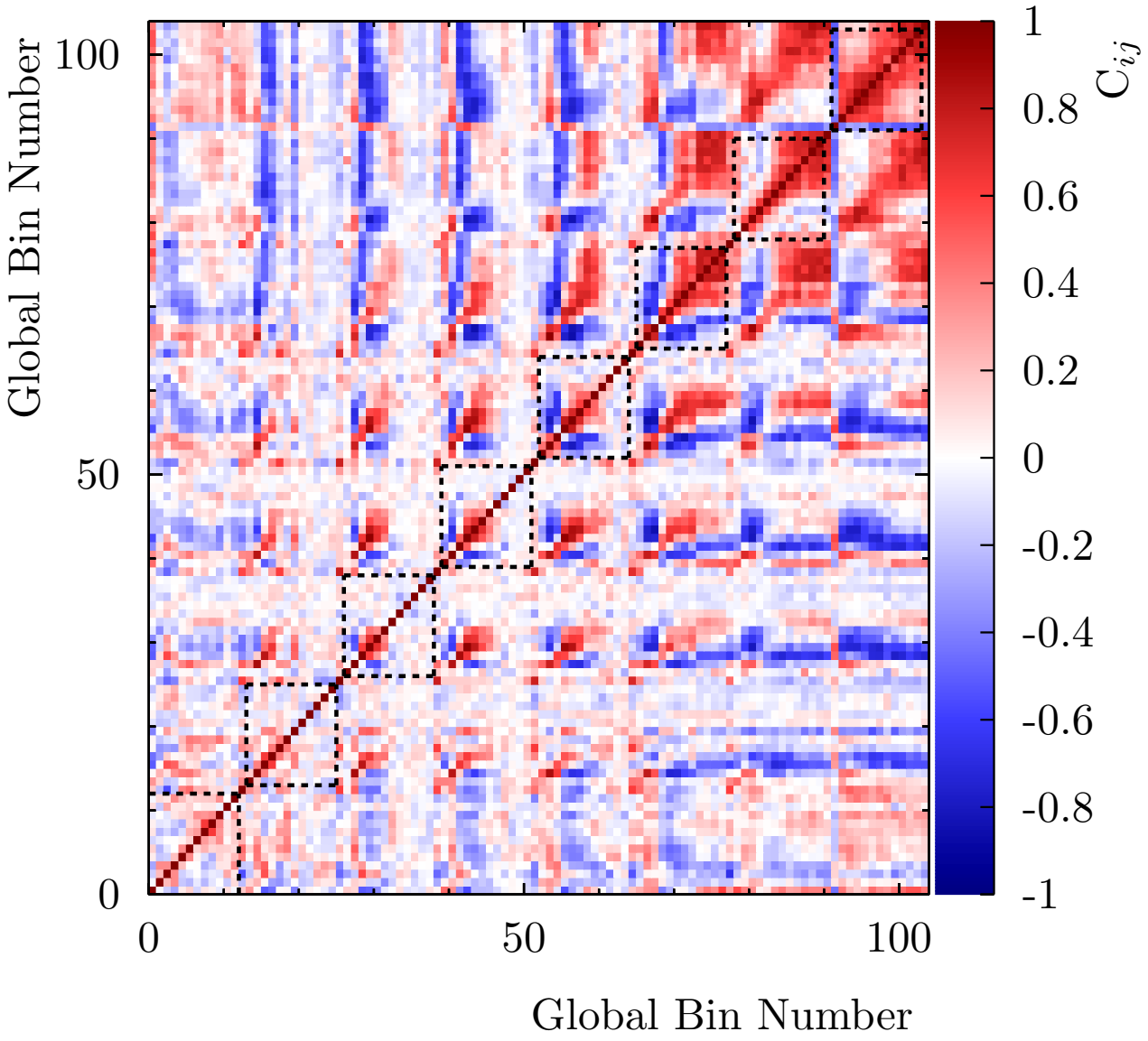


Figure 6.5: The total shape-only correlation matrix determined from all of the error sources described in the previous chapter. The demarcated areas correspond to the separate angular panes shown in Figure 6.4.

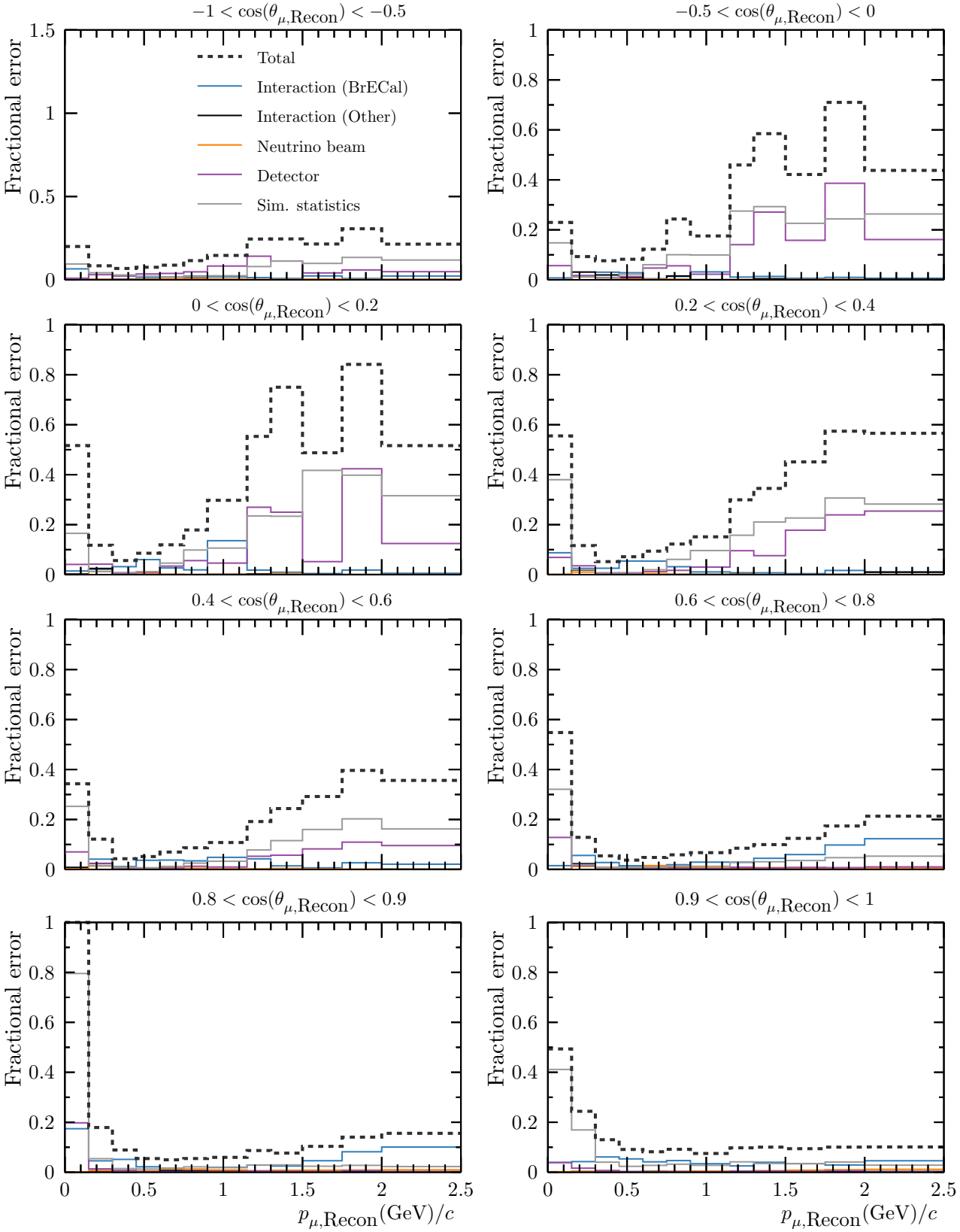


Figure 6.6: The fractional shape-only uncertainties, as a function of analysis bin, from the most important error sources described in §5.

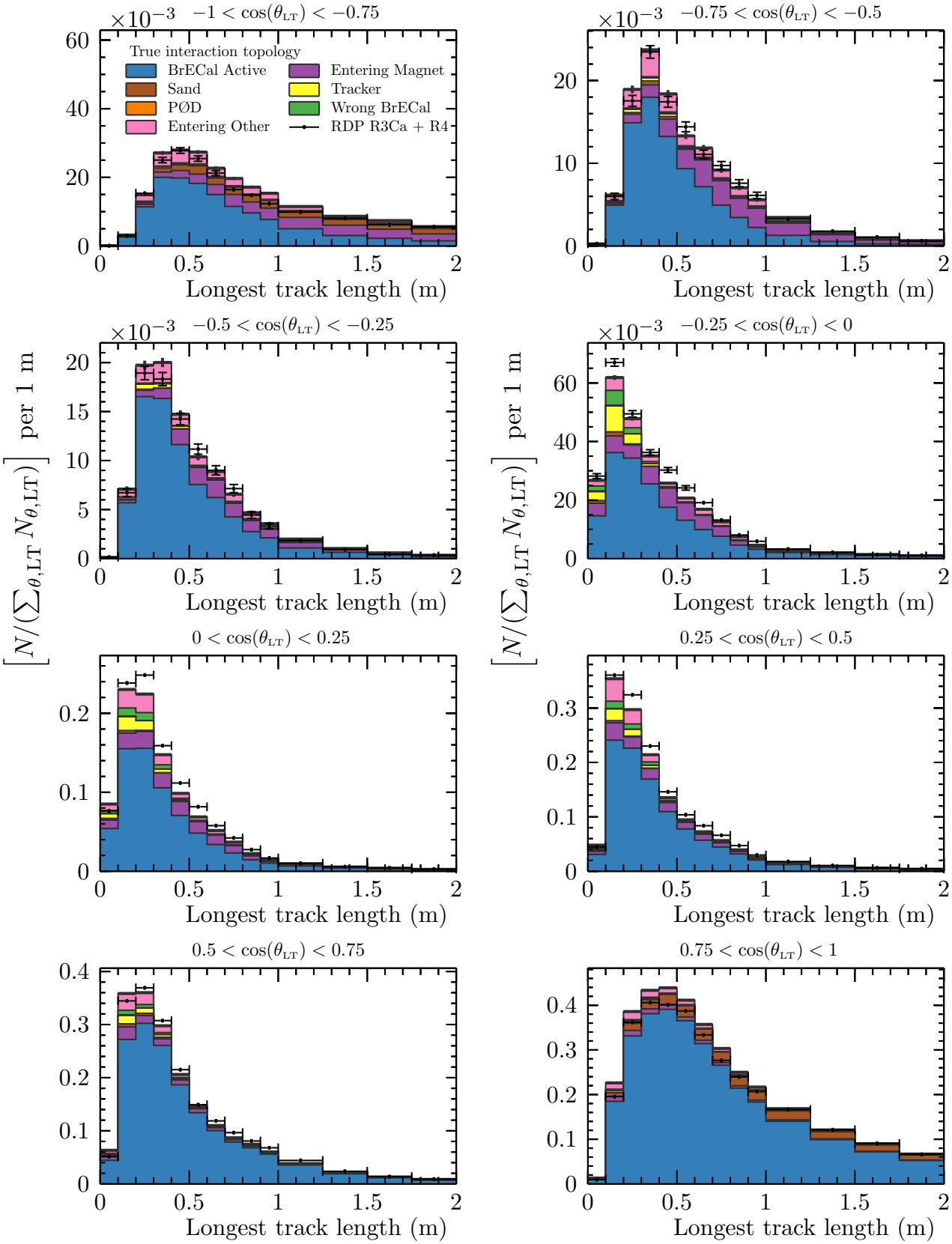


Figure 6.7: The shape-only distribution of longest reconstructed track properties for all selected event candidates in the silver sample. The distribution is normalised as in Figure 6.4. The non-sand components are tuned according to the interaction tuning described in § 5.2.

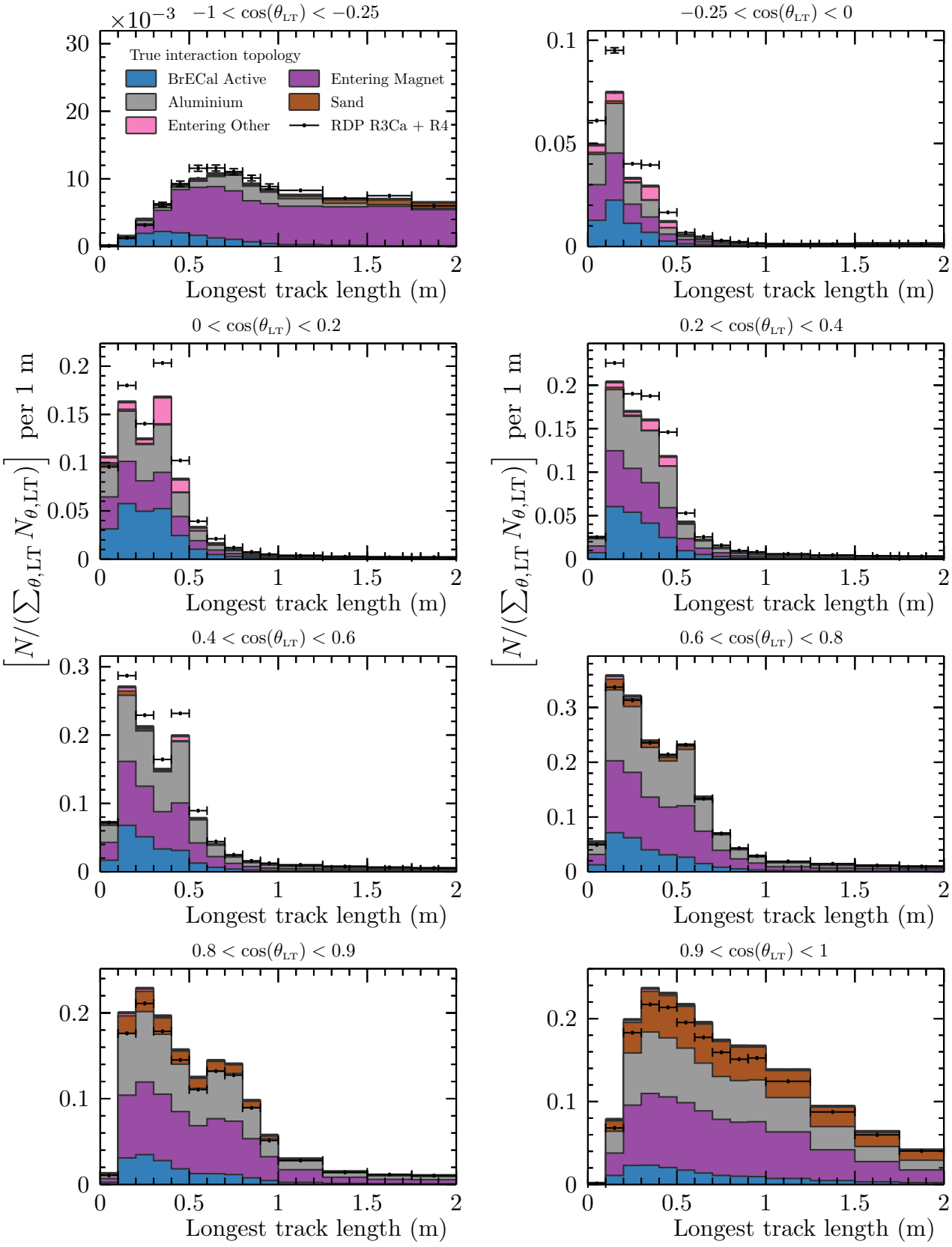


Figure 6.8: The shape-only distribution of longest reconstructed track properties for all selected entering background candidates in the entering magnet sample. The distribution is normalised as in Figure 6.4. The non-sand components are tuned according to the interaction tuning described in § 5.2.

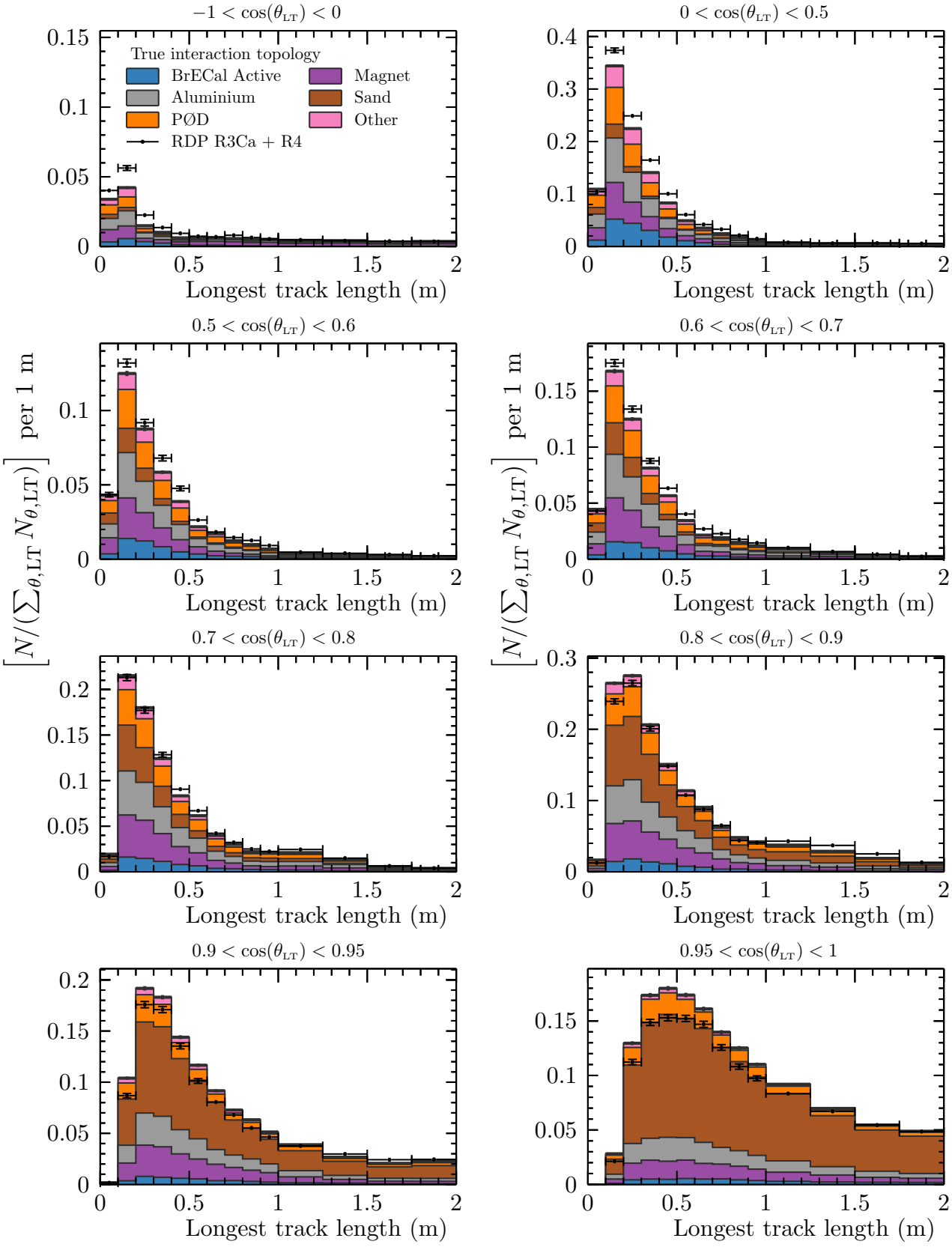


Figure 6.9: The shape-only distribution of longest reconstructed track properties for all selected entering background candidates in the entering upstream sample. The distribution is normalised as in Figure 6.4. The non-sand components are tuned according to the interaction tuning described in § 5.2.

current reconstruction, but this shape difference strongly motivates continued improvement to the reconstruction of backward-going final state particles.

A Pearson χ^2 test statistic was calculated to evaluate the compatibility of the prediction and associated uncertainties with the observed data. The test statistic was calculated from the simulated and observed data in Figure 6.4 and the associated covariance matrix as

$$\chi^2 = \sum_{i,j} \left(\vec{s} - \vec{d}\right)_i V_{ij}^{-1} \left(\vec{s} - \vec{d}\right)_j , \tag{6.1}$$

where $\vec{s}_i$ and $\vec{d}_i$ are the selected event rate in analysis bin i for the simulation and the observed data respectively. Each bin that contained at least one simulated event was included in the evaluation, this includes the momentum overflow bin. For the full reconstructed phase space, the test statistic was calculated to be $\chi^2/\text{NBins} = 245/104 = 2.35$. When restricting the reconstructed muon polar angle to $-0.5 < \cos\left(\theta_{\mu,\text{Recon.}}\right)$, in order to investigate the sub-sample with more reliable reconstruction, it was calculated as $\chi^2/\text{NBins} = 122/91 = 1.34$. This corresponds to reasonable agreement between the prediction and the observed data. Note that the treatment is shape-only; had the overall data set normalisations been included, the correspondence would be significantly less good because of the poorly-modelled veto fraction discussed in the last section. The discrepancies in both the backward bin and the overall normalisation highlight that more investigation is needed to be confident in the modelling of interactions occurring within the ECal.

In the other samples, the agreement is comparable to that seen for the Run 3Cb data set presented in previous chapters. For the forward-most bins, which contain significant sand muon contribution, the prediction is often an over-estimate of the observed data. For a given POT, the separation of the sand and ND280 interaction simulations may allow for extra candidates to be selected. Bunches where both a sand muon and a true neutrino interaction within the Barrel ECal leave a deposit in the same module should only contribute a single candidate to the sample, rather than one for each separate prediction, which is possible with the current simulation. This effect may also be partly to blame for the discrepancy in the second peak evident in the predicted distribution of deposit hit multiplicity (*c.f.* Figure 2.15a), which may

have contributed to the unsatisfactory results in the 2015 analysis. For the magnet-entering sample, the simulation significantly under-predicts the data in the high-angle and backwards-going bins.

The distribution of $E_{\text{rec}}^{\text{QE}}$ for a restricted phase space of muon production angle is shown in Figure 6.10. The reconstructed neutrino flux shape can be seen to vary across the detector. The overall correspondence between the simulation and the data is reasonable, but the relative peak rates and peak energies do differ outside of the shape-only uncertainties. Such a sample may provide a useful constraint of the off-axis variation of the T2K neutrino flux. While $E_{\text{rec}}^{\text{QE}}$ is a biased neutrino energy proxy for non-QE interactions, it combines the two reconstructed properties of the selected muons into a physically meaningful quantity. The counter-intuitive variation of the peak rate in each module is because of differences in the detector acceptance and golden sample selection efficiency between the barrel ECal modules. Robust efficiency and purity 'corrections' are outside the scope of this thesis, and quite probably not the preferred treatment for such a sample, as will be discussed in the final chapter. However, to recover the expected event rate variation across the Barrel ECal modules, the average selection efficiencies and purities from Figure 4.23 and Figure 4.28 were used to adjust the golden and silver integrated event rates for each module. The results of this adjustment are shown in Table. 6.2. Here the relative module rates are ordered as expected from Figure 2.10; the highest fiducial mass-normalised event rate is found in the bottom left module, which is closest to the axis of the neutrino beam. The top right and side right modules are predicted to have approximately the same mass-normalised rate, but in the adjusted data the rates differ significantly. More detailed study of the module-by-module event rates and an investigation into the constraining power of such a sample on the off-axis flux prediction would be of great interest.

For completeness, the distribution of reconstructed azimuthal angle for the golden sample is shown in Figure 6.11. The ND280 geometry and the golden sample requirement of an associated Tracker track confine the contributions from each module into different regions of azimuthal angular projection. No extra data–simulation discrepancies are apparent in this projection.

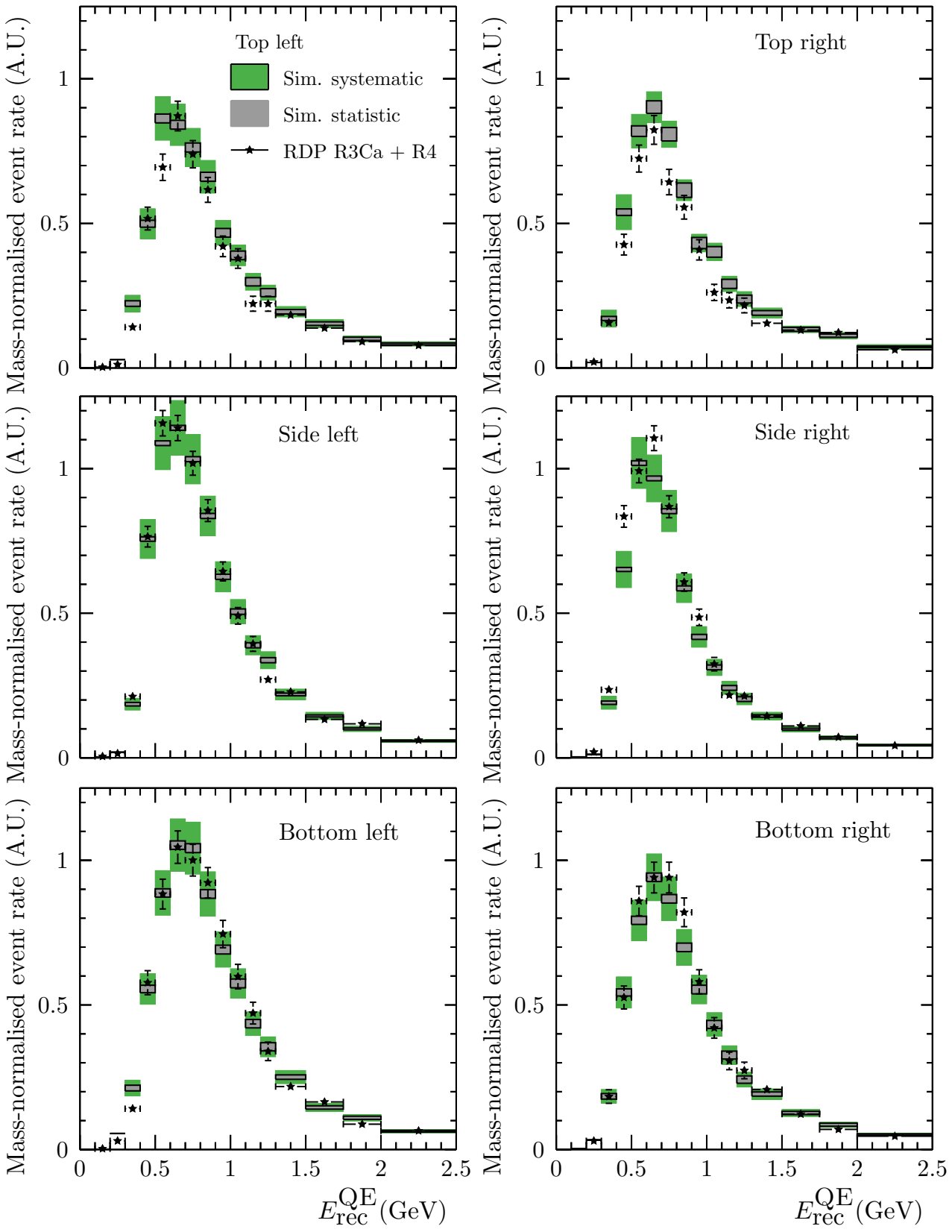


Figure 6.10: The fiducial mass-normalised, shape-only distribution of reconstructed neutrino energy under the assumption of CCQE kinematics. The reconstructed muon angle is restricted to $-0.5 < \cos(\theta_{\mu,\text{Recon.}})$ to investigate the regions of optimal reconstruction performance. The off-axis variation of the rate and width of the reconstructed energy distribution can be seen across the modules.

ECal Module	Adjusted number of selected interaction candidates			
	Sim.		Data	
	Count / 10^{20} kg POT	Fraction of Bottom left	Count / 10^{20} kg POT	Fraction of Bottom left
Top left	22.9(0)	0.51	22.2(1)	0.53
Top right	21.4(0)	0.47	19.7(0)	0.47
Side left	27.0(0)	0.60	24.3(0)	0.58
Side right	21.5(0)	0.47	20.6(0)	0.50
Bottom left	44.8(1)	1	41.2(1)	1
Bottom right	38.3(0)	0.85	35.8(1)	0.86

Table 6.2: The number of simulated true interactions per mass and POT in each of the Barrel ECal modules. The calculated efficiency and purity for the golden and silver samples in each module (Figure 4.23 and Figure 4.28) were used to adjust the observed event rate in a model-dependent way. The ordering of the mass-normalised module rates is as expected given that the axis of the beam passes closest to the bottom left module.

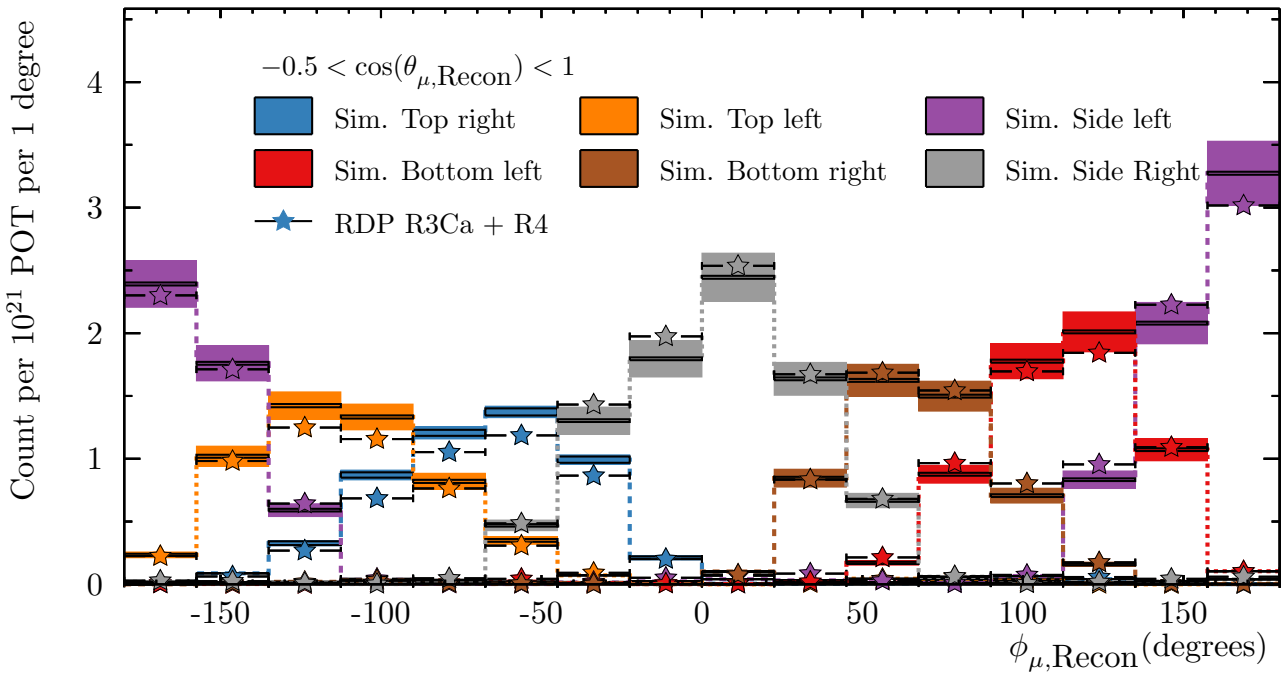


Figure 6.11: The shape-only distribution of reconstructed muon azimuthal angle for the golden sample. Selected muons that originate from each Barrel ECal module contribute to a characteristic region of phase space. The data–simulation correspondence is comparable to other projections presented previously.

Chapter 7

Conclusion

The analysis presented in this thesis was designed to sample neutrino interactions occurring within the electromagnetic calorimeters of the T2K near detector. Two previous measurements of the lead-target muon neutrino charged-current cross-section have been made using these detectors—Ref. [95] and Ref. [96]. Both analyses arrived at unsatisfactory results because of unresolved disagreement between the observed data and the simulation. In this analysis, extra emphasis was placed on remaining sensitive to localised regions of data–simulation discrepancy. The summation of distributions covering varying efficiencies and purities was approached with more care than in the 2015 analysis. Regions suspected of being problematic, such as the distributions of entering backgrounds, were investigated separately, rather than discarded. Where feasible, a statistically independent sample of data was used to motivate and cross-check choices made.

The ND280 reconstruction algorithms were adapted to incorporate the results of the ECal-isolated vertex reconstruction. This allowed information from the whole of ND280 to be used in the analysis. Four samples were developed and assessed, two interaction candidate samples, and two samples that aimed to characterise the distributions of entering background.

The observed event rate of selected neutrino interaction candidates was seen to differ between T2K data-taking periods. This was most likely caused by an increase in the rate of candidates being rejected because of upstream veto flags. The rate that candidates occurring within the

fiducial volume were vetoed was shown to increase linearly with beam intensity. As a result the overall data set normalisation was ignored for the final comparisons.

Uncertainties in the neutrino flux, neutrino interaction model, and detector response and reconstruction algorithm performance were assessed and the shape-only test statistic for the golden interaction sample was $\chi^2/\text{NBins} = 245/104 = 2.35$. The backward-most angular bin shows significantly worse simulation–data correspondence than is seen in the other angular bins and contains predominantly mis-reconstructed forward-going muons. While this discrepancy hints at modelling deficiency which would be interesting to investigate, when characterising the simulation–data correspondence for well-reconstructed event candidates, this bin may be temporarily ignored. The figure of merit for the golden interaction sample in a restricted phase space ($-0.5 < \cos(\theta_{\mu,\text{Recon}}) < 1$) was $\chi^2/\text{NBins} = 122/91 = 1.34$, which corresponds to good agreement between the simulation and the data within the estimated systematic uncertainties. Determining the failure modes of the reconstruction that result in this backward migration should be a focus for future improvement efforts.

The most important conclusion for analyses at ND280 is that the effect of using single, run-averaged beam intensities in the simulation was far from negligible for this analysis. The use of this shortcut, as well as the separation of the sand muon and ND280 interaction simulations, should be re-assessed by T2K. While the rate that sand muons are selected for Tracker-based analyses is generally insignificant, the poorly simulated effects from sand muon coincidence on selections may well be a significant source of error. This will only become more problematic as the J-PARC beam power rises.

7.1 Future improvements

The four areas most-ready for improvement in the analysis presented here are: the reconstruction for backward particles, the inclusion and systematic assessment of using track time of flight in the golden selection, the use of ECal PID, and the re-assessment of the veto flags.

While the overall rate of reconstructed backward tracks is similar between simulation and

data, the distribution of reconstructed momenta within the backward-most bin was observed to differ significantly. As seen in Figure 4.26, the vast majority of simulated tracks that were reconstructed as backward-going were truly forward-going. It would be of great interest to resolve the cause of such migration and allow the investigation of truly backward events, which are interesting as high-angle and backwards muons are sensitive to nuclear effects.

The use of track end time to re-position vertices may allow for significantly increased efficiency and purity of the golden event sample (*c.f.* § 3.4). This feature was not included here as it would warrant an involved systematic consideration and the apparent data–simulation discrepancy would need to be well-understood first. However, an example of using the time of flight to naively reposition vertex candidates is presented in Figure 7.1. For candidates which are associated to a Tracker track which appears to 'end' before it 'starts', the vertex candidate is placed at the opposite end of the track before the selection takes place. This allows vertex candidates that were determined to be reconstructed at the wrong track end to be reconsidered as entering background or as a candidate signal interaction originating in another Barrel ECal module. By including the time of flight, the number of simulated interaction candidates in the golden sample increased from 52.4(8) ND280 and 1.00(7) sand muon events per 10^{18} POT, to 54.0(3) and 0.84(6) selected events. The number of barrel ECal interactions in which a true muon was selected increased from 42.2(7) per 10^{18} POT to 45.3(7), which corresponds to a rise in purity from 79.(0)% to 82.(6)%. The majority of extra candidates were reconstructed as backward-going, which further validates approaching this selection technique with care. However, robustly incorporating the track time of flight would be a welcome improvement toward using such a sample to study events which produce high angle and backward final state particles.

The ECal PID algorithms were not designed or tuned to process the output of the vertex reconstruction. However, the separation of minimally-ionising and highly-ionising vertex tracks should, in principle, be possible. If effective, this would allow the interaction candidate samples to be further sub-divided by the existence of a pion. Such samples may provide new constraints for the near detector fit, which is performed as part of the T2K oscillation analyses, and could be interesting for interaction model builders. Samples developed to select exclusive final-state

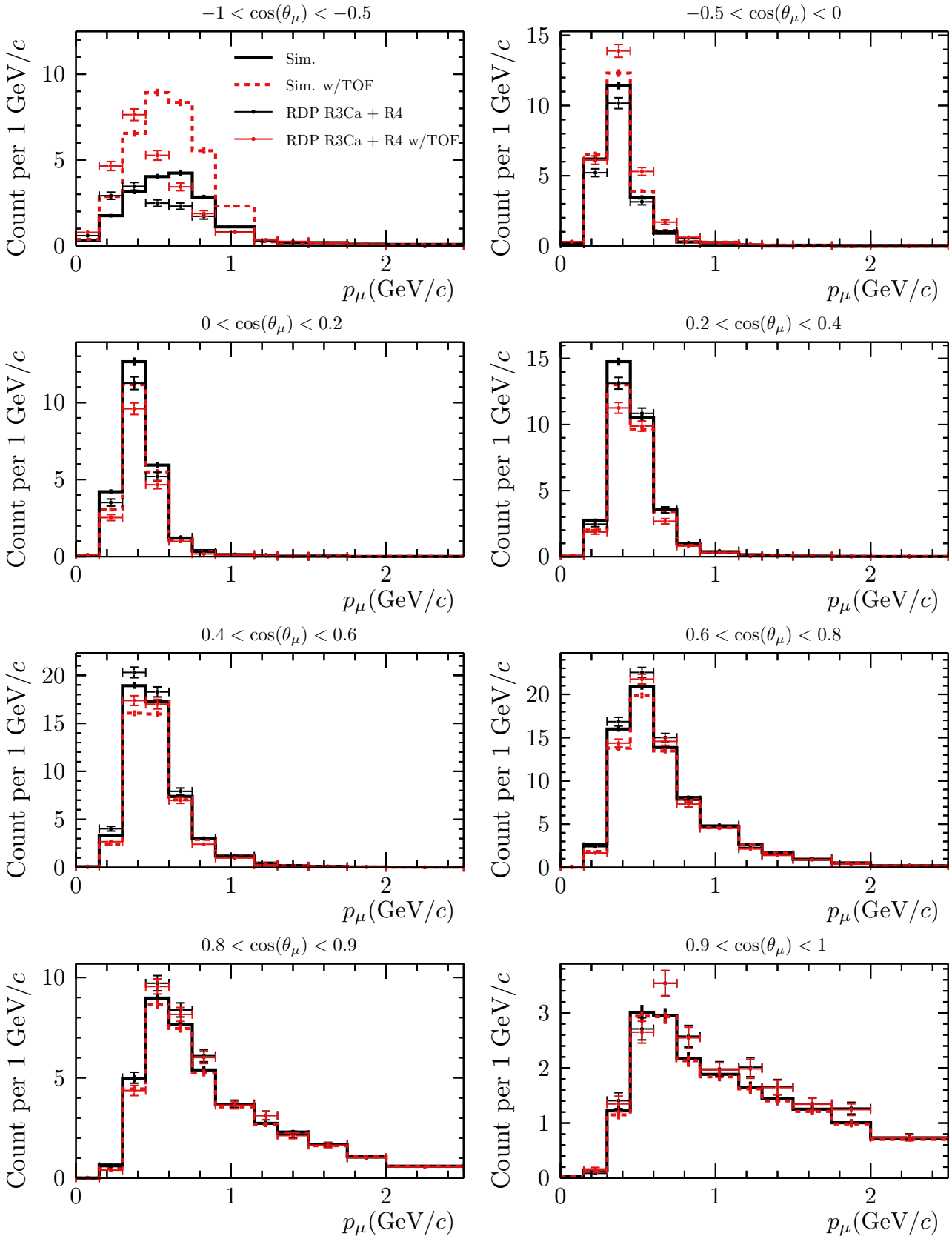


Figure 7.1: The golden interaction candidate sample where the Tracker track time of flight has been used to reposition reconstructed vertex positions to the earlier end of the track. The simulation has been rescaled so that the prediction has the same overall normalisation as the corresponding observed data.

topologies such as these may provide new insights into the effects of using nuclear targets [104].

Finally, the veto flags that contributed to the observed run period-dependent rate discrepancy may have been overly conservative. The harsh vetos were included to increase the purity of the interaction target sample, and while they do result in a higher purity sample, a robust rate comparison would provide new, useful information.

The data set used in this analysis constitutes one third of the currently recorded neutrino-mode POT. With the inclusion of the suggested modifications above, the use of more data would be worthwhile. Furthermore, there is a comparable recorded POT of processed anti-neutrino-mode data that were not used in this analysis, but in principle could be similarly investigated. To use the anti-neutrino data, studies into extra sources of background and poorly-reconstructed event topologies would be needed, however, the simultaneous analysis of neutrino mode and anti-neutrino mode data may provide a unique test of the flux prediction.

7.2 Future uses for an ECal-target sample

It is hard to envision a robust way to produce a lead-target exclusive cross section with such a sample. It is not possible to separate the interaction target on an event by event basis and so any published cross-section would necessarily be using a mixed carbon and lead target. It may be possible to use an FGD sample for carbon subtraction, but the differences in reconstruction efficiency, detector acceptance and neutrino flux shape exposure make this a very unattractive solution. Furthermore, the reconstruction and selection efficiencies for golden candidate muons are not flat in any relevant phase space dimension. As a result, the presentation as a one or two dimensional mixed-target cross-section measurement will always include model-dependent averaging over important dimensions of phase space. It would be challenging to present the four samples, in six separate ECal modules, in a way that was simple to consistently compare to. However, comparisons made at reconstructed-property level that have the access to the full detector simulation can mitigate these problems.

The T2K near detector fit provides a constraint of the cross-section model and unoscillated

neutrino flux for oscillation analysis. It uses the full reconstructed event candidates for a number of mutually exclusive event topologies and, as a result, can account for the correlated uncertainties and provide constraints on a large number of important parameters in the T2K oscillation analysis. In recent analyses, the best-fit flux prediction at ND280 is significantly increased relative to the nominal model [20]. If the golden and silver sample normalisations can be well understood, a sample of Barrel ECal interaction candidates may be able to provide a complementary constraint of the T2K flux prediction. This is largely because of the span of off-axis angles sampled and the high event rate due to the large active mass. Even if this sample is not ready to be included in the near detector fit, it provides a good model of one way to use the samples presented in this thesis for the benefit of neutrino physics. A description of the T2K near detector fit methodology can be found in Ref. [56].

One such compromise between performing a simultaneous fit of multiple samples and being able to publish a useful subset of the data, is to perform a template fit in reconstructed variables. In Analysis I of Ref. [75], the reconstruction smearing was encapsulated in templates that map between true particle kinematics and reconstructed particle kinematics. These templates can be presented as a smearing, or detector response, matrix, similar to those shown in Figure 4.25 and Figure 4.26. By including a column in the matrix that corresponds to simulated events with given true properties that were not reconstructed, the effect of the detector acceptance and reconstruction efficiencies can also be accounted for. Such a matrix might be said to characterise the "smearceptance" of the detector. These matrices can be used to apply an approximation of the detector effects to a simulated neutrino interaction vector and thus allow comparison of models to observed reconstruction-level data. Furthermore, the development of such matrices can be performed with particle guns and in many phase space dimensions, which reduces the ways in which model dependence can enter into published results. Methods of effectively including the effect of correlated systematic uncertainties into published acceptance and response functions needs more investigation. One simple method might be to produce a number of smearceptance matrices, each corresponding to a toy throw of the detector systematic parameters. However, to publish enough information so that future models could be compared to reconstruction-level data more consistently than a cross section measurement, the smearcep-

tance matrices would need to be calculated for a large number of true final state topologies (*e.g.* μ, $\mu + p$, $\mu + \pi^-$, $\mu + \mathrm{p} + \pi^+$, $\mu + 2\mathrm{p} + \gamma$, etc...). For each of these final state topologies, a full description would require three dimensions per final state particle and three dimensions for the detector position of the interaction—for example, a full description of a $\mu + p$ topology would require a nine dimensional histogram of final state properties to be well populated in regions where current and future models are likely to generate events. Without more investigation, it is not clear that this is computationally feasible. However, it may be an interesting way to present results that are more accessible and have greater longevity than unfolded cross-section measurements.

Appendix A

ECal-isolated vertexing data–simulation comparisons

A.1 ECal vertex track comparisons

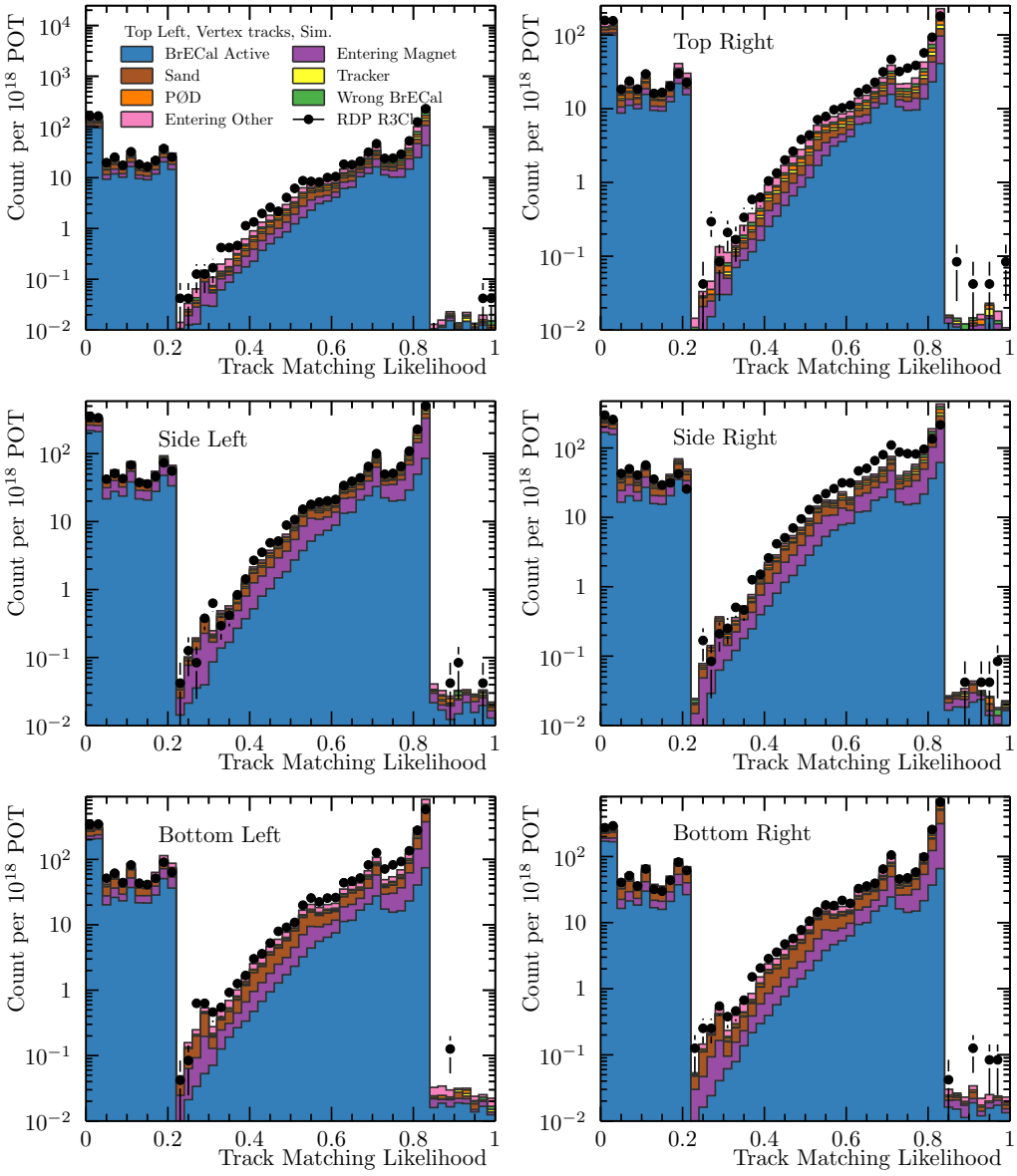


Figure A.1: The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

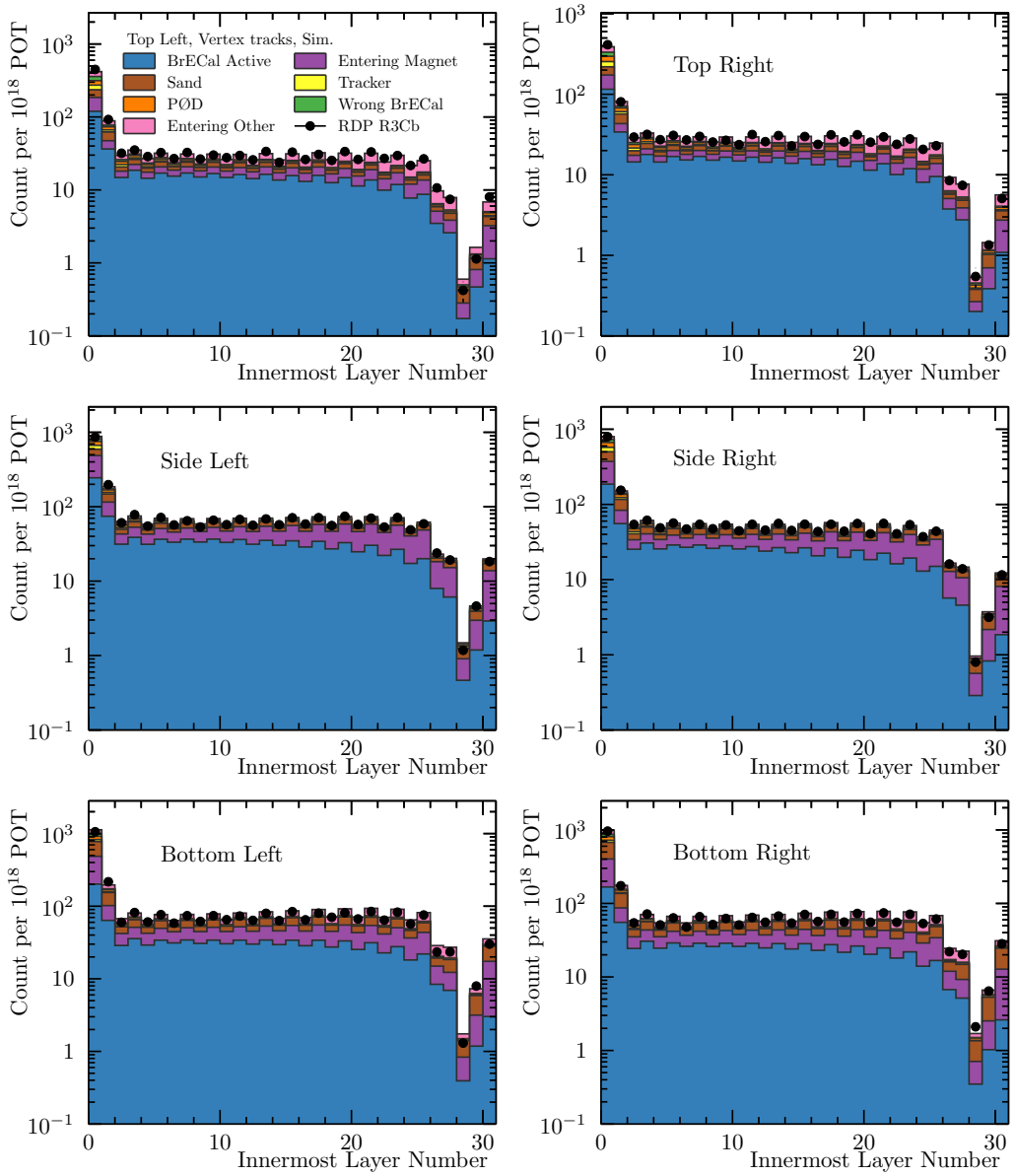


Figure A.2: The POT-normalised, layer numbers of the innermost constituent hit for all reconstructed vertex tracks reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

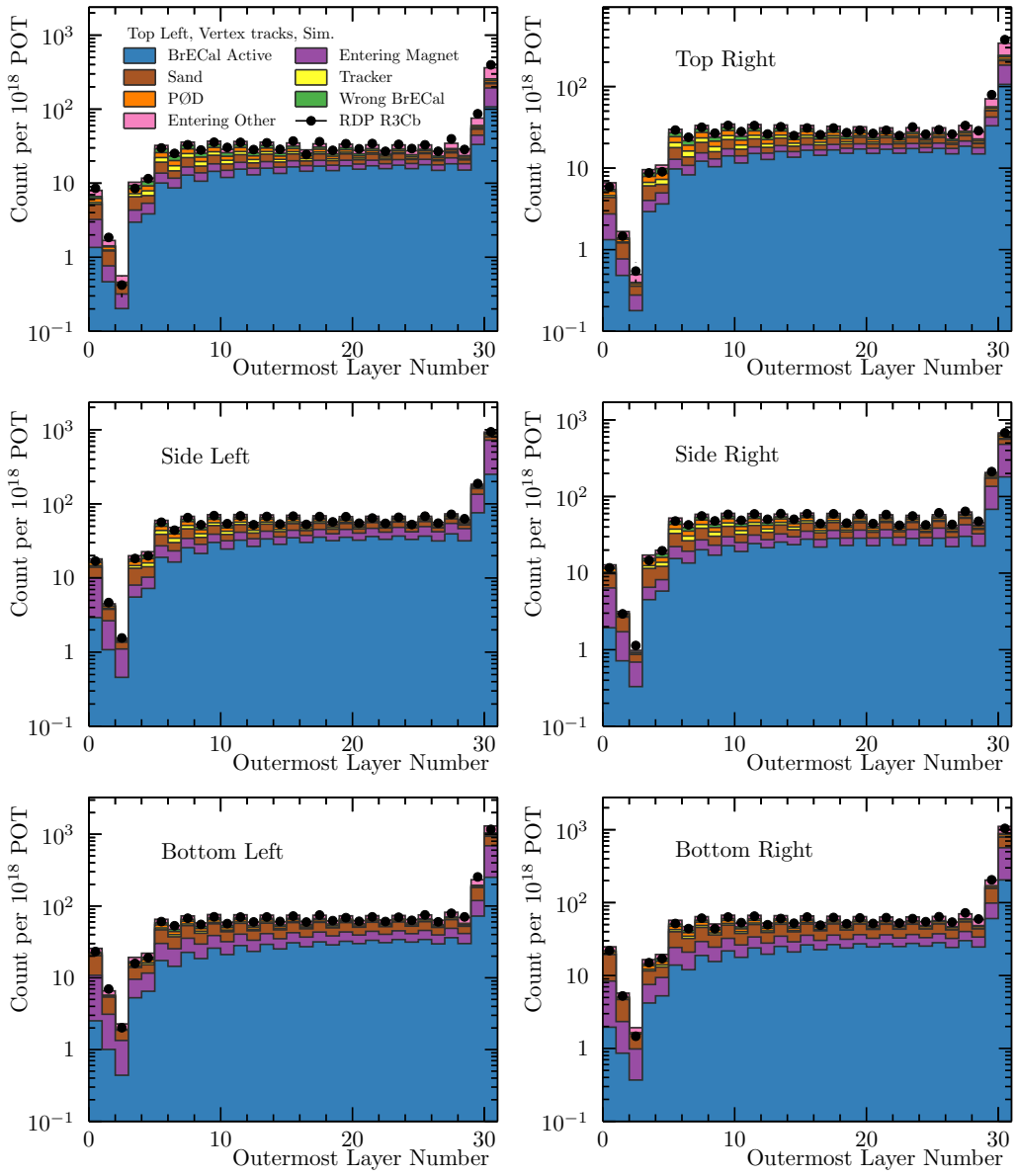


Figure A.3: The POT-normalised, layer numbers of the outermost constituent hit for all reconstructed vertex tracks reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

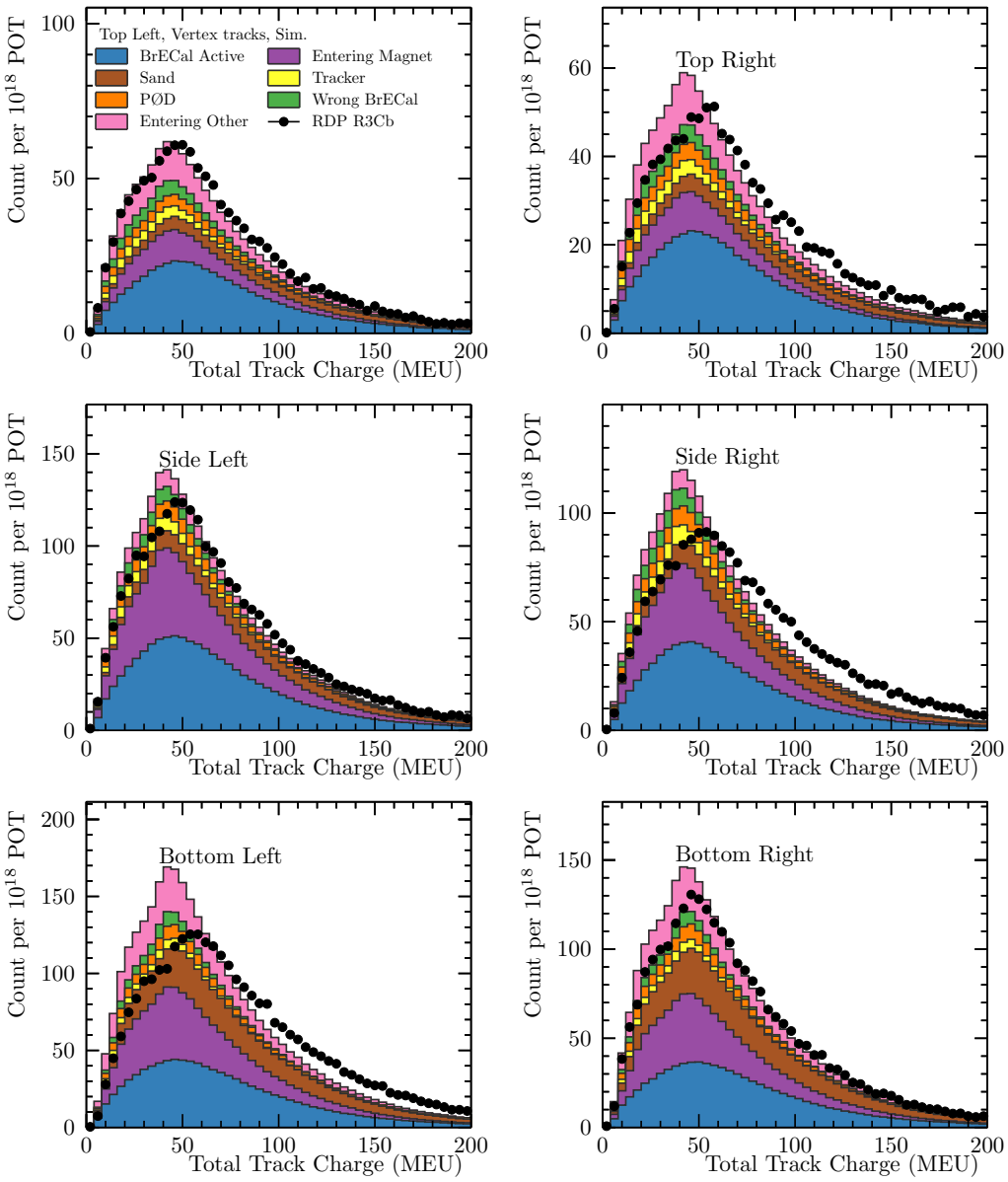


Figure A.4: The POT-normalised, total track charge in MIP-equivalent units for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

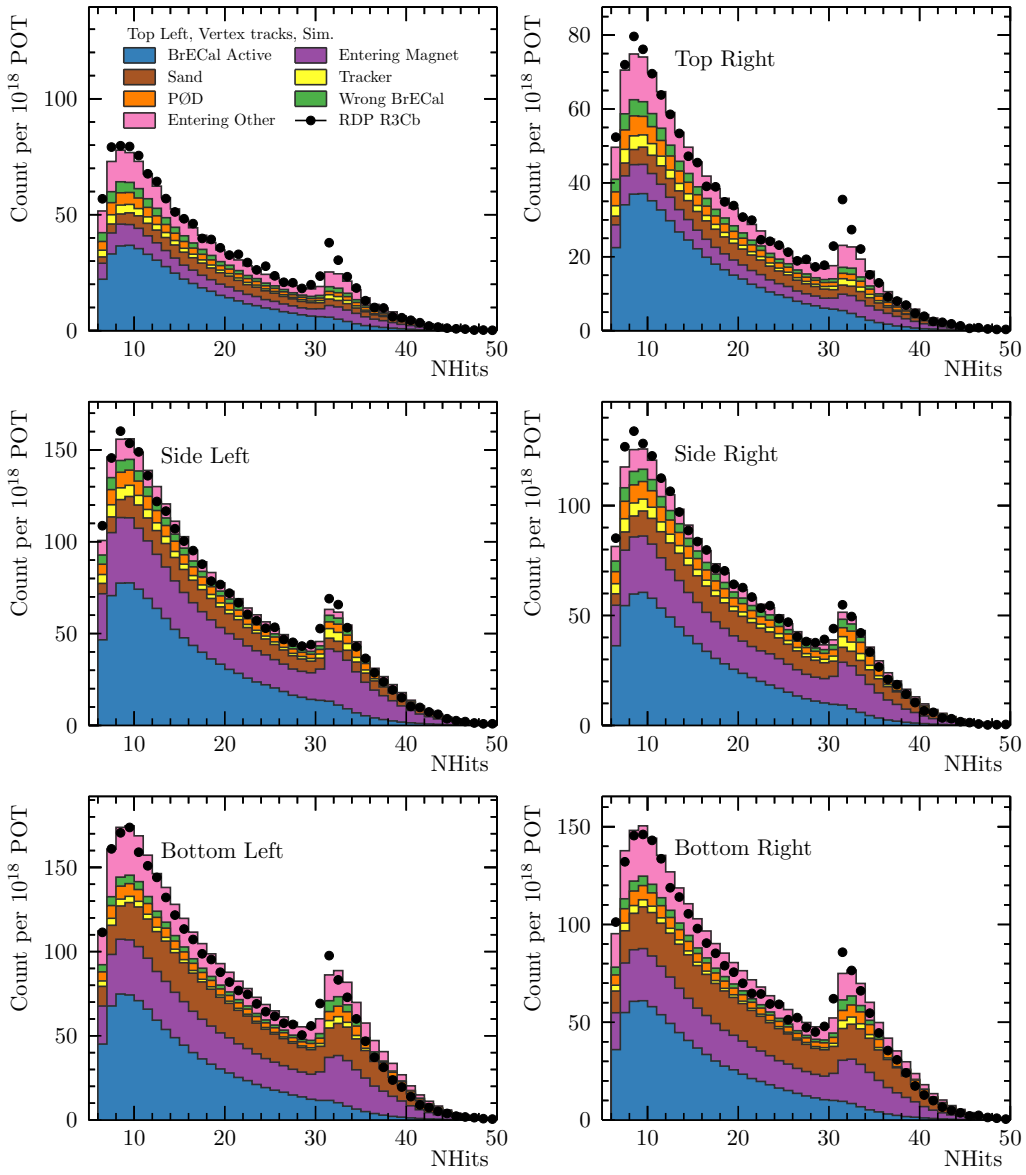


Figure A.5: The POT-normalised, number of constituent hits for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

A.2 ECal vertex candidate comparisons

The distributions of the number of hits associated to a vertex candidate, and the number of unused associated hits, ΔNHits, are shown in Figure A.6 and Figure A.7 respectively. The definition of ΔNHits used here is different to that used in [96, §7.6.1.3]. Here it is the number of hits assigned to a vertex candidate less the number of hits used in reconstructed ECal tracks. Unused hits left over from either the Hough transform, or by tracks that were rejected during the vertex candidate reconstruction are allocated to the closest candidate. This results in a difference for vertex clusters that were reconstructed into multiple vertex candidates but more-accurately presents the distribution of unused hits around each vertex-like deposit.

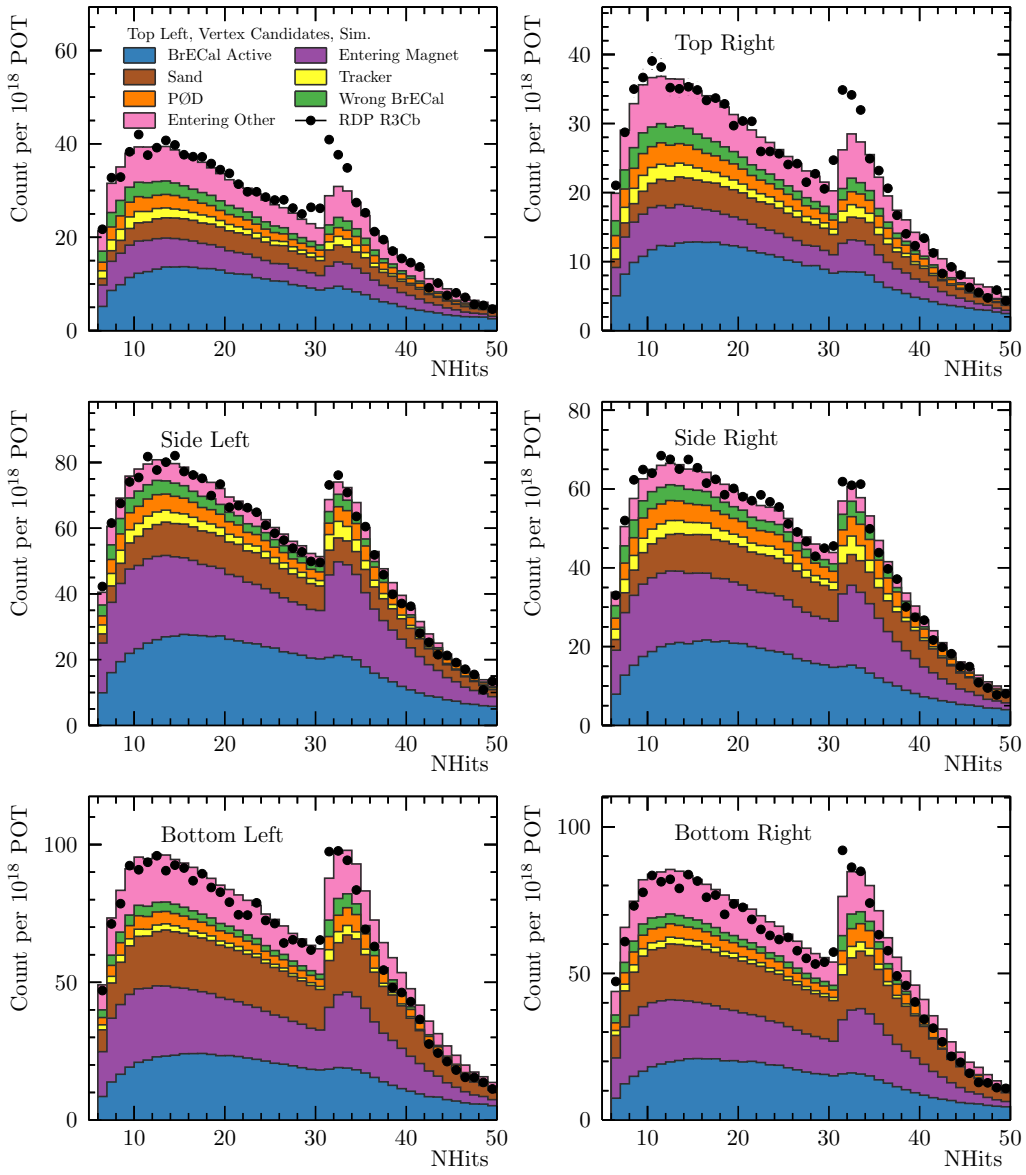


Figure A.6: The POT-normalised, number of hits multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

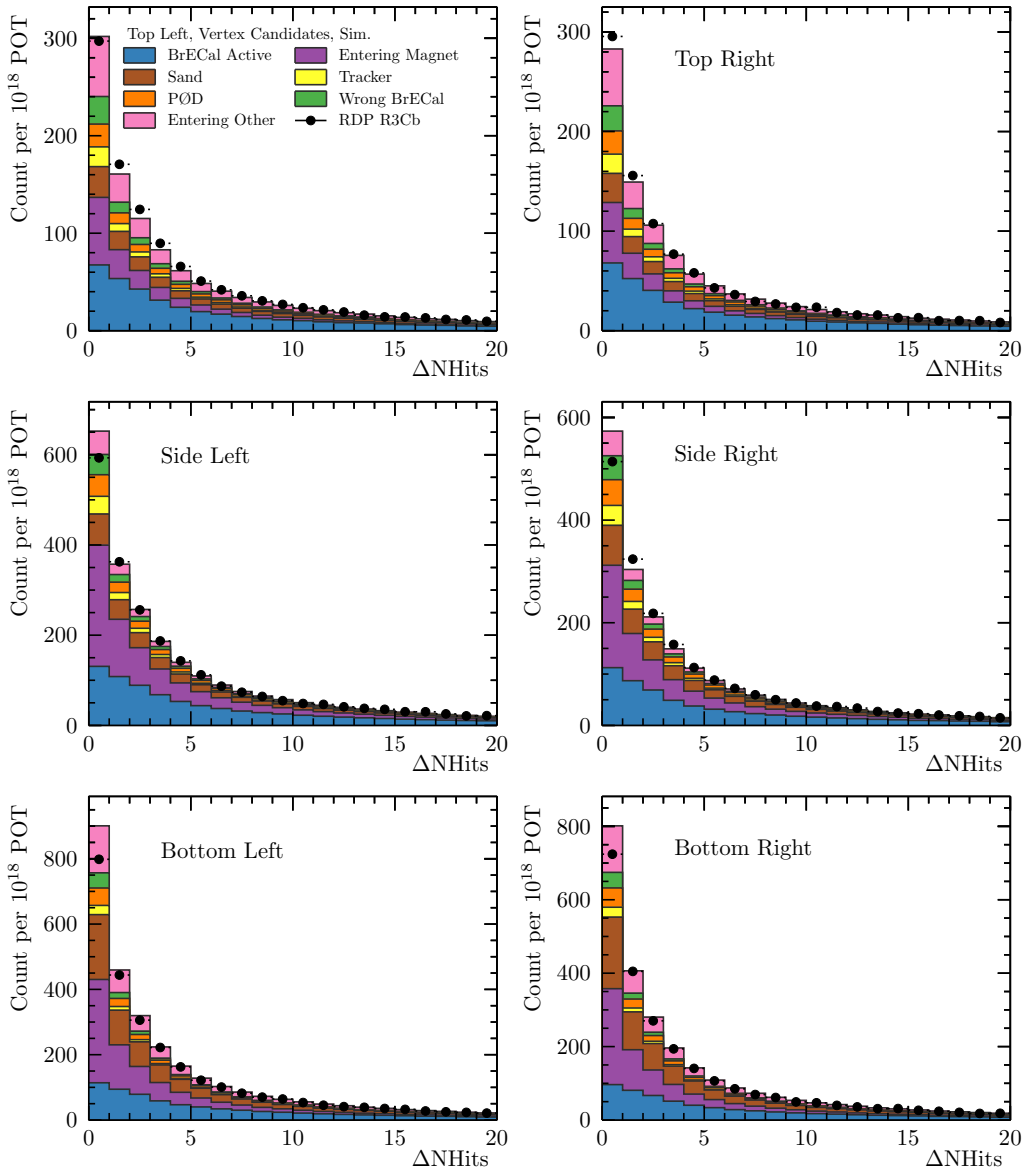


Figure A.7: The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each of the six barrel ECal modules. The distribution drawn from the simulation is separated by the true interaction position. The data overlay is calculated from the Run3C RDP control sample.

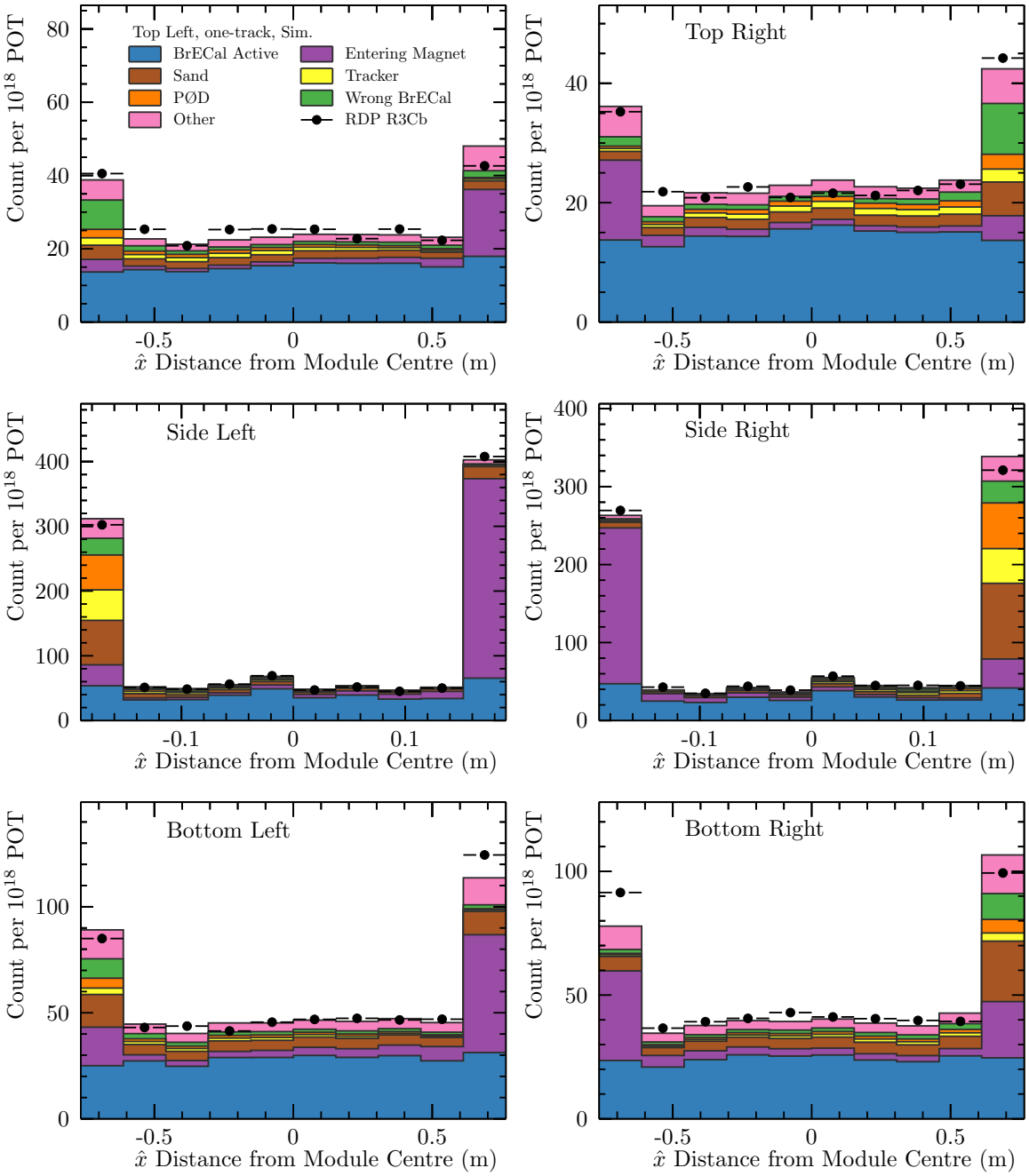


Figure A.8: The true interaction category of one-track ECal-isolated reconstructed vertex candidates as a function of the x distance from the centre of the module. Due to the orientation of the detector, the left side of the side left module corresponds to $\Delta x > 0$. This is the large module face near the magnet flux return.

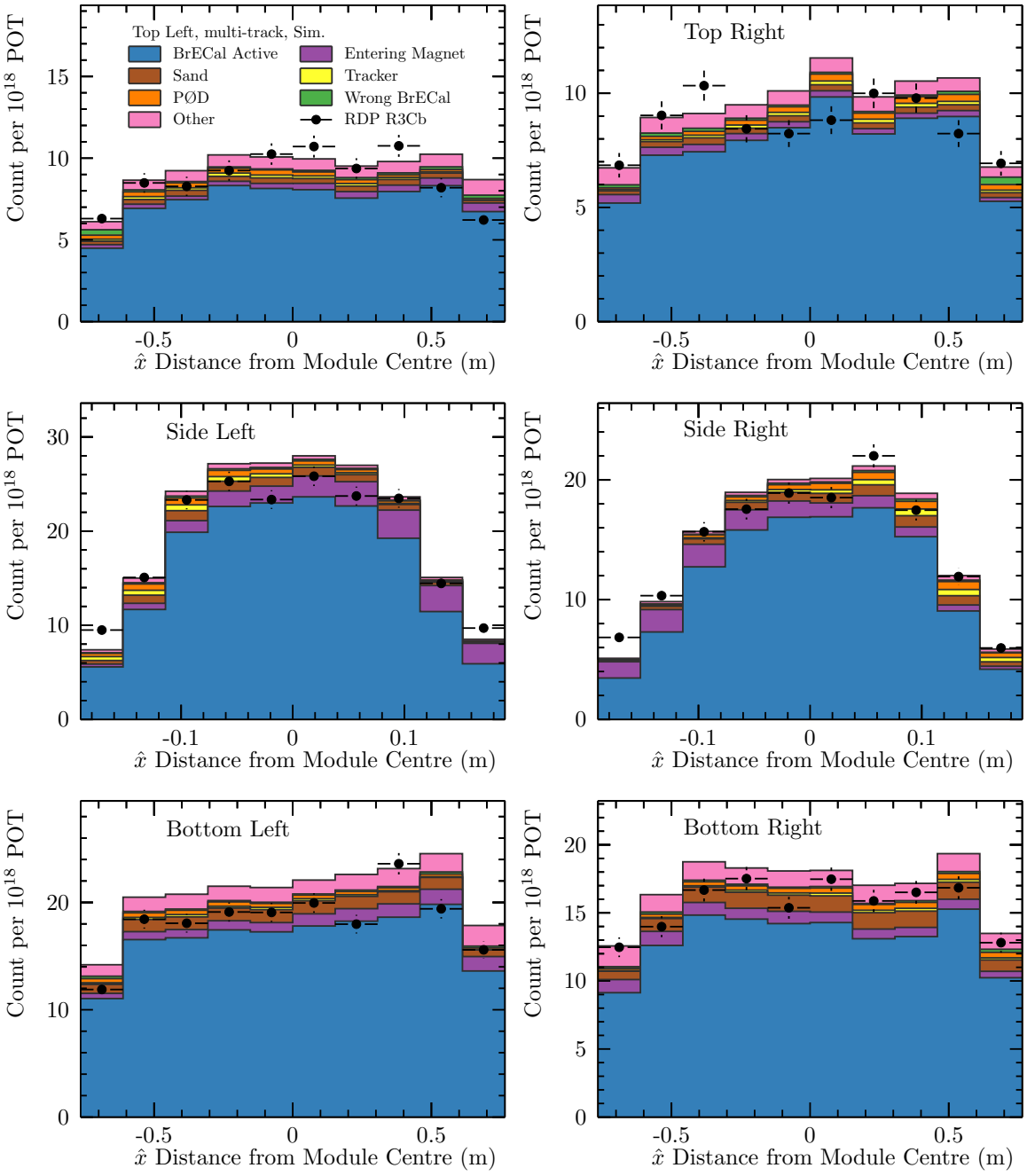


Figure A.9: The true interaction category of multi-track ECal-isolated reconstructed vertex candidates as a function of the x distance from the centre of the module. The track reconstruction efficiency can be seen to fall at the edges of the modules where less true ECal interactions are reconstructed as multi-prong than in the bulk of the detector.

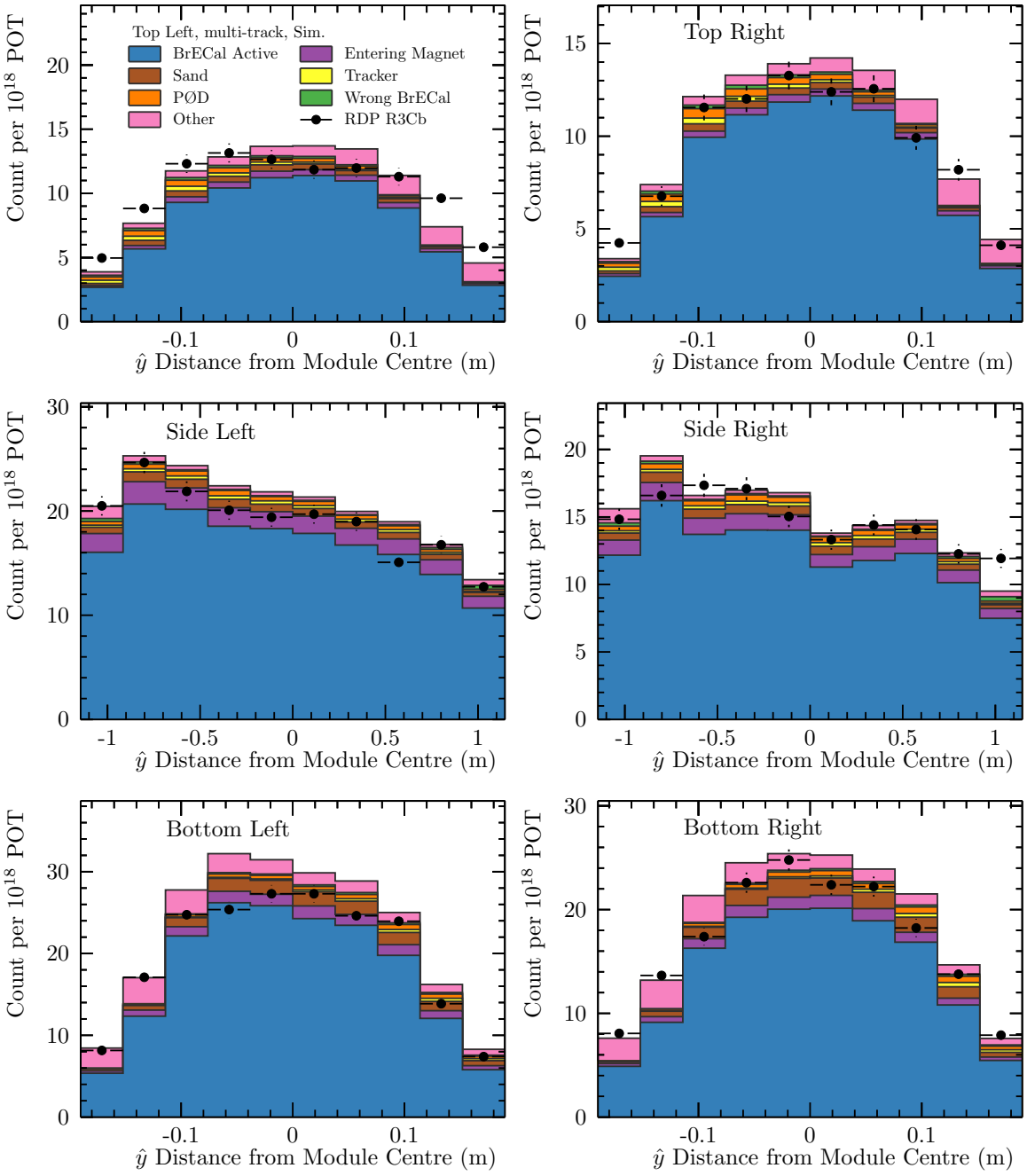


Figure A.10: The true interaction category of multi-track ECal-isolated reconstructed vertex candidates as a function of the y distance from the centre of the module. The track reconstruction efficiency can be seen to fall at the edges of the modules where less true ECal interactions are reconstructed as multi-prong than in the bulk of the detector.

Appendix B

NuWro ReWeight

The developments discussed in this appendix will be published in the proceedings of the *The XXVII International Conference on Neutrino Physics and Astrophysics* (Neutrino 2016). The following is an adaptation of the submitted manuscript, of which an eprint is available in Ref. [105].

B.1 Event Reweighting

Simulating particle interactions is an often inefficient process where interaction 'events' are generated according to some cross-section model. The correct distribution of the event properties is achieved by Monte Carlo techniques. Event generators often use rejection sampling, a technique in which thrown interaction properties are rejected with a probability proportional to the predicted cross section for producing such an event.

When generators are used as tools in particle physics analyses, it is sometimes useful to tune the simulation prediction through comparison to data. Free parameters in the models can be varied to maximise a likelihood formed between the prediction and observed data. This process often requires a large number of parameter variations, each of which requires a recalculation of the

simulation prediction. In the most simple cases, it may be possible to tune the model parameters by evaluating the cross section directly, without generating discrete events. However, models are often parameterised by kinematic variables that are not possible to precisely reconstruct in the case of neutrino–nucleus scattering. For example, in general the neutrino energy and four-momentum transfer cannot be exactly reconstructed. Furthermore, experimental event selections often place restrictions on hadronic final states and no analytic models exist to fully predict the distributions of leptonic and hadronic final states. As a result, the use of event generators is invaluable in interpreting the observed data.

As mentioned, event generation is often computationally inefficient, but it is sometimes possible to 'reweight' previously generated events in order to determine the response of a prediction to variations in the model parameters. Each generated event comes with an associated 'weight' that is proportional to the probability for that event. Reweighting is the process of calculating a scaling factor that can be combined with the original weight to determine the weight that an event would have were it originally generated with a different set of model parameters. This process involves no Monte Carlo techniques and so no computational time is 'wasted', and as a result, it can be many orders of magnitude faster than re-generation. This significant boost in calculation efficiency makes more involved studies of the compatibility between models and data sets feasible.

Using reweighting to determine the response to parameter variations often only requires a re-evaluation of the cross section for the original event kinematics. In this way, the event scale factor, SF, and event weight, W, are given by

$$\text{SF}\left(\mathbf{x}', \mathbf{k}\right) = \frac{\sigma\left(\mathbf{x}', \mathbf{k}\right)}{\sigma\left(\mathbf{x}, \mathbf{k}\right)} \text{ and} \tag{B.1}$$

$$\text{W}\left(\mathbf{x}', \mathbf{k}\right) = \text{SF}\left(\mathbf{x}', \mathbf{k}\right) \times \text{W}\left(\mathbf{x}, \mathbf{k}\right), \tag{B.2}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are the original set and the varied set of model parameters respectively, $\mathbf{k}$ is the vector of event properties required to evaluate the event cross section, $\sigma\left(\mathbf{x}', \mathbf{k}\right)$.

This is a mechanically simple process but requires careful validation.

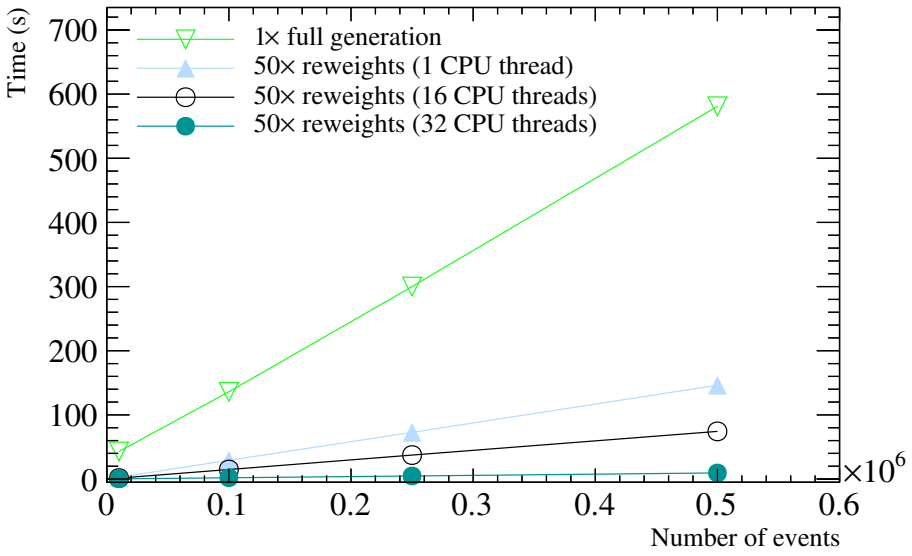


Figure B.1: The wall time taken to generate 5×10^5 carbon-target events with the NuMI on-axis flux compared to the time taken to reweight the same events 50 times. When reweighting, the majority of the CPU-time is spent evaluating cross section predictions, this is trivial to parallelise and scales well with multiple CPU cores

B.2 NuWro Reweight

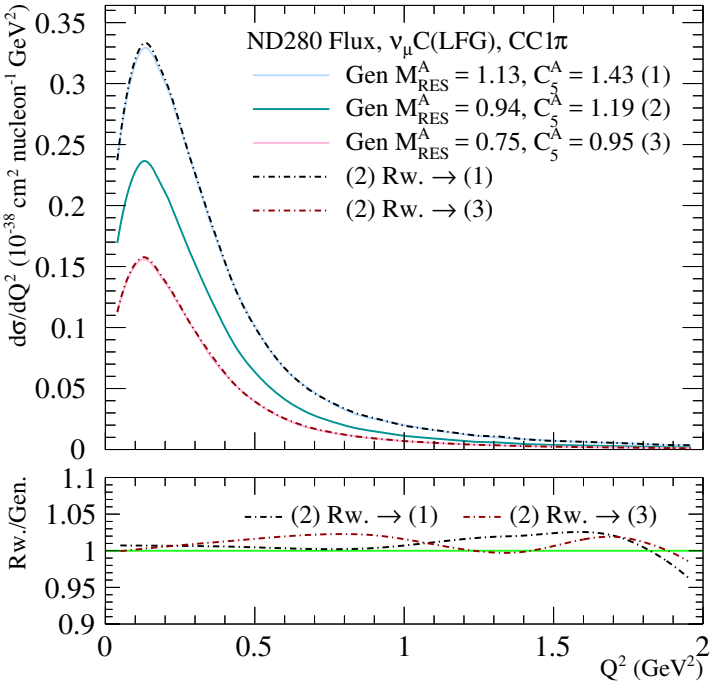
NuWro [106] is a neutrino interaction generator capable of producing predictions for neutrino-nucleus interactions for neutrinos of energies between 0.1 and 100 GeV. NuWro is written and maintained by the Wrocław neutrino group and contains a wide variety of models and variable parameters. Until now it has had no facility to perform event reweighting. This has limited its use in neutrino scattering analysis for anything more than final cross-section prediction comparisons. NuWro ReWeight was written to facilitate the use of NuWro at current and future neutrino interaction experiments.

NuWro ReWeight was written to be a fast and scalable tool for investigating prediction–data correspondence. An example of the time needed to generate and reweight events is shown in Figure B.1. For parameter variations that can be assessed through reweighting, reweighting can be seen to be orders of magnitude more efficient.

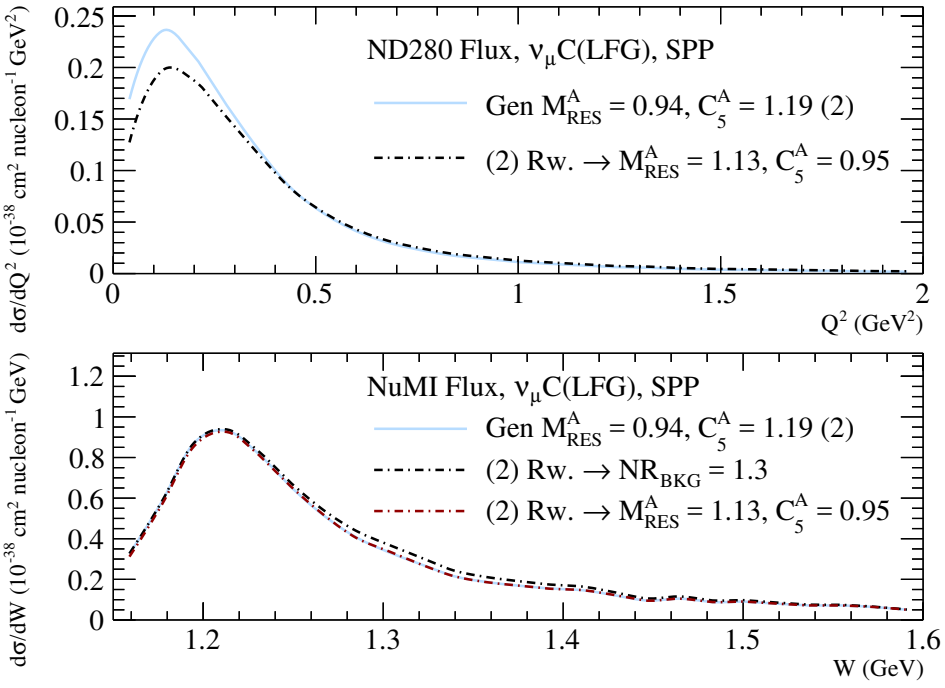
NuWro ReWeight supports variations of free-nucleon interaction model parameters for Charged Current Quasi Elastic (CCQE) [33] and single pion production (SPP) events—the most relevant interaction channels at current beam energies (*c.f.* Figure 1.5).

As the SPP reweighting was developed and validated by the author, the remaining discussion will focus there. Three variable model parameters are available for SPP events: $C_5^{\mathrm{A}}(0)$ and $M_{\mathrm{RES}}^{\mathrm{A}}$ are free parameters in nucleon resonance form factors[107]—most importantly the Delta resonance which is the only nucleon resonance explicitly simulated by NuWro [32, 108]. The other free parameter, $\mathrm{NR}^{\mathrm{BKG}}$, is a scale factor for the cross section of SPP through processes that do not explicitly involve a nucleon resonance. The effect of varying these parameters can be seen in Figure B.2b. The resonance cross-section model is parameterised in terms of the four-momentum transfer squared, $Q^2 = -\left(p_{\nu_\ell} - p_\ell\right)^2$, and the invariant hadronic mass of the final state, $W = \left|\sum_{\mathrm{hadrons}} p\right|$. In Q^2, $C_5^{\mathrm{A}}(0)$ and $M_{\mathrm{RES}}^{\mathrm{A}}$ have a significant effect on the normalisation, with shape effects most evident at low four momentum transfer. Opposing variations of $C_5^{\mathrm{A}}(0)$ and $M_{\mathrm{RES}}^{\mathrm{A}}$ result in a small change in the predicted cross section over a large fraction of the Q^2 range ($Q^2 \gtrsim 0.4$ GeV^2), which suggests that the parameters are strongly anti-correlated. The Delta resonances have a mass peak centred about $W = 1232\,\mathrm{GeV}$—as can be seen in Figure B.2b. The non-resonant SPP scale acts to fill in the transition region between Delta resonance production and multiple pion production and deep inelastic scattering interactions, both of which result in final states with greater W.

Validation of the reweighting was performed by comparing re-generated and reweighted predictions. The reweighting can be seen to accurately reproduce the generated predictions, as shown in Figure B.2a.



(a)



(b)

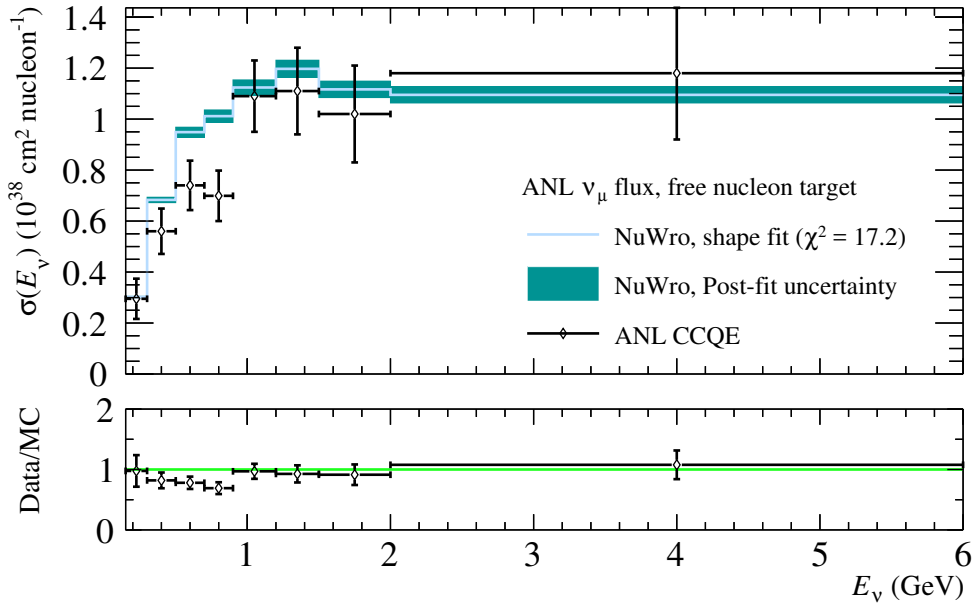
Figure B.2: (a) SPP parameter reweighting validation. Solid (dashed) lines show the generated (reweighted) prediction. The reweighting of parameter set (2) to parameter sets (1) and (3) can be seen to reproduce the generated predictions well. (b) Response of the carbon-target flux-averaged SPP cross section to variations of the SPP model parameters, as a function of Q^2 and W.

B.3 Free Nucleon Model Tuning

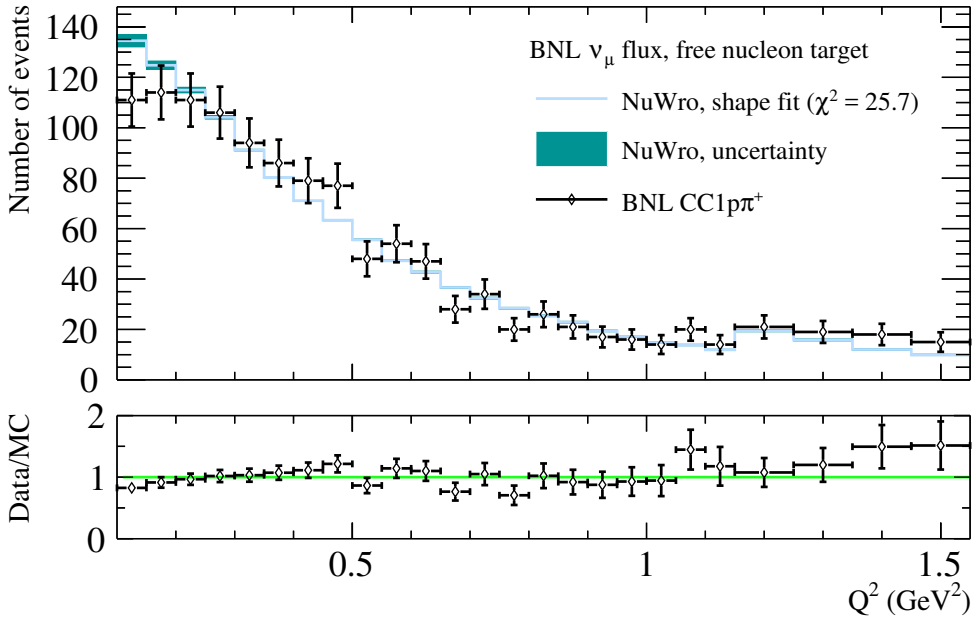
To test the event reweighting and the current predictions of NuWro, the NuWro CCQE and SPP models were tuned to historic bubble chamber data from ANL [109, 110] and BNL [111, 112]. Comparison to bubble chamber data is important because interactions on deuteron targets are expected to only exhibit weak final state interaction effects [113]. Subsequent comparison to nuclear-target data can be used to tune theoretical models of nuclear effects.

The global neutrino cross-section comparison framework, NUISANCE [100], was used to jointly fit a number of published projections of muon neutrino CCQE and SPP event selections. A binned Pearson χ^2 test between the generated (and reweighted) events and the released data was extremised to find the best fit parameter values. Event rate distributions, such as BNL CC1π^+ Q^2, Figure B.3b, were included in a shape-only way. Cross sections, such as ANL CCQE flux-unfolded $\sigma\left(E_\nu\right)$, Figure B.3a, were also used in the fit. A consistent goodness of fit test, such as χ^2 per number of degrees of freedom, is difficult to define because the bin-to-bin covariances were not published along-side the data. The pre- and post-fit parameter values are presented in Table. B.1. The fit converged and the best fit values of $M_{\mathrm{QE}}^{\mathrm{A}}$ and $M_{\mathrm{RES}}^{\mathrm{A}}$ were within the uncertainties of the nominal NuWro values. The best fit for $C_5^{\mathrm{A}}(0)$ was found to be lower than in Ref. [107], however, this fit allowed the non-resonant background contribution to vary through NR^{BKG} and included a different subset of the available data. A similar value of $C_5^{\mathrm{A}}(0)$ was found in Ref. [114]. This preliminary investigation did not include any hadronic mass projections for which the response of $C_5^{\mathrm{A}}(0)$ and NR^{BKG} might be less correlated.

As part of the minimisation, MINUIT [115] computes an approximate parameter error matrix—the default errors produced by MINUIT are renownedly un-reliable [115]. More involved procedures can be run, such as the MINOS error estimation procedure, but were not for this example fit. Uncertainty bands were generated by throwing toy sets of parameter values, distributed according to the post-fit error matrix (*c.f.* § 5). The post-fit parameter correlation matrix is presented in Figure B.4; it shows that $C_5^{\mathrm{A}}(0)$ and $M_{\mathrm{RES}}^{\mathrm{A}}$ are strongly anti-correlated. These anti-correlations render unique minimisation difficult because a variation in one parameter can result in a similar response of the test statistic as some variation in other other. Efficient



(a)



(b)

Figure B.3: Example best fit distributions and post-fit uncertainty bands for projections used in the free-nucleon interaction parameter tune—(a) and (b) show the comparisons to ANL CCQE flux-unfolded $\sigma(E_\nu)$ [109] and BNL CC 1 proton, $1\,\pi^+$ [112] event rate distribution respectively. Naively, the test statistic for both distributions suggest reasonable agreement between the prediction and the data, but because the covariance was not published, the use of a Pearson χ^2 is poorly motivated.

Param	NuWro nominal	Best fit
$M^{\mathrm{A}}_{\mathrm{QE}}$/GeV	1.03	1.05 ± 0.03
$M^{\mathrm{A}}_{\mathrm{RES}}$/GeV	0.94	0.93 ± 0.03
$C^{\mathrm{A}}_{5}(0)$	1.19	0.94 ± 0.03
$\mathrm{NR}^{\mathrm{BKG}}$	0.00	1.35 ± 0.13

Table B.1: The best fit CCQE and SPP parameter values from preliminary tune to ANL and BNL data compared with the NuWro nominal. The axial mass values remain near their nominal values while the other SPP parameters are pulled away significantly.

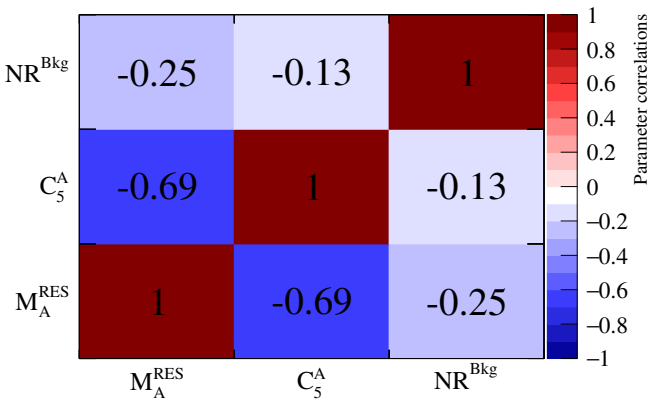


Figure B.4: The post-fit parameter correlation matrix. Variations of each of the SPP parameters are anti-correlated with each other. The two parameters governing Delta resonance productions are strongly anti-correlated as might have been expected from Figure B.2b (Bottom).

reweighting is an important tool in investigating models with correlated free parameters and their efficacy at reproducing the data.

Event reweighting has been added to the NuWro event generator. This facilitates more sophisticated investigations into the compatibility of models with data, as well as determination of well motivated, correlated model uncertainties for use in neutrino scattering analyses.

Appendix C

A lightyear of lead

At the time of writing, many neutrino physicists semi-frequently receive email purporting to disprove the observation of neutrino oscillations and thus claim that the 2015 Nobel prize, for the discovery that neutrinos are massive particles, was awarded in error. Among the many dubious comments presented in these correspondences, one of the more laughable claims is that physicists couldn't possibly have sampled a beam of neutrinos because everyone knows that at least a lightyear of lead is needed to stop a neutrino. From Table. 5.5, the ND280 Barrel ECals contain $25,200$ kg of lead, this contains approximately 1.22×10^5 mol, or 1.216×10^{29} lead atoms, each with a Van de Waals radius of 2.02×10^{-10} m. If placed in a row, one atom thick and one atom wide, this would form a column approximately 1300 lightyears long. It seems that everyone would agree that ND280 is the detector in which to perform lead-target neutrino physics.

List of Figures

1.1 Current world-leading measurements of $|\Delta m^2_{32}|$ and $\sin^2(\theta_{23})$. 17

1.2 A schematic of the Neutrinos at the Main Injector (NuMI) proton beam target and two-horn focussing system. 18

1.3 The off-axis effect on neutrino energy spectra from boosted charged pion decays. 19

1.4 Example Feynman diagrams for three charged-current neutrino scattering channels. 21

1.5 NuWro prediction of the muon neutrino charged current cross section as a function of neutrino energy overlaid on relevant experimental neutrino flux spectra. . 22

1.6 Example neutrino-induced, non-resonant single pion production diagrams used in the HNV model. 24

1.7 (a) The NuWro predicted muon neutrino CCQE and 2p2h cross section as a function of neutrino energy, compared with the corrected data from the NOMAD and MiniBooNE experiments. (b) Cartoon of a 2p2h interaction. 25

1.8 The quasi-elastic reconstructed energy bias for true CCQE interactions and other non-QE channels with no final-state pions. 27

1.9 The flux-unfolded, charged-current muon neutrino inclusive cross section for lead and hydrocarbon targets, as predicted by NEUT and NuWro. 29

2.1 Schematics of the primary and secondary beamlines that are involved in the production of the J-PARC neutrino beam. 32

2.2 (a) The protons per spill and the total number of protons on target up to the end of T2K Run 4. (b) The neutrino mode flux predictions at ND280, sub-divided by neutrino species. 34

2.3 A diagram depicting the various sub-detectors that make up the T2K off-axis near detector. 36

2.4 (a) Exploded view of the ND280 ECal construction. (b) ND280 event display showing a candidate interaction in FGD1. 37

2.5 FGD hit efficiencies. 39

2.6 (a) A ND280 TPC module. (b) ND280 TPC transverse momentum resolution. . 40

2.7 ND280 TPC energy deposition as a function of charged particle ID. 41

2.8 ND280 ECal hit efficiency and module timing resolution. 44

2.9 (a) An example νPRISM mock-up of the prediction oscillated muon neutrino flux at Super-K. (b) J-PARC muon neutrino flux spectrum as a function of off-axis angle. 48

2.10 The module mass-normalised rate of reconstructed deposits in the ND280 Barrel. 49

2.11 Previous lead-target muon neutrino charged-current scattering measurements from the CHORUS and MINERνA experiments. 50

2.12 The ECal fiducial volume definitions used in the 2015 analysis (§ 7.6.1.1)[96]. . . 52

2.13 The pre- and post-fit ECal-module selected event rates from the 2015 analysis. . 53

2.14 (a) The uncertainty correlation matrix for the ECal-module selected event rate from the 2015 analysis. (b) The background-enhanced template for the bottom left barrel module from the 2015 analysis. 54

2.15 (a) The hit multiplicity of reconstructed tracks passing the fiducial volume cut in the 2015 analysis. (b) The background to signal ratio used to justify the two dimensional cut in track charge and hit multiplicity in the 2015 analysis. 55

3.1 The simulated energy deposit from an ECal interaction occurring in the Top Right Barrel ECal. 58

3.2 A summary flow chart of the ECal-isolated vertexing algorithm. 60

3.3 The Hough transform of a single, activated bar in Hough track-parameter space. 63

3.4 (a) A two-dimensional event display of a simulated neutrino interaction in the ECal. (b) The Hough transform of the hits left by the simulated interaction in (a). 63

3.5 The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle momentum. 68

3.6 The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle polar angle cosine. 69

3.7 The muon tracking efficiency for fake interaction vertices occurring within the Barrel ECal fiducial volume as a function of the true particle azimuthal angle. . 70

3.8 The distribution of track position pulls for the non-running coordinates for the top, bottom, and side Barrel ECal modules. 72

3.9 The distribution of relevant track slope pulls for the top, bottom, and side Barrel ECal modules. 72

3.10 The distributions of the residuals for reconstructed vertex candidates with vertices occurring within the Barrel ECal active regions. 75

3.11 The migration of the number of true charged particles to the number reconstructed Hough tracks. 76

3.12 The POT-normalised, polar track angle for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. 77

3.13 The POT-normalised, azimuthal track angle for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. The distribution drawn from the simulation is subdivided by the true interaction position. 78

3.14 The POT-normalised, track length for all vertex tracks reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. 79

3.15 The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each of the six Barrel ECal modules. 80

3.16 Muon tracking efficiencies as a function of true momentum at the entry point to the TPC subdetector. 82

3.17 Muon tracking efficiencies as a function of true polar angle at the entry point to the TPC subdetector. 82

3.18 Muon tracking efficiencies as a function of true azimuthal angle at the entry point to the TPC sub-detector. 83

3.19 TREx muon tracking efficiency as a function of the true azimuthal angle at the entry point to the TPC subdetector for muons entering from each of the Barrel ECal modules. 84

3.20 The reconstructed position and momentum information from the TPC-exiting end of a track is extrapolated to the active region of an ECal to be used in inter-detector object matching. 86

3.21 The χ^2 distribution for attempted Tracker–Barrel ECal object matches. 87

3.22 The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of momentum at the entry point to the TPC. 89

3.23 The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of polar cosine angle at the entry point to the TPC. 89

3.24 The ECal-Tracker matching efficiency for muons entering from a Barrel ECal module as a function of azimuthal angle at the entry point to the TPC. 90

3.25 The results of the reconstruction on the simulated ECal interaction shown in Figure 3.1. 92

3.26 The momentum residuals for globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. 94

3.27 The residuals of the polar angle cosine for globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. 95

3.28 The momentum residuals for the TPC component of globally reconstructed tracks associated to global vertices in each of the Barrel ECal modules. 95

4.1 The rate of reconstructed one-track ECal-isolated vertex candidates as a function of the z distance from the centre of the module. 101

4.2 The rate of reconstructed multi-track ECal-isolated vertex candidates as a function of the z distance from the centre of the module. 102

4.3 The rate of reconstructed one-track ECal-isolated vertex candidates as a function of the y distance from the centre of the module. 103

4.4 The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function x distance from the module centre. 105

4.5 The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function y distance from the module centre. 106

4.6 The ratio of the vertex candidates reconstructed with one track to the number reconstructed with more than one track as a function z distance from the module centre. 107

4.7 The fiducial volume definition used in this analysis. 108

4.8 Examples of golden and silver interaction candidate topologies. 109

4.9 The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon momentum. 111

4.10 The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon cosine polar angle. 111

4.11 The TPC-active acceptance fraction for muons starting in one of the Barrel ECal module's active regions as a function of muon azimuthal angle. 112

4.12 A flow schematic describing the selection steps leading to the four analysis samples: magnet- and upstream-entering backgrounds and silver and golden ECal interaction candidates. 113

4.13 The effect of the tracker upstream veto on the golden candidate sample. 116

4.14 The effect of the upstream PØD veto on the golden candidate sample. 117

4.15 The effect of the upstream PØDECal veto on the golden candidate sample. . . . 118

4.16 The effect of the upstream Barrel ECal veto on the golden candidate sample. . . 119

4.17 The effect of the upstream SMRD veto on the golden candidate sample. 120

4.18 The effect of the upstream 'other' veto on the golden candidate sample. 121

4.19 The effect of the fiducial volume cut in the x direction on the golden candidate sample. 122

4.20 The effect of the fiducial volume cut in the y direction on the golden candidate sample. 123

4.21 The effect of the fiducial volume cut in the z direction on the golden candidate sample. 124

4.22 The effect of the fiducial volume cut in the x direction on the golden candidate sample. All other selection cuts have been applied except for the requirement of a Tracker track association. 125

4.23 The effect of each cut on the efficiency and purity of the golden selection. 127

4.24 The contents of the golden interaction candidate sample projected onto muon momentum and polar angle. 129

4.25 The migration matrix describing the smearing between true and reconstructed muon momentum. 130

4.26 The migration matrix describing the smearing between true and reconstructed muon polar angle. 131

4.27 The contents of the silver interaction candidate sample projected onto longest track length and polar angle. 134

4.28 The effect of each cut on the efficiency and purity of the silver selection. 135

4.29 The definitions of the two entering background samples in each module which aim to select tracks entering from the magnet and the upstream ECal face. . . . 137

4.30 The contents of the magnet entering background sample projected onto longest track length and polar angle. 138

4.31 The contents of the upstream entering background sample projected onto longest track length and polar angle. 139

4.32 The time of flight for the muon-like global track of selected golden event candidates. 141

4.33 The true stopping volume for muons produced in Barrel ECal interactions. . . . 143

5.1 The fractional uncertainties in the muon neutrino flux prediction at ND280. . . 145

5.2 Covariance matrix describing the correlated uncertainties in the 'neutrino mode' muon neutrino flux at the ND280. 146

5.3 The effect of the neutrino beam uncertainties propagated to the golden interaction sample. 148

5.4 The correlation matrix determined from the neutrino beam uncertainties in golden sample analysis bin number. 149

5.5 The effect of the neutrino interaction, nuclear model, and hadronic re-interaction uncertainties propagated to the golden interaction sample. 153

5.6 The correlation matrix determined from the combined effect of all interaction uncertainties in golden sample analysis bin number. 157

5.7 The reconstructed momentum variations, induced by the TPC momentum scale, momentum resolution, and magnetic field distortion systematic assessment, for 1000 toy throws of the simulated golden event candidate sample. 159

5.8 The reconstructed starting and ending positions of tracker segments of global tracks. 160

5.9 The ECal–Tracker matching control sample projected onto the Tracker segment 'search' direction polar angle and the reconstructed TPC momentum. 162

5.10 The distribution of the ECal–Tracker matching and ECal tracking efficiency-like metric. 163

5.11 The combined effect of the ND280 detector uncertainties propagated to the golden interaction sample. 168

5.12 The correlation matrix calculated from the propagation of the ND280 detector uncertainties in golden sample analysis bin number. 169

6.1 The golden sample reconstructed muon kinematic distribution for the simulation and the two sets of real data used in this analysis. 171

6.2 The effect of the silver selection cuts, and combinations thereof, on the total rate discrepancy between the different data taking periods. 174

6.3 The veto fraction as a function of proton beam intensity. 175

6.4 The shape-only distribution of reconstructed muon kinematics for all selected events in the golden sample. 177

6.5 The total shape-only correlation matrix determined from all of the error sources described in §5. 178

6.6 The fractional shape-only uncertainties, as a function of analysis bin, from the most important error sources described in §5. 179

6.7 The shape-only distribution of longest reconstructed track properties for all selected event candidates in the silver sample. 180

6.8 The shape-only distribution of longest reconstructed track properties for all selected entering background candidates in the entering magnet sample. 181

6.9 The shape-only distribution of longest reconstructed track properties for all selected entering background candidates in the entering upstream sample. 182

6.10 The fiducial mass-normalised, shape-only distribution of reconstructed neutrino energy under the assumption of CCQE kinematics. 185

6.11 The shape-only distribution of reconstructed muon azimuthal angle for the golden sample. 186

7.1 The golden interaction candidate sample where the Tracker track time of flight has been used to reposition reconstructed vertex positions to the earlier end of the track. 190

A.1 The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each barrel ECal module. . . . 195

A.2 The POT-normalised, layer numbers of the innermost constituent hit for all reconstructed vertex tracks reconstructed by the ECal-isolated vertexing for each barrel ECal module. 196

A.3 The POT-normalised, layer numbers of the outermost constituent hit for all reconstructed vertex tracks reconstructed by the ECal-isolated vertexing for each barrel ECal module. 197

A.4 The POT-normalised, total track charge in MIP-equivalent units for all vertex tracks reconstructed by the ECal-isolated vertexing for each barrel ECal module. 198

A.5 The POT-normalised, number of constituent hits for all vertex tracks reconstructed by the ECal-isolated vertexing for each barrel ECal module. 199

A.6 The POT-normalised, number of hits multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each barrel ECal modules. 201

A.7 The POT-normalised, reconstructed track multiplicity for all vertex candidates reconstructed by the ECal-isolated vertexing for each barrel ECal modules. . . . 202

A.8 The true interaction category of one-track ECal-isolated reconstructed vertex candidates as a function of the x distance from the centre of the module. 203

A.9 The true interaction category of multi-track ECal-isolated reconstructed vertex candidates as a function of the x distance from the centre of the module. 204

A.10 The true interaction category of multi-track ECal-isolated reconstructed vertex candidates as a function of the y distance from the centre of the module. 205

B.1 A comparison of the wall time taken to re-generate and reweight a large number of events with NuWro. 208

B.2 Examples of the SPP parameter response as functions of Q^2 and W. 210

B.3 Example best fit distributions and post-fit uncertainty bands for projections of ANL and BNL event data used in the free-nucleon interaction parameter tune . 212

B.4 The post-fit parameter correlation matrix from the NuWro ReWeight SPP parameter tune. 213

List of Tables

1.1 Summary of the long-baseline neutrino oscillation experiments mentioned in chapter 1. 16

4.1 The restrictions placed on the integrated-over dimensions in the distributions of $\Delta\{x, y, z\}$ that are used to motivate the choice of fiducial volume. 100

4.2 The fiducial masses of the two Barrel ECal module types. 104

5.1 The number of selected Barrel ECal interactions in a sample of Run3Cb POT-equivalent simulated neutrino beam data 147

5.2 The nominal values, widths, and applicability limits for the most relevant neutrino interaction parameters. 154

5.3 The number of selected Barrel ECal interactions in a sample of Run 3Cb POT-equivalent simulated neutrino beam data. 155

5.4 The dimensions and manufacturing tolerances of the Barrel ECal scintillator bars and lead absorber sheets. 165

5.5 The component masses for the two Barrel ECal module orientations, as calculated from the detector simulation geometry and the design specifications given in Ref. [67]. 166

6.1 The number of selected candidates per 10^{18} POT analysed for each sample. . . . 172

6.2 The number of simulated true interactions per mass and POT in each of the Barrel ECal modules. 186

B.1 The best fit CCQE and SPP parameter values from preliminary tune to ANL and BNL data compared with the NuWro nominal. 213

Bibliography

[1] C. L. Cowan et al., "Detection of the Free Neutrino: a Confirmation", Science vol. 124 (3212) pp. 103–104 (1956), doi:10.1126/science.124.3212.103

[2] The ALEPH Collaboration et al., "Precision electroweak measurements on the Z resonance", Physics Reports vol. 427 (5 - 6) pp. 257 – 454 (2006), doi:10.1016/j.physrep.2005.12.006

[3] Y. Fukuda et al. (Super-Kamiokande collaboration), "Evidence for oscillation of atmospheric neutrinos", Phys.Rev.Lett. vol. 81 pp. 1562–1567 (1998), doi:10.1103/PhysRevLett.81.1562

[4] Q.R. Ahmad et al. (SNO collaboration), "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory", Phys.Rev.Lett. vol. 89 p. 011301 (2002), doi:10.1103/PhysRevLett.89.011301

[5] K. Eguchi et al. (KamLAND collaboration), "First results from KamLAND: Evidence for reactor anti-neutrino disappearance", Phys.Rev.Lett. vol. 90 p. 021802 (2003), doi:10.1103/PhysRevLett.90.021802

[6] F.P. An et al. (DAYA-BAY collaboration), "Observation of electron-antineutrino disappearance at Daya Bay", Phys.Rev.Lett. vol. 108 p. 171803 (2012), doi:10.1103/PhysRevLett.108.171803

[7] K. Abe et al. (T2K collaboration), "Observation of Electron Neutrino Appearance in a Muon Neutrino Beam", Phys.Rev.Lett. vol. 112 p. 061802 (2014), doi:10.1103/PhysRevLett.112.061802, 1311.4750

[8] Boris Kayser, "B-Meson and Neutrino Oscillation: A Unified Treatment", (2011), arXiv:1110.3047 [hep-ph]

[9] R N Mohapatra et al., "Theory of neutrinos: a white paper", Reports on Progress in Physics vol. 70 (11) p. 1757 (2007), http://stacks.iop.org/0034-4885/70/i=11/a=R02

[10] D. G. Michael et al. (MINOS Collaboration), "Observation of Muon Neutrino Disappearance with the MINOS Detectors in the NuMI Neutrino Beam", Phys. Rev. Lett. vol. 97 p. 191801 (2006), doi:10.1103/PhysRevLett.97.191801

[11] K. Abe et al. (T2K), "The T2K Experiment", Nucl. Instrum. Meth. vol. A 659 pp. 106–135 (2011), doi:10.1016/j.nima.2011.06.067

[12] Jianming Bian, "The NOvA Experiment: Overview and Status", (2013), arXiv:1309.7898 [physics.ins-det]

[13] P. Adamson et al. (MINOS Collaboration), "Study of muon neutrino disappearance using the Fermilab Main Injector neutrino beam", Phys. Rev. D vol. 77 p. 072002 (2008), doi:10.1103/PhysRevD.77.072002

[14] Justin Evans (MINOS), "The MINOS experiment: results and prospects", Adv. High Energy Phys. vol. 2013 p. 182537 (2013), doi:10.1155/2013/182537

[15] M.G. Aartsen et al., "The IceCube Neutrino Observatory: instrumentation and online systems", Journal of Instrumentation vol. 12 (03) p. P03012 (2017), http://stacks.iop.org/1748-0221/12/i=03/a=P03012

[16] M. G. Aartsen et al. (IceCube Collaboration), "Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data", Phys. Rev. D vol. 91 p. 072004 (2015), doi:10.1103/PhysRevD.91.072004

[17] "Expression of Interest for an Extended Run at T2K to 20×10^{21} POT", (2016), https://j-parc.jp/researcher/Hadron/en/pac_1601/pdf/EoI_2016-10.pdf

[18] K. Abe et al. (Hyper-Kamiokande Working Group), "A Long Baseline Neutrino Oscillation Experiment Using J-PARC Neutrino Beam and Hyper-Kamiokande", (2014), arXiv:1412.4673

[19] R. Acciarri et al. (DUNE), "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)", (2016), arXiv:1601.05471 [physics.ins-det]

[20] K. Abe, J. Amey, C. Andreopoulos, and M. Antonova (T2K Collaboration), "Combined Analysis of Neutrino and Antineutrino Oscillations at T2K", Phys. Rev. Lett. vol. 118 p. 151801 (2017), doi:10.1103/PhysRevLett.118.151801

[21] "J-PARC TDR, KEK-Report 2002-13 and JAERI-Tech 2003-044", https://cds.cern.ch/record/747209/files/34072617.pdf

[22] Christopher W. Walter, "The Super-Kamiokande Experiment", (2008), arXiv:0802.1041 [hep-ex]

[23] K. Abe et al. (T2K collaboration), "Neutrino oscillation physics potential of the T2K experiment", Progress of Theoretical and Experimental Physics vol. 2015 (4) (2015), doi:10.1093/ptep/ptv031

[24] Gersende Prior, "Status of the Neutrino Factory accelerator design studies", Journal of Physics: Conference Series vol. 408 (1) p. 012013 (2013), http://stacks.iop.org/1742-6596/408/i=1/a=012013

[25] J. C Dusseux, J. B M. Pattison, and G. Ziebarth, "The CERN magnetic horn (1971) and its remote-handling system", (1972)

[26] L. Aliaga et al. (MINERvA), "Design, Calibration, and Performance of the MINERvA Detector", Nucl. Instrum. Meth. vol. A743 pp. 130–159 (2014), doi:10.1016/j.nima.2013.12.053

[27] Kirsty Duffy, Measurement of the neutrino oscillation parameters $\sin^2\theta_{23}$, Δm^2_{32}, $\sin^2\theta_{13}$, and δ_{CP} in neutrino and antineutrino oscillation at T2K, Ph.D. thesis, University of Oxford (2017), (https://www.t2k.org/docs/thesis/078)

[28] Omar Benhar and Noemi Rocco, "Nuclear Effects in Neutrino Interactions and Their Impact on the Determination of Oscillation Parameters", Adv. High Energy Phys. vol. 2013 p. 912702 (2013), doi:10.1155/2013/912702, 1310.3869

[29] L. Wolfenstein, "Neutrino oscillations in matter", Phys. Rev. D vol. 17 pp. 2369–2374 (1978), doi:10.1103/PhysRevD.17.2369

[30] Alexei Yu. Smirnov, "Neutrino mass hierarchy and matter effects", (2013), arXiv:1312.7309 [hep-ph]

[31] U. Mosel, "Neutrino Interactions with Nucleons and Nuclei: Importance for Long-Baseline Experiments", Annu. Rev. Nucl. Part. Sci. vol. 66 pp. 171–95 (2016)

[32] J A Nowak, "Wrocław neutrino event generator", Physica Scripta vol. 2006 (T127) p. 70 (2006), http://stacks.iop.org/1402-4896/2006/i=T127/a=025

[33] C. H. Llewellyn Smith, "Neutrino Reactions at Accelerator Energies", Phys. Rept. vol. 3 pp. 261–379 (1972), doi:10.1016/0370-1573(72)90010-5

[34] E. Hernández, J. Nieves, and M. Valverde, "Weak pion production off the nucleon", Phys. Rev. D vol. 76 p. 033005 (2007), doi:10.1103/PhysRevD.76.033005

[35] Dieter Rein and Lalit M Sehgal, "Neutrino-excitation of baryon resonances and single pion production", Annals of Physics vol. 133 (1) pp. 79 – 153 (1981), doi:10.1016/0003-4916(81)90242-6

[36] O. Lalakulich, T. Leitner, O. Buss, and U. Mosel, "One pion production in neutrino reactions: Including nonresonant background", Phys. Rev. D vol. 82 p. 093001 (2010), doi:10.1103/PhysRevD.82.093001

[37] U. Mosel and K. Gallmeister, "Muon Neutrino-Induced Charged Current Pion Production on Nuclei", (2017), arXiv:1702.04932 [nucl-th]

[38] Arie Bodek and Un-ki Yang, "Axial and Vector Structure Functions for Electron- and Neutrino- Nucleon Scattering Cross Sections at all Q^2 using Effective Leading order Parton Distribution Functions", (2010), arXiv:1011.6592 [hep-ph]

[39] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands, "PYTHIA 6.4 physics and manual", Journal of High Energy Physics vol. 2006 (05) p. 026 (2006), http://stacks.iop.org/1126-6708/2006/i=05/a=026

[40] A. A. Aguilar-Arevalo et al. (MiniBooNE Collaboration), "Measurement of neutrino-induced charged-current charged pion production cross sections on mineral oil at $E_\nu \sim 1$ GeV", Phys. Rev. D vol. 83 p. 052007 (2011), doi:10.1103/PhysRevD.83.052007

[41] V. Lyubushkin et al. (NOMAD), "A study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment", The European Physical Journal C vol. 63 (3) pp. 355–381 (2009), doi:10.1140/epjc/s10052-009-1113-0

[42] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, "Inclusive charged-current neutrino-nucleus reactions", Phys. Rev. C vol. 83 p. 045501 (2011), doi:10.1103/PhysRevC.83.045501

[43] M. Martini, M. Ericson, G. Chanfray, and J. Marteau, "Neutrino and antineutrino quasielastic interactions with nuclei", Phys. Rev. C vol. 81 p. 045502 (2010), doi:10.1103/PhysRevC.81.045502

[44] O. Benhar, A. Fabrocini, S. Fantoni, and I. Sick, "Spectral function of finite nuclei and scattering of GeV electrons", Nuclear Physics A vol. 579 (3) pp. 493 – 517 (1994), doi:10.1016/0375-9474(94)90920-2

[45] J. Nieves, J. E. Amaro, and M. Valverde, "Inclusive quasielastic charged-current neutrino-nucleus reactions", Phys. Rev. C vol. 70 p. 055503 (2004), doi:10.1103/PhysRevC.70.055503

[46] J. Ranft, "Hadron production in hadron-nucleus and nucleus-nucleus collisions in a dual parton model modified by a formation zone intranuclear cascade", Zeitschrift für Physik C Particles and Fields vol. 43 (3) pp. 439–446 (1989), doi:10.1007/BF01506540

[47] N. Metropolis et al., "Monte Carlo Calculations on Intranuclear Cascades. I. Low-Energy Studies", Phys. Rev. vol. 110 pp. 185–203 (1958), doi:10.1103/PhysRev.110.185

[48] K. Abe et al. (T2K collaboration), "Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam", Phys. Rev. Lett. vol. 112 p. 181801 (2014), doi:10.1103/PhysRevLett.112.181801

[49] P. Adamson et al. (NOvA Collaboration), "First measurement of muon-neutrino disappearance in NOvA", Phys. Rev. D vol. 93 p. 051104 (2016), doi:10.1103/PhysRevD.93.051104

[50] K. Abe et al. (T2K), "Measurement of the electron neutrino charged-current interaction rate on water with the T2K ND280 π^0 detector", Phys. Rev. vol. D91 p. 112010 (2015), doi:10.1103/PhysRevD.91.112010

[51] K. Abe et al. (T2K), "Measurement of the ν_μ charged current quasielastic cross section on carbon with the T2K on-axis neutrino beam", Phys. Rev. vol. D91 (11) p. 112002 (2015), doi:10.1103/PhysRevD.91.112002

[52] Yoshinari Hayato, "A neutrino interaction simulation program library NEUT", Acta Phys. Polon. vol. B40 pp. 2477–2489 (2009)

[53] K. Abe et al. (T2K collaboration), "T2K neutrino flux prediction", Phys. Rev. D vol. 87 p. 012001 (2013), doi:10.1103/PhysRevD.87.012001

[54] K. Abe et al. (T2K collaboration), "Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam", Phys.Rev.Lett. vol. 107 p. 041801 (2011), doi:10.1103/PhysRevLett.107.041801

[55] K. Abe et al. (T2K collaboration), "Measurement of Neutrino Oscillation Parameters from Muon Neutrino Disappearance with an Off-axis Beam", Phys.Rev.Lett. vol. 111 p. 211803 (2013), doi:10.1103/PhysRevLett.111.211803

[56] K. Abe et al. (T2K Collaboration), "Measurements of Neutrino Oscillation in Appearance and Disappearance Channels by the T2K Experiment With 6.6×10^{20} Protons on Target", Phys. Rev. D vol. 91 p. 072010 (2015), doi:10.1103/PhysRevD.91.072010

[57] Ko Abe et al. (T2K), "Measurement of Muon Antineutrino Oscillations with an Accelerator-Produced Off-Axis Beam", Phys. Rev. Lett. vol. 116 (18) p. 181801 (2016), doi:10.1103/PhysRevLett.116.181801

[58] K. Abe et al., "Measurements of the T2K neutrino beam properties using the INGRID on-axis near detector", Nuclear Instruments and Methods in Physics Research A vol. 694 pp. 211–223 (2012), doi:10.1016/j.nima.2012.03.023

[59] K. Abe et al. (T2K), "Measurement of the inclusive ν_μ charged current cross section on iron and hydrocarbon in the T2K on-axis neutrino beam", Phys. Rev. vol. D90 (5) p. 052010 (2014), doi:10.1103/PhysRevD.90.052010

[60] K. Abe et al. (T2K), "Measurement of the muon neutrino inclusive charged-current cross section in the energy range of 1-3 GeV with the T2K INGRID detector", Phys. Rev. vol. D93 (7) p. 072002 (2016), doi:10.1103/PhysRevD.93.072002

[61] T. Sekiguchi et al., "Development and operational experience of magnetic horn system for T2K experiment", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 789 pp. 57 – 80 (2015), doi:10.1016/j.nima.2015.04.008

[62] K. Suzuki et al., "Measurement of the muon beam direction and muon flux for the T2K neutrino experiment", Progress of Theoretical and Experimental Physics vol. 2015 (5) p. 053C01 (2015), doi:10.1093/ptep/ptv054

[63] S. Bhadra et al., "Optical transition radiation monitor for the T2K experiment", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 703 pp. 45 – 58 (2013), doi:10.1016/j.nima.2012.11.044

[64] M. Friend et al., "Flux Prediction and Uncertainty Updates with NA61 2009 Thin Target Data and Negative Focussing Mode Predictions", Tech. Rep. 217, T2K (2016), Available upon request., http://www.t2k.org/docs/technotes/217

[65] 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 vol. 637 (1) pp. 25 – 46 (2011), doi:10.1016/j.nima.2011.02.036

[66] P.-A. Amaudruz et al., "The T2K Fine-Grained Detectors", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 696 pp. 1 – 31 (2012), doi:10.1016/j.nima.2012.08.020

[67] D. Allan et al. (T2K UK), "The Electromagnetic Calorimeter for the T2K Near Detector ND280", JINST vol. 8 p. P10019 (2013), doi:10.1088/1748-0221/8/10/P10019, 1308.3445

[68] 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 vol. 686 pp. 48 – 63 (2012), doi:10.1016/j.nima.2012.05.028

[69] S. Aoki et al., "The T2K Side Muon Range Detector (SMRD)", Nucl. Instrum. Meth. vol. A698 pp. 135–146 (2013), doi:10.1016/j.nima.2012.10.001

[70] K. Abe et al. (T2K), "Measurement of the inclusive ν_μ charged current cross section on carbon in the near detector of the T2K experiment", Phys. Rev. vol. D87 (9) p. 092003 (2013), doi:10.1103/PhysRevD.87.092003

[71] K. Abe et al. (T2K), "Measurement of the Inclusive Electron Neutrino Charged Current Cross Section on Carbon with the T2K Near Detector", Phys. Rev. Lett. vol. 113 (24) p. 241803 (2014), doi:10.1103/PhysRevLett.113.241803

[72] K. Abe et al. (T2K), "Measurement of the ν_μ charged-current quasielastic cross section on carbon with the ND280 detector at T2K", Phys. Rev. vol. D92 (11) p. 112003 (2015), doi:10.1103/PhysRevD.92.112003

[73] K. Abe et al. (T2K collaboration), "Measurement of the Inclusive NuMu Charged Current Cross Section on Carbon in the Near Detector of the T2K Experiment", Phys.Rev. vol. D87 p. 092003 (2013), doi:10.1103/PhysRevD.87.092003

[74] K. Abe et al. (T2K), "Measurement of the intrinsic electron neutrino component in the T2K neutrino beam with the ND280 detector", Phys. Rev. vol. D89 p. 092003 (2014), doi:10.1103/PhysRevD.89.092003

[75] Ko Abe et al. (T2K), "Measurement of double-differential muon neutrino charged-current interactions on C_8H_8 without pions in the final state using the T2K off-axis beam", Phys. Rev. vol. D93 (11) p. 112012 (2016), doi:10.1103/PhysRevD.93.112012

[76] K. Abe et al. (T2K Collaboration), "First measurement of the muon neutrino charged current single pion production cross section on water with the T2K near detector", Phys. Rev. D vol. 95 p. 012010 (2017), doi:10.1103/PhysRevD.95.012010

[77] M. Yokoyama et al., "Development of Multi-Pixel Photon Counters", eConf vol. C0604032 p. 0126 (2006), arXiv:physics/0605241 [physics.ins-det]

[78] Georges Charpak, J. Derre, Y. Giomataris, and P. Rebourgeard, "MICROMEGAS, a multipurpose gaseous detector", Nucl. Instrum. Meth. vol. A478 pp. 26–36 (2002), doi:10.1016/S0168-9002(01)01713-2

[79] Pip Hamilton, A Study of Neutrino Interactions in Argon Gas, Ph.D. thesis, Imperial College London (2015), (https://www.t2k.org/docs/thesis/062)

[80] M. Antonello et al. (LAr1-ND, ICARUS-WA104, MicroBooNE), "A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam", (2015), arXiv:1503.01520 [physics.ins-det]

[81] R. Acciarri et al., "Design and construction of the MicroBooNE detector", Journal of Instrumentation vol. 12 (02) p. P02017 (2017), http://stacks.iop.org/1748-0221/12/i=02/a=P02017

[82] A. Vacheret et al., "Characterization and simulation of the response of Multi-Pixel Photon Counters to low light levels", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 656 (1) pp. 69 – 83 (2011), doi:10.1016/j.nima.2011.07.022

[83] I. Antcheva et al., "ROOT –?? A C++ framework for petabyte data storage, statistical analysis and visualization", Computer Physics Communications vol. 180 (12) pp. 2499 – 2512 (2009), doi:10.1016/j.cpc.2009.08.005

[84] G Battistoni et al., "Hadron production simulation by FLUKA", Journal of Physics: Conference Series vol. 408 (1) p. 012051 (2013), http://stacks.iop.org/1742-6596/408/i=1/a=012051

[85] C. Zeitnitz and T.A. Gabriel, "The GEANT-CALOR interface and benchmark calculations of ZEUS test calorimeters", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 349 (1) pp. 106 – 111 (1994), doi:10.1016/0168-9002(94)90613-0

[86] N. Abgrall et al. (The NA61/SHINE Collaboration), "Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c", Phys. Rev. C vol. 84 p. 034604 (2011), doi:10.1103/PhysRevC.84.034604

[87] R. Brun et al., "GEANT3", (1987)

[88] C. Andreopoulos et al., "The GENIE Neutrino Monte Carlo Generator", Nucl. Instrum. Meth. vol. A614 pp. 87–104 (2010), doi:10.1016/j.nima.2009.12.009, 0905.2517

[89] S. Agostinelli et al., "Geant4 – a simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 506 (3) pp. 250 – 303 (2003), doi:10.1016/S0168-9002(03)01368-8

[90] A. Cervera-Villanueva, J.J. Gómez-Cadenas, and J.A. Hernando, ""RecPack" a reconstruction toolkit", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 534 (1) pp. 180 – 183 (2004), doi:10.1016/j.nima.2004.07.074

[91] S. Bhadra et al. (nuPRISM), "Letter of Intent to Construct a nuPRISM Detector in the J-PARC Neutrino Beamline", (2014), arXiv:1412.3086 [physics.ins-det]

[92] M. Betancourt et al. (MINERvA), "Direct Measurement of Nuclear Dependence of Charged Current Quasielastic-like Neutrino Interactions using MINERvA", (2017), 1705.03791

[93] "Measurement of the Z/A dependence of neutrino charged-current total cross-sections", The European Physical Journal C - Particles and Fields vol. 30 (2) pp. 159–167 (2003), doi:10.1140/epjc/s2003-01292-3

[94] B. G. Tice et al. (MINERvA), "Measurement of Ratios of ν_μ Charged-Current Cross Sections on C, Fe, and Pb to CH at Neutrino Energies 2-20 GeV", Phys. Rev. Lett. vol. 112 (23) p. 231801 (2014), doi:10.1103/PhysRevLett.112.231801

[95] Mark Scott, Measuring charged current neutrino interactions in the electromagnetic calorimeters of the ND280 detector, Ph.D. thesis, Imperial College London (2013), (http://www.t2k.org/docs/thesis/036/mscottthesis)

[96] Dominic Bevan Brailsford, Measurement of the charged current inclusive muon neutrino interaction cross-section on lead using the T2K ND280 electromagnetic calorimeters, Ph.D. thesis, Imperial College London (2016), (http://www.t2k.org/docs/thesis/066)

[97] R Mankel, "Pattern recognition and event reconstruction in particle physics experiments", Reports on Progress in Physics vol. 67 (4) p. 553 (2004), http://stacks.iop.org/0034-4885/67/i=4/a=R03

[98] Laura Monfregola, Muon neutrino charged current quasi-elastic interactions in the T2K off-axis near detector, Ph.D. thesis, Universtitat de València (2015), (https://www.t2k.org/docs/thesis/063)

[99] C. Wilkinson et al., "Testing charged current quasi-elastic and multinucleon interaction models in the NEUT neutrino interaction generator with published datasets from the MiniBooNE and MINERνA experiments", Phys. Rev. vol. D93 (7) p. 072010 (2016), doi:10.1103/PhysRevD.93.072010

[100] P. Stowell et al., "NUISANCE: a neutrino cross-section generator tuning and comparison framework", Journal of Instrumentation vol. 12 (01) p. P01016 (2017), http://stacks.iop.org/1748-0221/12/i=01/a=P01016

[101] E. J. Moniz et al., "Nuclear Fermi Momenta from Quasielastic Electron Scattering", Phys. Rev. Lett. vol. 26 pp. 445–448 (1971), doi:10.1103/PhysRevLett.26.445

[102] L.L. Salcedo, E. Oset, M.J. Vicente-Vacas, and C. Garcia-Recio, "Computer simulation of inclusive pion nuclear reactions", Nuclear Physics A vol. 484 (3) pp. 557 – 592 (1988), doi:10.1016/0375-9474(88)90310-7

[103] Patrick de Perio, Joint Three-Flavour Oscillation Analysis of Numu Disappearance and Nue Appearance in the T2K Neutrino Beam, Ph.D. thesis, University of Toronto (2014), (https://www.t2k.org/docs/thesis/046)

[104] X.-G. Lu et al., "Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy", Phys. Rev. C vol. 94 p. 015503 (2016), doi:10.1103/PhysRevC.94.015503

[105] Luke Pickering, Patrick Stowell, and Jan Sobczyk, "Event reweighting with the NuWro neutrino interaction generator", in "27th International Conference on Neutrino Physics and Astrophysics (Neutrino 2016) London, United Kingdom, July 4-9, 2016", (2016), 1610.07053, URL https://inspirehep.net/record/1494066/files/arXiv:1610.07053.pdf

[106] Tomasz Golan, Cezary Juszczak, and Jan T. Sobczyk, "Effects of final-state interactions in neutrino-nucleus interactions", Phys. Rev. C vol. 86 p. 015505 (2012), doi:10.1103/PhysRevC.86.015505

[107] Krzysztof M. Graczyk and Jan T. Sobczyk, "Form factors in the quark resonance model", Phys. Rev. D vol. 77 p. 053001 (2008), doi:10.1103/PhysRevD.77.053001

[108] Cezary Juszczak, Jaroslaw A. Nowak, and Jan T. Sobczyk, "Simulations from a new neutrino event generator", Nucl. Phys. Proc. Suppl. vol. 159 pp. 211–216 (2006), doi:10.1016/j.nuclphysbps.2006.08.069

[109] S. J. Barish et al., "Study of neutrino interactions in hydrogen and deuterium: Description of the experiment and study of the reaction $\nu + d \to \mu^- + p + p_s$", Phys. Rev. D vol. 16 pp. 3103–3121 (1977), doi:10.1103/PhysRevD.16.3103

[110] G. M. Radecky et al., "Study of single-pion production by weak charged currents in low-energy νd interactions", Phys. Rev. D vol. 25 pp. 1161–1173 (1982), doi:10.1103/PhysRevD.25.1161

[111] N. J. Baker et al., "Quasielastic Neutrino Scattering: A Measurement of the Weak Nucleon Axial Vector Form-Factor", Phys. Rev. vol. D23 pp. 2499–2505 (1981), doi:10.1103/PhysRevD.23.2499

[112] T. Kitagaki et al., "Charged-current exclusive pion production in neutrino-deuterium interactions", Phys. Rev. D vol. 34 pp. 2554–2565 (1986), doi:10.1103/PhysRevD.34.2554

[113] Jia-Jun Wu, T. Sato, and T.-S. H. Lee, "Incoherent pion production in neutrino-deuteron interactions", Phys. Rev. C vol. 91 p. 035203 (2015), doi:10.1103/PhysRevC.91.035203, URL https://link.aps.org/doi/10.1103/PhysRevC.91.035203

[114] E. Hernández, J. Nieves, M. Valverde, and M. J. Vicente Vacas, "$N-\Delta(1232)$", Phys. Rev. D vol. 81 p. 085046 (2010), doi:10.1103/PhysRevD.81.085046, URL https://link.aps.org/doi/10.1103/PhysRevD.81.085046

[115] F. James, "MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1", (1994)

,